The Treasury - SOFR Swap Spread Puzzle Explained

David Wu[*] Robert Jarrow[†]

Abstract

SOFR swap rates have been consistently lower than US Treasury rates since their publication in October 2018. This difference is puzzling because it represents a potential arbitrage opportunity that is easily replicated in repo and bond markets. Furthermore, a traditional explanation for LIBOR-swap spreads, credit risk, no longer applies because SOFR is a near riskless rate. This paper examines market frictions and regulations in swap, bond, and repo markets to show that these swap spreads can be explained by three factors: bond market price quotes, repo transaction costs, and funding costs due to Basel III regulations.

[*]Operations Research and Information Engineering, Cornell University, Ithaca, N.Y. 14853. email: dbw78@cornell.edu.

[†]Samuel Curtis Johnson Graduate School of Management, Cornell University, Ithaca, N.Y. 14853. email: raj15@cornell.edu

1 Introduction

Interest rate markets are some of the largest and most liquid in the world. In the U.S. alone, government Treasuries had more than 24 trillion USD outstanding notional as of June 2023 SIFMA (2024). During the same time frame, interest rate swaps (IRS) - agreements that allow one party to exchange fixed interest rate payments with an agreed-upon floating reference rate - had notionals exceeding 465 trillion USD worldwide, making IRS the largest over-the-counter derivative market in the world BIS (2023). Under the traditional no arbitrage frictionless and competitive market assumptions, an investor can borrow funds in the repurchase agreement (repo) markets and buy Treasury bonds to synthetically construct a swap. Thus, after the inclusion of minimal market frictions, we expect the rates in both markets to be similar. However, since the 2008 financial crisis, IRS fixed rates have been consistently lower than those offered in the bond markets. Such a difference is puzzling: why have arbitrageurs not stepped in to profit and in the process, drive spreads to zero?

A recent and major change to interest rate markets was the introduction of the Secured Overnight Financing Rate (SOFR) as the dominant benchmark rate. Historically, interest rate derivatives were based on the London Interbank Offered Rate (LIBOR) index, which was a weighted average of quotes on borrowing rates from major global banks. This presented a difficulty when replicating the swap floating leg, because LIBOR was restricted to private loans issued by financial institutions. Some earlier works, such as Liu et al. (2002), used this to explain swap spreads. However, due bank manipulation concerns, the Dodd-Frank Act in the U.S. and various other regulations in Europe resulted in LIBOR being phased out, see Hou and Skeie (2014) and Wheatley (2012). In 2017, the Alternative Reference Rate Committee (ARRC) designated SOFR as a replacement for LIBOR. SOFR measures the cost of borrowing money overnight collateralized by U.S. Treasuries, and it is the rate associated with transactions at the 50th percentile of transaction volume. The SOFR phase-in was completed in June 2023, which marked the cessation of all USD LIBOR panels. For more details on SOFR, see New York Fed (2022).

In theory, the shift from LIBOR to SOFR makes it easier to replicate interest rate swaps and should reduce spreads. However, Figure 1 shows that swaps are still consistently negative. This difference is more significant for longer tenors. In particular, for 30 year swaps with \$100 notional, the swap fixed rate is around \$0.6 less than an equivalent Treasury per year.

This difference results in a present value discrepancy that can reach more than $12 per $100 notional.

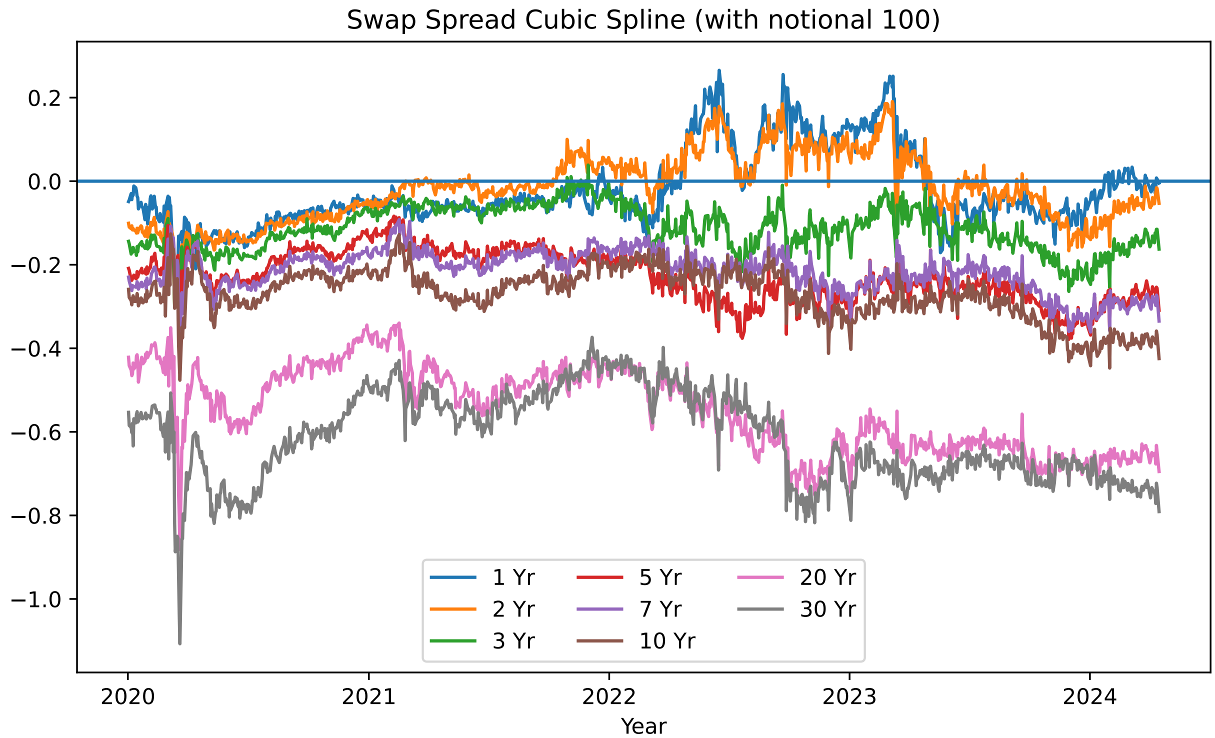

Figure 1: Spreads between the swap fixed rate and a theoretical no-arbitrage fixed rate implied by U.S. Treasury bond prices. These graphs show rate differences per swap maturity. The left axis shows the dollar value difference per $100 notion.

The previous literature has examined the drivers of swap spreads. One strand of this literature examines the the supply and demand imbalances in swap and Treasury markets, for example, see Jermann (2020), Hanson et al. (2024), Klingler and Sundaresan (2019), and Du et al. (2023). We instead build a term structure model accounting for frictions to understand the dynamics of swap spreads. Previous works following this approach include Duffie and Singleton (1997), Lang et al. (1998), Collin-Dufresne and Solnik (2001), Feldhütter and Lando (2008), Dubecq et al. (2016) Augustin et al. (2021), among others. Almost all the previous work focus exclusively on LIBOR swaps, instead of the newer SOFR swaps.

For example, Feldhütter and Lando (2008) and Dubecq et al. (2016) attribute part of the observed swap spread to credit risk, i.e. the difference in default risk of the financial sector and U.S. Treasuries. However, with SOFR replacing LIBOR, spreads driven by credit risk disappear because SOFR is a near riskless rate. Overnight borrowing rates that contribute to the SOFR calculation are collateralized by U.S. Treasuries and thus, have the same risk as U.S. Treasuries.

Johannes and Sundaresan (2007) provide theoretical and empirical evidence of the effects of marking-to-market and collateralization on swap spreads. However, as we will show later, marking-to-market actually makes swaps more similar to the daily roll-overs in the repo markets required to finance the bond purchase in the synthetic reconstruction of a swap.

The previous papers typically abstract away from market micro-structure frictions in both swaps and repo markets. For example, they do not include estimates for transaction costs in bond markets nor operational risks in repo markets. However, there is empirical evidence that liquidity and market conditions impact swap trades. Fritsche et al. (2020) and Riggs et al. (2020). In this paper, we show how these additional costs can reconcile swap spreads with an arbitrage free framework.

Unless otherwise specified, our time period covers Jan 1, 2020 to April 15, 2024. Cleared SOFR swaps only began trading on the CME starting Oct 1, 2018 CME (2018). We exclude the first year of data to allow markets to stabilize post transition. Our sample period includes both low interest rates due to the COVID pandemic and the post COVID rate hikes. Interestingly, swap spreads remain mostly constant over this period, regardless of the interest rate environment.

We first show that absent market frictions, there is a clear mispricing in swap markets relative to the U.S. Treasury and repo markets. Conventionally, swap spreads are defined loosely as the difference between the swap fixed rate and the yield on a Treasury with the same maturity. However, a single U.S. Treasury bond/note does not reflect the cash flows needed to replicate a swap because the timing of the cashflows are mismatched. This timing mismatch results in reinvestment risk. We calculate swap spreads using a no arbitrage framework and show that the observed spreads are too large to be explained by estimation noise alone. We show that these spreads can be attributed to three main factors.

The first is the valuation of the fixed cashflows. We document that even within the U.S. Treasury market, the same cash payment can have different prices due to market frictions.

In particular, U.S. coupon STRIPS tend to be cheaper than U.S. Treasuries, which are in turn cheaper than principal STRIPS. This has been observed as early as Jordan et al. (2000). This means that there are "multiple" yield curves that can be used to price future cash flows. Since the actual synthetic construction of a swap requires purchasing bonds, we use a notional weighted average approach to calculate a single yield curve and we also include bid-ask spreads in the valuation.

Second, we model the costs of the using repo markets. Repo transactions are typically negotiated over-the-counter (OTC) and require a developed network. This can create entry costs and risks of not being able to secure short-term funding. Such a risk is realized in practice, as can be seen from the use of the Fed Discount Window, see Fed (2022), which provides collateralized short-term funding at the primary rate (roughly 25bps above SOFR) for depository institutions. This funding is always available, but more expensive than other overnight rates. In well-functioning/frictionless markets, no firm would use the discount window funding. Furthermore, repo markets have counterparty risks when there are delivery fails. Such failures occur daily and total failures per day range between 20 billion USD to 25 billion USD. Defaults are expensive and most repo participants account for these costs in their transactions Wallerstein (2023). However, these costs are hidden from quoted SOFR rates. In this work, we explicitly model these costs to help explain swap spreads.

Last, the Basel III based regulations introduced supplementary leverage ratio (SLR) constraints. These constraints require that firms must maintain a certain amount of capital relative to assets. While swaps and repo transactions have the same cashflows (and risks), they are treated separately under the SLR framework. Swaps have a significantly lower impact on SLR, which creates a preference among investors to purchase swaps (despite the higher market prices).

Similar frictions have previously been documented as a factor in swap spreads. For example, Jermann (2020) shows that LIBOR swaps spreads can be explained if there are costs for holding bonds. Boyarchenko et al. (2018) show how hypothetical SLR constraints can greatly reduce the return on bank equity. In a similar vein, Andersen et al. (2019) argue that swap prices reflect a funding adjustment caused by debt overhang. Our work is the first to analyze funding constraints in conjunction with other bond and repo market costs.

We show that these three factors completely explain swap spreads, with the exception of a few market-wide liquidity crises. The costs due to SLR constraints have the largest effect. Furthermore, we include various other relevant considerations, such as forward rate curve construction, transaction costs in swaps, timing mismatches in synthetically constructed cashflows, etc. to show that they are not the cause of the swap spreads.

An outline of this paper is as follows. Section 2 presents the mechanics of a swap trades. Sections 3 and 4 study fixed and floating leg construction of a swaps cash flows. Section 5 introduces the impact of the Basel III regulation, and section 6 concludes.

2 The Mechanics of a Swap Transaction

An interest rate swap is a contractual agreement where two counterparties exchange a series of fixed interest rate payments with floating rate payments on some notional. The dominate floating interest rate benchmark in the U.S. is SOFR, and the swap's fixed rates are calculated so that the swap has zero value at inception.

The transition to SOFR also changed a number of the LIBOR swap characteristics. Previously, USD LIBOR swaps typically traded with semi-annual/annual payments on the fixed leg against quarterly payments on the USD LIBOR 3 month floating leg. With the SOFR swap, both legs have annual payments. This reduces the credit risk of the swap, and from a variation margin perspective, reduces large changes in the mark-to-market value associated with quarterly payments. Furthermore, SOFR is an overnight rate. To calculate the annual floating leg payment, the daily SOFR is compounded in arrears using a Act/360 day count convention.

2.1 Swap Market Participation

Swap markets generally consist of end users and intermediaries. End users of swaps arrive to the market trying to enter receive fixed or pay fixed positions. Intermediaries, usually swap dealers or hedge funds, must accommodate the net demand by taking the reverse position and hedging the position with Treasuries. The main swap end-users, as explained by the Treasury Borrowing Advisory Committee TBAC (2021), are shown in Table 1.

During our sample period, overall there was a net demand to receive fixed TBAC (2021). This is consistent with the direction of swap spreads. The net demand puts downward

End User	Net Position	Typical Use
Banks	Receive	Typically large positions, uses vary.
Hedge Fund	Pay	Hedge interest rate risk , swap spread trades.
Insurance & Pensions	Receive	Asset liability management, hedge interest rate risk.
Corporates	Receive	Transform fixed rate loans to floating , or conversely.
Government	Pay	Uses vary.

Table 1: Typical uses of participants in the swap market. Following market convention, "receive/pay" refers to the fixed leg of the swap.

pressure on fixed rates to entice intermediaries to offset positions. Thus, part of the swap spread can be seen as compensation for providing liquidity in the market.

While interest rate swaps have a maximum tenor of 50 years, most of the trading activity is concentrated at the short-term. Using Bloomberg's Swap Data Repository, Figure 2 shows that on average, trading activity for tenors 1.5 years or less dominate all swap trades. Volume for 20 and 30 year swaps, which have the largest swap spreads in our sample, are generally much smaller. This is consistent with general market trends: less liquid products tend to have larger mispricings.

2.2 No Arbitrage Pricing

While the previous analysis begins to shed light on the nature of swap spreads, it does not fully explain the swap spread puzzle. In particular, if a firm, either a bank or corporation, wants to enter a pay-float receive-fixed swap (expensive), why do they not synthetically construct the swap using bonds? Are there costs supporting these large spreads, especially at longer maturities? And if so, what are they?

To answer these questions, we focus on constructing the receive-fixed pay-float SOFR swap. We first consider the frictionless market, no-arbitrage price of a swap as a baseline and then modify the formula to account for transaction costs. Let $P(t, T)$ represent the time t price of a default-free zero-coupon bond paying 1 at time T, and let c be the fixed rate of the swap.

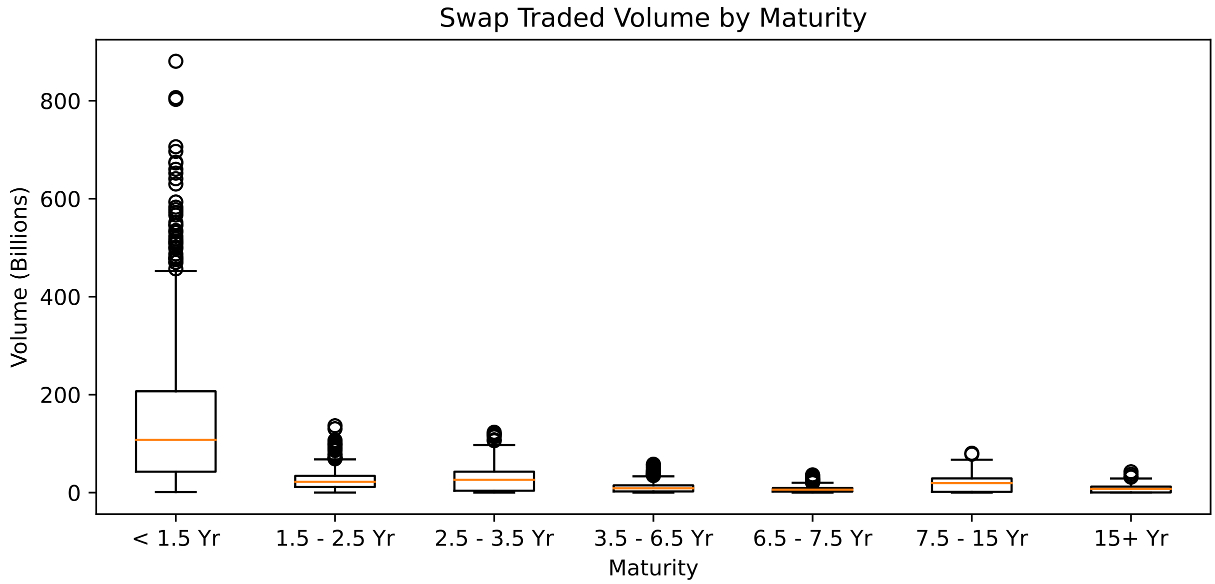

Figure 2: Daily trading volume (measured in swap notional) of USD OIS (overnight indexed swaps) from Jan 2020 - Apr 15, 2020 grouped by tenor. A majority of these swaps have floating rates indexed by SOFR, but a small portion are indexed by the Fed Funds rate. In each column, the box extends from the first quartile to the third quartile and the orange horizontal line corresponds to the median. The lines extend to the maximum and minimum values, excluding outliers (any value lying outside 1.5×Interquartile Range).

Consider a T tenor interest rate swap indexed by SOFR with notional 1. For the standard SOFR swap, payments are exchanged annually, and to construct the fixed leg, one needs to go long c zero coupon bonds maturing at times $k \in \{1, ..., T\}$. The cost is

$$\sum_{k=1}^{T} cP(0, k).$$

To construct the floating leg, one needs to borrow \$1 in the repo markets, and long one zero-coupon bond maturing at time T with cost

$$P(0, T) - 1.$$

To see why this works, one can borrow \$1 in the repo market and roll it over daily. After a year, this account owes \$1 plus interest accumulated at the SOFR rate. The investor can pay off the interest (which is equal to the SOFR swap payment), and start the next year again with \$1. One can repeat this procedure, paying annually the floating payment until year T. At year T, the \$1 in the repo account and bond at maturity cancel out. There are no cashflows related to the notional.

The swap rate c solves the equation

$$\sum_{k=1}^{T} cP(0,k) + \Big[P(0,T) - 1\Big] = 0$$

i.e. c equates the values of the floating and fixed payments, so that no cashflows occur at time 0. This argument can be proved formally as shown in Jarrow and Li (2023).

To generate Figure 1, we fit the zero-coupon bond prices $P(t,T)$ using cubic splines from Jan 1, 2020 to April 15, 2024 using data from the U.S. Treasury Bond Par Yields Release on the Federal Reserve's H.15 Selected Interest Rates section Fed (2024). For the algorithm we used to fit $P(0,k)$, see Wu and Jarrow (2024). Market swap rates are obtained from Bloomberg. Fitting the zero-coupon bond prices $P(t,T)$ with cubic splines is not crucial; in fact, swap spreads remain regardless of the forward curve parameterization.

2.3 Margin Requirements

There are margin complications in both the synthetic construction of a swap and trading swaps. This complication could have had an effect on the previous LIBOR swap as examined in Johannes and Sundaresan (2007). However, with the switch to SOFR swaps, the additional margin requirements actually make trading swaps more similar to its synthetic counterpart.

The synthetic construction of a \$1 notional swap requires a *bond portfolio* consisting of c long zero-coupon bonds at time $\{0, 1..., T\}$ and 1 long zero-coupon bond at time T. Let V_t represent the value of this bond portfolio. By construction, $V_0 = 1$. Thus, the value V_0 of the bonds can be used as collateral to borrow \$1 in the repo market, which is needed to construct the floating leg of the swap.

The collateral in repos are subject to haircuts of approximately 2%. Hu et al. (2021) [1]

[1] General Collateral Finance (GCF) Repos do not require haircuts but participants must post collateral to

Thus, V_0 can only be used to borrow \$0.98, and the firm needs to provide an additional \$0.02 investment to construct the bond portfolio. This \$0.02 saves on repo interest owed, and there is no opportunity cost, if the participant is not cash constrained.

Similarly, swaps require initial margin, which insures against the costs of replacing the swap in the case of a default. Using data from the CME Margin Matrix, initial margins typically range between 1% to 8% depending on swap's tenor and rate. Cash collateral earns interest at rates similar to SOFR with exact rates published on CME (2024). Consequently, there is no opportunity cost of providing the funding, if the participant is not cash constrained.

A second margin related cash-flow arises from daily re-balancing in repo markets. For example, if interest rates rise, V_t decreases. Let B_t (typically ≥ 1) be the value of the repo loan at time t. Repos are collateralized loans, so only $V_t \leq B_t$ can be borrowed, and any excess $B_t - V_t$ of the loan must be repaid. In return the investor saves overnight interest owed on $B_t - V_t$.

Similarly, swaps post variation margin (daily mark-to-market). If rates rise, the net present value of the swap decreases, and investors have to pay additional funds $B_t - V_t$. In return, they receive a price alignment interest CME (2014) determined by

$$\text{PAI} = \text{Adjusted NPV(prev bus. day)} \times \text{Latest Overnight Funding Rate} \times (\text{Days}/360).$$

where NPV = net present value. This again matches closely with how the repo market collateral functions.

The similarity of these cash flows and margin/collateral means that investors should not prefer synthetic construction over trading swaps. While there are calculation differences, these margin/collateral payments themselves are inherently small, and the slight differences between these two cannot explain the magnitude of the swap spreads. Because of this equivalence, we do not focus on margin/collateral differences in the subsequent analysis.

a clearing fund. For our purposes, the two are equivalent.

3 Fixed Leg Construction

Replicating the fixed leg of a n year swap requires purchasing n cashflows. The previous calculation does not account for bid-ask spreads in trading bonds. Furthermore, the zero-coupon bond's price $P(0, T)$ was calculated using CMT yields instead of actual bond prices. We address these two issues in this section.

Since payments for swaps are on an annual basis (while bonds have semi-annual coupon payments), to perfectly hedge the timing of the fixed payments, we need to use STRIPS (Separate Trading of Registered Interest and Principal of Securities). STRIPS separate the principal and coupon payments of U.S. Treasury notes or bonds, so that they can be traded as individual securities. For example, a 10 year bond can be stripped into 21 separate securities (one principal payment due at maturity and 20 semi-annual coupon payments) each with its own CUSIP. This process can also be reversed: if one collects the 21 securities, they can be reconstituted back into the U.S. Treasury security.

We obtain the bid and ask prices for all U.S. Treasury STRIPS from Refinitiv and U.S. Treasury bond prices from Treasury Direct. For our fixed leg construction, we use ask prices because the swap replication strategy requires buying bonds.

3.1 STRIPS Markets

Interestingly, principal STRIPS usually cost more than their coupon counterparts, even when the maturity dates are the same. The price difference is often quite large with principal STRIPS being up to \$3 more expensive on a \$100 notional.

To get a clearer sense of STRIPS versus bond markets, we reconstruct all Treasury bonds with principal STRIPS and then compute the price differences. This is not the same as reconstituting the bonds using STRIPS because we do not use any coupon STRIPS. However, this construction gives a portfolio with the same cashflows as a U.S. Treasury bond. Figure 3 shows the results. Particularly for long-term Treasuries (20+ years), the pricing errors are statistically similar to white noise plus a mean term. This means that there is a consistent overvaluation of principal STRIPS.

Repeating the analysis with coupon STRIPS, we get the reverse result: generally coupon STRIPS are cheaper than U.S. Treasuries. This phenomena has been documented previously

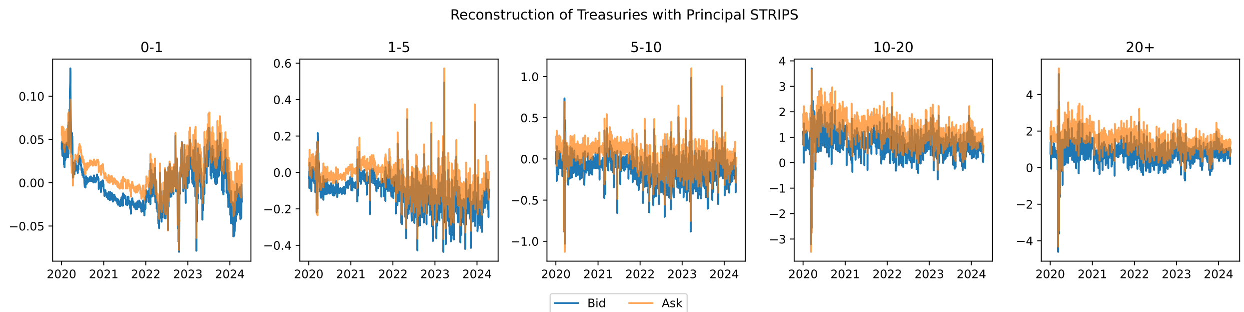

Figure 3: Time series graphs showing the reconstructed bond price using principal STRIPS minus the U.S. Treasury's price. The U.S. Treasuries are grouped into bins based on maturity. The blue graph plots differences in bid prices while and orange shows differences in ask prices.

in Jordan et al. (2000). One possible explanation for these differences are the costs (both exogeneous and endogenous - liquidity issues) in the reconstruction process, see Daves and Ehrhardt (1993). Any coupon STRIPS of the same maturity, regardless of which security it was stripped from, can be used to reconstitute a strippable Treasury security. However, the principal STRIPS in a reconstitution must be from the same underlying bond or note.

3.2 Modelling the Fixed Leg

For our swap analysis, we want to price of a set of known future cashflows. This is complicated by the fact that there are multiple "correct" prices depending on which underlying Treasury security is used as a reference and liquidity differences between coupon and principal STRIPS: some quoted prices have more volume than others.

Using the data release from the Monthly Statement of the Public Debt, we calculate the amount of outstanding coupon and principal STRIPS. Results for Jan 2020 are shown in Figure 4. For example, there are very few long-term coupon STRIPS. Thus, even though they have cheaper ask prices, the prices may not represent those at which trades could be executed.

To account for these liquidity differences, we construct the zero-coupon price curve using a two step process. First, we use a cubic spline interpolation between observation dates to get a continuous price function. We do this separately for principal and coupon STRIPS giving us two price functions $f_p(t)$ and $f_c(t)$, respectively. This process guarantees that the fitted

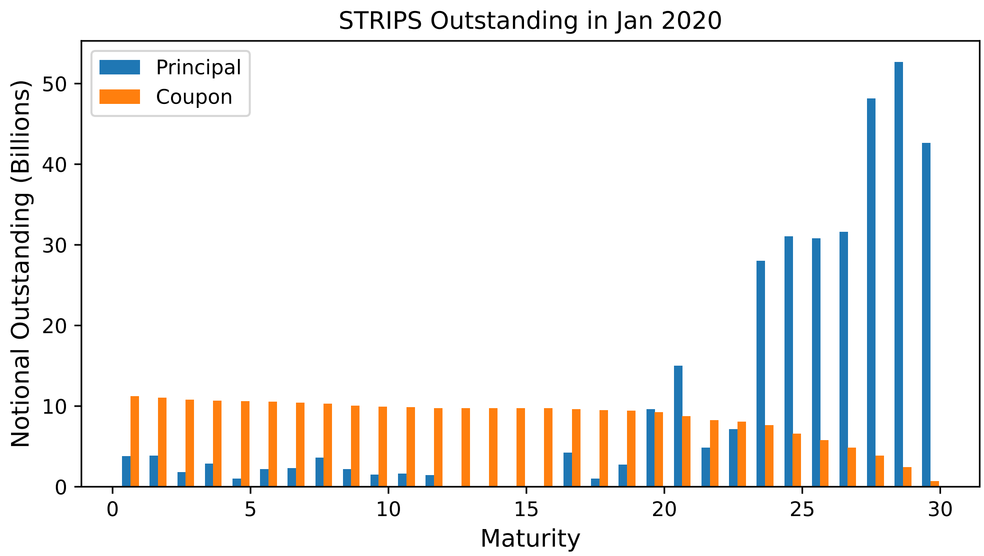

Figure 4: Amount of outstanding coupon and principal STRIPS grouped by maturity date. All cash flows in one year are grouped together.

price functions match all market prices on dates with observations. Second, we weigh the two functions according to their outstanding notional. In particular, we calculate a weight function

$$w(t) := \frac{\text{Principal Notional}}{\text{Total Strip Notional}}$$

in Jan 2020 (and similarly for subsequent months) using the data from Figure 4 with cubic spline interpolation between observation points. The final price curve is a weighted average of $f_p(t)$ and $f_c(t)$ given by

$$w(t)f_p(t) + [1 - w(t)]\, f_c(t).$$

The effects of using these new price functions on swap spreads are mixed as shown in Figure 5. In particular, 10 year swap spreads increased while 30 year swap spreads decreased (in magnitude). This is not surprising given the data on STRIPS markets. Shorter term swaps are mainly reconstructed using the cheaper coupon STRIPS. This further lowers the value of the fixed leg, and thus increases spreads. In contrast, to replicate 30 year swaps, more expensive principal STRIP payments dominate the market. Long dated swap payments were previously shown to be undervalued.

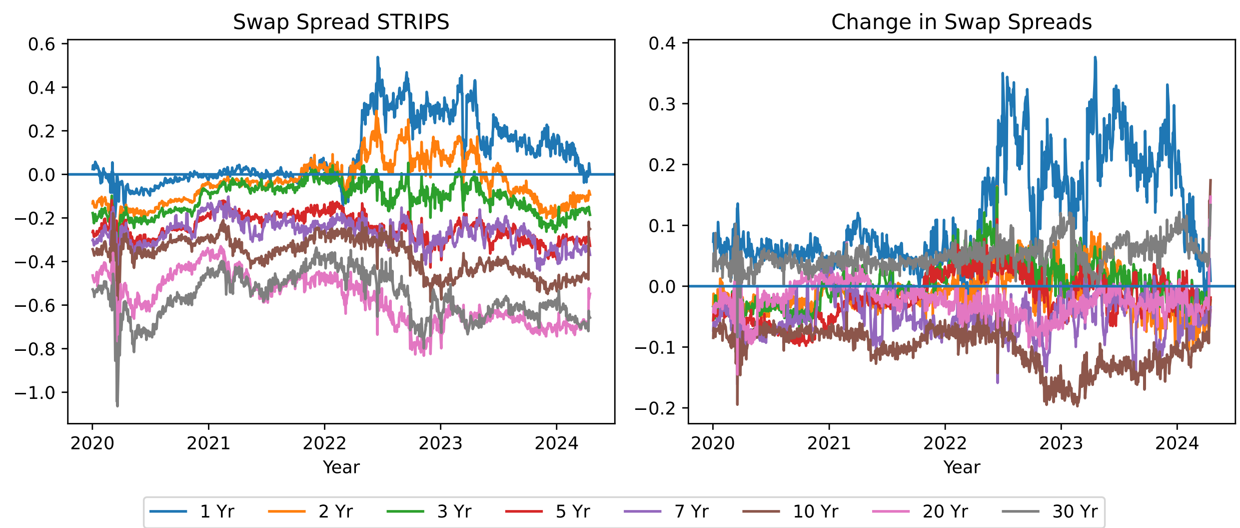

Figure 5: Left Panel: New swap spreads after using market prices and accounting for bid-ask spreads. Right Panel: Change in swap spreads from the baseline model in Figure 1

3.3 Timing Mismatches

Swap payments typically occur at different times than U.S. Treasury or STRIPS payments. We handled this issue by using interpolation, but those interpolated prices are not tradeable. Here, we show that this is an acceptable approximation. In particular, if timing mismatches between cash flow dates is a significant factor in swap spreads, we would expect that swap spreads should be smaller on those days where cashflows can be perfectly matched.

We test this hypothesis conditional on the swap's maturity. First, we fix a swap tenor and consider day i. Standing at day i, let T_{ij} represent the time (in years) until payment j. We calculate the total time gap between swap payments and U.S. Treasury cash flow dates given by

$$\text{Time mismatch}_i := \sum_j \min \left| T_{ij} - \text{Cashflow Dates for US Treasuries} \right|.$$

Scatterplots between Time mismatch$_i$ and swap spreads show little correlation regardless of the tenor. This provides evidence that swap spreads are not driven by timing mismatches in the cash flow payment dates. To provide a statistical test of this hypothesis, we fit the regression model

$$\text{Swap Spread}_i = b + m * \text{Time mismatch}_i \tag{1}$$

for each tenor i. We conduct the hypothesis test with the null hypothesis $H_0 : m = 0$ and the alternative hypothesis $H_a : m \neq 0$. Table 2 shows the P-values for each regression and the associated R^2 of the model. Overall, the amount of variation in swap spreads explained by Time Mismatch$_i$ is very low and we fail to reject the null hypothesis for all but one of the tenors at the 95% significance level. The apparent significance for 10 year swaps should not be overstated because we are doing sequential hypothesis testing. By pure chance, we expect 1 in 20 tests to be significant even in the absence of any relationship.

The overall trend is very clear: although a continuum of bonds do not trade, using interpolated prices is not a crucial factor in swap spreads.

	1 Year	2 Year	3 Year	5 Year	7 Year	10 Year	20 Year	30 Year
R^2	0.001	0.002	0.001	0.004	0.003	0.008	0.000	0.000
P-value	0.432	0.159	0.254	0.076	0.091	0.007 *	0.564	0.787

Table 2: R^2 and P-values from fitting the regression model in Equation 1. A " * " indicates significance at the 95% level.

4 Floating Leg Construction

While a IRS trade is a buy and hold transaction, the repo construction requires daily roll-overs to construct the compounded overnight rate, i.e. approximately 252 trades (one per trading day) are needed to replicate each floating payment. But the SOFR index ignores the transaction costs in rolling over a repo position and the counterparty risks involved in repo transactions. These add to the cost of synthetic construction which we examine in this section.

4.1 SOFR and Repo Markets

A repo transaction is typically used for short-term funding, with the most common tenor being overnight. Repos involve the sale of a security, combined with an agreement to repurchase the security at a predetermined date and at the same price plus interest. In essence, a repo is a short-term collateralized loan; the only distinction involves counterparty default considerations Duffie and Skeel (2012). There are three types of repos: bilateral, tri-party, and general collateral finance (GCF). The SOFR rate is calculated as the volume-weighted median [2] of all tri-partyrepo data collected from the Bank of New York Mellon, GCF repo data, and bilateral Treasury repo data cleared by the Fixed Income Clearing Corporation (FICC). The transaction data is filtered to remove specials (repos trading at a lower rate because of demand for a particular Treasury serving as collateral). We now explain each type of repo in more detail.

First, bilateral repos are negotiated between two counterparties and settled on a Delivery versus Payment (DVP) basis. Each party is responsible for valuing and managing the collateral, which involves operational complexities and leads to operational costs Copeland et al. (2012). DVP repos settle on a specific-security basis, which means they can be used for security-borrowing. For example, a cash provider may want a specific security to cover a short position - this cannot be done with other types repos. These two factors generally result in DVP repos having lower rates than the other two types of repos.

Next, tri-party repos outsource post-trade processing (collateral selection, payment and deliver, custody of collateral) to a clearing bank. In the U.S., the two clearing banks are the Bank of New York Mellon and JP Morgan Chase. The tri-party clearing bank does not provide a trading venue or perform market making. Instead, after a transaction has been negotiated, both parties notify the clearing bank to help with processing ICMA (2024). Tri-party repos can only be used for funding purposes because they transact on asset classes (for example, U.S. Treasuries) instead of specific securities. Furthermore, tri-party repo securities cannot be repledged outside the tri-party platform. This protects against fails to settle on a closing leg, in contrast with bilateral repos Baklanova et al. (2015).

Last, a special class of tri-party repos is the GCF repo offered by the FICC. Only firms deemed eligible by the FICC can negotiate GCF trades. GCF repos are negotiated on a blind basis, meaning that the parties of the transaction do not know each other's identity. Once a trade is brokered, FICC guarantees settlement by becoming the counterparty to every trade. GCF repos generally have the lowest operational costs and enhance liquidity in the inter-dealer repo market Agueci et al. (2014).

[2] The volume-weighted median is the rate associated with transactions occurring at the 50th percentile of transaction volume.

To replicate the floating leg of a swap, payments need to be rolled over on a daily basis. Here, the timing of the repo transactions cause an intraday counterparty risk. New repos are typically negotiated in the early morning with more than 60% of tri-party and GCF repos negotiated between 7:00am (when the market opens) and 8:30am Clark et al. (2021). However, settlement and unwinding of the previous day's repos start at 3:30pm TMPG (2022). If the counterparty for one repo fails to deliver, it also impacts the firm's ability to fulfill the roll over transaction. Table 3 summarizes the operational and counterparty risks for each type of repo.

DVP	Tri - Party	GCF
Opening Leg Risks		
Fail to Deliver Treasury (self)	Fail to Deliver Treasury (self)	Fail to Deliver Treasury (self)
Do not Receive Funds (counterparty)	Do not Receive Funds (counterparty)	NA (FICC guarantee)
Closing Leg Risks		
Fail to Return Treasury (counterparty)	NA	NA (FICC guarantee)
Default on Funds (self)	Default on Funds (self)	Default on Funds (self)

Table 3: Counterparty and operational (self) risks involved in the overnight repo markets from the perspective of an investor replicating the floating leg of a swap.

Since all loans are collateralized, loss from counterparty defaults are small. However, there is risk that the counterparty fails to provide funding or deliver the securities, and operational risk that the firm itself has issues preventing delivery, see Wallerstein (2023). These events are costly, and the defaulting party has to pay large (relative to transaction costs) fines. As a reference, the fine is on the order of \$500 per occurrence FICC (2021). Even if the firm itself can avoid the fines themselves, if the counter-party fails to deliver, the firm itself needs to find new sources of funding, likely at a larger cost.

4.2 Modeling the Operational and Counterparty Risks

Because the SOFR rate ignores operational costs and counterparty risk, the floating leg of a IRS is actually cheaper than constructing the swap synthetically. This again is consistent with the direction of swap spreads. Investors are willing to receive a lower fixed rate because they save on constructing the floating payments.

To explicitly calculate the magnitude of this impact, we use the volume-weighted mean interest rate of all repos in the FICC's GCF Repo Service as an approximation of the true rate of borrowing. The data is available at OFR (2024). As discussed above, GCF repos remove all counterparty risk and they have the lowest collateral management costs. Not surprisingly, the GCF rate is usually slightly higher than the tri-party and DVP repo rates.

Even so, there are still costs/risks of not being able to secure funding since all transactions are negotiated over-the-counter. As an upper bound for the cost of borrowing overnight, we use the Fed Discount Window's primary credit rate which is available to most depository intuitions. This rate (around 25bs higher than SOFR) is a collateralize loan from the Fed that is always available to authorized institutions. As such, it does not have any price discovery, settlement risk, or counterparty risks.

Figure 6 shows the difference between GCF/primary credit rates and SOFR.

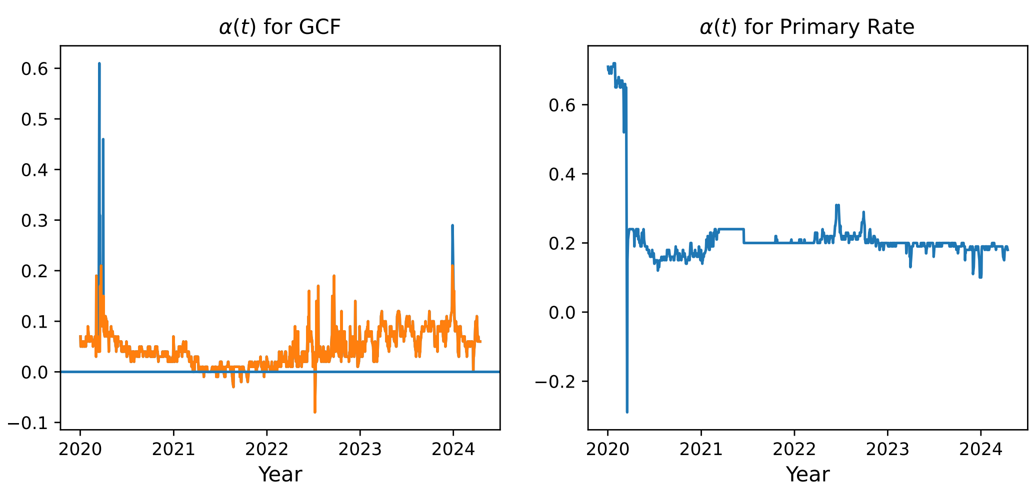

Figure 6: Left Panel: $\alpha(t) := \text{GCF} - \text{SOFR rate}$ (blue) and the same data after filtering for outliers (orange). Right Panel: $\alpha(t) := \text{Primary Credit} - \text{SOFR rate}$. All rates in basis points.

For GCF rates, we remove outliers that are more than 4 standard deviations from the sample mean. Most of the removed data points correspond to the beginning of the Covid crisis when there was a liquidity crisis O'Hara and Zhou (2021). For the primary rate (right panel), the sharp drop in March 2020 is due to the Fed deciding to (permanently) lower the primary rate for borrowing beginning March 20, 2020 Fed (2023). For model estimation, we only use primary rates after March 20, 2020.

4.2.1 Floating Leg Replication Cost

Let r_t represent the overnight rate (SOFR) and let α_t represent the spread between the GCF/primary rate and SOFR. α_t can be interpreted as the transaction cost for participating in the market. Let $B(T) := e^{\int_0^t r_u du}$ [3] represent the cumulative value of the money market account. We assume that bond markets are arbitrage free and complete. This implies that there exists a unique risk-neutral measure, $\mathbb{Q}$, in which discounted bond prices are martingales. In particular, the time 0 price of a zero-coupon bond maturing at T is

$$P(0,T) = E^{\mathbb{Q}}\left[\frac{1}{B_T}\right].$$

Consider a single swap payment that accumulates from time t to time T. The swap floating leg payment at time T is $S_t := e^{\int_t^T r_u du}$. The repo replication of the floating leg includes the overnight rate r_t and the additional cost α_t for a total of $R_t := e^{\int_t^T (r_u+\alpha_u) du}$. Thus, the additional accumulated costs at time T from a synthetic replication are

$$R_t - S_t = e^{\int_t^T (r_s+\alpha_s)ds} - e^{\int_t^T r_s ds} = e^{\int_t^T r_s ds}\left(e^{\int_t^T \alpha_s ds} - 1\right).$$

For analytic convenience and as a reasonable first approximation, we assume that r_t and α_t are conditionally independent given information at time t. This means that given the state of the market at time t, the realizations of r_s and α_s for $s > t$ are statistically independent. Then, under this assumption, the present value of the replication cost is

[3] Actual SOFR is reset daily. However, to simplify calculations, we use continuously compounded rates r_t.

$E^{\mathbb{Q}}\left[\frac{1}{B_T}e^{\int_t^T r_s ds}\left(e^{\int_t^T \alpha_s ds}-1\right)\right]$, which can be decomposed as

$$\begin{aligned} E^{\mathbb{Q}}\left[e^{\int_t^T \alpha_s ds}-1\right]E^{\mathbb{Q}}\left[\frac{e^{\int_t^T r_s ds}}{B_T}\right] &= E^{\mathbb{Q}}\left[e^{\int_t^T \alpha_s ds}-1\right]E^{\mathbb{Q}}\left[e^{-\int_0^t r_s ds}\right] \\ &= E^{\mathbb{Q}}\left[e^{\int_t^T \alpha_s ds}-1\right]P(0,t). \end{aligned}$$

Summing together the n swap payments, we get that the present value of the total floating payments, subtracting the additional accumulated replication costs are:

$$\underbrace{P(0,T)-1}_{\text{Swap Payment}} \quad - \quad \underbrace{\sum_{i=0}^{n-1} E^{\mathbb{Q}}\left[e^{\int_i^{i+1} \alpha_s ds}-1\right]P(0,i).}_{\text{Replication Costs}}$$

4.2.2 Estimating the Additional Costs α_t

We model the additional costs α_t using an Ornstein Uhlenbeck (OU) process given by

$$d\alpha(t) = k[\theta - \alpha(t)]dt + \sigma dW(t)$$

under the statistical probability measure where k, θ, σ are constants.

To facilitate estimation, we assume that there is no risk premium associated with α_t, so that this evolution is that under the risk neutral measure as well. Using historical time series of α_t (either from GCF or primary rate), we fit the parameters according to the following result.

Theorem 4.1. *Given daily time series observations $\{r_0, r_2, ..., r_N\}$ with the time step of $\delta = 1$ day, the maximum likelihood estimates for k, θ and σ are*

$$\hat{k} = -\frac{1}{\delta}\log\left(\frac{N\sum_{i=1}^N r_i r_{i-1} - \sum_{i=1}^N r_i \sum_{i=1}^N r_{i-1}}{N\sum_{i=1}^N r_{i-1}^2 - \left(\sum_{i=1}^N r_{i-1}\right)^2}\right)$$

$$\hat{\theta} = \frac{\sum_{i=1}^N \left(r_i - e^{-\hat{k}\delta} r_{i-1}\right)}{n(1-e^{-\hat{k}\delta})}$$

$$\widehat{\sigma^2} = \frac{2\hat{k}\sum_{i=1}^{N}\left(r_i - e^{-\hat{k}\delta}r_{i-1} - \hat{\theta}(1-e^{-\hat{k}\delta})\right)^2}{n\left(1-e^{-2\hat{k}\delta}\right)}$$

Proof. Straightforward calculation using the results from Brigo and Mercurio (2006). □

The fitted parameters are shown in Table 4 with standard errors estimated using the asymptotic results given in Tang and Chen (2009).

	$\hat{k}$	$\hat{\theta}$	$\hat{\sigma}$
GCF	77.8381 (8.1379)	0.00046 (3.189×10^{-5})	0.004241 (1.082×10^{-6})
Primary Rate	21.158 (4.0128)	0.001983 (5.041×10^{-5})	0.001780 (1.978×10^{-5})

Table 4: Estimated parameters for α_t using GCF repo rate and primary rate at the Fed discount window. The quantities in brackets represents the standard errors.

4.2.3 Impact on Swap Spreads

Under the above assumptions, we can explicitly calculate the term $E^{\mathbb{Q}}\left[e^{\int_i^{i+1}\alpha_s ds}\right]$ as given in the following theorem.

Theorem 4.2. *Given the assumptions above, we have that* $E\left[e^{\int_t^T \alpha(s)ds}\right] = \exp\left(\mu + \frac{\Sigma_1^2+\Sigma_2^2}{2}\right)$ *where*

$$\mu = \frac{\alpha(0)-\theta}{k}\left(e^{-kt} - e^{-kT}\right) + \theta(T-t)$$

$$\Sigma_1^2 = \frac{\sigma^2}{k^3}\left[\frac{1}{2}e^{2k(t-T)} + e^{-k(T+t)} - e^{k(t-T)} - \frac{1}{2}e^{-2kt} - \frac{1}{2}e^{-2kT} + \frac{1}{2}\right]$$

$$\Sigma_2^2 = \frac{\sigma^2}{k^2}\left[T - t - \frac{2}{k}\left(1-e^{-k(T-t)}\right) + \frac{1}{2k}\left(1-e^{-2k(T-t)}\right)\right]$$

Proof. The proof is included in Appendix A. □

To find the new arbitrage-free swap rate, after including the additional costs α_t, we again set the present value of the floating leg (after accounting for transaction costs) equal to the fixed leg, i.e. we solve for c such that

$$c\sum_{k=1}^{T}P(0,k)+\underbrace{P(0,T)+\sum_{i=0}^{n-1}E^{\mathbb{Q}}\left[e^{\int_i^{i+1}\alpha_s ds}-1\right]P(0,i)-1}_{\text{Floating Leg}}=0. \tag{2}$$

We use the ask prices for the long positions in the zero-coupon bonds as described in Section 3. The average improvement of swap spreads is shown in Table 5.

	1 Yr	2 Yr	3 Yr	5 Yr	7 Yr	10 Yr	20 Yr	30 Yr
GCF	0.0501	0.0472	0.0473	0.0452	0.0422	0.0336	0.0628	0.0483
Primary Rate	0.2059	0.2030	0.2031	0.2012	0.1983	0.1899	0.2196	0.2050

Table 5: Average changes, relative to Section 3, in swap spreads after considering repo costs across maturities.

These costs help to explain a portion of the swap spreads. However, even if a firm always borrows at the Fed's primary rate, there are still arbitrage profits possible by synthetically constructing a swap, given the observed swap spreads. For example, 30 year spreads are typically at least -40bp per year, but the primary rate based additional costs only account for 20bps of this difference. Thus, if there are no arbitrage opportunities, there must be other frictions still influencing the swap markets.

4.3 Transaction Fees

In addition to the operational and counterparty risk costs analyzed in the previous section, there are also various transaction fees that need to be included in the costs of synthetic replication of the swap. In tri-party repos, clearing banks charge a fee for their services. Also, collateral is usually delivered to the counterparty via the Fedwire Securities service. These costs are generally orders of magnitudes smaller than swap spreads, but we briefly discuss them here for completeness.

The FICC charges usually a few cents per million dollar notional in transaction processing fees. For example, for DVP repos the fee is \$0.04 per million for dealer accounts and 0.01 -

0.08 basis points for GCF repos FICC (2021). Fedwire fees are charged per transaction and are less than \$1. Given the large notional of repo transactions, these costs are negligible. Even with daily roll-over, based on the numbers above, we estimate the annual transaction fees to be on the order of $\$10^{-3}$ for a \$100 notional. This is too small to have a significant impact the swap spreads.

5 Regulation and Basel III

In addition to transaction costs incurred in both legs when replicating a swap, institutions must also consider funding costs. Following the financial crisis of 2007 - 2009, additional banking regulations were imposed to address the risks realized during the crisis. These regulations resulted in increased capital and regulatory requirements implemented based on the Basel III accord. In particular, the new regulations introduced the supplementary leverage ratio (SLR), a non-risk weighted leverage ratio. While easier to implement than tradition risk-weighted ratios, the new SLR treats all assets/portfolios as equally risky, which results in certain nearly risk-free (hedged) strategies actually requiring more capital requirements than risky ones.

In particular, the SLR is calculated as a ratio between Tier 1 Capital (mostly comprising of common equity) and total leverage exposure (includes on-balance sheet assets and off-balance sheet exposures such as over-the-counter derivatives and repo transactions) using the formula

$$\mathrm{SLR} = \frac{\text{Tier 1 Capital}}{\text{Total Leverage Exposure}}.$$

Banks are required to maintain a 3% minimum SLR, while the 8 U.S. banks identified as global systematically important are required to maintain an SLR of at least 5%.

The most relevant change for our purposes is that interest rate swaps and U.S. Treasury portfolios are calculated separately under the SLR framework. Even though they have the same cashflows, the capital requirements for derivatives and Treasuries are different Haynes et al. (2018). Indeed, consider an institution with an exogenous demand for a pay-float receive-fixed interest rate swap. To construct the desired cashflows, they have two options:

1. enter a swap agreement (this is expensive given the swap spreads), or

2. synthetically construct the swap using bonds and the repo market. (this is cheap given the swap spreads are negative).

Cochran et al. (2023) and Allahrakha et al. (2018) provide evidence suggesting that the SLR is binding. This means that entering large swap positions requires capital/leverage adjustments[4]. Such costs have long been considered as a factor in banking decisions, although not reported. More recently, major banks have started disclosing their funding value adjustments on their balance sheets as described in Albanese and Andersen (2014). In this section, we calculate the funding costs and incorporate them into the swap spread calculation.

5.1 Leverage Exposure: Swaps

Swap exposures are counted under the derivative exposure (off-balance sheet) section. According to Section 20 of Basel (2014), the leverage measure for swaps with a tenor of i years, L_i, is calculated according to

$$\text{exposure measure } (L_i) = \text{replacement cost} + \text{add-on (PFE)}$$

where PFE is potential future exposures. The replacement cost is the price of the swap obtained by marking to market, and is effectively zero for cleared swaps because variation margin is posted daily. The add-on accounts for PFE is calculated by multiplying an add-on factor (published in Paragraphs 1 and 3 of the Annex in Basel (2014)) to the notional of the swap. Combining the factors together, entering a 100 notional swap increases exposure by the amounts shown in Table 6.

	1 Year	2 Year	3 Year	5 Year	7 Year	10 Year	20 Year	30 Year
Exposure	0	0.5	0.5	0.5	1.5	1.5	1.5	1.5

Table 6: Change in Total Leverage Exposure, L_i, per 100 notional

Basel III allows derivatives to be netted for calculating exposure if they satisfy a master

[4]Banks are also subject to a minimum 4% U.S. leverage ratio. The key difference is that SLR takes into account both on-balance sheet and off-balance sheet exposures, while the U.S. leverage ratio only considers on-balance sheet. Thus, SLR is usually more restrictive DavisPolk (2014).

netting agreement. Since swap exposures are already small (relative to Treasuries), we do not account for any netting among swaps. Key to our argument is that swaps do not net with Treasuries or repo market transactions.

5.2 Leverage Exposure: Treasuries

Leverage calculations for repo transactions are detailed under the Securities-financing transaction (SFT) section where the bank is acting as principal (because they are buying/selling from their own inventory). Exposure is calculated as

$$\text{Exposure} = \text{Gross (receivable) SFT assets} + \text{Leverage Exposure Measure.}$$

To replicate a swap with a tenor of i years, the firm needs to enter a repo and sell the bond portfolio. The amount receivable (first term) is equal to the value of the fixed leg. Let E_i be the fair value of assets lent to the counterparty and C_i the fair value of the assets received from the counterparty at the start of the synthetic construction. The leverage exposure measure is calculated using $\max\{0, E_i - C_i\}$. In repo transactions, $C_i \approx 0.98E_i$ because bonds are sold with a haircut. Thus, to replicate an 100 notional swap,

$$\text{Exposure} = E_i + \max\{0, E_i - C_i\} \approx 1.02E_i = 102.$$

of capital is needed at time 0. In general, cash payables and receivables with the same counterparty, same settlement date, etc... may be measured in net. While this may reduce leverage requirements, forecasting future netting is difficult, and institutions cannot consistently rely on netting to cancel out their exposures.

5.3 Funding Costs

In order to enter a swap (either market purchase or synthetic construction), the bank needs to adjust their capital structure. However, in synthetic replication, the bank faces an additional opportunity cost of

$$1.02E_i - L_i$$

in asset holdings at time 0. Notice that future exposure, in general, is a random variable. However, in practice banks do not adjust their capital structure for each repo trade. Instead, the bank estimates the amount of repo trades it wishes to conduct on an on-going basis and adjusts its capital structure once, see Duffie (2018). Thus for our purposes, we view the opportunity cost as a constant. We use $E_i = 100$ (the initial value of the bond portfolio) and L_i from table 6.

We estimate a bank's return on assets using historical data on all banks provided by FDIC (2024). Since swaps have a maximum maturity of 50 years, we use data from the past 50 years (1974 - 2023) for estimation purposes. We find that the sample mean of a bank's return on assets is $a := 0.873\%$, continuously compounded. This estimate does not change significantly when using between 20 to 50 years of historical data.

Synthetically constructing a swap causes the firm to face an opportunity cost of

$$\left(e^{aT} - 1\right)(1.02E_i - L_i)$$

dollars, due to the increased capital required. If the firm chooses to purchase a swap, they can invest in other assets and receive a in expectation. Thus, the funding cost of construction is

$$C_{SLR} := E^{\mathbb{Q}}\left[\frac{\left(e^{aT} - 1\right)(1.02E_i - L_i)}{e^{\int_0^T r(t)dt}}\right] = \left(e^{aT} - 1\right)(1.02E_i - L_i)\, P(0,T).$$

After accounting for funding costs, the new arbitrage-free swap rate is

$$c\sum_{k=1}^{T} P(0,k) + \underbrace{P(0,T) + \sum_{i=0}^{n-1} E^{\mathbb{Q}}\left[e^{\int_i^{i+1}\alpha_s ds} - 1\right] P(0,i) - 1}_{\text{Repo}} + \underbrace{\left(e^{aT} - 1\right)(1.02E_i - L_i)\, P(0,T)}_{\text{Funding Costs}} = 0. \tag{3}$$

Results of estimating this theoretical swap rate are shown in Figure 7. Recall that any spread above zero means that there is no arbitrage. Notice that on four occasions, the 30 year swap spreads have reached the 0 mark before quickly rebounding. This is consistent with our arbitrage free argument: at any rate lower than 0, profits can be made and arbitrageurs step in and prices re-balance quickly.

Our results shows that leverage costs have the largest impact on swap spreads, and this illuminates some of the unexpected consequence of Basel III. The lack of netting across asset classes creates financial frictions that fragment the interest rate markets. More importantly, SLR may provide incentives that increase risk. This is observed in Kiema and Jokivuolle (2014) and Allahrakha et al. (2018) who argue that SLR cause firms to substitute low-risk low-reward strategies in favor of higher-risk higher-reward strategies. In our context, SLR makes hedging interest rate swaps less desirable. The simplest hedging strategy requires Treasury purchases/sales. This entails funding costs. A cheaper (but more risky) strategy is to simply not hedge. Thus, SLR actually encourages the opposite of risk management.

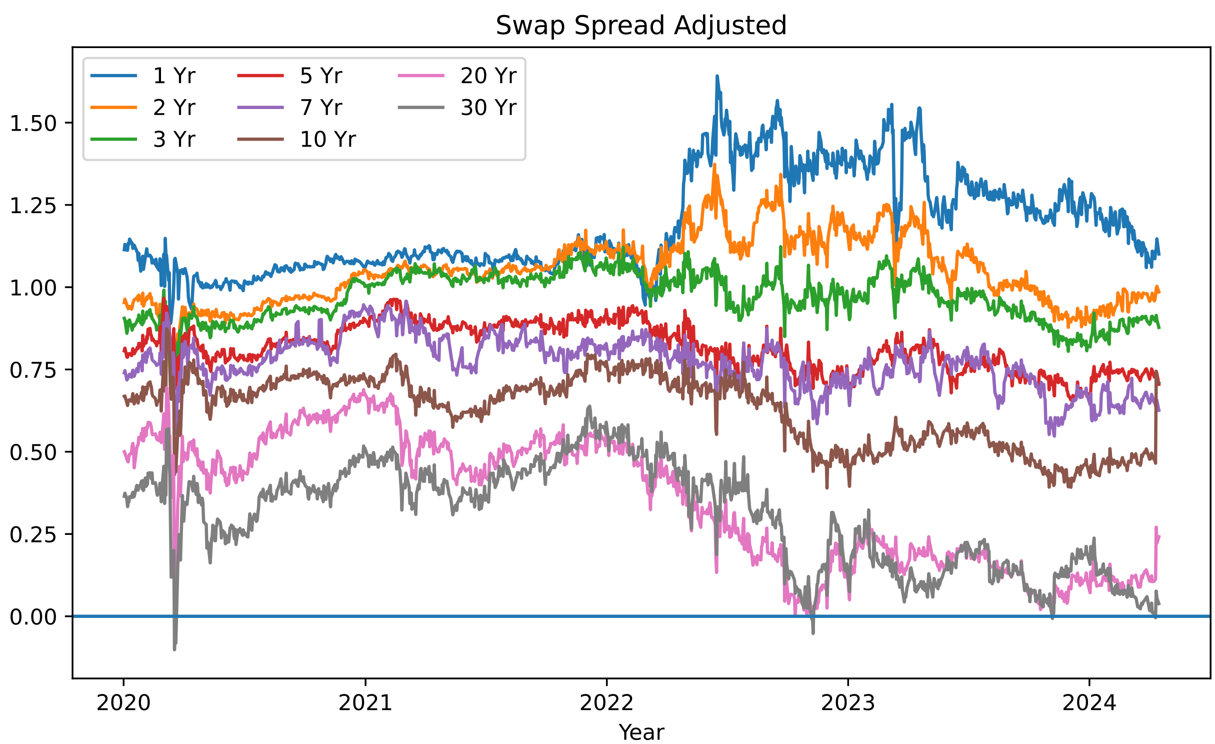

Figure 7: The swap spread after accounting for bid-ask spreads in STRIP markets, repo (GCF) costs, and SLR constraints.

While leverage costs have the largest impact on swap spreads, we emphasize that STRIPS markets and repo costs also play an important role. In particular, we repeated our analysis only accounting for the terms arising from SLR. We find that in this case, 30 year swap spreads are still be negative for much of October 2020 and onwards. This is inconsistent with an arbitrage free market.

6 Conclusion

In this paper, we analyze the swap spread in SOFR markets. The switch to SOFR reference rates actually makes synthetic swap construction easier. It removes credit risk and marking-to-market differences, which were previously attributed to LIBOR swap spreads. However, swap spreads based on SOFR have been persistent and stable during our Jan 2020 - Apr 15, 2024 sample period.

In this paper, we attribute swap spreads to three main factors: treasury markets, repo operational costs, and leverage requirements. We construct accurate yield curves from U.S. STRIPS ask prices and show that swap spreads are not correlated with the timing of U.S. Treasury cashflows. Second, we use the GCF Repo rate as an approximation for the true cost of borrowing after accounting for counterparty risk and operational costs, and the Fed Discount Window's primary rate as an upper bound on short-term overnight borrowing. Finally, we analyze funding opportunity costs imposed by the SLR ratio. We find that, consistent with previous works, funding constraints have the largest impact on swap spreads. However, in addition, bond prices and repo costs also play a role.

References

Paul Agueci, Leyla Alkan, Adam M Copeland, Isaac Davis, Antoine Martin, Kate Pingitore, Caroline Prugar, and Tyisha Rivas. A primer on the gcf repo® service. *FRB of New York Staff Report*, (671), 2014. 4.1

Claudio Albanese and Leif BG Andersen. Accounting for otc derivatives: Funding adjustments and the re-hypothecation option. *Available at SSRN 2482955*, 2014. 5

Meraj Allahrakha, Jill Cetina, and Benjamin Munyan. Do higher capital standards always reduce bank risk? the impact of the basel leverage ratio on the us triparty repo market. *Journal of Financial Intermediation*, 34:3–16, 2018. 5, 5.3

Leif Andersen, Darrell Duffie, and Yang Song. Funding value adjustments. *The Journal of Finance*, 74(1):145–192, 2019. 1

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

Viktoria Baklanova, Adam M Copeland, and Rebecca McCaughrin. Reference guide to us repo and securities lending markets. *FRB of New York Staff Report*, (740), 2015. 4.1

Basel. Basel iii leverage ratio framework and disclosure requirements. *Bank for International Settlements*, 2014. 5.1

BIS. OTC derivative statistics. https://data.bis.org/topics/OTC_DER/tables-and-dashboards/BIS,DER_D7,1.0?time_period=2023-S1, 2023. Accessed: May 6, 2024. 1

Nina Boyarchenko, Pooja Gupta, Nick Steele, and Jacqueline Yen. Negative swap spreads. *Economic Policy Review*, 24(2):1–14, 2018. 1

Damiano Brigo and Fabio Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*, volume 2. Springer, 2006. 4.2.2

Kevin Clark, Adam Copeland, Robert Jay Kahn, Antoine Martin, Mark E Paddrik, Benjamin Taylor, et al. Intraday timing of general collateral repo markets. Technical report, Federal Reserve Bank of New York, 2021. 4.1

CME. Irs clearing operational workflows. http://www.mexder.com.mx/wb3/wb/MEX/MEX_Repositorio/_vtp/MEX/1eee_cursos/_rid/21/_mto/3/9_s_cutinho_IRS_Clearing_Products_Valuation_Margin.pdf, 2014. 2.3

CME. OTC SOFR swaps clearing. https://www.cmegroup.com/education/files/otc-sofr-swaps-product-overview-10-18.pdf, 2018. 1

CME. Cash interest rates and non-cash collateral fees. https://www.cmegroup.com/clearing/financial-and-collateral-management/cash-interest-rates-and-non-cash-collateral-fees.html, 2024. 2.3

Paul Cochran, Sebastian Infante, Lubomir Petrasek, Zack Saravay, and Mary Tian. Dealers' treasury market intermediation and the supplementary leverage ratio. *Board of Governors of the Federal Reserve System*, 2023. 5

Pierre Collin-Dufresne and Bruno Solnik. On the term structure of default premia in the swap and libor markets. *The Journal of Finance*, 56(3):1095–1115, 2001. 1

Adam Copeland, Darrell Duffie, Antoine Martin, and Susan McLaughlin. Key mechanics of the us tri-party repo market. 2012. 4.1

Phillip R Daves and Michael C Ehrhardt. Liquidity, reconstitution, and the value of us treasury strips. *The Journal of Finance*, 48(1):315–329, 1993. 3.1

DavisPolk. Supplementary leverage ratio (slr) visual memorandum. *Davis Polk and Wardwell LLP*, 2014. 4

Wenxin Du, Benjamin Hébert, and Wenhao Li. Intermediary balance sheets and the treasury yield curve. *Journal of Financial Economics*, 150(3):103722, 2023. 1

Simon Dubecq, Alain Monfort, Jean-Paul Renne, and Guillaume Roussellet. Credit and liquidity in interbank rates: A quadratic approach. *Journal of Banking & Finance*, 68: 29–46, 2016. 1

Darrell Duffie. Financial regulatory reform after the crisis: An assessment. *Management Science*, 64(10):4835–4857, 2018. 5.3

Darrell Duffie and Kenneth J Singleton. An econometric model of the term structure of interest-rate swap yields. *The Journal of Finance*, 52(4):1287–1321, 1997. 1

Darrell Duffie and David A Skeel. A dialogue on the costs and benefits of automatic stays for derivatives and repurchase agreements. *U of Penn, Inst for Law & Econ Research Paper*, (12-02), 2012. 4.1

FDIC. Bankfind suite: Find annual historical bank data. https://banks.data.fdic.gov/explore/historical, April 2024. 5.3

Fed. Discount window lending. https://www.federalreserve.gov/regreform/discount-window.htm, 2022. 1

Fed. Discount window lending. https://www.federalreserve.gov/regreform/discount-window.htm, Dec 2023. 4.2

Fed. Selected interest rates (daily) - h.15. https://www.federalreserve.gov/releases/h15/, 2024. 2.2

Peter Feldhütter and David Lando. Decomposing swap spreads. *Journal of Financial Economics*, 88(2):375–405, 2008. 1

FICC. Guide to the ficc gsd fee schedule. https://www.dtcc.com/-/media/Files/Downloads/Clearing-Services/FICC/GOV/FICC-GOVfeeschedule.pdf, Jan 2021. 4.1, 4.3

Jan Philipp Fritsche, Michael Grill, and Claudia Lambert. Bond market liquidity and swap market efficiency â what role does the repo market play? *European Central Bank*, 2020. 1

Samuel G Hanson, Aytek Malkhozov, and Gyuri Venter. Demand-and-supply imbalance risk and long-term swap spreads. *Journal of Financial Economics*, 154:103814, 2024. 1

Richard Haynes, Lihong McPhail, and Haoxiang Zhu. Assessing the impact of the basel iii leverage ratio on the competitive landscape of us derivatives markets: Evidence from options. *CFTF Policy Brief*, 2018. 5

David Hou and David R Skeie. Libor: Origins, economics, crisis, scandal, and reform. *FRB of New York Staff Report*, (667), 2014. 1

Grace Xing Hu, Jun Pan, and Jiang Wang. Tri-party repo pricing. *Journal of Financial and Quantitative Analysis*, 56(1):337–371, 2021. 2.3

ICMA. What is a tri-party repo? https://www.icmagroup.org/market-practice-and-regulatory-policy/repo-and-collateral-markets/icma-ercc-publications/frequently-asked-questions-on-repo/24-what-is-tri-party-repo, Jan 2024. 4.1

Robert Jarrow and Siguang Li. Interest rate swaps: a comparison of compounded daily versus discrete reference rates. *Review of Derivatives Research*, 26(1):1–21, 2023. 2.2

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

Michael Johannes and Suresh Sundaresan. The impact of collateralization on swap rates. *The Journal of Finance*, 62(1):383–410, 2007. 1, 2.3

Bradford D Jordan, Randy D Jorgensen, and David R Kuipers. The relative pricing of us treasury strips: Empirical evidence. *Journal of Financial Economics*, 56(1):89–123, 2000. 1, 3.1

Ilkka Kiema and Esa Jokivuolle. Does a leverage ratio requirement increase bank stability? *Journal of Banking & Finance*, 39:240–254, 2014. 5.3

Sven Klingler and Suresh Sundaresan. An explanation of negative swap spreads: Demand for duration from underfunded pension plans. *The journal of finance*, 74(2):675–710, 2019. 1

Larry HP Lang, Robert H Litzenberger, and Andy Luchuan Liu. Determinants of interest rate swap spreads. *Journal of Banking & Finance*, 22(12):1507–1532, 1998. 1

Jun Liu, Francis A Longstaff, and Ravit E Mandell. The market price of credit risk: An empirical analysis of interest rate swap spreads. *National Bureau of Economic Research Cambridge, Mass., USA*, 2002. 1

New York Fed. Additional information about reference rates administered by the new york fed. https://www.sifma.org/resources/research/us-fixed-income-securities-statistics/, Jan 2022. 1

OFR. Ofr short-term funding monitor - data sets. https://www.financialresearch.gov/short-term-funding-monitor/datasets/repo-single/?mnemonic=REPO-GCF_AR_TOT-P, 2024. 4.2

Maureen O'Hara and Xing Alex Zhou. Anatomy of a liquidity crisis: Corporate bonds in the covid-19 crisis. *Journal of Financial Economics*, 142(1):46–68, 2021. 4.2

Lynn Riggs, Esen Onur, David Reiffen, and Haoxiang Zhu. Swap trading after dodd-frank: Evidence from index cds. *Journal of Financial Economics*, 137(3):857–886, 2020. 1

SIFMA. US treasury securities issuance, trading volume, outstanding, holders, yield curve rates. https://www.sifma.org/resources/research/us-fixed-income-securities-statistics/, 2024. 1

Cheng Yong Tang and Song Xi Chen. Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics*, 149(1):65–81, 2009. 4.2.2

TBAC. TBAC technical report. https://home.treasury.gov/system/files/221/TBACCharge1Q32021.pdf, 2021. 2.1, 2.1

TMPG. Clearing and settlement in the secured financing transaction market-trade flow mapping. https://www.newyorkfed.org/medialibrary/Microsites/tmpg/files/CS_SFT_Maps.pdf, 2022. 4.1

Eric Wallerstein. How a hack shook wall streetâs multitrillion-dollar foundations. *Wall Street Journal*, 2023. 1, 4.1

Martin Wheatley. The wheatley review of libor: final report. Technical report, Discussion paper HM Treasury, London, 2012. 1

David Wu and Robert Jarrow. Fitting dynamically consistent forward rate curves: Algorithm and comparison. *Working Paper*, 2024. 2.2

A Proof of Theorem 4.2

The OU process implies that

$$\alpha(s) = \alpha(0)e^{-ks} + \theta\left(1 - e^{-ks}\right) + \sigma \int_0^s e^{-k(s-u)} dW(u)$$

We need to calculate

$$\int_t^T \alpha(s)ds = \left[\int_t^T \alpha(0)e^{-ks} + \theta\left(1 - e^{-ks}\right) ds\right] + \left[\int_t^T \int_0^s \sigma e^{-k(s-u)} dW(u)ds\right]$$

The quantity in the first bracket evaluates to

$$\mu := \frac{\alpha(0) - \theta}{k}\left(e^{-kt} - e^{-kT}\right) + \theta(T - t)$$

The quantity in the second bracket is equivalent to

$$
\begin{aligned}
&\int_0^T \int_0^T \mathbb{I}(u < s)\mathbb{I}(t < s)\sigma e^{-k(s-u)} dW(u)ds \\
&= \sigma \int_0^T \int_0^T \mathbb{I}(u < s)\mathbb{I}(t < s) e^{-k(s-u)} ds dW(u) \\
&= \sigma \int_0^T \int_{\max\{u,t\}}^T e^{-k(s-u)} ds dW(u) \\
&= \sigma \int_0^T \frac{e^{ku}}{k} \left(e^{-k\max\{u,t\}} - e^{-kT}\right) dW(u) \\
&= \sigma \int_0^t \frac{e^{k(u-t)} - e^{k(u-T)}}{k} dW(u) + \sigma \int_t^T \frac{1}{k} \left(1 - e^{k(u-T)}\right) dW(u)
\end{aligned}
$$

The stochastic integrals are independent and each follow a normal distribution with mean 0 and variance Σ_1^2 and Σ_2^2, respectively. Thus $e^{\int_t^T \alpha(s)ds}$ follows a lognormal distribution and

$$E\left[e^{\int_t^T \alpha(s)ds}\right] = \exp\left(\mu + \frac{\Sigma_1^2 + \Sigma_2^2}{2}\right)$$