Granular Treasury Demand with Arbitrageurs

Kristy A.E. Jansen* Wenhao Li† Lukas Schmid‡

September 2024

Abstract

We collect a novel dataset encompassing the vast majority of the portfolio holders of the U.S. Treasury market and estimate granular demand functions for the Federal Reserve and preferred habitat investors such as commercial banks, mutual funds, insurance companies, and foreign investors. We embed these estimated demand functions in an equilibrium model of the U.S. Treasury market where risk-averse strategic arbitrageurs interact with preferred habitat investors and the Federal Reserve. Our tractable model accommodates cross-elasticities across different maturities, macroeconomic dynamics, and both conventional and unconventional monetary policies. We quantify arbitrageurs' risk aversion using data on dealers' and hedge funds' Treasury positions. We find (1) that the Treasury market is elastic because of low estimated arbitrageur risk aversion that significantly weakens demand impact; (2) a positive term premium response to monetary policy tightening due to high estimated cross-elasticities, rationalizing excess sensitivity of long rates; (3) a weak effect of Fed purchases on bond yields unless the Fed credibly commits to a persistent expansion of its balance sheet.

Keywords: Treasury demand; financial intermediaries; arbitrage; monetary policy; quantitative easing.

We thank Daniel Andrei (discussant), Adrien d'Avernas (discussant), Darrell Duffie, William Diamond, Francois Gourio, Erica Jiang, Arvind Krishnamurthy, Moritz Lenel, Jane Li, Tom Sargent, Selale Tuzel, and seminar and conference participants at the Federal Reserve Bank of San Francisco, BI-SHoF Conference on Asset Pricing and Financial Econometrics, Junior Valuation Workshop, Zurich Quantitative Macroeconomics Workshop, UBC Summer Finance Conference, USC, Federal Reserve Bank of Chicago, and Stanford Junior Macro-Finance Conference for very useful comments and discussions. We gratefully acknowledge financial support from NBER Financial Market Frictions and Systemic Risks initiative. We thank Winston Chen for excellent research assistance. Views expressed are those of the authors and do not necessarily reflect official positions of DNB or the Eurosystem.

*Marshall School of Business, University of Southern California & DNB; kjansen@marshall.usc.edu

†Marshall School of Business, University of Southern California; liwenhao@marshall.usc.edu

‡Marshall School of Business, University of Southern California & CEPR; lukas@marshall.usc.edu

1. Introduction

The U.S. Treasury market is a major backbone of the financial system, influencing the execution and transmission of monetary and fiscal policies, affecting global investment flows, and serving as a benchmark for pricing other financial instruments. Most recently, the U.S. Treasury market has taken center stage in the swift policy responses to both the global financial crisis and the COVID-19 pandemic. In response to the Great Recession, for example, the Federal Reserve aggressively lowered short-term rates, while it targeted the long-end of the yield curve by purchasing long-term bonds in the context of quantitative easing (QE) policies. By contrast, the Fed announced unwinding some of the Treasury positions and selling bonds by means of quantitative tightening (QT) starting in 2017, and then again in 2022. The impact of such interventions in the Treasury market critically depends on the extent to which bond investors are willing to accommodate such sales or purchases, given their risk capacity and mandates, and the extent to which they are willing to substitute alternate securities. Against this backdrop, we ask: Who are the investors that are willing to absorb Treasuries? How elastic is investor demand for Treasuries? What is the role of the Federal Reserve and that of arbitrageurs in the Treasury market?

In this paper, we propose to tackle these questions by combining insights from two influential literatures: the burgeoning literature on demand-based asset pricing, in the spirit of Koijen and Yogo (2019), and the preferred habitat view of the term structure of interest rates as in Vayanos and Vila (2021). We start by linking investors' Treasury holdings to observed Treasury characteristics and macro dynamics to trace out a set of empirically tractable demand curves based on a novel granular dataset on Treasury portfolio holdings covering most of the market, including commercial banks, mutual funds, and insurance companies, and foreign investors, among others. Alongside, we estimate the Federal Reserve's demand curves and link them to its monetary policy stance. Next, we embed these estimates into a dynamic equilibrium model to assess the pricing implications by introducing a risk-averse arbitrageur that clears the market. This framework then allows us to examine yield effects coming from investors' demand shocks, changes in macroeconomic conditions, government debt supply, and the monetary policy stance through counterfactual analysis from an equilibrium perspective. Our model accommodates cross-elasticities across different maturities, as well as conventional and unconventional monetary policy. Critically, our granular dataset allows us to identify and quantify arbitrageur risk-aversion based on dealers' and hedge funds' holdings.

Our results provide a novel perspective on the dynamics of Treasury yields and inform the debate regarding the effects of policies implemented through the U.S. Treasury market. Empirically, our dataset reveals that major investors in Treasuries, as well as the Fed, behave akin to preferred habitat investors and exhibit downward-sloping demand curves, and notably, exhibit strong cross-substitution across maturities. These empirical estimations are incorporated into our quantitative equilibrium model, which is estimated to fit the historical evolution of Treasury yields and the average holding of arbitrageurs. We have three main findings. (1) The Treasury market is elastic because arbitrageurs have a low estimated risk aversion that significantly weakens demand impact, but a spike of arbitrageur risk aversion can cause a sharp increase of long-term yields and significantly larger price impact of demand shocks. (2) Term premia rise in response to a monetary policy tightening, since preferred habitat investors exhibit high estimated cross-elasticities so that they reduce long-term bond positions accordingly, forcing arbitrageurs to absorb them and bear more risks. This is in sharp contrast to Vayanos and Vila (2021) and rationalizes the excess sensitivity of long rates to monetary policy shocks. (3) The effects of Fed purchases on bond yields are weak unless the Fed credibly commits to a persistent expansion of its balance sheet, thereby rationalizing a slow unwinding of the Fed's unconventional monetary policies in the context of QT. Throughout, we provide analytical and quantitative results, and a tractable model that can easily be taken to the data.

As a first step, we create a rich novel dataset on Treasury holdings across a wide range of institutions, such as insurance companies, mutual funds, banks, broker-dealers, foreign investors, and the Federal Reserve, among others, at the security or maturity bucket level. Our dataset covers over 70% of the total Treasury amount outstanding at any given point in time over the 2010Q1-2022Q4 period (Figure 1). We classify preferred habitat investors as commercial banks, insurance companies and pension funds (ICPFs), money market funds (MMFs), mutual funds, and foreign investors, while the arbitrageurs in our setting are the broker-dealers and hedge funds (PDHFs), mainly for two reasons: First, as shown by Hanson and Stein (2015) and Du et al. (2023b), broker-dealers and hedge funds behave as the opposite of yield-seeking investors, taking lower positions when the term spread is higher. Second, broker-dealers and hedge funds generally have better access to a wider range of trading instruments and platforms, allowing them to deploy sophisticated arbitrage strategies[1], which other institutions, especially heavily regulated sectors such as pension funds, insurance companies, and banks, could not easily implement.

While we consider the Fed separately, we find empirically that it behaves similarly to the preferred habitat investors. The preference for maturity is evident in the data: MMFs hold a large amount of the total Treasury outstanding with maturities below one year, while insurance companies, mutual funds, and the Fed have a greater demand for longer maturities. Building on the instrument in Koijen and Yogo (2020) and Fang et al. (2022), we identify own and cross yield sensitivities at the investor-level based on a panel dataset of Treasury holdings across maturity buckets and time.

[1]For example, broker-dealers and hedge funds take significantly negative net positions in Treasuries at certain periods, while we do not observe negative Treasury positions for other investors.

Our empirical approach adds to the extant demand-based asset pricing literature in three ways. First, we incorporate cross elasticities in our demand estimation. Treasuries of different maturities are substitutes in providing liquidity to investors, and naturally, there are significant cross-substitution effects. Indeed, coefficients on other-maturity yields have the opposite sign (i.e., a lower demand for current maturity when other-maturity yields are high), an empirical regularity borne by most preferred-habitat investors.

Second, we include macroeconomic variables that are relevant to the dynamics of real interest rates, inflation, and monetary policy, which are important state variables driving investors' expectation of future nominal rates.

Third, we estimate a dynamic demand system that leverages information from both the cross section and time series. This approach addresses the finding in Van der Beck (2022) that changes in portfolio holdings contain valuable information that cannot be fully captured by examining only the cross section.

With our estimated demand functions and dynamics of macro variables, we embed them into an equilibrium model of the Treasury market, in the spirit of the term structure framework of Vayanos and Vila (2021). The model is discrete-time and infinite horizon with three types of agents: preferred habitat investors whose demand depends on asset prices, macroeconomic factors, and latent demand shocks, a Fed that sets monetary policy and QE policy, and a representative risk-averse arbitrageur that optimizes over a mean-variance objective. We lump demand from preferred habitat investors and the Fed QE rule into an aggregate "preferred-habitat demand" since they carry similar functional forms.

To capture the rich economics in the Treasury market, our model features three key innovations. First, we allow for cross-substitution directly in the preferred habitat investor demand.[2] This is an important force that generates a positive reaction of term premia to a monetary policy tightening, contrasting the negative reaction in the standard setup of Vayanos and Vila (2021) without cross elasticity.[3] Second, we include a monetary-policy rule that depends on macroeconomic dynamics rather than treating the short-term interest rate as purely exogenous on its own. This provides a reasonable estimation of monetary shocks and also allows us to more comprehensively capture how macro variables, such as inflation and GDP gap, affect the demand and pricing of Treasuries. Third, we incorporate latent outside assets held by the arbitrageurs, to add the element of realism that prices of risk are not entirely driven by arbitrageur's Treasury portfolio. We let the data inform us to what extent a Treasury supply or demand shock triggers a change in risk premia.

[2] See Chaudhary et al. (2022) for estimations of cross elasticities in the corporate bond market, and An and Huber (2024) for estimations of cross elasticities in the currency market.

[3] Kekre et al. (2024) introduces arbitrageur's wealth effect to Vayanos and Vila (2021) and also generates a positive reaction of term premium to monetary policy tightening. We do not incorporate such a channel since the nonlinearity due to the wealth effect causes numerical challenges that are beyond our paper.

To gain intuition, we derive theoretical implications in a simplified version of the full quantitative model. First, we show that the monetary policy rate plays a dominant role for short-maturity Treasuries, while macro shocks and latent demand shocks become more important for long-maturity Treasuries. Second, we find that a higher arbitrageur risk aversion increases price impact and equilibrium yields, so it is crucial to estimate a reasonable risk aversion. Third, we find that higher cross elasticity of preferred-habitat demand lowers equilibrium Treasury prices. Fourth, the model implies that whether Treasury risk premium positively or negatively responds to monetary policy shocks depends on the magnitude of cross elasticity. Finally, QE reduces long-term yields.

Guided by the theoretical predictions, we proceed to quantify the model based on our empirical estimates. We estimate the model by minimizing the model fitting error on the dynamics of Treasury yields and the difference between the model versus data on average arbitrageurs' Treasury holding. The latter carries a high weight, so the average holding is accurately matched. We implement this procedure by modifying the algorithm used in the standard affine term structure literature (Duffie and Kan 1996; Duffee 2002; Joslin et al. 2011). The key difference is that the pricing kernel is endogenously determined by market clearing, adding another layer of complexity. Critically, arbitrageur risk-aversion is mainly identified by matching average holdings by arbitrageurs, which in the data include dealers and hedge funds.

Using the model, we explore the drivers of Treasury pricing by analyzing contributions of the monetary policy rate, macroeconomic variables, and latent demand. We use the Shapley value of R^2 as the measure of the relative contribution of different factors. We find that the monetary policy rate is the dominant driver of short-maturity Treasury yield, but it becomes dominated by macroeconomic fluctuations and latent demand shocks at the long-maturity end. The importance of latent demand shocks is less than 5% for short-maturity yields but becomes around 20% for long-maturity yields.

In view of our estimates of a low risk aversion, we find that the price impact of a latent demand shock at short maturities (below one year) is much larger if we remove arbitrageurs from the model. The reason is that arbitrageurs aggressively arbitrage between a short-term interest rate tightly controlled by the Fed and other short-maturity Treasuries. This arbitrage force becomes weaker at longer maturities due to larger risks. On average, the presence of arbitrageurs reduces the price impact of latent demand shocks by more than 100 times. Moreover, in terms of aggregate market elasticity, we find that a \$1 billion dollar extra latent demand of the overall Treasury market increases total Treasury valuation by \$0.27 billion, indicating a multiplier of 0.27. In comparison, the multiplier is 5 in the equity market (Gabaix and Koijen 2021) and 3.5 at rating-level corporate bond markets (Chaudhary et al. 2022). Nevertheless, absent from arbitrageurs in the Treasury market, Treasury-market multiplier becomes larger than that of equity and corporate bonds. Therefore, the Treasury market is elastic in the presence of arbitrageurs but inelastic otherwise.

Next, we evaluate the impact of monetary policy shocks. In Vayanos and Vila (2021), monetary policy tightening leads to a lower risk premium, because the higher interest rates lead to larger preferred-habitat investor holdings and smaller arbitrageur holdings, for which the arbitrageur requires less compensation. In sharp contrast, our results are the opposite because of a significant cross elasticity revealed by the data. This high cross elasticity reduces preferred-habitat investors' holdings of each maturity bucket in response to higher yields from other maturities, forcing arbitrageurs to hold more Treasuries and charge a higher risk premium. Our results are consistent with the literature that empirically identifies term premium responses to monetary policy shocks (Hanson and Stein 2015; Bauer et al. 2023). We show that once we shut off cross elasticities, our quantitative model generates the opposite result, and therefore, we confirm that cross elasticities are the key to resolving the puzzle in Vayanos and Vila (2021).

Finally, we study the response of the Fed balance sheet to demand shocks, and the impact of QE. We find that the Fed is the primary absorber of permanent demand shocks of long-term Treasuries. Quantifying the impact of QE, we discover that if the purchase is expected to be transient, the impact on Treasury yields is little. However, the response of Treasury yields becomes much more prominent if investors expect QE to be a permanent demand shift of the Fed. Therefore, we conclude that the impact of Fed purchases on bond yields is weak unless the Fed credibly commits to a persistent expansion of its balance sheet. Our results allows for more granular analysis of QE policies and contribute to a growing literature that quantifies the impact of QE (Krishnamurthy and Vissing-Jorgensen 2011; d'Amico et al. 2012; Swanson 2021; Selgrad 2023; Jiang and Sun 2024; Haddad et al. 2024).

Related Literature

Our paper contributes to a growing literature that analyzes the granular asset demand in fixed-income markets, building on the seminal work by Koijen and Yogo (2019). Specifically, Bretscher et al. (2020), Chaudhary et al. (2022), Siani (2022), and Darmouni et al. (2022) apply the demand system to corporate bond markets, Fang et al. (2022) to global government bond markets, Koijen et al. (2021) to the euro area government bond market, Jansen (2022) to the Dutch government bond market, and Jiang et al. (2022) to international bond and currency markets. Allen et al. (2020) analyze the demand of T-bill auctions and find that auction format matters for portfolio allocations. Doerr et al. (2023) and Stein and Wallen (2023) provide granular analysis on the demand of money-market funds for near-money assets. Closest to ours, Eren et al. (2023) apply the demand system to the overall U.S. Treasury market using Flows of Funds data. Consistent with their study, we also find that investment funds and banks are more price elastic than ICPFs within the U.S. Treasury market. We contribute to this literature by using granular data on U.S. Treasury holdings by different institutions including the Fed, and analyzing cross-elasticities in the U.S. Treasury market.

Furthermore, our paper is related to the preferred habitat view of the term structure of interest rates, e.g., Culbertson (1957), Modigliani and Sutch (1966), Guibaud et al. (2013), Greenwood and Vayanos (2014), and Vayanos and Vila (2021). Recent papers start to build a tighter connection between data and theory. Droste et al. (2021) identify demand shocks from Treasury auctions and calibrate the model in a New Keynesian framework to study the impact of QE. Hanson et al. (2024) quantify the demand and supply shocks in the interest-rate swap market. Khetan et al. (2023) leverages more detailed data on interest-rate swaps and find a high level of segmentation. Bahaj et al. (2023) utilize transaction-level data on UK inflation swaps to quantify a model of inflation risks. Our contribution is to build and estimate a quantitative version of Vayanos and Vila (2021) in the Treasury market that accounts for empirically estimated demand functions and arbitrageurs' Treasury holdings.

Our estimation of preferred-habitat demand is consistent with the hypothesis of "yield-oriented investors" in Hanson and Stein (2015), where investors consider not only long long-term yields but also short-term yields for long-term bond investment. We both theoretically and quantitatively confirm the rationale in Hanson and Stein (2015) that cross-substitution drives the positive term premium response to monetary policy tightening. This also addresses a broad literature that shows risk premium overall rises with monetary policy tightening (Bernanke and Kuttner 2005; Gertler and Karadi 2015; Bekaert et al. 2013; Kekre et al. 2024).

Our paper is also related to the recent literature on the specialty of U.S. government debt. Krishnamurthy and Vissing-Jorgensen (2012) show that there is a downward-sloping aggregate demand curve for the convenience provided by Treasuries. The literature shows that Treasury convenience yield is closely connected to financial crises (Del Negro et al. 2017; Li 2024), monetary policy (Nagel 2016; Drechsler et al. 2018; Diamond and Van Tassel 2021), exchange rates (Jiang et al. 2021), pricing of stocks (Di Tella et al. 2023), hedging properties of Treasuries (Acharya and Laarits 2023), banking (Diamond 2020; Li et al. 2023; Krishnamurthy and Li 2023), and financial regulation (Payne et al. 2022; Payne and Szőke 2024). Jiang et al. (2024b) provide evidence that the aggregate U.S. government debt is overvalued relative to common measures of fundamentals. Brunnermeier et al. (2024) show that Treasuries insure against idiosyncratic risks, which can be a resolution to the valuation puzzle. Jiang et al. (2024a) characterize government debt valuation and debt capacity in a dynamic model with hedging and convenience. We contribute to the above literature by unpacking the demand of Treasuries and sources of its variations across investors, and relating it to macroeconomic factors and monetary policy dynamics.

Finally, arbitrageurs are important in our analysis, in the same spirit as a growing literature that focuses on the role of financial intermediaries (He and Krishnamurthy 2013; Adrian et al. 2014; He et al. 2017; Du et al. 2018; Haddad and Sraer 2020; Wallen 2020; Jermann 2020; Haddad and Muir 2021; Fang and Liu 2021; Kargar 2021; Favara et al. 2022; Du et al. 2023a; Diamond et al. 2024; An and Huber 2024). d'Avernas and Vandeweyer (2023) and d'Avernas et al. (2023) provide theories of how different types of intermediaries together with the central bank affect the Treasury market dynamics. Duffie et al. (2023) uses dealer-level data on Treasury holdings to show that dealer balance sheet utilization is important for Treasury pricing. Du et al. (2023b) quantitatively show that balance sheet frictions of intermediaries are important in pricing Treasuries. A key contribution relative to this literature is that we cover the majority of the Treasury market beyond intermediaries, and explicitly link how the intermediary pricing kernel is related to its intermediation activities in the Treasury market.

2. Data

This section describes the various data sources that we use in our study, the merger of these data sets, and descriptive statistics.

2.1. Holdings Data Sources

The standard dataset for U.S. Treasury holders is from the Flow of Funds of the Federal Reserve, which has been used for Treasury demand analysis in Krishnamurthy and Vissing-Jorgensen (2007) and Eren et al. (2023). However, although the data is at the sector level, this dataset only contains total Treasury holdings across all maturities and therefore does not allow for more granular analysis, especially cross-substitution across maturities. Therefore, we collect several data sources for each investor type and a summary of data sources is provided in Table 1.

We now turn to a detailed description and cleaning procedure of each data source.

A. Banks - CALL Reports

We obtain banks' holdings of U.S. Treasuries at the maturity bucket level from CALL reports. CALL reports are regulatory filings required for all U.S. banks and include detailed information on a bank's assets, liabilities, income, and expenses. The CALL reports are filed on a quarterly basis and cover the period from the first quarter of 1976 to the end of 2022. Banks report their aggregate U.S. Treasury holdings and their holdings in different maturity buckets of U.S. Treasuries and U.S. Agency bonds combined. The maturity buckets are: $\tau < 3M$, $3M \leq \tau < 1Y$, $1Y \leq \tau < 3Y$,

Table 1. Summary of Data Sources

Investor Type	**Data Source**	**Frequency**	**Period**	**Detail**
Banks	CALL Reports	Quarterly	1976Q1-2022Q4	Maturity bucket
ETFs	ETF Global	Daily/Monthly	2012M1-2022M12	Security
Fed	Federal Reserve	Weekly	2003W1-2022W52	Security
Foreign Holders	Public TIC	Monthly	2010M1-2022M12	T-bill/non T-bill
Hedge Funds	CFTC	Weekly	2006W26-2022W52	Maturity bucket
Insurers and Pension Funds	eMAXX	Quarterly	2010Q1-2022Q4	Security
Money Market Funds	IMoneyNet	Monthly	2011M8-2022M12	Security
Money Market Funds	Flow of Funds	Quarterly	1993Q1-2022Q4	T-bill/non T-bill
Mutual Funds	Morningstar	Monthly/Quarterly	2000M1-2022M12	Security
Primary Dealers	Federal Reserve	Daily	1998D28-2022D365	Maturity bucket

$3Y \leq \tau < 5Y$, $5Y \leq \tau < 15Y$, $\tau \geq 15Y$. To obtain their allocation to U.S. Treasuries for different maturities, we assume that the fraction of Treasuries versus Agency bonds is fixed across maturities at a given point in time. Hence, at each point in time, we multiply the total maturity bucket holdings by the fraction of Treasuries relative to the sum of Treasuries and Agency bonds.

B. ETFs - ETF Global

We obtain the holdings of U.S. Exchange Traded Funds (ETFs) at the security (CUSIP) level from ETF Global. ETF Global contains extensive coverage of securities held by U.S. ETFs and in our analysis we focus on fixed-income funds. Funds either report daily or monthly, and to maintain consistency with the other datasets we use data at quarter ends. As U.S. ETFs only hold a small fraction of U.S. Treasuries outstanding, we merge them with the U.S. mutual fund sector.

C. Fed - Federal Reserve

The Federal Reserve System Open Market Account (SOMA) reports security holdings that are acquired through open market operations by the Fed. These data are obtained through the website of the Federal Reserve Bank of New York.[4] The holdings are at the security (CUSIP) level and reported on a weekly basis since the start of 2003.

D. Foreign Investors - Public TIC

From January 2010 to September 2011, we obtain monthly aggregate U.S. Treasury holdings by foreign investors from TIC Form S. As Tabova and Warnock (2021) point out, there are issues with TIC S that lead to substantial misreporting of U.S. Treasury holdings, but only after 2012.

[4] `https://www.newyorkfed.org/markets/soma-holdings`

Hence, we obtain monthly aggregate U.S. Treasury holdings from the public TIC Form SLT after September 2011, the starting date of this form. As of September 2011, TIC also provides a breakdown of the total amount held in T-bills versus non-Tbills.

Moreover, to avoid double counting, we subtract from the TIC data the holdings of foreign mutual funds that we obtain through Morningstar.

E. Hedge Funds - CFTC

We infer the hedge fund Treasury holdings from their futures positions, because hedge funds usually engage in the Treasury-futures basis trade on a hedged basis. Hence, the Treasury position is the opposite of the futures position (Du et al. 2023b). Hedge funds report their futures positions for conventional maturity buckets on a weekly basis since mid 2006. The deliverable maturity buckets are: $1\frac{3}{4}Y \leq \tau \leq 2Y$ (2-Year T-Note), $4\frac{1}{6}Y \leq T \leq 5\frac{1}{4}Y$ (5-Year T-Note), $6\frac{1}{2}Y \leq \tau \leq 7\frac{3}{4}Y$ (10-Year T-Note), $9\frac{5}{12}Y \leq \tau \leq 10Y$ (Ultra 10-Year T-note), $15Y \leq \tau \leq 25Y$ (T-Bond), $25Y \leq \tau \leq 30Y$ (Ultra T-Bond). The net futures position for each deliverable maturity is defined as a hedge fund's long position minus its short position. Therefore, the Treasury position of a hedge fund is the opposite of the sum of its net futures positions across the deliverable maturity buckets.

F. Insurers and Pension Funds - eMAXX

eMAXX provides a comprehensive coverage of fixed income holdings of institutional investors at the security (CUSIP) level. The database predominantly covers the holdings of insurance companies, mutual funds, and pension funds (Becker and Ivashina 2015; Bretscher et al. 2020). We only use the data on insurance companies and pension funds, and rely on Morningstar for mutual funds. Due to the voluntary nature of reporting by pension funds, the coverage of pension funds in eMAXX is limited, unlike the mandatory reporting by insurance companies. Additionally, we focus on the U.S. eMAXX database, which covers the holdings of North American investors. The holdings data are quarterly and cover the period from the first quarter of 2010 to the end of 2022.

G. Money Market Funds - IMoneyNet and Flow of Funds

IMoneyNet contains a wide coverage of asset holdings (predominately fixed income and cash) by money market funds (MMFs) at the security (CUSIP) level. We focus on both holdings reported by MMFs domiciled in the U.S. as well as on their offshore holdings. The holdings are reported on a monthly basis since August 2011.

To obtain a longer history of U.S. MMF holdings, we use Flow of Funds (FoF) data from the Federal Reserve. Using our security-level database, we verify that on average 99.6% of MMF holdings are in either T-bills or U.S. Treasuries with remaining time to maturity below 1 year. Hence, we can reasonably assume that MMF Treasury holdings reported in FoF have remaining maturities below 1 year.

H. Mutual Funds - Morningstar

We obtain holdings data on domestic and foreign mutual funds and foreign ETFs from Morningstar, Inc. The funds report all their positions including stocks, bonds, and cash at the security (CUSIP) level. We focus on both fixed-income and allocation funds. Funds either report monthly or quarterly, and to maintain consistency across the funds and other data sets we use data at quarter ends. Figure 12 reports the aggregate holdings in USD (trillions) over time. These aggregates align closely with the numbers reported in Maggiori et al. (2020).

I. Primary Dealers - Federal Reserve

To maintain transparency of primary dealers trading activities, their daily positions are made available through the website of the Federal Reserve Bank of New York.[5] Primary dealers report their holdings for conventional maturity buckets since early 1998. However, the specific maturity buckets reported change over time. The time frames with the same reporting standards are: January 1998 to June 2001, July 2001 to March 2013, April 2013 to December 2014, January 2015 to December 2021, and from January 2022 onwards. Generally, more recent data reports finer maturity buckets. To be consistent across time, we treat July 2001 to March 2013 as the baseline and aggregate the maturity buckets of subsequent periods to match that of this time frame. The positions in coupon paying bonds and notes (Coupons) are separately reported from inflation-linked bonds (TIPS). The final maturity buckets are: T-Bills, Coupons: $\tau \leq 3Y$, $3Y < \tau \leq 6Y$, $6Y < \tau \leq 11Y$, $\tau > 11Y$, TIPS: $\leq 2Y$, $2Y < \tau \leq 6Y$, $6Y < \tau \leq 11Y$, $\tau > 11Y$. As an example, for January 2022 and on, the maturity bucket $6Y < \tau \leq 11Y$ for Coupons is obtained by summing the net positions of $6Y < \tau \leq 7Y$ and $7Y < \tau \leq 11Y$.

2.2. Macro Data

We obtain four macro variables from the Federal Reserve Economic Data (FRED) that are drivers of monetary policy, real interest rates, inflation dynamics, and total government debt supply in classical macroeconomic models. First, we include GDP gap and core inflation to capture the reaction of monetary policy to macroeconomic dynamics. These variables are also important

[5] https://www.newyorkfed.org/markets/counterparties/primary-dealers-statistics

in determining the general equilibrium real rates and inflation dynamics. Second, we include debt/GDP to capture the dynamics of government debt and its connection to other macroeconomic variables, including the interaction with inflation according to the fiscal theory of price level. Third, we include credit spread, which is the nexus between financial markets and the macro-economy as shown by Gilchrist and Zakrajšek (2012).

2.3. Data Aggregation

First, to maintain consistency across datasets, we analyze data at the quarterly frequency from 2010Q1 to 2022Q4 when aggregating across all maturities and from 2011Q3 to 2022Q4 when analyzing holdings at the maturity-bucket level.

For the data sources in Table 1 that are at the security level, we have the corresponding CUSIP identifiers that we use to match the holdings data with the CRSP U.S. Treasury Database. The CRSP U.S. Treasury Database contains detailed bond-level information on U.S. Treasuries, including bond yields, prices, bond type, coupon rate, maturity date, issue date, and issuance size. We use the bond prices to convert nominal holdings to market values. For the sectors that report at a more aggregate level (banks, foreign investors, hedge funds, and primary dealers), we use their reported market value holdings directly.

We focus on three maturity buckets: $\tau < 1Y$, $1Y \leq \tau < 5Y$, and $\tau \geq 5Y$. This choice is determined by the data granularity limitations for some sectors. For investors that report at the CUSIP level it is straightforward to divide their holdings in the respective maturity buckets. For banks, we aggregate maturity bucket $\tau < 3M$ and $3M \leq \tau < 1Y$ to obtain the first bucket, $1Y \leq \tau < 3Y$ and $3Y \leq \tau < 5Y$ for the second bucket, and $5Y \leq \tau < 15Y$ and $\tau \geq 15Y$ for the third bucket. We follow a similar approach for the primary dealers and hedge funds. For foreign investors, we only know the fraction that is held in T-bills versus non T-bills. To allocate the foreign holdings to different maturity buckets, we first multiply the T-bill holdings by one divided by the fraction of TAO in maturity bucket 1 that is in T-bills for each point in time. The reason is that on average only 60% of TAO in maturity bucket 1 are T-bills, while the remaining 40% are bonds and notes with remaining time to maturity below 1 year.[6] We then subtract the additional fraction we attribute to maturity bucket 1 from the total non T-bill holdings to compute the total holdings in the remaining maturity buckets. To further determine the fraction in maturity bucket 2 versus 3, we choose the fraction such that the average duration of the foreign investors' Treasury portfolio is consistent with Tabova and Warnock (2021) at each point in time.

Figure 1 shows the fraction of the total amount outstanding of U.S. Treasuries by investor

[6] Our results do not depend on this assumption. We find similar results, both qualitatively and quantitatively, when we assume that T-bills are the only securities held in bucket 1.

type from 2010Q1 to 2022Q4. On average, our dataset contains about 78% of the holders of U.S. Treasuries. Based on FoF data, the remaining 22% consists of U.S. households ($\approx 11\%$), pension funds ($\approx 5\%$), local governments ($\approx 4\%$), and non-financial corporations ($\approx 2\%$). Figure 2 breaks down the holdings at the maturity bucket level, starting in 2011Q3, the first quarter we know the T-bill versus non T-bill split for foreign investors. Some striking patterns emerge. First, MMFs are only active in maturity bucket 1 and held between 10-35% of the short-term Treasuries outstanding. At the other end of the spectrum, ICPFs barely hold short-term Treasuries, but hold around 5% of the Treasuries with maturities beyond 5 years. The Fed holds substantially more of the intermediate and long-term bonds outstanding as opposed to short-term bonds. Mutual funds hold little short-term bonds, but are equally spread among maturity bucket 2 and 3. The net positions of primary dealers and hedge funds are mostly positive at longer maturities.

3. Empirical Results

3.1. Empirical Methodology

We start our empirical analysis by estimating a demand system for each sector i. Inspired by our model specified in Section 5, we estimate the demand for U.S. Treasuries using the following regression:

$$Z_t^i(m) = \theta_0^i + b_1^i y_t(m) + b_2^i y_t(-m) + (b_3^i)'\mathbf{x}_t(m) + (b_4^i)'\mathbf{Macro}_t + u_t^i(m), \tag{1}$$

where $Z_t^i(m)$ equals the total Treasury market value holding of sector i in maturity bucket m at time t (in billions), $y_t(m)$ is the yield for maturity bucket m, $y_t(-m)$ equals the weighted-average yield of the other maturity buckets, $\mathbf{x}_t(m)$ is a vector of value-weighted bond characteristics for maturity bucket m: coupon, maturity bucket fixed effects, bid-ask spread, and $\mathbf{Macro}_t$ equals a set of macro variables, including GDP gap, debt/GDP, core inflation, and credit spread. We also add the repo spread to the set of bond characteristics to control for the fact that some investors reduce U.S. Treasury holdings and enter the repo market to obtain indirect Treasury exposure in response to a rising repo rate relative to Treasury rates.

We focus on dollar value of holdings rather than portfolio weights, because dynamics in total portfolio demand is crucial for the term structure of interest rates – modeling only portfolio weights is not sufficient. Moreover, our model in Section 5 indicates that market values are what investors should care about, so our specification in (1) has a direct mapping to our dynamic quantitative model. Since discount rates on Treasuries during the sample period we analyze are on average low, our empirical estimations are similar if we replace all market values with face values of Treasury holdings.

To ensure that demand is stationary, we standardize $Z_t^i(m)$ by potential GDP. Specifically, we divide $Z_t^i(m)$ by the ratio of nominal potential GDP at t over the baseline value of nominal potential GDP at the beginning of our sample period. We use nominal values so that the scaling adjusts for the inflation effect. Moreover, using a GDP adjuster rather than just inflation ensures that we account for the growing scale of the economy. Finally, we use nominal potential GDP rather than nominal GDP to avoid cyclical fluctuations in nominal GDP that causes mechanical correlations among the variables due to the scaling. Our empirical results will be similar if we simply adjust for an exponential growth trend with the constant growth rate matching the overall economic growth during our sample period.

We provide summary statistics for the set of bond characteristics and macro variables in Table 2 and the correlation table is in Appendix Table 9. We denote Treasury maturity as τ and divide Treasuries into three maturity buckets: $\tau < 1Y, 1Y \leq \tau \leq 15, \tau > 15$, and denote these maturity buckets as $m \in \{1,2,3\}$. The choice for these three maturity buckets is motivated by two reasons: First, this division is a common set across portfolio holding data that we have for different investors. Second, we need sufficient cross-sectional variation across maturity buckets to apply our instrument. Using more granular maturity buckets complicates identification due to a reduction in variation across buckets.

We control for both own yield and other yield to account for the substitution effect between Treasuries of different maturities. If we only include own yield, but investors decrease demand when the yield in the other maturity buckets is higher, we would uncover a coefficient that is downward biased. The reason is that own yield and other yield are correlated, while demand increases if own yield goes up, but decreases when other yield goes up. Hence, when not accounting for other yield, b_1^i picks up both the positive and negative effect, leading to a coefficient that is biased towards zero.[7]

In our specification, we assume that the macro variables are exogenous to investors, as in Fang et al. (2022) and Koijen and Yogo (2020). That is, investor (latent) demand does not contemporaneously affect macro variables. In addition, we also assume that bond characteristics, except for yields, are exogenous to latent demand.

[7] In Appendix Table 13, we indeed show that the coefficient on own yield is attenuated closer to zero when not accounting for other yield.

Instrument

We could estimate the demand system specified in Equation (1) by GMM if it satisfies the moment condition:

$$\mathbb{E}[u_t^i(m)|y_t(m),\mathbf{x}_t(m),\mathbf{Macro}_t]=0 \tag{2}$$

The concern with this moment condition is that the error term may not be orthogonal to yields. For instance, if sectors have a large demand for Treasuries unrelated to bond characteristics or macro variables, then this latent demand is likely to also affect the yield. As such, we need an instrument for bond yields.

We build on the instrument used in Koijen and Yogo (2020) and Fang et al. (2022) and use the following three step procedure. First, we estimate demand for each investor type as in Equation (1), but excluding the yield. We then in a second step extract the predicted values $\hat{Z}_{it}(m)$. We also follow step (1) and (2) for the market value of Treasury supply at each maturity bucket, whereby we regress it on the FFR and macro variables, consistent with the specification of our US Treasury model introduced in Section 5. In a third step, we impose market clearing and extract the imposed yield that sets the implied demand equal to the implied market value of supply:

$$\sum_i \hat{Z}_{it}(m)=\frac{\hat{S}_t(m)}{(1+\tilde{y}_t(m))^{\tau(m)}}, \tag{3}$$

where $\hat{S}_t(m)$ is the predicted nominal value of the debt outstanding of maturity bucket m, and $\tau(m)$ the corresponding maturity. We take $\tau(m)$ as the average bond duration for maturity bucket m. We then extract pseudo yield $\tilde{y}_t(m)$ that clears the market at each point in time t and use it as instrument for the actual yield $y_t(m)$:

$$I_t(m)=\tilde{y}_t(m) \tag{4}$$

In summary, the idea behind the instrument is that the pseudo yield isolates the exogenous component of the yield. This instrument is valid under the identifying assumption that bond characteristics and macro variables are exogenous to investor latent demand, as we assume throughout. By construction, the instrument is independent of investors' latent demand in the second stage (Equation (1)) so that the exclusion restriction in (5) is satisfied. More formally, we can weaken moment condition (2) to:

$$\mathbb{E}[u_t^i(m)|I_t(m),\mathbf{x}_t(m),\mathbf{Macro}_t]=0. \tag{5}$$

3.2. Demand Functions of Preferred Habitat Investors

The first stage estimates of the demand system are summarized in Appendix Table 11. The corresponding Kleibergen-Paap statistic to test for weak instruments is 10.9, above the threshold of 10 for rejecting weak instruments (Stock and Yogo 2005).[8]

Table 3 shows the results using the IV methodology outlined in the previous section.[9] We find that all investors have downward sloping demand curves, except for the foreign sector, although the coefficient is insignificant. That is, investors demand more U.S. Treasuries of maturity bucket m when the yield (price) is high (low). In addition, investors load negatively on the yield of other maturity buckets, meaning that their demand for maturity bucket m decreases when the yield for other buckets are high. Generally, we find that other elasticity is slightly higher than own elasticity, but the order of magnitude between the coefficients is similar. This is consistent with the findings in Chaudhary et al. (2022). They find a ratio between cross-elasticity and own-elasticity of close to 1 at the CUSIP level and for portfolios at the rating $\times$ quarter-to-maturity level for corporate bonds, the latter aggregation closely resembling ours. This ratio implies that own and cross-elasticity have the same magnitude, but with opposite sign. This finding is also consistent with Eren et al. (2023), who study the aggregate elasticity of the U.S. Treasury market and find that investors load negatively on the 5Y German yield ('other yield') that has the opposite sign of their loadings on the 8Y U.S. yield ('own yield'), but with similar magnitudes. Finally, Table 12 shows that the coefficients on own and other yield remain virtually unchanged when not controlling for macro variables.

In addition, ICPFs, MMFs, and other U.S. investors increase demand for Treasuries when coupons are high, while the other sectors reduce their demand. ICPFs have a higher demand for Treasuries when the bid-ask spreads are high, that is, when Treasuries are less liquid. Furthermore, ICPFs have a high demand for long-term Treasuries, while other U.S. investors have a lower demand for the longest Treasuries. By means of the investment mandates of MMFs, they only operate in the shortest maturity bucket. Supply negatively affects foreign investor holdings, while it positively affects demand for other sectors, especially for MMFs. Moving to the macro variables, we find that banks, MFs U.S., and MFs ROW increase their demand for Treasuries when the GDP gap is high, while MMFs reduce their demand. This finding is consistent with Fang et al. (2022), who find that most investors increase demand for developed government debt when GDP growth in those countries is high. ICPFs reduce their demand for Treasuries when core inflation is high, while banks increase their demand. Finally, we find that all sectors except for foreign investors

[8] The first stage is the same for all sectors, except MMFs, for which the statistic equals 5.2. The reason is that for MMFs we only have one maturity bucket, because they do not invest in maturities beyond 1 year. As such, the instrument cannot exploit heterogeneity across maturities and we should interpret their result with care.

[9] The results of the OLS estimates are in Appendix 10.

increase demand for Treasuries when supply (measured by debt/GDP) is high.

By looking at holdings in market values, Table 3 does not allow to compare the price elasticities across investor types. As such, we standardize the holdings such that for each sector they have a mean of zero and a standard deviation of one. Figure 3a plots the coefficients on own and other yield for each investor type. Interestingly, mutual funds and MMFs appear to be most price elastic, followed by banks. ICPFs and foreign investors are least price elastic. This finding is broadly consistent with Eren et al. (2023), who also find that banks and investment funds are more price elastic than the ICPF sector.

3.3. Demand Functions of the Fed

For the Fed, we estimate their demand curves maturity-by-maturity bucket. The reason is that the Fed targets the long-end of the yield curve with it's unconventional monetary policies, not the shorter-end. We should therefore expect the Fed to respond to long-term yields only, while we should find them to be rather inelastic at the shorter-end. This is opposite from the preferred habitat investors, where we do not have a strong prior that they have significantly different responses to yields across maturities.

Table 4 summarizes the results. Interestingly, in the long-term bucket, the Fed operates as preferred habitat investors, whereby they increase demand for Treasuries when the yield is high, while they reduce demand for maturity bucket m when the yield in other buckets is high. This revealed behavior is consistent with Fed's policy goals. The goal of QE is explicitly to reduce long-term yields, so the Fed actively increases its balance sheet when long-term yields are high. On the other hand, the Fed achieves a consistency of conventional and unconventional monetary policies by increasing the short-term yields and reducing long-term Treasury holdings at the same time, generating a negative cross elasticity as we discovered in the regressions.

Also as we expected, the Fed becomes entirely inelastic to yields in its demand for short- and medium-term Treasuries. Additionally, like the other preferred habitat investors, their total holdings also on average load positively on total debt supply. The Fed also increases demand for long-term bonds when core inflation is high, while it reduces demand for medium- and long-term bonds when GDP gap is high. Across all maturities, the Fed is not responsive to changes in the credit spread.

Figure 3b shows the relative yield sensitivities of the Fed across maturity buckets. Clearly, the Fed is inelastic with respect to yield in the short- and medium-term bucket, while it's elasticity resembles those of banks in the long maturity bucket.

3.4. Summary

In summary, our estimated demand curves reveal three key findings. First, preferred habitat investors and the Fed have downward sloping demand curves. Second, the total demand from preferred-habitat investors exhibit strong cross substitution. Third, the Fed's demand for short- and medium-maturity is not significantly affected by Treasury yields, but its long-maturity Treasury demand significantly increases with long-maturity Treasury yield but decreases with other yields, similar to the overall behavior of preferred-habitat investors.

4. An Equilibrium Model of the Treasury Market

In this section, we set up a model where strategic arbitrageurs interact with preferred habitat investors and the Federal Reserve in the market for Treasuries. Model dynamics are driven by macroeconomic shocks, monetary policy shocks, and latent demand shocks. After we set up the model, we provide a simplified version that allow us to derive analytical results to obtain intuition regarding the fundamental mechanisms.

To capture the rich economics in the Treasury market, we deviate from Vayanos and Vila (2021) in three aspects. First, we allow for cross-substitution in the preferred habitat investor demand. This is an important force that generates reasonable term premium responses to monetary policy shocks. Second, we include a monetary-policy rule that depends on macroeconomic dynamics rather than treating the short-term interest rate as exogenous. This allows us to quantify the magnitude of monetary policy shocks. Third, we incorporate latent outside assets held by the arbitrageurs, to add the element of realism that risk prizes are not entirely driven by arbitrageur's Treasury portfolio. We let the data speak to the extent that Treasury supply or demand shocks trigger changes in risk premia.

4.1. Model Setup

The model is discrete-time and infinite-horizon. There are four types of agents in the economy: a competitive arbitrageur sector, the Federal reserve, a set of preferred-habitat investors, and the government. We only explicitly model the strategic decisions by arbitrageurs while we capture the behavior of other agents by policy rules. We model the Treasury market explicitly by market clearing. Economic fluctuations are driven by macroeconomic shocks and monetary policy shocks.

Consider bonds of maturities $\tau \in \{1, 2, \cdots, N\}$ that all pay a face value of 1 at maturity. Denote by $P_t^{(\tau)}$ and $y_t^{(\tau)}$, respectively the time-t price and yield of the bond with maturity τ. Define the log price vector as

$$p_t = \left(\log(P_t^{(1)}), \log(P_t^{(2)}), \cdots, \log(P_t^{(N)})\right)'. \tag{6}$$

For simplicity, we denote the yield of an one-period bond as r_t, defined as

$$r_t = -\log(P_t^{(1)}). \tag{7}$$

We consider r_t as directly controlled by monetary policy. All other bond yields and prices are endogenously determined in equilibrium.

The dynamics of the economy is driven by K-dimensional vector of macro factors,

$$\beta_t = (\beta_{1,t}, \beta_{2,t}, \cdots, \beta_{K,t})'. \tag{8}$$

which follows a VAR(1) process,

$$\beta_{t+1} = \bar{\beta} + \Phi(\beta_t - \bar{\beta}) + \Sigma^{1/2}\varepsilon_{t+1}. \tag{9}$$

In the above equation, ε_{t+1} is a K-dimensional vector that follows an i.i.d. standard normal distribution, and Φ is a matrix that determines the long-run dynamics.

We interpret the vector β_t as macro states of the economy that drive the monetary policy stance in equilibrium and also expectations regarding future economic states. Monetary policy depends on contemporaneous economic variables,

$$r_{t+1} = \bar{r} + \phi_r'(\beta_{t+1} - \bar{\beta}) + \rho_r r_t + \sigma_r \varepsilon_{t+1}^r, \tag{10}$$

where ρ_r captures monetary policy inertia, as discussed, for example, in Clarida et al. (2000), and ε_{t+1}^r reflects monetary policy shocks. We assume that monetary policy shocks ε_{t+1}^r are independent from ε_{t+1}, i.e., monetary policy shocks are not subsumed by public information on macro dynamics.

Denote the set of institutions excluding arbitrageurs as $\mathscr{I}$. Sector-i ($i \in \mathscr{I}$) demand for bonds with maturity $\tau \in \{1, \cdots, N\}$ is

$$Z_t^i(\tau) = \theta_0^i(\tau) - \alpha^i(\tau)' p_t - \theta^i(\tau)'\beta_t + u_t^i(\tau), \tag{11}$$

where the parameter $\alpha^i(\tau)$ reflect the price sensitivity to maturity τ itself but also cross-elasticities with maturity $\tau' \neq \tau$. We use "prime" to denote transpose of vectors and matrices, and all vectors are column vectors. We lump the demand for bonds from preferred habitat investors and the Federal Reserve together, and refer to it either as the "non-arbitrageur demand" or "preferred-habitat demand", defined as

$$Z_t(\tau) = \sum_{i \in \mathscr{I}} Z_t^i(\tau). \tag{12}$$

Accordingly, we define $\theta_0(\tau)$, $\alpha(\tau)$, $\theta(\tau)$, and $u_t(\tau)$ as the sums of corresponding values from each sector $i \in \mathscr{I}$. We use column vector forms to express our setup in a more convenient and compact notation. In vector form, we can write (12) as

$$Z_t = \theta_0 - \alpha p_t - \theta \beta_t + u_t, \tag{13}$$

where

$$\theta_0 = (\theta_0(1), \theta_0(2), \cdots, \theta_0(N))' \tag{14}$$

$$\alpha = (\alpha(1), \alpha(2), \cdots, \alpha(N))' \tag{15}$$

$$\theta = (\theta(1), \theta(2), \cdots, \theta(N))'. \tag{16}$$

Thus, α is a $N \times N$ matrix, θ is a $N \times K$ matrix, and θ_0 is a N-dimensional vector.

The unobservable, maturity-specific latent demand shock, $u_t = (u_t(1), u_t(2), \cdots, u_t(N))'$, reflects the non-systematic component of demand shocks. We assume that u_t is i.i.d., with mean zero and a covariance matrix Σ^u.

We can interpret the dependence of habitat-demand on the vector of Treasury prices, reflected by $\alpha(\tau)$, as reflecting either the notion that agents infer other investors' latent demand shocks from prices, or exhibiting convenience demand over Treasuries. Both interpretations naturally suggest investors' own sensitivity to prizes to be negative ($-\alpha(\tau, \tau) < 0$), so that a higher Treasury price (lower yield) implies lower demand. The first interpretation suggests that a high Treasury price might be the result of a positive latent demand, so agents expect the price to drop and thus they lower their current demand (i.e., sell Treasuries). Such beliefs can be rational under certain information structures, but for our purposes, we remain agnostic whether such a behavior is rational or not. The second interpretation suggests that there is fundamental demand for U.S. government debt beyond just financial returns, and thus not captured by the fundamental state variables, leading to a downward-sloping demand curve (e.g., Krishnamurthy and Vissing-Jorgensen). The second interpretation also suggests a natural substitution effect among Treasuries of different maturities, leading to cross-elasticities with opposite signs ($-\alpha(\tau, \tau') > 0$, $\tau' \neq \tau$), i.e., when Treasuries of other maturities become more expensive, the demand for maturity-τ Treasuries increases.

On the supply side, we assume that government supplies Treasuries according to macroeconomic conditions and the monetary policy rate. Accordingly, we specify the aggregate value of government bond supply, or, more precisely, the supply to the public market, i.e. marketable

Treasury securities, as

$$S_t(\tau) = \bar{S}(\tau) + \zeta(\tau)'\beta_t + \zeta_r(\tau)r_t, \tag{17}$$

or, in vector form, as

$$S_t = \bar{S} + \zeta\beta_t + \zeta_r r_t, \tag{18}$$

where

$$\zeta = (\zeta(1), \zeta(2), \cdots, \zeta(K))' \tag{19}$$

is an $N \times K$ matrix. We can interpret Equation (17) as coming from a budget equation of the government, where Treasury supply needs to adjust to the need for government financing and government debt management, which in turn are driven by macroeconomic conditions and the prevailing interest rate.

Next, we setup the strategic decision problem of arbitrageurs. Denote the total return from holding a Treasury of maturity τ as

$$R_{t+1}^{(\tau)} = \frac{P_{t+1}^{(\tau-1)} - P_t^{(\tau)}}{P_t^{(\tau)}}. \tag{20}$$

Accordingly, the total return of one-period Treasury is

$$R_{t+1} = R_{t+1}^{(1)} = \exp(r_t) - 1 \approx r_t, \tag{21}$$

to which we tie our specification of the Fed's monetary policy stance.

We assume that the arbitrageur takes advantage of investment opportunities in bonds with all maturities, with positions $X_t(\tau)$, as well as an outside asset, with position $\tilde{X}_t$. We view modeling outside assets as adding an important element of realism to models in the spirit of Vayanos and Vila (2021), as arbitrageurs risk-bearing capacity in the Treasury market plausibly depends on their positions in other markets.

Accordingly, we model the arbitrageurs' wealth dynamics as

$$W_{t+1} = W_t(1 + R_t) + \sum_{\tau=2}^{N} X_t(\tau)(R_{t+1}^{(\tau)} - R_t) + \tilde{X}_t(\tilde{R}_{t+1} - R_t). \tag{22}$$

We assume that the return of the outside asset is normally distributed and depends on the state of the economy, in that

$$\tilde{R}_{t+1} = \tilde{\phi}'\beta_t + \tilde{\phi}_r r_t + \tilde{\sigma}'\varepsilon_{t+1} + \tilde{\sigma}_r'\varepsilon_{t+1}^r, \tag{23}$$

where $\tilde{\phi}$ is an $K \times 1$ vector and $\tilde{\sigma}$ is an $K \times 1$ vector.

We follow Vayanos and Vila (2021) by assuming that arbitrageurs exhibit mean-variance preferences over wealth. Accordingly, their objective function at time t is

$$\max_{\{X_t^\tau\}_\tau, \tilde{X}_t} E_t[W_{t+1}] - \frac{\gamma}{2} Var_t(W_{t+1}), \tag{24}$$

which arbitrageurs aim to maximize subject to the wealth dynamics specified in (22).

Finally, for each maturity τ, there is a market-clearing condition,

$$Z_t(\tau) + X_t(\tau) = S_t(\tau). \tag{25}$$

We conjecture that there is an affine equilibrium where the log prices of bonds are affine functions of state vector β_t, the short-rate r_t, and the latent demand shock u_t,

$$p_t = A\beta_t + A_r r_t + A_u u_t + C, \tag{26}$$

where $A = (A(1), A(2), \cdots, A(N))'$ is an $N \times K$ matrix, $A_r = (A_r(1), A_r(2), \cdots, A_r(N))'$ is an $N \times 1$ vector, $A_u = (A_u(1), A_u(2), \cdots, A_u(N))'$ is an $N \times N$ matrix, $C = (C(1), C(2), \cdots, C(N))'$ is an $N \times 1$ vector.

4.2. A Simplified Version with Analytical Solutions

To gain intuition regarding the mechanisms at play in the model, we analyze a simplified version of the model in this subsection. In particular, we assume $N = 2$, so there are only two maturities for consideration that represent "short" and "long". We assume that the preferred-habitat demand has a simple structure with the matrix of demand response to price (see Equation (13)) as

$$\alpha = \begin{pmatrix} a & -b/2 \\ -b & a/2 \end{pmatrix}, \tag{27}$$

so that the matrix of demand response to Treasury yields is

$$\begin{pmatrix} a & -b \\ -b & a \end{pmatrix}. \tag{28}$$

We assume that both a and b are positive, so that Treasury demand increases in its own yield, but decreases in the other-maturity yield, which is the case for the aggregate preferred-habitat demand as we uncovered in Section 3.

We set $K = 1$ so that the macro factor β_t is only one dimensional, and we interpret this single-dimension factor as "supply" factor that drives the total debt supply. We also set $\phi_r = 0$ so that monetary policy process does not depend on the macro factor, and $\bar{r} = 0$ for simplicity. We further set $\zeta_r = 0$ so that debt supply is

$$S_t^{(\tau)} = \bar{S}^{(\tau)} + \zeta(\tau)\beta_t \tag{29}$$

for $\tau = \{1, 2\}$. We impose a regularity condition that $\zeta(2) > -\theta(2)$ so that any supply expansion does not automatically get overshadowed by the expansion of demand in response to such supply expansion. Finally, for simplicity, we shut off all outside portfolio exposure by setting $\tilde{X}_t = 0$.

Next, we solve the model and arrive at the following unique equilibrium solution for log prices,

$$\begin{aligned} p_t^{(1)} &= -r_t \\ p_t^{(2)} &= -\frac{1+\rho_r+\gamma\sigma_r^2 b}{1+\frac{a}{2}\gamma\sigma_r^2} r_t - \frac{\gamma\sigma_r^2(\zeta(2)+\theta(2))}{1+\frac{a}{2}\gamma\sigma_r^2}\beta_t + \frac{\gamma\sigma_r^2}{1+\frac{a}{2}\gamma\sigma_r^2}u_t(2) + \frac{\frac{1}{2}-\gamma\bar{S}^{(2)}+\gamma\theta_0(2)}{\frac{1}{\sigma_r^2}+\frac{a}{2}\gamma}, \end{aligned} \tag{30}$$

where the first equation comes from monetary policy controlling the short rate, and the second equation comes from arbitrageurs accommodating the imbalance between Treasury supply and preferred habitat demand subject to risk aversion. Detailed derivations are provided in Appendix B.1, which also contains proofs of all the following propositions in this section.

Using Equation (30), we summarize the drivers of Treasury price variation in the following proposition.

Proposition 1 (Decomposition of Treasury Pricing)**.** *Monetary policy rate r_t plays a dominant role for short-maturity Treasuries, while macro shocks and latent demand shocks become more important for long-maturity Treasuries.*

Proposition 1 is an intuitive result by simply observing Equation (30). The more general message is that the relative importance of macro factors and latent demand increases as the maturity of Treasuries increase, because the arbitrage force gets weaker at longer maturities. Furthermore, Proposition 1 also implies that a demand shock, either latent demand or permanent demand, has a larger price impact if it comes from longer maturities, because shorter-maturity demand shocks are better accommodated by lower-risk arbitrage. Taking the limit, the one-period arbitrage is perfect and the short rate is not affected by any demand shock.

Next, we analyze how arbitrageur risk aversion γ affects Treasury pricing.

Proposition 2 (Impact of Arbitrageur Risk Aversion)**.** *For long-term Treasuries, a larger arbitrageur risk aversion γ increases the magnitude of Treasury price sensitivity to the macro factor β_t, to latent demand u_t, and to permanent demand $\theta_0(2)$. Moreover, in the steady state, if the arbitrageur holds a positive amount of long-term Treasuries (i.e., $\bar{S}^{(2)} > \theta_0(2)$), higher risk aversion γ increases long-term yields and also the term spread.*

Proposition 2 states that higher arbitrageur risk aversion makes Treasury price more sensitive to many sources of variations in the model, which is intuitive given that arbitrageurs accommodate order imbalances subject to risk aversion. Moreover, if arbitrageurs take a long position in long-term Treasuries, they will demand a lower price and higher yield to compensate for bearing such risks when they become more risk averse. Since the short-term rate is tightly controlled by the Fed, this yield increase also translates into a higher term spread, making long-term Treasuries particularly cheap compared to short-term Treasuries.

It is useful to consider two extreme cases. In the first case, we take $\gamma \to \infty$, so that arbitrageurs "drop out" from the market. Then the long-term Treasury price becomes

$$p_t^{(2)} = -\frac{2b}{a} r_t - \frac{2}{a}(\zeta(2) + \theta(2))\beta_t + \frac{2}{a} u_t(2) + \frac{2}{a}(\theta_0(2) - \bar{S}^{(2)}) \tag{31}$$

This is a case where Treasury prices are entirely driven by supply and demand absent from strategic arbitrage. Therefore, there is no distinction among a temporary latent demand shock u_t, a permanent demand shock $\theta_0(2)$, or a supply shock $\zeta(2)\beta_t$ – all of them have the same price impact. Moreover, short-term rate r_t has an impact on long-term Treasury price $p_t^{(2)}$ only if cross substitution b is not zero.

In the second case, we take $\gamma \to 0$, so that arbitrageurs are risk neutral and arbitrage to the full extent, leading to

$$p_t^{(2)} = -(1+\rho_r) r_t + \frac{1}{2}\sigma_r^2 \tag{32}$$

which is the log Treasury price under the expectation hypothesis (the second term is the Jensen's term after taking logarithm). Intuitively, the current short rate is r_t and in expectation the next period short rate is $\rho_r r_t$, leading to a log price of $-(1+\rho_r) r_t$ plus a convexity adjustment.

According to the preferred-habitat demand in (13) and the simplified elasticity matrix in (27), the equilibrium non-arbitrageur Treasury holding is

$$Z_t^{(2)} = \theta_0(2) + b p_t(1) - \frac{a}{2} p_t(2) - \theta(2)\beta_t + u_t(2). \tag{33}$$

Expanding the above using (30) and collecting the terms involving r_t, β_t, u_t, and the intercept, we obtain preferred-habitat holding (including the Fed) as

$$Z_t^{(2)} = \frac{\theta_0(2)\frac{1}{\sigma_r^2} - \frac{1}{4}a + \frac{a}{2}\gamma \bar{S}^{(2)}}{\frac{1}{\sigma_r^2} + \frac{a}{2}\gamma} + \frac{\frac{a}{2}(1+\rho_r) - b}{1 + \frac{a}{2}\gamma\sigma_r^2} r_t + \frac{\frac{a}{2}\gamma\sigma_r^2(\zeta(2) + \theta(2))}{1 + \frac{a}{2}\gamma\sigma_r^2}\beta_t + \frac{1}{1 + \frac{a}{2}\gamma\sigma_r^2} u_t(2). \tag{34}$$

Equation (34) reflects equilibrium non-arbitrageur demand adjustments from two sources: one is directly from how demand depends on the macro factor β_t and latent demand u_t absent from a

Treasury price effect, and the other is from the response to the Treasury price change, both from the own yield and the other yield, due to cross elasticity. We note that if $\gamma \to \infty$, the above expression converges to $Z_t^{(2)} \to \bar{S}^{(2)} + \zeta(2)\beta_t$, i.e., the total debt supply in this simplified model as in (29). In this extreme case, the arbitrageur holding of long-term Treasury becomes zero. Generally, arbitrageur holding is $X^{(2)} = S_t^{(2)} - Z_t^{(2)}$, which in this simplified model is

$$X_t^{(2)} = \frac{\frac{1}{\sigma_r^2}\bar{S}^{(2)} + \frac{1}{4}a - \theta_0(2)\frac{1}{\sigma_r^2}}{\frac{1}{\sigma_r^2} + \frac{a}{2}\gamma} - \frac{\frac{a}{2}(1+\rho_r) - b}{1 + \frac{a}{2}\gamma\sigma_r^2} r_t + \frac{\zeta(2) + \theta(2)}{1 + \frac{a}{2}\gamma\sigma_r^2}\beta_t - \frac{1}{1 + \frac{a}{2}\gamma\sigma_r^2} u_t(2). \tag{35}$$

With both equilibrium prices in (30) and equilibrium portfolio allocations in (33) and (35), we analyze the effect of cross elasticity b in the following proposition.

Proposition 3 (Impact of Cross Elasticity)**.** *A higher cross elasticity in preferred-habitat demand leads to lower preferred-habitat investor holdings and thus higher equilibrium arbitrageur positions, which lowers equilibrium Treasury prices and increases equilibrium Treasury yields. On the other hand, under a sufficient condition*[10] $1 - \rho_r + \gamma\sigma_r^2(a-b) > 0$*, a higher cross elasticity binds Treasury yields of different maturities by reducing the gap of their interest-rate sensitivities.*

Proposition 3 highlights the two distinct effects of the cross elasticity: it lowers the "bearing capacity" of the market and depresses equilibrium Treasury prices, but also serves to bind Treasury yields of different maturities. The first part is perhaps surprising, since the usual intuition is that cross substitution reflects arbitrage, and when there is stronger arbitrage, the equilibrium price impact is also smaller. Indeed, when we lower γ, as shown by Proposition 2, price impact is smaller and arbitrage is stronger. We want to highlight that this is because the mean-variance model explicitly ties the own-elasticity with cross-substitution using a risk preference, while conceptually they consist of two dimensions. Modeling the preferred-habitat demand as captured by two dimensions, own elasticity and cross elasticity, thus reveals the richness in the preferred-habitat demand that could not be uncovered if we imposed a mean-variance optimization problem. The spirit of this separation is similar to the separation between risk aversion and the intertemporal elasticity of substitution.

We next discuss how the yield curve responds to monetary policy.

Proposition 4 (Monetary Policy and the Yield Curve)**.** *If* $2b/a > 1 + \rho_r$ *(strong cross elasticity), a positive monetary policy shock increases the term premium and causes over-reaction of long-term yield compared to the expectation hypothesis. On the other hand, if* $2b/a < 1 + \rho_r$ *(weak cross elasticity), we obtain the opposite result and there is under-reaction of long-term yield.*

[10] The condition $1 - \rho_r + \gamma\sigma_r^2(a-b) > 0$ is well satisfied according to our empirical estimations, mainly because both γ and σ_r are very small, while a and b are of very similar magnitudes.

We note that Proposition 4 sharply contrasts with the typical results in Vayanos and Vila (2021) type of models without cross elasticity. As Proposition 2 of Vayanos and Vila (2021) shows, there is under reaction of long-term yield compared to the expectation hypothesis. The basic intuition is that when the monetary policy rate rises, long-term Treasuries are cheaper due to the expectation effect, which induces preferred-habitat investors to hold more of them and reduces the holding of arbitrageurs, therefore reducing the risk premium of long-term Treasuries and dampening the yield increase in the first place. When there is strong cross substitution, there is another force that preferred habitat investors tend to reduce long-term Treasury holdings when short-term rate is higher, which then forces the arbitrageurs to increase their long-term Treasury holdings (see Equation (35)). This counteracts with the first force and may cause the yield to be even higher than according to the expectation hypothesis.

In Section 3, we show that for most sectors, the cross elasticity is of a similar order of magnitude as the own elasticity. After aggregating all the sectors, we find that $2b/a = 1.84$ across maturities, while $\rho_r = 0.78$, so the strong cross elasticity is supported in the data. As a result, our data imply overreaction of long-term yield compared to the expectation hypothesis, which is consistent with the literature (Bekaert et al. 2013; Hanson and Stein 2015; Gertler and Karadi 2015; Kekre et al. 2024). Kekre et al. (2024) generates overreaction by introducing wealth effects for arbitrageurs, while we achieve the same result by allowing for cross elasticities.

To further understand the intuition, we can also consider the two extreme cases in Equation (31) and (32). We find that we obtain the same price sensitivity to the short rate r_t in these two cases if and only if $2b = (1+\rho_r)a$, which is the cutoff in Proposition 4. Intuitively, an intermediate risk aversion $\gamma \in (0,\infty)$ represents a case between those two extremes. Therefore, if $2b > (1+\rho_r)a$, the pure-demand case is stronger so that there is overreaction to the policy rate for $\gamma \in (0,\infty)$. When $2b < (1+\rho_r)a$, the pure-demand case is weaker so that there is underreaction to the policy rate for $\gamma \in (0,\infty)$.

Apart from traditional monetary policy, unconventional monetary policy can also be analyzed within the framework. We interpret QE as a demand shift, i.e., a higher $\theta_0(2)$.

Proposition 5 (QE and Treasury Pricing)**.** *QE increases Treasury prices and reduces Treasury yields.*

Proposition 5 indicates the pivotal role of Fed's demand in the Treasury market. With a persistent QE in place (higher $\theta_0(2)$), the Fed permanently increases Treasury prices and lowers Treasury yields.

We note that in this two-maturity model, there is no difference between a temporary demand shock $u_t(2)$ and a permanent demand shock $\theta_0(2)$, since after one period, the two-period bond becomes one period and the price is fully determined by the monetary policy rate. In the full model, we expect the effect to be stronger for permanent shocks, and we will examine this hypothesis quantitatively using the full model.

Finally, we want to caution readers that although Propositions 2 to 5 provide very sharp characterizations regarding the roles of cross elasticities, price responses, and arbitrageur positions, these predictions are obtained under a drastic simplification of the full model. In the significantly richer full model, we consider more than two maturities, so the risk premium on the macro factors β_t will be priced into long-term Treasury pricing, and the demand elasticity matrix α is more complicated than the one in Equation (27). More importantly, the full model accounts for arbitrageur's outside portfolio which is affected by all the important factors including r_t and β_t, so predictions about how the short-rate r_t and macro factors β_t affect the Treasury yield curve are more complicated than the simple predictions in this section. Nevertheless, we believe this simple model still provides useful intuition that guides and helps us interpret our quantitative analysis in the following sections.

5. Model Solution and Estimation

In this section, we solve and estimate our equilibrium model, and evaluate model fit and steady-state implications.

5.1. Model Solution

We start with holding returns,

$$\begin{aligned} r_{t+1}^{(\tau)} &= p_{t+1}^{(\tau-1)} - p_t^{(\tau)} \\ &= A(\tau-1)'(\bar{\beta} + \Phi(\beta_t - \bar{\beta}) + \Sigma^{1/2}\varepsilon_{t+1}) + A_r(\tau-1)(\bar{r} + \phi_r'(\Phi(\beta_t - \bar{\beta}) + \Sigma^{1/2}\varepsilon_{t+1}) + \rho_r r_t + \sigma_r \varepsilon_{t+1}^r) \\ &\quad - A(\tau)' \cdot \beta_t - A_r(\tau) r_t + A_u(\tau-1)' u_{t+1} - A_u(\tau)' u_t + C(\tau-1) - C(\tau). \end{aligned} \tag{36}$$

We can approximate the total holding return as[11]

$$R_{t+1}^{(\tau)} = \exp(r_{t+1}^{(\tau)}) - 1 \approx r_{t+1}^{(\tau)} + \frac{1}{2} Var_t[r_{t+1}^{(\tau)}]. \tag{37}$$

[11] The approximation becomes exact when we take a continuous-time approach. Refer to Greenwood et al. (2023) for a more detailed discussion.

Since there is no uncertainty regarding the current short rate, this approximation also leads to Equation (21). We can thus express the total return as

$$
\begin{aligned}
R_{t+1}^{(\tau)} &= A(\tau-1)'(\bar{\beta}+\Phi(\beta_t-\bar{\beta})+\Sigma^{1/2}\varepsilon_{t+1})-A(\tau)'\cdot\beta_t+C(\tau-1)-C(\tau) \\
&+\frac{1}{2}\left(A(\tau-1)'+A_r(\tau-1)\phi_r'\right)\Sigma\left(A(\tau-1)+\phi_r A_r(\tau-1)\right) \\
&+A_r(\tau-1)(\bar{r}+\phi_r'(\Phi(\beta_t-\bar{\beta})+\Sigma^{1/2}\varepsilon_{t+1})+\rho_r r_t+\sigma_r\varepsilon_{t+1}^r)-A_r(\tau)r_t \\
&+\frac{1}{2}(A_r(\tau-1)\sigma_r)^2+A_u(\tau-1)'u_{t+1}-A_u(\tau)'u_t+\frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1).
\end{aligned} \tag{38}
$$

To simplify expressions, we denote

$$
\hat{A}(\tau-1)=A(\tau-1)+\phi_r A_r(\tau-1), \tag{39}
$$

so that $\hat{A}(\tau-1)'=A(\tau-1)'+A_r(\tau-1)\phi_r'$. Therefore, Equation (38) can be simplified as

$$
\begin{aligned}
R_{t+1}^{(\tau)} &= A(\tau-1)'(\bar{\beta}+\Phi(\beta_t-\bar{\beta})+\Sigma^{1/2}\varepsilon_{t+1})-A(\tau)'\cdot\beta_t+C(\tau-1)-C(\tau) \\
&+\frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1)+A_u(\tau-1)'u_{t+1}-A_u(\tau)'u_t+\frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1) \\
&+A_r(\tau-1)(\bar{r}+\phi_r'(\Phi(\beta_t-\bar{\beta})+\Sigma^{1/2}\varepsilon_{t+1})+\rho_r r_t+\sigma_r\varepsilon_{t+1}^r)-A_r(\tau)r_t+\frac{1}{2}(A_r(\tau-1)\sigma_r)^2.
\end{aligned} \tag{40}
$$

We note that the return $R_{t+1}^{(\tau)}$ in (40) mainly has four components. The first one is the change in macroeconomic factors β_t. The second one is the the change in latent demand u_t. The third one is the change in the monetary policy rate r_t. The final component are the Jensen terms for each type of risk, including the macroeconomic shocks, monetary policy shocks, and latent demand shocks.

Wealth thus evolves as

$$
W_{t+1}=W_t(1+r_t)+\sum_{\tau=2}^{N}X_t(\tau)(R_{t+1}^{(\tau)}-r_t)+\tilde{X}_t(\tilde{R}_{t,t+1}-r_t). \tag{41}
$$

To simplify notation, it is convenient to define the expected return on Treasuries of maturity τ as $\mu_t^{(\tau)}$, with

$$
\mu_t^{(\tau)}\equiv E_t[R_{t+1}^{(\tau)}] \tag{42}
$$

Next, we solve for the mean-variance problem in (24) and derive arbitrageur's first-order condition regarding Treasury holdings. For tractability, we make a simplifying assumption that the idiosyncratic latent demand shocks are not priced and do not carry a risk premium. This is a typical result in most asset pricing models. It is important to note that this assumption does not imply no price impact by latent demand shocks, since u_t can still directly affect price via demand pressure. In Appendix C.4, we examine this assumption numerically.

Then we obtain the first-order condition as

$$\begin{aligned}\mu_t^{(\tau)} - r_t = {} & \hat{A}(\tau-1)' \underbrace{\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\Sigma\hat{A}(\hat{\tau}-1)X_t(\hat{\tau})\right) + \Sigma^{1/2}\tilde{\sigma}\tilde{X}_t\right)}_{\lambda_{\beta,t}} \\ & + A_r(\tau-1)' \underbrace{\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\sigma_r^2 A_r(\hat{\tau}-1)X_t(\hat{\tau})\right) + \sigma_r\tilde{\sigma}_r\tilde{X}_t\right)}_{\lambda_{r,t}},\end{aligned} \tag{43}$$

where $\lambda_{\beta,t}$ is the price of risk of macroeconomic shocks and $\lambda_{\lambda,t}$ is the price of risk of monetary policy shocks. For the Treasury price exposure to macroeconomic shocks, $\hat{A}(\tau-1)$, the expected return $\mu_t^{(\tau)} - r_t$ needs to provide compensation, and the compensation per unit of exposure is reflected by $\lambda_{\beta,t}$. Similarly, the exposure of the Treasury price to interest-rate risks, $A_r(\tau-1)$, requires compensation as reflected by $\lambda_{r,t}$.

Moreover, Equation (43) implies that the price of risk in this model is driven by arbitrageur's total portfolio exposure to various risk factors, including macroeconomic risks and interest-rate risks. Part of such exposure is driven by arbitrageur's Treasury positions $X_t^{(\tau)}$, but it is also affected by the "outside asset" position $\tilde{X}_t$, and its risk exposure. As we will show later, the model allows us to quantify the relative contributions of each component to the price-of-risk variations.

Next, we solve for X_t^τ using the market clearing Equation (25) and replace $Z_t(\tau)$ with (12), $S_t(\tau)$ with (17), thereby pinning down the equilibrium arbitrageur holdings[12] as

$$X_t(\tau) = \left(\bar{S}(\tau) + \zeta(\tau)'\beta_t + \zeta_r(\tau)'r_t\right) - \left(\theta_0(\tau) - \alpha(\tau)'p_t - \theta(\tau)'\beta_t + u_t(\tau)\right). \tag{44}$$

Plugging the definition of $\mu_t^{(\tau)}$ in (42) and (40) as well as the equilibrium arbitrageur holdings X_t^τ in (44) into the pricing equation (43), we obtain

$$\begin{aligned}& A(\tau-1)'\left(\bar{\beta} + \Phi(\beta_t - \bar{\beta})\right) - A(\tau)'\beta_t + \frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1) + C(\tau-1) - C(\tau) \\ & + A_r(\tau-1)(\bar{r} + \phi_r'\Phi(\beta_t - \bar{\beta}) + \rho_r r_t) - A_r(\tau)r_t + \frac{1}{2}(A_r(\tau-1)\sigma_r)^2 \\ & - A_u(\tau)'u_t + \frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1) - r_t\end{aligned} \tag{45}$$

[12]Since arbitrageurs accommodate demand shocks, they are not price takers. Therefore, we cannot view this equation as a typical "demand function".

$$
\begin{aligned}
=&\hat{A}(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\Sigma\hat{A}(\hat{\tau}-1)\left(\begin{array}{c}(\bar{S}(\hat{\tau})+\zeta(\hat{\tau})'\beta_t+\zeta_r(\tau)'r_t)\\-(\theta_0(\hat{\tau})-\alpha(\hat{\tau})'p_t-\theta(\hat{\tau})'\beta_t+u_t(\hat{\tau}))\end{array}\right)\right)+\Sigma^{1/2}\tilde{\sigma}\tilde{X}_t\right)\\
+&A_r(\tau-1)\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\sigma_r^2A_r(\hat{\tau}-1)\left(\begin{array}{c}(\bar{S}(\hat{\tau})+\zeta(\hat{\tau})'\beta_t+\zeta_r(\tau)'r_t)\\-(\theta_0(\hat{\tau})-\alpha(\hat{\tau})'p_t-\theta(\hat{\tau})'\beta_t+u_t(\hat{\tau}))\end{array}\right)\right)+\sigma_r\tilde{\sigma}_r\tilde{X}_t\right).
\end{aligned}
$$

Note that we do not have sufficient degrees of freedom to pin down all dynamic parameters of $\tilde{X}_t$ (since we do not know the asset returns or positions of the outside asset). Instead, we assume that they can be spanned by β_t and r_t, so that

$$
\begin{aligned}
\Sigma^{1/2}\tilde{\sigma}\tilde{X}_t &= \Psi\beta_t+\Lambda r_t+\psi\\
\sigma_r\tilde{\sigma}_r\tilde{X}_t &= \Psi_r\beta_t+\Lambda_r r_t+\psi_r.
\end{aligned}
\tag{46}
$$

The extra parameters Ψ, ψ, Λ, Λ_r, Ψ_r, and ψ_r are useful to fit the arbitrageur model, but these parameters cannot be directly observed or estimated. The parameter Ψ is a $K\times K$ matrix and ψ is a $K\times 1$ vector. The loading Ψ_r is a $1\times K$ vector and ψ_r is a scalar.

With the assumption in (46), and the affine expression of p_t in (26), we can rewrite the equilibrium condition in (45) purely in terms of β_t, r_t, and u_t. Because the equation holds for all values of these variables, the coefficients in front of them must all be matched. Matching the coefficients of β_t leads to

$$
\begin{aligned}
&A(\tau-1)'\Phi-A(\tau)'+A_r(\tau-1)\phi_r'\Phi\\
&=\hat{A}(\tau-1)'\underbrace{\gamma\left(\left(\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\left(\zeta(\hat{\tau})'+\alpha(\hat{\tau})'A+\theta(\hat{\tau})'\right)\right)+\Psi\right)}_{\lambda_{\beta,\beta}}\\
&+A_r(\tau-1)\underbrace{\gamma\left(\left(\sum_{\hat{\tau}=2}^{N}\sigma_r^2A_r(\hat{\tau}-1)\left(\zeta(\hat{\tau})'+\alpha(\hat{\tau})'A+\theta(\hat{\tau})'\right)\right)+\Psi_r\right)}_{\lambda_{\beta,r}}.
\end{aligned}
\tag{47}
$$

Matching the coefficient of r_t leads to

$$
\begin{aligned}
A_r(\tau-1)'\rho_r-A_r(\tau)-1&=\hat{A}(\tau-1)'\underbrace{\gamma\left(\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\left(\zeta_r(\tau)'+\alpha(\hat{\tau})'A_r\right)+\Lambda\right)}_{\lambda_{r,\beta}}\\
&+A_r(\tau-1)'\underbrace{\gamma\left(\sum_{\hat{\tau}=2}^{N}\sigma_r^2A_r(\hat{\tau}-1)\left(\zeta_r(\tau)'+\alpha(\hat{\tau})'A_r\right)+\Lambda_r\right)}_{\lambda_{r,r}}
\end{aligned}
\tag{48}
$$

Matching the coefficient of u_t leads to

$$-A_u(\tau)' = \hat{A}(\tau-1)'\gamma\Sigma\left(\left(\sum_{\hat{\tau}=2}^{N}\hat{A}(\hat{\tau}-1)\alpha(\hat{\tau})'A_u\right) - \left(0,\hat{A}(1),\ldots,\hat{A}(N-1)\right)\right)$$
$$+A_r(\tau-1)'\gamma\sigma_r^2\left(\left(\sum_{\hat{\tau}=2}^{N}A_r(\hat{\tau}-1)\alpha(\hat{\tau})'A_u\right) - (0,A_r(1),\ldots,A_r(N-1))\right) \tag{49}$$

Matching the constant term leads to

$$\begin{aligned}
&A(\tau-1)'(I-\Phi)\bar{\beta} + A_r(\tau-1)(\bar{r}-\phi_r'\Phi\bar{\beta}) + \frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1) \\
&+\frac{1}{2}(A_r(\tau-1)\sigma_r)^2 + \frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1) + C(\tau-1) - C(\tau) \\
&= \hat{A}(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\left(\bar{S}(\hat{\tau}) - \left(\theta_0(\hat{\tau}) - \alpha(\hat{\tau})'C\right)\right) + \psi\right) \\
&+A_r(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\sigma_r^2 A_r(\hat{\tau}-1)\left(\bar{S}(\hat{\tau}) - \left(\theta_0(\hat{\tau}) - \alpha(\hat{\tau})'C\right)\right) + \psi_r\right).
\end{aligned} \tag{50}$$

The common structure in the iterative Equations (47), (48), (49), and (50) is that they are all related to the habitat-demand price elasticity $\alpha(\tau)$. Therefore, the price elasticity of habitat-demand is central in driving the pricing of Treasury securities. Moreover, the arbitrageur risk aversion coefficient γ also shows up in all equations, indicating that arbitrageurs are critical for understanding the price fluctuations in the Treasury market, as already shown in Proposition 2.

5.2. Estimation Methodology

The model has a very high-dimension parameter space. To quantify the model, we directly set parameters having a direct mapping to the data, and leave the rest for estimation. In particular, we use the instrumented regressions of demand estimation to directly uncover the preferred-habitat demand in (13), including parameters θ_0, α, and θ that we summarize in Table 14. Then we estimate the dynamics of Treasury supply in (18) and obtain coefficients $\bar{S}$, ζ, and ζ_r. Next, we use a regression to uncover monetary policy rules in (10), setting $\bar{r}$, ϕ_r, ρ_r, and σ_r. Finally, we implement a vector autoregression analysis on macro factors to uncover $\bar{\beta}$, Φ, and Σ in Equation (9). We provide details on all of these steps in Appendix B.3.

A key observation is that since Ψ, Ψ_r, Λ, and Λ_r are unknown and have the same dimensions as $\lambda_{\beta,\beta}$, $\lambda_{\beta,r}$, $\lambda_{r,\beta}$, $\lambda_{r,r}$, respectively, we can directly estimate the latter set of parameters, and then back out the former after fitting the model. Moreover, the iterative equations in (47), (48), and (49) indicate that A_u, which is the loading on the latent demand vector, does not enter the pricing of macroeconomic factors and the interest-rate factor. Therefore, we can solve A_u after obtaining A and A_r from (47) and (48). Finally, given ψ, ψ_r, A_u, A, A_r, and γ, we can obtain C from the Equation in (50).

We find that Equations (47), (48), (49), and (50) do not exhibit sufficiently many degrees of freedom to pin down all parameters. In particular, we need an extra moment for γ. We find that the steady-state arbitrageur's holding of long-term Treasuries as a fraction of long-term Treasury outstanding is a moment that is very sensitive to γ. Denote the data moment as h^o, which is 2% for Treasuries above 1Y maturity in our main sample from 2011 to 2022. Denote the model-implied moment as h. Denote the actual Treasury yield as $y_t^o(\tau)$ for maturity τ, and the model-implied Treasury yield as $y_t(\tau)$. Then the estimation problem is as follows:

$$\min_{\{\lambda_{\beta,\beta},\lambda_{\beta,r},\lambda_{r,\beta},\lambda_{r,r},\gamma,\psi,\psi_r\}} \mathbb{E}\left[M\cdot(h-h^o)^2+\sum_t\sum_\tau(y_t(\tau)-y_t^o(\tau))^2\right], \tag{51}$$

subject to the restriction in (47), (48), (49), and (50). We pick M to be large so that the average intermediary Treasury holding is matched really well. In our implementation, we pick $M=1000000$ and increasing M further does not change the results. For convenience, we rewrite (47) and (48) as

$$A(\tau)'=A(\tau-1)'\Phi+A_r(\tau-1)\phi_r'\Phi-\hat{A}(\tau-1)'\lambda_{\beta,\beta}-A_r(\tau-1)\lambda_{\beta,r} \tag{52}$$

$$A_r(\tau)=A_r(\tau-1)\rho_r-1-\hat{A}(\tau-1)'\lambda_{r,\beta}-A_r(\tau-1)\lambda_{r,r}, \tag{53}$$

where $\hat{A}(\tau-1)$ is a function of $A(\tau-1)$ and $A_r(\tau-1)$ as defined in (39). The model implies that $y_t(\tau)=-p_t(\tau)/\tau$, where $p_t(\tau)$ satisfies the pricing equation in (26), with u_t being unobservable with mean 0, variance Σ_u, and uncorrelated with β_t and r_t. As a result, we can rewrite the objective function as

$$\min_{\{\lambda_{\beta,\beta},\lambda_{\beta,r},\lambda_{r,\beta},\lambda_{r,r},\gamma,\psi,\psi_r\}} \mathbb{E}\left[M\cdot(h-h^o)^2+\sum_t\sum_\tau(A\beta_t+A_r r_t+C+\tau y_t^o(\tau))^2\right], \tag{54}$$

where we assume that u_t does not depend on β_t and its history so that the terms involving u_t drop out. In other words, we require the latent demand shocks to be non-systematic.

Next, with $\lambda_{\beta,r}$, $\lambda_{\beta,\beta}$, $\lambda_{r,\beta}$, $\lambda_{r,r}$, γ, and A_u solved, we can recover the intermediary outside asset risk loadings Ψ, Ψ_r, Λ, Λ_r, from the definitions of $\lambda_{\beta,r}$, $\lambda_{\beta,\beta}$, $\lambda_{r,\beta}$,and $\lambda_{r,r}$ in Equations (47) and (48). Knowing these "outside-asset" loadings inform us about the contribution of Treasury holdings in the variation of price of risk by arbitrageurs.

5.3. Model Estimation and Fit

We choose the macro state vector as $\beta_t =$ (credit spread, GDP gap, core inflation, debt/GDP), in line with our empirical analysis. The variables GDP gap and core inflation together reflect aggregate demand and supply fluctuations in the economy, and they are also the variables that drive monetary policy in the Taylor rule. Debt/GDP is important for keeping track of the total debt burden and could reflect fiscal shocks that are orthogonal to aggregate demand and supply shocks. The credit spread reflects credit market conditions and is also an important indicator for aggregate macroeconomic fluctuations (Gilchrist and Zakrajšek (2012), Krishnamurthy and Muir (2017)). Adding additional macroeconomic variables does not increase the explanatory power of the model for Treasury yield dynamics but could introduce overfitting problems, so we choose this set of four macro variables. We estimate a VAR of the form (9) using the same sample period as our main empirical analysis. We find that core inflation and debt/GDP are both highly persistent. Nevertheless, the maximum absolute value of the eigenvalue is 0.89, so macro variables converge to their long-run average.

To fit the monetary policy rule, we have to use a longer time period, because the monetary policy rate does not exhibit much variation during our main sample period. In particular, we use the post-Volcker period (1990 to 2024) excluding the zero lower bound (ZLB) period (2008-2015). We start from 1990 because that is when the Fed gained credibility in its fight of inflation. We find that the Fed lowers the interest rate if the credit spread is high, the GDP gap (GDP deviation from potential GDP) is low and tightens the interest rate if inflation is high. The coefficients on GDP gap and inflation have the same signs as in the classical Taylor rule (Taylor 1993). Moreover, there is a moderate amount of monetary policy inertia reflected by the coefficient of 0.78 on the lagged policy rate. This dependence on lagged policy rate generates an impact of monetary policy rate on long-term yields from the expectations effect and is critical to understand how yield curve responds to monetary policy shocks ε^r_{t+1}.

Due to economic growth, portfolio holdings in dollar values will not be stationary. For stationarity, we scale all quantities by the ratio of potential GDP in that particular quarter over the baseline value at the beginning of our sample period. For example, the ratio of potential GDP in 2021 Q3 to that in 2011 Q3 is 1.46. The dollar value of total debt supply in 2021 Q3 is 22.2 trillion, but we use a scaled value, namely $22.2/1.46 = 15.2$ trillion. Our scaling accounts not only for inflation but also economic growth in general. The underlying assumption is that after accounting for the scaling effect, all quantities are stationary in the fundamental state variables.

We estimate problem (54) on our main data sample from 2011 Q3 to 2022 Q4. In Figure 4, we show that the model can fit the the term structure pretty well, both across maturities and over time. Note that these results are achieved by having only fundamental economic variables as state variables, which is much more challenging than a typical affine term structure model that includes latent factors coming directly from Treasury yields. The equilibrium restrictions in the model impose tight restrictions on how flexibly the model can explain the dynamics of Treasury yields.

5.4. Quantifying Arbitrageur's Risk Aversion

Our model identifies arbitrageur's risk aversion by including the moment of long-term Treasury holdings by arbitrageurs. The data moment is 2%, i.e., arbitrageurs hold about 2% of total outstanding of > 1Y Treasuries. The model accurately matches this moment with less than 10^{-6} of error, and the resulting risk-aversion parameter is $\gamma = 0.053$. This value is much smaller than the calibrated value in Vayanos and Vila (2021)[13], which is about 7. The difference is more than 100 times! There are two main reasons for this large gap. First, we accommodate rich macro dynamics beyond interest-rate risks or supply risks. Second, we allow for dynamics in the arbitrageur's outside portfolio exposure. Both of these features of our model indicate that we do not need a large risk aversion to generate volatile long-term Treasury yields in the data, which is needed if there are fewer sources of risks and arbitrageurs' only source of risks is Treasury holding.

We want to highlight that this value of γ is still significantly different from zero, because the equilibrium price impact of demand shocks is still non-negligible, despite smaller than other financial markets. We will discuss this point in Section 6.2.

The novelty of our approach is that it only relies on quantities to pin down the arbitrageur's risk aversion. As a result, the model can generate realistic quantity allocations across sectors and build a tight linkage between quantities and prices. Our approach relies on granular data that allow us to distinguish arbitrageurs explicitly versus preferred habitat investors.

5.5. Steady State Yield Curve and Portfolio Holdings

In Figure 5 panel (a), we illustrate the steady-state yield curve. This steady-state yield curve is upward-sloping and mainly reflects the average shape of the yield curve in our model estimation period.

In Figure 5 panel (b), we illustrate the steady-state portfolio allocations across different sectors. The model implies that in the long run, foreign investors are still the largest holder among all groups of investors, and the Fed also plays an important role. Insurance and pension funds are not large holders, but they predominantly hold long-term Treasuries. Finally, as targeted by

[13]Table I of Vayanos and Vila (2021) reports a demand slope of 5.21 and risk aversion multiplying demand slope as 35.3, indicating a risk aversion of about 7.

the calibration, arbitrageurs' longer-term (>1Y) Treasury holding is 2% of the total longer-term Treasuries outstanding.

6. Dissecting the Treasury Market

In this section, we put our estimated model to work, and illustrate its basic mechanics and implications by dissecting the Treasury market. In particular, we decompose Treasury yields into different driving forces, quantify the impact of arbitrageur risk aversion, and evaluate the role of cross elasticities. Our quantitative analyses are guided by intuitions from Proposition 1 to 3.

6.1. Decomposition of Treasury Pricing

We start by taking guidance from the model to decompose Treasury pricing in a number of different ways.

First, we use the model-implied state variables to estimate the regressions in (26) and examine the contribution of each factors. We use the Shapley value of R^2 to measure the relative contribution of each variable, including macro variables (taken together), monetary policy, and latent demand. A key advantage of our approach is that from our granular demand data, we explicitly recover the latent demand component, which is treated as latent factors in the the literature (Ang and Piazzesi 2003; Bikbov and Chernov 2010; Joslin et al. 2014). We express the contribution of each variable as the Shapley value of R^2, which is calculated as the marginal contribution of each variable to the R^2 among all possible sets of dependent variable combinations. The Shapley value of macro variables is taken as the sum of Shapley values on all macro variables. Finally, we divide each individual Shapley value by the total Shapley value so that the scaled values add up to 100%.

As shown in Figure 6 panel (a), the relative contribution of different economic forces varies across the term structure. The patterns are exactly in line with the theoretical predictions of Proposition 1. For short-maturity Treasuries, monetary policy plays the dominant role, explaining the vast majority of variations, while macro variables play a secondary role. We note that latent demand almost has no explanatory power for short-maturity Treasury yields. As the maturity increases, the relative importance of the FFR declines while the relative importance of both macro variables and latent demand shocks expand. In particular, for long-maturity Treasuries, macroeconomic variables can explain about half of the variation in yields. The same exercise on prices (not reported) gives very similar results. Overall, these results confirm the intuition provided in Proposition 1.

Next, we use the model to decompose Treasury price variations into supply and demand shocks. The idea is that arbitrageurs take the order imbalances in the market and price Treasuries, so factors affecting the supply and demand imbalance play an important role in Treasury pricing. Define the sector-i specific shock as the following:

$$\Delta Z_t^i = -\theta(\tau)\Sigma^{1/2}\varepsilon_t + u_t,$$

which also includes shocks to the Fed's portfolio. Similarly, we can denote the supply shock as

$$\Delta S_t = \zeta'\varepsilon_t + \zeta_r\sigma_r\varepsilon_t^r.$$

With these supply and demand shocks, we can decompose their contribution to Treasury yield changes, $\Delta y_t(\tau) = y_t(\tau) - y_{t-1}(\tau)$ on demand and supply shocks and calculate the Shapley value in the same way as in the previous decomposition.

The second decomposition is conceptually distinct from the first one and focuses on different aspects of Treasury yield variations. The first one tells us how state variables and latent demand drive the overall variations in Treasury yields. In contrast, the second one focuses on the unpredictable components of supply and demand.

As shown in Figure 6 panel (b), the decomposition varies across the maturity structure of Treasuries. We find that debt supply shocks mainly play an important role at long maturity, while Fed portfolio shocks contribute significantly to yield variations at the short maturity. Although banks' holdings are very small compared to the entire market (half a trillion as shown in Figure 5), banks play a sizable role in transmitting shocks to the Treasury market, especially the medium maturity bucket (1$\sim$5 years). Foreign investors have significant shocks to both short maturity and medium maturity Treasuries, but not long maturity Treasuries. Although the total contribution of foreign investor demand shocks is similar to that of banks, foreign holdings are much larger (about 5 trillion as shown in Figure 5) than bank holdings, reflecting that foreign investors are mostly passive and not shocking the Treasury market.

Other U.S. investors' demand, which is the residual sector's demand in our estimation, mainly shocks the long-maturity bucket, as indicated by Figure 2. Money market funds only transmit shocks to short maturity Treasuries, which is what we expect given their mandate of holding only money-market instruments. Finally, insurance companies and pension funds (ICPF) significantly contribute to variations in long-maturity yields, consistent with their segmentation in the long-maturity Treasury market.

Overall, we find that the contribution of each sector to demand shocks could be very different from the amount of holdings, since the latter mainly reflects a stable preference, not how actively a sector trades on shocks. Therefore, decompositions in Figure 6(a) and Figure 6(b) reflect different aspects of what drives Treasury pricing.

6.2. The Impact of Arbitrageur Risk Aversion

As shown by Proposition 2, the risk aversion of arbitrageurs is a critical parameter that drives how Treasury yields respond to different types of shocks. To quantitatively show the effects of γ, we will vary γ slightly around the estimated value and calculate sensitivities of various prices and elasticities. We interpret this variation as reflecting the impact of bank regulations or financial disruptions on the functioning of the Treasury market, including the tighter bank regulation post-GFC and also the Treasury market disruptions in crisis episodes such as March 2020 (Duffie 2020; He et al. 2022).

The results are shown in Figure 7. In panel (a), we analyze the response of the steady-state yield curve to a change of risk aversion γ. A value of 1 implies that if risk aversion increases by 1%, then steady-state yields also increase by 1%. We find that there is a slightly negative change for yields for medium and short maturities, but a dramatic increase for long maturities. The strong positive response of long-maturity Treasury yields is in line with the intuition in Proposition 2.

In panel (b), we analyze how γ affects the price sensitivity to the policy rate, which is A_r in Equation (26). We find that across all maturities, a higher risk aversion increases the price sensitivity to the policy rate. As a result, a higher γ implies that Treasury yields are more responsive to monetary policy changes. Since the persistence parameter, ρ_r, in (10) remains unchanged, this change in A_r is due to the risk premium response to monetary policy rather than the expectation component.

In panel (c), we analyze how γ affects the price sensitivity to latent demand shocks, which is A_u in Equation (26). Since A_u is an $N \times N$ matrix, we take averages across the rows and use the average price response across different types of shocks, i.e., we plot $\frac{\partial}{\partial \log(\gamma)} \log(\frac{1}{N} A_u \cdot \mathbf{1})$. We note that across all maturities, the response becomes stronger, and it is most prominent at the long maturity.

In panel (d), we show how γ affects the price sensitivity to permanent demand averaged across permanent demand shocks across different maturities, i.e., we plot $\frac{\partial}{\partial \log(\gamma)} \log(\frac{1}{N} \frac{\partial p}{\partial \theta_0} \cdot \mathbf{1})$. We find that when arbitrageurs become more risk averse, the price impact of permanent demand shocks is more prominent for medium maturities, but much less prominent for long maturities. Since we have already shown in Section 4.2 that this value should be positive under the simplified model, the negative reaction should come from how risk aversion changes the impact of macroeconomic risks or outside portfolio risks.

Next, to highlight the importance of explicitly modeling an arbitrageur, i.e., setting γ to a finite value to match a moment rather than setting $\gamma = \infty$ and thus excluding the arbitrageur, we will compare the impact of demand shocks under the two scenarios.

The model prediction is quite different from the demand systems that do not explicitly capture arbitrageurs. When arbitrageurs are removed from the system, market clearing implies

$$p_t = \alpha^{-1}\left((\theta_0 - \theta\beta_t + u_t) - \left(\bar{S} + \zeta\beta_t + \zeta_r r_t\right)\right). \tag{55}$$

Therefore, absent the arbitrageur, the equilibrium price response to a demand shock is simply α^{-1}. On the other hand, with arbitrageurs, the equilibrium demand elasticity is also affected by arbitrageur risk-aversion γ, the volatility of macroeconomic shocks Σ, monetary policy uncertainty σ_r and inertia ρ_r, and the persistence of macroeconomic dynamics Φ.

In Table 5, we illustrate the equilibrium Treasury price response (in %) to a \$100 billion latent demand shocks on a certain maturity bucket. In Panel A, we report the values of the price elasticity in the full model with arbitrageurs. Since in our empirical investigation, we separate all Treasuries into three buckets, we only illlustrate the responses for these three maturity buckets, picking the average duration as the representing maturity. We find that for the same shock (i.e., same row of the table), long maturity price response is much stronger. For example, the response of long-maturity Treasury price is 16 times larger than that of short-maturity Treasury price response, given a demand shock on short-maturity Treasuries. Moreover, demand shocks on longer-maturity Treasuries is more powerful. For example, a \$100 billion dollar demand shock on long-maturity Treasuries increases the valuation of long-maturity Treasuries by 0.688%.

In Panel B, we take out the arbitrageur (i.e., setting $\gamma = \infty$), and examine the corresponding price elasticity. Without arbitrageur, the market clearing price is given by (55). The price elasticity assuming no arbitrageur is generally one to two orders of magnitude larger than the case with arbitrageur. Clearly, without accounting for arbitrageurs, the price impact on T-bills is too large and would never occur with a Federal Reserve that tightly controls the money market. In our model with arbitrageurs, since the Fed keeps a tight control on the monetary policy rate by actively accommodating any demand shocks in the one-period (one quarter) Treasury market, and with arbitrageur, such a force is present across the entire term structure and becomes weaker at longer maturities. Therefore, demand shocks on shorter-maturity Treasuries have a smaller effect on the yield curve.

This stronger arbitrage is clearly illustrated in Panel C of Table 5, where we present the ratio of price impact between the case without arbitrageur and the case with arbitrageur. On average, the price impact in the case without arbitrageur is more than 100 times the price impact with arbitrageur. In Appendix **??**, we also show the impact of permanent demand shocks with and without arbitrageurs and we obtain a similar conclusion.

We note that the results in Table 5 do not contradict those in Figure 7, where we locally vary the risk aversion parameter γ. Due to the nonlinearity in the model, locally the price sensitivity to permanent demand shocks at certain maturities decreases with γ, but globally, the predictions are consistent with the intuition in Proposition 2 derived from our simplified model.

6.3. How Elastic is the Treasury Market?

Given the price responses of each maturity bucket in Table 5, we can calculate the aggregate Treasury market elasticity. In particular, we define a Treasury market multiplier in the same spirit as (Chaudhary et al. 2022) and (Gabaix and Koijen 2021), which is the total valuation change of the whole Treasury market for a $1 billion dollar demand shock that reflects the overall Treasury market.

Using the values in Table 5 and average dollar outstanding of each maturity bucket as weights, we find that the total market multiplier for a representative latent demand shock is 0.27. This implies that for a $100 billion dollar demand shock on the whole Treasury market, the total Treasury valuation increases by $27 billion dollars. On the other hand, the multiplier is 1.00 for a representative permanent demand shock. [14]

In contrast, the multiplier for corporate bonds is 3.5 at the rating level aggregation (Chaudhary et al. 2022) and 5 in the stock market (Gabaix and Koijen 2021). As a result, price impact in the Treasury market is weaker, which suggests that the Treasury market is elastic.

However, in the same spirit as the previous section, once we remove the arbitrageur (i.e., setting $\gamma = \infty$), Treasury market multiplier increases to 18.00, which is much bigger than the multipliers of corporate bonds and equity.[15] Given this, we conclude that the Treasury market is elastic because of low arbitrageur risk aversion.

6.4. The Impact of Cross Elasticities

Allowing for cross elasticities in preferred-habitat demand is a major innovation in our framework compared to Vayanos and Vila (2021). According to Proposition 3, higher cross elasticities in the preferred-habitat demand lower equilibrium Treasury prices, but reduce the gap of interest-rate sensitivities across maturities. In this subsection, we evaluate whether these predictions are also present in our quantitative model.

In Figure 8, we show the impact of the cross elasticities. Panel (a) illustrates that the arbitrageur positions increase as cross elasticities increase. Quantitatively, a 1% increase of all cross elasticities will increase the arbitrageur position by about 4%.

[14] See Appendix C.1 for more details on calculating market multipliers for both latent and permanent demand shocks.
[15] See Appendix C.1 for more details.

As shown by Proposition 3, such an increase in arbitrageur holdings increase the risk premium and depress Treasury prices. Indeed, as shown by Panel (b), equilibrium Treasury prices fall, and the price reaction is much more prominent at medium and long maturities. Quantitatively, a 1% increase of all cross elasticities will depress Treasury prices by about 20 bps in the medium and long maturity buckets, but less than 1 bps for Treasuries below 1Y. This tapering effect at the short-maturity end is consistent with short-maturity demand shocks having lower price impact, as shown in Table 5.

Since prices decrease with cross elasticities, yields increase accordingly, as shown in Panel (c). We find that the increase of yields is most prominent for the medium maturity bucket. Consequently, higher cross elasticities will render the yield curve more hump-shaped.

Finally, in Panel (d), we illustrate the interest-rate sensitivity gap, defined as the difference of yield sensitivities between short-maturity and long-maturity Treasuries. Since short-maturity yields are more responsive than long-maturity yields, this gap is always positive. We plot how this gap deviates from its baseline when we increase cross elasticities. We find that a one percentage increase of all cross elasticities decreases the interest-rate sensitivity gap by about 2%.

In summary, according to our quantitative analysis, higher preferred-habitat cross elasticity increases arbitrageur Treasury holding, decreases Treasury prices, and shrinks interest-rate sensitivity gaps, all consistent with Proposition 3.

7. Conventional and Unconventional Monetary Policies

Our estimated model provides a suitable laboratory to provide quantitative guidance regarding how conventional and unconventional monetary policies affect the Treasury market and the yield curve. In this section, we analyze both monetary policy and quantitative easing/tightening. Given our granular data and quantitative model, we zoom in on Fed's portfolio adjustments. We show Fed policy rules, i.e., how policies depend on the underlying economic state, affect the Treasury yield curve and portfolio adjustments.

7.1. The Impact of Conventional Monetary Policy

In this subsection, we analyze the impact of a one-standard deviation shock to monetary policy, $\varepsilon_r = 1$, at the steady state. Given an annualized standard deviation of 0.75%, this is equivalent to a positive shock to the short rate by 0.75%. In Figure 9, we illustrate how the term structure responds to an increase in the monetary policy rate. The left panel illustrates the full response of the yield curve, as well as the expected component. Absent from risk premium changes, the expected future short rate change is $E_t[\Delta r_{t+h}] = \sigma_r \rho_r^h$, and the expectation component of the yield curve change for maturity τ is

$$\frac{1}{\tau}\sum_{h=0}^{\tau-1} \sigma_r \rho_t^h. \tag{56}$$

As shown in panel (a) of Figure 9, this expectation component declines quickly over maturity, approaching almost zero at around a 15 year maturity. On the other hand, the full response implied by the model strongly reacts to the monetary policy shock even at a 30 year maturity. Due to the difference, the risk premium component, which is the difference between the full response and the expectation component, positively responds to a monetary policy shock, as shown in panel (b). The positive response of term premium to monetary policy shocks and the "excessive reaction" of very-long maturity Treasuries are well-documented facts in the term structure literature (Hanson and Stein 2015).

Why do the term premia "over react" to a monetary policy shock, contrary to the intuition of Vayanos and Vila (2021)? In Vayanos and Vila (2021), a higher monetary policy rate reduces Treasury prices and thus boosts preferred-habitat demand, which causes the arbitrageurs to reduce their positions in Treasuries, leading to a lower risk premium and thus a lower term premium. On curve again in Figure 10. Consistent with the model intuition, once the cross elasticity weakens, there is under-reaction of the yield curve compared to the expectation hypothesis, and we are back to the baseline results in Vayanos and Vila (2021). As a result, obtaining the correct cross elasticities in preferred-habitat demand functions is crucial for getting the right response of the term premium to monetary policy shocks.

Our quantitative result of positive term premium response to monetary policy tightening is consistent with findings in Hanson and Stein (2015) that long-term real rates are much more sensitive to monetary policy shocks than implied by the expectation hypothesis. The explanation in Hanson and Stein (2015) is that bond investors are yield-oriented, comparing long-term yields with short-term yields when they make investment decisions of long-term bond holdings. Our estimation of cross elasticity in preferred-habitat demand is consistent with this explanation. Moreover, our evidence from a quantitative model of the whole Treasury market goes beyond the evidence of only commercial banks in Hanson and Stein (2015).

Next, we analyze the portfolio adjustment to the same Fed policy shock. Fed policy shocks may change the total supply of government debt, which has to be accounted for in the portfolio allocation. In Table 6, we illustrate sector-level portfolio adjustments in response to the monetary policy shock. The general pattern is that the government supplies more short-maturity bonds while reduces long-term bond supply, with a total net change of Treasuries outstanding of zero. The zero total response of debt supply to a contemporaneous monetary policy rate shock is because by definition, total debt supply is driven by the contemporaneous macro variable Debt/GDP, which causes the aggregate loadings on other factors, including the current monetary policy rate, to be close to zero. In response to both a higher short-term interest rate and a larger supply of short-term Treasuries, almost all investors increase their positions in short-term Treasuries, especially the money market funds (MMF) and arbitrageurs.

We observe a very notable change in the Fed portfolio. The model implies that the Fed will significantly reduce long-term Treasury holdings, which can be interpreted as quantitative tightening. This adjustment is consistent with a coherent monetary policy stance that when the Fed tightens monetary policy, it uses both the short rate and its balance sheet.

Given that Fed sales of long-term Treasuries go beyond the shrinking of the corresponding supply, other investors need to absorb a net increase of long-term Treasuries. The most significant expansion comes from "U.S. other", which is the sector that contains U.S. corporations, certain pension funds, and households. We also note that foreign investors are aggressively selling Treasuries in response to a monetary tightening. This is perhaps somewhat puzzling compared to the typical international capital flows in a Fed tightening cycle. We want to emphasize that the foreign investor demand coefficients on own yield and other yield are not significant and the signs are sensitive to how we infer the medium and long maturity positions.

Overall, a one standard deviation monetary policy shock coincides with the Fed tightening its balance sheets, which induces large portfolio adjustments in the market, causing MMF and arbitrageurs to significantly increase their positions. Moreover, the duration of total Treasury supply significantly decreases with a higher policy rate.

7.2. How Does Fed Balance Sheet React to Treasury Market Shocks?

Next, we analyze how the Fed balance sheet reacts to sell shocks and the role of the cross elasticity of the Fed's "demand function". We will implement three exercises. First, we introduce a temporary latent demand shock and examine the response of the Fed balance sheet together with other sectors. Second, we conduct a similar analysis for a permanent demand shock. Finally, we study the role of the cross elasticity in the Fed's demand function and its implications on portfolio allocations.

We use granular sector-level demand functions to study how each sector adjusts to latent demand shocks at the average state in the model. We start with the equilibrium with a zero latent demand shock at the steady state, so that the benchmark price is $\bar{p} = A\bar{\beta} + A_r\bar{r} + C$. Next, we introduce a demand shock Δu to the model, and then analyze how each sector absorbs this shock. The new price will be $p = A\bar{\beta} + A_r\bar{r} + A_u\Delta u + C$, with corresponding change $\Delta p = p - \bar{p} = A_u\Delta u$. Each sector i, including the Fed, absorbs this shock according to the demand specifications in equation (11), namely

$$\Delta Z^i = -\sum_{i \in \mathscr{I}} \alpha^{(i)} A_u \Delta u, \tag{57}$$

where we exclude the shock Δu itself and only consider adjustments due to price changes.

In Table 7, we illustrate the portfolio reallocation across different sectors in the case of a temporary \$100 billion dollar sell shock at either the short-maturity bucket or the long-maturity bucket. We find that the Fed balance sheet barely changes in response to either a short-maturity or a long-maturity latent demand shock. This is in line with the notion that the Fed as a central bank should not react to temporary market fluctuations but rather focus on long-term policy goals. We find that for a short-maturity sell shock, arbitrageurs are the only sector that significantly responds. On the other hand, for a long-maturity sell shock, other U.S. investors significantly absorb quantities, while MMFs and foreign investors amplify the shock.

Next, we analyze the impact of permanent demand shocks. We introduce a change in permanent demand $\Delta\theta_0$, and the then solve the model again to obtain the change of steady-state yield curve Δy^{steady}, as well as portfolio changes in response to such permanent changes in Treasury yields.

In Table 8, we illustrate the portfolio reallocation across different sectors when there is a permanent \$100 billion dollar sell shock at either the short-maturity bucket or the long-maturity bucket. We find that the Fed balance sheet barely changes in response to a short-maturity permanent demand shock, similar to other sectors, while arbitrageurs entirely absorb the demand shock, similar to the case of a latent demand shock on the short-maturity bucket. However, for a shock on the long-maturity bucket, the Fed significantly expands its balance sheet by buying 67 billion dollars of long-maturity Treasuries while keeping the position of other Treasuries almost unchanged. Consequently, the Fed is the main absorber of the long-maturity permanent demand shocks in the Treasury market.

Table 8 also shows that money market funds again amplify the long-maturity Treasury sell shock by selling more short-maturity Treasuries, together with the foreign investors, while other U.S. investors together with arbitrageurs help with absorbing the shock.

Overall, our model suggests that the Fed is the primary absorber for long-term permanent Treasury demand shocks, but does not significantly react to other types of permanent demand shocks or any transient latent demand shocks. Moreover, shutting off the cross elasticity in Fed's demand function makes the Fed more accommodating in response to long-term Treasury demand shocks and lowers the balance sheet pressure on arbitrageurs, consistent with the intuition in Proposition 3.

7.3. The Impact of Quantitative Easing

Next, we evaluate the impact of quantitative easing (QE) policies. Through the lens of our model, we can think of QE policies as changes in Fed's demand, either a temporary change or a permanent change. According to Proposition 5, QE increases Treasury pries and decreases Treasury yields.

In Figure 11, we show the impact of both transient QE and permanent QE. In both cases, we consider the steady state yield as the baseline scenario and then introduce a \$100 billion extra demand on each of the three maturity buckets, respectively. The dollar value is converted to stationary units as in Section 5.3. Then we show the change of yields in response to the demand shock. The key difference between the two is that a transient QE is equivalent to increasing u_t while a permanent QE is equivalent to increasing θ_0.

In panel (a) of Figure 11, we show how transient QE changes yields from the baseline. First, we find that the maturity of the purchase plays an important role. For QE on short-maturity Treasuries, there is little reaction of the yield curve, because dealers elastically arbitrage between short-maturity Treasuries and the one-period rate controlled by monetary policy. As maturity increases, the yield curve is more reactive, since arbitraging becomes weaker for longer-maturity demand shocks due to the extra risks involved in long-maturity Treasuries. These results justify why the Fed usually purchase long-term Treasuries in the QE programs.

In panel (b) of Figure 11, we illustrate how permanent QE changes yields. We find that a permanent QE on long-maturity Treasuries has a much larger impact on the yield curve everywhere than the same purchase of short-maturity Treasuries, a feature shared with the latent demand shocks in panel (a). Again, this confirms the rationale of long-term Treasury purchases. Moreover, there is a strong localization effect of QE policy, in that QE on a specific maturity bucket affects that maturity-bucket yield more strongly than others, which is a feature present in the framework based on Vayanos and Vila (2021) with more than one risk factors (Greenwood et al. 2023). The idea is that due to the presence of multiple sources of risks that are heterogeneously present at different maturities, the arbitrageur will not aggressively trade against a permanent demand shock, causing a localization of the price impact.

To compare the model-implied results with actual QE impact in event studies, we have to consider details of the QE implementation: First, the duration of QE purchases range between 3 to 10 years, so the average effect is in between our maturity bucket 2 and 3, i.e., between the solid orange line and dotted red line in Figure 11. Second, the expected duration of the QE purchase is between one quarter (panel (a) of Figure 11) and permanent (panel (b) of Figure 11). As a rough approximation, we use the average value of bucket 2 and 3, which implies that the impact of a \$100 billion purchase generates 5 bps to 20 bps yield change of 10-year Treasuries, depending on the expected persistence of QE. This is of similar order of magnitude to the 4.5 bps reported in Gulati and Smith (2022), which summarize results in many different studies in the literature, including Krishnamurthy and Vissing-Jorgensen (2011); Swanson (2011).

The advantage of our analysis is that we are able to provide granular results on how Treasury yields change according to the implementation details of the QE policy, including the expected persistence of such policy and the maturity of Treasuries being purchased. Our results highlight that QE policies are more effective when the average maturity of Treasury being purchased is longer and expected persistence of QE policy is larger.

Overall, these results suggest that an unexpected, one-time intervention of the Fed in Treasury markets, as represented by a temporary Fed demand shock in our model, has relatively weak effects on bond yields. Plausibly, in the early stages of QE in the aftermath of the Great Financial Crisis, market participants were unsure about the nature of the Fed's unconventional monetary policies and only learned about the Fed's stance over time in view of the repeated interventions through Q1 to Q3, and then Q4. A commitment by the Fed to intervene in Treasury markets and its willingness to maintain a large balance sheet can then be more realistically represented by a permanent demand shock in our model. The effects of permanent demand shocks on bond yields in the model are indeed large, especially on the long end of the yield curve. A plausible narrative of the Fed's stance in recent years is that by providing forward guidance regarding its future policies, it credibly committed to future interventions, and thereby shaped investor expectations in the bond market and boost bond prices.

8. Conclusion

In this paper, we estimate an equilibrium model of the Treasury market that features preferred habitat investors, the Federal Reserve, and a risk-averse arbitrageur. For this purpose, we collect a novel dataset that covers the vast majority of the Treasury market with granular holdings of different investors, including insurance companies, banks, money market funds, broker-dealers, and the Federal Reserve. We estimate the demand functions of each investor group, and we find that there are strong empirical regularities in price elasticities and cross elasticities, as well as how investor demand depends on macroeconomic dynamics. Our empirical approach adds to the extant demand-based asset pricing literature by incorporating cross elasticities, modeling macroeconomic dynamics, and extracting demand information from both cross section and time series.

We use our rich empirical estimates as inputs to the equilibrium model, and estimate the model to fit Treasury yield curve dynamics, as well as average holdings by arbitrageurs. Our model features three key innovations. First, the model accounts for cross substitution of demand in a tractable way. Second, the model captures monetary policy rules and central bank balance sheet as functions of macro states, and how they drive Treasury demand across different sectors. Third, we allow for outside assets by arbitrageurs, which is critical for the equilibrium model to generate reasonable yield curve dynamics.

Our quantitative analyses reveal three main findings. First, we find a very low risk aversion of arbitrageurs, which contributes to an elastic Treasury market. Absent from risk-tolerant arbitrageurs, price impact becomes 14 times higher on average. Second, contrary to Vayanos and Vila (2021), we find that monetary policy tightening increases risk premia, because of high estimated cross elasticities of preferred-habitat demand. Third, we find that market expectations about the persistence of Fed QE purchase is critical for the impact of QE.

We view our paper as building a framework for combining novel data with equilibrium demand-based models to investigate important macro-finance issues in the government bond market. Future research can build on our approach and either incorporate this demand view of Treasury pricing into macroeconomic models that endogenize fiscal capacity or connect to granular data on other fixed-income assets such as corporate bonds to investigate the origin of the Treasury convenience yields.

References

Acharya, V. V. and Laarits, T. (2023). When do treasuries earn the convenience yield?: A hedging perspective.

Adrian, T., Etula, E., and Muir, T. (2014). Financial intermediaries and the cross-section of asset returns. *The Journal of Finance*, 69(6):2557–2596.

Allen, J., Kastl, J., and Wittwer, M. (2020). Estimating demand systems for treasuries. Technical report, Working paper.

An, Y. and Huber, A. (2024). Intermediary elasticity. *Available at SSRN 4825151*.

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary economics*, 50(4):745–787.

Bahaj, S., Czech, R., Ding, S., and Reis, R. (2023). The market for inflation risk. Technical report, Bank of England.

Bauer, M. D., Bernanke, B. S., and Milstein, E. (2023). Risk appetite and the risk-taking channel of monetary policy. *Journal of Economic Perspectives*, 37(1):77–100.

Becker, B. and Ivashina, V. (2015). Reaching for yield in the bond market. *The Journal of Finance*, 70(5):1863–1902.

Bekaert, G., Hoerova, M., and Duca, M. L. (2013). Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60(7):771–788.

Bernanke, B. S. and Kuttner, K. N. (2005). What explains the stock market's reaction to federal reserve policy? *The Journal of finance*, 60(3):1221–1257.

Bikbov, R. and Chernov, M. (2010). No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics*, 159(1):166–182.

Bretscher, L., Schmid, L., Sen, I., and Sharma, V. (2020). Institutional corporate bond pricing. *Swiss Finance Institute Research Paper*, (21-07).

Brunnermeier, M. K., Merkel, S., and Sannikov, Y. (2024). Safe assets. *Journal of Political Economy*.

Chaudhary, M., Fu, Z., and Li, J. (2022). Corporate bond multipliers: Substitutes matter. *Available at SSRN 4292266*.

Clarida, R., Gali, J., and Gertler, M. (2000). Monetary policy rules and macroeconomic stability: evidence and some theory. *The Quarterly Journal of Economics*, 115(1):147–180.

Culbertson, J. (1957). The term structure of interest rates. *The Quarterly Journal of Economics*, 71(4):485–517.

d'Amico, S., English, W., López-Salido, D., and Nelson, E. (2012). The federal reserve's large-scale asset purchase programmes: rationale and effects. *The Economic Journal*, 122(564):F415–F446.

Darmouni, O., Siani, K., and Xiao, K. (2022). Nonbank fragility in credit markets: Evidence from a two-layer asset demand system. *Available at SSRN 4288695*.

d'Avernas, A., Petersen, D., and Vandeweyer, Q. (2023). The central bank's balance sheet and treasury market disruptions.

d'Avernas, A. and Vandeweyer, Q. (2023). Treasury bill shortages and the pricing of short-term assets. *The Journal of Finance*.

Del Negro, M., Eggertsson, G., Ferrero, A., and Kiyotaki, N. (2017). The great escape? a quantitative evaluation of the fed's liquidity facilities. *American Economic Review*, 107(3):824–857.

Di Tella, S., Hébert, B., Kurlat, P., and Wang, Q. (2023). The zero-beta rate. Technical report, Working paper, Stanford Graduate School of Business.

Diamond, W. (2020). Safety transformation and the structure of the financial system. *The Journal of Finance*, 75(6):2973–3012.

Diamond, W., Jiang, Z., and Ma, Y. (2024). The reserve supply channel of unconventional monetary policy. *Journal of Financial Economics*, 159:103887.

Diamond, W. and Van Tassel, P. (2021). Risk-free rates and convenience yields around the world. *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Doerr, S., Eren, E., and Malamud, S. (2023). Money market funds and the pricing of near-money assets. *Swiss Finance Institute Research Paper*, (23-04).

Drechsler, I., Savov, A., and Schnabl, P. (2018). A model of monetary policy and risk premia. *The Journal of Finance*, 73(1):317–373.

Droste, M., Gorodnichenko, Y., and Ray, W. (2021). Unbundling quantitative easing: Taking a cue from treasury auctions. *Manuscript, UC Berkeley*.

Du, W., Hébert, B., and Huber, A. W. (2023a). Are intermediary constraints priced? *The Review of Financial Studies*, 36(4):1464–1507.

Du, W., Hébert, B., and Li, W. (2023b). Intermediary balance sheets and the treasury yield curve. *Journal of Financial Economics*, 150(3):103722.

Du, W., Tepper, A., and Verdelhan, A. (2018). Deviations from covered interest rate parity. *The Journal of Finance*, 73(3):915–957.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance*, 57(1):405–443.

Duffie, D. (2020). Still the world's safe haven? Redesigning the US Treasury market after the COVID-19 crisis'. Hutchins Center on Fiscal and Monetary Policy at Brookings.

Duffie, D., Fleming, M. J., Keane, F., Nelson, C., Shachar, O., and Van Tassel, P. (2023). Dealer capacity and us treasury market functionality. Technical report, Staff Report.

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates. *Mathematical finance*, 6(4):379–406.

Eren, E., Schrimpf, A., and Xia, F. D. (2023). The demand for government debt. *Available at SSRN 4466154*.

Fang, X., Hardy, B., and Lewis, K. K. (2022). Who holds sovereign debt and why it matters. Technical report, National Bureau of Economic Research.

Fang, X. and Liu, Y. (2021). Volatility, intermediaries, and exchange rates. *Journal of Financial Economics*, 141(1):217–233.

Favara, G., Infante, S., and Rezende, M. (2022). Leverage regulations and treasury market participation: Evidence from credit line drawdowns. *Working Paper*.

Gabaix, X. and Koijen, R. S. (2021). In search of the origins of financial fluctuations: The inelastic markets hypothesis. Technical report, National Bureau of Economic Research.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Gilchrist, S. and Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American economic review*, 102(4):1692–1720.

Greenwood, R., Hanson, S., and Vayanos, D. (2023). Supply and demand and the term structure of interest rates. Technical report, National Bureau of Economic Research.

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

Guibaud, S., Nosbusch, Y., and Vayanos, D. (2013). Bond market clienteles, the yield curve, and the optimal maturity structure of government debt. *The Review of Financial Studies*, 26(8):1914–1961.

Gulati, C. and Smith, A. L. (2022). The evolving role of the fed's balance sheet: Effects and challenges. *Economic Review (01612387)*, 107(4).

Haddad, V., Moreira, A., and Muir, T. (2024). Asset purchase rules: How qe transformed the bond market.

Haddad, V. and Muir, T. (2021). Do intermediaries matter for aggregate asset prices? *The Journal of Finance*, 76(6):2719–2761.

Haddad, V. and Sraer, D. (2020). The banking view of bond risk premia. *The Journal of Finance*, 75(5):2465–2502.

Hanson, S. G., Malkhozov, A., and Venter, G. (2024). Demand-and-supply imbalance risk and long-term swap spreads. *Journal of Financial Economics*, 154:103814.

Hanson, S. G. and Stein, J. C. (2015). Monetary policy and long-term real rates. *Journal of Financial Economics*, 115(3):429–448.

He, Z., Kelly, B., and Manela, A. (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 126(1):1–35.

He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. *American Economic Review*, 103:732–770.

He, Z., Nagel, S., and Song, Z. (2022). Treasury inconvenience yields during the covid-19 crisis. *Journal of Financial Economics*, 143(1):57–79.

Jansen, K. A. (2022). Long-term investors, demand shifts, and yields. *Available at SSRN 3901466*.

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

Jiang, W., Sargent, T. J., Wang, N., and Yang, J. (2024a). A p theory of government debt and taxes. *The Journal of Finance*.

Jiang, Z., Krishnamurthy, A., and Lustig, H. (2021). Foreign safe asset demand and the dollar exchange rate. *The Journal of Finance*, 76(3):1049–1089.

Jiang, Z., Lustig, H., Van Nieuwerburgh, S., and Xiaolan, M. Z. (2024b). The us public debt valuation puzzle. *Econometrica*, 92(4):1309–1347.

Jiang, Z., Richmond, R. J., and Zhang, T. (2022). A portfolio approach to global imbalances. Technical report, National Bureau of Economic Research.

Jiang, Z. and Sun, J. (2024). Quantitative tightening with slow-moving capital. Technical report, National Bureau of Economic Research.

Joslin, S., Priebsch, M., and Singleton, K. J. (2014). Risk premiums in dynamic term structure models with unspanned macro risks. *The Journal of Finance*, 69(3):1197–1233.

Joslin, S., Singleton, K. J., and Zhu, H. (2011). A new perspective on gaussian dynamic term structure models. *The Review of Financial Studies*, 24(3):926–970.

Kargar, M. (2021). Heterogeneous intermediary asset pricing. *Journal of Financial Economics*, 141(2):505–532.

Kekre, R., Lenel, M., and Mainardi, F. (2024). Monetary policy, segmentation, and the term structure. Technical report, National Bureau of Economic Research.

Khetan, U., Li, J., Neamtu, I., and Sen, I. (2023). The market for sharing interest rate risk: Quantities and asset prices. Technical report, Working papers, SSRN.

Koijen, R. and Yogo, M. (2019). A demand system approach to asset pricing. *Journal of Political Economy*, 127(4):1475–1515.

Koijen, R. S., Koulischer, F., Nguyen, B., and Yogo, M. (2021). Inspecting the mechanism of quantitative easing in the euro area. *Journal of Financial Economics*, 140(1):1–20.

Koijen, R. S. and Yogo, M. (2020). Exchange rates and asset prices in a global demand system. *National Bureau of Economic Research*.

Krishnamurthy, A. and Li, W. (2023). The demand for money, near-money, and treasury bonds. *The Review of Financial Studies*, 36(5):2091–2130.

Krishnamurthy, A. and Muir, T. (2017). How credit cycles across a financial crisis. Technical report, National Bureau of Economic Research.

Krishnamurthy, A. and Vissing-Jorgensen, A. (2007). The demand for treasury debt.

Krishnamurthy, A. and Vissing-Jorgensen, A. (2011). The effects of quantitative easing on interest rates: channels and implications for policy.

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

Li, W. (2024). Public liquidity and financial crises. *American Economic Journal: Macroeconomics*.

Li, W., Ma, Y., and Zhao, Y. (2023). The passthrough of treasury supply to bank deposit funding. *Columbia Business School Research Paper, USC Marshall School of Business Research Paper*.

Maggiori, M., Neiman, B., and Schreger, J. (2020). International currencies and capital allocation. *Journal of Political Economy*, 128(6):2019–2066.

Modigliani, F. and Sutch, R. (1966). Innovations in interest rate policy. *The American Economic Review*, 56(1/2):178–197.

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

Payne, J. and Szőke, B. (2024). Convenience yields and financial repression.

Payne, J., Szőke, B., Hall, G., and Sargent, T. J. (2022). Costs of financing us federal debt under a gold standard: 1791-1933.

Selgrad, J. (2023). Testing the portfolio rebalancing channel of quantitative easing.

Siani, K. (2022). Raising bond capital in segmented markets. *Available at SSRN 4239841*.

Stein, J. C. and Wallen, J. (2023). The imperfect intermediation of money-like assets. Technical report, Working Paper.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear iv regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, chapter 5. Cambridge: Cambridge Univ. Press.

Swanson, E. T. (2011). Let's twist again: a high-frequency event-study analysis of operation twist and its implications for qe2. *Brookings Papers on Economic Activity*, 2011(1):151–188.

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

Tabova, A. M. and Warnock, F. E. (2021). Foreign investors and us treasuries. Technical report, National Bureau of Economic Research.

Taylor, J. B. (1993). Discretion versus policy rules in practice. In *Carnegie-Rochester conference series on public policy*, volume 39, pages 195–214. Elsevier.

Van der Beck, P. (2022). On the estimation of demand-based asset pricing models. *Swiss Finance Institute Research Paper*, (22-67).

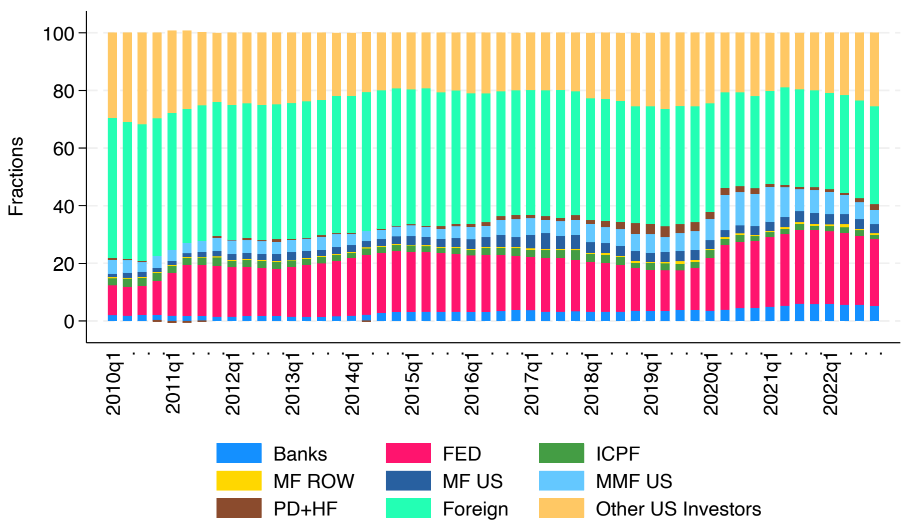
Vayanos, D. and Vila, J.-L. (2021). A preferred-habitat model of the term structure of interest rates. *Econometrica*, 89(1):77–112.

Wallen, J. (2020). *Markups to financial intermediation in foreign exchange markets*. Stanford University.

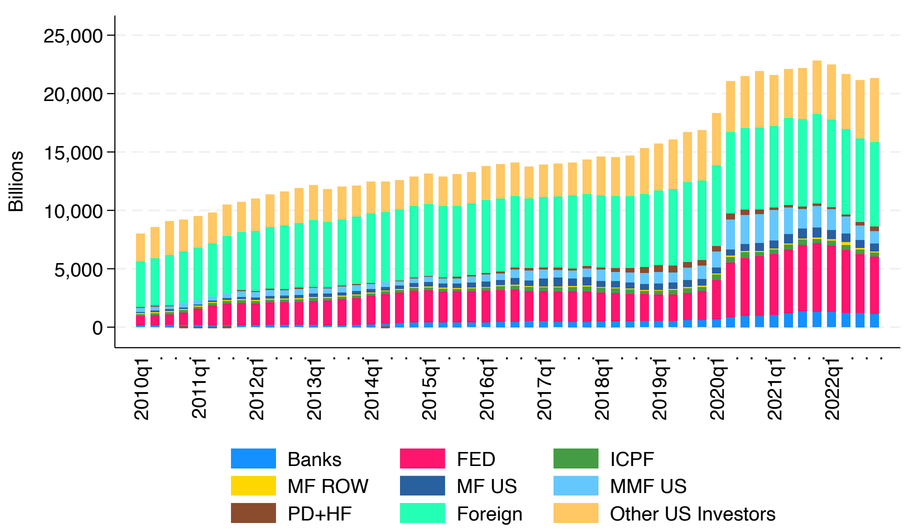
Figures and Tables

Figure 1. **Holdings of US Treasuries by Investor Type**

Panel (a) shows the fraction of the total US Treasuries' outstanding held by each investor type over time. Panel (b) shows the total dollar value of US Treasuries (billions) held by each investor type over time. The sectors are US banks (Banks), Fed, US insurance companies and pension funds (ICPF), mutual funds outside the US (MF ROW), US mutual funds (MF US), US money market funds (MMF US), US primary dealers and hedge funds (PD+HF), foreign investors (Foreign), and the residual sector (Residual). We subtract the MF ROW holdings from the foreign investor holdings. The residual sector is defined as the total US Treasuries' outstanding minus the holdings of all the other sectors. We report market values and the quarterly sample period is 2010Q1-2022Q4.



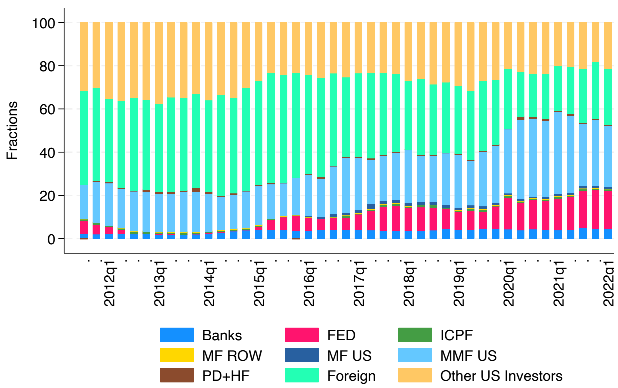
(a) % Total Holdings of TAO



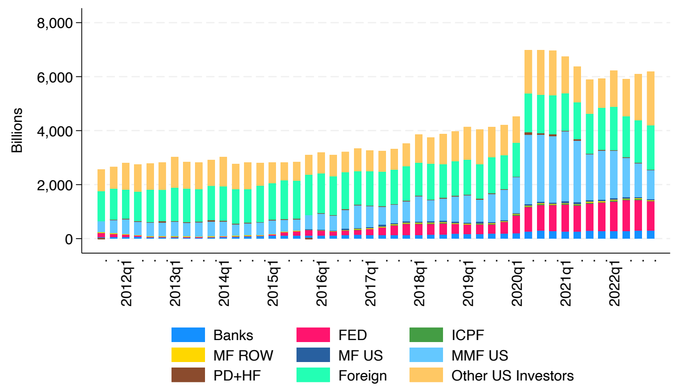
(b) Total Holdings (billions)

Figure 2. **Holdings of US Treasuries by Maturity Bucket**

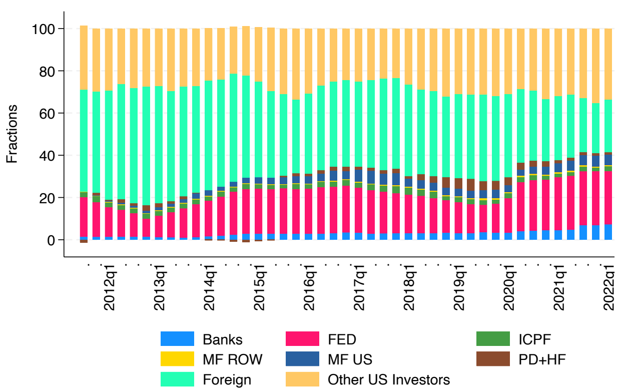
These graphs display the fraction of total US Treasuries' outstanding held by each investor type by maturity buckets: $\tau < 1Y$, $1 \leq \tau < 5Y$, and $\tau \geq 5Y$. We also report the total dollar value US Treasuries (billions) held by each investor type. The sectors are US banks (Banks), Fed, US insurance companies and pension funds (ICPF), mutual funds outside the US (MF ROW), US mutual funds (MF US), US money market funds (MMF US), US primary dealers and hedge funds (PD+HF), foreign investors (Foreign), and the residual sector (Residual). We subtract the MF ROW holdings from the foreign investor holdings. The residual sector is defined as the total US Treasuries' outstanding minus the holdings of all the other sectors. We report market values and the quarterly sample period is 2011Q3-2022Q4.



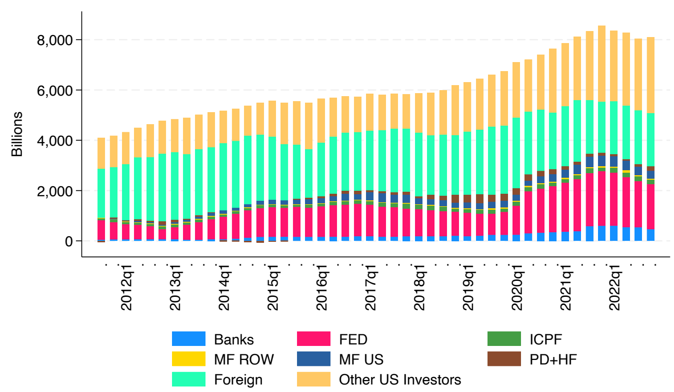
(a) $\tau < 1Y$: % Total Holdings of TAO



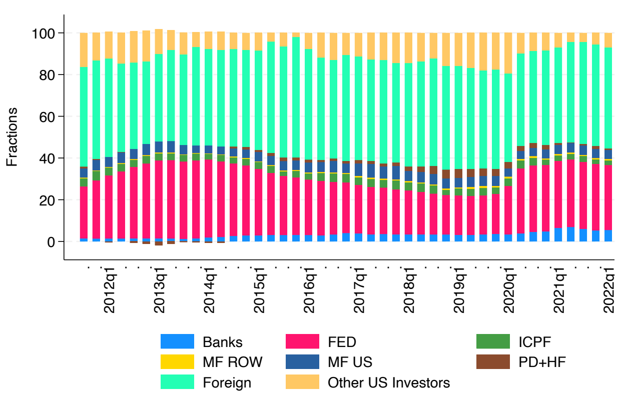
(b) $\tau < 1Y$: Total Holdings (billions)



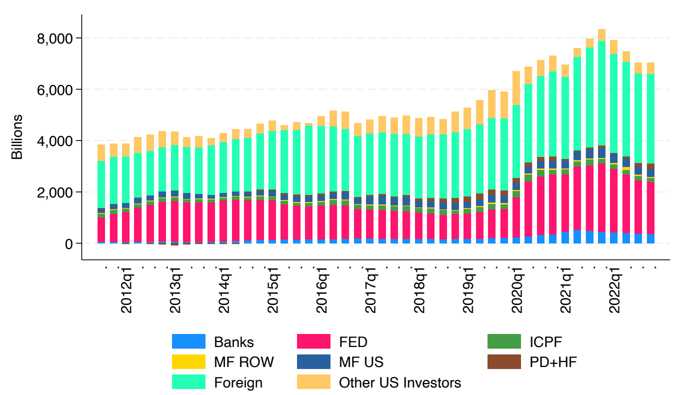
(c) $1 \leq \tau < 5Y$: % Total Holdings of TAO



(d) $1 \leq \tau < 5Y$: Total Holdings (billions)



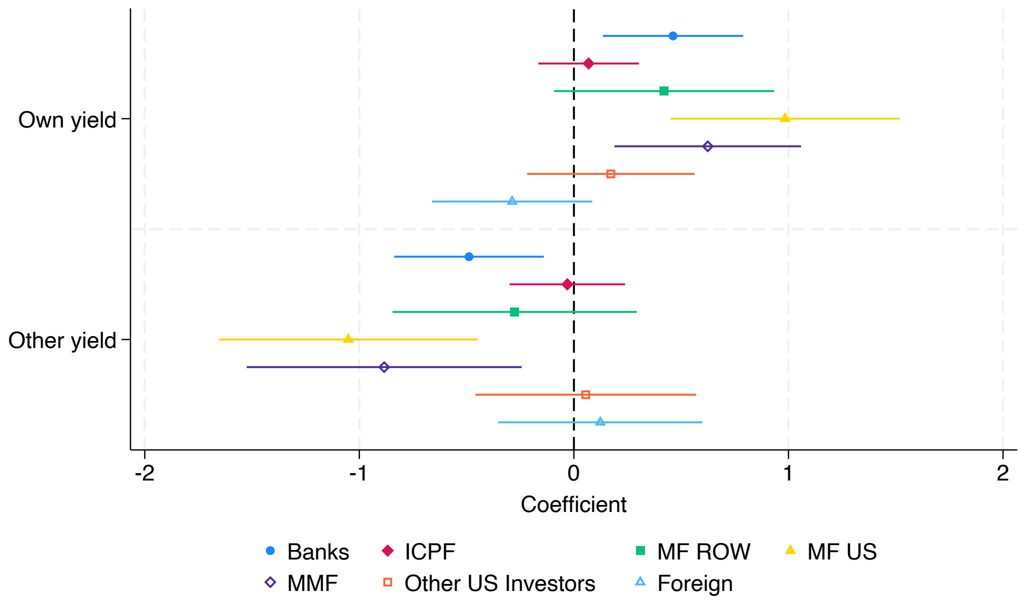
(e) $\tau \geq 5Y$: % Total Holdings of TAO



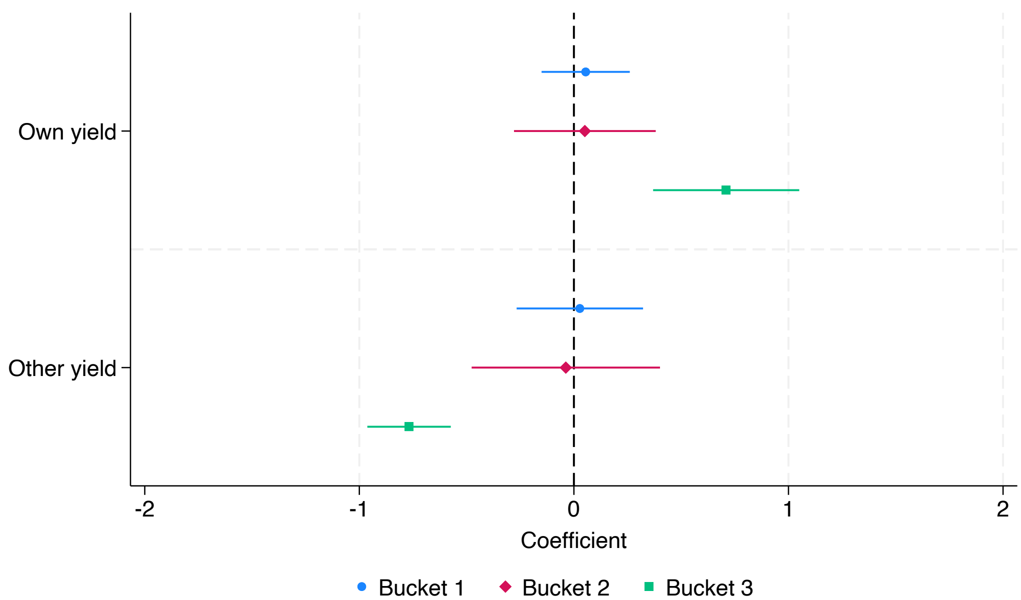
(f) $\tau \geq 5Y$: Total Holdings (billions)

Figure 3. **Yield Elasticities by Investor Type**

Panel (a) plots the coefficients on own and other yield for different preferred habitat investors, using standardized holdings to allow for comparison of coefficients across investor types. The sectors are US banks, insurance companies and pension funds (ICPF), mutual funds outside the US (MF ROW), US mutual funds (MF US), US money market funds (MMF US), other US investors (Other US investors), and foreign investors (Foreign). Panel (b) shows the yield sensitivities for the Federal Reserve by maturity bucket. We use market values and the quarterly sample period is 2011Q3-2022Q4.



(a) Preferred Habitat Investors



(b) Federal Reserve

Figure 4. **Model Fit on the Dynamics of Treasury Yields.**

Model predicted yields are constructed using equation (26) without latent demand (setting $u_t = 0$).

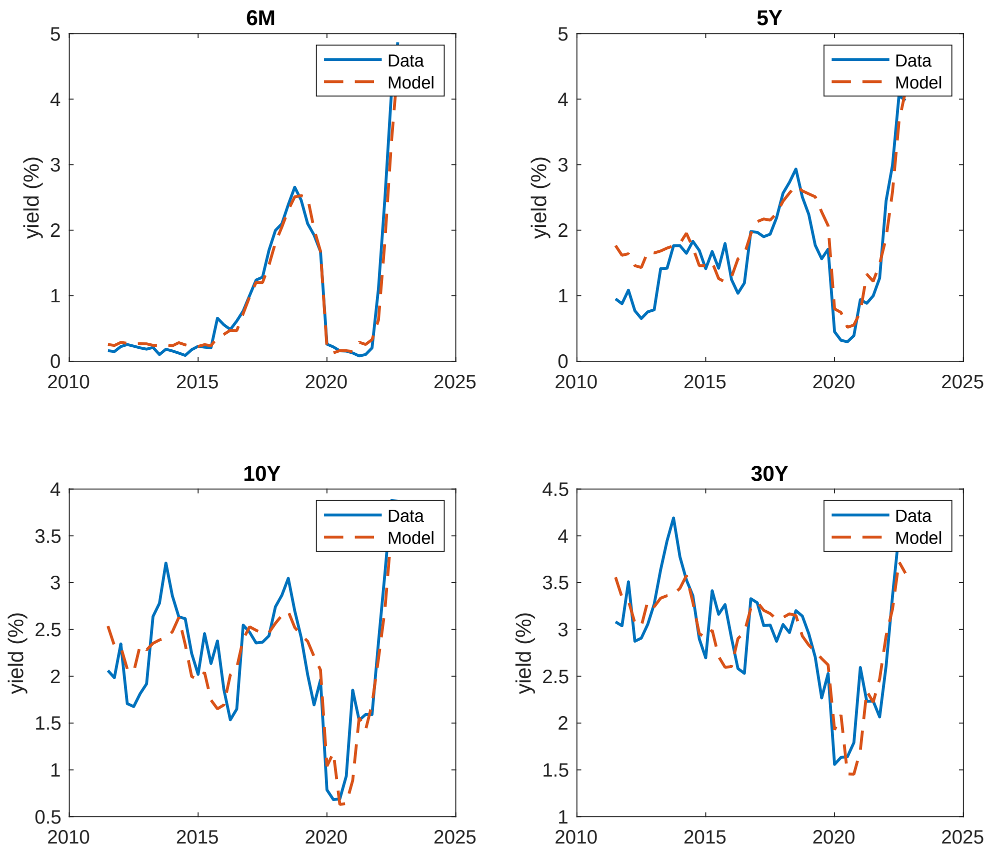
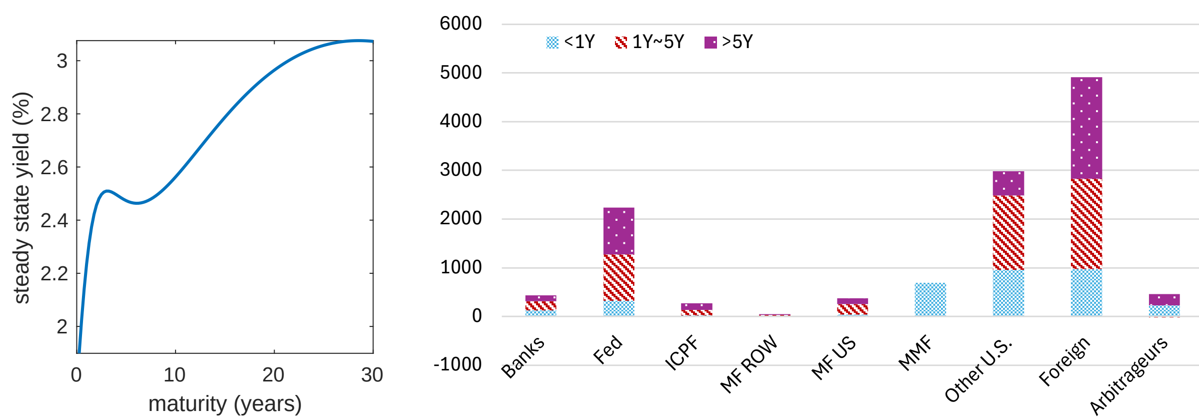


Figure 5. **Steady State.**

This figure illustrates the yield curve and portfolio allocations at the steady state, defined as the state where all shocks are zero. the left panel illustrates the steady-state yield curve. The right panel illustrates the steady-state portfolio holdings (in billions of dollars) for each group of investors and maturity bucket.

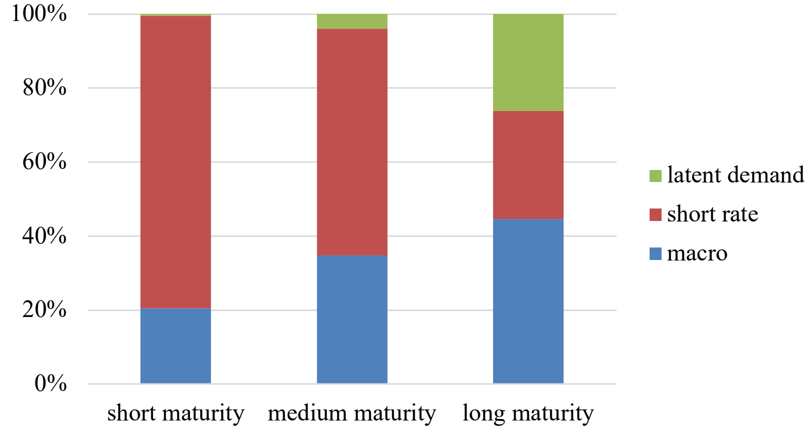


(a) Steady state yield curve

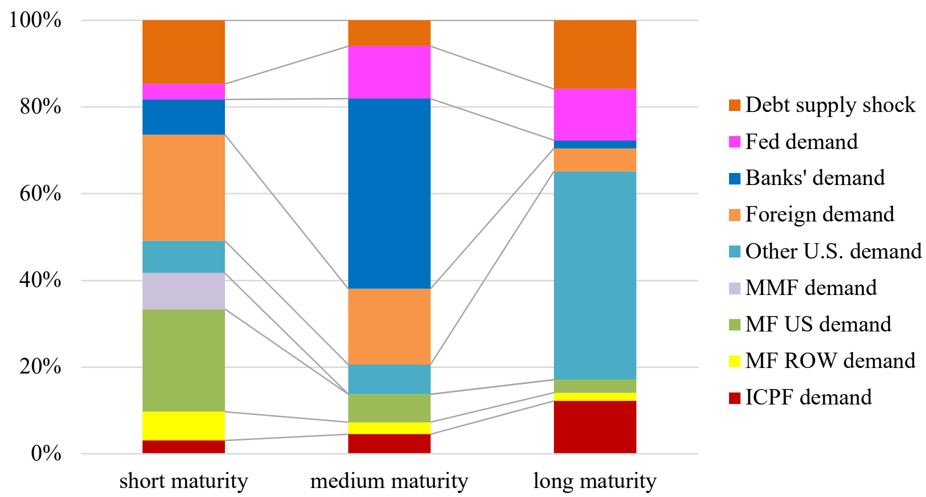
(b) Steady state portfolio allocations

Figure 6. **Decomposition of Treasury Yield Variation.**

In this figure, we decompose Treasury yield variations. In panel (a), we show the relative contribution of macroeconomic factors, FFR, and latent demand to the variation in Treasury yields, using the relative magnitude of their Shapley values of R^2, which is calculated as the average marginal contribution of each variable to the R^2 among all possible sets of dependent variable combinations. In panel (b), we focus on how these different factors take effect through the supply and demand forces in the model, by regressing one-quarter difference in Treasury yields on shocks to aggregate supply and sector-level demand. Note that panel (b) removes predictable components of supply and demand and only focus on shocks.



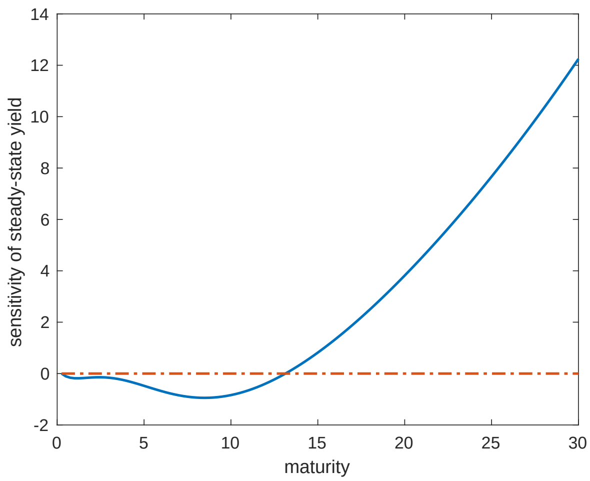
(a) Decomposition into macroeconomic, short rate, and latent demand variations.



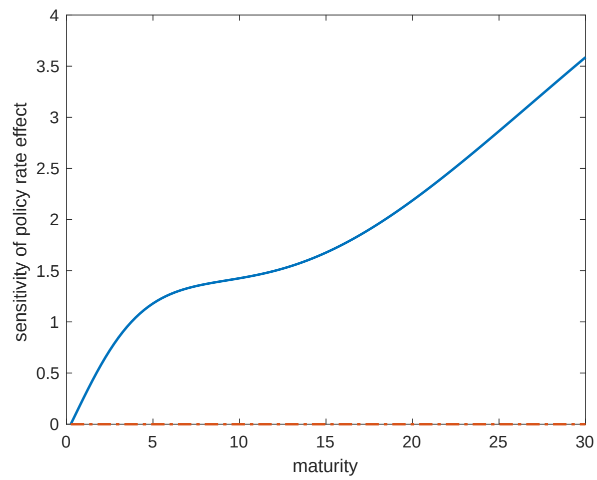
(b) Contributions of sector-level demand and aggregate supply shocks.

Figure 7. **Impact of Arbitrageur Risk Aversion.**

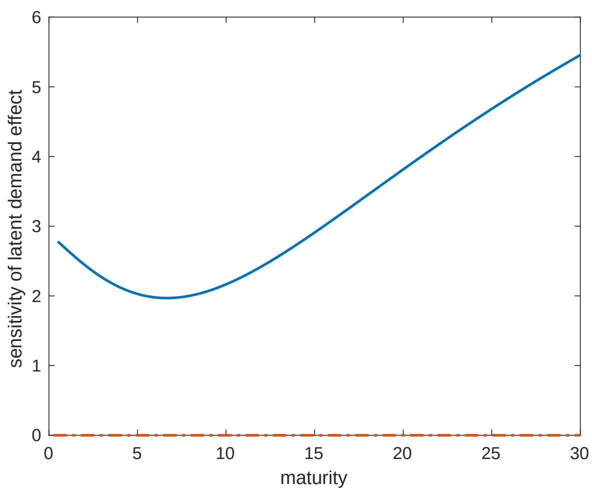
We illustrate how arbitrageur risk aversion γ affects the equilibrium. In particular, we increase γ by 1%, and illustrate the percentage changes of various yields and sensitivities. Panel (a) shows $\frac{\partial \log(y^{\text{steady}})}{\partial \log(\gamma)}$, the sensitivity of steady state yield curve. Panel (b) illustrates $\frac{\partial \log(|A_r|)}{\partial \log(\gamma)}$, the price elasticity to interest rate r. Panel (c) shows $\frac{\partial}{\partial \log(\gamma)} \log(\frac{1}{N} A_u \cdot \mathbf{1})$, the price elasticity to latent demand averaged across latent demand shocks of different maturities. Panel (d) shows $\frac{\partial}{\partial \log(\gamma)} \log(\frac{1}{N} \frac{\partial p}{\partial \theta_0} \cdot \mathbf{1})$, the price elasticity to permanent demand averaged across permanent demand shocks of different maturities.



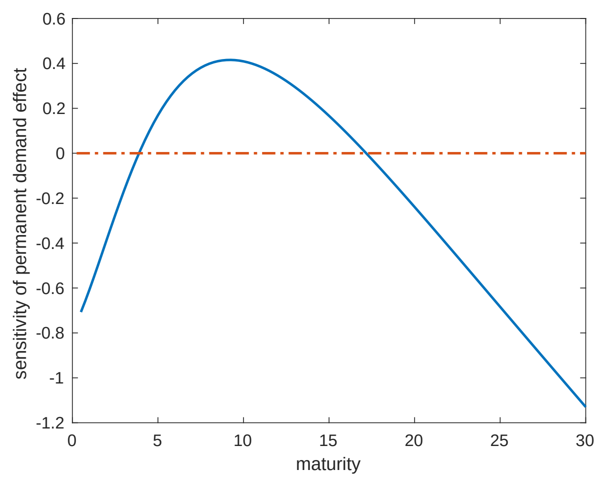
(a) Sensitivity of steady-state yield



(b) Sensitivity of policy rate effect



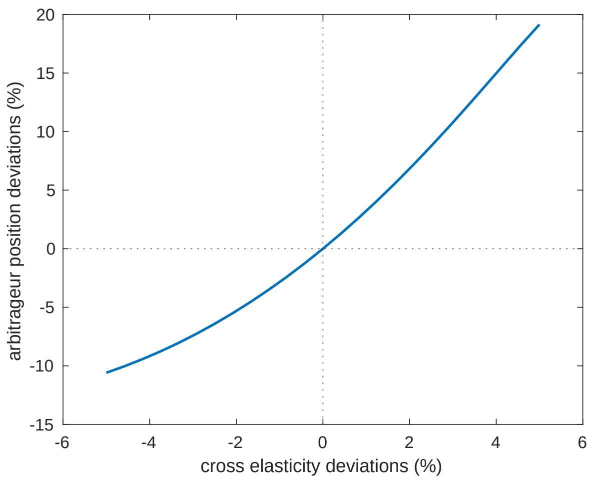
(c) Sensitivity of latent demand effect



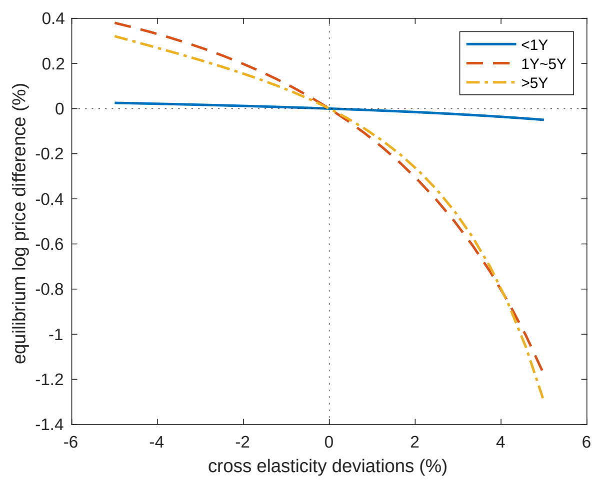
(d) Sensitivity of permanent demand effect

Figure 8. **Impact of Cross Elasticity.**

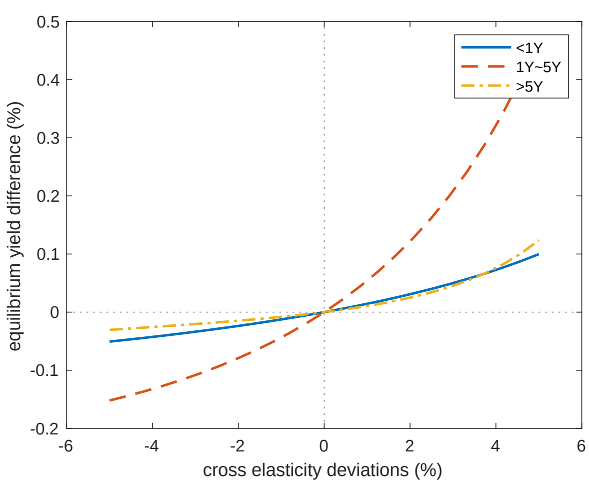
We illustrate the impact of cross elasticities. In particular, we change all cross elasticities by percentage deviations (x-axis), and illustrate the percentage deviations of various quantities and yields. In panel (a), we show the percentage deviation of arbitrageur steady-state Treasury holdings from baseline. In panel (b), we show the log price difference from baseline, expressed in %. In panel (c), we illustrate equilibrium yield difference from baseline, expressed in %. In panel (d), we measure the interest-rate sensitivity gap, which is the difference of yield sensitivities between short-maturity and long-maturity Treasuries (always positive because short-maturity yield is more sensitive to the short rate). We show percentage deviations of this measure from the baseline.



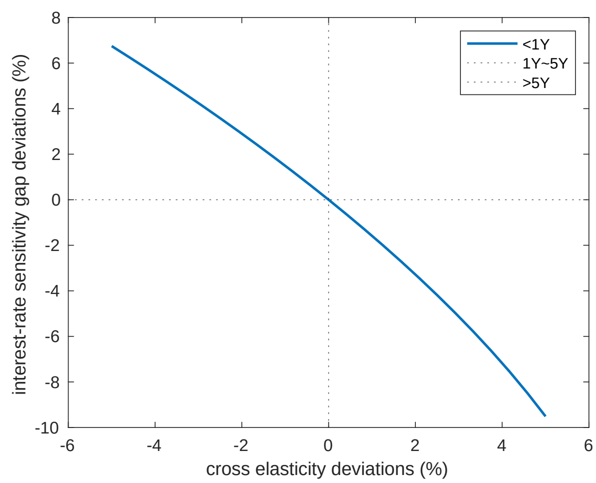
(a) Arbitrageur holding



(b) Steady-state log price (difference in %)



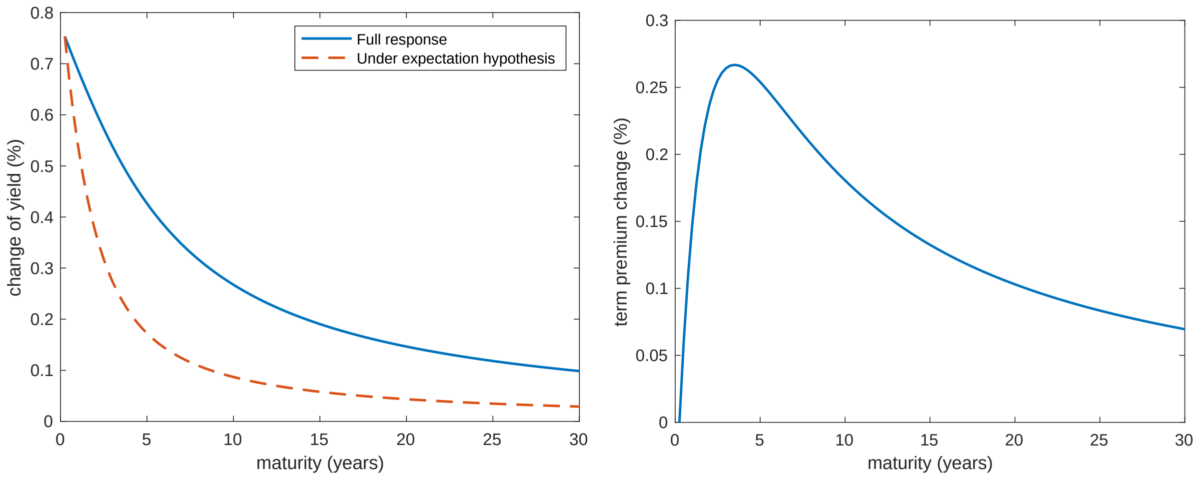
(c) Steady-state yield



(d) Interest-rate sensitivity gap

Figure 9. **Contemporaneous Yield Curve Response to a Monetary Policy Shock.**

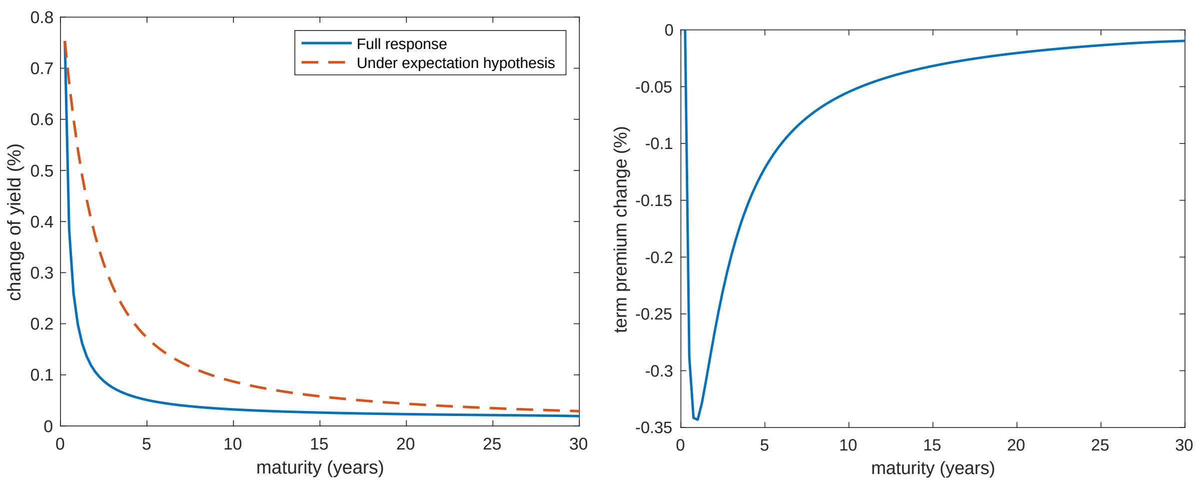
This figure illustrates impact of a one standard deviation monetary policy shock ($\varepsilon_t^r = 1$). The left panel illustrates the full response of the yield curve and also the component due to expected future short rate. The right panel illustrates the response of term premium, which is the difference between the two curves in the left panel.



(a) Yield curve response (b) Term premium response

Figure 10. **Contemporaneous Yield Curve Response to a Monetary Policy Shock, Shutting off Cross Elasticity.**

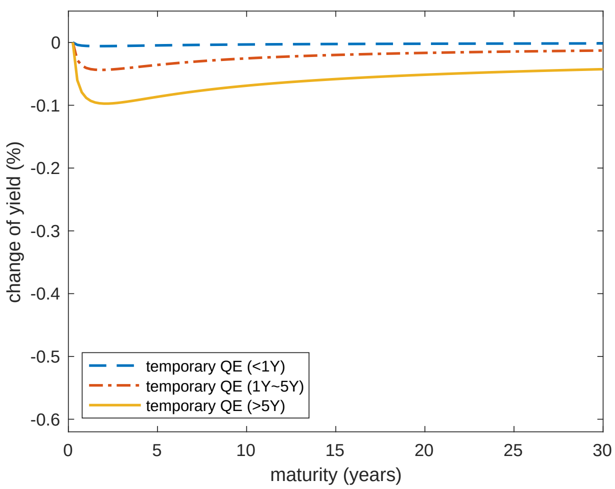
This figure illustrates impact of a one standard deviation monetary policy shock ($\varepsilon_t^r = 1$), but shutting off cross elasticity by setting $\alpha(\tau,\tau') = 0$ for any $\tau \neq \tau'$. The left panel illustrates the full response of the yield curve and also the component due to expected future short rate. The right panel illustrates the response of term premium, which is the difference between the two curves in the left panel.



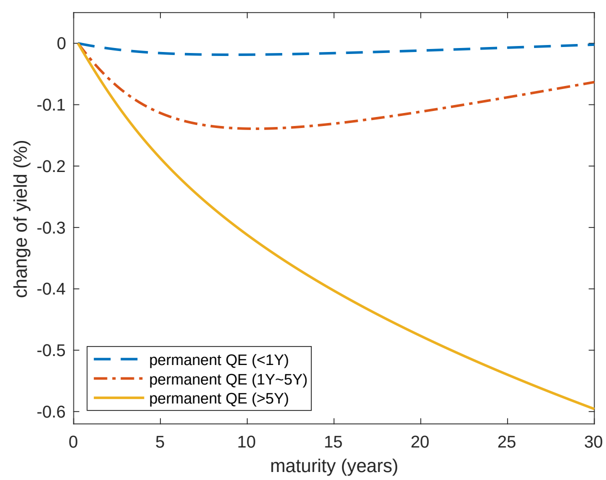
(a) Yield curve response (b) Term premium response

Figure 11. **Impact of QE Shocks on Treasury Yields.**

This figure illustrates how a \$100 billion QE shock on different maturity buckets, either temporary (left panel, increasing latent demand u_t) or permanent (right panel, increasing permanent demand θ_0), affects Treasury yields. For dollar values, we use the stationary model unit as described in Section 5.3.



(a) Temporary QE Policy



(b) Permanent QE Policy

Table 2. **Summary Statistics**

This table provides summary statistics of the main variables of interest: $y_t(m)$, which is the value-weighted yield of maturity bucket m, $y_t(-m)$, which is the value-weighted yield of the other maturity buckets excluding maturity bucket m, coupon rate, bid-ask spread, Supply (measured as Debt/GDP), Credit Spread, GDP gap, Core Inflation, and the Repo Spread.

	mean	sd	min	max
$y_t(m)$	1.387	1.079	0.041	4.291
$y_t(-m)$	1.456	0.898	0.132	4.289
Coupon	2.053	0.896	0.750	4.258
Bid-Ask Spread	0.046	0.028	0.010	0.096
Credit Spread	0.954	0.233	0.550	1.490
Supply	0.760	0.096	0.640	0.974
GDP Gap	-1.530	1.815	-9.207	1.342
Core Inflation	2.450	1.309	1.173	6.429
Repo Spread	-0.180	0.173	-0.456	0.230

Table 3. **Demand System Results - IV**

This table shows the IV estimates of our demand system specified in Equation (1). The dependent variable is the market value of US Treasuries held by sector i in maturity bucket m at time t, adjusted for GDP growth. The endogenous variables are: $y_t(m)$, which is the value-weighted yield of maturity bucket m, $y_t(-m)$, which is the value-weighted yield of the other maturity buckets excluding maturity bucket m. We instrument own and other yield using pseudo yields specified in Equation (4). Additional variables include Coupon rate, Bid-Ask Spread, indicator variable if the holdings are in maturity bucket 2 ($\mathbb{1}\{1Y \leq \tau < 5\}$), indicator variable if the holdings are in maturity bucket 3 ($\mathbb{1}\{\tau \geq 5\}$), Credit Spread, Supply (measured as Debt/GDP), Credit Spread, GDP gap, Core Inflation, and the Repo Spread. We orthogonalize the coupon and the bid-ask spread with respect to the maturity fixed effects. The quarterly sample period is from 2011Q3-2022Q4. HAC standard errors with optimal lags are reported in brackets; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Banks	**ICPF**	**MF ROW**	**MF US**	**MMF**	**Other US Investors**	**Foreign**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$y_t(m)$	36.408**	4.698	5.995	91.634***	200.521**	82.352	-149.749
	[15.743]	[7.247]	[4.315]	[29.848]	[80.576]	[110.642]	[115.887]
$y_t(-m)$	-39.208**	-1.355	-3.711	-96.909***	-249.203**	26.208	62.506
	[16.632]	[8.175]	[4.837]	[34.195]	[125.315]	[142.032]	[149.515]
Coupon	-78.355***	0.641	-11.225*	-85.442**	55.895	114.547	-414.631***
	[21.228]	[11.714]	[5.742]	[36.352]	[269.002]	[159.535]	[159.146]
Bid-Ask Spread	5.703	7.933***	0.870	1.147	61.998	-22.293	-1.293
	[4.399]	[2.640]	[1.568]	[10.225]	[61.720]	[39.002]	[43.439]
$\mathbb{1}\{1Y \leq \tau < 5\}$	34.773***	90.672***	11.849***	111.611***		536.661***	961.296***
	[9.245]	[2.673]	[2.388]	[16.976]		[57.879]	[85.441]
$\mathbb{1}\{\tau \geq 5\}$	-36.630	109.861***	8.455	10.485		-502.807**	1207.448***
	[28.337]	[12.925]	[8.636]	[57.911]		[235.195]	[227.333]
Credit Spread	9.362	-11.268	3.294	-24.194	-313.294***	128.570	70.563
	[12.513]	[8.084]	[2.418]	[24.696]	[83.748]	[111.048]	[121.835]
Supply	404.715***	38.963	62.416***	48.745	4125.721***	1649.323***	-937.677**
	[50.499]	[31.106]	[13.035]	[94.955]	[436.312]	[387.438]	[461.683]
GDP Gap	5.186**	-1.183	1.248**	8.458**	-29.528***	1.483	-6.433
	[2.245]	[1.223]	[0.493]	[3.731]	[8.441]	[16.080]	[14.612]
Core Inflation	13.638***	-4.657**	-0.832	-4.791	-34.297	-41.928	-19.302
	[4.587]	[2.351]	[0.930]	[7.389]	[32.818]	[27.019]	[37.736]
Repo Spread	16.748	-32.311***	-7.745	-21.335	-409.850***	-141.323	77.286
	[18.287]	[10.757]	[5.456]	[36.091]	[140.149]	[140.951]	[172.406]
Constant	-187.444***	7.136	-41.223***	110.327	-1879.445***	-545.681	1800.327***
	[46.326]	[27.119]	[13.011]	[97.928]	[428.062]	[408.287]	[423.657]
Observations	138	138	138	138	46	138	138
Eigenvalue Kleibergen-Paap statistic (*first stage*)	10.9	10.9	10.9	10.9	5.2	10.9	10.9

Table 4. **Demand System Results - Fed**

This table shows the OLS and IV estimates of our demand system specified in Equation (1) for the Fed. The dependent variable is the market value of US Treasuries held by sector i in maturity bucket m at time t, adjusted for GDP growth. The endogenous variables are: $y_t(m)$, which is the value-weighted yield of maturity bucket m, $y_t(-m)$, which is the value-weighted yield of the other maturity buckets excluding maturity bucket m. We instrument own and other yield using pseudo yields specified in Equation (4). Additional variables include Coupon rate, Bid-Ask Spread, Credit Spread, Supply (measured as Debt/GDP), Credit Spread, GDP gap, Core Inflation, and the Repo Spread. Column (1) shows the results for $\tau < 1Y$, Column (2) for $1Y \leq \tau < 5$, and Column (3) for $\tau \geq 5$. The quarterly sample period is from 2011Q3-2022Q4. HAC standard errors with optimal lags are reported in brackets; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	$\mathbb{1}\{\tau < 1Y\}$	$\mathbb{1}\{1Y \leq \tau < 5\}$	$\mathbb{1}\{\tau \geq 5\}$
	(1)	(2)	(3)
$y_t(m)$	4.555	80.302	342.187***
	[50.368]	[94.079]	[98.963]
$y_t(-m)$	57.529	-127.840	-373.290***
	[72.548]	[128.993]	[55.647]
Coupon	14.787	-1464.432***	14.916
	[127.408]	[164.344]	[167.977]
Bid-Ask Spread	130.564***	99.498***	-83.438**
	[39.448]	[36.702]	[36.420]
Credit Spread	-33.292	92.033	-25.731
	[47.733]	[58.457]	[105.429]
Supply	2248.161***	-670.480*	2614.501***
	[195.833]	[403.624]	[679.746]
GDP Gap	-4.885	-35.976***	-26.617**
	[4.553]	[8.680]	[12.336]
Core Inflation	27.370	26.071	76.759***
	[19.645]	[23.303]	[21.406]
Repo Spread	-145.130***	458.134***	-47.098
	[46.739]	[162.043]	[143.665]
R-squared	0.960	0.883	0.875
Observations	46	46	46
Kleibergen-Paap statistic (*first stage*)	5.2	15.69	29.99

Table 5. **Impact of Latent Demand Shocks on Treasury Prices with and without Arbitrageurs.**

We illustrate the impact of latent demand shocks with and without arbitrageurs. A value of 1 indicates that \$100 billion less latent demand of Treasuries increases the price by 1%.

Panel A: With Arbitrageur

	Price change (%) of		
	short maturity	medium maturity	long maturity
shock on short maturity	0.002	0.014	0.032
shock on medium maturity	0.014	0.105	0.253
shock on long maturity	0.030	0.237	0.688

Panel B: Without Arbitrageur

shock on short maturity	0.592	1.891	10.385
shock on medium maturity	2.203	10.731	48.664
shock on long maturity	0.499	2.006	13.032

Panel C: Price Impact Ratio (Panel B/Panel A)

shock on short maturity	308.49	133.33	320.70
shock on medium maturity	157.94	101.91	192.48
shock on long maturity	16.77	8.46	18.93

Table 6. **Portfolio Adjustment to Monetary Policy Shocks.**

This table illustrates portfolio adjustments to a one standard deviation shock to monetary policy, which is a 0.75% increase in one-period rate. We also report the model-implied change of Treasury supply in response to this shock. All units are in billions of dollars.

Sector	$\tau < 1$	$1 \leq \tau < 5$	$\tau \geq 5$	Total Change
Banks	10.2	2.8	-15.6	-2.7
ICPF	2.9	2.0	0.3	5.2
MF ROW	2.8	1.7	-0.8	3.7
MF U.S.	26.4	7.8	-38.2	-4.0
MMF	42.2	0.0	0.0	42.2
Other U.S.	71.1	58.5	37.2	166.8
Foreign	-83.2	-56.3	1.9	-137.6
Fed	24.2	5.9	-125.3	-95.2
Arbitrageurs	32.6	-52.8	41.8	21.6
total supply	129.1	-30.4	-98.7	0.0

Table 7. **Model-Implied Sector-Level Portfolio Adjustment to Latent Demand Shocks.**

This table illustrates how each sector adjusts their portfolio positions in response to a $100 billion sell shock of Treasuries at different maturity buckets, excluding the sell shock in the first place. The shock is temporary and will revert in one period.

Panel A: A $100 billion Sell Shock of $\tau < 1$ Treasuries

Sector	$\tau < 1$	$1 \leq \tau < 5$	$\tau \geq 5$	Total Change
Banks	0.0	0.1	-0.1	0.0
ICPF	0.0	0.0	0.0	0.0
MF ROW	0.0	0.0	0.0	0.0
MF U.S.	-0.1	0.2	-0.2	-0.1
MMF	-0.3	0.0	0.0	-0.3
Other U.S.	0.4	0.6	0.4	1.4
Foreign	-0.3	-0.6	-0.2	-1.1
Fed	0.1	0.1	-0.7	-0.5
Arbitrageurs	100.2	-0.3	0.7	100.6

Panel B: A $100 billion Sell Shock of $\tau \geq 15$ Treasuries

Sector	$\tau < 1$	$1 \leq \tau < 5$	$\tau \geq 5$	Total Change
Banks	-1.0	1.0	-0.8	-0.8
ICPF	0.2	0.4	0.2	0.7
MF ROW	0.1	0.3	0.1	0.5
MF U.S.	-2.4	2.6	-1.8	-1.6
MMF	-8.3	0.0	0.0	-8.3
Other U.S.	7.0	9.5	7.5	24.0
Foreign	-3.8	-10.3	-4.8	-18.9
Fed	2.0	1.2	-7.4	-4.2
Arbitrageurs	6.3	-4.6	106.9	108.6

Table 8. **Model-Implied Sector-Level Portfolio Adjustment to Permanent Demand Shocks.**

This table illustrates how each sector adjusts their portfolio positions in response to a permanent \$100 billion sell shock of Treasuries at different maturity buckets, excluding the sell shock in the first place.

Panel A: A \$100 billion Sell Shock of $\tau < 1$ Treasuries

Sector	$\tau < 1$	$1 \leq \tau < 5$	$\tau \geq 5$	Total Change
Banks	-0.5	-0.1	0.4	-0.2
ICPF	0.0	0.0	0.1	0.1
MF ROW	0.0	0.0	0.1	0.1
MF U.S.	-1.2	-0.2	1.1	-0.3
MMF	-3.3	0.0	0.0	-3.3
Other U.S.	0.5	1.1	1.7	3.3
Foreign	0.7	-0.8	-2.4	-2.5
Fed	0.0	0.1	2.8	2.9
Arbitrageurs	103.9	-0.2	-3.8	99.8

Panel B: A \$100 billion Sell Shock of $\tau \geq 15$ Treasuries

Sector	$\tau < 1$	$1 \leq \tau < 5$	$\tau \geq 5$	Total Change
Banks	-7.6	-3.8	9.1	-2.2
ICPF	-0.2	0.2	1.4	1.4
MF ROW	-0.7	-0.1	1.7	0.9
MF U.S.	-18.7	-9.1	23.1	-4.7
MMF	-48.4	0.0	0.0	-48.4
Other U.S.	6.4	13.3	28.2	47.8
Foreign	10.9	-3.2	-44.0	-36.3
Fed	0.4	-0.3	63.9	63.9
Arbitrageurs	57.9	3.0	16.7	77.6

Appendix

A. Additional Empirical Analysis

Table 9. **Correlation Table**

This table provides the correlation table of the main variables of interest: $y_t(m)$, which is the value-weighted yield of maturity bucket m, $y_t(-m)$, which is the value-weighted yield of the other maturity buckets excluding maturity bucket m, coupon rate, bid-ask spread, Supply (measured as Debt/GDP), Credit Spread, GDP gap, Core Inflation, and the Repo Spread. We orthogonalize the coupon and the bid-ask spread with respect to the maturity fixed effects.

	$y_t(m)$	$y_t(-m)$	Coupon	Bid-Ask Spread	Credit Spread	Supply	GDP Gap	Inflation	Repo Spread
$y_t(m)$	1								
$y_t(-m)$	0.581	1							
Coupon	-0.107	-0.296	1						
Bid-Ask Spread	0.021	-0.022	-0.308	1					
Credit Spread	-0.023	-0.041	0.309	-0.061	1				
Supply	-0.080	-0.128	-0.565	0.444	-0.161	1			
GDP Gap	0.428	0.507	-0.362	0.234	-0.308	0.131	1		
Inflation	0.403	0.490	-0.478	0.010	-0.021	0.429	0.461	1	
Repo Spread	-0.028	-0.047	-0.239	0.205	-0.386	0.263	0.247	-0.174	1

Figure 12. **Morningstar Aggregate Holdings by Domestic and Foreign Bond Funds**. This graph shows the aggregate holdings of US and foreign bond funds in USD (trillions) over time.

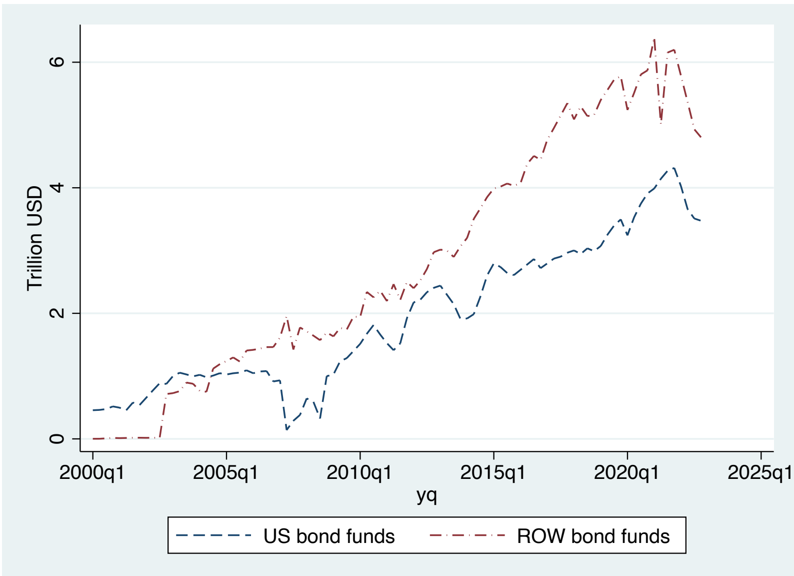


Table 10. **Demand System Results - OLS**

This table shows the OLS estimates of our demand system specified in Equation (1). The dependent variable is the market value of US Treasuries held by sector i in maturity bucket m at time t. The independent variables are: $y_t(m)$, which is the value-weighted yield of maturity bucket m, $y_t(-m)$, which is the value-weighted yield of the other maturity buckets excluding maturity bucket m, Coupon rate, Bid-Ask Spread, indicator variable if the holdings are in maturity bucket 2 ($\mathbb{1}\{1Y \leq \tau < 15\}$), indicator variable if the holdings are in maturity bucket 3 ($\mathbb{1}\{\tau \geq 15\}$), Credit Spread, Supply (measured as Debt/GDP), GDP gap, Core Inflation, and the Repo Spread. We orthogonalize the coupon and the bid-ask spread with respect to the maturity fixed effects. The quarterly sample period is from 2011Q3-2022Q4. HAC standard errors with optimal lags are reported in brackets; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Banks	**ICPF**	**MF ROW**	**MF US**	**MMF**	**Residual**	**Foreign**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$y\_t(m)$	33.693***	-1.344	2.715	24.270**	15.343	27.609	-189.358***
	[9.355]	[3.550]	[1.938]	[10.012]	[36.182]	[54.724]	[55.082]
$y\_t(-m)$	-34.052***	3.248	-0.567	-20.418**	60.959	18.223	104.212
	[9.753]	[4.315]	[2.312]	[9.815]	[48.134]	[70.659]	[81.179]
Coupon	-73.739***	5.533	-8.064**	-11.198	238.157	119.960	-373.536***
	[16.009]	[7.642]	[3.820]	[19.950]	[193.425]	[109.558]	[124.091]
Bid-Ask Spread	5.732	8.699***	1.220	7.116	24.230	-10.290	2.550
	[4.472]	[2.896]	[1.574]	[10.565]	[44.858]	[35.481]	[44.048]
$\mathbb{1}\{1Y \leq \tau < 5\}$	35.493***	92.933***	13.015***	134.400***		561.913***	975.010***
	[7.832]	[3.018]	[1.617]	[13.511]		[53.055]	[72.735]
$\mathbb{1}\{\tau \geq 5\}$	-30.545**	120.245***	14.390***	137.828***		-431.544***	1280.822***
	[15.386]	[7.879]	[4.154]	[17.075]		[123.500]	[111.837]
Credit Spread	8.184	-12.014	2.713	-39.377**	-176.795***	135.373	62.528
	[11.646]	[7.392]	[2.230]	[19.608]	[67.486]	[111.003]	[121.973]
Supply	421.600***	38.644	65.801***	183.985**	4948.067***	1372.296***	-876.191**
	[40.424]	[29.718]	[10.134]	[79.781]	[202.533]	[305.338]	[415.031]
GDP Gap	4.955**	-0.932	1.313***	8.455***	-26.426***	9.296	-6.071
	[2.216]	[1.194]	[0.487]	[3.032]	[7.884]	[15.477]	[14.723]
Core Inflation	12.963***	-3.724	-0.554	-3.305	-96.491***	-15.857	-17.270
	[4.411]	[2.283]	[0.866]	[5.807]	[18.468]	[27.299]	[38.702]
Repo Spread	16.195	-30.480***	-7.038	-12.125	-415.439***	-102.575	84.159
	[18.096]	[10.431]	[5.518]	[26.151]	[90.107]	[137.171]	[172.047]
R-squared	0.900	0.918	0.748	0.846	0.963	0.846	0.842
Observations	138	138	138	138	46	138	138

Table 11. **First Stage IV**

This table shows the first stage estimates of the IV methodology specified in Equation (1). The dependent variable in Column (1) is $y_t(m)$, the value-weighted yield of maturity bucket m and in Column (2) is $y_t(-m)$, the value-weighted yield of the other maturity buckets $-m$. We instrument own and other yield using pseudo yields specified in Equation (4). Additional variables include Coupon rate, Bid-Ask Spread, indicator variable if the holdings are in maturity bucket 2 ($\mathbb{1}\{1Y \leq \tau < 15\}$), indicator variable if the holdings are in maturity bucket 3 ($\mathbb{1}\{\tau \geq 15\}$), Credit Spread, Supply (measured as Debt/GDP), GDP gap, Core Inflation, and the Repo Spread. We orthogonalize the coupon and the bid-ask spread with respect to the maturity fixed effects. The quarterly sample period is from 2011Q3-2022Q4. HAC standard errors with optimal lags are reported in brackets; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	$y_t(m)$	$y_t(-m)$
	(1)	(2)
$I_t(m)$	0.700***	0.434***
	[0.049]	[0.046]
$I_t(-m)$	0.627***	0.808***
	[0.113]	[0.074]
Coupon	-0.421**	-0.913***
	[0.206]	[0.157]
Bid-Ask Spread	-0.155**	-0.189***
	[0.069]	[0.065]
$\mathbb{1}\{1Y \leq \tau < 5\}$	-1.046***	-0.153
	[0.316]	[0.221]
$\mathbb{1}\{\tau \geq 5\}$	-0.042	-0.855***
	[0.261]	[0.196]
Credit Spread	0.841***	0.891***
	[0.223]	[0.175]
Supply	1.508**	0.452
	[0.721]	[0.647]
GDP Gap	-0.211***	-0.205***
	[0.049]	[0.030]
Core Inflation	0.101	0.094*
	[0.063]	[0.049]
Repo Spread	-0.055	-0.137
	[0.287]	[0.247]
R-squared	0.882	0.894
Observations	138	138

Table 12. **Demand System Results - IV no macro**

This table shows the IV estimates of our demand system specified in Equation (1). The dependent variable is the market value of US Treasuries held by sector i in maturity bucket m at time t. The endogenous variables are: $y_t(m)$, which is the value-weighted yield of maturity bucket m, $y_t(-m)$, which is the value-weighted yield of the other maturity buckets excluding maturity bucket m. We instrument own and other yield using pseudo yields specified in Equation (4). Additional variables include Coupon rate, Bid-Ask Spread, indicator variable if the holdings are in maturity bucket 2 ($\mathbb{1}\{1Y \leq \tau < 15\}$), indicator variable if the holdings are in maturity bucket 3 ($\mathbb{1}\{\tau \geq 15\}$), and the Repo Spread. We orthogonalize the coupon and the bid-ask spread with respect to the maturity fixed effects. The quarterly sample period is from 2011Q3-2022Q4. HAC standard errors with optimal lags are reported in brackets; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Banks	**ICPF**	**MF ROW**	**MF US**	**MMF**	**Other US Investors**	**Foreign**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$y_t(m)$	90.027***	2.433	10.710**	94.762***	719.878***	166.869	-261.507**
	[23.719]	[6.559]	[4.754]	[30.714]	[269.867]	[123.217]	[130.615]
$y_t(-m)$	-97.818***	-4.778	-9.991*	-88.024***	-1372.927***	-169.309	195.784
	[26.674]	[7.875]	[5.121]	[32.051]	[360.334]	[145.503]	[144.247]
Coupon	-211.056***	1.989	-23.726***	-88.998***	-2997.786***	-153.830	-126.599
	[29.789]	[9.612]	[5.967]	[27.749]	[847.152]	[138.921]	[162.769]
Bid-Ask Spread	12.563	9.309***	3.150*	7.505	720.362***	28.829	-15.835
	[9.284]	[2.737]	[1.873]	[9.958]	[278.030]	[40.274]	[40.735]
$\mathbb{1}\{1Y \leq \tau < 5\}$	16.416	92.014***	10.343***	109.359***		517.622***	998.492***
	[14.828]	[3.062]	[2.676]	[18.624]		[59.785]	[79.383]
$\mathbb{1}\{\tau \geq 5\}$	-136.946***	111.388***	-0.884	10.285		-708.308***	1421.644***
	[46.060]	[12.157]	[9.500]	[58.648]		[256.301]	[244.401]
Repo Spread	-3.837	-19.439**	-4.208	22.537	-318.349	-103.738	66.997
	[33.789]	[8.774]	[6.406]	[34.030]	[583.965]	[125.698]	[143.605]
Observations	138	138	138	138	46	138	138
Kleibergen-Paap statistic (*first stage*):	17.56	17.56	17.56	17.56	5.76	17.56	17.56

Table 13. **Demand System Results - IV no other yield**

This table shows the IV estimates of our demand system specified in Equation (1), excluding other yield. The dependent variable is the market value of US Treasuries held by sector i in maturity bucket m at time t. The endogenous variable is $y_t(m)$, which is the value-weighted yield of maturity bucket m. We instrument own and other yield using pseudo yields specified in Equation (4). Additional variables include Coupon rate, Bid-Ask Spread, indicator variable if the holdings are in maturity bucket 2 ($\mathbb{1}\{1Y \leq \tau < 5\}$), indicator variable if the holdings are in maturity bucket 3 ($\mathbb{1}\{\tau \geq 5\}$), Credit Spread, Supply (measured as Debt/GDP), Credit Spread, GDP gap, Core Inflation, and the Repo Spread. We orthogonalize the coupon and the bid-ask spread with respect to the maturity fixed effects. The quarterly sample period is from 2011Q3-2022Q4. HAC standard errors with optimal lags are reported in brackets; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Banks	**ICPF**	**MF ROW**	**MF US**	**MMF**	**Other US Investors**	**Foreign**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$y_t(m)$	12.428*	3.869	3.725**	32.362***	55.119***	98.381***	-111.519***
	[7.545]	[3.046]	[1.683]	[10.998]	[17.090]	[36.755]	[42.923]
Coupon	-42.546***	1.879	-7.835**	3.064	218.783	90.612	-471.718***
	[13.644]	[8.268]	[3.222]	[20.031]	[200.838]	[95.222]	[90.384]
Bid-Ask Spread	6.606	7.964***	0.956	3.380	26.289	-22.897	-2.733
	[5.439]	[2.678]	[1.532]	[10.462]	[45.160]	[37.316]	[44.366]
$\mathbb{1}\{1Y \leq \tau < 5\}$	41.736***	90.913***	12.508***	128.821***		532.007***	950.196***
	[9.785]	[2.522]	[1.915]	[15.164]		[53.010]	[79.684]
$\mathbb{1}\{\tau \geq 5\}$	14.205	111.618***	13.267***	136.132***		-536.787***	1126.406***
	[11.778]	[5.240]	[3.053]	[18.583]		[80.011]	[85.718]
Credit Spread	0.688	-11.568	2.473	-45.633**	-207.283***	134.368	84.391
	[12.751]	[7.953]	[2.442]	[22.226]	[58.420]	[116.356]	[125.931]
Supply	518.931***	42.911	73.228***	331.050***	4795.023***	1572.976***	-1119.765***
	[52.662]	[28.658]	[10.713]	[97.921]	[187.138]	[328.926]	[433.022]
GDP Gap	3.851	-1.230	1.122**	5.157*	-27.575***	2.375	-4.304
	[2.356]	[1.261]	[0.517]	[3.076]	[7.532]	[16.329]	[14.103]
Core Inflation	9.921*	-4.786*	-1.184	-13.978**	-84.271***	-39.444	-13.376
	[5.325]	[2.551]	[0.922]	[6.995]	[16.570]	[27.196]	[34.666]
Repo Spread	14.688	-32.383***	-7.940	-26.426	-414.490***	-139.946	80.570
	[19.231]	[10.758]	[5.693]	[30.200]	[94.252]	[141.619]	[175.441]
Observations	138	138	138	138	46	138	138
Kleibergen-Paap statistic (*first stage*):	101.34	101.34	101.34	101.34	871.87	101.34	101.34

B. Model Derivations and Estimation

B.1. Proofs of Results in the Simple Model

Since the simple model is a special case of the main model, we can use the derivations for the main model to help with proofs in the simple model. In particular, we will rely on the iteration equations in (47), (48), (49), and (50).

Derivations of Equilibrium Treasury Prices in Equation (30)

First, we note that due to perfect arbitrage, we must have $p_t^{(1)} = -r_t$, so that $A(1) = 0$, $A_r(1) = -1$, $C(1) = 0$, and $A_u(1)' = (0,0)$. The holding return for 2-period Treasury bond as in (38) can be simplified as

$$R_{t+1}^{(2)} = -A(2)\cdot\beta_t + C(1) - C(2) - (\rho_r r_t + \sigma_r \varepsilon_{t+1}^r) - A_r(2) r_t + \frac{1}{2}\sigma_r^2 - A_u(2)' u_t \tag{58}$$

Next, we set $\tau = 2$ in the iteration equation for A_r in (48), which leads to

$$A_r(1)\rho_r - A_r(2) - 1 = A_r(1)\gamma\left(\sigma_r^2 A_r(1)\alpha(2)' A_r\right)$$

$$-\rho_r - A_r(2) - 1 = -\gamma\left(-\sigma_r^2(-b,\frac{a}{2})\begin{pmatrix} -1 \\ A_r(2)\end{pmatrix}\right)$$

$$-\rho_r - A_r(2) - 1 = \gamma\sigma_r^2\left(b + \frac{a}{2}A_r(2)\right).$$

Therefore,

$$A_r(2) = -\frac{1+\rho_r+\gamma\sigma_r^2 b}{1+\frac{1}{2}\gamma\sigma_r^2 a}.$$

To obtain $A(2)$, we set $\tau = 2$ in the iteration equation for A in (47),

$$\begin{aligned}
-A(2) &= A_r(1)\gamma\left(\sigma_r^2 A_r(1)(\zeta(2) + \alpha(2)'A + \theta(2))\right) \\
&= -\gamma\left(-\sigma_r^2(\zeta(2) + (-b,\frac{a}{2})\begin{pmatrix} 0 \\ A(2)\end{pmatrix} + \theta(2))\right) \\
&= \gamma\sigma_r^2(\frac{a}{2}A(2) + \zeta(2) + \theta(2))
\end{aligned}$$

which leads to

$$A(2) = -\frac{\gamma\sigma_r^2(\theta(2)+\zeta(2))}{1+\gamma\sigma_r^2\frac{a}{2}}.$$

Next, we solve for A_u. For $\tau = 2$, equation (49) leads to

$$-A_u(2)' = -\gamma\sigma_r^2 \left(-\alpha(2)' \begin{pmatrix} A_u(1)' \\ A_u(2)' \end{pmatrix} - (0, A_r(1)) \right)$$

$$-A_u(2)' = -\gamma\sigma_r^2 \left(-(-b, \frac{a}{2}) \begin{pmatrix} A_u(1)' \\ A_u(2)' \end{pmatrix} - (0, -1) \right)$$

$$A_u(2)' = \gamma\sigma_r^2 \left(-\frac{a}{2} A_u(2)' - (0, -1) \right)$$

$$A_u(2)' = \frac{1}{1 + \gamma\sigma_r^2 \frac{a}{2}} (0, \gamma\sigma_r^2)$$

Consequently, we obtain the A_u matrix as

$$A_u = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\gamma\sigma_r^2}{1+\gamma\sigma_r^2 \frac{a}{2}} \end{pmatrix}.$$

Then, we solve for $C(2)$ via setting $\tau = 2$ in equation (50),

$$\frac{1}{2}\sigma_r^2 + C(1) - C(2) = A_r(1)\gamma \left(\sigma_r^2 A_r(1)(\bar{S}^{(2)} - \theta_0(2) + \alpha(2)'C) \right)$$

$$\frac{1}{2}\sigma_r^2 - C(2) = \gamma\sigma_r^2 \left(\bar{S}^{(2)} - \theta_0(2) + (-b, \frac{a}{2}) \begin{pmatrix} 0 \\ C(2) \end{pmatrix} \right)$$

$$\begin{aligned} C(2) &= \frac{\frac{1}{2}\sigma_r^2 - \gamma\sigma_r^2 \bar{S}^{(2)} + \gamma\sigma_r^2 \theta_0(2)}{1 + \gamma\sigma_r^2 \frac{a}{2}} \\ &= \frac{\frac{1}{2} - \gamma\bar{S}^{(2)} + \gamma\theta_0(2)}{\frac{1}{\sigma_r^2} + \gamma\frac{a}{2}} \end{aligned}$$

Summarizing all the above, we obtain

$$p_t^{(2)} = -\frac{1 + \rho_r + \gamma\sigma_r^2 b}{1 + \frac{a}{2}\gamma\sigma_r^2} r_t - \frac{\gamma\sigma_r^2(\zeta(2) + \theta(2))}{1 + \frac{a}{2}\gamma\sigma_r^2} \beta_t + \frac{\gamma\sigma_r^2}{1 + \frac{a}{2}\gamma\sigma_r^2} u_t(2) + \frac{\frac{1}{2} - \gamma\bar{S}^{(2)} + \gamma\theta_0(2)}{\frac{1}{\sigma_r^2} + \frac{a}{2}\gamma},$$

which is identical to equation (30).

Proof of Proposition 1

According to equation (30), $p_t^{(1)}$ is entirely explained by r_t, while $p_t^{(2)}$ are also explained by β_t and $u_t(2)$. As a result, macro shocks and latent demand shocks are more important for long-maturity

Treasuries.

Proof of Proposition 2

To prove Proposition 2, we derive three important sensitivities.

$$\frac{\partial p_t^{(2)}}{\partial \beta_t} = -\frac{\gamma\sigma_r^2(\zeta(2)+\theta(2))}{1+\frac{a}{2}\gamma\sigma_r^2}$$
$$\frac{\partial p_t^{(2)}}{\partial u_t} = \frac{\gamma\sigma_r^2}{1+\frac{a}{2}\gamma\sigma_r^2}$$
$$\frac{\partial p_t^{(2)}}{\partial \theta_0(2)} = \frac{\gamma\sigma_r^2}{1+\frac{a}{2}\gamma\sigma_r^2}$$

These three sensitivities clearly increase in magnitude with γ. To show how price reacts to γ, we express the steady state (note that we assume $\bar{r}=0$ at the steady state) as

$$p^{(2)} = -\frac{\gamma\sigma_r^2(\zeta(2)+\theta(2))}{1+\frac{a}{2}\gamma\sigma_r^2}\bar{\beta} + \frac{\frac{1}{2}-\gamma\bar{S}^{(2)}+\gamma\theta_0(2)}{\frac{1}{\sigma_r^2}+\frac{a}{2}\gamma}.$$

Since the magnitude of the sensitivities over $\bar{\beta}$ decrease in γ, a higher γ contributes to lower $p^{\text{steady}}(2)$ through the first two components. If we further have $\bar{S}^{(2)} > \theta_0(2)$, the last term can be expressed as

$$\frac{\frac{1}{2}}{\frac{1}{\sigma_r^2}+\frac{a}{2}\gamma} - \frac{\bar{S}^{(2)}-\theta_0(2)}{\frac{1}{\gamma\sigma_r^2}+\frac{a}{2}},$$

which decreases with γ. Taken all of them together, we conclude that $p^{(2)}$ decreases with γ and therefore increases the long-term yield. Since the short rate is fixed at r_t, this increase of long-term yield also implies a higher term spread.

Finally, we did not state the results about interest-rate sensitivity, since the results depend on parameter cases. The price sensitivity to interest rate is

$$\begin{aligned}\frac{\partial p_t^{(2)}}{\partial r_t} &= -\frac{1+\rho_r+\gamma\sigma_r^2 b}{1+\frac{a}{2}\gamma\sigma_r^2}\\ &= -\left(\frac{1}{1+\frac{a}{2}\gamma\sigma_r^2}\underbrace{(1+\rho_r)}_{\text{expectation hypothesis}} + \frac{\frac{a}{2}\gamma\sigma_r^2}{1+\frac{a}{2}\gamma\sigma_r^2}\underbrace{\frac{2b}{a}}_{\text{pure habitat demand}}\right)\end{aligned}$$

As a result, whether this increases or decreases with γ depends on the relative magnitude of the sensitivity in the expectation hypothesis versus the pure demand case.

Proof of Proposition 3

According to equation (34) and (35), equilibrium portfolio allocations in response to cross elasticity are

$$\frac{\partial Z_t^{(2)}}{\partial b} = -\frac{1}{1+\frac{a}{2}\gamma\sigma_r^2} r_t < 0,$$

$$\frac{\partial X_t^{(2)}}{\partial b} = \frac{1}{1+\frac{a}{2}\gamma\sigma_r^2} r_t > 0.$$

Therefore, a higher cross elasticity in preferred-habitat demand leads to lower preferred-habitat investor holding and thus higher equilibrium arbitrageur position. he impact on equilibrium Treasury price according to (30) is

$$\frac{\partial p_t^{(2)}}{\partial b} = -\frac{\gamma\sigma_r^2}{1+\frac{a}{2}\gamma\sigma_r^2} r_t < 0,$$

so a higher cross elasticity lowers the equilibrium Treasury price $p_t^{(2)}$.

Next, the interest-rate sensitivity gap between long- and short-maturity Treasuries is

$$\begin{aligned}
|\frac{\partial y_t^{(2)}}{\partial r_t} - \frac{\partial y_t^{(1)}}{\partial r_t}| =& |\frac{\partial \left(-\frac{p_t^{(2)}}{2}\right)}{\partial r_t} - 1| \\
=& |\frac{1+\rho_r+\gamma\sigma_r^2 b}{2+a\gamma\sigma_r^2} - 1| \\
=& |\frac{1-\rho_r+\gamma\sigma_r^2(a-b)}{2+a\gamma\sigma_r^2}|
\end{aligned}$$

By assumption, $1-\rho_r+\gamma\sigma_r^2(a-b) > 0$, so when b increases, this gap is smaller.

Proof of Proposition 4

The expectation component of the long-term Treasury yield is

$$\bar{y}_t^{(2)} = \frac{1+\rho_r}{2} r_t$$

Using $y_t^{(2)} = -p_t^{(2)}/2$ and equation (30), we get the term premium expression

$$
\begin{aligned}
& y_t^{(2)} - \bar{y}_t^{(2)} \\
=& \left(\frac{1+\rho_r+\gamma\sigma_r^2 b}{2+a\gamma\sigma_r^2} - \frac{1}{2}(1+\rho_r) \right) r_t + \frac{\gamma\sigma_r^2(\zeta(2)+\theta(2))}{2+a\gamma\sigma_r^2}\beta_t - \frac{\gamma\sigma_r^2}{2+a\gamma\sigma_r^2}u_t(2) - \frac{\frac{1}{2} - \frac{\bar{r}}{\sigma_r^2} - \gamma\bar{S}^{(2)} + \gamma\theta_0(2)}{2+a\gamma\sigma_r^2} \\
=& \frac{b-\frac{1}{2}(1+\rho_r)a}{\frac{2}{\gamma\sigma_r^2}+a} r_t + \frac{\gamma\sigma_r^2(\zeta(2)+\theta(2))}{2+a\gamma\sigma_r^2}\beta_t - \frac{\gamma\sigma_r^2}{2+a\gamma\sigma_r^2}u_t(2) - \frac{\frac{1}{2} - \frac{\bar{r}}{\sigma_r^2} - \gamma\bar{S}^{(2)} + \gamma\theta_0(2)}{2+a\gamma\sigma_r^2}
\end{aligned}
$$

As a result,

$$
\frac{\partial(y_t^{(2)} - \bar{y}_t^{(2)})}{\partial r_t} = \frac{b-\frac{1}{2}(1+\rho_r)a}{\frac{2}{\gamma\sigma_r^2}+a}
$$

while the baseline response according to the expectation hypothesis is

$$
\frac{\partial \bar{y}_t^{(2)}}{\partial r_t} = \frac{1+\rho_r}{2} > 0
$$

Consequently, the full response is

$$
\frac{\partial y_t^{(2)}}{\partial r_t} = \underbrace{\frac{1+\rho_r}{2}}_{\text{expectation hypothesis}} + \underbrace{\frac{b-\frac{1}{2}(1+\rho_r)a}{\frac{2}{\gamma\sigma_r^2}+a}}_{\text{change of term premium}}
$$

When $2b > (1+\rho_r)a$, the term premium component is positive, so that the long-term Treasury yield over-reacts to monetary policy shock compared to the expectation hypothesis. When $2b < (1+\rho_r)a$, the term premium component is negative, so that the long-term Treasury yield under-reacts to monetary policy shock compared to the expectation hypothesis.

Proof of Proposition 5

The impact of QE on Treasury price, as reflected by the increase of the permanent demand $\theta_0(2)$, is as follows,

$$
\frac{\partial p_t^{(2)}}{\partial \theta_0(2)} = \frac{\gamma\sigma_r^2}{1+\frac{a}{2}\gamma\sigma_r^2} > 0.
$$

Therefore, Treasury prices increase with QE, which implies a decrease of Treasury yields.

B.2. Derivations for the Full Model

Wealth thus evolves as

$$
\begin{aligned}
W_{t+1} &= W_t(1+r_t) + \sum_{\tau=2}^{N} X_t^{(\tau)}(R_{t+1}^{(\tau)} - r_t) + \tilde{X}_t(\tilde{R}_{t,t+1} - r_t) \\
&= W_t(1+r_t) + \tilde{X}_t(\tilde{R}_{t,t+1} - r_t) + \frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1) \\
&\quad + \sum_{\tau=2}^{N} X_t^{(\tau)} \begin{pmatrix} A(\tau-1)'(\bar{\beta} + \Phi(\beta_t - \bar{\beta}) + \Sigma^{1/2}\varepsilon_{t+1}) - A(\tau)'\cdot\beta_t + \frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1) \\ +A_u(\tau-1)'u_{t+1} - A_u(\tau)'u_t + \frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1) \\ +A_r(\tau-1)(\bar{r} + \phi_r'(\Phi(\beta_t - \bar{\beta}) + \Sigma^{1/2}\varepsilon_{t+1}) + \rho_r r_t + \sigma_r \varepsilon_{t+1}^r) - A_r(\tau) r_t \\ +C(\tau-1) - C(\tau) + \frac{1}{2}(A_r(\tau-1)\sigma_r)^2 - r_t \end{pmatrix} \\
&= W_t(1+r_t) + \sum_{\tau=2}^{N} X_t^{(\tau)} \begin{pmatrix} A(\tau-1)'\left(\bar{\beta} + \Phi(\beta_t - \bar{\beta})\right) - A(\tau)'\beta_t + \frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1) \\ -A_u(\tau)'u_t + \frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1) + C(\tau-1) - C(\tau) \\ +A_r(\tau-1)(\bar{r} + \phi_r'\Phi(\beta_t - \bar{\beta}) + \rho_r r_t) - A_r(\tau) r_t + \frac{1}{2}(A_r(\tau-1)\sigma_r)^2 - r_t \end{pmatrix} \\
&\quad + \left(\sum_{\tau=2}^{N} X_t^{(\tau)} \left(A(\tau-1)'\Sigma^{1/2} + A_r(\tau-1)\phi_r'\Sigma^{1/2} \right) + \tilde{X}_t \tilde{\sigma}' \right) \varepsilon_{t+1} + \left(\sum_{\tau=2}^{N} X_t^{(\tau)} A_r(\tau-1)\sigma_r + \tilde{X}_t \tilde{\sigma}_r' \right) \varepsilon_{t+1}^r \\
&\quad + \left(\sum_{\tau=2}^{N} X_t^{(\tau)} A_u(\tau-1)' \right) u_{t+1} + \tilde{X}_t(\tilde{\phi}'\beta_t + \tilde{\phi}_r r_t - r_t)
\end{aligned}
\tag{59}
$$

To simplify notatations, it is convenient to define the expected return on Treasuries of maturity τ as

$$
\begin{aligned}
\mu_t^{(\tau)} &= A(\tau-1)'\left(\bar{\beta} + \Phi(\beta_t - \bar{\beta})\right) - A(\tau)'\beta_t + \frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1) - A_u(\tau)'u_t + C(\tau-1) - C(\tau) \\
&\quad + \frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1) + A_r(\tau-1)(\bar{r} + \phi_r'\Phi(\beta_t - \bar{\beta}) + \rho_r r_t) - A_r(\tau) r_t + \frac{1}{2}(A_r(\tau-1)\sigma_r)^2.
\end{aligned}
\tag{60}
$$

In that case, we obtain

$$
E_t[W_{t+1}] = W_t(1+r_t) + \sum_{\tau=2}^{N} X_t(\tau)\left(\mu_t^{(\tau)} - r_t\right) + \tilde{X}_t(\tilde{\phi}'\beta_t + \tilde{\phi}_r r_t - r_t),
$$

and we can write the variance as

$$
\begin{aligned}
Var_t(W_{t+1}) &= \left(\sum_{\tau=2}^{N} X_t(\tau)\hat{A}(\tau-1)'\Sigma^{1/2} + \tilde{X}_t\tilde{\sigma}'\right)\left(\sum_{\tau=2}^{N} X_t(\tau)\Sigma^{1/2}\hat{A}(\tau-1) + \tilde{X}_t\tilde{\sigma}\right) \\
&+ \left(\sum_{\tau=2}^{N} X_t(\tau)A_r(\tau-1)\sigma_r + \tilde{X}_t\tilde{\sigma}_r'\right)^2 + \left(\sum_{\tau=2}^{N} X_t(\tau)A_u(\tau-1)'(\Sigma^u)^{1/2}\right)\left((\Sigma^u)^{1/2}\sum_{\tau=2}^{N} X_t(\tau)A_u(\tau-1)\right) \\
&= \sum_{\tau=2}^{N} \hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1)(X_t(\tau))^2 + 2\sum_{\hat{\tau}\neq\tau} \hat{A}(\tau-1)'\Sigma\hat{A}(\hat{\tau}-1)X_t(\tau)X_t(\hat{\tau}) \\
&+ 2\sum_{\tau=2}^{N} \hat{A}(\tau-1)'\Sigma^{1/2}\tilde{\sigma}\cdot(X_t(\tau)\tilde{X}_t) + \tilde{\sigma}'\tilde{\sigma}(\tilde{X}_t)^2 + \left(\sum_{\tau=2}^{N} X_t(\tau)A_r(\tau-1)\sigma_r + \tilde{X}_t\tilde{\sigma}_r'\right)^2 \\
&+ \sum_{\tau=2}^{N} A_u(\tau-1)'\Sigma^u A_u(\tau-1)(X_t(\tau))^2 + 2\sum_{\hat{\tau}\neq\tau} A_u(\tau-1)'\Sigma^u A_u(\hat{\tau}-1)X_t(\tau)X_t(\hat{\tau}).
\end{aligned}
$$

Consequently, we can write the FOC of arbitrageurs as

$$
\begin{aligned}
\mu_t^{(\tau)} - r_t &= \gamma\left(\sum_{\hat{\tau}=2}^{N} \hat{A}(\tau-1)'\Sigma\hat{A}(\hat{\tau}-1)X_t(\hat{\tau}) + \hat{A}(\tau-1)'\Sigma^{1/2}\tilde{\sigma}\tilde{X}_t\right) \\
&+ \gamma\left(\sum_{\hat{\tau}=2}^{N} A_r(\tau-1)\sigma_r^2 A_r(\hat{\tau}-1)X_t(\hat{\tau}) + A_r(\tau-1)'\sigma_r\tilde{\sigma}_r\tilde{X}_t\right) \\
&+ \gamma\left(\sum_{\hat{\tau}=2}^{N} A_u(\tau-1)'\Sigma^u A_u(\hat{\tau}-1)X_t(\hat{\tau})\right) \\
&= \hat{A}(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\Sigma\hat{A}(\hat{\tau}-1)X_t(\hat{\tau})\right) + \Sigma^{1/2}\tilde{\sigma}\tilde{X}_t\right) \\
&+ A_r(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\sigma_r^2 A_r(\hat{\tau}-1)X_t(\hat{\tau})\right) + \sigma_r\tilde{\sigma}_r\tilde{X}_t\right) \\
&+ A_u(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N} \Sigma^u A_u(\hat{\tau}-1)X_t(\hat{\tau})\right)
\end{aligned}
$$

$$
\tilde{\phi}'\beta_t + \tilde{\phi}_r r_t - r_t = \gamma\left(\sum_{\tau=2}^{N} A(\tau-1)'\Sigma^{1/2}\tilde{\sigma}\cdot X_t(\tau) + \tilde{\sigma}'\tilde{\sigma} + \sum_{\tau=2}^{N} A_r(\tau-1)'\sigma_r\tilde{\sigma}_r\cdot X_t(\tau) + (\tilde{\sigma}_r)^2\right) \tag{61}
$$

Defining the prices of risk as

$$\lambda_{\beta,t} = \gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\Sigma\hat{A}(\hat{\tau}-1)X_t(\hat{\tau})\right)+\Sigma^{1/2}\tilde{\sigma}\tilde{X}_t\right), \tag{62}$$

$$\lambda_{r,t} = \gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\sigma_r^2 A_r(\hat{\tau}-1)X_t(\hat{\tau})\right)+\sigma_r\tilde{\sigma}_r\tilde{X}_t\right), \tag{63}$$

$$\lambda_{u,t} = \gamma\left(\sum_{\hat{\tau}=2}^{N}\Sigma^u A_u(\hat{\tau}-1)X_t(\hat{\tau})\right), \tag{64}$$

We can further simplify the first-order condition by expressing them in terms of matching coefficients, that is, by replacing the definition of $\mu_t^{(\tau)}$ in (60). Accordingly, we have

$$\begin{gathered} A(\tau-1)'\left(\bar{\beta}+\Phi(\beta_t-\bar{\beta})\right)-A(\tau)'\beta_t+\tfrac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1)+A_r(\tau-1)(\bar{r}+\phi_r'\Phi(\beta_t-\bar{\beta})+\rho_r r_t)\\ +C(\tau-1)-C(\tau)-A_r(\tau)r_t+\tfrac{1}{2}(A_r(\tau-1)\sigma_r)^2-A_u(\tau)'u_t+\tfrac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1)-r_t\\ =\hat{A}(\tau-1)'\lambda_{\beta,t}+A_r(\tau-1)\lambda_{r,t}+A_u(\tau-1)'\lambda_{u,t}. \end{gathered} \tag{65}$$

Ultimately, these coefficients are pinned down in equilibrium, that is, when markets clear. The market clearing condition is

$$Z_t(\tau)+X_t(\tau)=S_t(\tau).$$

As a result, our model implies that the price of risk λ_t varies over time, and depends on both the quantity of Treasury supply $S_t(\tau)$ and preferred habitat demand $Z_t(\tau)$, as well as the outside portfolio returns $\tilde{X}_t$.

As a next step, we plug in the expressions for $Z_t(\tau)$ in (12) and $S_t(\tau)$ in (17) and can express the equilibrium arbitrageur holdings as

$$X_t(\tau)=\left(\bar{S}(\tau)+\zeta(\tau)'\beta_t+\zeta_r(\tau)'r_t\right)-\left(\theta_0(\tau)-\alpha(\tau)'p_t-\theta(\tau)'\beta_t+u_t(\tau)\right). \tag{66}$$

In the main text, we impose the assumption that $A_u(\tau-1)'\lambda_{u,t}$ is dominated by the other two terms, $\hat{A}(\tau-1)'\lambda_{\beta,t}+A_r(\tau-1)\lambda_{r,t}$, which we will verify in Appendix C.4. The idea is that idiosyncratic latent demand shocks do not affect price of risks. Under this simplification assumption, plugging (66) into the pricing equation (65), we obtain

$$A(\tau-1)'\left(\bar{\beta}+\Phi(\beta_t-\bar{\beta})\right)-A(\tau)'\beta_t+\frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1)+C(\tau-1)-C(\tau) \tag{67}$$

$$A_r(\tau-1)(\bar{r}+\phi_r'\Phi(\beta_t-\bar{\beta})+\rho_r r_t)-A_r(\tau)r_t+\frac{1}{2}(A_r(\tau-1)\sigma_r)^2-A_u(\tau)'u_t+\frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1)-r_t$$

$$=\hat{A}(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\Sigma\hat{A}(\hat{\tau}-1)\left((\bar{S}(\hat{\tau})+\zeta(\hat{\tau})'\beta_t+\zeta_r(\tau)'r_t)-(\theta_0(\hat{\tau})-\alpha(\hat{\tau})'p_t-\theta(\hat{\tau})'\beta_t+u_t(\hat{\tau}))\right)\right)+\Sigma^{1/2}\tilde{\sigma}\tilde{X}_t\right)$$

$$+A_r(\tau-1)\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\sigma_r^2 A_r(\hat{\tau}-1)\left((\bar{S}(\hat{\tau})+\zeta(\hat{\tau})'\beta_t+\zeta_r(\tau)'r_t)-(\theta_0(\hat{\tau})-\alpha(\hat{\tau})'p_t-\theta(\hat{\tau})'\beta_t+u_t(\hat{\tau}))\right)\right)+\sigma_r\tilde{\sigma}_r\tilde{X}_t\right) \tag{68}$$

With the assumption in (46), and the affine expression of p_t in (26), we rewrite the equilibrium condition in (45) as follows:

$$A(\tau-1)'\left(\bar{\beta}+\Phi(\beta_t-\bar{\beta})\right)-A(\tau)'\beta_t+\frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1)+C(\tau-1)-C(\tau)$$

$$+A_r(\tau-1)(\bar{r}+\phi_r'\Phi(\beta_t-\bar{\beta})+\rho_r r_t)-A_r(\tau)r_t+\frac{1}{2}(A_r(\tau-1)\sigma_r)^2-A_u(\tau)'u_t+\frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1)-r_t$$

$$=\hat{A}(\tau-1)'\gamma\left(\begin{array}{l}\sum_{\hat{\tau}=2}^{N}\left(\Sigma\hat{A}(\hat{\tau}-1)\left(\begin{array}{c}(\bar{S}(\hat{\tau})+\zeta(\hat{\tau})'\beta_t+\zeta_r(\tau)'r_t)\\-(\theta_0(\hat{\tau})-\alpha(\hat{\tau})'(A\beta_t+A_r r_t+A_u u_t+C)-\theta(\hat{\tau})'\beta_t)-u_t(\hat{\tau})\end{array}\right)\right)\\+\Psi\beta_t+\Lambda r_t+\psi\end{array}\right)$$

$$+A_r(\tau-1)\gamma\left(\begin{array}{l}\sum_{\hat{\tau}=2}^{N}\left(\sigma_r^2 A_r(\hat{\tau}-1)\left(\begin{array}{c}(\bar{S}(\hat{\tau})+\zeta(\hat{\tau})'\beta_t+\zeta_r(\tau)'r_t)\\-(\theta_0(\hat{\tau})-\alpha(\hat{\tau})'(A\beta_t+A_r r_t+A_u u_t+C)-\theta(\hat{\tau})'\beta_t)-u_t(\hat{\tau})\end{array}\right)\right)\\+\Psi_r\beta_t+\Lambda_r r_t+\psi_r\end{array}\right) \tag{69}$$

Matching the coefficients on β_t, r_t, u_t, and the constant term, we obtain an iteration equation as follows:

$$
\begin{aligned}
&A(\tau-1)'\Phi - A(\tau)' + A_r(\tau-1)\phi_r'\Phi \\
&= \hat{A}(\tau-1)'\underbrace{\gamma\left(\left(\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\left(\zeta(\hat{\tau})'+\alpha(\hat{\tau})'A+\theta(\hat{\tau})'\right)\right)+\Psi\right)}_{\lambda_{\beta,\beta}} \\
&+ A_r(\tau-1)\underbrace{\gamma\left(\left(\sum_{\hat{\tau}=2}^{N}\sigma_r^2 A_r(\hat{\tau}-1)\left(\zeta(\hat{\tau})'+\alpha(\hat{\tau})'A+\theta(\hat{\tau})'\right)\right)+\Psi_r\right)}_{\lambda_{\beta,r}}
\end{aligned}
\tag{70}
$$

$$
\begin{aligned}
A_r(\tau-1)'\rho_r - A_r(\tau) - 1 &= \hat{A}(\tau-1)'\underbrace{\gamma\left(\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\left(\zeta_r(\tau)'+\alpha(\hat{\tau})'A_r\right)+\Lambda\right)}_{\lambda_{r,\beta}} \\
&+ A_r(\tau-1)'\underbrace{\gamma\left(\sum_{\hat{\tau}=2}^{N}\sigma_r^2 A_r(\hat{\tau}-1)\left(\zeta_r(\tau)'+\alpha(\hat{\tau})'A_r\right)+\Lambda_r\right)}_{\lambda_{r,r}}
\end{aligned}
\tag{71}
$$

$$
\begin{aligned}
-A_u(\tau)' &= \hat{A}(\tau-1)'\gamma\Sigma\left(\left(\sum_{\hat{\tau}=2}^{N}\hat{A}(\hat{\tau}-1)\alpha(\hat{\tau})'A_u\right)-\left(0,\hat{A}(1),...,\hat{A}(N-1)\right)\right) \\
&+ A_r(\tau-1)'\gamma\sigma_r^2\left(\left(\sum_{\hat{\tau}=2}^{N}A_r(\hat{\tau}-1)\alpha(\hat{\tau})'A_u\right)-(0,A_r(1),...,A_r(N-1))\right)
\end{aligned}
\tag{72}
$$

$$
\begin{aligned}
&A(\tau-1)'(I-\Phi)\bar{\beta} + A_r(\tau-1)(\bar{r}-\phi_r'\Phi\bar{\beta}) + \frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1) \\
&+ \frac{1}{2}(A_r(\tau-1)\sigma_r)^2 + \frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1) + C(\tau-1) - C(\tau) \\
&= \hat{A}(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\left(\bar{S}(\hat{\tau})-\left(\theta_0(\hat{\tau})-\alpha(\hat{\tau})'C\right)\right)+\psi\right) \\
&+ A_r(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\sigma_r^2 A_r(\hat{\tau}-1)\left(\bar{S}(\hat{\tau})-\left(\theta_0(\hat{\tau})-\alpha(\hat{\tau})'C\right)\right)+\psi_r\right).
\end{aligned}
\tag{73}
$$

B.3. Setting Model Parameters

The model is quite flexible accounting for the rich dependence of investor demand on macroeconomic factors and Treasury prices, as well as dynamics in the state variables. In this subsection, we provide details of how we use data to directly inform model parameters.

We take the average duration as the maturity for each maturity bucket, obtaining $\tau_1 = 2$, $\tau_2 = 10$, and $\tau_3 = 42$ (all in quarters). For each maturity bucket, we sum up the coefficients of preferred-habitat investors' demand in Table 3 and 4. To convert regression results to the model format, we express the demand for each maturity bucket separately, and use the intercept term to capture maturity-bucket fixed effects. We then add the maturity-by-maturity bucket estimates of the Fed to the preferred habitat investor demand to obtain total non-arbitrageur demand. For simplicity, our model does not capture characteristic-based demand (i.e., loadings on coupon rate and bid-ask spread) and the repo spread, so we take the average of these components and add them to the intercept of preferred-habitat demand. The above leads to the following pooled preferred-habitat demand:

Table 14. **Model inputs**

	$\tau < 1Y$	$1 \leq \tau < 5Y$	$\tau \geq 5Y$
$\theta_0(m)$	-2082	3847	741
$y_t(m)$	305	92	332
$y_t(-m)$	-302	-65	-355
Credit Spread	-178	253	182
GDP Gap	-30	-11	-22
Inflation	-43	-81	28
Supply	7372	1176	3791

Moreover, in the model, the demand is expressed as a function of prices, not yields, so we need to convert the yield sensitivity into price sensitivity, using the chain rule,

$$\frac{\partial Z(\tau)}{\partial p^\tau} = \frac{\partial Z(\tau)}{\partial y^\tau}\frac{\partial y^\tau}{\partial p^\tau} = -\frac{1}{\tau}\frac{\partial Z(\tau)}{\partial y^\tau} \tag{74}$$

Second, we estimate the supply dynamics in Equation (18). We implement a linear regression of the Treasury total supply in each maturity bucket and then recover the loadings on macro factors, the short rate, and the intercept $\bar{S}$. Similar to the demand estimation, we concentrate the supply into three maturities that represent the average duration of three maturity buckets. In Figure 13, we illustrate that the model fits the total supply well.

Third, we estimate the monetary policy dynamics in (10). We rewrite the monetary policy equation as

$$r_{t+1} = (\bar{r} - \phi_r'\bar{\beta}) + \phi_r'\beta_{t+1} + \rho_r r_t + \sigma_r \varepsilon_{t+1}^r, \tag{75}$$

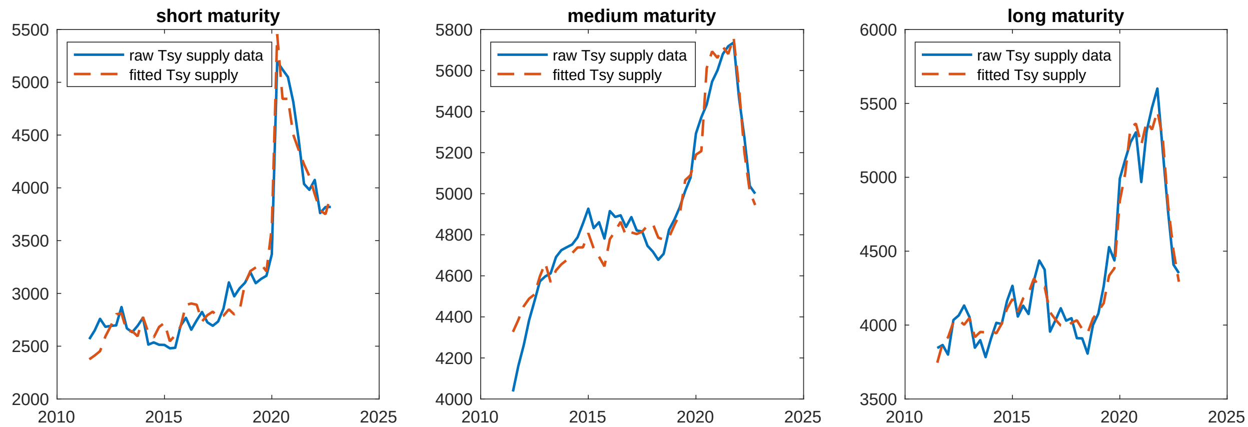


Figure 13. Treasury Supply: Data versus Model Fitting.

where the intercept term is identified as a whole. To fit the monetary policy rule, we have to use a longer time period, because monetary policy rate does not have much variation during our main sample period. In particular, we use the post-Volcker period (1990 to 2024) excluding the zero lower bound (ZLB) period (2008-2015). We start from 1990 because it is when the Fed gained credibility in its fight of inflation. The resulting monetary-policy equation is:

$$\begin{aligned} r_{t+1} = {} & 1.9 - 1.36 * \text{credit spread}_{t+1} + 0.06 * \text{GDP gap}_{t+1} + 0.22 * \text{core inflation}_{t+1} \\ & - 1.13 * \text{debt/GDP}_{t+1} + 0.78 * r_t + 0.75 * \varepsilon^r_{t+1} \end{aligned} \tag{76}$$

Equation (76) suggests that the Fed lowers the interest rate if credit spread is high, GDP gap (GDP deviation from potential GDP) is low and tightens interest rate if inflation is high. The coefficient on GDP gap and inflation have the same signs as the classical Taylor rule (Taylor 1993) but much smaller coefficients. Moreover, there is moderate amount of monetary policy inertia reflected by the coefficient of 0.78 on lagged policy rate. This dependence on lagged policy rate generates an impact of monetary policy rate on long-term yields from the expectation effect and is critical to understand how yield curve responds to monetary policy shocks ε^r_{t+1}.

Fourth, we estimate the dynamics of macro factors in Equation (9). It is important to get the long-run average of macroeconomic factors correct. Therefore, we take the sample average of macro factors directly as $\bar{\beta}$. Denote the demeaned macro factors as $\hat{\beta}_t$. Then we recover the coefficients with the following regression:

$$\tilde{\beta}_{t+1} = \Phi \tilde{\beta}_t + \Sigma^{1/2} \varepsilon_{t+1}. \tag{77}$$

Alternatively, we could directly run a linear regression with an intercept to uncover $\bar{\beta}$ and Φ simultaneously. We find that the estimations of Φ are similar between the two approaches, but the simultaneous estimation of $\bar{\beta}$ and Φ gives unreasonable long-run average of macro variables. The matrix Σ is estimated as the covariance matrix of the regression residuals in (77).

B.4. Model Estimation

Estimation of the model involves high dimensionality and requires a reasonable initialization of model parameters. Our high-level idea is to segregate the model into different components and separately initialize each component.

At the first step, we solve for the following simpler optimization problem with unconstrained C,

$$\min_{\{\lambda_{\beta,\beta},\lambda_{\beta,r},\lambda_{r,\beta},\lambda_{r,r},C\}} \mathbb{E}\left[\sum_t \sum_\tau (y_t(\tau) - y_t^o(\tau))^2\right], \tag{78}$$

subject to

$$A(\tau)' = A(\tau-1)'\Phi + A_r(\tau-1)\phi_r'\Phi - \hat{A}(\tau-1)'\lambda_{\beta,\beta} - A_r(\tau-1)\lambda_{\beta,r} \tag{79}$$

$$A_r(\tau) = A_r(\tau-1)\rho_r - 1 - \hat{A}(\tau-1)'\lambda_{r,\beta} - A_r(\tau-1)\lambda_{r,r} \tag{80}$$

where $\hat{A}(\tau-1)$ is a function of $A(\tau-1)$ and $A_r(\tau-1)$ as defined in (39). The model implies that $y_t(\tau) = -p_t(\tau)/\tau$, where $p_t(\tau)$ satisfies the pricing equation in (26), with u_t being unobservable with mean 0, variance Σ_u, and uncorrelated with β_t and r_t. As a result, we can rewrite the objective function as

$$\min_{\{\lambda_{\beta,\beta},\lambda_{\beta,r},\lambda_{r,\beta},\lambda_{r,r},C\}} \sum_t \sum_\tau \left(\frac{A\beta_t + A_r r_t}{\tau} + \frac{C}{\tau} + y_t^o(\tau)\right)^2, \tag{81}$$

We note that this problem does not explicitly involve arbitrageur risk aversion γ, because that is embedded in the solution of risk premium $\lambda_{\beta,\beta}$, λ_β, r, λ_r, β, λ_r, r and the intercept C.

In the estimation, the dimension of β is $K = 6$, and the dimension of r_t is 1. The vector C is 120×1 (quarterly frequency of 30 years gives rise to 120 maturities). Therefore, the total degree of freedom is 7*7+120 = 169. This is a very high dimensional optimization problem. Similar to a typical affine term structure estimation, it is important to find a good initial point for the algorithm. We leverage on an important insight from the affine term structure literature, which is to use regressions to initialize the coefficient matrix.

In particular, we start with a linear regression problem:

$$\min_{A,A_r,C} \sum_t \sum_\tau (A\beta_t + A_r r_t + C + \tau y_t^o(\tau))^2,$$

Solving this estimation on A, A_r, and C is equivalent to regress the log-price vector

$$(y_t^o(1), 2y_t^o(2), \cdots, Ny_t^o(N))$$

on β_t and r_t, where C serves as the intercept term.

Next, knowing the values of the matrices A, A_r, we can view the iteration equations in (79) and (80) as another set of regressions. Rewriting (79) and (80) in a regression form,

$$\begin{aligned}\underbrace{A(\tau)' - A(\tau-1)'\Phi}_{\text{left hand side}} &= \underbrace{A_r(\tau-1)}_{\text{dep var}}(\phi_r'\Phi - \lambda_{\beta,r}) - \underbrace{\hat{A}(\tau-1)'}_{\text{dep var}}\lambda_{\beta,\beta}, \\ \underbrace{-A_r(\tau) + A_r(\tau-1)\rho_r - 1}_{\text{left hand side}} &= \underbrace{\hat{A}(\tau-1)'}_{\text{dep var}}\lambda_{r,\beta} + \underbrace{A_r(\tau-1)}_{\text{dep var}}\lambda_{r,r}.\end{aligned} \tag{82}$$

where the regression coefficients are $\phi_r'\Phi - \lambda_{\beta,r}$, $\lambda_{\beta,\beta}$, $\lambda_{r,\beta}$, and $\lambda_{r,r}$. Note that ϕ_r and Φ are directly estimated in the data. Consequently, we can use regressions to initialize all of the four price of risk matrices and also the constant term C.

Once we initialize the problem, we use built-in minimization solvers to find the solution to the optimization problem (81) to find better initial values.

Next, we note that for any given γ, and the solved matrix A, A_r, and $\hat{A}$, we can uniquely pin down the latent-demand impact matrix A_u. Equation (49) can be viewed as a linear matrix equation for A_u. To see that, we denote

$$\begin{aligned}\hat{A}^{\text{shift}} &= \big(0, \hat{A}(1), \dots, \hat{A}(N-1)\big)' \\ A_r^{\text{shift}} &= (0, A_r(1), \dots, A_r(N-1))'\end{aligned} \tag{83}$$

which are "shifts" of the original $\hat{A}$ and A_r matrices. Then we can stack all different τ in equation (49) to get the matrix equation

$$\begin{aligned}&\left(\hat{A}^{\text{shift}}\gamma\Sigma\sum_{\hat{\tau}=2}^{N}\hat{A}(\hat{\tau}-1)\alpha(\hat{\tau})' + A_r^{\text{shift}}\gamma\sigma_r^2\sum_{\hat{\tau}=2}^{N}A_r(\hat{\tau}-1)\alpha(\hat{\tau})' + I\right)A_u \\ =&\hat{A}^{\text{shift}}\gamma\Sigma(\hat{A}^{\text{shift}})' + A_r^{\text{shift}}\gamma\sigma_r^2(A_r^{\text{shift}})'\end{aligned} \tag{84}$$

which is simply a linear equation for A_u that can be solved immediately.

To initialize the arbitrageur risk-aversion γ and the intercepts ψ and ψ_r of arbitrageur's outside portfolios, we will solve the subset of the main objective function in (54). Given initializations of A, A_r, A_u, we solve for

$$\min_{\{\gamma,\psi,\psi_r\}} \mathbb{E}\left[(h-h^o)^2\right], \tag{85}$$

subject to (50). We find that arbitrageur holding of long-term Treasuries in the model is very sensitive to γ, so this optimization problem has very high power in identifying γ. Moreover, to speed up the algorithm and ensure a solution, we express the equation for C as a linear system rather than a iterative system. In particular, we rewrite (73) as

$$\begin{aligned}
&A(\tau-1)'(I-\Phi)\bar{\beta}+A_r(\tau-1)(\bar{r}-\phi_r'\Phi\bar{\beta})+\frac{1}{2}\hat{A}(\tau-1)'\Sigma\hat{A}(\tau-1)\\
&+\frac{1}{2}(A_r(\tau-1)\sigma_r)^2+\frac{1}{2}A_u(\tau-1)'\Sigma^u A_u(\tau-1)-\hat{A}(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\left(\bar{S}(\hat{\tau})-\theta_0(\hat{\tau})\right)+\psi\right)\\
&-A_r(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\sigma_r^2A_r(\hat{\tau}-1)\left(\bar{S}(\hat{\tau})-\theta_0(\hat{\tau})\right)+\psi_r\right)\\
=&\hat{A}(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\alpha(\hat{\tau})'\right)C+A_r(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\sigma_r^2A_r(\hat{\tau}-1)\alpha(\hat{\tau})'\right)C+C(\tau)-C(\tau-1)
\end{aligned} \tag{86}$$

The left-hand side is a single value and denote it as $C_0(\tau)$. Also denote the vector

$$\tilde{A}(\tau)'=\hat{A}(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\alpha(\hat{\tau})'\right)+A_r(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\sigma_r^2A_r(\hat{\tau}-1)\alpha(\hat{\tau})'\right)+(\mathbf{1}_\tau-\mathbf{1}_{\tau-1})'$$

where $\mathbf{1}_\tau$ is an N-dimensional vector that is one for element τ but zero otherwise. Then equation (86) can be simplified as

$$C_0(\tau)=\tilde{A}(\tau)'C$$

for all $\tau\in\{2,3,\cdots,N\}$. For $\tau=1$, we know that $p_t^{(1)}=-r_t$, which implies that $C(1)=0$. Stacking all of the equations for $\tau\in\{1,2,3,\cdots,N\}$, we get

$$\begin{pmatrix}0\\C_0(2)\\\vdots\\C_N(1)\end{pmatrix}=\begin{pmatrix}\mathbf{1}_1'\\\tilde{A}(2)'\\\vdots\\\tilde{A}(N)'\end{pmatrix}C,$$

which is a linear system that can be easily solved.

Finally, with $\lambda_{\beta,r}$, $\lambda_{\beta,\beta}$, $\lambda_{r,\beta}$, $\lambda_{r,r}$, γ, and A_u solved, we can recover the intermediary outside asset risk loadings Ψ, Ψ_r, Λ, Λ_r, from the definitions of $\lambda_{\beta,r}$, $\lambda_{\beta,\beta}$, $\lambda_{r,\beta}$,and $\lambda_{r,r}$ in equations (47) and (48). In particular, we have

$$\Psi=\frac{1}{\gamma}\lambda_{\beta,\beta}-\sum_{\hat{\tau}=2}^{N}\Sigma\hat{A}(\hat{\tau}-1)\left(\zeta(\hat{\tau})'+\alpha(\hat{\tau})'A+\theta(\hat{\tau})'\right)$$

$$\Psi_r = \frac{1}{\gamma}\lambda_{\beta,r} - \left(\sum_{\hat{\tau}=2}^{N} \sigma_r^2 A_r(\hat{\tau}-1) \left(\zeta(\hat{\tau})' + \alpha(\hat{\tau})' A + \theta(\hat{\tau})' \right) \right)$$

$$\Lambda = \frac{1}{\gamma}\lambda_{r,\beta} - \sum_{\hat{\tau}=2}^{N} \Sigma \hat{A}(\hat{\tau}-1) \left(\zeta_r(\tau)' + \alpha(\hat{\tau})' A_r \right)$$

$$\Lambda_r = \frac{1}{\gamma}\lambda_{r,r} - \sum_{\hat{\tau}=2}^{N} \sigma_r^2 A_r(\hat{\tau}-1) \left(\zeta_r(\tau)' + \alpha(\hat{\tau})' A_r \right)$$

C. Additional Quantitative Results

C.1. Calculating Aggregate Treasury Market Elasticity

For comparison with the literature, we will use Treasury market multiplier, which is the inverse of Treasury market elasticity. We calculate the aggregate Treasury market multiplier in following steps: First, we construct a "representative demand shock", accounting for the weights of different maturity buckets. Second, we calculate the total price response in each maturity bucket due to the demand shock. Finally, we use the weights of different maturity bucket to aggregate bucket-level price responses to an aggregate Treasury valuation change. The ratio between this valuation change and the total size of the demand shock is what we define as market multiplier, consistent with the literature.

Denote the average outstanding (in \$ billion amount scaled to the year of 2011) of three maturity buckets as $\bar{\omega}$, which is $\bar{\omega} = (3998, 6103, 5389)'$. Thus, the weight vector of three maturity buckets is

$$\omega = \frac{\bar{\omega}}{\sum_i \bar{\omega}_i} = (0.258, 0.394, 0.348)'.$$

Next, we construct a representative demand shock, by multiplying \$1 billion with the weight ω, so that the demand vector is simply ω. Denote the set of durations for three maturity buckets as $\mathscr{T}$. Then for the three maturity buckets, the log price response matrix is $\hat{A}_u \equiv A_u(\mathscr{T}, \mathscr{T})$. Thus, the log price response vector to the representative latent demand shock is $\hat{A}_u \omega$. Finally, we need to multiply the value of three maturity buckets to get a billion dollar valuation change $\bar{\omega}' \hat{A}_u \omega$, which is the aggregate Treasury market multiplier. Using the transpose of $\hat{A}_u$ matrix as shown in Panel A of Table 5, we obtain a multiplier of 0.27, i.e., a \$1 billion representative demand shock on the Treasury market increases total Treasury valuation by \$0.27 billion.

Following the same procedure, we can use Panel B of Table 5 to calculate the Treasury market multiplier in the case without arbitrageur, and we find that is 18, which is a very large number. Therefore, low risk-aversion arbitrageurs are crucial for an elastic Treasury market.

Next, we show the aggregate multiplier for permanent demand shocks. In Table 15, we show the price impact of permanent demand shocks in the case with and without arbitrageurs. This table has the same format as our main Table 5. Using Panel A of 15, we can calculate the Treasury market multiplier for permanent demand shock as 1.00, which is higher than the case of latent demand shock, because permanent demand shocks significantly change the risk premium.

Panel B of Table 15 is identical to Panel B of Table 5, because absent from arbitrageurs, latent demand shocks and permanent demand shocks are treated the same by preferred-habitat investors. Consequently, in the case without arbitrageurs, Treasury market multipliers to permanent demand shock and to latent demand shock are identical.

Table 15. **Impact of Permanent Demand Shocks on Treasury Prices with and without Arbitrageurs.**

We illustrate the impact of permanent demand shocks with and without arbitrageurs. A value of 1 indicates that \$100 billion extra permanent demand of Treasuries increases the price by 1%.

Panel A: With Arbitrageur

		Price change (%) of	
	short maturity	medium maturity	long maturity
shock on short maturity	0.0007	0.0051	0.0093
shock on medium maturity	0.0235	0.1741	0.3481
shock on long maturity	0.1284	1.0586	3.3735

Panel B: Without Arbitrageur

shock on short maturity	0.5921	1.8907	10.3854
shock on medium maturity	2.2033	10.7307	48.6644
shock on long maturity	0.4988	2.0057	13.0317

Panel C: Price Impact Ratio (Panel B/Panel A)

shock on short maturity	830.32	369.30	1112.04
shock on medium maturity	93.82	61.62	139.81
shock on long maturity	3.88	1.89	3.86

C.2. Dissecting Arbitrageur's Outside Portfolio

In Figure 14, we show how changes in arbitrageur's outside portfolio affect equilibrium Treasury pricing, in a similar way as Figure 7. We view this exercise as illustrating how other markets that arbitrageurs participate in can transmit to the Treasury market via arbitrageurs. In particular, we increase all the outside-portfolio exposures, Ψ, Ψ_r, Λ, Λ_r, ψ, ψ_r by 1%, and we analyze the percentage changes of various yields and sensitivities.

C.3. The Impact of Fed's Cross Elasticity

We evaluate the role of cross elasticity in the Fed's demand function. As shown by Table 4, the Fed has a strong cross-substitution especially on long-maturity Treasuries. Given the long-term Treasury yield, if short-term Treasury yield becomes higher, Table 4 indicates that the Fed will significantly reduce long-term Treasury holdings. This response can also be interpreted as a coherent monetary tightening, when the Fed both increases the short-term interest rate and shrinks its long-term Treasury holdings.

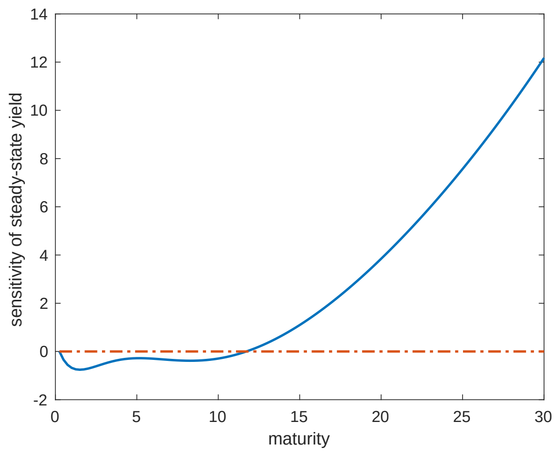
With the above intuitions, we will analyze what will happen to the Treasury market if the Fed does not tighten its long-term Treasury holding when short-term yield goes up, i.e., we set the cross-elasticity coefficients of Fed's demand to zero and estimate the model again following the same setup in equation (51). With this alternative model, we calculate how sector-level portfolio adjusts to a permanent demand shock similar to the exercise in Table 8, and we show the results in Table 16. We find that a permanent demand shock on short-term Treasuries leads to similar results as in Table 4, but results are significantly different in response to a permanent demand shock on long-term Treasuries, with the Fed increasing its long-term Treasuries holdings by about \$100 billion, at par with the primary sell shock size of \$100 billion. Due to a more accommodation Fed, arbitrageurs do not need to expand its balance sheet aggressively, so the expansion of arbitrageur's balance sheet becomes smaller.

C.4. Latent Demand and Risk Premium

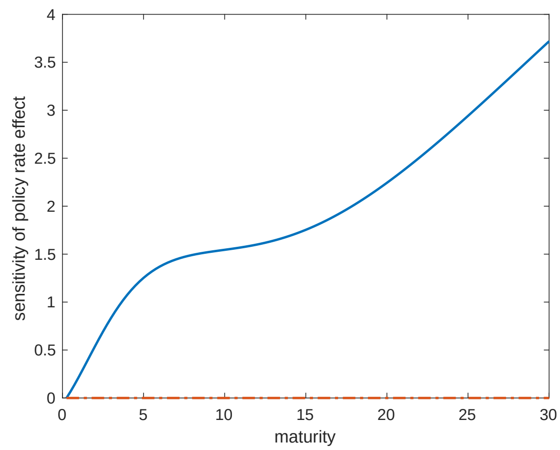
In our quantitative analysis, we assume that the risk premium by latent demand shocks u_t, as in (64), is negligible compared to that of macroeconomic factors β_t and monetary policy rate r_t. This assumption greatly simplifies model estimation. To evaluate whether this assumption is reasonable, we calculate the contribution of each components of the risk premium to expected

Figure 14. **Impact of Outside Portfolio Exposure.**

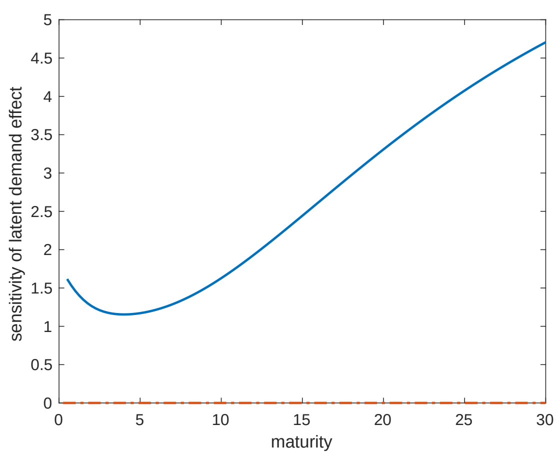
We illustrate how outside portfolio exposure affects the equilibrium. We increase arbitrageur outside portfolio risk exposures (Ψ, Ψ_r, Λ, Λ_r, ψ, ψ_r) all by 1%, and illustrate the percentage changes of various yields and sensitivities. Panel (a) shows $\frac{\partial \log(y^{\text{steady}})}{\partial \log(\gamma)}$, the sensitivity of steady state yield curve. Panel (b) illustrates $\frac{\partial \log(A_r)}{\partial \log(\gamma)}$, the price elasticity to interest rate r. Panel (c) shows $\frac{\partial}{\partial \log(\gamma)} \log(\frac{1}{N} A_u \cdot \mathbf{1})$, the price elasticity to latent demand averaged across latent demand shocks of different maturities. Panel (d) shows $\frac{\partial}{\partial \log(\gamma)} \log(\frac{1}{N} \frac{\partial p}{\partial \theta_0} \cdot \mathbf{1})$, the price elasticity to permanent demand averaged across permanent demand shocks of different maturities.



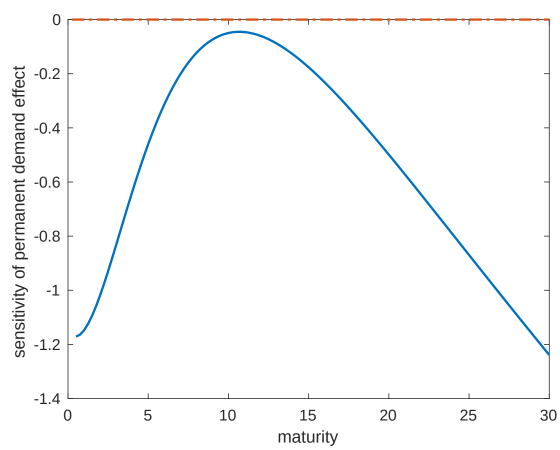
(a) Sensitivity of steady-state yield



(b) Sensitivity of policy rate effect



(c) Sensitivity of latent demand effect



(d) Sensitivity of perm. demand effect

Table 16. **Model-Implied Sector-Level Portfolio Adjustment to Permanent Demand Shocks without Fed Cross Elasticity.**

This table shows the same counterfactual results as Table 8, but shutting off the cross elasticity of Fed's demand.

Panel A: A \$100 billion Sell Shock of $\tau < 1$ Treasuries

Sector	$\tau < 1$	$1 \leq \tau < 5$	$\tau \geq 5$	Total Change
Banks	-0.5	0.0	0.3	-0.2
ICPF	0.0	0.0	0.1	0.1
MF ROW	0.0	0.0	0.1	0.1
MF U.S.	-1.1	0.1	0.7	-0.3
MMF	-3.0	0.0	0.0	-3.0
Other U.S.	3.0	0.9	0.0	3.9
Foreign	-1.9	-0.8	-0.4	-3.1
Fed	0.0	0.3	5.3	5.7
Arbitrageurs	103.5	-0.6	-6.0	96.9

Panel B: A \$100 billion Sell Shock of $\tau \geq 15$ Treasuries

Sector	$\tau < 1$	$1 \leq \tau < 5$	$\tau \geq 5$	Total Change
Banks	-6.5	-3.0	7.6	-1.9
ICPF	-0.2	0.2	1.2	1.2
MF ROW	-0.6	0.0	1.4	0.8
MF U.S.	-16.1	-7.3	19.3	-4.1
MMF	-41.9	0.0	0.0	-41.9
Other U.S.	40.6	27.5	-18.0	50.2
Foreign	-25.6	-19.3	4.8	-40.1
Fed	0.3	2.3	95.0	97.6
Arbitrageurs	50.1	-0.4	-11.4	38.3

returns as follows. Define

$$\mu_{\beta,t}(\tau) = \hat{A}(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\Sigma\hat{A}(\hat{\tau}-1)X_t^{(\hat{\tau})}\right)+\Sigma^{1/2}\tilde{\sigma}\tilde{X}_t\right)$$
$$\mu_{r,t}(\tau) = A_r(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\left(\sigma_r^2 A_r(\hat{\tau}-1)X_t^{(\hat{\tau})}\right)+\sigma_r\tilde{\sigma}\tilde{X}_t\right)$$
$$\mu_{u,t}(\tau) = A_u(\tau-1)'\gamma\left(\sum_{\hat{\tau}=2}^{N}\Sigma^u A_u(\hat{\tau}-1)X_t^{(\hat{\tau})}\right)$$

Then the expected excess return of Treasuries of maturity τ is

$$\mu_t^{(\tau)} - r_t = \mu_{\beta,t}(\tau)+\mu_{\beta,t}(\tau)+\mu_{u,t}(\tau)$$

Next, we calculate the contribution to the average risk premium by u_t as

$$\mathscr{F}(\tau) = \left|\frac{\frac{1}{T}\sum_{t=1}^{T}\mu_{u,t}(\tau)}{\frac{1}{T}\sum_{t=1}^{T}\left(\mu_{u,t}(\tau)+\mu_{\beta,t}(\tau)+\mu_{r,t}(\tau)\right)}\right|$$

For the three maturity buckets, this ratio is 0.47%, 0.89%, and 1.34%. As a result, we conclude that the contribution of the latent demand shock to risk premium on average is small, confirming our simplification assumption.