Journal of Financial Economics 150 (2023) 103723

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Intermediary balance sheets and the treasury yield curve ☆

Wenxin Du [a], Benjamin Hébert [b,*], Wenhao Li [c]

[a] Columbia Business School, NBER, and CEPR, United States of America
[b] Stanford University Graduate School of Business and NBER, United States of America
[c] University of Southern California, Marshall School of Business, United States of America

ARTICLE INFO

JEL classification:
G1
G2
F3
E5

Keywords:
Treasury bonds
Covered interest parity
Arbitrage
Intermediary asset pricing

ABSTRACT

We document a regime change in the Treasury market post-Global Financial Crisis (GFC): dealers switched from net short to net long Treasury bonds. We construct "net-long" and "net-short" curves that account for balance sheet and financing costs, and show that actual yields moved from the net short curve pre-GFC to the net long curve post-GFC. Our theory shows the regime shift caused negative swap spreads and co-movement among swap spreads, dealer positions, and covered-interest-parity violations. Furthermore, the effects of various monetary and regulatory policies are regime-dependent. We highlight Treasury supply as a plausible driver of this regime shift.

0. Introduction

Prior to the GFC, yields on Treasury bonds were below the fixed rates on interest rate swaps (specifically, overnight index swaps (OIS)) of the same maturity. This positive swap-Treasury spread ("swap spread," for short) pre-GFC was in part due to the "convenience yield" of the Treasury bonds (à la Krishnamurthy and Vissing-Jorgensen (2012)). However, post-GFC, yields on U.S. Treasury bonds have been consistently above the corresponding swap rates, suggesting that Treasury bonds are now in some sense "inconvenient." Likewise, prior to the GFC, CIP deviations were almost non-existent, but have emerged as significant and persistent following the crisis. The emergence of CIP deviations and the change in the sign of swap spreads are both well-documented in the literature.[1]

We provide new empirical evidence to link swap spreads and CIP deviations with *regimes* in which primary dealers are net long or net short with respect to Treasury bonds. Fig. 1 illustrates a significant change in the Treasury market over the past twenty years. Pre-GFC, primary dealers on net maintained a short position in Treasury bonds, and the swap spread was positive. Post-GFC, primary dealers switched to a net long position in Treasury bonds, and the swap spread became negative. That is, dealers were short Treasury bonds when those bonds had yields

☆ Nikolai Roussanov was the editor for this article. We would like to thank Adrien Auclert, Patrick Augustin, Bruno Biais, Rohan Kekre, Sven Klingler, Darrell Duffie, Michael Fleming, Thierry Foucault, Valentin Haddad, Terrence Hendershott, Sebastian Infante, Arvind Krishnamurthy, Ralph Koijen, Toomas Laarits, Tim Landvoigt, Tyler Muir, Stefan Nagel, Vania Stavrakeva, Dominik Supera, and Fabrice Tourre for insightful comments, as well as seminar and conference participants at Bank of England, BIS, Boston College, Chicago Fed, Cleveland Fed, CICF, CTFC, Cornell, CUHK, EFA, Fixed Income and Financial Institutions Conference, Florida International University, Washington University in St. Louis, NBER LTAP, HEC Paris, HKU, Duke, Michigan, Treasury Office for Financial Research, University of British Columbia, UC Berkeley, USC, and UT Austin Junior Macro. We in particular would like to thank Ken Singleton for both his insightful feedback and for sharing his term structure estimation code. We thank Filippo Cavaleri for excellent research assistance. We also thank the editor and an anonymous referee for helpful suggestions. A portion of this research was conducted at the Becker-Friedman Institute, whose hospitality B. Hébert gratefully acknowledges. This paper was largely written while Du was at the Federal Reserve Bank of New York. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of New York or the Federal Reserve System. All remaining errors are our own.

* Corresponding author.
E-mail addresses: wenxin.du@columbia.edu (W. Du), bhebert@stanford.edu (B. Hébert), liwenhao@marshall.usc.edu (W. Li).

[1] On the sign switch for long-maturity LIBOR swap spreads, see Klingler and Sundaresan (2019) and Jermann (2020). On the sign switch at all maturities for OIS swaps, which are swaps based on the effective federal funds rate, see Augustin et al. (2021) and Fleckenstein and Longstaff (2021). On the CIP deviations, see Du et al. (2018).

Available online 10 October 2023
https://doi.org/10.1016/j.jfineco.2023.103723

Received 22 December 2022; Received in revised form 26 September 2023; Accepted 26 September 2023

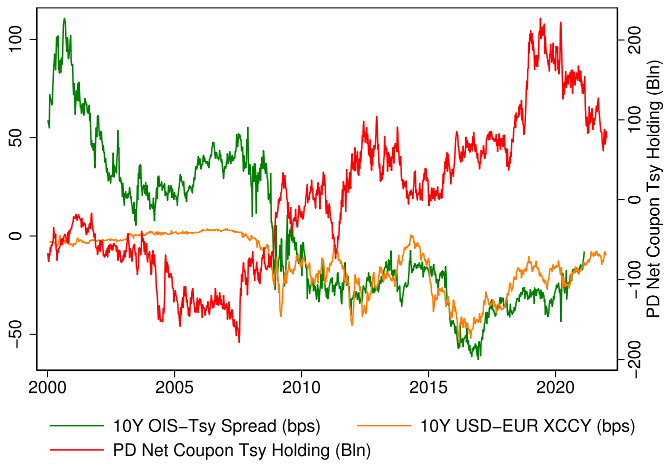

Fig. 1. Primary Dealer Treasury Holdings, Swap Spreads, and Cross-Currency Basis. *Notes*: This figure plots the spread between the 10-year OIS-Treasury swap spread (in green), and the 10-year USD-EUR cross-currency basis (in orange), and primary dealers' net holdings of coupon Treasury bonds. The Treasury yield and cross-currency basis are from Bloomberg, the 10-year OIS rate is from J.P. Morgan Markets, and the primary dealer position data are from the publicly available primary dealer statistics published by the Federal Reserve Bank of New York. The quote on the cross-currency basis swap effectively measures the direct dollar interest rate minus the synthetic dollar interest by swapping EUR interest rate into dollars (Du et al. (2018)). (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

lower than swap rates, and became long Treasury bonds once those bonds had higher yields than swap rates. The sign switch in the swap spread exactly coincided with the sign switch in dealers' net position.

We explain the collection of evidence in Fig. 1 via a unified framework that incorporates the convenience (or lack thereof) of Treasury bonds, covered interest rate parity (CIP) deviations, and primary dealers' net Treasury positions. Our framework emphasizes the role of balance-sheet constrained dealers in the Treasury market. This framework naturally leads to regimes in which dealers are net long, net short, or flat with respect to Treasury bonds. Intuitively, in the long regime, Treasury yield should be above the swap rate of the same maturity to compensate for balance sheet costs, while in the short regime, both balance sheet costs and financing spreads lower Treasury yields relative to swap rates.

Our paper shows that the Treasury market regime is central to understanding pricing dynamics and the effects of policy implications. Specifically, relative to the existing literature on swap spreads and CIP deviations, our main contributions are four-fold. First, we quantitatively demonstrate that the same balance sheet costs that rationalize CIP violations can also rationalize swap spreads, both before and after the GFC (i.e. across regimes), after properly accounting for the differences between the financing rates of long and short positions. Second, we show that dealers' actual net Treasury bond positions co-move with their incentive to act as arbitrageurs between Treasury bonds and swaps. That is, dealers' net Treasury demand curves are motivated by arbitrage. Third, we propose a framework to endogenize dealers' position, and suggest that Treasury supply is a significant driver of the regime shift and hence of negative OIS swap spreads at all maturities (complementing the pension demand channel of Klingler and Sundaresan (2019), which is relevant primarily at the 30yr maturity). Lastly, we show that effects of a wide range of monetary policies and regulatory policies are regime-dependent, and use our framework to explain why Treasury yields fell relative to other interest rates during the Global Financial Crisis (GFC) but rose during the March 2020 COVID-19 crisis (He et al. (2022) and Duffie (2020)).

Our formal analysis proceeds in the following steps. We first derive what we call the "net-long" and "net-short" curves. The net long curve describes a Treasury yield above which a dealer would always want to be net long Treasury bonds, regardless of its beliefs about future Treasury yields. Similarly, the net short curve provides a yield below which a dealer would always be willing to be net short the bond. We construct these curves taking the interest rate swap curve and the structure of CIP violations as given. Our construction uses interest-rate swaps to measure discount rates free from balance sheet costs, CIP violations to measure balance sheet costs, and takes into account the spreads between financing rates and discount rates. The interpretation of these costs is consistent with the view of Jermann (2020) that swap spreads are affected by "bond holding costs," but more specific in that these costs affect both swap spread and CIP deviations.

The use of swap rates and CIP violations is motivated by a single assumption: that all zero-cost, zero-balance-sheet trading strategies are at least weakly unprofitable under a common stochastic discount factor (SDF).[2] This assumption allows us to compare the strategy of financing the purchase of a Treasury bond with the strategy of borrowing unsecured and engaging in CIP arbitrage (the difference between these is a zero-cost, zero-balance-sheet strategy). We use this comparison to derive our net long curve, and construct our net short curve in a similar manner.

The net long and net short curves differ for two reasons: financing rates and balance sheet costs. In particular, because taking either a net short or net long position increases the size of a dealer's balance sheet, higher balance sheet costs will increase the net long yield but decrease the net short yield. We find that actual Treasury yields are quite close to the estimated net short curve pre-GFC and close to the estimated net long curve post-GFC, across a variety of maturities. This finding is consistent with the data on dealer net positions (i.e. dealers were net short pre-GFC and net long post-GFC), providing evidence in support of our view of dealers as arbitrageurs and our notion of Treasury market regimes.

We next construct a two-period, two-market supply-demand model for Treasury bonds and synthetic dollar lending in the foreign exchange (FX) swap market. Dealers arbitrage the spread between synthetic dollar lending rate and the swap rate (i.e., the CIP deviation), and the spread between swap rate and Treasury yield (i.e., the swap spread). In the model, the CIP deviation endogenously reflects the shadow cost of dealer balance sheet. Apart from dealers, we model two types of clients for Treasury bonds. One type is "real-money" investors such as domestic pension funds who do not rely on dealers' balance sheets to fund their positions. Real-money investors decide between holding Treasury bills versus long-term Treasury bonds, and their demand for bonds is an increasing function of the expected excess returns on the Treasury bonds. Second, we model foreign investors who invest in U.S. Treasury bonds financed by "synthetic dollars" (foreign currency converted to dollars using short-dated FX swaps).[3] We assume that foreign investors' demand for bonds increases in the expected return on the Treasury bond net the synthetic dollar rate.

We close the model by imposing market clearing conditions for Treasury bonds and synthetic dollars, as well as an intermediary balance sheet constraint. The model has a unique equilibrium, which falls into one of the three regimes: intermediaries are either long, short, or flat with respect to Treasury bonds. The regime itself is determined by the difference between Treasury bond supply and client demand; that is,

[2] This assumption is implicit in the models of Jermann (2020) and Du et al. (2022). Under this assumption, the SDF prices zero-cost and zero-balance-sheet strategies (e.g. derivatives including swaps with negligible balance sheet cost and zero initial cost), but does not directly determine the price of cash securities such as Treasuries.

[3] FX-hedged foreign investors do not use leverage but nevertheless rely on dealer balance sheets to obtain FX swap hedging. A canonical example of an FX-hedged foreign investor is a Japanese life insurance company, with liabilities denominated in yen and substantial U.S. dollar fixed-income assets.

dealers act as "buyer of last resort" in the long regime and a "short-seller of last resort" in the short regime.

In our framework, the change in sign of dealer positions leads naturally to a change in sign in the swap spread. Our quantitative model shows that the change in net dealer positions from short to long and associated change in financing costs is on its own sufficient to explain the sign changes in swap spreads, even in the absence of changes in the balance sheet costs. However, the large increase in the balance sheet cost post-GFC is essential to quantitatively explain the size of the observed change in swap spreads and the post-GFC co-movement of swap spreads and CIP violations, even though the balance sheet cost alone cannot directly explain the sign switch in dealers' Treasury position. More specifically, the spread between the financing rate and the floating rate in the interest rate swap is a significant determinant of the positive swap spreads pre-GFC, whereas balance sheet costs are the main driver of the negative swap spread post-GFC (Appendix Figure A9 illustrates a decomposition of these forces within our quantitative model). The former is consistent with an earlier literature on swap spreads, e.g. Grinblatt (2001) and Feldhütter and Lando (2008).

Our framework also allows us to analyze the drivers of the pre- to post-GFC regime change. The model generates a clear prediction that an increase in the Treasury supply increases dealers' Treasury position, other things being equal. Since the GFC, the supply of Treasury bonds, net of Federal Reserve purchases, has experienced a four-fold increase, which acts as an important driver for the regime shift in dealers' position and the sign switch in the swap spread. Meanwhile, as discussed in Klingler and Sundaresan (2019), underfunded pension funds significantly increased their demand for duration through interest rate swaps post-GFC. As dealers are natural counterparties for the swap demand of pension funds, an increase the swap spread can in theory increase dealers' positions. Our regression analysis suggests that interest rate swap demand from pensions is indeed important in explaining the swap spread at the 30-year maturity, while the Treasury supply factor is significant at all maturities. Our discussion on the role of Treasury supply is also consistent with the evidence documented in Fleming et al. (2022) that an increase in Treasury issuance leads to an increase in dealer inventory.

Furthermore, the model helps explain a new empirical fact we document in the paper: dealer's net long position in Treasury bonds is negatively correlated with the term spread between the long-term Treasury bond yield and short rate in the post-GFC period. That is, dealers hold more bonds when the price is high relative to the short rate (i.e. a small term spread) and fewer bonds when the price is low relative to the short rate (i.e. a large term spread). This behavior at first appears puzzling, on the grounds that a larger term spread predicts larger expected returns of bonds vs. the short rate (Campbell and Shiller (1991)). Our view of dealers as arbitrageurs absorbing the excess supply/demand for Treasury bonds can rationalize this fact: during the post-GFC period, when the term spread is low, client demand for Treasury is low. To clear the market, dealers must increase their holdings, which requires that Treasury bonds offer an attractive yield relative to interest rate swaps (after accounting for financing rates and balance sheet costs). In contrast, the existing literature has often emphasized instead returns-seeking behavior (as in the intermediary asset pricing literature, He and Krishnamurthy (2013)) and the role of dealers as over-the-counter market makers (as in Duffie et al. (2005)). These perspectives are not mutually exclusive with our view; however, they do not naturally lead to the prediction that dealers should buy high and sell low in the Treasury market, as suggested by the data.

In the last section of the paper, we demonstrate that the effects of a variety policy interventions are regime-dependent, using our two-period model. We also show that the effects of dealer distress differ across regimes, offering an explanation for why Treasury yields went down relative to swap rates in the GFC and up during March 2020. These differences arise from differences in the comparative statics of our model across regimes, which ultimately are a consequence of the difference in the way balance sheet costs affect Treasury yields. The model, with minimal modification, is capable of speaking to the effects of quantitative easing and tightening, inter-central-bank swap lines, regulatory exemptions to the supplementary leverage ratio (SLR), and interest rate policy. We discuss the effects of each of these policies, and emphasize how they differ in the long and short regimes. Finally, we discuss the implications of the framework for the ongoing tightening cycle, and draw a parallel with the experience of 2017-2019 tightening cycle.

Our paper is most closely related to Jermann (2020) and Du et al. (2022), in that we model swap spreads (Jermann (2020)) and CIP violations (Du et al. (2022)) as arising from constraints on intermediaries. Our paper combines these perspectives quantitatively and shows that the same balance sheet costs can jointly explain these arbitrage opportunities and data on dealers' positions. One key difference between our paper and Jermann (2020) is that Jermann (2020) treats Treasury yields as exogenous, and as a result predicts that dealer positions are increasing in the term spread, in contrast to our framework and the data. Related to this, Favara et al. (2022) show evidence that SLR shocks have reduced large banks' participation in the U.S. Treasury market, and Klingler and Sundaresan (2023) provide evidence that dealer balance sheet constraints contribute to the diminishing convenience of T-bills. Du et al. (2018), Hébert (2020), and Du et al. (2022) argue that CIP deviations can proxy for the shadow cost of the these constraints. The strong co-movement of CIP violations and swap spreads post-GFC we document is consistent with these perspectives.

Recent work by Liu (2019) and Siriwardane et al. (2021) examines co-movement across different near-arbitrages. Liu (2019) argues in favor of a significant common component across arbitrage opportunities, whereas Siriwardane et al. (2021) provide evidence suggestive of market segmentation. In both papers, the correlation between CIP deviations and the swap spread stands out being among the highest, which supports the use of CIP deviations as a balance sheet cost proxy for Treasury trading activities. Our view is closer to that of Liu (2019), in that we expect (especially at lower frequencies) a strong degree of co-movement across arbitrage spreads that require dealer balance sheet.

Considering swap spreads and CIP deviations together helps address a puzzle: why would long-maturity (e.g. 30 year) swap spreads be affected by fluctuations in current balance sheet costs? The answer suggested by Du et al. (2022) is that there is a substantial risk premium associated with the risk that balance sheet costs increase; our quantitative analysis confirms that this risk premium plays a significant role in long maturity swap spreads. Our quantitative term-structure framework is substantially more general than the models employed in these papers, and our two-market equilibrium model is able to address a broader range of policy questions.

The key limitation of our static model is that it treats swap rates as exogenous, and we do not attempt to explain the behavior of dealers' swap counter-parties. In contemporaneous work, Hanson et al. (2022) start from a similar set of motivating facts and adopt an approach broadly similar to our two-period model to explain the way in which shocks to the demand for interest rate swaps affect swap spreads. Hanson et al. (2022) focus on the swap market, treating the Treasury market as exogenous; combining the two approaches is an interesting direction for future research. Their interest is in separating supply and demand shocks, whereas our focus is on providing estimates of the dealer net long and short Treasury curves, validating the notion of Treasury market regimes, and analyzing policy interventions.

The convenience yield literature argues that Treasury securities are valued not only by their fundamental cash flows, but also by the superior liquidity and safety (Longstaff (2004); Krishnamurthy and Vissing-Jorgensen (2012); Greenwood et al. (2015); Joslin et al. (2021)), with macro-finance implications of this convenience discussed in Drechsler et al. (2018), Jiang et al. (2021), and Krishnamurthy and Li (2023). Our results highlight that no convenience demand is needed to explain intermediaries' demand for Treasury bonds. More recently, the dislocation of the Treasury bond market during the height of the COVID-19 pandemic in March 2020 has led some authors to question whether U.S. Treasury bonds remain convenient (e.g. Duffie (2020) and He et al. (2022)). We highlight instead the importance of the regime change in the dealers' net position and the interaction between dealer balance sheet constraints and client demand for Treasury bonds.

Complementary to our analysis on the role of dealers, the role of hedge funds in the Treasury market has been examined in Barth and Kahn (2021) and Kruttli et al. (2021), particularly regarding their activities in the Treasury cash-futures basis arbitrage funded by dealers' balance sheets. We develop the net long and net short curve from the perspective of a securities dealer, but argue in Appendix Section F.1 that dealers will in effect transmit their balance sheet costs to hedge funds and other levered clients, consistent with Boyarchenko et al. (2018a). As a result, the curves we develop are applicable for these levered clients as well.

The negative correlation between dealers' net position and Treasury yield curve slope we document is the opposite to the pattern observed in typical (excluding the largest) commercial bank portfolios (Haddad and Sraer (2020)). This contrast emphasizes the importance of distinguishing between the Treasury activities of securities dealers and commercial bank subsidiaries. Fluctuations in dealers' inventory have also been linked to overall liquidity conditions (Goldberg (2020)).

The structure of the paper is as follows. We provide institutional background on Treasury trading by dealers in Section 1. We derive and estimate the net long and the net short curves in Section 2. We introduce the demand from real-money investors and build an equilibrium model for Treasury market dynamics in Section 3. We analyze policy implications in Section 4, and in Section 5 we conclude.

1. Institutional background

In this section, we outline the mechanics about how dealers go long or short Treasury securities and hedge with swaps. We then discuss when these strategies are arbitrages, and compare them with CIP arbitrage. For interest rates, yields, and returns, capital letters denote gross rates, and lowercase letters denote log rates.

1.1. The long treasury vs. swap arbitrage

Suppose that the dealer goes long (buys) a zero-coupon Treasury bond of maturity n at time t. At the onset of the trade, the dealer buys the Treasury bond at gross yield $Y_{n,t}$, and finances the entire position period-by-period at the time $t+j$ gross financing rate I^l_{t+j} (for $j = 0, ..., n-1$). This trade is zero-cost, in the sense that the entire purchase cost of the bond is financed through a mix of secured (repo) and unsecured debt financing (I^l_{t+j} is the effective one-period gross financing rate of this mix at time $t+j$; see Internet Appendix Section A for more details.).

A dealer can hedge the interest rate risk of this trade by entering into a swap contract.[4] Interest rate swaps exchange a fixed interest payment ($R_{n,t}$) for a sequence of floating interest payments (R_{t+j}).[5] To hedge a net long bond position, the dealer pays the fixed rate and receives the floating rate in a swap, whose maturity is set to match the bond (n in our example).

[4] Dealers have an incentive to hedge their net interest rate risk, due to risk-based capital requirements, but will typically do so at the trading desk level or the whole book level as opposed to trade-by-trade.

[5] Our analysis will focus on overnight index swaps (OIS), in which one party pays a fixed rate of interest in exchange for a series of floating payments indexed to the overnight interbank federal funds rate. Prior to the GFC, swaps indexed to LIBOR were more commonly used, and recently swaps indexed to SOFR (Secured Overnight Financing Rate) have been introduced. OIS rates are available for our entire sample and are similar to LIBOR swap rates pre-GFC and to SOFR swap rates in the recent period.

After hedging, the dealer will receive the spread between the floating rate R_{t+j} and the financing rate I^l_{t+j} each period. If the dealer holds the bond and swap to maturity, the dealer will earn the yield $Y_{n,t}$ on the bond and pay the fixed rate $R_{n,t}$ on the swap; the difference of these two is the negative of the "swap spread" $R_{n,t} - Y_{n,t}$. Thus, if the floating rate is guaranteed to exceed the financing rate ($R_{t+j} \geq I^l_{t+j}$ for all j), and the swap spread is negative, the dealer is guaranteed to earn a profit on this trade if held to maturity. If in addition the spread $R_{t+j} - I^l_{t+j}$ is constant, the dealer's profit is certain ex-ante (because $R_{n,t} - Y_{n,t}$ is known at the trade's inception). In this sense, the swap hedges the bond. If instead $R_{t+j} - I^l_{t+j}$ is not constant, there is residual "basis risk," but negative swap spreads are nevertheless an arbitrage opportunity if $R_{t+j} - I^l_{t+j}$ is guaranteed to be positive. During this trade, the dealer will have a larger balance sheet, approximately equal to the value of the Treasury net long position.[6]

1.2. The short treasury vs. swap arbitrage

Short-selling a Treasury bond and hedging with a swap works in an essentially identical way, with the directions of the trades reversed. When a dealer short-sells a Treasury bond, the dealer borrows the bond from its owner, and offers the owner cash as collateral. The dealer raises that cash by selling the borrowed bond, and each period receives an effective interest rate of I^s_{t+j} on this cash. See Internet Appendix Section A.2 for more details.

The dealer can hedge with an interest rate swap by receiving the fixed rate $R_{n,t}$ and paying the floating rate R_{t+j}. If the spread between R_{t+j} and I^s_{t+j} is bounded above, and the swap spread $R_{n,t} - Y_{n,t}$ exceeds this bound, the dealer is guaranteed to make a profit, assuming the position is held to maturity. This arbitrage also increases the size of the dealer's balance sheet. The asset in this case is a "loan" to the original owner of the bond (the dealer has given the original owner cash in exchange for Treasury collateral), and the liability is a security to be delivered.

Thus, the long Treasury-swap and short Treasury-swap arbitrage strategies are both zero-cost, n-period strategies that use balance sheet and, under certain assumptions, deliver a guaranteed profit if held to maturity. We will compare these strategies to another n-period, zero-cost, balance-sheet-intensive arbitrage opportunity: CIP arbitrage.

1.3. CIP arbitrage

One-period USD-EUR CIP arbitrage involves borrowing dollars in the cash market, converting the dollars to euros in the spot foreign exchange market, lending out the euros, and using a forward contract to lock in the exchange rate at which the euro proceeds of the loan are converted back to dollars. Suppose the dealer borrows dollars in unsecured funding markets at the rate R_t (the same rate used in the interest rate swap), and define the "synthetic" dollar rate R^{syn}_t as the synthetic dollar lending rate in the FX swap market obtained by converting the euro lending rate into dollars. The profit of the one-period CIP arbitrage is the difference between these two rates, $R^{syn}_t - R_t$, and is positive for most major USD-pairs in the post-GFC period (Du et al. (2018)). This trading strategy is also zero-cost, in that the synthetic dollar lending is entirely financed by unsecured dollar borrowing.

A dealer who plans to engage in the CIP arbitrage for n periods can lock in the profits of this trade using interest rate and cross-currency

[6] To a first approximation, the interest rate swap is entirely off-balance-sheet. More precisely, trading interest rate swaps can increase the size of the balance sheet slightly (Boyarchenko et al., 2018b). The total exposure includes initial and variation margins (typically a couple percent of total notional), and an additional 0-1.5% of the swap notional calculated for off-balance sheet interest rate derivative exposure using the Current Exposure Method, depending on the maturity of the interest rate swaps. We ignore the additional balance sheet costs of trading derivatives to simplify our analysis.

basis swaps (the details of which are unimportant at this stage of our discussion). However, engaging in CIP arbitrage will increase the size of the dealer's balance sheet; in this case, the asset is the euro loan and the liability is the dollar borrowing.

1.4. Comparing arbitrages

Each of these three arbitrages lasts for n periods if held to maturity, has zero initial cost, and uses balance sheet.[7] The core of our analysis focuses on perturbations in which a dealer does less CIP arbitrage and more of one of the two Treasury vs. swap arbitrages. These perturbations can be viewed as zero-cost, zero-balance sheet trading strategies.

Empirically, the gross rate at which dealers finance the long Treasury trade (I^l_{t+j}) is higher than the gross rate at which they finance the short Treasury trade (I^s_{t+j}). We document in Internet Appendix Section A.2 that the annualized rate the original bond owners ("security lenders") pay to dealers is on average roughly 20 basis points below I^l_{t+j}, even for securities that are not "special", using data from Markit Securities Finance.[8] Because of this difference, the long and short Treasury vs. swap arbitrages will never both be appealing to dealers. If swap spreads are sufficiently low/negative, as in the post-GFC period, the long arbitrage will be profitable, while if swap spreads are sufficiently large, as in the pre-GFC period, the short arbitrage will be profitable.

Each of these arbitrages is subject to market-to-market fluctuations. For example, the two-period CIP arbitrage has a risk-premium associated with the possibility that the one-period CIP arbitrage becomes larger next period (Du et al. (2022)). To account for this, we construct a model, based on a single critical assumption: that all zero-cost, zero-balance-sheet trading strategies are weakly unprofitable under a common SDF. Using this assumption, we carefully compare the swap-Treasury and CIP arbitrages.

2. The long and short treasury yield curves

In this section, we construct what we call "net long" and "net short" yield curves. These yield curves represent arbitrage bounds at which a dealer would be willing to go net long or net short Treasury bonds. In frictionless models (in which arbitrage capacity is unlimited and the gross financing rates I^l_t and I^s_t are equal), there is a single Treasury yield curve. At yields above this yield curve, dealers would want to go net long bonds, while at lower yields, dealers would want to go net short bonds. In our model, two frictions create a wedge between the yield at which dealers would go net long and the yield at which dealers would go net short.

The first friction is balance sheet costs. As emphasized in the previous section, going either net long or net short a Treasury bond increases the size of dealer balance sheets. This has an opportunity cost, and this opportunity cost acts like a tax on both these trades. It raises the yield dealers require when going net long, while lowering the yield they require when going net short. We proxy for balance sheet costs using CIP deviations.

The second friction is the difference in financing rates of the long and short Treasury positions. When a dealer goes net long, they finance the position at a relatively high rate (I^l_t); when a dealer goes net short, they receive a relatively low rate on their cash (I^s_t). The difference between these two rates creates another wedge between the net long and net short curves.

[7] There are differences between the strategies with regards to the timing of payments. We develop our term structure model to account for these kinds of issues.

[8] When a dealer makes a repo loan to a hedge fund client (typically at a rate above I^l_{t+j}), the client selects the Treasury collateral. A dealer who wants to short a specific bond cannot rely on this kind of lending to find the bond. Instead, the dealer borrows the bond from a security lender at the lower rate I^s_{t+j}.

Our analysis in this section is partial equilibrium. We take swap rates, financing rates, and Treasury yields as given, and ask if dealers are willing to go net long or net short Treasury bonds. We will then compare our answer to this question, which is based on price data, with our quantity data on dealers' net Treasury positions.

We proceed in four steps. First, we will introduce a very simple model of dealer behavior, to illustrate the trade-off a dealer faces between going net long or short Treasury bonds and other arbitrage activities. Second, we generalize the ideas illustrated in this simple model, and construct the net long and net short curves. Third, we discuss the assumptions under which these curves represent arbitrage bounds. Finally, we build a term structure model to estimate the net long and net short curves.

2.1. One-period net-long and net-short treasury yields

We first consider the problem of a risk-neutral dealer who can choose between trading a single n-period zero-coupon Treasury bond and a one-period CIP arbitrage, subject to a fixed balance sheet constraint. We make these stark assumptions (risk-neutrality, only one bond, fixed balance sheet) to illustrate the main idea; they are not imposed in the more general analysis that follows. The dealer makes this choice at a single date in this simplified example; for this reason, we will omit time subscripts.

Let q^{bond} be the dealer's bond position (in dollars, not notional), with negative values implying short-selling. Let q^{syn} be the dealer's "synthetic dollar" position (in dollars), which we define as the currency hedged investment leg of the CIP arbitrage (the other leg is borrowing dollars). We assume that the gross synthetic dollar rate (e.g. the currency-hedged euro rate) is above the gross dollar borrowing rate ($R^{syn} > R$), so that the direction of the arbitrage is to lend in synthetic dollars and finance in actual dollars.

Suppose the dealer faces a fixed balance sheet constraint,

$$q^{syn} + |q^{bond}| = \bar{q}, \tag{1}$$

where $\bar{q} > 0$ is the balance sheet capacity for Treasury and CIP arbitrage. Note that the absolute value of q^{bond} enters this expression, reflecting the fact that both long and short positions require balance sheet.

We assume, following the discussion above, that if the dealer buys the bond, the dealer will finance its position with at the gross interest rate I^l. If the dealer instead short-sells the bond, it receives a gross return I^s on its cash. We define ER^{bond} as the expected gross bond return. In this simple example, we will treat this expected return as exogenous. The dealer will buy the bond if ER^{bond} is sufficiently high, and short the bond if ER^{bond} is sufficiently low.

The dealer's problem is

$$\begin{aligned}\max_{q^{bond},q^{syn}} & \Big(\underbrace{ER^{bond}}_{\text{return of owning bond for one period}} - \underbrace{I^l}_{\text{financing of long position}} \Big) \cdot \max\{q^{bond},0\} \\ & + \Big(\underbrace{I^s}_{\text{return earned on cash collateral}} - \underbrace{ER^{bond}}_{\text{return of shorting bond for one period}} \Big) \cdot \max\{-q^{bond},0\} \\ & + \Big(\underbrace{R^{syn} - R}_{\text{CIP arbitrage spread}} \Big) \cdot q^{syn}\end{aligned} \tag{2}$$

subject to the balance sheet constraint in (1). That is, the dealer chooses between the long Treasury trade, the short Treasury trade, and the CIP arbitrage. We have omitted the swap part of the Treasury-swap arbitrages from these calculations. This simplification is justified under the assumption that the swaps are fairly priced from the perspective of the dealer.

We assume that the profit of the Treasury trade does not dominate the synthetic lending trade, so that the synthetic lending amount is always non-zero. This assumption eliminates the corner case of $q^{syn}=0$, which has less empirical relevance, and implies that the shadow cost of balance sheet constraint is measured by CIP deviations. We also assume, as in the data, that $I^s < I^l$.

Solving the problem, we find three different regimes. The first regime features $q^{bond}>0$ and $q^{syn}>0$. In this regime, the dealer must be indifferent to reallocating balance sheet capacity between the long Treasury and synthetic lending trades (simultaneously increasing q^{bond} and decreasing q^{syn} by equal amounts, or the reverse). That is, the first-order conditions require

$$ER^{bond} = I^l + (R^{syn} - R).$$

To further build intuition, we log-linearize the expected bond return:

$$ER^{bond} = E[\exp(ny - (n-1)y')] \approx 1 + ny - (n-1)E[y'],$$

where y is the log bond yield and y' is the next period log bond yield.

The log yield y in this regime must equal the "net long yield" y^l, which can be approximated as

$$y^l \approx \frac{n-1}{n}E[y'] + \frac{1}{n}(i^l + r^{syn} - r). \tag{3}$$

Here and in what follows, for interest rates, we use lower-case letters to indicate log rates (e.g. $i^l = \ln(I^l)$ and $r^{syn} = \ln(R^{syn})$).

The second regime features $q^{bond}<0$ and $q^{syn}>0$. In this regime, indifference at the margin between CIP arbitrage and the short Treasury trade requires that

$$ER^{bond} = I^s - (R^{syn} - R).$$

The log yield y in this regime is equal to the "net short yield" y^s, which can be approximated as

$$y^s \approx \frac{n-1}{n}E[y'] + \frac{1}{n}(i^s - r^{syn} + r). \tag{4}$$

Comparing these two regimes, we can see that the expected return is lower in the short regime than in the long regime, for two reasons. First, the funding rate is lower ($I^s < I^l$), and second, the opportunity cost of balance sheet ($R^{syn} - R$) is subtracted from the short funding rate but added to the long funding rate. As a result, the short yield is below the long yield, reflecting (intuitively) that the dealer requires a low yield/high price to justify a net short position and a high yield/low price to justify a net long position.

The third possible regime features $q^{bond}=0$ and $q^{syn}>0$. In this regime, the Treasury bond yield has to satisfy $y^s \le y \le y^l$. That is, the profit on the long and short Treasury strategies does not justify the balance sheet cost, so the dealer chooses to have a zero net Treasury position.

To gain further intuition, consider the case of a one-period bond that matures next period ($n=1$). Defining $r^{cip} = r^{syn} - r$, (3) and (4) can be written in the $n=1$ case as

$$y^l \approx\ r - (r - i^l) + r^{cip},$$

$$y^s \approx\ r - (r - i^s) - r^{cip}.$$

The net long yield can differ from r (which we will think of as the one-period swap rate) for two reasons. First, holding Treasury bonds takes up bank balance sheet, so the yield has to be higher by an amount equal to the opportunity cost of the balance sheet (measured by r^{cip}). Second, if dealers' financing rate is lower than their unsecured funding cost, $i^l < r$, then there is a financing benefit to owning the Treasury bond, which makes the dealer willing to accept a lower yield.

The net short yield can differ from r for similar reasons. The impact of the opportunity cost of balance sheet affects the sell yield with a negative sign. The sell yield has to be lower (the price to be higher) to justify dealer's short position, which also takes balance sheet. The sell yield is further lowered if the return on the cash collateral is lower than the dealer's borrowing cost, $i^s < r$.

2.2. Multi-period net long and net short curves

We next extend the logic of these yields to a more general, multi-period setting, constructing what we will call the "net long curve" and "net short curve."

We make three key assumptions: (i) there is an SDF that prices interest rate swaps and cross-currency basis swaps, (ii) some synthetic dollar lending occurs in equilibrium, and (iii) all zero-cost, zero-balance-sheet trading strategies are weakly unattractive under this SDF (i.e. $0 \ge E[MR]$ in the standard notation). We will interpret this SDF as a dealer's SDF, under the usual intermediary asset pricing assumption that dealers are active in all of these markets. Let $\mathbb{Q}$ denote the risk-neutral measure associated with this SDF.

The first assumption is justified by the observation that essentially all of the arbitrages documented in the literature involve the use of balance sheet. This is true, for example, of all of the arbitrages considered in Boyarchenko et al. (2018a) and Siriwardane et al. (2021). That is, no-arbitrage between derivatives appears to hold in the data. The second is justified by the empirical observation that currency dealers actively engage in CIP arbitrage.

The third assumption generalizes the first-order condition of our simple model. In that model, the strategy of increasing or decreasing q^{bond}, financed at the appropriate financing rate, and offsetting the change in balance sheet by changing q^{syn}, is weakly unattractive for the dealer. Our generalization of this condition assumes that all feasible zero-cost, zero-balance sheet perturbations are weakly unattractive.

These assumptions are much more general than the simple model of the previous section. In particular, we have made no assumptions on the pricing of balance-sheet-increasing or balance-sheet-reducing strategies, and hence are agnostic about why dealers find it costly to increase the size of their balance sheet. We have said nothing about how many maturities of Treasury bonds are available, or what other assets are also traded by dealers. We have not taken a stance on whether (and if so, why) dealers are risk-averse; the dealer's risk-neutral measure $\mathbb{Q}$ could depend on shareholder preferences, managerial preferences, value-at-risk constraints, and other considerations.

In the equations that follow, each period is a month and all rates and yields are monthly. Consider the strategy of going long the Treasury bond at log yield $y_{n,t}$, financing via repo, and offsetting the balance sheet effect by reducing CIP activity. As in our simple model, the expected bond return cannot exceed the cost of financing a long-position, adjusted for balance sheet costs:

$$\underbrace{I_t^l}_{\text{secured financing}} + \underbrace{(R_t^{syn} - R_t)}_{\text{forgone CIP profits}} \ge \underbrace{E_t^{\mathbb{Q}}[e^{ny_{n,t} - (n-1)y_{n-1,t+1}}]}_{\text{expected bond return}}. \tag{5}$$

The left-hand side of this expression represents costs paid at time $t+1$ per dollar invested at time t. The financing of the bond purchase must be repaid (the first term), and the profits of the forgone CIP arbitrage are lost (the second term). These must be weighed against the expected bond return (the payoff at time $t+1$ per dollar invested at time t). Note that because both sides of this equation are defined in terms of $t+1$ payoffs, the discount rate associated with the SDF is irrelevant.

Define the log interest rate

$$x_{1,t} = \ln(I_t^l + R_t^{syn} - R_t), \tag{6}$$

and iterate:

$$y_{n,t} \le y_{n,t}^l = -\frac{1}{n}\ln(E_t^{\mathbb{Q}}[\exp(-\sum_{j=0}^{n-1} x_{1,t+j})]) \approx E_t^{\mathbb{Q}}[\frac{1}{n}\sum_{j=0}^{n-1}(i_{t+j}^l + r_{t+j}^{syn} - r_{t+j})]. \tag{7}$$

The yield curve $y_{n,t}^l$ defines what we call the "net long curve." This curve represents the point at which dealers would be willing to switch from CIP arbitrage activity to taking a net long position in a Treasury bond. Since zero-cost, zero-balance-sheet strategies are weakly unattractive, this net long curve is, by induction, an upper bound on Treasury yields (under the assumption that dealers engage in CIP arbitrage).

This net long curve can also be viewed as a lower bound on swap spreads (defined as the difference between swap rates and Treasury yields of matching maturity). Let $r^{cip}_{n,t}$ be the log n-period CIP violation (the n-period synthetic dollar rate minus the n-period swap rate). Linearizing (7) and recalling that $r_{n,t}$ is the log n-period swap rate,

$$r_{n,t} - y_{n,t} \geq r_{n,t} - y^l_{n,t} \approx -\underbrace{r^{cip}_{n,t}}_{\text{n-period CIP violation}} + E^{\mathbb{Q}}_t[\frac{1}{n}\sum_{j=0}^{n-1}\underbrace{(r_{t+j} - i^l_{t+j})}_{\text{financing benefit}}]. \quad (8)$$

The difference between the short-maturity swap rate r_{t+j} and the financing rate i^l_{t+j} is generally small and stable over time. As a result, equation (8) implies that if yields are close to the net long curve, we should expect swap spreads to be negative and close in absolute value to the matched-maturity CIP violation.

We should also emphasize the important role of risk premia that is hidden in this expression. The n-period CIP violation $r^{cip}_{n,t}$ is well above the physical ($\mathbb{P}$) measure expectation of future short-maturity CIP violations, as emphasized by Du et al. (2022). That is, net long curve yields are higher than swap rates both because of expectations of non-zero future balance sheet costs and because of the risk associated with the possibility that these costs become larger.

We develop the net short curve via similar logic. Consider the strategy of short-selling the Treasury bond at yield $y_{n,t}$, borrowing the bond against cash collateral, and offsetting the balance sheet effect by reducing CIP activity. Again as in our simple model, the return on the cash collateral cannot exceed the bond return after accounting for the opportunity cost of balance sheet:

$$\underbrace{I^s}_{\text{gross return on cash collateral}} \leq \underbrace{E^{\mathbb{Q}}_t[e^{ny_{n,t}-(n-1)y_{n-1,t+1}}]}_{\text{expected bond return}} + \underbrace{(R^{syn}_t - R_t)}_{\text{forgone CIP profits}}. \quad (9)$$

The left-hand side of this expression is the cash generated at date $t+1$ per dollar invested at t by selling the bond and posting the cash as collateral. The right-hand side represents the costs of this trade at date $t+1$ per dollar invested at t, including both the cost of repurchasing the bond and the forgone CIP profits. Define the log interest rate

$$x_{2,t} = \ln(I^s_t + R_t - R^{syn}_t) \quad (10)$$

and iterate:

$$y_{n,t} \geq y^s_{n,t} = -\frac{1}{n}\ln(E^{\mathbb{Q}}_t[\exp(-\sum_{j=0}^{n-1} x_{2,t+j})]) \approx E^{\mathbb{Q}}_t[\frac{1}{n}\sum_{j=0}^{n-1}(i^s_{t+j} - r^{syn}_{t+j} + r_{t+j})]. \quad (11)$$

The yield curve $y^s_{n,t}$ defines what we call the "net short curve." This curve represents the point at which dealers would be willing to switch from CIP arbitrage activity to taking a net short position in a Treasury bond. The net short curve is a lower bound on Treasury yields (again under the assumption that dealers engage in CIP arbitrage).

The net short curve can also be viewed as an upper bound on swap spreads. Linearizing (11),

$$r_{n,t} - y_{n,t} \leq r_{n,t} - y^s_{n,t} \approx \underbrace{r^{cip}_{n,t}}_{\text{n-period CIP violation}} + E^{\mathbb{Q}}_t[\frac{1}{n}\sum_{j=0}^{n-1}\underbrace{(r_{t+j} - i^s_{t+j})}_{\text{security lending spread}}]. \quad (12)$$

The difference between the short-maturity swap rate and the security lending rate is positive, and the n-period CIP violation in our definition is also positive and proxies for the balance sheet cost. Taking these two forces together, equation (12) implies that if yields are close to the net short curve, we should expect positive swap spreads.

Our assumptions are sufficient to determine an upper and lower bound on Treasury yields (or, equivalently, on swap spreads), but are not enough to pin down Treasury yields themselves. In a frictionless world, the net long and net short curves converge to one curve and thus exactly pin down Treasury yields. In the presence of frictions, yields can fall anywhere in between the net short and net long curves.

2.3. Discussion

Before proceeding, we elaborate on the interpretation of these curves in pre-GFC and post-GFC periods, the role of interest rate swap hedges, the connection between yields and positions, and a broader interpretation of intermediaries as dealers and their levered clients.

Pre-GFC and post-GFC Pre-GFC, CIP violations were roughly zero and security lending rates (i^s_t) were below one-month swap rates (r_t) by an average of 25 basis points annualized.[9] It follows (by (12)) that net short curve yields were significantly lower than matched-maturity swap rates, which is to say there was a significant positive swap spread. In contrast, the net-long financing rate (i^l_t) is only about five basis points (annualized) below swap rates (during both the pre-GFC and post-GFC periods), leading to a net-long curve that is only slightly above the swap curve (by (8)).

Post-GFC, synthetic lending rates are well above swap rates ($r^{cip}_t = r^{syn}_t - r_t \gg 0$). By (8), this leads to a significant "negative swap spread" between swap rates and the net long curve. In contrast, the net short curve now features yields far below swap rates, reflecting both the spread between the short financing rate and one-month swap rates and the CIP violations.

Actual swap spreads went from being large and positive pre-GFC to negative post-GFC. This fact, combined with the discussion above, previews our result that Treasury yields went from being close to the net short curve pre-GFC to close to the net long curve post-GFC. We discuss the consequences of this shift in Sections 3 and 4.

The net short and net long curves as arbitrage bounds The bound $y_{n,t} \in [y^s_{n,t}, y^l_{n,t}]$ is an arbitrage bound if $y_{n-1,t+1} \in [y^s_{n-1,t+1}, y^l_{n-1,t+1}]$ with probability one, which follows from our key assumptions. This observation gives rise to the following interpretation: the net long yield is a yield at which a dealer would be willing to go long even if the dealer believed all dealers would be net long the bond in the future with probability one. Likewise, the net short yield is the yield at which a dealer would be willing go short, even if the dealer believed all dealers would be net short in the future with probability one. Yields falling in between these bounds can be loosely interpreted as related to the probability that dealers will be either net long or net short in the future.

In Internet Appendix Section F.2, we show that if $y_{n,t} > y^l_{n,t}$, dealers will perceive an arbitrage opportunity, even if future yields could be outside the bounds (i.e. if it is possible that $y_{n-1,t+1} > y^l_{n-1,t+1}$). The net short curve does not share this property. However, we are able to derive a lower value for yields which we call "partial equilibrium net short curve" such that the dealer will always be willing to go short, and show that the net short curve described in the main text is a close approximation of this partial equilibrium net short curve.

Hedging with swaps It is standard practice to hedge a Treasury position with an interest rate swap. Prior to the GFC, hedging using LIBOR swaps was typical; following the GFC, OIS rose in prominence, and more recently swaps referencing repo rates (SOFR swaps) have replaced LIBOR swaps following the benchmark reference rate reform. Our analysis will focus on OIS swaps, which are available for a reasonably long sample (unlike SOFR swaps) and are insensitive to credit risk (unlike LIBOR swaps). See Internet Appendix Section A for more details.

[9] This spread is time-varying, and correlated with yield curve factors. As a result, there is a non-trivial risk premium that lowers the net short curve even further in our quantitative model.

Hedging Treasury bonds with interest rates swaps leaves the dealer exposed to mark-to-market risk associated with fluctuations in the Treasury-swap spread. In our log-linearized equations (8) and (12), we can see that the swap spreads for the net-long and net-short curve depend on the risk-neutral expectations of future balance sheet constraints and residual basis spreads between money market rates.

Thus, the net long curve $y^l_{n,t}$ can be hedged with a synthetic dollar swap (by swapping EUR rates to dollars), under the assumption that the financing spread $r_t - i^l_t$ stable. The net short curve $y^s_{n,t}$ can be hedged via a combination of swaps and synthetic dollar swaps, under the assumption that the financing spread $r_t - i^s_t$ is stable.

If one assumes that a Treasury yield will always be at one end of these boundaries, then the Treasury bond itself can be hedged. But note that the hedge will differ depending on which of the two boundaries is assumed to apply, and in neither case is the hedge a single swap. The intuition for this result is that balance sheet costs matter, and hedging the Treasury bond with an interest rate swap hedges interest rate risk but does not hedge balance sheet risk.

Dealers and levered clients In developing the net long and net short curves, we have adopted the perspective of a dealer. In Internet Appendix Section F.1, we argue that these curves are also arbitrage bounds from the perspective of the dealer's levered clients (i.e. hedge funds). Dealers' balance sheet costs are in effect transmitted to these clients via bilateral lending markets, a point emphasized by Boyarchenko et al. (2018a). As a result, balance sheets costs will influence a substantial segment of the Treasury market, even though dealers on their own hold a relatively small quantity of Treasury bonds on net. Internet Appendix Section F.1 also contains some suggestive evidence supporting this perspective. In what follows, we will treat dealers and their levered clients as a consolidated entity.

2.4. Term structure estimation

To estimate the net long and net short curves, we need to construct the risk-neutral expectations of $x_{1,t+j}$ and $x_{2,t+j}$. This is where our assumption that interest rate swaps and cross-currency basis swaps are priced by a common SDF applies. The risk-neutral expectations we need are determined primarily by the swap curve and the term structure of CIP violations. We will proceed by constructing an SDF (in particular, a term structure model) that fits swap rates and CIP violations, and then use that SDF to construct the net long and net short curves.

At the heart of our calculations is a comparison between a Treasury hedged with a swap and a CIP violation of the same maturity. A rough version of this comparison can be done without a model: one simply compares the swap spread with the CIP violation. Our term structure model allows for a more careful version of this comparison. First, it allows us to consider Treasury bonds with maturities and coupon structures that do not exactly line up with the available points of the swap and CIP curves. Second, it allows us to explicitly account for the residual basis risk between financing rates. Third, it allows us to model the unwinding of the trade when the bond has six months remaining maturity (which helps fit the short-end of the yield curve). Fourth, it smooths the swap and CIP curves, reducing micro-structure-induced noise. Lastly, it accounts for covariances that are omitted from the naive spread calculation; that is, we use the exact, non-linear equations from the preceding section, as opposed to their linearized approximations.

Our term structure model largely follows the standard approach in the no-arbitrage term structure model literature (Joslin et al., 2011, 2014). At first glance, this might seem strange, given that our model necessarily features arbitrage. In particular, our term structure model must match both swap rates and CIP violations, which is equivalent to matching dollar swap rates and synthetic dollar swap rates. Our term structure model therefore features two different short rates, and as a result two different yield curves, as in Augustin et al. (2021). We use a no-arbitrage approach (as opposed to other methods of yield curve interpolation) on the grounds that such an approach is consistent with our assumption of no-arbitrage across derivatives.

We will use lower case symbols to denote scalars or vectors and uppercase symbols to denote matrices. Let z_t denote a vector of N risk factors (our empirical exercise will have $N = 5$) for our term structure model. We assume that, under the physical (actual) probability measure $\mathbb{P}$, z_t follows a Gaussian AR(1) process,

$$z_{t+1} = k^{\mathbb{P}}_{0,z} + K^{\mathbb{P}}_{1,z} z_t + (\Sigma_z)^{1/2} \epsilon^{\mathbb{P}}_{z,t+1}, \epsilon^{\mathbb{P}}_{z,t+1} \sim N(0, I_N), \tag{13}$$

where I_N is the $N \times N$ identify matrix, $k^{\mathbb{P}}_{0,z}$ is an $N \times 1$ vector of constants, $K^{\mathbb{P}}_{1,z}$ is an $N \times N$ matrix, and Σ_z is a symmetric positive semi-definite $N \times N$ matrix. The intermediary's log stochastic discount factor that prices derivatives is

$$m_{t+1} = -(\delta_0 + \delta_1^T z_t) - \frac{1}{2} \lambda_t^T \lambda_t + \lambda_t^T \epsilon^{\mathbb{P}}_{z,t+1}, \tag{14}$$

where $\lambda_t = (\Sigma_z^{-1})(\lambda_0 + \Lambda_1 z_t)$ is the price of risk associated with each shock. We will assume that the profits of derivatives trades are discounted using the OIS curve, consistent with industry practice.[10] That is, $r_t = \delta_0 + \delta_1^T z_t$, where r_t is the log of the fixed rate associated with a one-month overnight index swap.

This standard specification leads to a risk-neutral ($\mathbb{Q}$) measure dynamics for the state vector z_t,

$$z_{t+1} = k^{\mathbb{Q}}_{0,z} + K^{\mathbb{Q}}_{1,z} z_t + (\Sigma_z)^{1/2} \epsilon^{\mathbb{Q}}_{z,t+1}, \quad \epsilon^{\mathbb{Q}}_{z,t+1} \sim N(0, I_N), \tag{15}$$

where the parameters $k^{\mathbb{Q}}_{0,z}$ and $K^{\mathbb{Q}}_{1,z}$ are functions of the physical measure parameters and the SDF parameters. It also leads zero coupon-swap rates that are affine in the state vector,

$$r_{n,t} = -\frac{1}{n} \ln(E^{\mathbb{Q}}_t[\exp(\sum_{j=0}^{n-1} -r_{t+j})]) = \frac{1}{n}(a^{swap}_n + (b^{swap}_n)^T z_t), \tag{16}$$

where $r_{n,t}$ denotes log of the annual fixed rate associated with an n-month swap.

The log synthetic dollar rate is likewise assumed to be affine,

$$r^{syn}_t = \hat{\delta}_0 + \hat{\delta}_1^T z_t, \tag{17}$$

which generates the recursion

$$r^{syn}_{n,t} = -\frac{1}{n} \ln(E^{\mathbb{Q}}_t[\exp(\sum_{j=0}^{n-1} -r^{syn}_{t+j})]) = \frac{1}{n}(a^{syn}_n + (b^{syn}_n)^T z_t). \tag{18}$$

We assume that the rates $x_t = (x_{1,t}, x_{2,t}, y^{bill}_t)$ are affine functions of our state variable, where $x_{1,t}$ and $x_{2,t}$ are the discount rates associated with our net short and net long curves, and y^{bill}_t is the monthly log yield on a six-month Treasury bill. Specifically,

$$x_t = \gamma_0 + \Gamma_1 z_t + (\Sigma_x)^{\frac{1}{2}} \epsilon_{x,t}, \quad \epsilon_{x,t} \sim N(0, I_3). \tag{19}$$

These additional variables can be thought of as akin to the "macro factors" often included in standard term-structure models. The key assumption is that the measurement errors $\epsilon_{x,t}$ do not enter the SDF (and hence have the same distribution under $\mathbb{P}$ and $\mathbb{Q}$). The coefficients γ_0 and Γ_1 can be identified from regressions of these factors on the state variables.[11]

[10] The choice to use OIS rather than some other discount rate does not substantially impact our results, as we use the SDF only to price zero-NPV derivatives, whose value is not sensitive to shifts in the level of the discount rate.

[11] The tri-party repo and secured lending rates (which are essentially the long and short financing rates) are overnight rates. Given data limitations, we use overnight tri-party repo rates and overnight security lending rates to construct x_t. The 1-month CIP basis data are available, but to avoid the quarter-end effect (Du et al., 2018), we instead use the 3-month CIP basis to obtain the synthetic rate in x_t. Our estimation reveals that there are unit-root elements in the z_t process. A more sophisticated approach is to require that the spreads $x_{1,t}$ −

Finally, we assume that the dealer unwinds their position when the remaining maturity reaches six months, at which point the Treasury bond is equivalent to a Treasury bill. We make this assumption to capture the empirical observation that, pre-GFC, bill yields fell below other short-term risk-free rates (see e.g. Nagel (2016)). Incorporating this effect is important when constructing bounds on a one-year bond (which will become equivalent to a bill relatively soon) but has a minimal effect on long-maturity bonds. The curves we construct are thus

$$y^{l}_{n,t} = -\frac{1}{n}\ln(E^{\mathbb{Q}}_{t}[\exp(-\sum_{j=0}^{n-7} x_{1,t+j})\exp(-6y^{bill}_{t+n-6})]), \tag{20}$$

$$y^{s}_{n,t} = -\frac{1}{n}\ln(E^{\mathbb{Q}}_{t}[\exp(-\sum_{j=0}^{n-7} x_{2,t+j})\exp(-6y^{bill}_{t+n-6})]). \tag{21}$$

Under the assumptions of our term structure model, these curves are also affine in the state variables.

2.5. Term structure estimation results and predictions

We obtain data from various public sources, as documented in Internet Appendix Section A. We then estimate the term structure model to fit dollar swap rates and synthetic dollar swap rates, using a standard maximum likelihood approach. The goal of our estimation procedure is to accurately fit and interpolate these curves, and the model is flexible enough to fit both sets of rates with a high degree of accuracy. For details on the estimation procedure and on related issues such as coupon vs. zero-coupon bonds, and for illustrations of model fit, see Internet Appendix Section D.6.

In Fig. 4, we show the model-implied net long and net short curves, in comparison with the Treasury yield curve. We show both levels and their differences with the OIS curve. For brevity, we only show the standard 2-year and 10-year maturities. Graphs with a richer set of maturities are shown in Appendix Figures A7 and A8.

Several patterns are immediately apparent. Prior to the GFC, bond yields were close to the net short curve, consistent with the net position data of Fig. 1. After the GFC, bond yields are close to the net long yield, suggesting that dealers will be long coupon bonds. This is again consistent with the position data.

Both funding spreads and balance sheet costs play an important quantitative role in the model. Appendix Figure A9 presents decompositions of the net long and net short curves, in which one of these forces have been turned off. The short funding spread $r_t - i^{s}_{t}$ plays a critical role in the pre-GFC period, while balance sheet costs (as proxied for by CIP violations) are essential to explain the post-GFC period.

We next compare Treasury yields relative to the net short and net long curves with position data. The net short and net long curves we construct are estimates of yields at which the dealer should be willing to take a net position in a Treasury bond, after accounting for financing and balance sheet costs. This leads naturally the prediction that if the yield is at the net long yield, dealers should be net long, and if the yield is at the net short yield, dealers should be net short.

We first define a "relative yield index" for each maturity n by

$$pos_{n,t} = 2 * \frac{y_{n,t} - y^{s}_{n,t}}{y^{l}_{n,t} - y^{s}_{n,t}} - 1. \tag{22}$$

This index takes on a value of one if the yield of the n-month maturity Treasury bond is equal to the net long yield, negative one if it is equal to the net short yield, and zero if it is equal to the average of the net long and net short yields.

We then construct a weighted average (across maturities) version of the relative yield index, where the weight for each maturity is the fraction of dealer Treasury bond holding in that maturity bucket over total dealer Treasury bond holdings. We obtain the net primary dealer coupon Treasury bond position from the FR2004 primary dealer statistics published by the Federal Reserve Bank of New York, and normalize the total value by the total assets of primary dealers.

Fig. 5 compares the weighted average relative yield index together with total dealer Treasury holding scaled by dealer balance sheet size. The bond position and relative yield index are plotted on different axes (because they are not in comparable units), with the zero points on each axis aligned. For a variety of reasons, we do not expect these series to perfectly align. First, the mapping between how close a yield is to the net long or net short curves and the predicted net dealer position in that maturity is unclear, and may change over time. Second, the arbitrage bounds we construct are motivated by trading strategies that (at least potentially) hold the bond almost to maturity. Dealers also intermediate bonds between clients, and may be willing to buy a bond for the purpose of selling it quickly even if they view the bond as overpriced (close to the net short yield). This kind of intermediation activity acts as a kind of noise in the relationship between net dealer positions and the relative yield index, and is likely accentuated when looking at specific maturity buckets as opposed to overall net dealer positions. Despite these caveats, there is a non-trivial correspondence between the relative yield index and actual positions, as shown in Fig. 5. See Appendix Figure A10 for maturity-specific relative yield indices and their comparison with actual dealer positions across four maturity buckets.

Summarizing our analysis thus far, Treasury yields have moved from being close to net short arbitrage bounds pre-GFC to being close to net long arbitrage bounds post-GFC, and net primary dealer positions have responded by switching from being net short to net long. The relationship between yields and positions we document is consistent with the view that balance-sheet-constrained dealers act as arbitrageurs between the Treasury and swap markets.[12] We next consider the implications of this perspective, with an emphasis on the causes and consequences of the regime shift we have documented.

3. A model of the treasury market

Thus far, we have said little about how the Treasury regime is determined. In this section, we build a supply-and-demand model to endogenize dealers' net position, as a function of their balance sheet constraint, demand for Treasury bonds from non-dealers, and the overall supply of Treasury bonds. This model helps explain the change in the Treasury regime pre- and post-GFC, the striking correlation between the slope of the yield curve and the dealer position in Fig. 2, and fragility of the Treasury market when dealers' balance sheet constraints are tight.

3.1. Model setup

The model has two dates (zero and one), and a single n-period Treasury bond. Date one exists only for the purpose of determining payoffs; all of our analysis will focus on date zero, and we will omit time subscripts for all date zero rates and yields. We will take as exogenous the date zero log interest rates y^{bill}, i^{l}, i^{s}, and r, as well as two different expectations concerning future bond yields. We define the dealer's date zero risk-neutral ($\mathbb{Q}$-measure) expectation of date one bond yields as

$$y_{\mathbb{Q}} \equiv E^{\mathbb{Q}}[y_{n-1,1}] = E^{\mathbb{Q}}[r_{n-1,1} - (r_{n-1,1} - y_{n-1,1})] \tag{23}$$

$r_t - r^{cip}_{t}$ (approximately the tri-party repo and fed funds spread), $x_{2,t} - r_t + r^{cip}_{t}$ (approximately the security lending rate and fed funds spread), and $x_{3,t} - r_t$ (approximately the T-bill and fed funds spread) are stationary, i.e., have zero loading on the unit-root element. Our main approach is the direct regression of x_t on z_t, but we show in the internet appendix that results are similar if we impose stationarity on the spreads. See Internet Appendix Section D for more details.

[12] We interpret our results as showing that Treasury yields are often at or near arbitrage bounds given swap prices. However, one could equally say that swap yields are at or near arbitrage bounds given Treasury yields, adopting the perspective of Hanson et al. (2022).

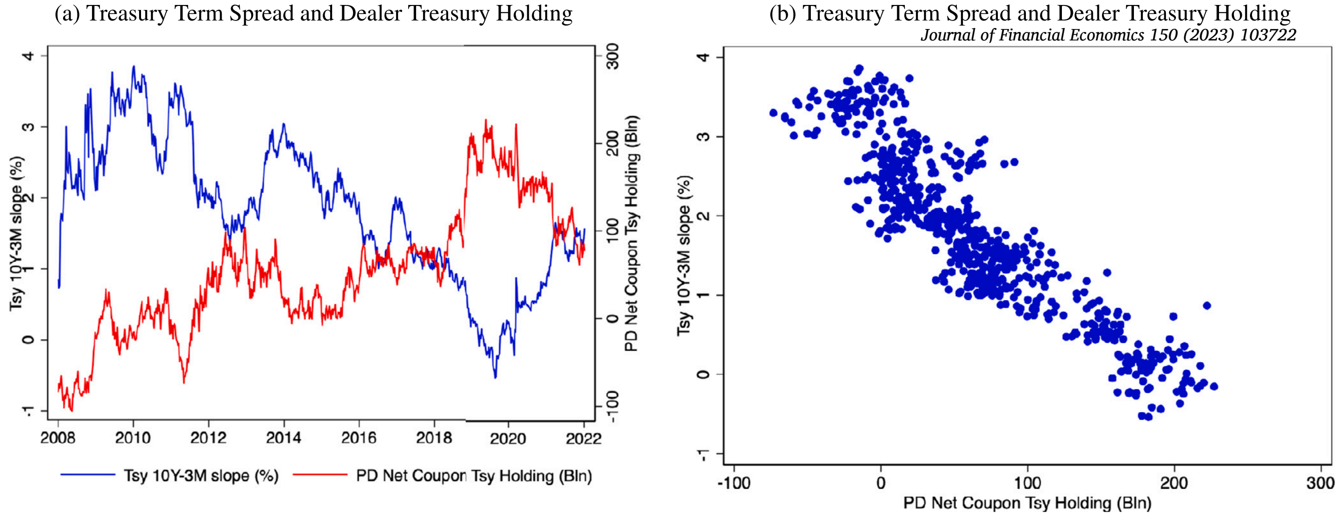

Fig. 2. Term Spreads and Primary Dealer Treasury Holdings.

where $r_{n-1,1}$ and $y_{n-1,1}$ denote the $(n-1)$-period log swap rates and log Treasury bond yields at date one. We define the corresponding physical measure ($\mathbb{P}$) counterpart as $y_{\mathbb{P}}$. Note that the dealer's SDF and associated risk-neutral measure $\mathbb{Q}$ price derivatives, but will generally not price positive-investment assets (such as Treasury bonds and bills).

We have written the definition of $y_{\mathbb{Q}}$ in this way to highlight the possible interpretations of comparative statics with respect to $y_{\mathbb{Q}}$. One interpretation, which we emphasize, is that a decrease in $y_{\mathbb{Q}}$ (holding all else, and in particular r, constant) represents a decrease in the swap term premium holding constant the risk-neutral expectation of future swap-Treasury spreads.[13] It is also important to distinguish between comparative statics that change $y_{\mathbb{Q}}$ holding $y_{\mathbb{P}}$ constant and comparative statics that change both $y_{\mathbb{P}}$ and $y_{\mathbb{Q}}$. We interpret the former as a change in swap term premia and the latter as a change in expected future rates.

These interest rates and expected bond prices will allow us to compute the dealer's net long and net short curves, using the approximations (3) and (4) defined in Section 2.1, with $y_{\mathbb{Q}}$ in the place of $E[y']$.[14] The key endogenous variables are y, the yield of an n-period zero-coupon bond at date zero, and r^{syn}, the one-period synthetic unsecured risk-free rate at date zero. We focus on the following asset prices that are closely related to the empirical motivations in Fig. 1.

- The Treasury term spread, $y - y^{bill}$.
- The synthetic–swap spread, $r^{syn} - r$, which maps to CIP deviations.
- The n-period swap spread $r_n - y$, with r_n approximated as

$$r_n \approx E^{\mathbb{Q}}[\frac{1}{n}r + \frac{n-1}{n}r_{n-1,1}]$$

- The swap term spread, $r_n - r$.

We will treat dealers and their levered clients as a single consolidated entity, based on the analysis of Internet Appendix Section F.1. In the Treasury market, dealers will interact with two kinds of investors. Hedged investors purchase the Treasury bond and swap it to their local currency. Unhedged investors choose between the Treasury bond and Treasury bills. Dealers also have other counterparties in the synthetic lending market; these other counterparties do not participate in the Treasury market. This structure implicitly assumes that the tri-party repo, bill, and interest rate swap markets are infinitely elastic, whereas the Treasury and synthetic borrowing markets are elastic but not infinitely so; we make these assumptions to simplify our exposition.

Our model of dealers is exactly that of Section 2.1, with the exception that we no longer require the simplification of risk-neutrality, and instead simply treat the $\mathbb{Q}$ measure as exogenous. Recall that q^{syn} is the quantity (in dollars) of synthetic loans supplied by the dealers at date zero, and q^{bond} be the quantity (in dollars) of bonds owned (positive) or short-sold (negative), and that the dealer balance sheet constraint is $q^{syn} + |q^{bond}| = \bar{q}$.

Let S^{bond} be the supply of bonds (in notional quantities) at date zero, and let D_U^{bond} and D_H^{bond} be the demand (in dollars) for bonds from unhedged and hedged investors, respectively. Market clearing in the bond market requires

$$q^{bond} + D_U^{bond} + D_H^{bond} = \exp(-ny)S^{bond}. \tag{24}$$

The bond price $\exp(-ny)$ enters this expression to convert the notional supply S^{bond} into dollars.

Unhedged investor demand is a continuously differentiable, strictly positive and increasing function of the expected log excess return of the bond over bills,[15]

$$D_U^{bond} = D_U(ny - y^{bill} - (n-1)y_{\mathbb{P}}). \tag{25}$$

Likewise, hedged investor demand is a continuously differentiable, strictly positive and increasing function of the expected log excess return in dollars[16] on the hedged strategy:

$$D_H^{bond} = D_H(ny - r^{syn} - (n-1)y_{\mathbb{P}}). \tag{26}$$

When a dealer helps a hedged investor exchange e.g. yen into dollars and hedge using forwards, the dealer will end up with a yen asset (cash) and a dollar liability. As a result, this activity increases the dealer's bal-

[13] That is, for the purpose of labeling the horizontal axis in Fig. 8 and for our subsequent discussion of policy, we assume a 1:1 relationship between changes in $y_{\mathbb{Q}}$ and changes in $E^{\mathbb{Q}}[r_{n-1,1}]$. An equally valid alternative interpretation of the comparative static with respect to $y_{\mathbb{Q}}$ is as a change in the risk-neutral expectation of future swap-Treasury spreads holding the swap curve constant.

[14] We use these approximations to simplify the analysis that follows, at the suggestion of an anonymous referee. A previous version of this paper used exact (non-linear) definitions of the buy and sell curves, and reached identical conclusions.

[15] As documented in Haddad and Sraer (2020), typical bank portfolios behave like unhedged investors in our model in that their position in long-term bonds is increasing the expected excess return of long-term Treasury bonds.

[16] For simplicity, we use the hedged return in dollars, as opposed to in local currency; this allows us to ignore second-order terms associated with the covariance between interest rates and exchange rates.

ance sheet, and is functionally equivalent, from the dealer's perspective, to lending synthetic dollars.

There is an important role for market segmentation in these equations. Both types of Treasury clients are assumed to care about $y_{\mathbb{P}}$ and not $y_{\mathbb{Q}}$, because they trade Treasury bonds but not swaps. Segmentation between constrained agents (dealers) and unconstrained agents (clients), whether endogenous (e.g., Alvarez and Jermann (2000); Chien et al. (2011); Biais et al. (2021)) or exogenous (e.g., Gertler and Kiyotaki (2010); He and Krishnamurthy (2013)), is necessary to generate Treasury-swap arbitrage.

Dealers can also lend synthetic dollars to other counterparties (hedged investors buying corporate bonds, for example). We assume that the demands of these other investors for synthetic dollars are $D^{syn}(r^{syn}-r)$, where D^{syn} is a continuously differentiable, non-negative and strictly decreasing function of the spread between synthetic dollars and risk-free rates. Note that $D^{syn}(\cdot)$ is decreasing in its argument (a spread), because the synthetic lending counterparties are paying that spread to dealers. In contrast, the other demand functions, D_H^{bond} and D_U^{bond}, are increasing in their arguments (expected returns), because investors are earning those expected returns. Market clearing in the synthetic dollar market requires that

$$q^{syn} = D_H^{bond} + D^{syn}(r^{syn}-r). \tag{27}$$

These market clearing conditions, together with the net long and net short bounds of our simple model (equations (3) and (4)) and their associated implications for dealer positions, define our model. To guarantee that an equilibrium exists in our model, we make the following assumptions.

Assumption 1. *We assume that the demand functions are well-behaved:*

- *Excess synthetic loan demand is possible:* $D^{syn}(0) > \bar{q}$.
- *Excess synthetic loan supply is possible: For all* y, $\lim_{r^{syn}\longrightarrow\infty} D^{syn}(r^{syn}-r) + D_H(ny - r^{syn} - (n-1)y_{\mathbb{P}}) = 0$.
- *Excess bond supply is possible:* $\lim_{y\longrightarrow-\infty} D_U(ny - y^{bill} - (n-1)y_{\mathbb{P}}) + D_H(ny - r - (n-1)y_{\mathbb{P}}) = 0$.

Note that excess bond demand is also possible without additional assumptions.[17]

Our static model features three possible equilibria, which we refer to as regimes: a long regime $q^{bond} > 0$, a short regime $q^{bond} < 0$, and an intermediate regime $q^{bond} = 0$.[18] We discuss the long and short regimes (which are of greater empirical interest) in the main text, and provide results for the intermediate regime in Internet Appendix Section B.

Our focus, when analyzing these regimes, will be on the spread between synthetic dollars and the swap rate, $r^{syn}-r$, which is endogenously determined. In our quantitative analysis, this spread was a key input to the model, and we measured it with CIP violations. In this two-market market, the rate r^{syn} should be understood as the risk-free return the dealer requires for assets held on balance sheet, as opposed to specifically a euro rate swapped to dollars. Under this interpretation, the spread $r^{syn}-r$ is the kind of financial intermediation spread that plays a key role in macro-finance models (Brunnermeier and Pedersen, 2009; He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; Gertler and Kiyotaki, 2010).

3.2. The long regime

We first consider a regime that dealers are long ($q^{bond} > 0$), in which case the Treasury yield y is equal to the dealer long curve y^l as in (3),

$$y = \underbrace{\frac{n-1}{n} y_{\mathbb{Q}} + \frac{1}{n}(i^l + r^{syn} - r)}_{\equiv y^l}. \tag{28}$$

This equation generates a dealer indifference condition between the two endogenous variables r^{syn} and y, with y strictly increasing in r^{syn}. As synthetic lending spread increases (higher r^{syn}), the dealer requires a better return on buying the Treasury bond (higher y).

Combining the market clearing conditions in (24) and (27), the intermediary balance sheet constraint in (1), and using $q^{bond} > 0$, we have

$$\bar{q} - e^{-ny} S^{bond} + D_U(ny - y^{bill} - (n-1)y_{\mathbb{P}}) = D^{syn}(r^{syn}-r). \tag{29}$$

The left-hand side of (29) represents the residual balance sheet available for synthetic lending, which in equilibrium must equal the residual demand for synthetic lending. Note that the Treasury bond demand from hedged investors does not appear in this equation, because the dealer balance sheet is unaffected by changes in their demand, holding all else constant. An increase in hedged investor demand reduces the quantity of Treasury bonds dealers must hold, but increases the amount of synthetic borrowing they must finance, and hence has no effect on their balance sheet usage, provided that dealers have a net long bond position. The left-hand side of (29) is strictly increasing in y, while the right-hand side is strictly decreasing in r^{syn}. Therefore, equation (29) generates a kind of market indifference condition, where r^{syn} strictly decreases with y. In Fig. 6, we illustrate this market indifference condition in a plot of the synthetic lending spread $r^{syn}-r$ against the Treasury term spread $y - y^{bill}$. Intuitively, a higher synthetic dollar lending spread leads to more dealer balance sheet allocation to synthetic dollar lending but less to bonds. This decline of bond demand from dealers increases the bond yield.

A long-regime equilibrium (y, r^{syn}) is a point where the dealer indifference curve in (28) intersects with the market indifference curve in (29) and $q^{bond} > 0$, which requires

$$D_H(ny - r^{syn} - (n-1)y_{\mathbb{P}}) + D_U(ny - y^{bill} - (n-1)y_{\mathbb{P}}) < e^{-ny} S^{bond}. \tag{30}$$

Because the two indifference curves have opposite slopes, such an equilibrium is unique if it exists, as illustrated in Fig. 6.

We next consider various comparative statics in a long-regime equilibrium.

Proposition 1. *In a long regime equilibrium, holding all else constant,*

1. *An increase in* S^{bond} *leads to an increase in* y *and an increase in* r^{syn},
2. *A decrease in* $\bar{q}$ *or a parallel decrease in* D_U *leads to an increase in* y *and an increase in* r^{syn},
3. *An increase in* $y_{\mathbb{Q}}$ *leads to an increase in* y *and a decrease in* r^{syn},
4. *A parallel increase in* $D^{syn}(\cdot)$ *increases both* y *and* r^{syn}.

Proof. See Internet Appendix Section E.1. □

Intuitively, an increase in the bond supply S^{bond} means that, holding yields constant, less dealer balance sheet is available for synthetic lending, causing a higher synthetic lending spread. As a result, the market indifference curve in Fig. 6 shifts upwards, which leads to an increase in both synthetic lending spreads and bond yields in equilibrium.

A decrease in either dealer balance sheet capacity $\bar{q}$ or in unhedged bond demand D_U is equivalent to an increase in the supply of bonds S^{bond}, and hence generates the same comparative statics.

An increase in $y_{\mathbb{Q}}$ holding $y_{\mathbb{P}}$ constant is an increase in swap term premium, which shifts the dealer indifference curve to the right in Fig. 6. This leads to an increase in y and a decrease in r^{syn}.

[17] If we take $y \to \infty$, the right hand side of (24) becomes zero while the non-dealer demand $D_U^{bond} + D_H^{bond}$ is strictly positive, causing an excess bond demand.

[18] In position data, dealers will never have an exactly zero net Treasury position, for reasons (for example, market-making activities) that are outside the scope of our model. We view the intermediate regime as describing a situation in which dealers are targeting a roughly net flat position.

Lastly, a parallel increase in the synthetic dollar demand $D^{syn}(\cdot)$ shifts the market indifference curve in Fig. 6 upwards, thus increasing both y and r^{syn}.

3.3. The short regime

We next consider a regime in which dealers are short ($q^{bond} < 0$). In this regime, the bond yield y must equal to the net short yield y^s as defined in (4),

$$y = \underbrace{\frac{n-1}{n} y_{\mathbb{Q}} + \frac{1}{n}(i^s - r^{syn} + r)}_{\equiv y^s}. \tag{31}$$

This equation generates a dealer indifference condition, where r^{syn} strictly decreases with y. The less attractive it is to sell the Treasury bond (higher y), the lower synthetic lending rate r^{syn} must be to generate indifference between these two activities. This relationship has the opposite sign compared to the long regime.

Combining the market clearing conditions in (24) and (27), balance sheet constraint (1), and using $q^{bond} < 0$, we obtain

$$\begin{aligned} &\bar{q} + e^{-ny} S^{bond} - D_U(ny - y^{bill} - (n-1)y_{\mathbb{P}}) - 2D_H(ny - r^{syn} - (n-1)y_{\mathbb{P}}) \\ &\quad = D^{syn}(r^{syn} - r). \end{aligned} \tag{32}$$

The left-hand side of (32) again represents the residual balance sheet available for synthetic lending, which in equilibrium must equal the residual demand for synthetic lending. Note that demand from hedged investors has a double impact on the dealer's balance sheet. All else equal, an increase in this demand will result in dealers taking larger short positions and providing more synthetic financing to hedged investors, both of which use up dealer balance sheet.

The left-hand side of (32) is strictly decreasing in y: unlike the long regime, more demand from investors and less supply require dealers to take larger short positions, using up more balance sheet. Equation (32) therefore generates a kind of market indifference condition, where r^{syn} strictly increases in y. In Fig. 7, we illustrate this market indifference condition together with the dealer indifference condition. The key difference from the long-regime illustration in Fig. 6 is that the slopes flip for both the dealer indifference curve and the market indifference curve.

A short-regime equilibrium (y, r^{syn}) is a point where the two indifference curves intersect and $q^{bond} < 0$, which requires

$$D_H(ny - r^{syn} - (n-1)y_{\mathbb{P}}) + D_U(ny - y^{bill} - (n-1)y_{\mathbb{P}}) > e^{-ny} S^{bond}. \tag{33}$$

Because the two indifference curves have opposite slopes, such an equilibrium is again unique if it exists, as illustrated in Fig. 7.

We study the same set of comparative statics, now in the context of the short regime.

Proposition 2. *In a short regime equilibrium, holding all else constant,*

1. *An increase in S^{bond} leads to an increase in y and a decrease in r^{syn};*
2. *An increase in $\bar{q}$ or a parallel decrease in D_U leads to an increase in y and a decrease in r^{syn};*
3. *An increase in $y_{\mathbb{Q}}$ leads to an increase in y and an increase in r^{syn};*
4. *A parallel increase in $D^{syn}(\cdot)$ increases r^{syn} and decreases y.*

Proof. See Internet Appendix Section E.2. □

Intuitively, an increase in bond supply S^{bond} in the short regime means that, holding the bond yield constant, more dealer balance sheet is available for synthetic lending, causing a lower synthetic lending spread. The market indifference curve in Fig. 7 shifts downwards, which leads to an increase in the bond yield (as in the long regime) but a decrease in the synthetic lending spread (unlike the long regime).

A decrease in unhedged bond demand D_U is equivalent to an increase in bond supply, and hence generates the same comparative statics. Unlike the long regime, an increase (rather than decrease) in dealer balance sheet capacity $\bar{q}$ is equivalent to an increase in supply S^{bond}, because dealers are short bonds instead of long bonds.

An increase in $y_{\mathbb{Q}}$ holding $y_{\mathbb{P}}$ constant shifts the dealer indifference curve to the right in Fig. 7, which leads to an increase in y (like the long regime) but a decrease in r^{syn} (unlike the long regime).

Lastly, a parallel increase in the synthetic dollar demand $D^{syn}(\cdot)$ shifts the market indifference curve in Fig. 7 upwards, thus increasing r^{syn} (like the long regime) but decreasing y (unlike the long regime).

3.4. Equilibrium

We next discuss several factors that can determine which of three regimes occur in equilibrium.

Proposition 3. *The equilibrium exists and is unique given the exogenous parameters. Holding all else constant,*

1. *There exists an $0 \le S^s \le S^l \le \infty$ such that a short regime equilibrium exists for all $S^{bond} < S^s$, a long regime equilibrium exists for all $S^{bond} > S^l$, and an intermediate regime equilibrium exists for $S^{bond} \in [S^s, S^l]$.*
2. *There exists an $-\infty \le y^l_{\mathbb{Q}} \le y^s_{\mathbb{Q}} \le \infty$ such that a short regime equilibrium exists for all $y_{\mathbb{Q}} > y^s_{\mathbb{Q}}$, a long regime equilibrium exists for all $y_{\mathbb{Q}} < y^l_{\mathbb{Q}}$, and an intermediate regime equilibrium exists for $y_{\mathbb{Q}} \in [y^l_{\mathbb{Q}}, y^s_{\mathbb{Q}}]$.*

Proof. See Internet Appendix Section E.4. □

Intuitively, when bonds are scarce, dealers will be short to meet client demands, while when bonds are abundant dealers will be long to fill the shortfall in client demand.

Less intuitively, when the term premium is large and the yield curve steep, dealers will be short bonds. Considering the dealer's bond position in isolation, this looks like a money-losing strategy. However, if the swap curve has a higher term premium than the Treasury curve and the dealer is hedging with swaps, then selling bonds and hedging is in fact a profitable strategy. Similarly, buying bonds when the curve is flat looks like a money-losing strategy, but is in fact profitable if the swap curve is even flatter and the dealer hedges. In our model, client demand for Treasury bonds is driven by the expected returns on those bonds; spreads must therefore move in a way that induces dealers to take the opposite position.

Fig. 8 illustrates the comparative statics of the model with respect to changes in the swap curve term premium, interpreted as a change in $y_{\mathbb{Q}}$ holding fixed $y_{\mathbb{P}}$. The long regime appears when $y_{\mathbb{Q}}$ is sufficiently low but the short regime appears when $y_{\mathbb{Q}}$ is sufficiently high. The swap-Treasury spread is an increasing function of the swap term premium, but the rate at which it increases depends on the regime. In contrast, the synthetic lending spread is U-shaped: it is high in both the long and short regimes, and low in the intermediate regime. Lastly, the dealer Treasury position as a fraction of its capacity $\bar{q}$ decreases with the swap term premium, consistent with the motivating facts in Fig. 2.

3.5. Model vs. data

We now return to the motivating facts illustrated in Figs. 1 and 2 and discuss them through the lens of our model. We first discuss the relationship between the yield curve slope and dealers position. We then discuss drivers of the regime change and examine specific features of the pre-GFC and post-GFC regimes.

Yield curve slope vs. dealer position The model is able to explain why dealers' net long position is larger when the term spread (slope of the yield curve) is lower. This pattern is documented in Fig. 2, and is at

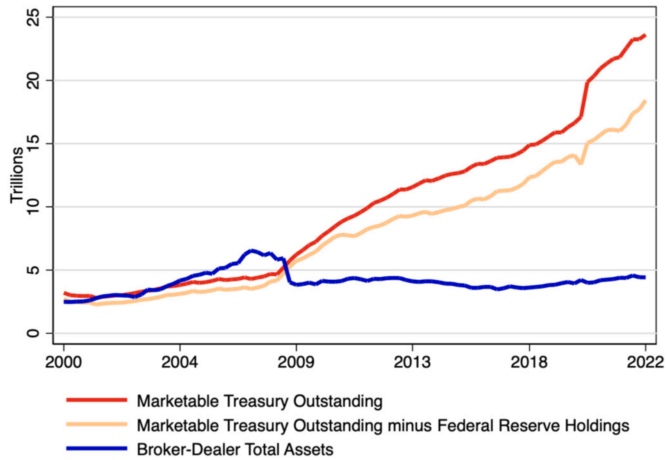

Fig. 3. Treasury Supply and Broker-Dealer Total Assets. *Notes*: This figure plots the total marketable Treasury securities outstanding (in red), the total marketable Treasury outstanding minus Federal Reserve holdings (in orange), and the financial assets of the U.S. broker-dealer sector (in blue) in trillions of dollars from Flow of Funds.

first puzzling in light of the fact that the slope positively predicts future bond returns (Campbell and Shiller (1991)). That is, if dealers held more Treasury bonds when the expected returns on those bonds are higher (as in Jermann (2020)), we would expect a positive relationship between bond holdings and the term spread.

Our model resolves the puzzle by assuming that client demand is increasing in expected excess bond returns. Dealers accommodate client demand by acting as arbitrageurs between the bond and swap markets, which leads naturally to the result that their Treasury positions declines in expected excess bond returns, and hence the term spread. Dealers justify this larger position by charging a larger swap spread.

Treasury supply, pension swap demand, and the sign switch in the swap spreads Since the GFC, the supply of Treasury securities has risen significantly. In our model, an increase in the supply of Treasury securities, without being met by real-money demand, can result in dealers switching from the net-short to net-long regime (Proposition 3 part 1), which leads to the sign-switch of the swap spread. In addition to the increase in the Treasury supply, an alternative (and not mutually exclusive) hypothesis is that an increased demand from under-funded defined-benefit (DB) pension to receive fixed and pay floating in the interest rate swap market caused the 30-year swap spread to turn negative post-GFC (Klingler and Sundaresan, 2019). In our model, swap demand enters as a decline in the swap term premium i.e. a decline in $y_{\mathbb{Q}}$. Specifically, swap demand will only affect swap spreads if dealers are risk-averse or face risk-based constraints. Under these assumptions, swap demand will affect their SDF and its associated risk-neutral measure, $\mathbb{Q}$, and hence enter our model via $y_{\mathbb{Q}}$.

Thorough the lens of our model, an increase in the Treasury supply and an increase in the interest rate swap demand, other things being equal, have the same effects on dealers' Treasury position, because both changes require dealers to increase their holdings of Treasury bonds and their exposure to interest rate swaps by paying fixed and receiving floating to hedge duration risk. Which channel plays a more dominant role in determining the regime is an empirical question.

Fig. 3 illustrates a sharp rise in the supply of U.S. Treasury securities since GFC. The outstanding of marketable Treasury securities grew from \$4.7 trillion in 2008 to \$22.5 trillion as of June 2022. This resulted in a large increase in the supply of Treasury bonds, even after accounting for

Table 1
OIS-Treasury Spreads, Pension Fund Underfunding, and Treasury Supply.

	Dependent variable: Δ(OIS-Tsy Spread)$_t$ of maturity			
	30Y	10Y	5Y	2Y
ΔUFR$_t$	-0.685^{**}	-0.175	0.163	-0.175
	(0.339)	(0.279)	(0.251)	(0.233)
Δlog(Net Tsy$_t$)	-1.405^{***}	-1.156^{***}	-0.970^{***}	-0.702^{***}
	(0.364)	(0.300)	(0.271)	(0.251)
Constant	0.018	0.016	0.013	0.012
	(0.013)	(0.011)	(0.010)	(0.009)
Observations	76	76	76	76
Adjusted R^2	0.218	0.165	0.126	0.092

Notes: This table reports regressions of quarterly changes in the OIS-Treasury spread for different maturities on the quarterly changes of dependent variables. Data are at quarterly frequency from 1996 Q4 (constrained by the availability of OIS data) to 2015 Q4 (same as Klingler and Sundaresan (2019)). ΔUFR$_t$ is the change in the underfunding ratio of private and local government defined-benefit pension funds, constructed in the same way as Klingler and Sundaresan (2019). The calculation methods from flow of funds changed from an accumulated benefit obligation basis to a projected benefit obligation basis on September 2018, and we use the updated data. This data update has small effects on the regression coefficients. $\Delta\log(\text{Net Tsy})_t$ is the growth of net Treasury supply, defined as total marketable Treasury securities minus Federal Reserve holdings. Numbers in parentheses are standard errors. Statistical significance: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

the Federal Reserve's bond purchases.[19] At the same time, the total size of the broker-dealer balance sheet declined relative to its pre-GFC peak. Meanwhile, as discussed in Klingler and Sundaresan (2019), defined-benefit pensions became increasingly underfunded post-GFC.[20] Klingler and Sundaresan (2019) argue that the underfunded ratio (UFR) is a proxy for pension swap demand (which is not directly observed).

To examine the effects of Treasury supply and defined benefit pensions interest rate swap demand on the swap spread, in Table 1, we regress changes in the OIS-Treasury swap spreads on changes in the UFR of the defined benefit pensions and the growth of Treasury supply. These regressions follow the specification of Klingler and Sundaresan (2019); the key changes are the use of OIS instead of LIBOR swap spreads and the inclusion of the growth in net Treasury supply.

Consistent with Klingler and Sundaresan (2019), the pension demand channel is relevant for 30yr swaps but much less so for other maturities, suggesting some segmentation of long term rates (on this point, see Greenwood and Vissing-Jorgensen (2018)). However, OIS swap spreads of essentially all maturities switched sign at the time of the GFC, as illustrated in Fig. 4.[21] In contrast, the growth of Treasury supply consistently has a large and significant explanatory power on OIS swap spreads of all maturities, consistent with Treasury supply as a driving force of the swap-Treasury spread. Including the growth of Trea-

[19] The Federal Reserve's holdings of Treasury securities grew by about \$5tn between the beginning of 2008 and June 2022, absorbing about 30% of the increase in Treasury supply.

[20] In September 2018, the Bureau of Economic Analysis and Flow of Funds changed the method used to estimate state and local government pension liabilities to a projected benefit obligation basis; previously, they were presented on an accumulated benefit obligation basis. In the original data of Klingler and Sundaresan (2019), pensions were overfunded pre-GFC and underfunded post-GFC. In the updated data, they were underfunded pre-GFC and became more so post-GFC.

[21] See Appendix Figure A8 for more maturities. Klingler and Sundaresan (2019) focus on LIBOR swaps, which exhibit negative swaps spreads only at the longest maturities. LIBOR swaps are affected by credit risk and (in the post-GFC period) by uncertainty related to the transition away from the LIBOR benchmark, which may explain why they differ substantially from OIS swaps in the post-GFC era.

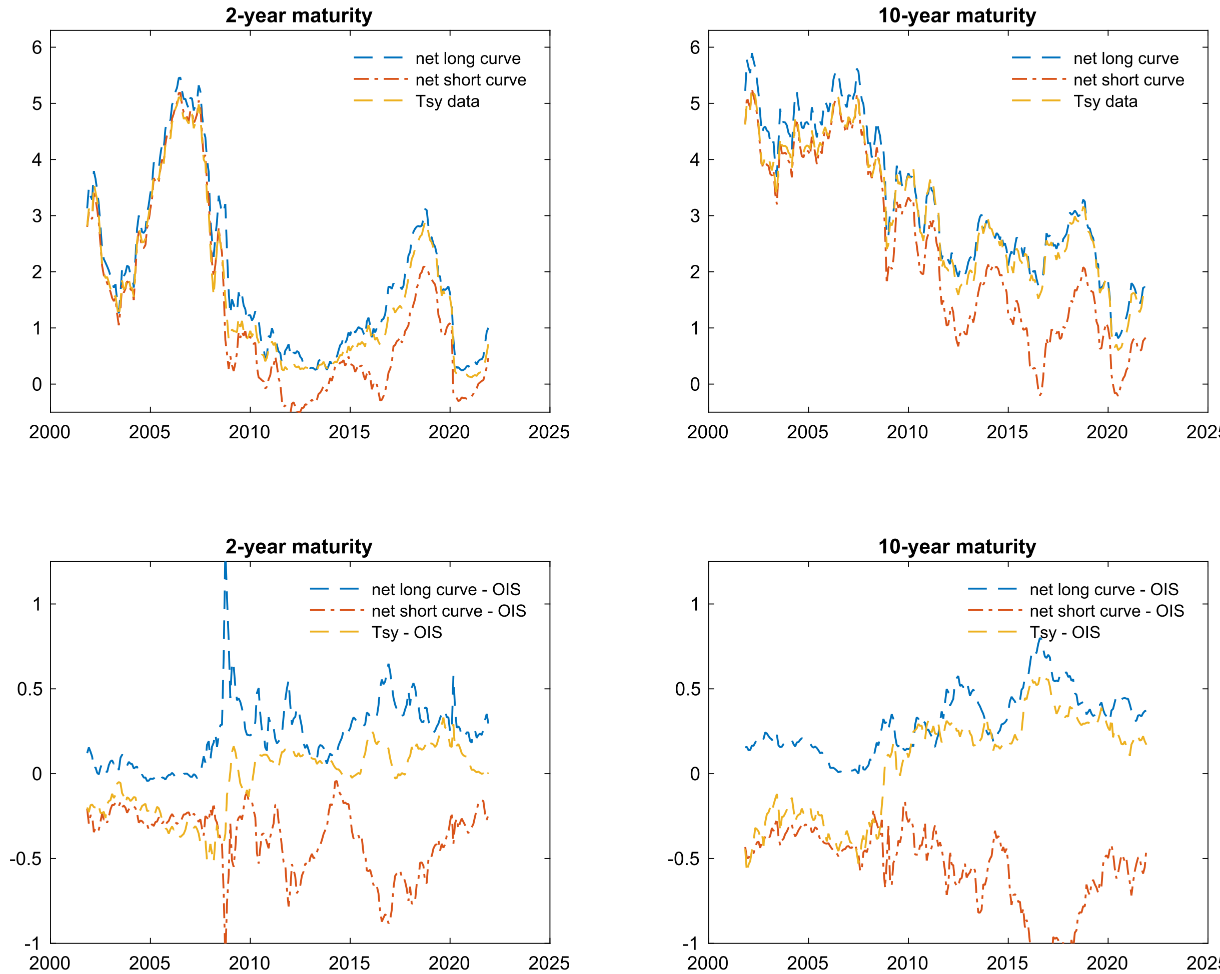

Fig. 4. Long and Short Yields. *Notes*: In this figure, we show the model-implied net-long and net-short curves for 2yr and 10yr Treasury coupon bonds, together with the actual Treasury yields. We plot levels (top row) and their differences with respect to OIS rates of comparable maturities (bottom row). Data are from 2003 to 2021. All yields are par yields. More details on the term structure model can be found in Section 2.4.

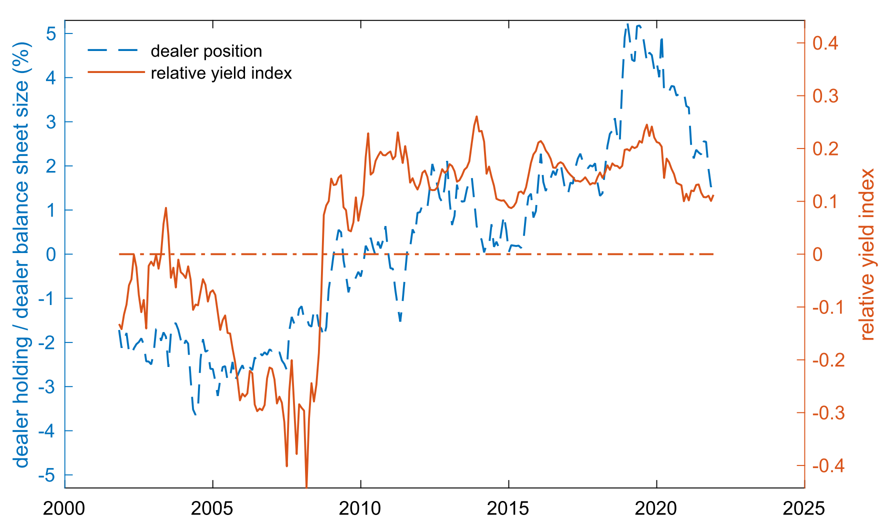

Fig. 5. Relative Yield and Position in Aggregate. *Notes*: In this figure, we plot the weighted average of the relative yield index across maturities vs. primary dealer Treasury positions. For each maturity, the relative yield index is defined as 2*(long curve − Treasury curve)/(long curve − short curve) − 1. Each relative yield is then weighted by the dealer sector's net coupon Treasury holdings for the corresponding maturity bucket. The primary dealer Treasury position is the ratio of total primary dealer net position in all coupon Treasury securities (published by Federal Reserve Bank of New York) to total financial assets of the broker-dealer sector (published by Flows of Funds). Data are from 2003 to 2021.

Journal of Financial Economics 150 (2023) 103722

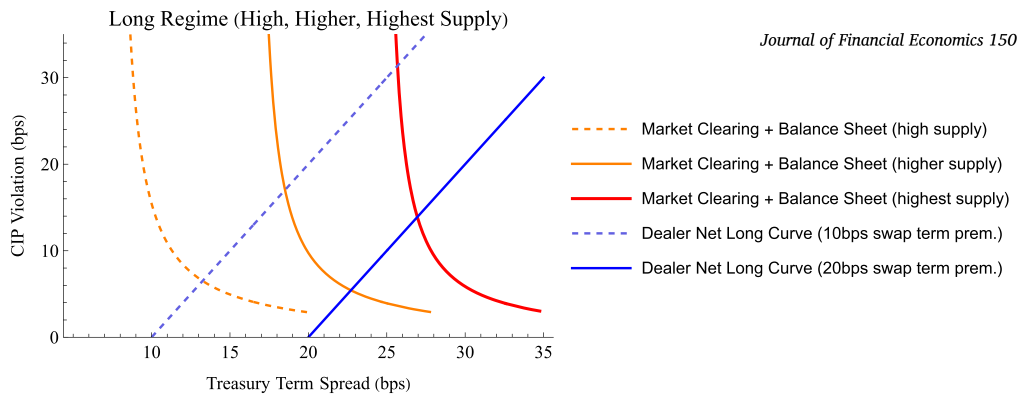

Fig. 6. Indifference Curves in the Long Regime. *Notes*: This figure illustrates the dealer indifference curve defined in (3) and the market indifference curve defined in (29), under three different levels of S^{bond} and two different levels of y_{Q}. The functional forms and parameters used to generate the figure are described in Internet Appendix Section C.

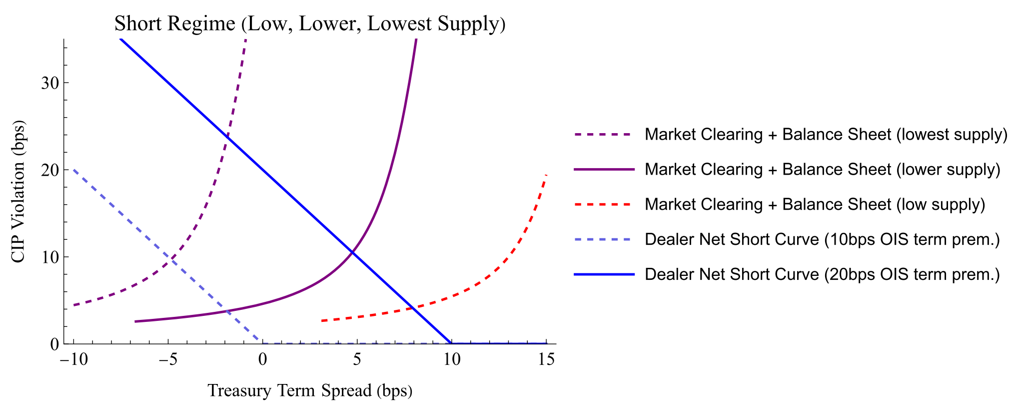

Fig. 7. Indifference Curves in the Short Regime. *Notes*: This figure illustrates the dealer indifference curve defined in (4) and the market indifference curve defined in (32), under three different levels of S^{bond} and two different levels of y_{Q}. The functional forms and parameters used to generate the figure are described in Internet Appendix Section C.

sury supply as an explanatory variable increases the R^2 of a regression on 30-year OIS-Tsy spread from 7% to 22%, and the R^2 of a regression on 10-year OIS-Tsy spread from almost zero to 17%. The significant explanatory power of Treasury supply does not diminish as we introduce additional control variables (see Appendix Table A2 for details).

For these reasons, we emphasize Treasury supply as a key driver of the regime change, although we cannot rule out a role for swap or Treasury demand. One implication of this view is that Treasury yields are substantially higher (relative to swap rates) than they could have been had the government not induced a Treasury market regime change via its issuance. Specifically, holding fixed balance sheet costs at their post-GFC levels, the gap between the net long and net short curves is often 100 basis points or more. Note that Treasury bonds are a small part of dealers' overall balance sheets, which is why we view this counterfactual as reasonable.

Pre-GFC dealer-short regime Prior the GFC, CIP violations were close to zero, swap spreads were positive, and dealers were net short in the Treasury market (Fig. 1). Our model reconciles these facts in the following way. First, since dealers had ample balance sheet capacity, CIP violations remained small (Proposition 2 part 2). Second, because dealers were net short, Treasury yields ended up on the net short curve, which is below the OIS swap curve due to funding spreads ($i^s < r$) and low Treasury bill yields (which lead to a lower y_{Q}). Consequently, swap spreads were positive at all maturities. Appendix Figure A9 shows that the funding spreads drive most of the positive swap spread pre-GFC. The T-bill convenience and balance sheet costs play minor role in explaining the pre-GFC swap spread.

Overall, given dealers' short position, ample dealer leverage, and low Treasury bill yields (relative to other rates), our model can replicate the stylized facts illustrated in Fig. 1 for the pre-GFC period.[22] Our model also allows us to consider the counter-factual scenario: what if balance sheet constraints were tight pre-GFC. In the dealer-short regime, a higher balance sheet cost would reduce the Treasury yield further, so that swap spreads would become even more positive pre-GFC.

Post-GFC dealer-long regime After the GFC, CIP violations were significant, swap spreads were negative, and there was a strong correlation between the two (Fig. 1). Dealers were net long, and the size of their position was strongly negatively correlated with the slope of the yield curve (Fig. 2). Dealers faced tight leverage constraints as part of the post-GFC regulatory regime (Duffie (2017)).

Our model reconciles these facts in the following way. First, because dealers had tight balance sheet constraints, CIP violations became large (Proposition 1 part 2). Given that dealers were net long, Treasury yields ended up on the net long curve, which is above the OIS swap curve

[22] In Internet Appendix Section C Figure A3, we replicate Fig. 8, with a relatively high balance sheet capacity for dealers, relatively low bill yields, and a relatively small supply of Treasury bonds. Our model, with these parameters, predicts a small and roughly constant CIP spread, a positive swap spread, and a negative dealer Treasury position.

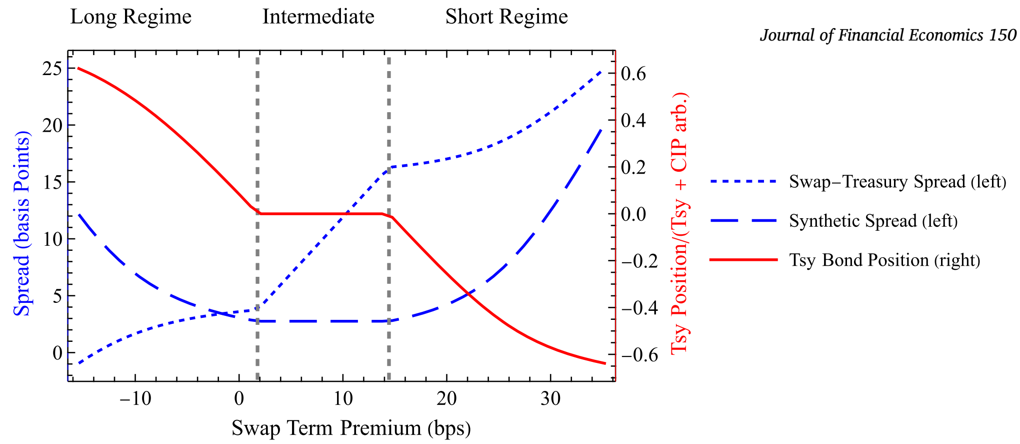

Fig. 8. All-Regimes and the Swap Term Premium. *Notes*: This figure plots the synthetic lending spread $r^{syn} - r$, swap-Treasury spread $r_n - y$, and the ratio of Treasury position over total Treasury and CIP arbitrage position $q^{bond}/\bar{q}$, as a function of the swap curve term premium $r_n - r$, holding fixed expected future rates. The functional forms and parameters used to generate the figure are described in Internet Appendix Section C.

mainly due to balance sheet costs.[23] Consequently, swap spreads are negative at all maturities. Moreover, because swaps spreads are now driven primarily by balance costs, they are now strongly correlated with other measures of balance sheet costs (CIP violations). Appendix Figure A9 confirms the importance of the balance sheet costs in driving the swap spread post-GFC. Thus, given a large bond supply, limited dealer leverage, and Treasury bill yields comparable to other rates, our model can replicate the stylized facts illustrated in Figs. 1 and 2 for the post-GFC period.[24]

3.6. Regimes and treasury market fragility

During the financial crisis of 2008-2009, Treasury yields fell by more than matched maturity swap rates. During the COVID-induced financial turmoil of March 2020, the reverse was true: Treasury yields did not fall by as much as swap rates, and in fact briefly rose (Duffie, 2020; Haddad et al., 2021; He et al., 2022). The different comparative statics across the long and short regimes in our model offer an explanation for this pattern.

In the short regime, an increase in balance sheet costs (as measured by the spread $r^{syn} - r$), all else equal, will lead to lower Treasury yields (Proposition 2). In contrast, in the long regime, an increase in balance sheet costs will lead, again all else equal, to higher Treasury yields (Proposition 1). Both crises were characterized by large increases in arbitrage spreads; the difference was that the market was in the short regime pre-GFC and in the long regime post-GFC.

He et al. (2022) attribute the differences between these two episodes to client demand for Treasury bonds (a dash-for-cash in COVID, a flight-to-safety in the GFC). Our story is compatible with theirs, in the sense that Treasury selling in COVID would increase balance sheet costs (the long regime) and Treasury buying in the GFC would also increase balance sheet costs (the short regime). However, our story does not rely customer demand for Treasury bonds as the causal factor behind the increase in balance sheet costs. In both the GFC and COVID episodes, even if clients had not bought or sold Treasury bonds on net, we expect that balance sheet costs would have risen, and as a result predict that Treasury yields would have moved (relative to swap rates) upwards in COVID but downwards in the GFC. Quantifying the role of Treasury demand, as opposed to other forces, is an interesting direction for future research.

[23] The funding spread $i^l < r$ is small, and in the post-GFC regime bill yields are close to r. As a result, the OIS-Treasury spread is primarily a function of balance sheet costs.

[24] In Internet Appendix Section C Figure A4, we replicate Fig. 8, but with a low balance sheet capacity for dealers, no spread between bill and funding rates, and a relatively large supply of Treasury bonds, as a way of capturing the post-GFC period. Our model, with these parameters, predicts significant CIP violations, negative swap spreads, co-movement between these two in response to shocks, and a positive dealer Treasury position.

Table 2
Summary of Policy Implications.

Policy Type	Long Regime		Short Regime	
	Tsy Yield	Lending Rate	Tsy Yield	Lending Rate
↓ Term premium	↓	↑	↓	↓
↑ Debt supply	↑	↑	↑	↓
Quantitative tightening	↑	↑	↑	↓
Swap lines	↓	↓	↑	↓
SLR exemptions	↓	↓	↑	↓

Notes: Tsy yield is the long-term Treasury yield y. Lending rate is the synthetic lending rate r^{syn}.

4. Implications for policy

In this section we consider a variety of fiscal, monetary, and regulatory policies, and study how the effects of those policies depend on the regimes we have identified in the Treasury market. We will emphasize the way in which the Treasury market regime determines the direction of the effects of these policies on Treasury yields and synthetic dollar rates; drawing quantitative conclusions would require estimating client demand curves. As discussed above, we view synthetic dollar rates as a proxy for financial intermediation spreads more generally.

All of the policies we consider have effects on inflation, real economic activity, and financial stability that are outside the scope of our model. Policies that increase arbitrage spreads and other financial market distortions can be justified on these grounds. However, combinations of the policies we discuss might achieve these same objectives while avoiding financial market distortions, and the goal of our analysis is to highlight these possibilities.

We view interest rate policy as shifting short term rates (r, i^s, i^l, y^{bill}) and the expectations of future rates ($y_{\mathbb{Q}}$ and $y_{\mathbb{P}}$). We will first discuss the effects of interest rate policies, and then under the assumption of constant interest rate policy, consider the effects of fiscal policy and quantitative easing/tightening (QE/QT), swap lines with other central banks, and supplementary leverage ratio (SLR) exemptions.

Table 2 summarizes our results, which we subsequently explain in more detail. Setting aside the specifics associated with each policy, our main message is that the effects of each of the policies depend on the Treasury market regime.

4.1. Interest rate policies and term premia

The effects of interest rate policy are somewhat nuanced. As a benchmark, consider a policy that causes all short-term rates (r, i^s, i^l, y^{bill}) and expected future rates ($y_{\mathbb{Q}}$ and $y_{\mathbb{P}}$) to shift in parallel, in conjunction with a fiscal policy that holds fixed the dollar value of the bond supply ($\exp(-ny)S^{bond}$). The dealer indifference conditions (equations (28) and (31)) and the client demand functions (equations (25) and (26)) in our model depend only on spreads, and not the level of rates. As a result, such a policy shifts the bond yield y in parallel with all other rates and otherwise has no effect on quantities and synthetic lending spreads.[25] Interest rate policies matter in our model only insofar as they deviate from this benchmark.

One potential deviation is a change in term premia. There is considerable evidence that interest rate policies can affect term premia (e.g. Campbell and Shiller (1991) and Hanson and Stein (2015)). If a policy reduces $y_{\mathbb{Q}}$ relative to $y_{\mathbb{P}}$ (i.e. a reduction in term premia), expected bond returns will decline, thereby reducing client demand (all else equal). In the long regime, dealers will absorb additional bonds, crowding out synthetic lending, whereas in the short regime dealers will supply fewer bonds via shorting, increasing their capacity for synthetic lending (Propositions 1 and 2). In both regimes, bond yields will fall by less than swap yields (see Fig. 8).

4.2. Fiscal policy and quantitative easing/tightening

We now discuss the effects of quantity changes: Treasury debt issuance and Federal Reserve Treasury purchases. We abstract from the effects of changes in expected future rates induced by these quantity changes, as the effects of changing term premium are discussed in the previous subsection.

The effects of increased debt issuance (fiscal policy) are described, for the long and short regimes, by the comparative statics of Propositions 1 and 2. In both regimes, intuitively, debt issuance raises bond yields. In the long regime, this increases synthetic lending spreads, as dealers are forced to absorb additional bond supply on balance sheet. In the short regime, synthetic lending spreads decrease, as dealers absorb bond supply by reducing the size of their short positions, which increases the amount of their balance sheet available for synthetic lending.

We define QE (QT) as the Federal Reserve's purchases (or sales/redemptions) of Treasury bonds in the secondary market.[26] The effects of QE/QT are similar to the effects of Treasury supply, but also involve changes in the levels of reserves held at the Fed. Here, we will separately consider the Treasury demand and reserve supply channels of QE/QT. We assume that no SLR exemption is applied to reserves or Treasury securities.

To isolate the effects of the Treasury demand channel, we consider a hypothetical version of QE in which the Fed purchases Treasury bonds in exchange for Treasury bills. This operation (which is somewhat akin to "Operation Twist") leaves the supply of reserves unchanged. We model this operation, which isolates the Treasury demand channel of QE, as a parallel outward shift in the demand curve for Treasury bonds (D_U) in our static model, which is equivalent (in both the long and short regimes) to a shift in bond supply. Holding fixed money market yields and swap rates, in the long regime quantitative easing will reduce both yields and synthetic lending spreads (see part two of Proposition 1). In contrast, in the short regime, quantitative easing will reduce yields while increasing synthetic lending spreads (see part two of Proposition 2). Both of these effects operate through the balance sheet mechanism; the regime matters because it determines whether balance sheet constraints are tightened or loosened by QE. If the goal of quantitative easing is to lower financial intermediation spreads, then our results imply that QE is effective via the Treasury demand channel in the long regime but not in the short regime.

To isolate the reserve supply channel, we consider a hypothetical purchase of Treasury bills in exchange for reserves, which leaves the supply of longer-maturity Treasury bonds unchanged. Note that the combination of these two operations is QE: the purchase of Treasury bonds in exchange for reserves. Note also that the effects of QT are exactly the opposite of those of QE. The effects of QE through the reserve supply channel are more complex, and depend in particular on whether the Fed's overnight reverse repo (ONRRP) facility is actively used.

With an actively used ONRRP facility, the reserve supply channel is muted. The ONRRP facility allows money market funds to make repo loans to the Federal Reserve. If the Fed exchanges reserves for bills, these funds will sell bills and receive deposits at their clearing banks, and then lend those deposits back to the Federal Reserve using the ONRRP facility. Thus, with an active ONRRP facility, the effects of QE operate entirely through the Treasury demand channel. In contrast, without an active ONRRP facility, the reserves created by the QE operation will end up on bank balance sheets, potentially offsetting the Treasury demand channel in the long regime, but amplifying its effect in the short regime.

4.3. Central bank swap lines

We next consider the policy of establishing swap lines between the Federal Reserve and other central banks. These swap lines allow foreign central banks to borrow dollars from the Federal Reserve, using their own currency as collateral. The foreign central banks then lend those dollars to their local banks, typically for the purpose of financing a position in dollar-denominated assets. In the discussion that follows, we hold all other monetary, fiscal, and regulatory policies constant.

Swap lines allow non-U.S. banks to borrow dollars, and are a substitute for borrowing synthetic dollars via a dealer. We assume that dealers are not themselves borrowing using the swap lines; providing subsidized dollar funding to dealers would have additional effects not described in this section. We therefore incorporate swap lines into our static model as equivalent to a demand shift in the synthetic lending market (a parallel shift in D^{syn}). The swap lines establish by the Federal Reserve generally have rates determined by policy. If the rate is higher than the prevailing market rate, the facility will go unused, and the equivalent demand shift is zero. If the rate is appealing, it is equivalent to a rate ceiling in the synthetic loan market, and hence to an endogenously sized decrease in the demand for synthetic dollars.[27]

Any demand decrease in the synthetic lending market will lead to reduced synthetic lending spreads in both the long and short regimes (see part four of Propositions 1 and 2). In the long regime, these reduced lending spreads will also lead to reduced Treasury yields (and hence swap spreads), whereas in the short regime reduced lending spreads would lead to increased Treasury yields. Again, both of these effects operate through the relaxation of balance sheet constraints, and the regime determines the relationship between balance sheet tightness and

[25] Without the offsetting fiscal expansion, bond yields will rise less than one-for-one with other rates, and synthetic lending spreads will change in a regime-dependent direction, because the increase in yields reduces the dollar supply of bonds. This effect is potentially offset by forces outside our model– for example, the holders of bonds might experience capital losses and decrease their demand. For this reason, we do not emphasize it as the main effect.

[26] The Federal Reserve and other central banks have at times purchased mortgage, corporate, and other bonds as part of quantitative easing programs. The model described thus far considers only Treasury bonds, and for this reason we restrict attention to Treasury purchases.

[27] During normal times, because of stigma associated with tapping central bank liquidity facility and moral suasion from central banks discouraging banks to use swap line to fulfill their routine funding needs, the take-up in the swap line is extremely low even when the swap line rate borrowing rate is temporarily below the implied dollar rate from the FX swap market.

Treasury yields. Note that in both cases, there would be an increase in the demand for Treasury bond financing (either tri-party repo or security lending). Our model assumes these rates are fixed, but in a more complex model these rates might also adjust.

We should note that, in our model, a swap line with a rate equal to the swap rate and large capacity could drive the synthetic lending spread to zero. This would break the link between CIP violations and balance sheet costs. Other financial intermediation spreads would decline (due to the relaxation of dealer balance sheet constraints), but would not go to zero.

4.4. Leverage ratio exemptions

We next consider changes to regulatory policy that involve temporarily exempting certain kinds of low-risk assets from the SLR calculation. We consider two possible exemptions: exempting Treasury bonds and repo loans against Treasury collateral (exempting Treasurys, for short), and exempting reserves. Similar policies were implemented during the most acute parts of the COVID-induced market disruptions in 2020.

Recall in our static model that we have consolidated dealers and their levered clients into a single entity, based on the analysis of Internet Appendix Section F.1. This consolidation is based on the fact that both the repo loans that a dealer provides to its levered clients to hold Treasury bonds and direct holdings of Treasury bonds increase the size of the dealer's balance sheet. For this reason, it is simpler to consider a policy that exempts both repo loans against Treasury bonds and Treasury bonds directly owned by dealers (exempting one but not the other would shift the net holdings of Treasury bonds from dealers to their levered clients or vice versa, in addition to relaxing balance sheet constraints). As in the previous subsection, we hold all other monetary, fiscal, and regulatory policies constant. In particular, we assume that the regulatory exemptions are temporary and do not affect $y_{\mathbb{Q}}$ or $y_{\mathbb{P}}$.

Exempting Treasurys will both free up dealer balance sheet capacity for synthetic lending and remove the need to reduce CIP arbitrage activity when taking a net position in Treasury bonds. This will lead to Treasury discount rates that are a function only of financing rates (i^l in the long regime, i^s in the short regime), and therefore have the effect of reducing yields in the long regime and increasing yields in the short regime. In both regimes, the SLR exemption will allow dealers to allocate the regulated portion of their balance sheet entirely to synthetic lending ($q^{syn} = \bar{q}$) and lead to a reducing in financial intermediation spreads.

Exempting reserves (or any other assets) frees up the dealer balance sheet space for Treasury holding and synthetic lending, and thus is equivalent to expanding the balance sheet capacity $\bar{q}$ in our static model. In the long regime, this would result in a decline in bond yields and synthetic lending spreads; in the short regime, bond yields would rise while synthetic lending spreads fall (see part two of Propositions 1 and 2).

4.5. The post-COVID period

In the two years following March 2020, the supply of Treasury bonds expanded substantially (Fig. 3). Despite this expansion of supply, balance sheet costs and swap spreads remained near their pre-COVID levels (see Fig. 4).[28] Our model predicts that, if all else were equal, the expansion of Treasury supply would have led to higher balance sheet costs and more negative swap spreads.

However, all else was not equal. The Federal Reserve expanded its quantitative easing program, absorbing a substantial portion of the additional bond issuance (Fig. 3). The Fed also expanded its swap lines with other central banks and temporarily relaxed regulatory constraints on banks' Treasury holdings. Our interpretation of the post-COVID period is that these policies were able to offset the effects of increased Treasury supply on swap spreads and CIP violations.

4.6. Implications for monetary policy tightening cycles

In June 2022, the Fed began increasing short-term interest rates substantially while engaging in quantitative tightening, initiating a "tightening cycle." Tightening cycles are often associated with flat or inverted Treasury yield curve, and low expected returns on long-term Treasury bonds. In our model, this dampens real money investors' demand for Treasury bonds, and it is particularly challenging for dealers and levered investors to accommodate a reduction in Fed holdings of Treasury bonds (QT) when client demand is weak.

Consider the experience of the 2017-2019 tightening cycle, in which the Fed normalized its balance sheet for the first time post-GFC and increased the short-term interest rates from the zero-lower-bound to 2.5 percent. During that tightening cycle, dealers' increased their Treasury holdings by about \$100 billion and hedge funds increased their holdings by about \$350 billion, together accounting for the entirety of the \$390 billion Fed balance sheet normalization from October 2017 to September 2019. The swap-Treasury spread and the Treasury cash-futures basis widened considerably over the period.[29]

Consistent with this experience, our model suggests that the combination of QT with an active ONRRP facility and a declining term premium, in the long regime, can lead to more negative swap spreads and higher financial intermediation spreads. SLR exemptions and the use of the swap lines established with foreign central banks have the potential to ameliorate these effects.

5. Conclusion

We have documented a regime change in the U.S. Treasury bond market. Prior to the 2008-2009 financial crisis, dealers were net short-sellers of Treasury bonds, swap spreads were positive, and CIP violations were small. Following the GFC, dealers became net long Treasury bonds, swap spreads turned negative, and covered interest parity violations emerged. Our analysis ties these observations together by constructing arbitrage bounds, the net short and net long curves, and providing evidence of dealers-as-arbitrageurs in the Treasury market.

We then discuss the causes and consequences of this regime change. We view the large increase in Treasury supply as the primary driver of this regime change. Using a stylized static model, we have argued that this regime shift has amplified the effects of quantitative easing and of the yield curve slope on borrowing spreads. In the post-GFC dealer-long regime our model predicts tighter dealer balance constraints in response to Fed quantitative tightening and a flat or inverted Treasury yield curve, and more elevated financial intermediation spreads. Our analysis suggests that other polices, including the use of swap lines and of exemptions to SLR calculations, can help offset these effects.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Code availability

The code and data packet for this article can be found at https://doi.org/10.17632/zfwr38f2bf.1.

[28] The spike in balance sheet costs during March 2020 is apparent in the 2-year maturity net long and net short curves, but CIP spreads and other balance sheet cost measures reverted rapidly to their pre-COVID levels.

[29] Crowded dealer balance sheets and the build-up of the levered investor positions may have contributed to the repo market distress of September 2019.

Data availability

The authors do not have permission to share data.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jfineco.2023.103722.

References

Alvarez, F., Jermann, U.J., 2000. Efficiency, equilibrium, and asset pricing with risk of default. Econometrica 68, 775–797.

Augustin, P., Chernov, M., Schmid, L., Song, D., 2021. Benchmark interest rates when the government is risky. J. Financ. Econ. 140, 74–100.

Barth, D., Kahn, R.J., 2021. Hedge funds and the treasury cash-futures disconnect, OFR WP, pp. 21–01.

Biais, B., Hombert, J., Weill, P.-O., 2021. Incentive constrained risk sharing, segmentation, and asset pricing. Am. Econ. Rev. 111, 3575–3610.

Boyarchenko, N., Eisenbach, T.M., Gupta, P., Shachar, O., Van Tassel, P., 2018a. Bank-intermediated arbitrage. FRB of New York Staff Report.

Boyarchenko, N., Gupta, P., Steele, N., Yen, J., et al., 2018b. Negative swap spreads. Federal Reserve Bank of New York Economic Policy Review 24, 1.

Brunnermeier, M.K., Pedersen, L.H., 2009. Market liquidity and funding liquidity. Rev. Financ. Stud. 22, 2201–2238.

Brunnermeier, M.K., Sannikov, Y., 2014. A macroeconomic model with a financial sector. Am. Econ. Rev. 104, 379–421.

Campbell, J.Y., Shiller, R.J., 1991. Yield spreads and interest rate movements: a bird's eye view. Rev. Econ. Stud. 58, 495–514.

Chien, Y., Cole, H., Lustig, H., 2011. A multiplier approach to understanding the macro implications of household finance. Rev. Econ. Stud. 78, 199–234.

Drechsler, I., Savov, A., Schnabl, P., 2018. A model of monetary policy and risk premia. J. Finance 73, 317–373.

Du, W., Hébert, B.M., Huber, A.W., 2022. Are intermediary constraints priced? Rev. Financ. Stud.

Du, W., Tepper, A., Verdelhan, A., 2018. Deviations from covered interest rate parity. J. Finance 73, 915–957.

Duffie, D., 2017. Post-crisis bank regulations and financial market liquidity. Banca d'Italia.

Duffie, D., 2020. Still the world's safe haven? Redesigning the US treasury market after the Covid-19 crisis'. Hutchins center on fiscal and monetary policy at brookings. Available online: https://www.brookings.edu/research/still-the-worlds-safe-haven. (Accessed 28 November 2021).

Duffie, D., Gârleanu, N., Pedersen, L.H., 2005. Over-the-counter markets. Econometrica 73, 1815–1847.

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

Feldhütter, P., Lando, D., 2008. Decomposing swap spreads. J. Financ. Econ. 88, 375–405.

Fleckenstein, M., Longstaff, F.A., 2021. Treasury richness. NBER Working Paper.

Fleming, M.J., Rosenberg, J.V., Nguyen, G., 2022. How do treasury dealers manage their positions? FRB of New York Staff Report.

Gertler, M., Kiyotaki, N., 2010. Financial Intermediation and Credit Policy in Business Cycle Analysis. Handbook of Monetary Economics, vol. 3. Elsevier, pp. 547–599.

Goldberg, J., 2020. Liquidity supply by broker-dealers and real activity. J. Financ. Econ. 136, 806–827.

Greenwood, R., Hanson, S.G., Stein, J.C., 2015. A comparative-advantage approach to government debt maturity. J. Finance 70, 1683–1722.

Greenwood, R.M., Vissing-Jorgensen, A., 2018. The impact of pensions and insurance on global yield curves. Harvard Business School Finance Working Paper, pp. 19–59.

Grinblatt, M., 2001. An analytic solution for interest rate swap spreads. Int. Rev. Finance 2, 113–149.

Haddad, V., Moreira, A., Muir, T., 2021. When selling becomes viral: disruptions in debt markets in the Covid-19 crisis and the fed's response. Rev. Financ. Stud. 34, 5309–5351.

Haddad, V., Sraer, D., 2020. The banking view of bond risk premia. J. Finance 75, 2465–2502.

Hanson, S.G., Malkhozov, A., Venter, G., 2022. Demand-supply imbalance risk and long-term swap spreads. Working paper.

Hanson, S.G., Stein, J.C., 2015. Monetary policy and long-term real rates. J. Financ. Econ. 115, 429–448.

He, Z., Krishnamurthy, A., 2013. Intermediary asset pricing. Am. Econ. Rev. 103, 732–770.

He, Z., Nagel, S., Song, Z., 2022. Treasury inconvenience yields during the Covid-19 crisis. J. Financ. Econ. 143, 57–79.

Hébert, B.M., 2020. Externalities as arbitrage. NBER Working Paper No. 27953.

Jermann, U., 2020. Negative swap spreads and limited arbitrage. Rev. Financ. Stud. 33, 212–238.

Jiang, Z., Krishnamurthy, A., Lustig, H., 2021. Foreign safe asset demand and the dollar exchange rate. J. Finance 76, 1049–1089.

Joslin, S., Li, W., Song, Y., 2021. The Term Structure of Liquidity Premium. USC Marshall School of Business Research Paper.

Joslin, S., Priebsch, M., Singleton, K.J., 2014. Risk premiums in dynamic term structure models with unspanned macro risks. J. Finance 69, 1197–1233.

Joslin, S., Singleton, K.J., Zhu, H., 2011. A new perspective on Gaussian dynamic term structure models. Rev. Financ. Stud. 24, 926–970.

Klingler, S., Sundaresan, S., 2019. An explanation of negative swap spreads: demand for duration from underfunded pension plans. J. Finance 74, 675–710.

Klingler, S., Sundaresan, S., 2023. Diminishing treasury convenience premiums: effects of dealers' excess demand and balance sheet constraints. J. Monet. Econ. 135, 55–69.

Krishnamurthy, A., Li, W., 2023. The demand for money, near-money, and treasury bonds. Rev. Financ. Stud. 36, 2091–2130.

Krishnamurthy, A., Vissing-Jorgensen, A., 2012. The aggregate demand for treasury debt. J. Polit. Econ. 120, 233–267.

Kruttli, M.S., Monin, P., Petrasek, L., Watugala, S.W., 2021. Hedge fund treasury trading and funding fragility: Evidence from the Covid-19 crisis.

Liu, J., 2019. Comovement in arbitrage limits. Jacobs Levy Equity Management Center for Quantitative Financial Research Paper.

Longstaff, F.A., 2004. The flight-to-liquidity premium in us treasury bond prices. J. Bus. 77, 511–526.

Nagel, S., 2016. The liquidity premium of near-money assets. Q. J. Econ. 131, 1927–1971.

Siriwardane, E., Sunderam, A., Wallen, J., 2021. Segmented arbitrage. Available at SSRN 3960980.