QT vs QE: Who is In When the Central Bank is Out?*

Iryna Kaminska
Bank of England

Alex Kontoghiorghes
Bank of England

Walker Ray
Federal Reserve Bank of Chicago and CEPR

June 30, 2026

Abstract

We analyze the role of preferred habitat demand in the transmission of quantitative easing (QE) and tightening (QT) programs. Combining granular bidding data from Bank of England QT and QE auctions with secondary market bond level transaction data, we find: (i) during QE, higher sales from pension funds and insurance companies led to higher dealer bid prices; (ii) in contrast, this pattern disappears during QT auctions. We rationalize this asymmetry in a preferred-habitat model of the bond market with supply-dependent demand: as bond supply increases, habitat demand becomes less elastic. Our results are vital for policymakers: as outstanding supply of government debt increases, the bond price effects of QT may become amplified.

Keywords: quantitative easing, quantitative tightening, central bank auctions, monetary policy, monetary transmission mechanism, preferred habitat, gilt market.

*The authors thank Lena Boneva, Laura Coroneo, Jean-Sébastien Fontaine, Ruggero Jappelli (discussant), François Koulischer (discussant), Wolfgang Lemke, Angelo Ranaldo, Matt Roberts-Sklar, Raman Uppal, Tarik Umar, Lia Vaz Cruz, the UK Debt Management Office (DMO), as well as participants at the 12th Asset Pricing Workshop, EFA 2025, AEA 2025, 2024 Financial Intermediation and Monetary Policy Conference at the Banque de France, EEA Congress 2023, and at the Bank of England, Swiss National Bank, University of Cyprus, and University of Liverpool seminars for their helpful comments and suggestions. The views expressed in this paper are solely those of the authors and should not be taken to represent the views of the Bank of England (or any of its committees), the Federal Reserve Bank of Chicago, or the Federal Reserve System.

1 Introduction

In 2021, after more than a decade of unconventional monetary policy easing, major central banks around the world returned to the "old-fashioned" ways of policy making. Besides increasing nominal rates above zero, central banks also began reducing the size of their balance sheets: after multiple rounds of asset purchases through a decade of "quantitative easing" (QE) rounds, central banks began to embark on the uncharted territory with "quantitative tightening" (QT) policies. Even after several years of QT implementation, we still do not have a clear understanding of how this fairly new monetary policy tool works, how it affects both financial markets and the broader macroeconomy, or how its workings compare to the QE transmission channels.

Our paper aims to shed light on an important transmission channel of QE and QT policies. By buying or selling long-duration securities, central banks can impact financial conditions through changing the composition of assets held in the market. But the precise way in which these large-scale asset operations affect asset prices depend on exactly *who* buys or sells these assets, and who *in equilibrium* are the marginal investors for pricing risk. To the extent the composition or behavior of these investors differ during QE vs. QT programs, then QT will have important state-dependent effects; in other words, QT might not simply be the "reverse" of QE.

In particular, the channels of unconventional policy will depend in part on the preferences of investors who have particularly specialized demand for bonds of specific maturities. Insurance companies and pension funds (ICPFs) are a classic example of such "preferred habitat" (PH) investors, as they have a special demand for longer maturity bonds. The theoretical importance of these investors for monetary policy has been shown in the seminal work of Vayanos and Vila (2009, 2021); however, the explicit role of preferred habitat demand for the transmission of QE and QT programs has been hard to pin down empirically, as the demand function of these investors is extremely hard to measure.

In this paper, we propose to overcome this obstacle by using the unique experience of the Bank of England (BoE). The active approach to QT chosen by the BoE is different to the passive approaches used by most other central banks (such as the Federal Reserve or the ECB, who thus far have simply allowed their holdings to unwind as assets mature). The BoE instead has been actively selling assets into the secondary markets via auctions since 2022.

Our approach combines and analyzes granular BoE auctions data jointly with detailed transaction level data around central bank auctions, focusing on potentially distinct patterns of trade between dealers and ICPFs during QE and QT monetary policy periods. The BoE auctions data provide rich insights into the bidding behavior of auction participants, allowing us to study the behavior of dealers during auctions and investigate whether their bidding has been affected by the identity and type of the dealers' trading counter-parties during these auctions. This setting gives us the unique opportunity to utilize central bank proprietary auctions data in order to study the role of dealer counter-party demand in QE/QT transmission during specific programs.

Our findings during the QE phase are consistent with the standard preferred habitat view. We show that ICPFs played a crucial role in the QE transmission process: ICPF trading volume is significantly higher during QE auction days; moreover, higher sales to dealers by ICPFs lead these dealers to place higher bids at auction. That is, consistent with habitat demand theory, when QE purchases draw more from habitat investors' holdings, the decline in yields is mitigated. These findings provide direct evidence for the importance of habitat investor demand coming from ICPFs during QE implementation.

Surprisingly, however, during the later QT phase, we do not detect any special role of ICPF investors in policy transmission. We find that the bidding behavior of dealers with greater interactions with ICPFs around QT auctions are not meaningfully different from other dealers' bids. This asymmetry between QE and QT is inconsistent with the predictions of standard habitat theory.

We explore a number of possible candidates that could account for the observed asymmetry. Our baseline estimates on the asymmetric effects of the habitat demand channel during QT vs. QE remain robust when we account for the potentially asymmetric roles of interest rate volatility and dealer intermediation capacity during the two periods. Instead, we find that habitat demand is particularly sensitive to the total outstanding supply of government bonds. Consistent with the idea that ICPF demand becomes "satiated" as supply increases, we find that ICPF demand transmission to dealer bids is weak precisely when outstanding gilt supply is high.

We find further support for our results by studying long maturity UK government gilt auctions by the Debt Management Office (DMO) during QT and pre-COVID pandemic periods (i.e. before gilt issuance has increased to record amounts). Theoretically, habitat demand transmission channels are similar during either Treasury auctions or central bank asset programs; hence DMO auctions provide a useful way to validate our findings studying QE and QT auctions directly (also see Ray et al. (2024)). Consistent with our baseline results, we find similar asymmetries in DMO auctions: compared to the pre-

2020 periods, ICPF demand transmission during QT periods is significantly weaker.

Our empirical findings present a challenge to the current views established in the theoretical literature. Standard portfolio rebalance models do not allow for state dependencies, and therefore QE and QT are predicted to be mirror images of one another. Based on our empirical results regarding the asymmetric behavior of habitat investors in particular, we propose an extension of Vayanos and Vila (2021) to rationalize our findings. A key assumption of standard preferred habitat models is constant demand elasticity of habitat demand; this assumption is critical for determining the implied symmetric nature of QT and QE impacts. Thus, we relax this assumption and allow for a more generalized demand elasticity model. In particular, we allow the elasticity of habitat investor demand to depend on the outstanding bond supply available to investors. As a result, we show that the decreased role of the preferred habitat demand channel during QT can be explained by the increased government bond issuance since the COVID-19 pandemic.

Our empirical and theoretical findings thus reaffirm the importance of the state-dependent nature of the transmission of unconventional monetary policy. The equilibrium effects of unconventional policy depend critically on *who* is buying what the central bank is selling. When the shape of demand depends on total outstanding bond supply, the asymmetries between the different phases of unconventional monetary policy is rationalized.

Our findings suggest a note of caution for central banks pursuing balance sheet normalization through quantitative tightening, when preferred habitat investors play a key role in the market. Policymakers have stated that QT will be uneventful: like "watching paint dry." However, the increased supply of bonds (both from fiscal policy as well as from QT itself) may lead to a deterioration in habitat demand. When this occurs, our model shows that the price impacts of QT can become asymmetrically larger, as dealers are forced to absorb a larger fraction of the issuance from QT. Thus, both our empirical and theoretical results are vital to policymakers as they indicate that monetary transmission mechanisms of QE and QT may differ.

Related literature. An increasingly large body of the monetary and finance literature is dedicated to studying the effects of official asset purchases or sales. Our work also belongs to this literature. However, there are some important distinguishing features that set us apart from the existing studies. In particular, our paper makes three main contributions to the literature.

First, our empirical focus is to explicitly evaluate the role of the preferred habitat demand in the QE and QT transmission mechanism through the use of the granular data from and around QE/QT auctions. The "preferred habitat" hypothesis of interest rates was developed by Modigliani and Sutch (1966); it suggests that some market participants have a "preference" for a particular maturity "habitat" in order to match the term structure of their assets and liabilities. For example, ICPFs are often thought to have a preference for bonds of specific longer maturities, because they tend to hold long-term assets in order to match their liabilities.

Several previous studies analyzed this type of preferred habitat investors. Giese et al. (2024) showed that the UK government bond market is characterized by the significant presence of PH investors (including ICPFs), whose gilt holdings are less sensitive to changes in the price of gilts, in line with the idea that such investors value these assets for non-pecuniary reasons. They found that PH behavior in the gilt market exists across the term structure and that foreign central banks (at shorter maturities) and ICPFs (at longer maturities) make up some of the PH investor groups. They also showed that some of these investors—including insurance companies—reduced their gilt holdings more than proportionately during the 2016 QE4 program, in line with these investors playing a key role in the transmission of QE. Similarly, Joyce et al. (2017) showed that ICPFs were selling gilts to the Bank of England as part of the QE1 and QE2 programs. However, although these studies find results consistent with preferred habitat demand for gilts, they are unable to say if this type of behavior was explicitly manifested during the QE auctions or whether it had significant impact on QE transmission to the gilt market as a result.[1]

Ours is the first study to address this question by using granular QE auctions data. The Bank of England auctions data provide rich insights into the bidding behavior of auction participants and hence are most suited for the analysis and evaluation of the portfolio balance channel of QE.[2] In this spirit, our paper is close to Ray et al. (2024),

[1]Some studies tried to shed light on the role of portfolio balance by analyzing how the impact of QE announcements varies across individual gilts, based on the presumption that if QE lowers gilt yields by shrinking the supply of gilts, then those gilts more likely to be purchased by the central bank should see larger drops in yields after announcements. For example, using this approach, McLaren et al. (2014) find that this was the case for the first two QE programs. From their estimates, they calculate that this 'local supply' effect explains around half of the total impact of the QE1 and QE2 announcements on medium and long-term gilt yields.

[2]The Bank of England QE auctions data were used previously to study other questions, such as QE liquidity impacts. For example, Boneva et al. (2022) use the granular offer-level data from the Corporate Bond Purchase Scheme auctions to construct proxy measures for the Bank of England's demand for bonds and auction participants' supply of bonds in order to control for any reverse causality from liquidity to purchases.

who exploit the structure of the primary market for U.S. Treasuries to proxy for QE transmission and isolate and evaluate demand shocks that are transmitted solely through PH demand channels. The advantage of our paper is the direct use of granular QE auctions data. This enables us to explicitly estimate the effects of preferred habitat demand during QE transmission.

Our second main contribution is the comparative empirical analysis of QE and QT transmissions. Various studies have investigated the impact of QE. Generally, the findings document meaningful effects of QE on financial markets (see, e.g., among many others, the evidence on the yield impact of QE announcements by Krishnamurthy and Vissing-Jorgensen (2012), Christensen and Rudebusch (2012), Swanson (2021)). In contrast, evidence on the impact of balance sheet unwinding is scarce, mainly because there have to date only been a few attempts where central banks have actively unwound their balance sheets. A couple of existing papers, such as Du et al. (2024) and D'Amico and Seida (2024), focus on the general impact of QT announcements by the Federal Reserve and other central banks on financial markets.[3] Joyce and Lengyel (2024) take a different approach and infer the effect of the Bank's initial QT program from an analysis of the yield curve impact of individual debt security announcements by the UK's debt management office. They find local supply transmission channels to be important at the short and the long end of the yield curve, which they argue provides indirect evidence for market segmentation and bond demand coming from preferred habitat investors.

Instead, given that the Bank of England is the first among the central banks to actively reduce its balance sheet in the form of asset sales back to the secondary market, we can focus exclusively on QT sales, zooming in on particular QT channels. This approach is cleaner than disentangling the announcement effects of QT from all other policy effects and information signals communicated at the same time. Moreover, not only do we provide the first study of direct effects from active QT sales on gilt markets, we also examine the role of the state dependency in the asymmetric effects of QE and QT. And while the notion of QE and QT state-contingency is not new (for instance, see Bailey et al. (2020), Haldane et al. (2016), Ray et al. (2024), or Kekre et al. (2024), on the importance of different states of financial market conditions, or Gertler and Karadi (2013), and Cantore and Meichtry (2024) on the role of zero lower bound for policy

[3]While here we focus on the asset side of the central bank balance sheet and on the impacts of asset purchases and sales on financial markets and real activity, there is a recent complementary study of QE/QT by Kumhof and Salgado-Moreno (2024), who develop a theoretical framework to focus on the liability side, i.e. on the effect of reserve issuance and policy rules for reserves on macroeconomic and financial stability.

rates), the nature of the state dependence uncovered in this paper is different to the alternatives already available in the literature.

The third contribution of this paper is theoretical. While recently there have been several notions in the literature advocating the important role of the limited asset supply for QE transmission, the theory of PH demand for safe bonds has been guided so far by the seminal work of Vayanos and Vila (2021). While the transmission mechanisms in their model is quite rich, the habitat demand function is relatively simple. Crucially, they assume a constant demand elasticity, which is not consistent with the empirical results we find. In this paper we propose to build on and extend the demand model of Vayanos and Vila (2021) (see also Jansen et al. (2024b) for another extension which enriches the demand structure of preferred habitat models). Our model is based on a more generic demand function, where the elasticity can depend on the quantity of the assets available to investors. This approach delivers non-linear and asymmetric market responses to asset purchase programs, consistent with our empirical findings.

The remainder of the paper is organized as follows. The institutional background on Bank of England QE and QT and the data are described in Section 2 and Section 3, respectively. We subsequently present our testable hypotheses, estimation methodology, and main empirical findings in Section 4. The theoretical model with time-varying elasticity of demand is presented in Section 5. Section 6 concludes.

2 Institutional Background: the Bank of England

The Bank of England traditionally relies on the policy rate to steer monetary conditions. However, when policy rates hit the effective lower bound in the aftermath of the Global Financial Crisis, the Bank of England switched to the unconventional monetary policy of gilt purchases. QE began in March 2009 and continued through late 2021, after which the Bank shifted to quantitative tightening in 2022, gradually selling the assets accumulated during the QE.

2.1 QE and QT Programs

The asset purchases were financed by the creation of central bank reserves and formally carried out by the Bank of England's Asset Purchase Facility Fund Limited (APF) which is a subsidiary of the BoE (for more details on the structure of the APF, see Busetto et al. (2022)). In about four years, from March 2009 to November 2012, the BoE removed

£375 billion of gilts from the market (QE1-QE3); in August 2016, as part of a package of monetary policy measures, the BoE announced an extra £60 billion of government bond purchases (QE4); additional government bond purchases were announced in March 2020 in response to the COVID-19 outbreak (QE5), eventually bringing the total BoE QE holdings up to over 40 percent of nominal U.K. GDP once completed. The stock of gilts held in the APF peaked at £875 billion towards the end of 2021. Between different rounds of QE, the stock of QE purchases was maintained through the reinvestment of principal from maturing gilts.

As monetary policy transitioned into a tightening phase, the Monetary Policy Committee (MPC) announced principles of QT in August 2021. Because the primary active tool for monetary policy tightening is the conventional bank rate rather than the balance sheet, QT sales were meant to be conducted in a relatively gradual and predictable manner in order to not disrupt the functioning of financial markets. The BoE put the principles into practice in February 2022, initially via passive QT (i.e. through ending the reinvestment of maturing bonds) and eventually announcing active gilt sales in September 2022, when the MPC voted to reduce the stock of gilts held in the APF by £80 billion in the first year. Subsequently, the MPC reviewed QT sales annually, voting to reduce the stock of gilts held in the APF for monetary policy purposes by £100 billion in 2023, and then again in 2024. After around two years and a half of implementation, QT reduced the size of the total APF by approximately 30 percent.

It is important to acknowledge that although the reduction in the APF is substantial, the QT policy has been implemented in the environment of the much larger universe of the outstanding gilts due to the post-COVID increase in the gilt issuance. Figure 1 displays the nominal value of conventional gilts, which as of May 2025 account for over 75% of debt issued by the DMO (£2.07tr).[4] Since the start of active QT, the nominal free-float of conventional gilts has increased by over 85% (£690bn to £1280bn).

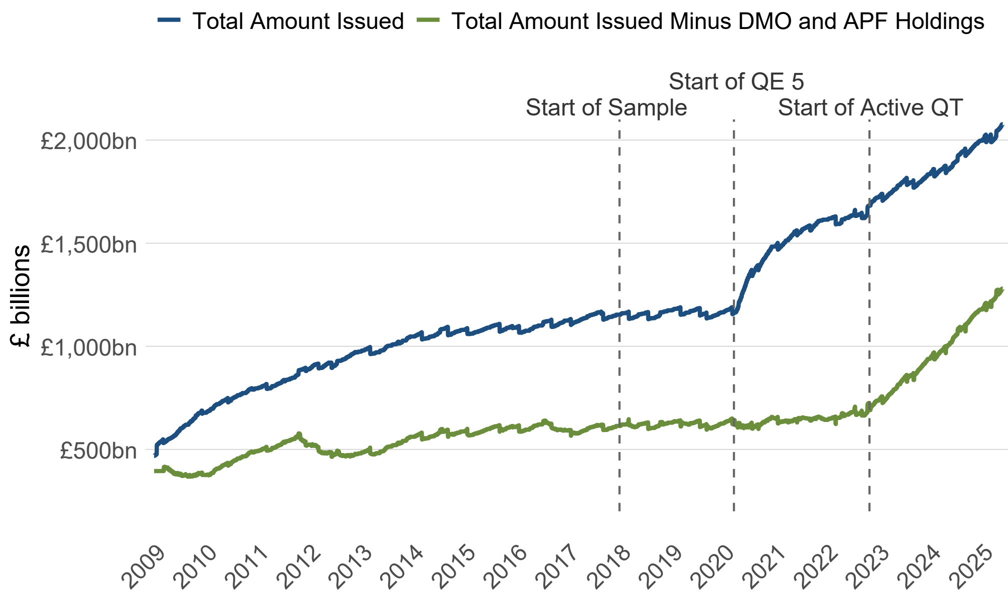
2.2 QE and QT Auctions

The Bank of England implemented QE through a series of 'reverse auctions' with a multi-object, multi-unit, discriminatory price auction format. In this subsection we outline how the auctions were conducted.

The Bank of England provided detailed schedules, operational notices, and eligibility documentation for all auction-related facilities on its website. To take part in gilt

[4] Inflation index-linked gilts account for the rest (£670bn).

Figure 1: UK nominal gilts in circulation



Note: This figure plots the total UK nominal gilt issuance, and total issuance minus DMO and APF holdings, in £billions since 2009. Sources: UK DMO, BoE.

auctions, market participants must be a Gilt-Edged Market Maker (GEMM), as listed on the website of the U.K.'s Debt Management Office (DMO).[5]

For each auction, participants were allowed to submit an unlimited number of offers containing a price, quantity, and the bond identifier (ISIN). Information about the offers is private. The BoE ran separate auctions for different maturity buckets and these auctions typically took place on different days of the week. Since the QE2 program, the BoE normally ran three auctions per week, for 3-7 (short), 7-15 (medium), and 15-50 (long) year buckets. In the week ahead of each QE auction, the BoE published a list of eligible bonds. On auction days, dealers were expected to submit their bids to sell eligible gilts to the BoE between 14:15 to 14:45 (London time). In March 2020, when the COVID pandemic shook the UK's economy and the markets, the maturity buckets and schedule were redefined so that the Bank could undertake three auctions per day, purchasing gilts with a residual maturity of 3-7 years at 12:15; 7-20 years at 13:15; and over 20 years at 14:15. The total amount purchased was aimed to be allocated evenly between

[5]For more information, see Market Operations Guide and Market Participants.

the maturity sectors.[6] In general, the purchases conducted by the APF excluded gilts with residual maturity less than three years, gilts of which the BoE already holds more than 70 percent of free float, index-linked gilts, and gilts with an issue size of £4 billion or less. Gilts that were newly issued by the DMO in the week preceding the auction or those re-opened by the DMO either one week before or after the auction were also excluded. After the auction close, the received offers were ranked according to the spread between the offered yield and the secondary market yield prevailing at the auction close. The BoE accepted the most attractive bids until the announced volume was reached. Successful auction participants received their offer price, meaning that all purchases were undertaken on a discriminatory price basis. After the close of each auction, the BoE published aggregate auction results including quantities offered and allocated. Individual offers remain private to bidders and the BoE. Before and during each auction, the BoE closely monitored the developments in the secondary market and reserved the right to exclude a gilt from the auction should unusual price developments occur.

The QT auctions have been conducted in a symmetric way to the QE format. The BoE set a schedule of auctions aiming to reduce the APF as evenly as possible across maturity sectors. Similarly to QE auctions, the QT sectors are defined as gilts with a residual maturity between: 3-7 years (short), 7-20 years (medium) and over 20 years (long). In line with the MPC's approach and given the principle that QT should be gradual, QT auctions have been less frequent than QE; typically, there has been one auction per week (Monday), focusing on one maturity sector only.[7]

2.3 Gilt Market Architecture

Gilt-edged Market Makers (GEMMs) are the UK government bond market's designated primary dealers, appointed by the UK DMO to support both primary issuance and secondary market liquidity. GEMMs are required to participate actively in gilt auctions, distribute newly issued gilts to end-investors, and provide continuous two-way prices in the secondary market, thereby acting as the core intermediaries between the DMO and investors and ensuring depth and stability in the gilt market. In return, they receive privileged access to DMO operations, including participation in auctions, syndications,

[6] See also the most recent market notice for details on the design of the QE and QT auctions, available at: Asset Purchase Facility: Gilt Purchases - Market Notice 5 August 2021 | Bank of England and Asset Purchase Facility: Gilt Sales – Market Notice 21 June 2024 | Bank of England.

[7] For more details on the BoE's approach to gilt sales see Asset Purchase Facility: Gilt Sales – Market Notice 1 September 2023 | Bank of England.

mini-tenders, and certain official market-management facilities.

In the primary market, GEMMs play a central role in ensuring auction coverage and price discovery. They are expected to bid in auctions, warehouse gilts on their balance sheets, and subsequently place them with investors domestically and internationally. Their feedback through formal consultations also informs the DMO's decisions on issuance size, maturity composition, and instrument design.

In the secondary market, GEMMs provide liquidity by making markets in conventional and index-linked gilts, facilitating trading across maturities and supporting the functioning of related markets such as gilt repo and STRIPS. Inter-dealer brokers allow GEMMs to redistribute risk anonymously among themselves, while end-investors typically access liquidity through GEMMs' dealer-to-client platforms. This intermediation function is central to gilt market resilience, particularly during periods of heavy issuance or heightened volatility.

GEMMs eligibility for APF auctions reflects their active engagement in gilt trading and their capacity to facilitate monetary policy operations. As key intermediaries between the Bank of England and the wider financial system, GEMMs help transmit policy signals and manage the flow of government securities.

To take part in APF auctions, GEMMs must be signed up to the gilt purchase or sale open market operations within the Sterling Market Facilities. Participation is limited to one firm per group unless distinct activities—such as banking and asset management—are undertaken. The number of participating GEMMs was relatively stable over time, with around 16 active during QE and QT auctions on average (more details on the auctions statistics are in Section 3.1). As dealers, GEMMs aggregate client flow, set offer prices and sizes across the BoE's QE and QT auctions, manage balance-sheet and risk exposures around the auctions. In doing so, GEMMs temporarily warehouse gilts ahead of QE auctions and after QT auctions, and subsequently redistribute inventory to the market. Consistent with this intermediation role, in the earlier study of QE auctions, Boneva et al. (2025) show that dealer participation in APF auctions can be sensitive to balance-sheet constraints and prior risk-taking in the gilt market.

3 Data

The UK bond transactions data used for the analysis are reported under the Markets in Financial Instruments Regulation (EU), as set out in Article 26 of UK MiFIR, and made available to the BoE by the Financial Conduct Authority (FCA). The transactions dataset begins in January 2018 (when the MiFID II framework was implemented) and is the successor to the Zen data set, which is maintained by the FCA. The data set contains transaction level data for every bond trade including price, quantity, buyer identity, and seller identity. Given the transaction data are available from 2018, we have restricted our sample period to be from January 2018 to April 2025.

We combine this data with QE and QT auctions data obtained from the Bank of England Markets desk, which carries out bond auctions. This data set contains all individual bid prices and quantities for each dealer for each bond at each auction, as well as the bid's spread to the benchmark yield (established by the BoE at each auction). These data, in combination with the secondary market transaction data, give us unique insights into the role of PH demand in the implementation of QE and QT.

Matching the two data sets allows us to link the amount of offers made by GEMMs at the BoE auctions to the trade flows between GEMMs and other investors around the auctions. In particular, we analyze the data by splitting transactions by different types of investors, such as dealers, hedge funds, ICPFs, and others. Finally, we also collect data on DMO auctions from the website of the DMO, which includes auction dates, auctioned bonds, as well as the issued amounts.

3.1 Gilt Auction Data

Our sample covers 538 gilt auctions run by the BoE between January 2018 and April 2025, of which there are 447 QE auctions (including 75 reinvestment auctions before the new round of QE launched in March 2020) and 91 QT auctions.[8]

Table 1 provides summary statistics on all auctions in our sample, split into QE and QT phases. Focusing on the auction participation, we observe that the number of participating GEMMs was relatively stable over the sample, with 17 and 16 active during QE and QT auctions respectively. Each of them was successful in their bids at least once. The average number of eligible gilts per auction is 8—roughly the same for QE5 and QT auctions—although the number of auctioned bonds was smaller on average for the reinvestment auctions. On average, the maturity of auctioned gilts was slightly lower during the QT sample than during QE sample (14.62 years vs. 17.22 years). The

[8] Since our focus is to analyze purchases for monetary policy implementation purposes, we excluded from the sample the BoE asset purchase auctions aimed to restore markets functioning during the LDI crisis in 2022. As Bandera and Stevens (2026) note, the 2022 intervention aimed to be temporary so that the monetary policy spillovers were as small as possible.

Table 1: Summary statistics of the auctions data

	QE Sample			QT Sample
All Maturity Auctions	All QE	QE Rein.	QE5	All QT
Number of Auctions	447	75	372	91
Participating Dealers	17	17	15	16
Total Auctioned Bonds	54	36	50	44
Avg. Auctioned Bonds	8	5	8	9
Bond Maturity (Yrs)	17.22	15.98	17.35	14.62
Bid Amount (£m)	58.26	75.39	56.48	-47.89
Allocated Nominal (£m)	18.14	22.67	17.67	-28.08
Allocated Proceeds (£m)	20.66	26.16	20.09	-20.81
Cover Ratio	3.42	6.46	2.81	2.28
Number of Bids	61.91	34.72	67.39	34.53
Spread to Benchmark (%)	-0.02	-0.03	-0.02	0.01
Long Maturity Auctions	All QE	QE Rein.	QE5	All QT
Number of Auctions	143	19	124	25
Participating Dealers	16	16	15	15
Total Auctioned Bonds	24	18	22	19
Avg. Auctioned Bonds	12	7	13	10
Bond Maturity (Yrs)	31.21	34.25	30.98	31.34
Bid Amount (£m)	40.42	62.66	38.68	-33.09
Allocated Nominal (£m)	13.16	12.99	13.18	-39.30
Allocated Proceeds (£m)	16.96	15.68	17.06	-19.32
Cover Ratio	3.57	12.47	2.28	1.71
Number of Bids	74.21	40.63	79.35	34.96
Spread to Benchmark (%)	-0.02	-0.05	-0.02	0.01

Note: This table displays the summary statistics for the main variables used in the regression analysis for the QE sample (2018–2021) and QT sample (2022–2025). PH refers to Preferred Habitat investors, HF refers to Hedge Funds, and QE Rein. refers to QE reinvestments. Source: BoE.

bid-to-cover ratio is on average reasonably stable over different subsamples, although it tends to be lower for QT auctions.

The lower section of Table 1 shows the same statistics focusing on long maturity bucket auctions only. Most of the participating dealers have been active in this maturity segment (the number of participating dealers in long maturity QT auctions and all maturity QT auctions is the same). In our sample, which consists of 143 QE auctions for long maturity gilts (out of 447 QE auctions) and 25 long maturity QT auctions (out of 91), the average auction size (*Allocated Proceeds*) has been typically smaller for the long

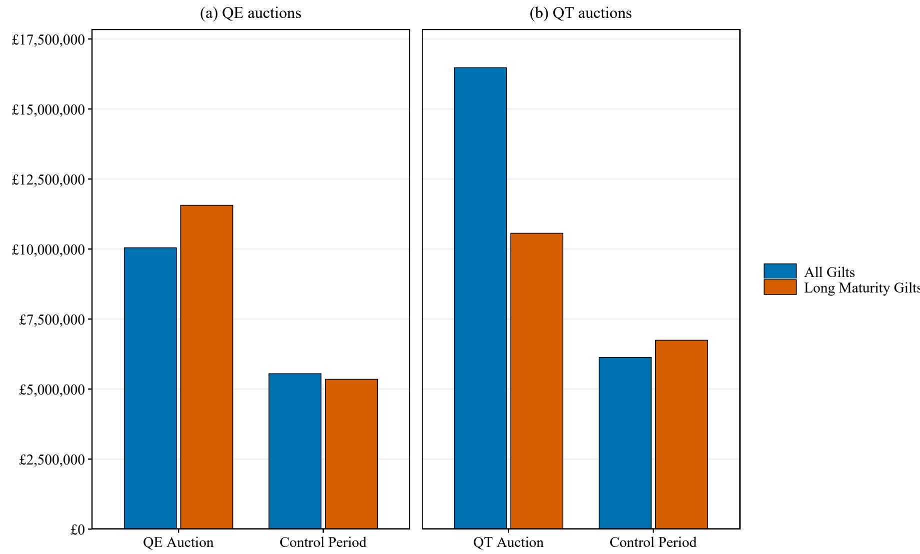


Figure 2: ICPF trades with GEMMs around QE and QT auctions

Note: This figure plots the average purchase value of ICPF investors for QE and QT auctioned bonds on auction days, alongside the corresponding values during the control period when no APF auctions were held. Average values are provided for all gilts auctioned and longer maturity gilts (over 20 years time to maturity) separately.

maturity auctions. The average number of eligible gilts per longer maturity auctions is also typically higher, ranging from 7 for the reinvestment auctions to 13 during QE5, comparing to 10 during QT. Correspondingly, we observe that the bid frequencies and sizes seem to differ across the QE and QT auctions. In the long maturity segment auctions, there was substantially higher average number of bids during QE rather than QT, and the average bid was also larger during QE.

3.2 Secondary Market Activity: Preliminary Analysis

Before going into the details of our approach, here we share some stylized facts about the ICPFs and dealer activity around the auctions.

Initial data analysis indicates that ICPFs were actively trading with dealers in the secondary market around the times of QE auctions (both during reinvestment and active purchase phases) as well as around QT auctions. The trading volumes by ICPF investors are displayed in Figure 2, which compares the trades on auction days to non-auction days.[9] The chart indicates the QE auctions were associated with higher trading activity

[9]As a control period for both the QE and QT comparison, we use transactions which took place

from preferred habitat investors. Indeed, the average transaction value of gilts sold by ICPF investors to primary dealers on QE auction days is twice their normal amount when compared to the control period (i.e. months without auctions). A similar increase in transactions is observed during QT, suggesting that part of the market trading can be related to central bank auctions.

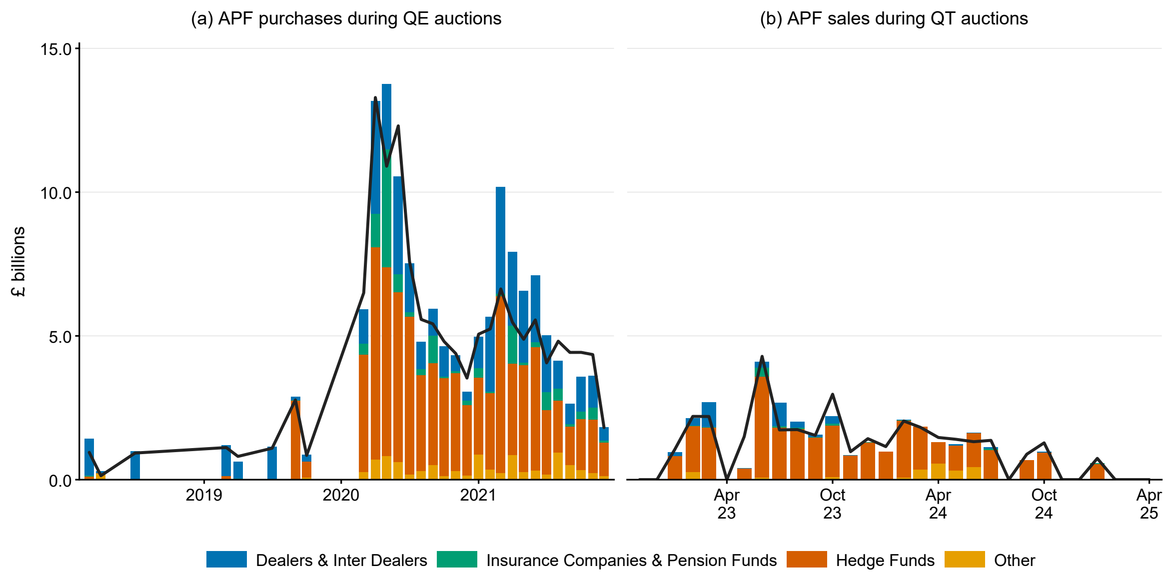


Figure 3: APF purchases during QE and APF sales during QT for long-maturity gilts

Note: Panels show (a) APF purchases during QE auctions and (b) APF sales during QT auctions for long-maturity (over 20 years) bonds. Stacked bars decompose the volumes bought from dealers by investor type, aggregated monthly. Solid line is the amount bought (sold) by the APF in the corresponding month.

As can be seen from Figure 3, the secondary market trades on auction days and the dealers' trades at the APF auctions are closely matching, providing further evidence of the link between them. These patterns are also consistent with the dealers' intermediary role; indeed, rather than being buy-and-hold investors, GEMMs tend to buy gilts from other investors in order to sell at QE auctions and tend to sell the purchased gilts after the QT auctions.

Figure 3 also shows clear evidence of some asymmetries regarding how APF purchases and sales interact with ICPF trade activity during QE and QT. Although ICPFs are active during both unconventional policy phases, the relative trade volume during QE is

between 2018 and 2020 before the start of QE5, and in months where no reinvestments took place, which is a total of 14 months.

significantly higher than during QT. Instead, the trade activity of dealers and especially hedge funds is relatively larger during QT than during QE, which fills in for the relative absence of ICPF trades during this period.

In the next section, we formally explore the asymmetries regarding the patterns of ICPF demand during QE and QT phases.

4 Testable Hypotheses, Methodology, and Results

In this section, we investigate how the trading volume of ICPFs investors with dealers affects the bidding behavior of dealers during auctions. For this, we first provide testable hypotheses that guide our empirical analysis.

Our first hypothesis is inspired by the literature that argues that ICPFs belong to the preferred habitat group of investors because they have a strong and less elastic preference for government bonds than other investors (see, for example, Giese et al. (2024); Vayanos and Vila (2021); Krishnamurthy and Vissing-Jorgensen (2012)). Our second hypothesis proposes that this effect would be higher for longer maturities, especially for very long maturities, where ICPFs display strongest preferred habitat (see Giese et al. (2024), and Vayanos and Vila (2021)). Correspondingly, we formulate the hypotheses as:

Hypothesis 1: The higher the trading volume between preferred habitat investors and dealers around the QE (or QT) auction day, the higher the dealer bid will be.

Hypothesis 2: The magnitude of the effect described by Hypothesis 1 will be larger for longer maturity bonds.

4.1 Estimation Methodology

In this subsection, we test our hypotheses and analyze the relationship between dealer bids at the APF auctions and the dealer trades with ICPFs using panel regressions.

Let i refer to the individual dealer, b refer to the specific bond, and t to the auction date. Our dependent variable is the average yield spread to benchmark yield, $Y_{i,b,t}$, calculated as the average offer yield across all bids for bond b made by dealer i at a given auction minus the corresponding benchmark yield established by the BoE for that bond at that auction. The main variable of interest is $PHRatio_{i,b,t}$, calculated as the ratio of a gilt b volume bought (sold) by a dealer i from ICPFs relative to the total volume of that gilt b bought (sold) by a dealer i on the secondary market on the auction date

t. Another variable of interest is $HFRatio_{i,b,t}$, similarly defined for the dealer i as the share of secondary market trades with Hedge Fund (HF) investors in the gilt b relative to total secondary market trade by the dealer i in gilt b on the day of the auctions.

To investigate the role played by the PH investors in shaping dealers' bidding behavior, we estimate the following panel data regression:

$$\mathrm{Y}_{i,b,t} = \beta_1 PHRatio_{i,b,t} + \beta_2 HFRatio_{i,b,t} + \beta Z_{i,b,t} + \xi_i + \mu_b + \psi_t + \epsilon_{i,b,t}, \tag{4.1}$$

where $Z_{i,b,t}$ is a vector of bond-, dealer-, and bond/dealer-specific control variables, including total bid quantity for a bond by a dealer, dealer balance sheet constraints (measured by the total amount of gilts accumulated by a dealer during the two days prior to the auction), and a bond-specific liquidity control (calculated as the total amount of that bond bought and sold by all investors relative to the total free float available). Finally, ξ_i refers to dealer fixed effects, μ_b to bond fixed effects, and ψ_t time fixed effects.

To estimate the regressions we focus on the most relevant trades in the secondary markets. In particular, for QE, we narrow the trades to only those in which GEMMs were the buyers. For QT, we narrow the trades to only those in which the dealer was the seller. We use all bids and do not restrict to only successfully allocated bids because *ex-ante* dealers did not know if their bids would be successful. We restrict our observations to only those in which preferred habitat investors traded with the dealers within the specific time horizon. We also include the associated bid quantity to control for any price affects associated with the size of bid.

Table 2 contains summary statistics of the variables used for the regressions on our QE and QT samples. The main variable of interest, $PHRatio_{i,b,t}$, which is the ratio of ICPFs trades on the auction dates, was higher during QE (12 percent) than during QT (8 percent), suggesting that ICPFs were more active around QE auctions. The relative role of HFs was stable over the two samples, accounting for 27 and 24 percent during QE and QT respectively, although on average hedge funds trades with dealers were larger during QT than during QE. Finally, the average maturity of the auctioned gilts have naturally been higher during QE (around 18 years) than during its subsequent unwind (around 14 years).

Turning to our second hypothesis, we constrain the regression in equation (4.1) to be estimated only on long maturity bonds; that is, on bonds with time to maturity

Table 2: Summary statistics of the regression variables

QE Sample	**Mean**	**Min**	**Median**	**Max**	**Std. Dev.**	**N**
Individual Bond Issuance (£m)	27,937	4,037	28,385	41,896	8,111	8,829
Individual Bond Free Float (£m)	15,902	1,234	14,921	28,481	5,420	8,829
PH Amount Sold (£m)	3	0	0	1,408	25	8,829
HF Amount Sold (£m)	34	0	0	2,559	124	8,829
Total Volume Traded (£m)	1,292	0	652	18,917	1,814	8,829
Volume Purchased Dealer (£m)	2,184	52	1,797	18,285	1,865	8,829
Time To Maturity (Years)	17.79	3.00	14.35	53.25	12.77	8,829
Bond Specialness (bps)	-7.88	-53.30	-6.75	50.00	8.52	8,754
Bond Return Volatility (%)	0.58	0.02	0.33	26.57	0.87	8,790
Duration (Years)	14.55	2.93	11.99	38.81	9.31	8,829
PH Ratio	0.12	0	0	1	0.29	8,002
HF Ratio	0.27	0	0	1	0.41	8,002
Dealer Ratio	0.38	0	0.07	1	0.44	8,002
QT Sample						
Individual Bond Issuance (£m)	32,959	12,854	32,007	44,643	6,377	1,421
Individual Bond Free Float (£m)	16,088	4,564	14,290	32,409	5,666	1,421
PH Amount Bought (£m)	2	0	0	430	16	1,421
HF Amount Bought (£m)	53	0	0	1,800	159	1,421
Total Volume Traded (£m)	1,259	0	707	9,428	1,504	1,421
Volume Sold Dealer (£m)	185	0	133	1,432	200	1,421
Time To Maturity (Years)	14.18	3.01	10.12	48.73	11.56	1,421
Bond Specialness (bps)	-17.91	-118.19	-6.31	7.66	25.98	1,421
Bond Return Volatility (%)	0.54	0.04	0.40	5.15	0.50	1,356
Duration (Years)	10.55	2.88	8.44	31.42	6.99	1,421
PH Ratio	0.08	0	0	1	0.25	1,301
HF Ratio	0.24	0	0	1	0.39	1,301
Dealer Ratio	0.41	0	0.14	1	0.44	1,301

Note: This table displays the summary statistics for the main variables used in the regression analysis for the QE sample (2018–2021) and QT sample (2022–2025) separately. PH refers to Preferred Habitat investors, and HF refers to Hedge Funds.

exceeding 20 years:

$$Y_{i,b,t|m>20y} = \beta_1 PHRatio_{i,b,t} + \beta_2 HFRatio_{i,b,t} + \beta Z_{i,b,t} + \xi_i + \mu_b + \psi_t + \epsilon_{i,b,t}. \qquad (4.2)$$

The estimated regression coefficients from equations (4.1) and (4.2) are shown in Table 3.

Table 3: Estimates of the PH demand impact during QE and QT auctions

Dependent Variable: Average Bid Spread to Benchmark Yield							
	QE				**QT**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PH Ratio	-0.008^{***}	-0.011^{**}	-0.005^{*}	-0.008^{**}	0.007	0.023	0.002
	(0.003)	(0.005)	(0.003)	(0.003)	(0.016)	(0.027)	(0.004)
HF Ratio	–	–	0.007^{**}	0.385	–	0.035	-0.010^{***}
			(0.003)	(0.271)		(0.033)	(0.003)
Obs.	1,901	1,110	1,901	1,368	182	182	1,301
Adj. R^2	0.231	0.356	0.233	0.192	−0.261	−0.220	0.13

Note: This table displays the coefficient estimates of equation (4.1) for QE auctions in columns one to four, and QT auctions in columns five to seven. The dependent variable for all specifications is the average bid to benchmark yield spread. Column 1 contains just the *PH Ratio* alongside all controls, column 2 uses all bonds with a time to maturity over 20 years, column 3 uses all bonds but includes the hedge fund ratio (*HF Ratio*), and column 4 includes all bonds but only auctions where the *HF Ratio* was below 5%. Column 5 for QT auctions includes only the *PH Ratio*, column 6 includes both *PH Ratio* and *HF Ratio*, and column 7 is the only specification where the condition to only include *PH Ratio* more than 0 is dropped. SEs are clustered at the bond level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.2 Estimation Results

We first discuss the estimation output for the case of QE auctions, including the estimated regression results for various specifications of equations (4.1) and (4.2) over several subsamples. Then, in the following subsection, we examine the significance of preferred-habitat demand pressures on the sample covering QT auctions.

QE Results. Before turning to each specification, we start by noting that the coefficients of interest (loadings on PH demand variables) are statistically significant (see Table 3 Columns 1-4) over the QE subsample. The coefficient β_1 is estimated to be negative in equation (4.1), which implies that increased preferred-habitat demand (or, equivalently, larger share of gilts sold to a dealer by ICPFs investors) is associated with lower average bid spread to benchmark yield at the QE auctions. The result holds for the full sample of QE purchased bonds (see Column 1), and for bonds with a time to maturity over 20 years (see Column 2). The findings confirm the importance of the demand from institutional investors during QE: higher ICPFs investor sales to dealers lead to stronger dealer bidding during QE auctions, consistent with habitat demand Hypothesis 1. Moreover, the magnitude of the effect is larger when we look at longer maturity bonds, providing support for habitat demand Hypothesis 2.

Table 4: Estimates of the PH demand impact using alternative samples

Dependent Variable: Average Bid Spread to Benchmark Yield						
	QE			**QT**		
	(1)	(2)	(3)	(4)	(5)	(6)
PH Ratio	−0.008***	–	–	0.007	–	–
	(0.003)			(0.016)		
PH Ratio + 1 Day	–	−0.006*	–	–	−0.008	–
		(0.003)			(0.010)	
PH Ratio + 2 Days	–	–	−0.004	–	–	−0.003
			(0.003)			(0.011)
Obs.	1,901	1,973	1,979	182	193	193
Adj. R^2	0.231	0.236	0.224	−0.261	−0.219	−0.224

Note: This table displays the coefficient estimates of equation (4.1) for QE auctions in columns one to four, and QT auctions in columns five to seven. The dependent variable for all specifications is the average bid to benchmark yield spread. The *PH Ratio* is estimated over alternative windows: Column 1 and 4 cover only trades on the day of the auctions, column 2 uses all trades with PH investors over the day of the QE auction and the preceding day, column 3 uses all trades during the QE auction day and preceding two days; Column 5 includes the *PH Ratio* calculated for the trades on day of the QT auction and a following day, column 6 includes the PH trades over the QT auction day and two subsequent days. SEs are clustered at the bond level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Next, analyzing the role of hedge funds (columns 3 and 4), we find that the coefficient of hedge fund trade volumes is positive, and in the full sample is strongly statistically significant. The preferred habitat demand effect remains negative and statistically significant, though the magnitude is slightly smaller (compare columns 1 and 3).

Because it could take some time for dealers to procure gilts for QE auctions and to sell gilts after QT auctions, we also estimate the regression where the right hand side variables of interest, *PH Ratio* and *HF Ratio*, are estimated over a wider window covering the adjacent days. Table 4 shows the coefficient estimates for these specifications. Similar to our baseline results, when we estimate the impact of trades with PH investors over a wider window, capturing not only the day of the auction but also the previous day, the coefficient on *PH Ratio* remains significant, although the significance is weaker. Finally, when we widen the window further, the significance fades away, which is understandable given that the wider window would capture a larger share of trades unrelated to the QE auctions.

In summary, estimation results over the QE sample do not reject the PH demand channel hypotheses stated in the beginning of the section. Higher preferred habitat trade flows with dealer around auction dates lead to lower average spreads to benchmark prices relative to other dealers, or in other words, higher auction bid prices, consistent with the hypothesis that preferred habitat investors value the bonds they invest in more than other investors such as hedge funds.

QT Results. We now turn to analyzing QT auctions. It is important to note that there have been relatively few QT auctions compared to QE auctions. However, our QT sample covers more than two years of active sales by the BoE. Over that period the APF portfolio has been unwound by around 25 percent. Thus, our results still can provide insights into potential asymmetries between QE and QT.

Table 3 reports our baseline QT results (Columns 5-7). The coefficient of interest *PH Ratio* is statistically insignificant, implying a weaker role of ICPF investors during the QT auctions. As a result this suggests that the preferred habitat demand part of portfolio-balance channel is not key for the policy transmission during this phase of the unconventional monetary policy. The dealers' willingness to buy at the QT auctions instead seems to be more driven by hedge fund activity (*HF Ratio* in Column 7). In summary, we find that insurance companies and pension funds have been less active in QT auctions compared to QE auctions, while hedge funds have absorbed in proportion more gilts than during QE.

We can think of various reasons why this might be the case. The structural transition from defined benefit (DB) to defined contribution (DC) pension provision has reduced the natural demand for UK gilts by shifting the ICPFs away from liability-driven investment. While DB schemes have historically been price-insensitive holders of long-term gilts to hedge guaranteed liabilities (typical of PH investors), DC schemes instead are less tied to balance-sheet liabilities and allocate primarily on return considerations, with limited demand for long-dated government bonds. Moreover, this compositional shift was accentuated by the 2022 LDI crisis (Alexander et al. (2023)), which prompted tighter regulation and a more cautious approach to gilt exposure among remaining DB schemes. While these structural changes matter, a plausible primary driver of the decline in excess demand for gilts has been the sharp increase in net gilt supply in the post-COVID period—reflecting elevated fiscal borrowing (see Figure 1). The relatively less appealing nature of holding gilts in the higher inflation environment and stronger preference for the newly issued gilts by DMO than for APF held gilts are two other potential drivers of the reduced demand from institutional investors. We explore these hypotheses and others in more detail in the next section.

4.3 Robustness

Accounting for Interest Rate Risk. Our analysis of QE and QT covers the sample between 2018 and 2025. This period was characterized by changing interest rate environment. While QE was implemented when policy rates were held close to the effective lower bound, the QT program has been launched at the time of rising policy rates, which may be associated with elevated interest rate risk. Dealers, who typically lend long term and borrow short term, are vulnerable to interest rate risk, which could affect their bidding behavior in QE and QT auctions. Moreover, amid elevated interest rate volatility, Value-at-Risk (VaR) limits on Treasury trading desks may have become more binding therefore limiting dealers' capacity to intermediate (see e.g. Duffie et al. (2023), or Li et al. (2025)). Furthermore, pension funds with long-term liabilities are highly sensitive to fluctuations in interest rates as well (see e.g. Jansen et al. (2024a)).

Therefore, to analyze the robustness of our results to changing interest rate risk and various market aspects it captures (dealers' constraints and ability to support market functioning, in particular), we introduce two additional controls. First, we include the standard deviation of bond specific yield changes from the previous five days on a rolling basis across the sample in order to capture bond specific volatility. Second, we account for the price sensitivity to interest rate risk as captured by a duration measure. Figure 4 plots the average modified duration of the APF bond portfolio since 2018, the start of our sample.

Table 5 builds on the regression results in Table 3 by adding the modified duration and interest rate volatility measures. The first and main observation is that neither measure is statistically significant at the 90% confidence interval in any specification. This suggests that any role of interest rate risk is either not statistically significant in affecting the bidding behavior of dealers, or that the effects are already being captured in the other controls and fixed effects. The magnitude of the key coefficient of interest *PH Ratio* is quantitatively similar as our baseline right in both the QE and QT sample. In particular, the key asymmetry between QE and QT we found in the baseline persists.

Dealer Intermediation Capacity and Hedge Funds. The importance of dealer intermediation capacity and its effects on market liquidity and functioning is well known (e.g. see Adrian et al. (2017), Duffie (2023); or, closer to our case, Boneva et al. (2025),

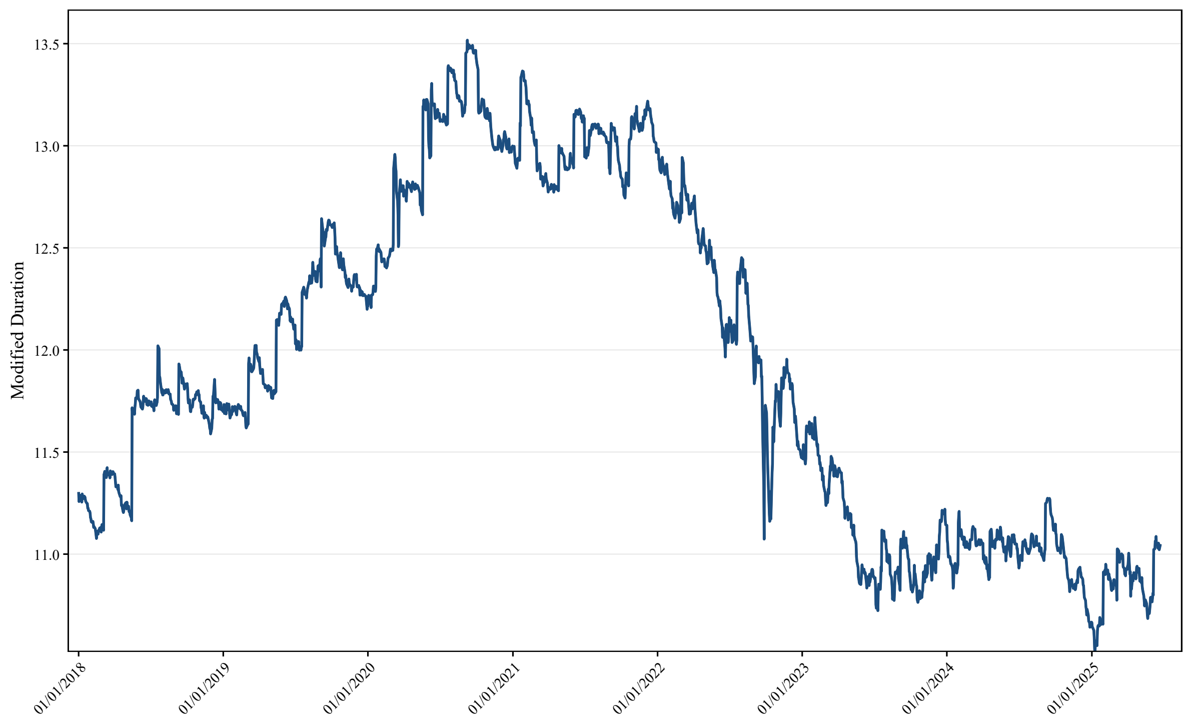


Figure 4: The average modified duration of the APF bond portfolio since 2018

Note: This figure plots the average modified duration of the bonds held within the APF bond portfolio since 2018.

who study dealers behavior at the earlier BoE QE auctions). Since only GEMM's are able to participate in APF auctions, the ability and willingness of these primary dealers to purchase gilts in the auctions and intermediate in secondary markets is of foremost importance, and therefore very relevant to our empirical study.

To investigate the importance of dealer intermediation capacity in dealer auction bids, we calculate a "*Dealer Ratio*," following the same methodology used to calculate the *PH Ratio* and *HF Ratio*. We then include this variable in our panel regression to investigate if our findings regarding the asymmetric role of habitat demand in QE and QT auctions was actually correlated with a given dealer's willingness to intermediate in gilt markets. In other words, the economic rationale for including the *Dealer Ratio* is to test whether the preferred habitat demand effect we have documented is in fact just capturing the reluctance (or inability) of dealers to warehouse gilts and intermediate in secondary markets.

Table 6 includes the *Dealer Ratio* alongside the *PH Ratio*. We find that both our QE and QT results regarding habitat demand are highly similar as our baseline results.

Table 5: Estimates of the impacts of PH demand and interest rate risk

Dependent Variable: Average Bid Spread to Benchmark Yield							
	QE				**QT**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PH Ratio	-0.007^{**}	-0.009^{*}	-0.005^{*}	-0.008^{**}	0.008	0.022	0.002
	(0.003)	(0.005)	(0.003)	(0.003)	(0.018)	(0.029)	(0.005)
HF Ratio	–	–	0.007^{*}	0.193	–	0.032	-0.011^{***}
			(0.003)	(0.130)		(0.035)	(0.003)
Duration	0.002	0.001	0.002	0.002	0.016	0.008	0.007
	(0.002)	(0.004)	(0.002)	(0.003)	(0.030)	(0.027)	(0.008)
IR Vol	0.097	0.032	0.106	0.164	−1.070	−0.879	0.559
	(0.093)	(0.143)	(0.093)	(0.132)	(1.652)	(1.831)	(0.219)
Obs.	1,897	1,107	1,897	1,365	167	167	1,239
Adj. R^2	0.246	0.381	0.247	0.207	−0.331	−0.306	0.122

Note: This table displays the coefficient estimates of equation (4.1) for QE auctions in columns one to four, and QT auctions in columns five to seven. The dependent variable for all specifications is the average bid to benchmark yield spread. Column 1 contains just the *PH Ratio* alongside all controls, column 2 uses all bonds with a time to maturity over 20 years, column 3 uses all bonds but includes the hedge fund ratio *HF Ratio*, and column 4 includes all bonds but only auctions where the *HF Ratio* was below 5%. Column 5 for QT auctions includes only the *PH Ratio*, column 6 includes *PH Ratio* and *HF Ratio*, and column 7 is the only specification where the condition to only include *PH Ratio* more than 0 is dropped. Bond and time specific proxies for interest rate risk are captured by modified duration ('Duration'), and the standard deviation of daily bond yield changes over the previous five days ('IR Vol'). SEs are clustered at the bond level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Additionally, the *Dealer Ratio* variable is also not statistically significant. This result suggests that at least during QE and QT auctions considered, the balance sheet capacity of dealers did not affect their gilt bidding behavior, which is therefore more likely driven by their clients' demand for the gilts being auctioned.

Linking Supply to Preferred Habitat Demand. Using our bond specific free-float variable which captures the amount of the bond is available to buy, relative to the total amount issued, we can test whether bonds which had relatively more available free float also had a relatively weaker preferred habitat demand channel. We define high free-float as a dummy which equals one if the bond-time observation was in the top half of all observations, and interact it with the our key variable of interest, the *PH Ratio*.

Table 7 shows that the more available free float, or supply available to investors, the weaker the preferred habitat demand channel. We find that the coefficient for *PH Ratio* is closer to zero when the outstanding bond supply is high.

Table 6: The effect of primary dealer trades

Dependent Variable: Average Bid Spread to Benchmark Yield					
	QE			**QT**	
	(1)	(2)	(3)	(4)	(5)
PH Ratio	−0.009***	−0.013**	−0.012*	−0.002	0.007
	(0.003)	(0.005)	(0.006)	(0.013)	(0.005)
Dealer Ratio	−0.004	−0.007	0.069	−0.036	0.006**
	(0.003)	(0.005)	(0.169)	(0.023)	(0.003)
Obs.	1,901	1,110	798	182	1,301
Adj. R^2	0.231	0.357	0.171	−0.227	0.121

Note: This table displays the coefficient estimates of equation (4.1) for QE auctions in columns one to three, and QT auctions in columns four to five, with the inclusion of *Dealer Ratio* in the place of *HF Ratio*. The dependent variable for all specifications is the average bid to benchmark yield spread. Column 1 contains the *PH Ratio* and *Dealer Ratio* alongside all controls, column 2 uses all bonds with a time to maturity over 20 years, and column 3 includes all bonds but only auctions where the *Dealer Ratio* was below 5%. Column 4 for QT auctions includes the *PH Ratio* and *Dealer Ratio*, column 5 is the only specification where the condition to only include *PH Ratio* more than 0 is dropped. SEs are clustered at the bond level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Taking stock, we find that the asymmetry between the habitat demand transmission mechanism during QE vs. QT is not driven by variation in interest rate risk or dealer constraints. Instead, our results suggest that increases in the outstanding supply of gilts are one of the key drivers in explaining the weakening habitat demand channel. We explore and develop this mechanism further in our theoretical model.

4.4 Additional Evidence: DMO Auctions

To further corroborate our findings regarding the role of the increased supply of gilts for the asymmetric PH demand impacts during QE and QT, in this section we investigate DMO gilt absorption by preferred habitat investors in the increasing gilt free float environment (Figure 1). For this, we compare the ratio of net purchases to auctioned gilt amount from Debt Management Office (DMO) auctions between the pre-QE period (April 2018 to March 2020), and the ongoing QT period (January 2023 onwards for this specific exercise). We focus on nominal gilts only (i.e. we exclude auctions for index-linked gilts) and consider both regular and syndicate auctions.

First, we calculate net volumes (total purchase volumes minus sales) relative to the auctioned bond amount, across a range of investor groups up to 22 trading days

Table 7: The role of supply on PH demand

Dependent Variable: Average Bid Spread to Benchmark Yield			
	(1)	(2)	(3)
PH Ratio	-0.008^{***}	-0.011^{***}	-0.008^{**}
	(0.003)	(0.004)	(0.003)
High Free Float	–	0.005	0.001
		(0.002)	(0.004)
PH Ratio × High Free Float	–	0.005^{**}	0.005^{**}
		(0.002)	(0.002)
HF Ratio	–	–	0.008^{**}
			(0.003)
Obs.	1,901	1,901	1,901
Adj. R^2	0.231	0.231	0.233

Note: This table displays the regression results when interacting our main variable of interest, *PH Ratio*, with a free-float dummy variable which equals one if the free-float of the given bond-time observation is above the sample median. SEs are clustered at the bond level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

(approximately one month) after a given DMO auction for long maturity gilts only. We then estimate the following dynamic difference-in-difference:

$$\text{Gilt Absorption Ratio}_{i,\tau} = \sum_{d=1}^{22} \beta_d \, \mathbb{1}[\tau = d] \times \text{QT Period} + \alpha_i \; + \; \alpha_t \; + \; \epsilon_{i,\tau}, \qquad (4.3)$$

where Gilt Absorption $\text{Ratio}_{i,\tau}$ is the ratio of net volumes to auctioned amounts for auctioned bond i, τ days after the auction where $\tau = 0, \ldots, 22$, QT Period is a dummy variable which equals one if the DMO auction was held from January 2023 onwards, β_d measures the trading day specific difference between the QT period ratio relative to the pre-QE5 period, α_i capture bond specific fixed effects, α_t is a date fixed effect, and $\epsilon_{i,\tau}$ a residual.

In line with the evidence presented from the auctions data in Section 4.2, Figure 5 shows that the ratio of absorbed gilts from preferred habitat investors for long maturity nominal gilts is significantly less in the QT period than during the pre-QE5 period, by as much as 50 percentage points in the immediate aftermath of the auction. This reduction persists even up to three weeks after the auction date, after which the coefficient estimate finishes negative but not statistically significant from zero. This provides very clear evidence that preferred habitat investor demand for nominal gilts is relatively less than

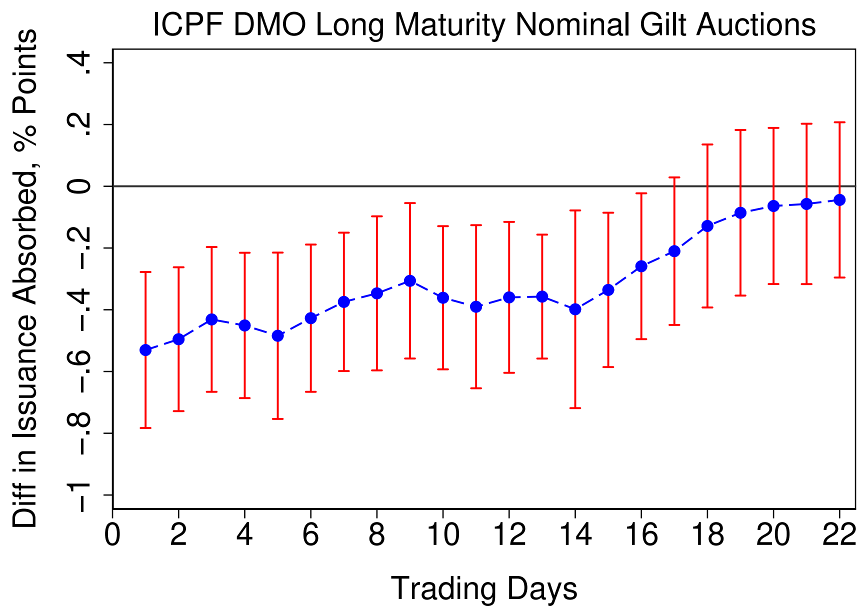


Figure 5: Difference in Gilt Absorption Ratio by ICPFs during QT and QE

Note: This figure plots the coefficient estimates of equation 4.3 for ICPF investors and long maturity gilts only. The control period is pre-QE5 and the treatment period is since January 2023. Bond and date fixed effects are included, and 95% confidence intervals are included vertically alongside coefficient estimates.

when the free float of gilts was smaller.

On the other hand, in Figure 6, we see that hedge fund net volumes are substantially larger during the QT period, with a difference of around 65% after one month. This is consistent with our APF auction results which show that hedge funds are more active in QT auctions than they were during QE auctions, especially for long maturity gilts.

In summary, our analysis of DMO auctions data and the relative net volumes absorbed by various investor groups one month after the auction date confirms that the PH demand has been significantly lower post COVID-19 than it was before gilt free-float increased to record amounts. Instead, a different type of investors—hedge funds, who are less traditional but more flexible and less regulatory constrained—are taking a more prominent role.[10]

To understand how a relatively weaker PH demand and a larger role of the hedge fund activity would impact monetary policy transmission mechanism, in the next section we present an equilibrium yield curve model for analyzing the interaction between PH investors, arbitrageurs, and the size of outstanding government debt.

[10] The increased role in hedge fund trading activity is also visible in the euro case. See here: ECB Blog.

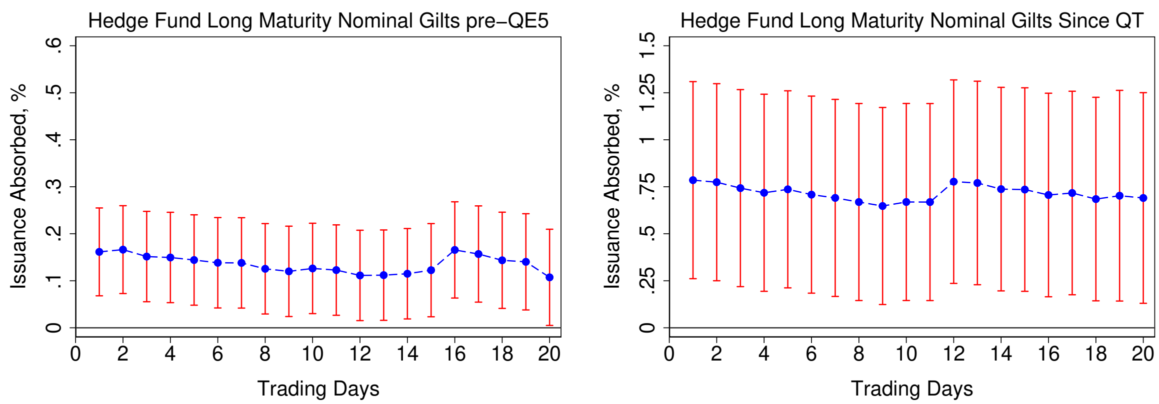


Figure 6: Gilt Absorption Ratio by Hedge Funds during QT and QE

Note: This figure plots the average gilt absorption expressed as a ratio of net volumes to DMO auctioned bonds between April 2018 to March 2020 (pre-QE5 period, left) and since January 2023 (QT period, right) for hedge fund investors for gilts with a time to maturity equal to above twenty years, up to twenty trading days after the auction date.

5 Theoretical Model

In order to make sense of our empirical findings, we model the determination of interest rates as arising from the interaction between fixed income arbitrageurs and preferred habitat investors. Arbitrageurs trade bonds of all maturities in order to satisfy a mean-variance portfolio allocation problem. Preferred habitat investors instead specialize in bonds of specific maturities, with elastic demand increasing in the yield of the specific bond. Our setup is identical to that of Vayanos and Vila (2021), except we allow for the demand elasticity of preferred habitat investors to vary over time as a function of the outstanding supply of bonds.

Formally, we assume there are a continuum of zero-coupon bonds with price $P_t^{(\tau)}$, yields $y_t^{(\tau)} \equiv -\frac{1}{\tau} \log P_t^{(\tau)}$ and maturities $0 \leq \tau \leq T \leq \infty$. The instantaneous nominal short rate is given by $r_t \equiv \lim_{\tau \to 0} y_t^{(\tau)}$.

Arbitrageurs solve the following portfolio allocation problem:

$$\max_{\{X_t^{(\tau)}\}} E_t \, \mathrm{d}W_t - \frac{a}{2} Var_t \, \mathrm{d}W_t \tag{5.1}$$

$$\text{subject to: } \mathrm{d}W_t = W_t r_t \, \mathrm{d}t + \int_0^T X_t^{(\tau)} \left[\frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} - r_t \, \mathrm{d}t \right] \mathrm{d}\tau \, . \tag{5.2}$$

As is clear from the budget constraint (5.2), arbitrageurs may always invest all of their wealth in the instantaneous risk-free rate. Relative to this case, arbitrageurs may seek out higher returns by investing a portion of their wealth $X_t^{(\tau)}$ in the bond carry trade across any maturity τ. The realized returns are represented by the integral term in the budget constraint. The risk aversion parameter a governs the trade-off between expected returns and the price risk inherent in conducting such carry trades.

Preferred-habitat demand is assumed to follow

$$Z_t^{(\tau)} = -\alpha_t(\tau) \log P_t^{(\tau)} - \beta_t^{(\tau)}. \tag{5.3}$$

The elasticity $\alpha_t(\tau) > 0$ is a function of maturity τ. Additionally, habitat elasticities vary over time. Time-variation arises as a function of the outstanding stock of bonds; in particular, we model the elasticity function in such a way that demand elasticities decrease as the outstanding stock of bonds increases. We make this assumption to capture the idea that habitat investors become "satiated" when the market is flooded with bonds, and thus require larger price movements to be induced to purchase additional bonds. The term $\beta_t^{(\tau)}$ represents additional sources of variation in habitat demand which are independent of movements in prices ("noise" demand).

In general, the dynamics of the economy are governed by a set of risk factors $\mathbf{q}_t \in \mathbb{R}^J$ which evolve according to the vector Ornstein-Uhlenbeck process

$$\mathrm{d}\mathbf{q}_t = -\mathbf{\Gamma} \mathbf{q}_t \, \mathrm{d}t + \boldsymbol{\sigma} \, \mathrm{d}\mathbf{B}_t \,, \tag{5.4}$$

where the dynamics and diffusion matrices $\mathbf{\Gamma}, \boldsymbol{\sigma}$ are primitives of the model. The vector $\mathbf{B}_t$ is a set of standard independent Brownian motions. For simplicity, we assume that the risk factors $\mathbf{q}_t$ are in terms of deviations from steady state. The short rate $r_t \equiv \mathbf{e}_r^\top \mathbf{q}_t$ is a function of the risk factors (and may be included in the set of risk factors, in which case the vector $\mathbf{e}_r$ is a standard normal basis vector).

The outstanding stock of bonds with maturity τ is given by

$$S_t^{(\tau)} \equiv \hat{\mathbf{\Theta}}(\tau)^\top \mathbf{q}_t, \tag{5.5}$$

where the vector function $\hat{\mathbf{\Theta}}(\tau)$ governs how movements in risk factors lead to changes in the supply of bonds with maturity τ. The habitat demand elasticities $\alpha_t(\tau)$ depend on the total supply of bonds $S_t \equiv \int_0^T S_t^{(\tau)}\,\mathrm{d}\tau$. Hence, market clearing is therefore given by

$$S_t^{(\tau)} = X_t^{(\tau)} + Z_t^{(\tau)} \tag{5.6}$$

$$\implies X_t^{(\tau)} = \alpha(\tau, \mathbf{q}_t) \log P_t^{(\tau)} + \mathbf{\Theta}(\tau, \mathbf{q}_t)^\top \mathbf{q}_t, \tag{5.7}$$

where in (5.7) we have written the habitat demand slope and intercept as functions of risk factors $\mathbf{q}_t$, by defining $\mathbf{\Theta}(\tau, \mathbf{q}_t)$ so that $\mathbf{\Theta}(\tau, \mathbf{q}_t)^\top \mathbf{q}_t = \beta(\tau, \mathbf{q}_t) + \hat{\mathbf{\Theta}}(\tau)^\top \mathbf{q}_t$.

5.1 Equilibrium

Because demand elasticities vary over time, unlike in Vayanos and Vila (2021) our model is no longer consistent with a time-homogenous (log) affine solution. Nevertheless, it is useful to decompose bond prices as

$$-\log P_t^{(\tau)} = \mathbf{A}(\tau, \mathbf{q}_t)^\top \mathbf{q}_t + C(\tau). \tag{5.8}$$

The coefficient functions $\mathbf{A}(\tau, \mathbf{q}_t)$ and $C(\tau)$ are endogenous; in this section, we characterize the equilibrium as the solution for these coefficient functions.

To simplify notation, we suppress notational dependence on maturity τ and risk factors $\mathbf{q}_t$; we denote partial derivatives as: $\partial_\tau \mathbf{A} \equiv \frac{\partial \mathbf{A}(\tau, \mathbf{q}_t)}{\partial \tau}$ and $\nabla_{\mathbf{q}} \mathbf{A} \equiv \nabla_{\mathbf{q}} \mathbf{A}(\tau, \mathbf{q}_t)$. Without loss of generality, we assume that log prices $\log P_t^{(\tau)}$ in the model are measured in terms of deviations from steady state, so that $C(\tau) = 0$ for all maturities. With an appropriate definition of the demand intercept function $\beta(\tau, \mathbf{q}_t)$, we therefore have that in steady state ($\mathbf{q}_t = \mathbf{0}$), supply $S(\tau, \mathbf{0}) = 0$ and habitat demand $Z(\tau, \mathbf{0}) = 0$.

Applying Ito's Lemma to (5.8) gives

$$\frac{\mathrm{d}P_t^{(\tau)}}{P_t^{(\tau)}} = \mu(\tau, \mathbf{q}_t)\,\mathrm{d}t + \boldsymbol{\sigma}(\tau, \mathbf{q}_t)\,\mathrm{d}\mathbf{B}_t \tag{5.9}$$

$$\text{where } \mu(\tau, \mathbf{q}_t) = \left(\partial_\tau \mathbf{A} + \mathbf{\Gamma}^\top \left[\mathbf{A} + \nabla_{\mathbf{q}} \mathbf{A}^\top \mathbf{q}_t\right]\right)^\top \mathbf{q}_t + \frac{1}{2}\mathcal{T}_t, \tag{5.10}$$

$$\boldsymbol{\sigma}(\tau, \mathbf{q}_t) = -\left[\mathbf{A} + \nabla_{\mathbf{q}} \mathbf{A}^\top \mathbf{q}_t\right] \boldsymbol{\sigma}, \tag{5.11}$$

and where the term the convexity term $\mathcal{T}_t$ depends on second-order derivatives.

The arbitrageur optimality conditions imply that

$$\mu(\tau, \mathbf{q}_t) - r_t = \boldsymbol{\sigma}(\tau, \mathbf{q}_t)\boldsymbol{\Lambda}_t, \tag{5.12}$$

$$\text{where } \boldsymbol{\Lambda}_t^\top = a \int_0^T X_t^{(\tau)} \boldsymbol{\sigma}(\tau, \mathbf{q}_t)\, \mathrm{d}\tau\,. \tag{5.13}$$

Note that the market price of risk $\boldsymbol{\Lambda}_t$ depends on arbitrageur risk aversion, and the total quantity of risk exposure (captured by the integral terms). This expression is the same as in Vayanos and Vila (2021). The difference is that, through market clearing, the risk exposure of arbitrageurs will depend on time-variation in demand elasticities.

Combining arbitrageur optimality conditions with market clearing (5.7), and imposing the minor approximation condition $\mathcal{T}_t \approx 0$,[11] we can characterize the equilibrium as follows:

$$\mathbf{0} = \partial_\tau \mathbf{A} + \mathbf{M}\left[\mathbf{A} + \nabla_{\mathbf{q}} \mathbf{A}^\top \mathbf{q}_t\right] - \mathbf{e}_r, \tag{5.14}$$

$$\text{where } \mathbf{M} \equiv \boldsymbol{\Gamma}^\top - \int_0^T \left[-\alpha \mathbf{A} + \boldsymbol{\Theta}\right]\left[\mathbf{A} + \nabla_{\mathbf{q}} \mathbf{A}^\top \mathbf{q}_t\right]^\top \mathrm{d}\tau\, \hat{\boldsymbol{\Sigma}}, \tag{5.15}$$

and where $\hat{\boldsymbol{\Sigma}} = a\boldsymbol{\sigma}\boldsymbol{\sigma}^\top$. As in Vayanos and Vila (2021), through market clearing, the risk-adjusted dynamics matrix $\mathbf{M}$ solves a complicated fixed-point problem which balances arbitrageur optimality conditions and habitat demand curves. Two additional complexities arise in our setting. First, the matrix $\mathbf{M}$ is not time-homogeneous, but instead depends on the value of risk factors $\mathbf{q}_t$ through the demand functions $\alpha(\tau, \mathbf{q}_t)$ and $\boldsymbol{\Theta}(\tau, \mathbf{q}_t)$. Second, equilibrium is characterized by a set of PDEs, rather than ODEs (after imposing the initial conditions $\mathbf{A}(0, \mathbf{q}_t) = \mathbf{0}$).

5.2 Analytical Results

To gain intuition, we first study a simplified single risk-factor version of the model. We suppose that

$$q_t = -\kappa_q q_t\, \mathrm{d}t + \sigma\, \mathrm{d}B_t\,, \tag{5.16}$$

$$r_t = e_r q_t, \quad S_t(\tau) = -\gamma(\tau) q_t. \tag{5.17}$$

[11] We discuss approximation results further in the next section.

That is, a single stochastic risk factor governs the behavior of both the short rate and outstanding supply $S_t = \int_0^T S_t(\tau)\,\mathrm{d}\tau$. We additionally consider deterministic shocks to bond demand of the form

$$\mathrm{d}\beta_t = -\kappa_\beta \beta_t \,\mathrm{d}t\,, \tag{5.18}$$

where note that shocks to β_t are unanticipated by arbitrageurs. Movements in β_t affect bond market clearing through

$$\delta(\tau)\beta_t - \gamma(\tau)q_t = X_t^{(\tau)} + Z_t^{(\tau)}. \tag{5.19}$$

Thus, an increase in $\beta_t > 0$ acts as a QT shock; while a decrease $\beta_t < 0$ acts as a QE shock.

Finally, we assume that the habitat demand elasticity function $\alpha(S_t, \tau) > 0$, $\frac{\partial \alpha(S_t,\tau)}{\partial S_t} < 0$; and $\alpha(S_t, \tau) > 0 \to \underline{\alpha}(\tau)$ as $S_t \to \infty$, while $\alpha(S_t, \tau) \to \bar{\alpha}(\tau)$ as $S_t \to -\infty$.

Analytical results: limiting cases. We first study the limiting behavior of the model. As $S_t \to \pm\infty$, the affine coefficient equation (5.8) approaches

$$-\log P_t^{(\tau)} \to \begin{cases} \bar{A}_q(\tau)q_t + \bar{A}_\beta\beta_t + \bar{C}(\tau) & \text{as } S_t \to \infty \\ \underline{A}_q(\tau)q_t + \underline{A}_\beta\beta_t + \underline{C}(\tau) & \text{as } S_t \to -\infty \end{cases}. \tag{5.20}$$

The following proposition characterizes the behavior of the model at these limits. All proofs are in Appendix A.

Proposition 1 (Limiting supply results)**.** *Assume risk aversion $a > 0$, risk $\sigma > 0$; and that the short rate and bond demand mappings satisfy $e_r > 0, \delta(\tau) > 0, \gamma(\tau) > 0$. In response to a negative bond demand shock $\beta_t > 0$, bond yields rise on impact at either limiting case: $\bar{A}_q(\tau) > 0, \underline{A}_q(\tau) > 0$. Under the sufficiency condition that $\|\underline{\alpha}\| < K$ and $\|\bar{\alpha}\|^{-1} < K$ for a constant K defined in (A1):*

- *The effect is amplified when $S_t \to \infty$, and mitigated when $S_t \to -\infty$: $\bar{A}_q(\tau) > \underline{A}_q(\tau)$.*
- *Arbitrageurs absorb more of the change in bonds when $S_t \to \infty$ relative to $S_t \to$*

$-\infty$:

$$\lim_{S_t \to \infty} \frac{\partial X_t^{(\tau)}}{\partial \beta_t} > \lim_{S_t \to -\infty} \frac{\partial X_t^{(\tau)}}{\partial \beta_t}.$$

Intuitively, Prop. 1 follows because a QT shock implies that arbitrageurs in equilibrium must hold more duration risk. When arbitrageurs are risk-averse, this requires compensation in the form of higher expected returns, which pushes up yields. However, the increase in yields induces price-elastic habitat investors to purchase some of the increased outstanding bonds. This interaction between arbitrageurs, investors, and supply is present in many preferred habitat models, both in partial equilibrium (Vayanos and Vila (2021), Greenwood and Vayanos (2014)) or in general equilibrium (Ray et al. (2024), Kamdar and Ray (2025)). What is novel in our setting is that when outstanding supply is very scarce ($S_t \to -\infty$), habitat investors are the most elastic; thus, this offsetting effect mitigates the price impact of QT. On the other hand, if outstanding supply is already large ($S_t \to \infty$), habitat investors require significantly larger yield movements to induce them to buy any more bonds. Thus, arbitrageurs absorb the majority of the QT shock, and therefore the price impact is amplified.

Analytical results: quadratic approximation. We next study the model near steady state, and employ second-order approximations to study the state-dependence of QE/QT shocks. We now write

$$-\log P_t^{(\tau)} = A_{1q}(\tau)q_t + A_{1\beta}(\tau)\beta_t + A_{2q}(\tau)q_t^2 + A_{2\beta}(\tau)\beta_t^2 + A_{2q\beta}(\tau)q_t\beta_t + C(\tau) + O(\xi^3), \tag{5.21}$$

where the term $O(\xi^3)$ represents higher-order terms. The habitat demand elasticity function can be approximated as

$$\alpha(S_t, \tau) = \alpha(\tau) + \alpha_q(\tau)q_t + O(\xi^2), \tag{5.22}$$

where $\alpha_q(\tau) > 0$ since when $\gamma(\tau) > 0$, $\frac{\partial S_t}{\partial q_t} < 0$.[12] The following proposition characterizes the first- and second-order behavior of the model near the steady state.[13]

[12] As shown in the proof, since habitat demand is zero in steady state, we only need to keep track of first-order terms in the elasticity function.

[13] We follow the limited arbitrage literature and use a "low risk, low risk-bearing capacity" perturbation approach (see e.g., Itskhoki and Mukhin (2021) or Gourinchas et al. (2025)).

Proposition 2 (Quadratic approximation results)**.** *Assume risk aversion $a > 0$, risk $\sigma > 0$; and that the short rate and bond demand mappings satisfy $e_r > 0, \delta(\tau) > 0, \gamma(\tau) > 0$. Additionally, assume that $a \equiv \xi^{-2}\hat{a}$ and $\sigma \equiv \xi\hat{\sigma}$, where ξ is a perturbation scalar. In response to a negative bond demand shock $\beta_t > 0$, at steady state, bond yields rise on impact: $A_{1\beta}(\tau) > 0$. Under the sufficiency condition that the steady state demand and supply functions are not too large ($\|\alpha\| < K$ and $\|\gamma\| < K$ for a constant K defined in* (A4)*):*

- *The effect is amplified when $S_t > 0$ and mitigated when $S_t < 0$: $A_{2q\beta}(\tau) < 0$.*
- *Arbitrageurs absorb more of the change in bonds when $S_t > 0$ relative to $S_t < 0$:*

$$\left.\frac{\partial X_t^{(\tau)}}{\partial \beta_t}\right|_{S_t>0} > \left.\frac{\partial X_t^{(\tau)}}{\partial \beta_t}\right|_{S_t<0}.$$

Prop. 2 shows that the previous intuition carries over even without extremely large or small outstanding supply. The result follows because of the duration risk which arbitrageurs are required to hold when the central bank conducts QT. Elastic habitat traders in essence reduce the total aggregate duration risk which arbitrageurs must hold in equilibrium. Because QT/QE shocks fundamentally work by either increasing or decreasing the duration risk held by arbitrageurs in equilibrium, when ex-ante outstanding supply is low (and habitat elasticity is therefore high), QT operations therefore have smaller effects. But when outstanding supply has increased and duration risk pricing is more extreme, additional QT operations will have larger price impacts.

Rationalizing our empirical findings, and policy implications. Taking stock, our analytical results provide both an explanation for our empirical findings and additional policy implications. First, when the outstanding stock of bonds is low, relatively small changes in bond prices induce elastic habitat investors to purchase more bonds. In particular this implies that when the central bank conducts balance sheet operations during periods of relative bond scarcity, a sizable portion of the purchases or sales are absorbed by these habitat investors. This helps rationalize our findings regarding QE and QT asymmetries: outstanding gilt supply during QT is much higher than during QE, and hence the habitat investor reactions are mitigated. This also explains the results we find for DMO auctions: habitat demand channels are weakened after gilt issuance has increased to record amounts.[14]

Second, the policy implications follow from the asymmetric reactions of habitat investors. When the absorption of central bank balance sheet policies by habitat investors is small, the changes in bond quantities must therefore hit arbitrageurs' balance sheets more heavily. Therefore, the required change in risk compensation is larger, and all else equal the yield effects of the program will be larger.

Note that our analysis thus far only considered a single risk factor. We now consider a multi-factor model with both short rate risk and supply risk.

5.3 Stylized Parameterization

Moving beyond limiting or approximation results, equations (5.14)-(5.15) characterize the equilibrium of our model, but solving this set of equations in general must be done numerically. In order to continue extracting intuitive insights from these complicated sets of expressions, we next examine a multi-factor version of the model under a simplified parameterization. First, we assume that

$$\boldsymbol{\Theta}(\tau, \mathbf{q}_t)^\top \mathbf{q}_t = \alpha(\tau, \mathbf{q}_t)\mathbf{q}_t^\top \nabla_{\mathbf{q}} \mathbf{A} \mathbf{q}_t + \tilde{\boldsymbol{\Theta}}(\tau, \mathbf{q}_t)^\top \mathbf{q}_t.$$

Second, we assume that the demand and supply functions take the form of dirac delta functions:

$$\alpha(\tau, \mathbf{q}_t) = \delta(\tau - T)\alpha(\mathbf{q}_t),$$
$$\tilde{\boldsymbol{\Theta}}(\tau, \mathbf{q}_t) = \delta(\tau - T)\tilde{\boldsymbol{\Theta}}(\mathbf{q}_t),$$

and take the limit as $T \to \infty$. Although these stylized assumptions remove some of the richness of the model described above, we derive a significantly simpler characterization of equilibrium below.

Let $\mathbf{B} \equiv \mathbf{B}(\tau, \mathbf{q}_t) = \mathbf{A} + \nabla_{\mathbf{q}} \mathbf{A}^\top \mathbf{q}_t$ (which captures the entire reaction of log bond prices to changes in the risk factors). Define the scalar

$$z \equiv z(\mathbf{q}_t) = \mathbf{B}^\top \hat{\boldsymbol{\Sigma}} \mathbf{B}. \tag{5.23}$$

[14]Note that our baseline model results are derived under a representative arbitrageur. Because some of our empirical regression results rely on within-dealer price variation, we consider a model extension in Appendix Section B which can map directly to our regression results in (4.1).

Then with the above assumptions in place, equations (5.14)-(5.15) become

$$\left[\mathbf{\Gamma}^{\top} + \alpha z \mathbf{I}\right] \mathbf{B} = \mathbf{e}_r + z\tilde{\mathbf{\Theta}}. \tag{5.24}$$

Conditional on the value of the scalar z (and the set of risk factors $\mathbf{q}_t$, which enter through the demand and supply functions), this is a simple linear system of equations which characterizes $\mathbf{B}$. Thus, for a given level of the risk factors $\mathbf{q}_t$, we can solve the equilibrium of the model as the solution to a scalar fixed point problem using (5.23) and (5.24).

5.4 Numerical Example

To explore the implications of the model, we consider a simple two-factor version. We assume that the risk factors $\mathbf{q}_t$ consist of the short rate r_t and a supply factor S_t (thus, we abstract from habitat noise demand risk). We assume that these factors are independent (so $\mathbf{\Gamma}$ and $\boldsymbol{\sigma}$ in (5.4) are diagonal). We choose the mean reversion of short rate and supply shocks to be $\kappa_r = 0.5$ and $\kappa_s = 0.2$, respectively. Volatilities are set to $a \cdot \sigma_r^2 = 0.05$ and $a \cdot \sigma_s^2 = 0.1$ (note that for the purposes of studying how bond prices react, we do not need to separately calibrate physical risk and risk aversion).

We assume that the demand elasticity varies as a function of the supply factor according to the top panel of Figure 7. The elasticity function α varies from values of 2.0 (when supply S_t is low) to 0.2 (when supply S_t is high).

The second panel of Figure 7 shows how equilibrium arbitrageur holdings of bonds reacts to increases in supply. Regardless of the level of outstanding supply, when additional bonds are issued, arbitrageurs in equilibrium increase their holdings of bonds. Preferred habitat investors also increase holdings: an increase in bond supply leads to a decrease in bond prices, which incentivizes habitat investors to buy a portion of the increase in bonds. But the strength of this channel is governed by habitat elasticity. When the outstanding supply of bonds increases, habitat elasticities decrease and therefore arbitrageurs end up holding a larger fraction of the increase in issuance.

Thus, in states of the world where bonds are scarce, an increase in supply is absorbed by both arbitrageurs and preferred habitat investors. However, when we enter a regime of saturated bond markets, only arbitrageurs absorb additional increases in supply.

The final panel of Figure 7 shows what this mechanism implies for the equilibrium response of (log) bond prices to changes in the short rate factor (blue line) or the supply factor (red line) as a function of the supply outstanding. We find that, as expected,

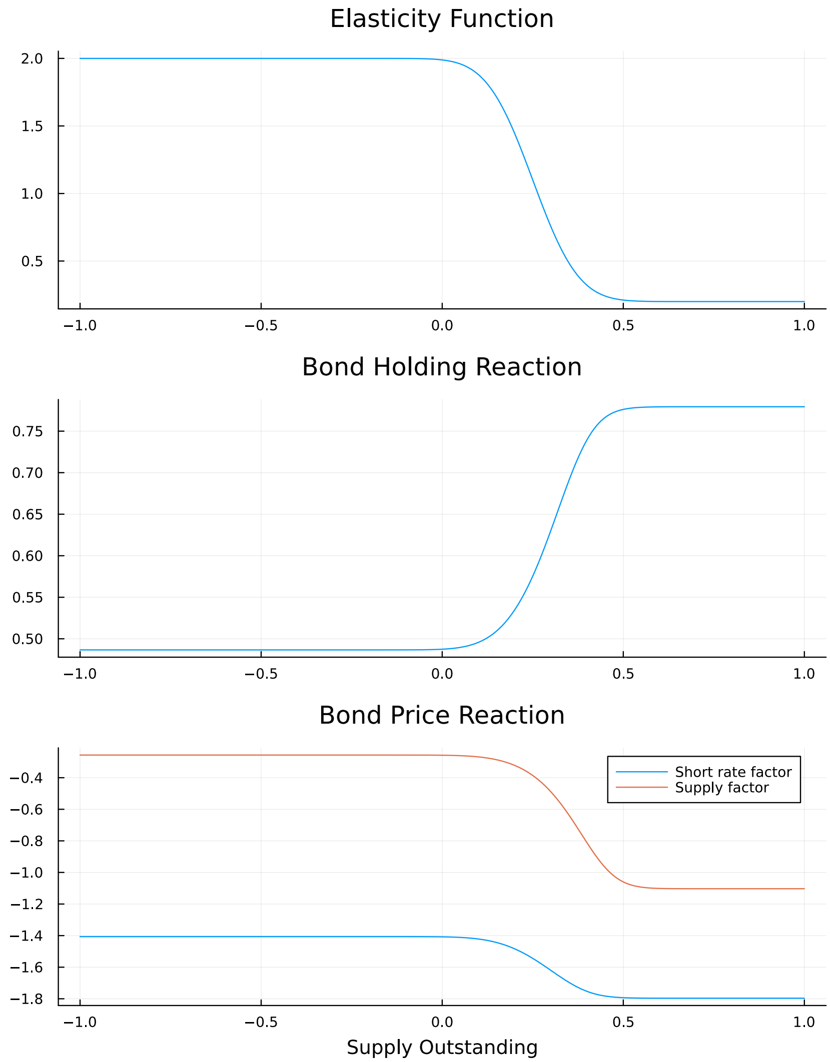


Figure 7: Two-factor model

Note: equilibrium model objects as a function of outstanding supply S_t (on the x-axis). The top panel shows the elasticity function used in the calibration. The middle panel shows how equilibrium arbitrageur bond holdings respond to supply shocks on impact. The bottom panel shows the on-impact equilibrium bond price reactions to short rate increases r_t (blue line) and supply increases S_t (red line).

increases in the short rate or supply decrease the price of bonds regardless of the total supply outstanding. However, as supply increases into the region in which the demand elasticity falls, the responsiveness of bond prices to these factors increases (in absolute magnitude). This change in responsiveness is much more salient for the supply factor: bond prices become more than twice as sensitive to changes in supply.

Consistent with our one-factor analytical results, the quantitative model both helps rationalize our key empirical findings regarding QE and QT asymmetries; and provides important implications for policymakers. In particular, our findings suggest a note of caution for central banks pursuing balance sheet normalization through quantitative tightening, when the preferred habitat investors play a key role in the market. Policymakers have stated that QT will be uneventful: like "watching paint dry." Figure 7 shows that the response to QT may be consistent and uneventful while outstanding bond supply remains low. However, once habitat investor demand is highly satiated, a decline in demand elasticities would lead to significantly larger price reactions to continued unwinding, implying potential disruptions in bond markets.

6 Conclusion

Using granular offer-level Bank of England auctions data and UK government bond transactions data, we find that increased trade volumes from preferred habitat investors to dealers during QE periods lead to stronger dealer bidding behavior at the auctions, with lower offer spreads to benchmark yields. During QT periods, preferred habitat investors have played a less significant role; instead, QT transmission has been more affected by a larger role for hedge fund activity. Therefore, our empirical results imply that the transmission mechanism of unconventional monetary policy is potentially asymmetric and state-dependent.

To explain the empirical findings, we build a theoretical model to show how the available supply of government debt could drive the results through the asymmetric and state-contingent role of demand for bonds. At the time of large debt supply, and hence relatively smaller role of buy-and-hold/preferred habitat investors, the bond price reactions to QT become more susceptible to the conditions of financial markets and the capacity of intermediaries to absorb the debt supply. Our results therefore suggest a note of caution for policymakers: QT is not a mechanical policy tool, and its effects warrant careful calibration, reflecting prevailing market conditions and the level of government debt.

References

Adrian, T., Boyarchenko, N., and Shachar, O. (2017). Dealer balance sheets and bond liquidity provision. *Journal of Monetary Economics*, 89(C):92–109.

Alexander, P., Fakhoury, R., Horn, T., Panjwani, W., and Roberts-Sklar, M. (2023). Financial stability buy/sell tools: a gilt market case study. *Bank of England Quarterly Bulletin*, 2023 Q4.

Bailey, A., Bridges, J., Harrison, R., Jones, J., and Mankodi, A. (2020). The central bank balance sheet as a policy tool: past, present and future. *Bank of England Staff Working Paper 899*.

Bandera, N. and Stevens, J. (2026). Monetary policy consequences of financial stability interventions: Assessing the uk gilt crisis and the central bank policy response. *European Economic Review*, 187:105360.

Boneva, L., Elliott, D., Kaminska, I., Linton, O., McLaren, N., and Morley, B. (2022). The impact of corporate QE on liquidity: evidence from the UK. *The Economic Journal*, 132(648):2615–2643.

Boneva, L., Kastl, J., and Zikes, F. (2025). Dealer balance sheets and bidding behavior in the bank of england's qe reverseauctions. *Journal of Financial Economics*, 174:104182.

Busetto, F., Chavaz, M., Froemel, M., Joyce, M., Kaminska, I., and Worlidge, J. (2022). QE at the Bank of England: a perspective on its functioning and effectiveness. *Bank of England Quarterly Bulletin*, 1.

Cantore, C. and Meichtry, P. (2024). Unwinding quantitative easing: State dependency and household heterogeneity. *European Economic Review*, 170:104865.

Christensen, J. H. and Rudebusch, G. D. (2012). The response of interest rates to US and UK quantitative easing. *The Economic Journal*, 122(564):F385–F414.

Du, W., Forbes, K., and Luzzetti, M. N. (2024). Quantitative tightening around the globe: What have we learned? Working Paper 32321, National Bureau of Economic Research.

Duffie, D. (2023). Resilience Redux in the U.S. Treasury Market. *Structural Shifts in the Global Economy, A Symposium Sponsored by Federal Reserve Bank of Kansas,*, Jackson Hole, Wyoming(August):77–119.

Duffie, D., Fleming, M., Keane, F., Nelson, C., Shachar, O., and Tassel, P. V. (2023). Dealer capacity and U.S. Treasury market functionality. *BIS Working Papers*, 1138.

D'Amico, S. and Seida, T. (2024). Unexpected Supply Effects of Quantitative Easing and Tightening. *The Economic Journal*, 134(658):579–613.

Gertler, M. and Karadi, P. (2013). QE 1 vs. 2 vs. 3. . . : A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool. *International Journal of Central Banking*, 9(1):5–53.

Giese, J., Joyce, M., Meaning, J., and Worlidge, J. (2024). Do preferred habitat investors exist? evidence from the uk government bondmarket. *Economics Letters*, 234:111462.

Gourinchas, P.-O., Ray, W., and Vayanos, D. (2025). A Two-Country New-Keynesian Model with Limited Arbitrage in Currency and Bond Markets. Technical report.

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

Haldane, A. G., Roberts-Sklar, M., Wieladek, T., and Young, C. (2016). QE: the story so far. *Bank of England Working Paper 624*.

Itskhoki, O. and Mukhin, D. (2021). Exchange rate disconnect in general equilibrium. *Journal of Political Economy*, 129(8):2183–2232.

Jansen, K., Klingler, S., Ranaldo, A., and Duijm, P. (2024a). Pension Liquidity Risk. Working Papers 801, DNB.

Jansen, K. A., Li, W., and Schmid, L. (2024b). Granular treasury demand with arbitrageurs. Technical report, National Bureau of Economic Research.

Joyce, M. and Lengyel, A. (2024). The yield curve impact of government debt issuance surprises and the implications for QT. *Bank of England Staff Working Paper No. 1,097*.

Joyce, M., Liu, Z., and Tonks, I. (2017). Institutional investors and the QE portfolio balance channel. *Journal of Money, Credit and Banking*, 49(6):1225–1246.

Kamdar, R. and Ray, W. (2025). Optimal Unconventional Policy in a New Keynesian Preferred Habitat Model. Technical report.

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

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

Kumhof, M. and Salgado-Moreno, M. (2024). Quantitative easing and quantitative tightening: the money channel. Bank of England working papers 1090, Bank of England.

Li, D., Petrasek, L., and Tian, M. H. (2025). Risk-averse dealers in a risk-free market - the role of trading desk risk limits. *FEDS Working Paper*, No. 2025-34.

McLaren, N., Banerjee, R. N., and Latto, D. (2014). Using changes in auction maturity sectors to help identify the impact of QE on gilt yields. *The Economic Journal*, 124(576):453–479.

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

Ray, W., Droste, M., and Gorodnichenko, Y. (2024). Unbundling quantitative easing: Taking a cue from treasury auctions. *Journal of Political Economy*, 132(9):3115–3172.

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.

Vayanos, D. and Vila, J.-L. (2009). A preferred-habitat model of the term structure of interest rates. *NBER Working Paper No. 15487.*

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

A Proofs

Proof of Proposition 1. As $S_t \to \pm\infty$, by assumption the elasticity function converges to the constants $\bar{\alpha}(\tau), \underline{\alpha}(\tau)$, respectively. This implies that the partial derivatives of the affine coefficients with respect to $\mathbf{q}_t$ approach zero; and thus the risk-adjusted dynamics matrix from (5.15) is independent of $\mathbf{q}_t$ and converges to the time-invariant matrices $\bar{\mathbf{M}}, \underline{\mathbf{M}}$, respectively. Letting $\mathbf{M} \in \{\bar{\mathbf{M}}, \underline{\mathbf{M}}\}$, and writing the demand elasticity function as $\hat{\alpha}\alpha(\tau)$ for a scalar $\hat{\alpha}$, we can therefore characterize the affine coefficients in either limiting case according to the following system of ODEs:

$$\begin{aligned}
\mathbf{0} &= \mathbf{A}'(\tau) + \mathbf{M}\mathbf{A}(\tau) - \mathbf{e}_r, \\
\mathbf{M} &\equiv \mathbf{\Gamma}^\top - \int_0^T \left[-\hat{\alpha}\alpha(\tau)\mathbf{A}(\tau) + \mathbf{\Theta}(\tau)\right] \mathbf{A}(\tau)^\top \, \mathrm{d}\tau \, \hat{\mathbf{\Sigma}},
\end{aligned}$$

and where in the single risk factor version of the model, we have

$$\mathbf{A}(\tau) = \begin{bmatrix} A_q(\tau) \\ A_\beta(\tau) \end{bmatrix}, \quad \mathbf{\Gamma} = \begin{bmatrix} \kappa_q & 0 \\ 0 & \kappa_\beta \end{bmatrix}, \quad \mathbf{e}_r = \begin{bmatrix} e_r \\ 0 \end{bmatrix}, \quad \mathbf{\Theta}(\tau) = \begin{bmatrix} -\gamma(\tau) \\ \delta(\tau) \end{bmatrix}, \quad \hat{\mathbf{\Sigma}} = \begin{bmatrix} a\sigma_q^2 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since β_t is deterministic in this model, we can separately solve for $A_q(\tau)$ and then for $A_\beta(\tau)$ using the same techniques as Vayanos and Vila (2021). This gives the following fixed point problem that characterizes the log-price coefficient functions:

$$\begin{aligned}
\mathbf{M} &= \begin{bmatrix} \kappa_q^* & 0 \\ \kappa_{q,\beta}^* & \kappa_\beta \end{bmatrix} \\
\kappa_q^* &= \kappa_q + a\sigma_q^2 \int_0^T \left(\hat{\alpha}\alpha(\tau)A_q(\tau) + \gamma(\tau)\right) A_q(\tau) \, \mathrm{d}\tau \\
\kappa_{q,\beta}^* &= a\sigma_q^2 \int_0^T \left(\hat{\alpha}\alpha(\tau)A_\beta(\tau) - \delta(\tau)\right) A_q(\tau) \, \mathrm{d}\tau \\
A_q(\tau) &= \frac{1 - e^{-\kappa_q^*\tau}}{\kappa_q^*} e_r \\
A_\beta(\tau) &= \frac{-\kappa_{q,\beta}^*}{\kappa_\beta - \kappa_q^*} \left(\frac{1 - e^{-\kappa_q^*\tau}}{\kappa_q^*} - \frac{1 - e^{-\kappa_\beta\tau}}{\kappa_\beta} \right) e_r
\end{aligned}$$

First, we have that $\kappa_q^* > 0$ and so that $A_q(\tau) > 0$: this follows immediately from the assumptions positivity assumptions in the statement of the proposition and the properties of the function $\frac{1-e^{-x}}{x}$.

Next, from the properties of the function $\frac{1}{y-x}\left(\frac{1-e^{-x}}{x}-\frac{1-e^{-y}}{y}\right)$, the sign of $A_\beta(\tau)$ is the opposite of $\kappa^*_{q,\beta}$ for $\tau > 0$. Then towards contradiction, suppose that $\kappa^*_{q,\beta} \geq 0$. But then all the terms in the integrand which defines $\kappa^*_{q,\beta}$ are weakly positive, and strictly positive for some $\tau > 0$. Thus, we must have $\kappa^*_{q,\beta} < 0$, which implies $A_\beta(\tau) > 0$ for all $\tau > 0$.

These inequalities hold strictly as $\hat{\alpha} \to 0$, since

$$\lim_{\hat{\alpha}\to 0} \kappa^*_q = \kappa_q + a\sigma_q^2 \int_0^T \gamma(\tau) A_q(\tau)\,\mathrm{d}\tau > 0$$
$$\lim_{\hat{\alpha}\to 0} \kappa^*_{q,\beta} = -a\sigma_q^2 \int_0^T \delta(\tau) A_q(\tau)\,\mathrm{d}\tau < 0$$

We now characterizing the limits of these objects as $\hat{\alpha} \to 0$ and $\hat{\alpha} \to \infty$. Let

$$\widehat{\kappa^*_q} = \hat{\alpha}^{-1/3}\kappa^*_q$$
$$\widehat{\kappa^*_{q,\beta}} = \hat{\alpha}^{2/3}\kappa^*_{q,\beta}$$
$$\widehat{A_q(\tau)} = \hat{\alpha}^{1/3} A_q(\tau) = \frac{1-e^{-\kappa^*_q \tau}}{\widehat{\kappa^*_q}} e_r$$
$$\widehat{A_\beta(\tau)} = \hat{\alpha} A_\beta(\tau) = \frac{-\widehat{\kappa^*_{q,\beta}}}{\hat{\alpha}^{-1/3}\kappa_\beta - \widehat{\kappa^*_q}} \left(\frac{1-e^{-\kappa^*_q\tau}}{\kappa^*_q} - \frac{1-e^{-\kappa_\beta \tau}}{\kappa_\beta}\right) e_r$$

We can re-write the fixed point problem for κ^*_q and $\kappa^*_{q,\beta}$:

$$\widehat{\kappa^*_q} = \hat{\alpha}^{-1/3}\kappa_q + a\sigma_q^2 \int_0^T \left(\alpha(\tau)\widehat{A_q(\tau)} + \hat{\alpha}^{-2/3}\gamma(\tau)\right)\widehat{A_q(\tau)}\,\mathrm{d}\tau$$
$$\hat{\alpha}^{-1/3}\widehat{\kappa^*_{q,\beta}} = a\sigma_q^2 \int_0^T \left(\alpha(\tau)\widehat{A_\beta(\tau)} - \delta(\tau)\right)\widehat{A_q(\tau)}\,\mathrm{d}\tau$$

We first conjecture that the objects $\widehat{\kappa^*_q}, \widehat{\kappa^*_{q,\beta}}, \widehat{A_q(\tau)}, \widehat{A_\beta(\tau)}$ converge to non-zero, finite limits as $\hat{\alpha} \to \infty$. Under this conjecture, we have

$$\widehat{A_q(\tau)} \to \lim_{\hat{\alpha}\to\infty} \frac{1}{\widehat{\kappa^*_q}} e_r$$
$$\widehat{A_\beta(\tau)} \to \lim_{\hat{\alpha}\to\infty} \frac{-\widehat{\kappa^*_{q,\beta}}}{\widehat{\kappa^*_q}} \left(\frac{1-e^{-\kappa_\beta\tau}}{\kappa_\beta}\right) e_r$$

Then we also have

$$\lim_{\hat{\alpha}\to\infty} \widehat{\kappa_q^*} \equiv \ell_q^* = \left(a\sigma_q^2 e_r^2 \int_0^T \alpha(\tau)\,\mathrm{d}\tau \right)^{1/3}$$

which follows from the dominated convergence theorem and algebraic manipulation of the above expression for $\widehat{\kappa_q^*}$.

Next, we have

$$\hat{\alpha}^{-1/3}\widehat{\kappa_{q,\beta}^*} \to 0 = \lim_{\hat{\alpha}\to\infty} \int_0^T \left(\alpha(\tau)\widehat{A_\beta(\tau)} - \delta(\tau) \right) \widehat{A_q(\tau)}\,\mathrm{d}\tau$$

Substituting the above expressions and solving for the limit of $\widehat{\kappa_{q,\beta}^*}$ gives

$$\lim_{\hat{\alpha}\to\infty} \widehat{\kappa_{q,\beta}^*} = -\frac{\int_0^T \delta(\tau)\,\mathrm{d}\tau}{\int_0^T \alpha(\tau) \left(\frac{1-e^{-\kappa_\beta \tau}}{\kappa_\beta} \right) \mathrm{d}\tau \, \frac{e_r}{\ell_q^*}}$$

Hence, the conjecture is internally consistent.

To show that these limits exist and are unique, take any sequence $\hat{\alpha}_n \to \infty$ as $n \to \infty$. Note that $\widehat{\kappa_{q_n}^*} > 0$ for all n, and from the dominated convergence theorem:

$$\begin{aligned} I_1 &= \int_0^T \alpha(\tau) \left(1 - e^{-\kappa_q^* \tau}\right)^2 \mathrm{d}\tau \to \int_0^T \alpha(\tau)\,\mathrm{d}\tau \equiv C_1 > 0 \\ I_2 &= \int_0^T \gamma(\tau) \left(1 - e^{-\kappa_q^* \tau}\right) \mathrm{d}\tau \to \int_0^T \gamma(\tau)\,\mathrm{d}\tau \equiv C_2 > 0 \end{aligned}$$

First, towards contradiction, suppose $\liminf_{n\to\infty} \widehat{\kappa_{q_n}^*} = 0$, then take a subsequence such that $\widehat{\kappa_{q_m}^*} \to 0$. But then

$$\begin{aligned} \widehat{\kappa_{q_m}^*}^3 &\to 0 \\ \hat{\alpha}_m^{-1/3} \kappa_q \widehat{\kappa_{q_m}^*}^2 &\to 0 \\ a\sigma_q^2 e_r^2 I_1 &\to a\sigma_q^2 e_r^2 C_1 > 0 \\ \hat{\alpha}_m^{-2/3} a\sigma_q^2 e_r I_2 \widehat{\kappa_{q_m}^*} &\to 0 \end{aligned}$$

and thus we reach a contradiction, since these limits imply $\widehat{\kappa_{q_m}^*} \to a\sigma_q^2 e_r^2 C_1 > 0$.

Second, towards contradiction, if $\limsup_{n\to\infty} \widehat{\kappa_{q_n}^*} = \infty$, then take a subsequence such that $\widehat{\kappa^*_{q_m}} \to \infty$. But then

$$\begin{aligned}\hat{\alpha}_m^{-1/3}\kappa_q\widehat{\kappa^*_{q_m}}^{-1} &\to 0\\ a\sigma_q^2 e_r^2 I_1\widehat{\kappa^*_{q_m}}^{-3} &\to 0\\ \hat{\alpha}_m^{-2/3}a\sigma_q^2 I_2\widehat{\kappa^*_{q_m}}^{-2} &\to 0\end{aligned}$$

and we again reach a contradiction, since these limits imply $\widehat{\kappa^*_{q_m}} \to 0$.

Finally, if $\liminf_{n\to\infty}\widehat{\kappa^*_{q_n}} = \ell_1$ and $\limsup_{n\to\infty}\widehat{\kappa^*_{q_n}} = \ell_2$ with $0 < \ell_1 < \infty, 0 < \ell_2 < \infty$, then the same algebraic argument above shows that $\ell_1 = \ell_2 = \ell$. Thus, the limit ℓ exists and is unique. Similar arguments also show existence and uniqueness of the limit for $\widehat{\kappa^*_{q\beta}}$.

Thus, if $\underline{\alpha}(\tau)$ is sufficiently small $\bar{\alpha}(\tau)$ is sufficiently large, the claim follows. That is, there exists a constant $K > 0$ such that

$$\exists K > 0 : \quad \|\underline{\alpha}\| < K, \quad \|\bar{\alpha}\|^{-1} < K \tag{A1}$$

for which the results hold.

The final claim of the proposition follows since

$$\lim_{\hat{\alpha}\to\infty} -\left(\hat{\alpha}\alpha(\tau)\widehat{A_\beta(\tau)} - \delta(\tau)\right) < \delta(\tau)$$

□

Lemma 1 (Quadratic approximation results)**.** *Take a general OU process for the state dynamics*

$$\mathrm{d}q_t = -\Gamma q_t\,\mathrm{d}t + \xi\cdot\hat{\sigma}\,\mathrm{d}B_t$$

where the scalar ξ is a perturbation parameter. Then the log-pricing equation satisfies

$$\begin{aligned}-\log P_t(\tau) &= A_1(\tau)^\top q_t + q_t^\top A_2(\tau)q_t + C(\tau) + O(\xi^3)\\ A_1'(\tau) + MA_1(\tau) - e_r &= 0\\ A_2'(\tau) + A_2(\tau)M^\top + MA_2(\tau) &= N(\tau)\end{aligned}$$

where $N(\tau)$ and M are defined in (A2) and (A3).

Proof of Lemma 1. Applying Ito's Lemma to $P_t(\tau)$, we find

$$
\begin{aligned}
\mu_t(\tau) &= A_1'(\tau)^\top q_t + q_t^\top A_2'(\tau) q_t + O(\xi^3) \\
&\quad + \left[A_1(\tau)^\top + 2q_t^\top A_2(\tau)\right] \Gamma q_t + O(\xi^3) \\
&\quad + C'(\tau) + \xi^2 \frac{1}{2} \operatorname{Tr}\left\{\sigma^\top \left(A_1(\tau)A_1(\tau)^\top - 2A_2(\tau)\right)\sigma\right\} + O(\xi^3) \\
\sigma_t(\tau) &= -\xi \cdot \left[A_1(\tau)^\top + 2q_t^\top A_2(\tau) + O(\xi^2)\right] \sigma
\end{aligned}
$$

Note that for the drift term $\mu_t(\tau)$, we keep terms up to second-order, while for the diffusion vector $\sigma_t(\tau)$, we keep only first-order terms. As we will see in the later part of the proof, this is sufficient for characterizing equilibrium.

Habitat demand and supply can be written

$$
\begin{aligned}
Z_t(\tau) &= -\left(\alpha(\tau) + \alpha_1(\tau)^\top q_t + O(\xi^2)\right) \cdot \log P_t(\tau) - \Theta_{z,1}(\tau)^\top q_t - \zeta(\tau) + O(\xi^3) \\
S_t(\tau) &= \Theta_{s,1}(\tau)^\top q_t
\end{aligned}
$$

where $\zeta(\tau) = -\alpha(\tau)C(\tau)$. Hence market clearing $S_t(\tau) = X_t(\tau) + Z_t(\tau)$ implies

$$
X_t(\tau) = -\left(\alpha(\tau) + \alpha_1(\tau)^\top q_t\right) \cdot \left[A_1(\tau)^\top q_t + q_t^\top A_2(\tau) q_t\right] + \Theta_1(\tau)^\top q_t + O(\xi^3)
$$

where $\Theta_1(\tau) = \Theta_{z,1}(\tau) + \Theta_{s,1}(\tau)$. Note that when $q_t = 0, X_t(\tau) = 0$. Define $\tilde{\Sigma} \equiv a\sigma\sigma^\top = \hat{a}\hat{\sigma}\hat{\sigma}^\top$. From the quadratic conjecture for $\log P_t(\tau)$, the first-order and second-order terms of the following expression are:

$$
\begin{aligned}
a \operatorname{Cov}_t\left(\frac{\mathrm{d}P_t(\tau)}{P_t(\tau)}, \frac{\mathrm{d}P_t(\tau')}{P_t(\tau')}\right) X_t(\tau') &= -\left[A_1(\tau)^\top \tilde{\Sigma} A_1(\tau')\right] \left[\alpha(\tau')A_1(\tau') - \Theta_1(\tau')\right]^\top q_t \\
&\quad - q_t^\top \left[A_1(\tau)^\top \tilde{\Sigma} A_1(\tau')\right] \cdot \left[\alpha(\tau') \cdot A_2(\tau') + \alpha_1(\tau') A_1(\tau')^\top\right] q_t \\
&\quad - 2q_t^\top \left[A_2(\tau')\tilde{\Sigma}A_1(\tau) + A_2(\tau)\tilde{\Sigma}A_1(\tau')\right] \left[\alpha(\tau')A_1(\tau') - \Theta_1(\tau')\right]^\top q_t + O(\xi^3)
\end{aligned}
$$

Note that because $\overline{X(\tau)} = 0$ in steady state, there are no zero-order terms.

Integrating:

$$
\int_0^T a \operatorname{Cov}_t\left(\frac{\mathrm{d}P_t(\tau)}{P_t(\tau)}, \frac{\mathrm{d}P_t(\tau')}{P_t(\tau')}\right) X_t(\tau') \, \mathrm{d}\tau' = A_1(\tau)^\top L_1 q_t + q_t^\top L_2(\tau) q_t + O(\xi^3)
$$

where

$$
\begin{aligned}
L_1 &= -\int_0^T \tilde{\Sigma} A_1(\tau') \left[\alpha(\tau') \cdot A_1(\tau') - \Theta_1(\tau')\right]^\top \mathrm{d}\tau' \\
L_2(\tau) &= -\int_0^T \left[A_1(\tau)^\top \tilde{\Sigma} A_1(\tau')\right] \cdot \left[\alpha(\tau') \cdot A_2(\tau') + \alpha_1(\tau') A_1(\tau')^\top\right] \mathrm{d}\tau' \\
&\quad - 2\int_0^T \left[A_2(\tau')\tilde{\Sigma} A_1(\tau) + A_2(\tau)\tilde{\Sigma} A_1(\tau')\right] \left[\alpha(\tau') A_1(\tau') - \Theta_1(\tau')\right]^\top \mathrm{d}\tau'
\end{aligned}
$$

or written in symmetric form,

$$
\frac{1}{2}\left(L_2(\tau) + L_2(\tau)^\top\right) \equiv N(\tau) + A_2(\tau) L_1 + L_1^\top A_2(\tau)
$$

where

$$
N(\tau) = \frac{1}{2}\left(N_1(\tau) + N_1(\tau)^\top\right) + \frac{1}{2}\left(N_2(\tau) + N_2(\tau)^\top\right) \tag{A2}
$$

and where

$$
\begin{aligned}
N_1(\tau) &= -\int_0^T \left[A_1(\tau)^\top \tilde{\Sigma} A_1(\tau')\right] \cdot \left[\alpha(\tau') \cdot A_2(\tau') + \alpha_1(\tau') A_1(\tau')^\top\right] \mathrm{d}\tau' \\
N_2(\tau) &= -2\int_0^T \left[A_2(\tau')\tilde{\Sigma} A_1(\tau)\right] \left[\alpha(\tau') A_1(\tau') - \Theta_1(\tau')\right]^\top \mathrm{d}\tau'
\end{aligned}
$$

The arbitrageur optimality conditions are still given by

$$
\begin{aligned}
\mu_t(\tau) - e_r^\top q_t &= a\sigma_t(\tau) \int \sigma_t(\tau')^\top X_t(\tau') \,\mathrm{d}\tau' \\
&= a \int \mathrm{Cov}_t\left(\frac{\mathrm{d}P_t(\tau)}{P_t(\tau)}, \frac{\mathrm{d}P_t(\tau')}{P_t(\tau')}\right) X_t(\tau') \,\mathrm{d}\tau'
\end{aligned}
$$

Equating the zero-order terms gives

$$
C'(\tau) + \xi^2 \frac{1}{2} \operatorname{Tr}\left\{\sigma^\top \left(A_1(\tau) A_1(\tau)^\top - 2A_2(\tau)\right)\sigma\right\} = 0
$$

Equating the first-order terms gives

$$
\begin{aligned}
& A_1'(\tau) + \Gamma^\top A_1(\tau) - e_r = L_1^\top A_1(\tau) \\
\iff\ & A_1'(\tau) + M A_1(\tau) - e_r = 0
\end{aligned}
$$

where

$$M \equiv (\Gamma - L_1)^\top \tag{A3}$$

Equating the second-order terms gives

$$\begin{aligned} A_2'(\tau) + A_2(\tau)\Gamma + \Gamma^\top A_2(\tau) &= \frac{1}{2}\left(L_2(\tau) + L_2(\tau)^\top\right) \\ \iff A_2'(\tau) + A_2(\tau)M^\top + MA_2(\tau) &= N(\tau) \end{aligned}$$

which completes the proof.

□

Proof of Proposition 2. We apply the results of Lemma 1, where now

$$\Gamma = \begin{bmatrix} \kappa_q & 0 \\ 0 & \kappa_\beta \end{bmatrix}, \quad \tilde{\Sigma} = \begin{bmatrix} \hat{\sigma}_q^2 & 0 \\ 0 & 0 \end{bmatrix}, \quad e_r = \begin{bmatrix} e_r \\ 0 \end{bmatrix}, \quad \Theta_1(\tau) = \begin{bmatrix} -\gamma(\tau) \\ \delta(\tau) \end{bmatrix}, \quad \alpha_1(\tau) = \begin{bmatrix} \alpha_{1q}(\tau) \\ 0 \end{bmatrix}$$

and the endogenous coefficients

$$A_1(\tau) = \begin{bmatrix} A_{1q}(\tau) \\ A_{1\beta}(\tau) \end{bmatrix}, \quad A_2(\tau) = \begin{bmatrix} A_{2q}(\tau) & A_{2q\beta}(\tau) \\ A_{2q\beta}(\tau) & A_{2\beta}(\tau) \end{bmatrix}$$

Under this simplifying case, we have

$$M \equiv \begin{bmatrix} M_{11} & 0 \\ M_{21} & \kappa_\beta \end{bmatrix}$$

where

$$\begin{aligned} M_{11} &= \kappa_q + \hat{\sigma}_q^2 \int_0^T \left[\alpha(\tau)A_{1q}(\tau) + \gamma(\tau)\right] A_{1q}(\tau)\, \mathrm{d}\tau \\ M_{21} &= \hat{\sigma}_q^2 \int_0^T \left[\alpha(\tau)A_{1\beta}(\tau) - \delta(\tau)\right] A_{1q}(\tau)\, \mathrm{d}\tau \end{aligned}$$

and

$$N(\tau) = A_{1q}(\tau) \cdot N \equiv A_{1q}(\tau) \cdot \begin{bmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{bmatrix}$$

where

$$N_{11} = -\hat{\sigma}_q^2 \int_0^T \Big\{ A_{1q}(\tau)\left[\alpha(\tau)A_{2q}(\tau) + \alpha_{1q}(\tau)A_{1q}(\tau)\right] + 2A_{2q}(\tau)\left[(\alpha(\tau)A_{1q}(\tau) + \gamma(\tau))\right] \Big\}\,\mathrm{d}\tau$$

$$N_{12} = -\hat{\sigma}_q^2 \int_0^T \Big\{ A_{1q}(\tau)\left[\alpha(\tau)A_{2q\beta}(\tau) + \frac{1}{2}\alpha_{1q}(\tau)A_{1\beta}(\tau)\right] + \left[(\alpha(\tau)A_{1\beta}(\tau) - \delta(\tau))\,A_{2q}(\tau) + (\alpha(\tau)A_{1q}(\tau) + \gamma(\tau))\,A_{2q\beta}(\tau)\right] \Big\}\,\mathrm{d}\tau$$

$$N_{22} = -\hat{\sigma}_q^2 \int_0^T \Big\{ A_{1q}(\tau)\left[\alpha(\tau)A_{2\beta}(\tau)\right] + 2A_{2q\beta}(\tau)\left[\alpha(\tau)A_{1\beta}(\tau) - \delta(\tau)\right] \Big\}\,\mathrm{d}\tau$$

Taking as given M, and using the initial conditions $A_{1q}(0) = 0, A_{1\beta}(0) = 0$, we have

$$A_{1q}(\tau) = \frac{1 - e^{-M_{11}\tau}}{M_{11}} e_r$$

$$A_{1\beta}(\tau) = \frac{-M_{21}}{\kappa_\beta - M_{11}} \left(\frac{1 - e^{-M_{11}\tau}}{M_{11}} - \frac{1 - e^{-\kappa_\beta \tau}}{\kappa_\beta} \right) e_r$$

Note that the same arguments as in the proof of Prop. 1 show that $M_{11} > 0$ and $M_{21} < 0$, and so $A_{1q}(\tau) > 0$ and $A_{1\beta}(\tau) > 0$ for all $\tau > 0$.

We now turn to the second-order unknown coefficient functions with initial conditions $A_{2q}(0) = 0, A_{2q\beta}(0) = 0, A_{2\beta}(\tau) = 0$. First, note that

$$A'_{2q}(\tau) + 2M_{11}A_{2q}(\tau) = A_{1q}(\tau)N_{11}$$

Taking as given M and N, the solution for the second-order term $A_{2q}(\tau)$ is

$$A_{2q}(\tau) = N_{11}\frac{A_{1q}(\tau)^2}{2e_r}$$

and using this solution, we can solve for N_{11}:

$$N_{11} = \frac{-\hat{\sigma}_q^2 \int_0^T \alpha_{1q}(\tau)A_{1q}(\tau)^2\,\mathrm{d}\tau}{1 + \hat{\sigma}_q^2 \int_0^T \frac{A_{1q}(\tau)^2}{2e_r}\left[A_{1q}(\tau)\alpha(\tau) + 2\left[(\alpha(\tau)A_{1q}(\tau) + \gamma(\tau))\right]\right]\mathrm{d}\tau}$$

which is negative given that $\alpha_{1q}(\tau) \geq 0$ for all τ. Thus $A_{2q}(\tau) < 0$ for all $\tau > 0$.

Next, we have

$$A'_{2q\beta}(\tau) + (M_{11} + \kappa_\beta)A_{2q\beta}(\tau) = A_{1q}(\tau)N_{12} - A_{2q}(\tau)M_{21}$$

and the solution for the second-order term $A_{2q\beta}(\tau)$ is

$$\begin{aligned} A_{2q\beta}(\tau) = & \frac{N_{12}e_r}{M_{11}}\left[\frac{1-e^{-\lambda\tau}}{\lambda} - \frac{e^{-M_{11}\tau} - e^{-\lambda\tau}}{\kappa_\beta}\right] \\ & - \frac{M_{21}N_{11}e_r}{2M_{11}^2}\left[\frac{1-e^{-\lambda\tau}}{\lambda} - \frac{2\left(e^{-M_{11}\tau} - e^{-\lambda\tau}\right)}{\kappa_\beta} + \frac{e^{-2M_{11}\tau} - e^{-\lambda\tau}}{\kappa_\beta - M_{11}}\right] \\ \equiv & N_{12}F(\tau) - G(\tau) \end{aligned}$$

with $\lambda = M_{11} + \kappa_\beta$. Note that $F(\tau) > 0, G(\tau) > 0$ since for all $\tau > 0$, $A_{1q}(\tau) > 0$, $A_{2q}(\tau) < 0$, and $M_{21} < 0$.

Using this solution, we can solve for N_{12}:

$$N_{12} = \frac{\hat{\sigma}_q^2 \int_0^T G(\tau)\left[2\alpha(\tau)A_{1q}(\tau) + \gamma(\tau)\right] - \left(\alpha(\tau)A_{1\beta}(\tau) - \delta(\tau)\right)A_{2q}(\tau) - \frac{1}{2}\alpha_{1q}(\tau)A_{1\beta}(\tau)A_{1q}(\tau)\,\mathrm{d}\tau}{1 + \hat{\sigma}_q^2 \int_0^T F(\tau)\left[2\alpha(\tau)A_{1q}(\tau) + \gamma(\tau)\right]\mathrm{d}\tau}$$

Note that the denominator is greater than 1 for any parametrization. Moreover, if $\alpha(\tau)$ and $\gamma(\tau)$ are sufficiently small, then the numerator is negative. That is, we have the following sufficient condition for $N_{12} < 0$:

$$\exists K > 0 : \quad \|\alpha\| < K, \quad \|\gamma\| < K \implies N_{12} < 0 \tag{A4}$$

and $N_{12} < 0$ is itself a sufficient condition for $A_{2q\beta}(\tau) < 0$ for all $\tau > 0$.

Finally, note that

$$\frac{\partial X_t(\tau)}{\partial q_t \beta_t} = -2\alpha(\tau)A_{2q\beta}(\tau) - \alpha_1(\tau)A_{1\beta}(\tau)$$

which is negative under the same sufficiency condition $\|\alpha\| < K$.

□

B Model Extension

This appendix provides an extension of the model in which arbitrageurs and habitat investors maintain relationships so that the law of one price may not hold. The purpose is to map the model more directly to the empirical regression specification equation (4.1) which relies on within-dealer price variation at each auction.

Extension setup. The setup of the model is otherwise identical to the baseline except that we track the arbitrageur $i \in \mathcal{I}$ and maturity $\tau \in [0, T]$ prices and quantities, where $\mathcal{I}$ is the set of arbitrageurs which we normalize to have unit measure. The arbitrageur problem is now:

$$\max_{\{X_{it}^{(\tau)}\}} E_t \,\mathrm{d}W_{it} - \frac{a}{2} Var_t \,\mathrm{d}W_{it} \tag{B1}$$

$$\text{subject to: } \mathrm{d}W_{it} = W_{it} r_t \,\mathrm{d}t + \int_0^T X_{it}^{(\tau)} \left[\frac{\mathrm{d}P_{it}^{(\tau)}}{P_{it}^{(\tau)}} - r_t \,\mathrm{d}t \right] \mathrm{d}\tau \,, \tag{B2}$$

and preferred-habitat demand for τ investors with a relationship with arbitrageur i is

$$Z_{it}^{(\tau)} = \alpha_t(\tau) \left[y_{it}^{(\tau)} - \beta_{it}^{(\tau)} \right] . \tag{B3}$$

Note that the term $\beta_{it}^{(\tau)}$ now also captures additional sources of within-dealer/habitat variation in noise demand. We assume that the dynamics of aggregate and idiosyncratic risk factors can be summarized in a vector $\mathbf{q}_{it}$ which follows the same class of vector OU processes considered in the baseline model.

Auction frictions. If there is frictionless trade among perfectly competitive arbitrageurs, then the law of one price will be enforced in equilibrium: $P_{it}^{(\tau)} = P_t^{(\tau)}$ for all arbitrageurs $i \in \mathcal{I}$, and the model is identical to the baseline. We instead assume that following a shock to the supply of bonds $\Delta S_t^{(\tau)}$ (from either the fiscal authority or the central bank conducting QE/QT), each arbitrageur purchases (if $\Delta S_t^{(\tau)} > 0$) or sells (if $\Delta S_t^{(\tau)} < 0$) an equal fraction of the change in supply. This simplifying assumption allows us to model within-dealer variation in prices with minimal departures from our baseline model. In particular, we are able to derive simple expressions for the model-implied regression coefficient in equation (4.1) and examine the qualitative changes as a function of habitat demand elasticities.

The purchases or sales can either be conducted via proprietary trades or on behalf of the habitat investor clients for each arbitrageur i. The market clearing conditions are thus:

$$\Delta X_{it}^{(\tau)} + \Delta Z_{it}^{(\tau)} = \Delta S_t^{(\tau)}$$
$$\int_{i\in\mathcal{I}} \left(\Delta X_{it}^{(\tau)} + \Delta Z_{it}^{(\tau)}\right) \mathrm{d}i = \Delta S_t^{(\tau)}.$$

Equilibrium. The baseline model explores the consequences of explicit time-variation in demand elasticities $\alpha_t(\tau)$; for this extension, we simplify and assume that demand elasticities are fixed over time but study the comparative static consequences of high and low demand elasticities for the regressions considered in Section 4.2. With fixed demand elasticities, the equilibrium asset prices satisfy

$$-\log P_{it}^{(\tau)} = \mathbf{A}(\tau)\top \mathbf{q}_{it} + C(\tau).$$

Note that the affine coefficients are fixed over time, and the within-dealer variation is summarized by the variation in risk factors $\mathbf{q}_{it}$. The arbitrageur FOCs imply

$$\mu_{it}^{(\tau)} - r_t = a\mathbf{A}(\tau)^\top \mathbf{\Lambda}_{it},$$
$$\mathbf{\Lambda}_{it}^\top \equiv \int_0^T X_{it}^{(\tau)} \mathbf{A}(\tau)\, \mathrm{d}\tau\,.$$

Regression coefficient. The model-implied regression coefficient is

$$\hat{\beta}^{auc} \equiv \frac{Cov(\Delta \hat{y}_{it}^{(\tau)}, \Delta Z_{it}^{(\tau)} - \Delta S_t^{(\tau)})}{Var(\Delta Z_{it}^{(\tau)} - \Delta S_t^{(\tau)})},$$

where the notation $\hat{y}_{it}^{(\tau)}$ denotes yield deviations from aggregate averages at time t. That is,

$$\Delta \hat{y}_{it}^{(\tau)} \equiv \Delta y_{it}^{(\tau)} - \int_{i\in\mathcal{I}} \Delta y_{it}^{(\tau)}\, \mathrm{d}i\,.$$

For simplicity, we assume a single factor structure for the idiosyncratic noise demand and supply: $\beta_{it}^{(\tau)} = \theta(\tau)\beta_{it}$ and $S_t^{(\tau)} = \gamma(\tau)s_t$, where $\beta_{it}, s_t \in \mathbf{q}_{it}$. This assumption is not necessary for our main results but simplifies computing the model-implied regression coefficient. Combined with affine asset prices, we have

$$\begin{aligned}
\Delta y_{it}^{(\tau)} &= \frac{1}{\tau}\mathbf{A}(\tau)^{\top}\Delta\mathbf{q}_{it} \\
&= \frac{1}{\tau}A_{\beta}(\tau)\Delta\beta_{it} + \frac{1}{\tau}A_{s}(\tau)\Delta s_{t}, \\
\Delta Z_{it}^{(\tau)} &= \alpha(\tau)\left[\Delta y_{it}^{(\tau)} - \Delta\beta_{it}^{(\tau)}\right] \\
&= \alpha(\tau)\left[\left(\frac{1}{\tau}A_{\beta}(\tau) - \theta(\tau)\right)\Delta\beta_{it} + \frac{1}{\tau}A_{s}(\tau)\Delta s_{t}\right],
\end{aligned}$$

and therefore

$$\begin{aligned}
\Delta\hat{y}_{it}^{(\tau)} &= \frac{1}{\tau}A_{\beta}(\tau)\Delta\beta_{it}, \\
\Delta Z_{it}^{(\tau)} - \Delta S_{t}^{(\tau)} &= \alpha(\tau)\left(\frac{1}{\tau}A_{\beta}(\tau) - \theta(\tau)\right)\Delta\beta_{it} + \left(\alpha(\tau)\frac{1}{\tau}A_{s}(\tau) - \gamma(\tau)\right)\Delta s_{t}.
\end{aligned}$$

So long as the demand shocks are idiosyncratic, then $Cov(\Delta\beta_{it}, \Delta s_{t}) = 0$ and the model-implied variance-covariances are given by

$$\begin{aligned}
Cov(\Delta\hat{y}_{it}^{(\tau)}, \Delta Z_{it}^{(\tau)} - \Delta S_{it}^{(\tau)}) &= \frac{1}{\tau}A_{\beta}(\tau)\alpha(\tau)\left(\frac{1}{\tau}A_{\beta}(\tau) - \theta(\tau)\right)Var(\Delta\beta_{it}), \\
Var(\Delta Z_{it}^{(\tau)} - \Delta S_{it}^{(\tau)}) &= \left[\alpha(\tau)\left(\frac{1}{\tau}A_{\beta}(\tau) - \theta(\tau)\right)\right]^{2} Var(\Delta\beta_{it}) \\
&\quad + \left(\alpha(\tau)\frac{1}{\tau}A_{s}(\tau) - \gamma(\tau)\right)^{2} Var(\Delta s_{t}).
\end{aligned}$$

Then the model-implied regression coefficient is therefore

$$\hat{\beta}^{auc} = \frac{\frac{1}{\tau}A_{\beta}(\tau)\alpha(\tau)\left(\frac{1}{\tau}A_{\beta}(\tau) - \theta(\tau)\right)Var(\Delta\beta_{it})}{\left[\alpha(\tau)\left(\frac{1}{\tau}A_{\beta}(\tau) - \theta(\tau)\right)\right]^{2} Var(\Delta\beta_{it}) + \left(\alpha(\tau)\frac{1}{\tau}A_{s}(\tau) - \gamma(\tau)\right)^{2} Var(\Delta s_{t})}.$$

Note that for $\alpha(\tau) = 0$, we have

$$\hat{\beta}^{auc} = 0,$$

and as $\alpha(\tau) \to \infty$, we have

$$\lim_{\alpha(\tau)\to\infty} \hat{\beta}^{auc} = \lim_{\alpha(\tau)\to\infty} \frac{\frac{1}{\tau}A_{\beta}(\tau)\left(\frac{1}{\tau}A_{\beta}(\tau) - \theta(\tau)\right)Var(\Delta\beta_{it})}{\alpha(\tau)\left(\frac{1}{\tau}A_{\beta}(\tau) - \theta(\tau)\right)^{2} Var(\Delta\beta_{it}) + \alpha(\tau)\left(\frac{1}{\tau}A_{s}(\tau)\right)^{2} Var(\Delta s_{t})}.$$

Then for large enough $\alpha(\tau)$, $\hat{\beta}^{auc} < 0 \iff \frac{1}{\tau}A_\beta(\tau) < \theta(\tau)$. This condition is a natural one, as it implies that in equilibrium, a positive demand shock for τ, i habitat investors leads to an increase in holdings $Z_{it}^{(\tau)}$.

Thus, variation in habitat demand in an extension of the model to allow for within-dealer pricing directly matches our empirical regression coefficients.