The Market for Inflation Risk☆

Saleem Bahaj
University College London
& Bank of England

Robert Czech
Bank of England

Sitong Ding
London School of Economics

Ricardo Reis
London School of Economics

June 2023

Abstract

This paper uses transaction-level data on the universe of traded UK inflation swaps to characterize who buys and sells inflation risk, when, and with what price elasticity. This provides measures of expected inflation cleaned from liquidity frictions. We first show that this market is segmented: pensions funds trade at long horizons while hedge funds trade at short horizons, with dealer banks as their counterparties in both markets. This segmentation suggests three identification strategies—sign restrictions, granular instrumental variables, and heteroskedasticity—for the demand and supply functions of each investor type. Across the three strategies, we find that (*i*) prices absorb new information within three days; (*ii*) the supply of long-horizon inflation protection is very elastic; and (*iii*) short-horizon price movements are unreliable measures of expected inflation as they primarily reflect liquidity shocks. Our counterfactual measure of long-horizon expected inflation in the absence of these shocks suggests that the risk of a deflation trap during the pandemic and of a persistent rise in inflation following the energy shocks were overstated, while since Autumn of 2022, expected inflation has been lower and falling more rapidly than conventional measures.

Keywords: asset demand system, monetary policy, anchored expectations, identification of demand and supply shocks.

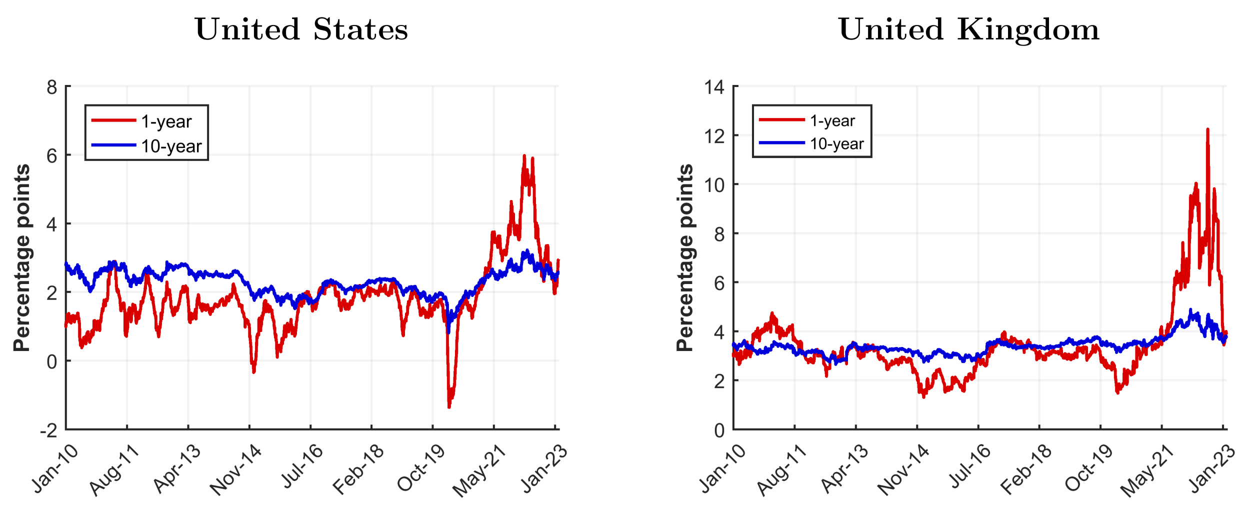
JEL Codes: E31, E44, G12, C30.

☆The authors would like to thank seminar audiences at the Bank of England, Princeton University, University of Cambridge, University of Kent, and University of Liverpool, as well as Markus Brunnermeier, Andrew Hauser, Misa Tanaka and Motohiro Yogo for their comments. E-mail contacts: `s.bahaj@ucl.ac.uk`, `robert.czech@bankofengland.co.uk`, `s.ding5@lse.ac.uk`, and `r.a.reis@lse.ac.uk`. The views expressed in this paper are those of the authors and are not necessarily those of the Bank of England or its committees. Reis is grateful to UKRI grant 109166 for financial support.

1 Introduction

A large and fast-growing market for inflation swaps allows investors to trade inflation risk. Prices in this market are closely watched because they provide a measure of what market participants expect inflation will be that complements (and in some ways is superior to) the measure from inflation-indexed government bonds (the break-even rates). Scores of papers in monetary policy have used these prices to report high-frequency measures of expected inflation at multiple horizons and to study how they respond to identified shocks (see, e.g., Haubrich, Pennacchi, and Ritchken, 2012; Beechey, Johannsen, and Levin, 2011). Speeches by policymakers often show plots of these prices over time, especially of forward swaps at long horizons, to assess whether monetary policy is credibly delivering price stability (see, e.g., Mann, 2022; Ramsden, 2022). More recently, the large gap between 1-year and 10-year measures of expected inflation in the US and the UK shown in Figure 1 has been interpreted as revealing that markets believe that current high inflation will be temporary and that expectations have stayed anchored.

Figure 1 PRICES OF INFLATION SWAPS IN THE RECENT PAST



NOTE: Prices shown are zero-coupon swap breakeven rates. SOURCE: Bloomberg.

This paper studies the relevant players and quantities behind these prices. Using detailed regulatory transaction-level data on every over-the-counter (OTC) inflation swap contract sold in the UK, we identify who buys and sells inflation insurance on a daily basis and study how the prices are formed. With a model of demand and supply, we decompose the price movements into fundamentals and liquidity shocks. From the joint variation in prices and quantities, we learn which shocks drive prices at different dates, and what are the slopes of demand and supply curves for inflation protection by different investor types. With this who and what, we then revisit the account of how inflation expectations moved during the pandemic, the subsequent period of increasing inflation, and the UK mini-budget crisis of September and October of 2022.

Contributions and outline: We provide three contributions. As the first contribution, Section 2 establishes three stylized facts on the UK market for inflation swaps. First, dealers are not neutral market makers, but they are net sellers of inflation protection, beyond their holdings of index-linked gilts. Moreover, increases in the sale of swaps are only weakly correlated with changes in the holdings of indexed bonds. Second, at long horizons (10 years or longer) the buyers of the swaps are mostly pension funds. They hold large, persistent, positive net positions, and much of the variation in sales and trading volume is driven by the actions of pension funds. Thus, pension funds are net receivers of inflation in this market while dealers are the net payers: when inflation rises unexpectedly, there is a direct flow of payments from dealers to pension funds. Third, at short horizons (3 years or less), informed traders like hedge funds hold small net positions, but actively trade, so that in any given day they can have a negative or positive net position. When short-term inflation rises, sometimes they win, sometimes they lose, and on the other side, dealer banks likewise have fluctuating net positions. In other words, we observe a remarkable *segmentation* of this market: pension funds barely trade in the short-horizon market while informed traders have most of their buying and selling activity in the short-horizon market.

Motivated by these facts, Section 3 provides a model of the market for inflation risk with two main characteristics. First, it is a portfolio choice model, as opposed to a pure model of broker-dealers, since dealers in this market hold persistent large net positions. Second, it features two markets where inflation risk is traded, with one agent — dealer banks — supplying the inflation swap contracts to meet the demand in both markets. On the other hand, the demand in each market arises from either hedge funds or pension funds depending on whether the market offers inflation swap contracts at short or long horizons, respectively. The model clarifies the meaning of demand and supply in these markets, provides a simple decomposition of inflation swap prices into expected inflation, compensation for risk, and liquidity frictions, and identifies the primitives and market frictions that can give rise to them.

Using the model, our second contribution is to propose three identification strategies in Section 4 to estimate the demand system for inflation protection across institutions. The first strategy exploits the high frequency of the data. It assumes that, within a day, hedge funds respond more to fundamental expectations of inflation than banks, and banks respond more than pension funds, while the supply by banks across the two segmented markets adjusts to each other with a one-day delay. These assumptions impose restrictions on the relative shifts of demand and supply functions in response to shocks that amount to sign restrictions on the structural responses of prices and quantities to various shocks. The second strategy exploits the cross-sectional variation in the data, since we have transaction-level data that allows us to observe the daily trading activities at an institutional level. This cross-sectional variation is highly granular, with some institutions taking larger positions that fit a power law distribution. We construct a granular instrumental variable, using institution-level disturbances as instruments for the aggregate demand for each class of agents. The final identification strategy exploits the heteroskedasticity in our time series data. We have nearly four years of data at a daily frequency. Releases of inflation data are regular and cause heightened volatility in the inflation swap market. Assuming that shocks to fundamentals drive most of the heteroskedasticity on these dates, we use the time-varying volatility to identify these shocks.

Section 5 presents our third and final contribution: estimates from the UK inflation swap market using daily data from January 2019 to February 2023. Consistently across our three identification strategies we find that:

1. Impulse responses to shocks become horizontal within two to three days. The inflation swap market seems to incorporate new information relatively quickly.

2. Supply of inflation protection by dealer banks to pensions funds at long horizons is very elastic, unlike supply to hedge funds at short horizons.

3. Most of the price variance of short-horizon inflation swaps is driven by liquidity shocks. In contrast, fundamentals dominate at long horizons. Across both markets, liquidity shocks to dealer banks account for only 10% to 30% of the variability in prices, while liquidity shocks to pension funds or hedge funds play a large role.

4. At the long horizon, we produce estimates of expected inflation driven by fundamentals, cleaned of liquidity frictions. During the pandemic, the uncleaned conventional measures overstated the risk of deflation, and during the energy crisis overstated the risk of inflation. From September 2022 to February 2023, 10-year expected inflation has been below what swap prices indicate, providing a more optimistic view on inflation expectations being anchored.

Connection to the literature: Our paper is related to five strands of the literature. First, we contribute to the literature on segmented markets that builds on Vayanos and Vila (2021). We highlight a particular market where the segmentation across horizons is clear. While this market for inflation risk is not as large as the markets for bonds and foreign exchange that the prior literature has focused on, it has the virtue that we have transaction-level data on a large sample of the market. Therefore, we can cleanly empirically identify the arbitrageurs and preferred-habitat agents that are described in those theories. We provide three complementary identification strategies that may be useful in other estimations of segmented market models.

Second, we estimate an asset demand system, in the footsteps of Koijen and Yogo (2019) or more recently Gabaix, Koijen, Mainardi, Oh, and Yogo (2022). While previous studies have used data on stocks (e.g., Koijen, Richmond, and Yogo, 2020), bonds (e.g., Koijen, Koulischer, Nguyen, and Yogo, 2021) or exchange rates (e.g., Koijen and Yogo, 2020), we focus on the large and liquid other-the-counter market for inflation swaps. Our approaches to identification are different, as is our focus on extracting measures of expected inflation from the prices of swaps. Within this literature, in our use of granular instrumental variables (Gabaix and Koijen, 2020), we are joined by Gabaix and Koijen (2021). Differently, we use trade-level positions across different investor types to build a granular instrumental variable for the demand of each type.

Third, and related to the previous two, Begenau, Piazzesi, and Schneider (2015), Hanson, Malkhozov, and Venter (2022), Jiang, Matvos, Piskorski, and Seru (2023) and McPhail, Schnabl, and Tuckman (2023) focus on the interest rate swap market to uncover the types of investors who bear interest rate risk and how their exposure to risk varies. Begenau, Piazzesi, and Schneider (2015), Jiang, Matvos, Piskorski, and Seru (2023) and McPhail, Schnabl, and Tuckman (2023) estimate the positions taken by banks, while Hanson, Malkhozov, and Venter (2022) assume a risk profile for the intermediating dealers as a proxy for their positions. Instead, we focus on inflation derivatives, and we directly observe the directional positions taken on by dealer banks and other investors. Our results, in contrast to the evidence by Jiang, Matvos, Piskorski, and Seru (2023) and McPhail, Schnabl, and Tuckman (2023) on interest rate swaps, show that banks do take large net positions in the inflation risk market. We also find that liquidity shocks affecting the clients of dealers play an important role in driving market prices alongside shocks to dealers themselves. Like Hanson, Malkhozov, and Venter (2022), we use an affine representation of the structural shocks and, in one of our identification strategies, we use sign restrictions, but we further complement this with other strategies so we can cross-validate the results, all pointing to the same conclusions on what drives the prices of inflation swaps.

Fourth, our focus on inflation risk and how it is priced by financial markets is shared with a long literature: see Cieslak and Pflueger (2023) and D'Amico and King (2023) for recent surveys.[1] A perennial concern in this literature is whether market prices reflect not just subjective expectations and compensation for risk, but also liquidity premia, a catch-all term for market imperfections that can be large and vary over time. We address this challenge head on, contributing to a better understanding of where these premia come from, and providing superior estimates of market-expected inflation.

Fifth, and finally, we use the regulatory EMIR Trade Repository (TR) data on trade-level derivative positions to shed light on OTC derivative markets (Abad et al., 2016). Cenedese, Della Corte, and Wang (2021) use the EMIR TR data on FX forwards and swaps to show that the leverage ratio requirement causes deviations from the covered interest parity. Hau, Hoffmann, Langfield, and Timmer (2021) also use the EMIR TR data on FX derivatives to document price discrimination against non-financial investors. Cenedese, Ranaldo, and Vasios (2020) show that

[1]Recent contributions include, among others, Campbell, Pflueger, and Viceira (2020), Reis (2020), Boons, Duarte, de Roon, and Szymanowska (2020) and Fang, Liu, and Roussanov (2022).

clients in the interest rate swap market pay a higher price when buying interest rate protection over-the-counter from a dealer bank, rather than using centrally-cleared trades.[2] To the best of our knowledge, our work is the first to use the granular EMIR TR data on inflation swaps, which enables us to reveal who is buying the swaps, who is selling them, and the associated gross and net positions.

This paper is organized as follows. Section 2 describes the granular transaction-level data and the set of novel stylized facts that characterize the UK inflation swap market. Section 3 presents the model and a formal characterization of the asset demand system. Section 4 discusses the set of three identification strategies and their empirical implementation. Section 5 presents our estimates of liquidity shocks, counterfactual measures of expected inflation as well as the slopes of the demand and supply functions in both the short and long horizon markets. Section 6 then concludes.

2 Data, summary statistics, and stylized facts

We start by describing the market, the source of our novel data, and some summary statistics for the UK inflation swap market. Then, we establish three main facts.

2.1 The inflation swap market

In this paper, we focus on the dealer-client segment of the UK inflation swap market, where dealer banks (dealers for short) sell the contracts to different financial institutions, including pension funds and liability driven investors (PFLDIs, or just pension funds, for short) to informed traders (i.e. hedge funds) and to others (including non-dealer banks and insurance companies among others). This is an over-the-counter (OTC) market, where the terms of a transaction are negotiated privately between the two counterparties involved in the trade.

Most contracts take the form of zero coupon swaps linked to an index measure of inflation. This is a contract where cash changes hands at the end of the swap contract on a legally binding settlement date. The floating rate payer pays one plus the total growth of the inflation index over the life of the contract times the gross notional. The fixed rate payer pays the compounded fixed rate times the notional. The fixed rate of the swap contract is set ex ante such that the net present value payoff of the swap to both counterparties is equal to zero at initiation — for this reason it is commonly referred to as the swap breakeven rate — while the realization of the floating inflation is uncertain, and so the liability associated with this leg is determined ex post.

[2] Other recent contributions include, for instance, Benos, Ferrara, and Ranaldo (2022), Czech, Della Corte, Huang, and Wang (2022) and Ferrara, Mueller, Viswanath-Natraj, and Wang (2022).

There is inflation risk since, when realized inflation rate deviates from the fixed rate, one counterparty becomes liable to pay insurance to the other counterparty through a net cash flow payment based on the notional written on the swap contract. The inflation measure for the UK inflation swap market is the retail price index (RPI), as this dominates nearly all swap contracts traded on UK inflation, partly because this is the index used in inflation-linked gilts. In the past decade, the RPI tends to be 1.5 percentage point or more above the CPI, which is the base of the 2% target of the Bank of England.[3]

2.2 The EMIR Trade Repository Data

Regulatory efforts to enhance the transparency of derivatives markets intensified after the Great Financial Crisis. During the G20 summit in September 2009, it was agreed that derivative trades should be reported to trade repositories, thus granting regulators access to high-quality and high-frequency data. In the UK, the commitment to increase the transparency of derivative markets was implemented as part of the European Market Infrastructure Regulation (EMIR), which makes it mandatory for UK legal entities to report the terms of any derivative transaction to a trade repository authorized by the Financial Conduct Authority (FCA) by the next business day.[4]

We rely on the EMIR Trade Repository (TR) data to obtain trade-level information on OTC inflation swaps. Our data sample spans the period from January 2019 to February 2023, and it consists of all trades submitted to DTCC Derivatives Repository Plc ('DTCC') in which at least one of the counterparties is a UK-regulated entity. Since this is the largest trade repository in terms of market share, we are confident that it captures a volume of derivatives trading data that is highly representative of the market (Abad et al., 2016). We use the DTCC's daily trade state reports to capture all outstanding inflation swap trades on a given day, as well as to obtain the flow of trading activities. We allocate investors to groups using a best-endeavor sectoral classification.

The raw data has approximately 3.5 billion observations. As expected, it is noisy. We identify the inflation swap contracts from all other interest rate derivatives and clean the data by eliminating: (i) duplicated transaction-level reports at the counterparty-trade ID-other counterparty level, (ii) intragroup transactions, (iii) compression trades, (iv) trades with implausible notional amounts (greater than $10bn and lower than $1000), and (v) trade reports that do not meet the

[3]The Appendix further discusses the comparison between price indices. In 2020, the UK chancellor announced that RPI is to be aligned with CPIH "no earlier than Feb 2030", with no compensation for holders of index-linked gilts. However, the market is yet to price this transition in its entirety, likely due to expectations of a delay or possible compensation. Given the slow-moving nature of the transition, our estimates of high-frequency movements in RPI swap prices should not be affected by this future change.

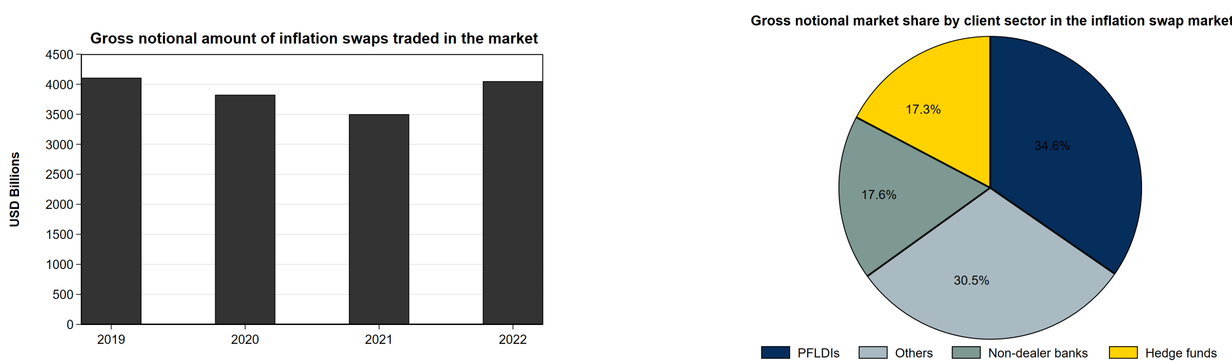
[4]Since the UK's departure from the European Union, the EMIR has been adapted into UK legislation under UK-EMIR. As of February 2023, there are four trade repositories regulated by the Financial Conduct Authority that are registered under UK-EMIR to operate in the UK. These are: DTCC Derivatives Repository Plc, UnaVista Limited, REGIS-TR UK Limited and ICE Trade Vault Europe Limited.

set of UK EMIR validation rules. Finally, we drop all observations with inconsistent values for either the reported notional, the identities of the counterparties, the counterparty side, the maturity date, or the underlying inflation index.[5] Finally, since the vast majority of trades are spot contracts (rather than forward contracts), we focus on this predominant segment of the inflation swap market. This leaves us with more than 25 million observations, on a daily basis, from 2^{nd} January 2019 to 10^{th} February 2023.

2.3 The size of the market and its main participants

The left panel of figure 2 shows the average gross notional amount outstanding across the sample. The market size is substantial, varying between \$3.5tn and \$4tn, which is roughly 110-130% of UK GDP. Approximately 60% of the total gross notional amount outstanding in the market is in the centrally-cleared market segment (CCP) that only clearing members (mainly dealer banks) have direct access to.[6] We instead focus on the dealer-client segment of the market, which is intermediated by 18 of the largest international banks. The right panel of Figure 2 shows the main sectors of the client side of the market: PFLDIs, hedge funds, non-dealer banks, and others, which includes insurance companies, non-financial and other types.[7]

Figure 2 GROSS NOTIONAL AMOUNTS: AVERAGE AND COMPOSITIONS



NOTE: The left figure shows the stock of inflation swap contracts outstanding, traded by all investors in the market in a given year, averaged across monthly trade state reports at month-end. The right figure shows the distribution of total gross notional traded by various client institutions, computed as an average across all monthly trade state files SOURCE: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to December 2022.

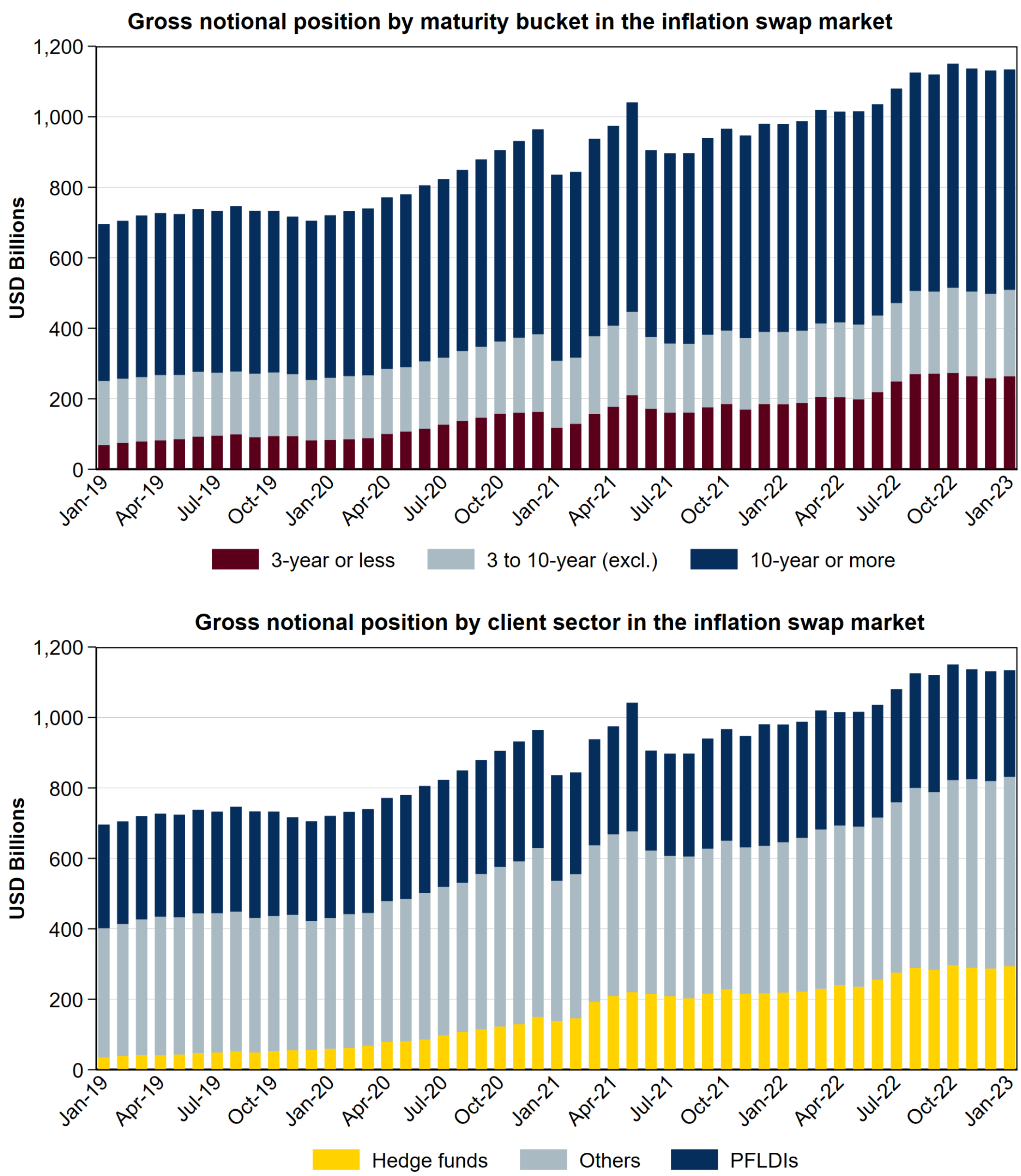
[5] We provide more details on the various cleaning steps in Section A.1 of the Appendix.

[6] The remaining ≈40% of trades are intermediated by these dealer banks, and can be divided into inter-dealer trades (17%) and dealer-client trades (22%). Transactions between a clearing member and the CCP appear as a unique trade in the dataset, although they may represent one half of a trade with another clearing member, as well as one or multiple trades between the clearing member and her client. By focusing on the dealer-client segment of the CCP market, we avoid double-counting.

[7] The full list of investor types also includes asset managers, sovereign wealth funds, trading services, proprietary trading firms, central banks, state and supranational institutions, but these tend to be small.

Figure 3 plots the gross notional positions outstanding in the dealer-client segment of the inflation swap market since 2019, split by maturity and by client sector. Following the rise in inflation since 2021, these have grown rapidly reaching a peak of around \$1.1tn in late 2022, as more agents have bought and sold protection against inflation. Across horizons, the market is dominated by contracts for inflation over a short horizon, from the present to three years away or less, and contracts for a long horizon of ten years from now or more. Across clients, hedge funds have steadily increased their notionals since the Covid-19 market turmoil in 2020 and the reappearance of inflation, from less than \$50bn in 2019 to around \$200bn in 2022.

Figure 3 THE QUANTITIES AND THE INSTITUTIONS BEHIND THE PRICES



NOTE: Gross notional positions outstanding that are traded in the dealer-client segment of the market across all price indices. SOURCE: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to December 2022.

2.4 Three facts on the segmented UK inflation swap market

To establish the main facts, we focus on net notional amounts. These are defined as the sum of net protection bought (sold) by counterparties that are net buyers (sellers) of inflation swaps.

The first fact is that dealer banks are not neutral market makers. Rather, as illustrated by the top panel of 4, since at least the second half of 2019, dealer banks have issued an amount of inflation protection in this market that is beyond their holdings of index-linked government bonds (the natural hedge against their inflation risk exposures).[8] The magnitude of their exposures is substantial, and reached a peak of almost $150bn in 2020. Rather than being driven by the risk-taking behavior of a few large dealer banks, this phenomenon is consistent across the entire sector: nearly all dealer banks sell inflation protection and thereby take on inflation risk.

The bottom panel of Figure 4 reveals that it is primarily PFLDIs that take the opposite position to dealers. They have persistently large and positive net notional positions in this market, with a maximum of approximately $100bn in our sample. Given their gross notionals of around $250bn in 2022 (see Figure 3), the PFLDI's high net-to-gross ratio emphasizes the largely one-directional appetite of PFLDIs for buying inflation protection.[9] Hedge funds, in contrast to PFLDIs, tend to switch between being net buyers and sellers of inflation swaps, consistent with prior studies documenting the role of hedge funds as arbitrageurs and informed traders in derivative markets (see, e.g., Czech, Della Corte, Huang, and Wang, 2022).

The second fact is that PFLDIs are buyers of inflation protection mainly at at long horizons. Figure 5 shows the net notional positions of both PFLDIs and dealers broken down by the initial time-to-maturity of the contracts. PFLDIs hold a sizable amount of inflation swap contracts of 10-year initial maturity or longer, relative to their overall net position. This persists over time, and is consistent with PFLDIs seeking to buy inflation protection to hedge their long-dated liabilities. Dealers make a market to meet this demand, and persistently hold a sizable amount of inflation risk in this segment.

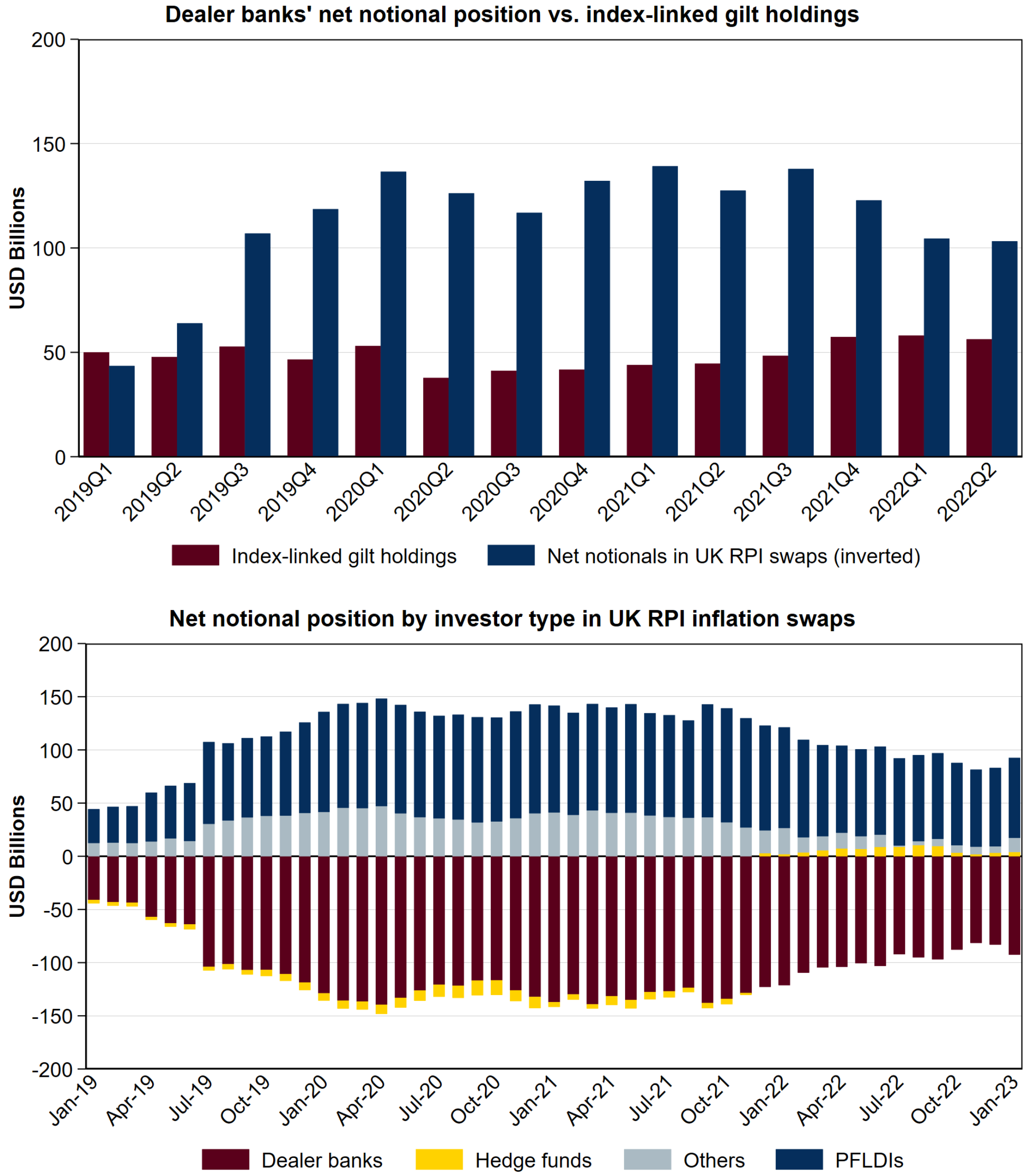
The third fact is that hedge funds trade inflation risk primarily in the short horizon markets. Figure 6 shows the net notional position taken on by hedge funds. They primarily trade inflation swaps with an initial maturity of 3 years or less. Further, they switch between being net buyers and net sellers, which is consistent a priori with them following informed trading strategies that try to exploit arbitrage opportunities.

Combined, these three facts imply a remarkable *segmentation* of the UK RPI market: PFLDI

[8]To obtain data on the index-linked gilt holdings of dealer banks, we use the UK banking system's Global Network of granular exposures, which captures roughly 90% of the UK banking system's total assets. The data and cleaning procedure is described in Covi, Brookes, and Raja (2022).

[9]The pronounced appetite of pension funds for inflation protection may be driven by the limited supply of index-linked gilts, which cover only a small fraction of pension funds' total liabilities, as shown in Figure E.2 in the Appendix.

Figure 4 DEALERS HAVE A NON-ZERO NET EXPOSURE TO INFLATION RISK

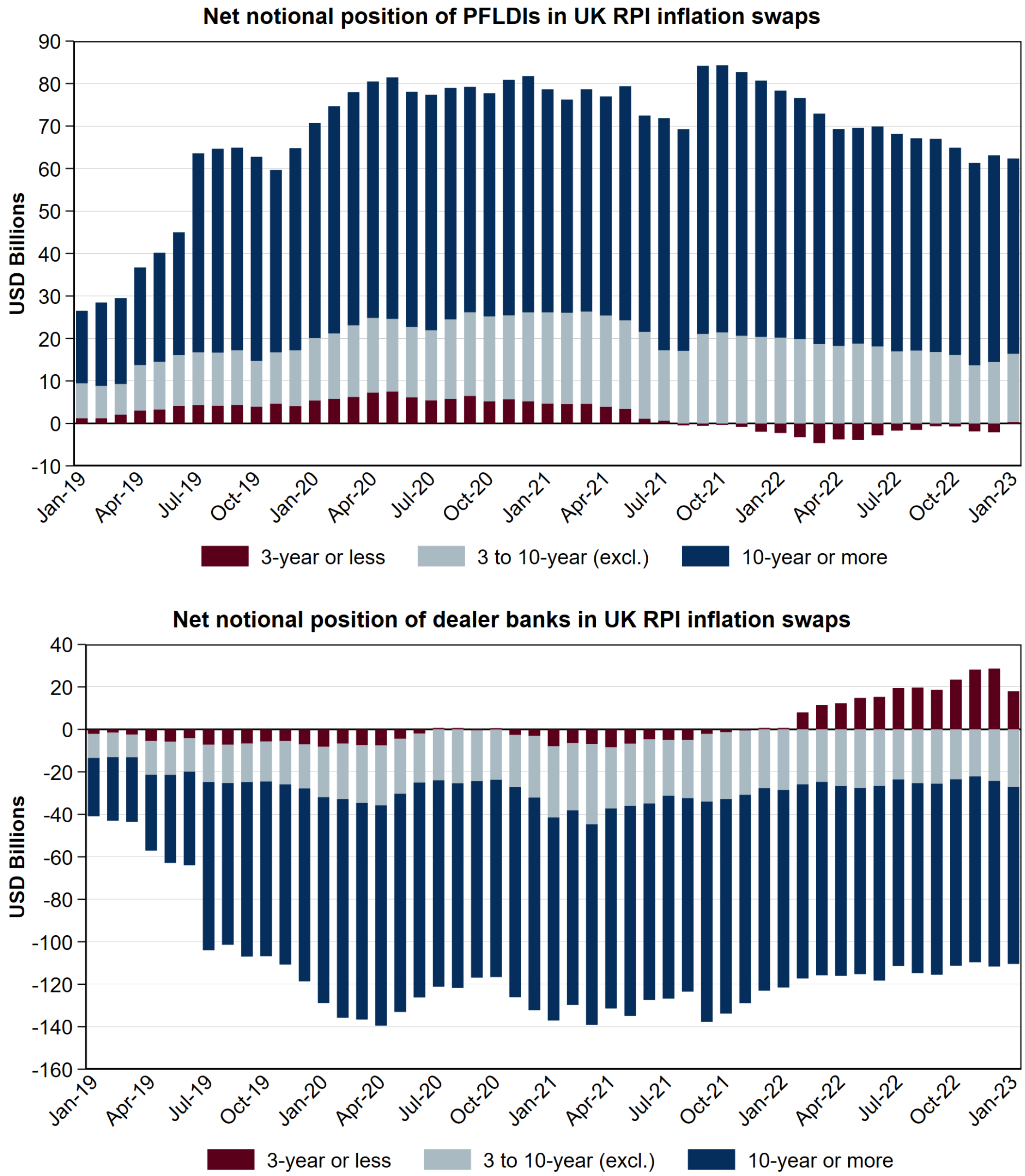


SOURCE: UK banking system's Global Network of granular exposures and DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

institutions primarily trade in the long horizon market where they hold persistently large positive net positions, hedge funds trade in the short horizon markets with fluctuating net positions, and dealer banks are the counterparties in both markets to both types of clients, so they trade actively in both. Dealers fit into the category of arbitrageurs across maturities while PFLDIs and hedge funds fit into the one for preferred-habitat investors in the spirit of Vayanos and Vila (2021).

Looking for more evidence of this market segmentation, we utilize the high frequency of the data. The top panel of Figure 7 shows the gross notional positions aggregated at the sectoral level at a daily frequency. Clearly, PFLDI's gross trading volume in the long horizon market dominates

Figure 5 Pension funds buy protection from dealers in the long horizon market

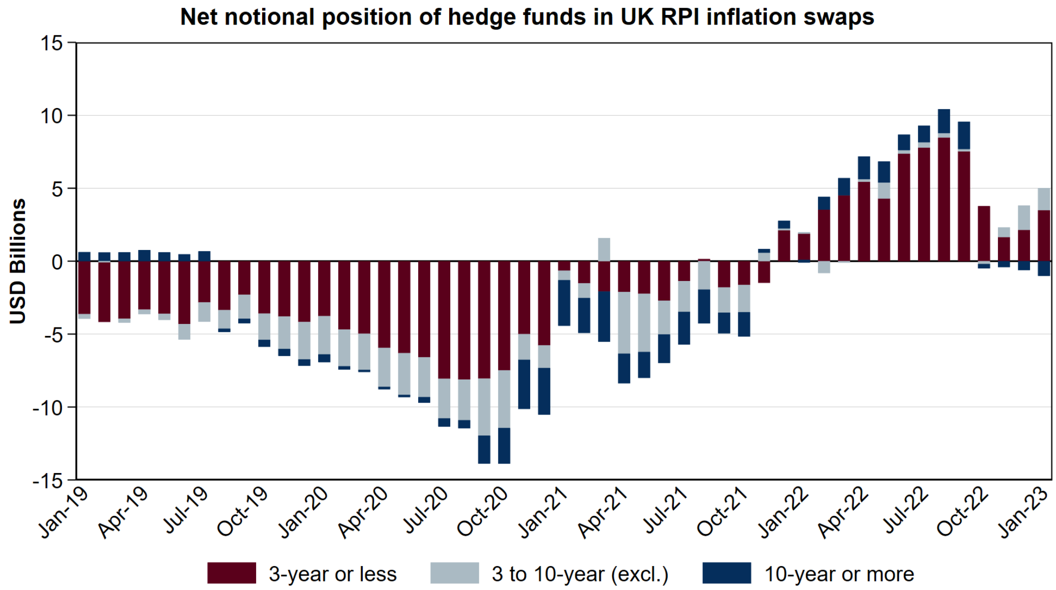


Source: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

that from the hedge funds, and vice versa in the short horizon market. There is some trading activity of hedge funds in the long horizon market, and of PFLDIs in the short horizon market, but these are small relative to the predominant type in the market, or relative to their large positions in the other market.[10] The bottom panel of Figure 7 instead calculates the median maturity of trades from both the PFLDI and the hedge fund sectors. The segmentation of the markets is also clear here.

[10] Appendix E.4 further plots gross notional positions and amounts traded across horizons by investor type. Even by this measure, PFLDIs are mostly active in the long horizon market, and hedge funds trade large amounts in the short horizon market. Hedge funds sometimes take positions in the long horizon market, but these are almost always small compared to their positions in the short horizon market.

Figure 6 HEDGE FUNDS ARE ACTIVE IN THE SHORT HORIZON MARKET



SOURCE: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

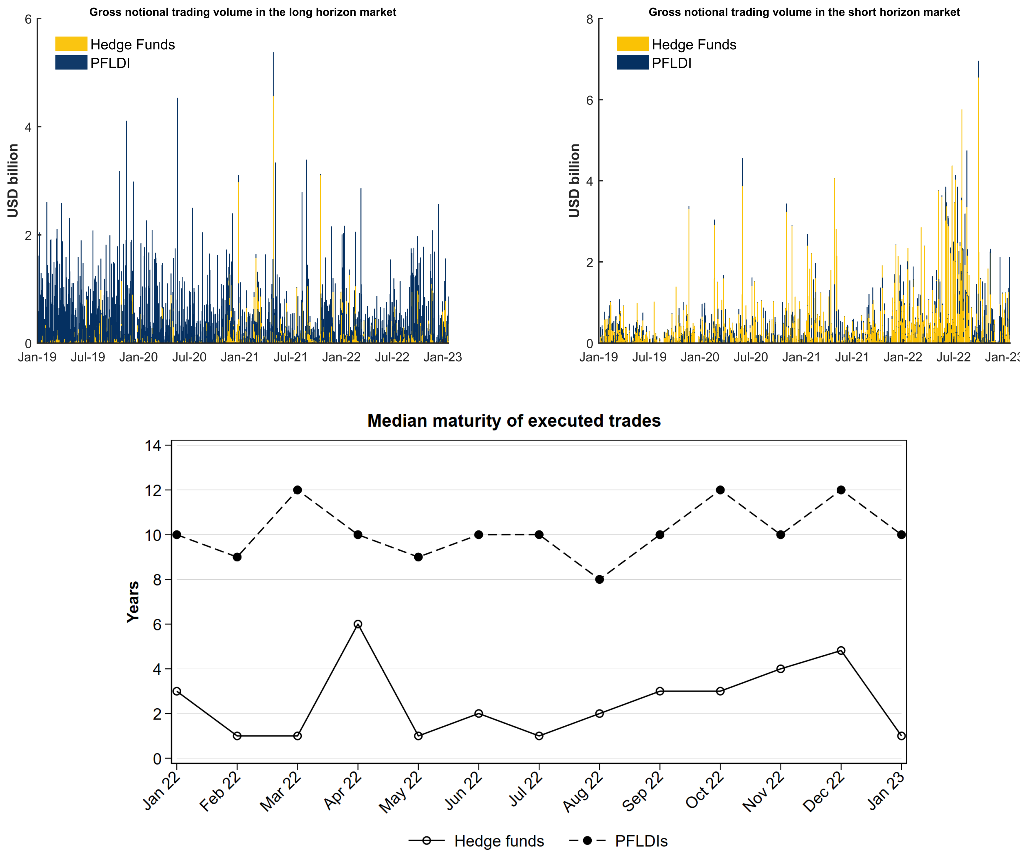
2.5 A robustness check: comparison to alternative data

To check that our trade repository dataset is representative of the trading activity in the OTC inflation swap market, we collect supervisory data on the derivative holdings of insurance companies that are regulated by the UK's Prudential Regulation Authority (PRA) and subject to the Solvency II Directive. Most insurers within scope of the Solvency II Directive are required to submit annual and quarterly returns, with the exception of some smaller firms with quarterly waivers. The reports include detailed information on the derivatives holdings of a given insurer, including the identity of the counterparty, the underlying security, the notional amount, and the derivative category (e.g., inflation swap). Given the supervisory nature of the reporting, we can assume that the Solvency II data provide an exact quarterly snapshot of the total inflation swap holdings in the insurance sector.

Figure 8 compares the average gross notional outstanding of the insurance sector in our dataset with the supervisory holdings in the Solvency II data for the period 2019 Q1 - 2022 Q4. The figure shows that the EMIR TR data cover the vast majority of trading activity in the OTC inflation swap market as reported to the Solvency II database. In 2022 Q4, for example, both datasets report a gross notional of around \$320bn for the UK insurance sector.[11] Throughout our sample period, the EMIR TR data cover almost 90% of the total inflation swap holdings reported to the Solvency II database. The improved coverage in the second half of our sample is likely due to the increased precision of the regulatory reporting in the EMIR TR data.

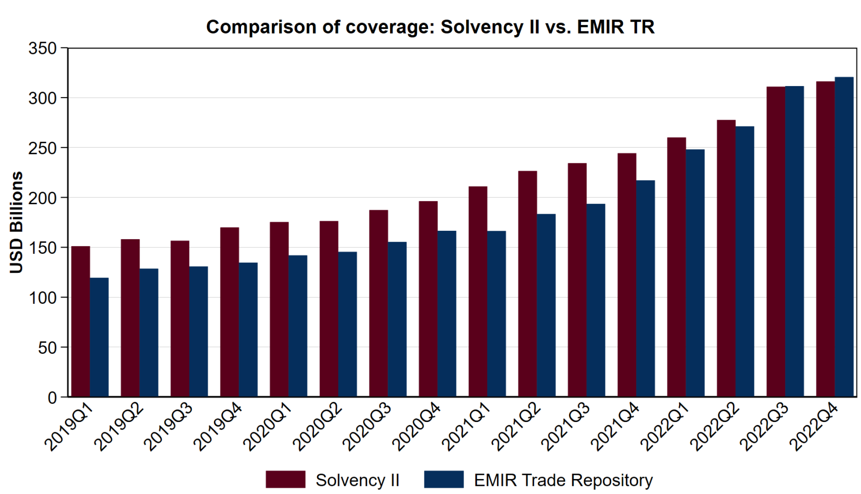
[11] Section D of the Appendix describes the mechanical nature of insurers' inflation swap trading in recent periods in more detail. Our baseline results remain stable when including insurers' trading volumes.

Figure 7 Market segmentation at higher frequencies



Note: The top panel plots the gross notional traded by both the hedge fund and PFLDI sectors at a daily frequency for a given execution date of the trade. This is aggregated from contract-level data pertaining to each institution belonging to both sectors. Source: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

Figure 8 Comparison of Solvency II insurance holdings and EMIR TR data



Source: Solvency II and DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

3 A model of the demand for inflation risk

A model that fits our facts should include three types of agents where all, including dealers, take non-zero net positions and hold on to them, and (at least) two separate markets, for both short-horizon and long-horizon inflation risk, with hedge funds in the first and pensions funds in the latter, while dealers are active in both. This section offers such a model. It will impose few restrictions, beyond some functional forms and a structure of uncertainty that provides analytical linear demand functions. The goal is not to provide a structural representation to directly estimate in the data, but rather to have a vehicle to clarify concepts, catalog different types of shocks, and later state clearly the identification assumptions.

3.1 Goals, beliefs, trading constraints and markets

There are three types of agents: two clientele traders that are the pension funds (f) and hedge funds (h), and a dealer bank sector (b) acting as their counterparties. There are many institutions within each group, indexed by i. The objective of each institution (and taking pension funds for concreteness) is to solve a static portfolio allocation problem where they maximize utility of terminal wealth $a'_{f,i}$ according to a constant absolute risk-aversion utility function:

$$\mathbb{E}_{f,i}\Big[-\exp\left(-\widetilde{\gamma}_{f,i}a'_{f,i}\right)\Big] \tag{1}$$

The Arrow-Pratt coefficient $\widetilde{\gamma}_{f,i} = \gamma_{f,i}/a_{f,i}$, so that it can scale with initial wealth $a_{f,i}$; otherwise, larger institutions would mechanically hold riskier positions.

We interpret the single period in our model as a trading day. An inflation swap contract for the short horizon is an asset that costs a fixed coupon p at the end of the day, and pays off π. This is not the realization of inflation per se, but rather the realization of the value of the floating portion of the swap contract at the end of the day. Swap contracts are subject to variation margin so that counterparties exchange cash to ensure that the contract's net value remains at zero at the end of the period, so these payments can be interpreted as meeting a margin requirement. An equally valid interpretation is that the position is closed down at day-end with the agent whose side of contract appreciated receiving a payment.

Normalizing the daily real safe rate to 1, the budget constraint of a pension fund is:

$$a'_{f,i} = a_{f,i} + (\pi - p)q_{f,i} + (d - s)e_{f,i} + y_{f,i} \tag{2}$$

Aside from swaps, the agent also trades a market asset that costs s and pays d. This could be the result of a portfolio choice involving many other assets, but as we will focus solely on the demand for inflation risk, taking $e_{f,i}$ to be the total amount invested in those assets, we do not need to solve for its components. Finally, $y_{f,i}$ is a background risk that cannot be traded due to incomplete markets, and which may be correlated with inflation.

Let $\Theta_f \in \mathbb{N}$ denote the set of institutions in the pension fund sector. Within this set, each institution has individual beliefs captured in the expectations operator $\mathbb{E}_{f,i}(.)$. For the expectations of asset payoffs and background risk, we assume they all agree that $\mathbb{E}_{f,i}[d] = \theta_d$ and $\mathbb{E}_{f,i}[y_{f,i}] = 0$, as this is not our focus.[12] Rather, we focus on disagreement over expected inflation. Namely:

$$\mathbb{E}_{f,i}(\pi) = \mu_{f,i}\pi^e \quad \text{with} \quad \sum_{i\in\Theta_f} \mu_{f,i} = 1, \tag{3}$$

such that the parameters $\mu_{f,i}$ capture this heterogeneity in inflation expectations and π^e is the *fundamental* expected inflation. Note that this is without loss of generality, as it simply defines what π^e is as the average across expectations.

We assume that all the institutions believe that returns are normally distributed so that they solve a mean-variance optimization problem. The variances of the three exogenous random variables are σ_π^2, σ_d^2, and $\sigma_{y_{f,i}}^2$, while we denote the covariances of expected inflation with market returns and background risk by $\sigma_{\pi,d}$ and $\sigma_{d,y}$, respectively, and their associated correlations by $\rho_{\pi,d}$ and $\rho_{d,y}$.[13]

Finally, we allow for relatively general capacity constraints on each institution's ability to take on inflation risk. These capture regulatory constraints, balance sheet constraints, investment mandates or limitations on short sales. The constraint is given by a continuous function:

$$G_f(q_{f,i}, z_{f,i}) \geq 0 \tag{4}$$

that measures the proximity of the pension fund to the capacity limits. The $z_{f,i}$ is an exogenous institution-specific shifter in the tightness of these financial constraints, while $q^*_{f,i}$ is the net notional exposure in equilibrium. We will use $\lambda_{f,i}$ to denote the Lagrange multiplier associated with this constraint at the optimal choice, and write $g_{f,i} \equiv \partial G_f(q^*_{f,i}, z_{f,i})/\partial q_{f,i}$ for brevity, omitting the arguments of the function.

A similar problem describes the actions of hedge funds in the short horizon market, where P, Π, and $Q_{h,i}$ are the prices, payoffs, and net holdings of short-horizon swaps, respectively, and $\Theta_h \in \mathbb{N}$ is the set of institutions in that sector. To be clear about the key restriction of *segmented markets* that we are imposing, and because we will refer to it later, the formal assumption is:

[12] Allowing these to be institution specific would complicate the algebra but leave all the propositions unchanged.

[13] Again to focus on inflation risk, and in order to reduce the length of the expressions, we set the correlation between market returns and background risk to zero: $\sigma_{d,y_{f,i}} = 0$.

Assumption 1. *(Segmented markets.) Pension funds do not participate in the short-horizon market $Q_{f,i} = 0$ and hedge funds do not participate in the long horizon market $q_{h,i} = 0$.*

Finally, for dealer banks, who are active in both the short and long horizon inflation markets, their budget constraints is:

$$a'_{b,i} = a_{b,i} + (\pi - p)q_{b,i} + (\Pi - P)Q_{b,i} + (d - s)e_{b,i} + y_{b,i} \tag{5}$$

so they choose both $q_{b,i}$ and $Q_{b,i}$. Following the large dealer net positions that we found in the data, especially in the long-horizon market, the dealers are solving a portfolio problem, as opposed to a market making problem alone.

As with the client institutions, dealer banks also have capacity constraints on their ability to take on risk. However, given the high-frequency nature of our transaction-level data, we will assume that within each dealer bank, the ability of traders on the short horizon desk to take on risk is not constrained by the risk taken by the other desk serving the long horizon market within a day, prior to books being balanced. In reality, the desks at a dealer that sell long-horizon and short-horizon inflation swaps can be separated, and are often manned by different traders. Books are only compared at the end of day, and while orders are made to have the position in one book to be potentially offset or constraining the position in the other book, this happens only a day after. Formally, this implies that dealer banks face two separate capacity constraints.

Assumption 2. *(Desk separation within the day.) Dealers face separate capacity constraints:*

$$G_b^S(Q_{b,i}, z_{b,i}) \geq 0 \quad \text{and} \quad G_b^L(q_{b,i}, z_{b,i}) \geq 0 \tag{6}$$

so that $\partial G_b^S(\cdot,\cdot)/\partial q_{b,i} = 0$ and $\partial G_b^L(\cdot,\cdot)/\partial Q_{b,i} = 0$.

3.2 Demand, supply and market clearing

Proposition 1, proven in appendix B.1, states the solution of the problem in the previous section in the form of an asset demand system of pension fund institutions.

Proposition 1. *Given market prices p^* and s^*, a pension fund's optimal demand for inflation protection $q^*_{f,i}$ scaled by size is given by:*

$$\frac{q^*_{f,i}}{a_{f,i}} = \underbrace{\frac{\mu_{f,i}\pi^e - p^*}{\gamma_{f,i}\sigma_\pi^2(1-\rho_{\pi,d}^2)}}_{\text{price and beliefs}} - \underbrace{\left(\frac{\sigma_d}{\sigma_\pi}\right)\left[\frac{\theta_d - s^*}{\gamma_{f,i}\sigma_d^2(1-\rho_{\pi,d}^2)}\right]\rho_{\pi,d}}_{\text{hedging demand}} - \underbrace{\left[\frac{1}{(1-\rho_{\pi,d}^2)\sigma_\pi^2}\right]\left(\frac{\sigma_{\pi,y_{f,i}}}{a_{f,i}} + \frac{\lambda_{f,i}g_{f,i}}{\gamma_{f,i}}\right)}_{\text{liquidity frictions}} \tag{7}$$

Demand for inflation swaps scales with the size of the institutions, and depends on three terms.

The first is a subjective expected Sharpe ratio: the difference between expected inflation and the price of the swap, scaled by risk aversion times overall uncertainty. If an institution expects inflation to be higher it will want to buy more inflation protection; if it is more uncertain about inflation or more risk averse, it will respond less to those expectations. The slope of the demand curve in the traditional quantity-price space will be higher the more risk averse the agent is.

The second factor is the hedging of market risk. The higher the correlation between expected inflation and the returns of the pension fund's portfolio (higher $\rho_{\pi,d}$), the less it will want to buy inflation protection, since now higher inflation also comes with higher returns in other investments. This hedging demand scales with the size of the fund's position in the market, which depends on its Sharpe ratio.

The third factor captures liquidity frictions, driven by two features of the model. The first is that a higher covariance of expected inflation with background income lowers the demand for protection against inflation because this income gives a natural hedge. The second is that a binding capacity constraint lowers demand relative to what the pension fund would like to obtain but is prevented by regulations or internal governance constraints.

On the other side of the market are dealer banks. Their asset demand system looks similar, and is stated in appendix B.2. Here, we present a slightly restricted version where their beliefs about how inflation at different horizons covaries with market returns follows a one-factor structure, so $\rho_{\pi,\Pi} = \rho_{\pi,d}\rho_{\Pi,d}$. Together with desk separation, this assumption leads to similar demand functions as those of the clients.

Proposition 2. *Given market prices P^*, p^* and s^*, the optimal allocation of risk in the two markets of a dealer bank is:*

$$\frac{q^*_{b,i}}{a_{b,i}} = \frac{\mu_{b,i}\pi^e - p^*}{\gamma_{b,i}\sigma^2_\pi(1-\rho^2_{\pi,d})} - \frac{\sigma_d}{\sigma_\pi}\left[\frac{\theta_d - s^*}{\gamma_{b,i}\sigma^2_d(1-\rho^2_{\pi,d})}\right]\rho_{\pi,d} - \left[\frac{1}{(1-\rho^2_{\pi,d})\sigma^2_\pi}\right]\left(\frac{\sigma_{\pi,y_{b,i}}}{a_{b,i}} + \frac{\lambda^L_{b,i}g^L_{b,i}}{\gamma_{b,i}}\right) \tag{8}$$

$$\frac{Q^*_{b,i}}{a_{b,i}} = \frac{\mu_{b,i}\Pi^e - P^*}{\gamma_{b,i}\sigma^2_\Pi(1-\rho^2_{\Pi,d})} - \frac{\sigma_d}{\sigma_\Pi}\left[\frac{\theta_d - s^*}{\gamma_{b,i}\sigma^2_d(1-\rho^2_{\Pi,d})}\right]\rho_{\Pi,d} - \left[\frac{1}{(1-\rho^2_{\Pi,d})\sigma^2_\Pi}\right]\left(\frac{\sigma_{\Pi,y_{b,i}}}{a_{b,i}} + \frac{\lambda^S_{b,i}g^S_{b,i}}{\gamma_{b,i}}\right) \tag{9}$$

Asset prices in equilibrium are pinned down by market clearing conditions:

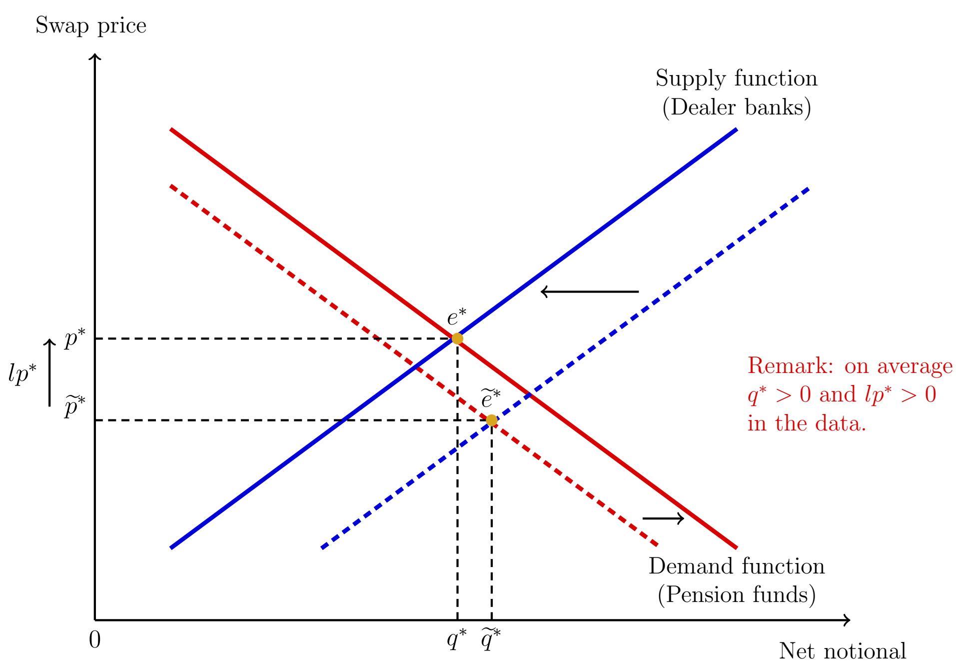
$$q^* \equiv \sum_{i\in\Theta_f} q^*_{f,i} = -\sum_{i\in\Theta_b} q^*_{b,i} > 0 \tag{10}$$

so that $q > 0$ means that pension funds have net positive notional holdings, and the negative of the sum of demand from dealer banks is the supply in the market.

3.3 The markets and the frictionless equilibrium

Figure 9 displays equilibrium in the long-horizon market with a simple graph that has the supply and demand curve as solid lines, and an equilibrium at the point e^*. In the data, we observed that $q^*_{b,i} < 0 < q^*_{f,i}$. The model can explain this as a result of four forces. First, pension funds may be more risk averse than banks, $\gamma_{f,i} > \gamma_{b,i}$. This is plausible, as they may be less well diversified and are forced by regulation to be more prudent. Second, pension funds have more hedging needs from their other assets than dealer banks, which makes sense given pension funds' large holdings of nominal bonds and the limited supply of index-linked bonds.[14] Third, pension funds are more exposed to background risk that covaries with inflation $\sigma_{\pi,y_{f,i}} > \sigma_{\pi,y_{b,i}}$, which again makes sense as many funds have liabilities denominated in real terms. Fourth, the regulatory environment for pension funds encourages them to buy inflation protection easing trading constraints: $\lambda_{f,i} < \lambda^L_{b,i}$, this is consistent with real world pension fund practices such as liability driven investments. Overall, therefore, the model can make sense of what we see in the data.

Figure 9 THE FRICTIONLESS EQUILIBRIUM



The *frictionless market equilibrium* arises when there are complete markets to fully insure institution-specific income risk, so $\sigma_{\pi,y_{b,i}} = \sigma_{\pi,y_{f,i}} = \sigma_{\pi,y_{h,i}} = 0$, and the regulatory, short-sale or other capacity constraints do not bind for any agent in either the long horizon market, $\lambda^L_{b,i} = \lambda_{f,i} = 0$, or the short horizon market, $\lambda^S_{b,i} = \lambda_{h,i} = 0$. Appendix B.3 solves for this counterfactual.

[14] For more details see, e.g., `https://prod.schroders.com/en/sysglobalassets/schroders/sites/ukpensions/pdfs/2016-06-pension-schemes-and-index-linked-gilts.pdf`

Lemma 1. *If $\widetilde{p}$ is the frictionless price of a long horizon inflation swap, in equilibrium, it is:*

$$\widetilde{p}^* = \underbrace{\left[\frac{\sum_{i\in\Theta_f}\widetilde{\gamma}_{f,i}^{-1}\mu_{f,i}}{\sum_{i\in\Theta_f}\widetilde{\gamma}_{f,i}^{-1}+\sum_{i\in\Theta_b}\widetilde{\gamma}_{b,i}^{-1}} + \frac{\sum_{i\in\Theta_b}\widetilde{\gamma}_{b,i}^{-1}\mu_{b,i}}{\sum_{i\in\Theta_f}\widetilde{\gamma}_{f,i}^{-1}+\sum_{i\in\Theta_b}\widetilde{\gamma}_{b,i}^{-1}}\right]}_{\text{size-weighted dispersion of beliefs}} \underbrace{\pi^e}_{\text{expected inflation}} - \underbrace{\frac{\theta_d - \widetilde{s}^*}{\sigma_d^2}\sigma_{\pi,d}}_{\text{risk premium}} \tag{11}$$

The frictionless price for swaps depends on expected inflation minus a compensation for risk. Importantly, the coefficient on π^e is a weighted function of each institution's beliefs about expected inflation that depend in turn on the size of their gross notional risk traded. Therefore, the first term on the right hand side does not equal π^e necessarily, as long as there is granularity in asset holdings with some institutions being much larger, so that their beliefs carry much more weight in pricing of the asset. We use ε_π to denote the fundamental shocks that drive the frictionless price.

Given this frictionless price, the liquidity premium lp^* is then the difference between it and the actual price:

Lemma 2. *The liquidity premium in the long horizon market is defined as $lp^* = p^* - \widetilde{p}^*$, and is driven by frictions in both the demand from pension funds and the supply from dealer banks:*

$$lp^* = \underbrace{-\frac{\sum_{i\in\Theta_b}\left\{\sigma_{\pi,y_{b,i}} + \frac{\lambda_{b,i}^L g_{b,i}^L}{\widetilde{\gamma}_{b,i}}\right\}}{\sum_{i\in\Theta_f}\widetilde{\gamma}_{f,i}^{-1}+\sum_{i\in\Theta_b}\widetilde{\gamma}_{b,i}^{-1}}}_{\equiv\varepsilon_b,\ \text{the supply friction from dealer banks}} + \underbrace{-\frac{\sum_{i\in\Theta_f}\left\{\sigma_{\pi,y_{f,i}} + \frac{\lambda_{f,i} g_{f,i}}{\widetilde{\gamma}_{f,i}}\right\}}{\sum_{i\in\Theta_f}\widetilde{\gamma}_{f,i}^{-1}+\sum_{i\in\Theta_b}\widetilde{\gamma}_{b,i}^{-1}}}_{\equiv\varepsilon_f,\ \text{the demand friction from pension funds}} \tag{12}$$

Figure 9 represents the frictionless supply and demand as dashed lines, and the equilibrium as $\widetilde{e}^*$. The literature has emphasized the capacity constraint facing dealers, in which case $\lambda_{b,i}^L g_{b,i}^L$ would be the dominant term, and the liquidity premium is positive.[15] This is the case depicted in the figure, where the actual supply is further to the left of the frictionless one than is the case with demand ($\varepsilon_b > \varepsilon_f$). However, shocks that are specific to the pension fund sector may sometimes make its capacity constraint bind sharply, $\lambda_{f,i} g_{f,i}$ jumps, and the liquidity premium can easily turn negative (ε_f dominates). Moreover, changes in the macroeconomy, in monetary policy, and in the policy regime driving inflation, which have been plentiful during our sample period, will change the perceived correlation between inflation and the income flows of pension funds and banks, so that $\sigma_{\pi,y_{f,i}}$ and $\sigma_{\pi,y_{b,i}}$ can be large and volatile, leading to large fluctuations in the liquidity premium. The observed prices of inflation swaps p^* can therefore be very far from actual risk-adjusted expected inflation $\widetilde{p}^*$, and move significantly over time, driven by market frictions and institutional-level constraints that shift both demand and supply.

Finally, the same ingredients drive the short-horizon inflation market where hedge funds and

[15]This corresponds to a *negative* supply shock.

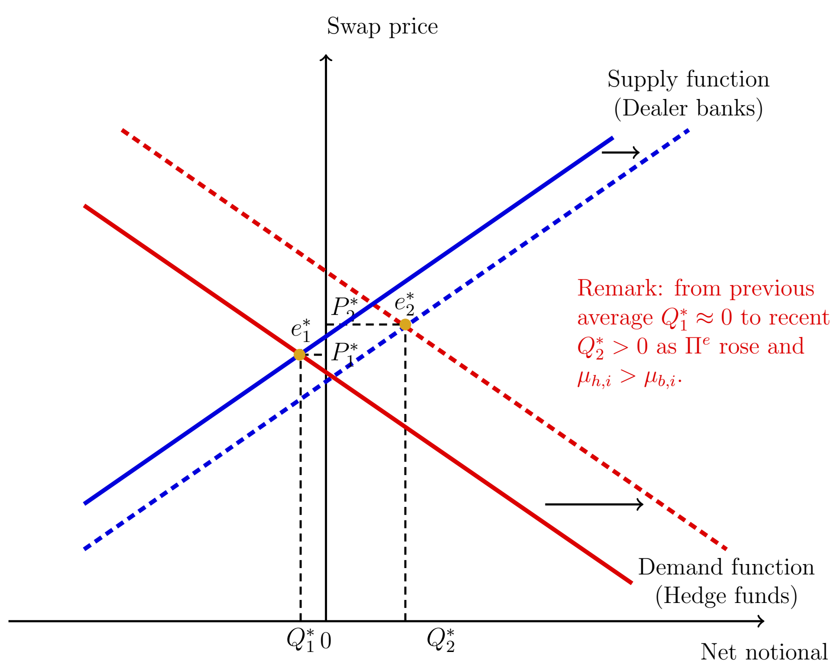
dealer banks meet and markets clear according to the condition:

$$Q^* = \sum_{i \in \Theta_h} Q^*_{h,i} = -\sum_{i \in \Theta_b} Q^*_{b,i} \tag{13}$$

The equilibrium is represented in Figure 10, starting from an initial situation with solid lines and subscript 1, to what happens after a shock to fundamental expected inflation shown with dashed lines and subscript 2.

This equilibrium differs from the one in the long horizon market in two ways suggested by the model that fit the facts that we documented. First, banks and hedge funds have on average similar beliefs regarding hedging demand, background risk, regulation. Therefore, the historical equilibrium net holdings Q^*_1 were on average close to zero, as in the data. Second, we can expect this market to have large swings, in quantities and prices, because changes in expected inflation π^e come with large changes in the dispersion of beliefs. This is because hedge funds, being more informed traders, would be expected to be more sensitive to inflation news than banks: $\mu_{h,i} > \mu_{b,i}$, matching the recent shift to a positive Q^*_2. At the same time, the pandemic and political uncertainty have shifted the liquidity frictions for both sides of the market and so shifted demand and supply as well. Therefore, observing the new price P^*_2, it is especially difficult to measure how much has expected inflation actually increases.

Figure 10 EQUILIBRIUM IN THE SHORT HORIZON INFLATION SWAP MARKET



4 Identification strategies

The identification problem is that the observed prices of inflation swaps p and P can move because of shocks to fundamentals or shocks to the liquidity frictions. As the model clarified, there are three separate sources of frictions, from each of the three sectors trading in the market. There are four shocks in column vector of shocks is: $\boldsymbol{\varepsilon} = (\varepsilon_h, \varepsilon_f, \varepsilon_b, \varepsilon_\pi)'$ and usually only two observables.

We, instead, have daily data on prices *and* quantities from the 2^{nd} January 2019 to the 10^{th} February 2023, for 879 observations. Therefore, unlike previous studies, we have four variables corresponding to the following aggregated data series: (i) q: net purchases of UK RPI inflation swaps by PFLDIs with initial time-to-maturity 10 years or more; (ii) p: a weighted-average daily price of UK RPI zero coupon inflation swaps of initial time-to-maturity 10 years or more, where the weights are gross notionals traded in each long maturity category by PFLDI institutions as a share of the total across the data sample; (iii) Q: the net purchases of swaps by hedge funds with initial time-to-maturity 3 years or less; and (iv) P, the weighted-average daily price of UK RPI zero coupon inflation swaps with weights equal to the share of gross notional amount traded in each maturity category by hedge fund institutions in this market. Let $\mathbf{Y} = (Q, P, q, p)$ be the column vector with these variables.

Having as many variables as shocks is progress, but still not enough for identification. The fundamental shocks shift all demand and supply curves, while the liquidity shocks shift the demand for either pension funds, hedge funds, or dealers, separately depending on their source. Formally:

$$\mathbf{Y} = \boldsymbol{\Psi}\boldsymbol{\varepsilon} \tag{14}$$

and we need to pin down the elements of the 4×4 matrix $\boldsymbol{\Psi}$ to fully identify the system, or at least four elements in its inverse to extract from the data $\mathbf{Y}$ the fundamental expected inflation ε_π. In this section we describe how we use the high-frequency, the cross-section, and the time length of our data to achieve identification through hree separate strategies that exploit each of these three features.

4.1 First identification strategy: heterogeneity in reactivity

It seems plausible to believe that in response to fundamental news, dealers are more informed (or attentive, or reactive) than pension funds. After all, the dealers trade the inflation risk and see all sides of the market. As such, they have more precise posterior information about inflation and they respond more to news. In this case, following a shock to the fundamental that raises π^e (and so frictionless $\widetilde{p}$), the upward shift in the supply function dominates the shift in the demand function. Therefore, we expect the new equilibrium to have a higher price and a lower net notional amount traded. That is, p should rise and q should fall.

Instead, in the short horizon market, the informed hedge funds are likely more reactive to fundamentals than dealer banks, as we discussed in Figure 10. Therefore, now it is the upward shift in the demand function that dominates leading to a higher P and a rise in Q.

This identifies a fundamental shock as the one that satisfies these sign restrictions on quantities following a rise in prices (and symmetrically for a fall). Formally, the identification assumption is:

Assumption 3. *(Differential reactiveness to fundamental news about inflation.) Dealer banks respond more to fundamental long-horizon expected inflation than pension funds but less to fundamental short-horizon expected inflation than hedge funds:*

$$\frac{\sum_{i\in\Theta_h}\widetilde{\gamma}_{h,i}^{-1}\mu_{h,i}}{\sum_{i\in\Theta_h}\widetilde{\gamma}_{h,i}^{-1}+\sum_{i\in\Theta_b}\widetilde{\gamma}_{b,i}^{-1}} > \frac{\sum_{i\in\Theta_b}\widetilde{\gamma}_{b,i}^{-1}\mu_{b,i}}{\sum_{i\in\Theta_f}\widetilde{\gamma}_{f,i}^{-1}+\sum_{i\in\Theta_b}\widetilde{\gamma}_{b,i}^{-1}} > \frac{\sum_{i\in\Theta_f}\widetilde{\gamma}_{f,i}^{-1}\mu_{f,i}}{\sum_{i\in\Theta_f}\widetilde{\gamma}_{f,i}^{-1}\sum_{i\in\Theta_b}\widetilde{\gamma}_{b,i}^{-1}} \tag{15}$$

Next, suppose that there is a shock to the liquidity component of demand from pension funds that leads to a rightward shift of the demand function. In equilibrium, this would raise p and q simultaneously in the long horizon market. Therefore, for the same increase in p, we can use the equilibrium responses in quantities q to distinguish between a liquidity demand shock and a fundamental shock. If instead the shock were to the liquidity component of dealers' supply of inflation protection, this would shift the supply curve downward, and so lead to a higher p and lower q, just as with a fundamental shock. To separately identify the two, we bring in the price and quantity information from the short-horizon market. Since the shock to the dealers would *also* appear in the short-horizon market, we should observe P rising and Q falling.[16] Observing both markets, we can therefore distinguish between the two shocks again using information on Q. In this case, the identification comes from assumption 1 on segmented markets.

There is a final barrier to overcome for identification. If the change in the supply of inflation swaps in the long market following a fundamental shock causes the capacity constraints in the short market to bind, this would lead to a reduction in supply in the short horizon market. Therefore, P would rise and Q fall, making the supply shocks indistinguishable from a fundamental shock. This is ruled out by the desk-separation assumption 2: within any given trading day, this spillover from quantity supplied in one market into constraints in the other market does not happen.

Note that all three assumptions, and especially desk separation, exploit the high frequency in the data. We should definitely not expect that desks are separated within a month, or even a week. Also, whatever information differences between banks, hedge funds, and pensions funds may well

[16] Formally, recall that the shock $z_{b,i}$ to dealer bank i's ability to take on inflation risk affects the capacity constraints in both markets: $G_b^S(Q_{b,i}, z_{b,i})$ and $G_b^S(Q_{b,i}, z_{b,i})$.

be gone within a few days; in fact, we will find it is so. Combining the discussion of demand and supply shifters in the model, the three assumptions imply the following sign restrictions on the response of observables to shocks within a day:

$$\mathbf{A} = \begin{pmatrix} + & 0 & - & + \\ + & 0 & + & + \\ 0 & + & - & - \\ 0 & + & + & + \end{pmatrix} \tag{16}$$

These sign restrictions set identify the four shocks, and the response of each variable to them.

4.2 Second identification strategy: cross-sectional granularity

An alternative way to go about identifying the fundamental shocks uses the buying and selling behavior of specific market participants. If the previous strategy exploited the high frequency of the data, this strategy exploits its cross-sectional richness. Assumption 3 on differential reactiveness within a day is replaced by a new assumption on cross-sectional granularity.

Before we state this assumption, we first show that the asset demand system in equation (7) derived under Proposition 1 can be written as an interactive fixed effects factor model. We introduce the following two definitions on $\mu_{f,i,t}$ and $\varepsilon_{f,i,t}$, which are the time-varying subjective beliefs of institution i and its demand shock arising from liquidity frictions respectively:

$$\mu_{f,i,t} = \kappa^{\mu}_{f,i}\mu_{f,t} \tag{17}$$

$$\varepsilon_{f,i,t} = -\frac{1}{(1-\rho^2_{\pi,d})\sigma^2_{\pi}}\left(\frac{\sigma_{\pi,y_{f,i}}}{a_{f,i}} + \frac{\lambda_{f,i}g_{f,i}}{\gamma_{f,i}}\right) = \frac{\kappa^{lp}_{f,i}lp^*_t}{\gamma_{f,i}\sigma^2_{\pi}(1-\rho^2_{\pi,d})} + \tilde{\varepsilon}_{f,i,t} \tag{18}$$

In equation (17), $\kappa^{\mu}_{f,i}$ measures the disagreement of expected inflation by pension fund i relative to its entire sector. In equation (18), we define its demand shock to be the sum of an *idiosyncratic* disturbance $\tilde{\varepsilon}_{f,i,t}$ and a fund-specific component contributing to the market-wide liquidity premium lp^*_t defined in equation (12). This component is captured by a coefficient $\kappa^{lp}_{f,i}$. Given these definitions, the demand system in equation (7) appended with time subscripts can be rewritten as:

$$\frac{q_{f,i,t}}{a_{f,i,t}} = \boldsymbol{\omega}'_{f,i}\mathbf{F}_t + \widetilde{\varepsilon}_{f,i,t} \tag{19}$$

where $\mathbf{F}_t$ denotes the unobserved common factors and $\boldsymbol{\omega}_{f,i}$ are the fund-specific factor loadings.

These are given by:

$$\mathbf{F}_t = \begin{pmatrix} \pi_t^e \\ lp_t^* \end{pmatrix}, \quad \boldsymbol{\omega}_{f,i} = \begin{pmatrix} \frac{(\kappa_{f,i}^{\mu}\mu_{f,t}-\Lambda_t)}{\gamma_{f,i}\sigma_\pi^2(1-\rho_{\pi,d}^2)} \\ \frac{\kappa_{f,i}^{lp}-1}{\gamma_{f,i}\sigma_\pi^2(1-\rho_{\pi,d}^2)} \end{pmatrix} \tag{20}$$

Appendix C provides a extended derivation of this factor structure, showing how these reduced-form variables and coefficients map to the model objects.

In the implementation, we estimate this regression with an interactive fixed effects model following Bai (2009) to extract the idiosyncratic component of *each* pension fund's demand. We subsequently construct an instrument for sector-wide shocks to pension fund demand $\varepsilon_{f,t}$ defined in equation (12), exploiting the granularity of their swap positions in this market. This granular instrument is constructed as a weighted sum of the residuals:

$$GIV_{f,t} = \sum_{i \in \Theta_f} a_{f,i,t}\tilde{\varepsilon}_{f,i,t} \tag{21}$$

Given that $\mathbf{F}_t$ spans all the determinants of demand by the pensions funds, we know that the residuals are independent of fundamentals $\mathbb{E}(GIV_{f,t}\varepsilon_{\pi,t}) = 0$ and of the determinants of the supply by banks: $\mathbb{E}(GIV_{f,t}\varepsilon_{b,t}) = 0$. Further, because of the segmentation of markets (assumption 1), we know that $\mathbb{E}(GIV_{f,t}\varepsilon_{h,t}) = 0$. Therefore, the exclusion restriction for $GIV_{f,t}$ to be an instrument for $\varepsilon_{f,t}$ is satisfied.

For the instrument to be relevant though requires $GIV_{f,t}$ to be correlated with $\varepsilon_{f,t}$. Yet, from the properties of estimated regression residuals, we know that: $\sum_{i\in\Theta_f} \tilde{\varepsilon}_{f,i,t} = 0$. If the data is granular in the sense that the the average of these residuals is no longer zero once they are weighted by the size of each firm, then the relevance condition can be satisfied. This is because the shocks to some of the large firms can drive the average. The same applies to $GIV_{h,t}$ and $GIV_{b,t}$ for the liquidity shocks to hedge funds and banks, respectively. We then have three valid instruments, which by the properties of a system that has four variables and four shocks, implies that the shocks to fundamentals are point identified.

The intuition behind this approach is that because some institutions are larger, individual shocks to their demand function will drive the aggregate demand in the market. Since we have many such institutions, we can measure the idiosyncratic changes in their demand. This provides an aggregate demand shock. This is possible because we have institution-specific transaction data on $q_{f,i,t}$ and $a_{f,i,t}$ and so we can identify these large institutions.[17]

More precisely, the assumption of granularity is:

Assumption 4. *(Granularity of the institutions) The data on asset positions* $a_{f,i,t}$, $a_{h,i,t}$ *and* $a_{b,i,t}$

[17] A more technical discussion of the identification strategy and details of its implementation is in Appendix C.

is granular in that:

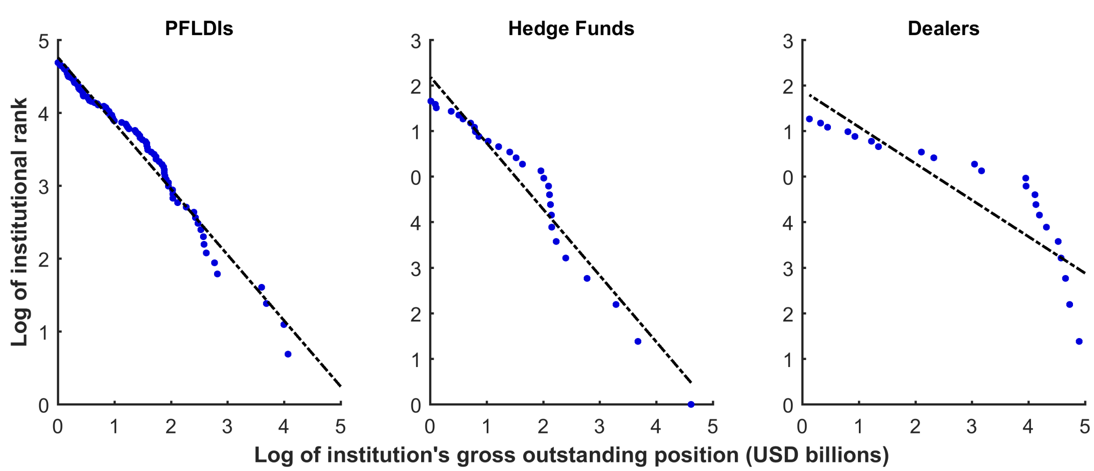
$$\mathbb{E}(GIV_{f,t}\varepsilon_{f,t}) \neq 0 \ \ and \ \ \mathbb{E}(GIV_{b,t}\varepsilon_{b,t}) \neq 0 \ \ and \ \ \mathbb{E}(GIV_{h,t}\varepsilon_{h,t}) \neq 0 \tag{22}$$

We can verify the plausibility of this assumption in the data. In our sample, there are 210 PFLDIs. Figure 11 shows the plot of the rank of each PFLDI institution against their outstanding gross notional positions. The Pareto parameter on their outstanding gross notional position is 0.13, strongly supporting the granularity assumption. For institutions within the PFLDI sector with an outstanding gross notional position to-date larger than $1bn, we estimated a power law regression of (the log of) their rank on (the log of) their gross notional positions (Gabaix, 2016). The fit of the regression is also in the figure. The estimated power law coefficient is -0.9, with a standard error of 0.013 and an R^2 of 0.979. Therefore, the size of PFLDI's gross inflation risk exposures comes close to satisfying Zipf's law, which is a particular power law distribution with a power law coefficient of -1.

There are fewer hedge funds (30) and banks (18), so results are more imprecise. Still, the power law exponent points estimates for hedge funds and dealer banks are -0.728 and -0.402 with standard errors 0.035 and 0.058, respectively. Although these estimates are far from satisfying Zipf's law, they still support granularity.

An alternative approach to check the relevance of the instruments is through standard first-stage F-statistics. They are: 18.1 for $GIV_{f,t}$, 65.8 for $GIV_{h,t}$ and 38.2 for $GIV_{b,t}$.

Figure 11 Institutional rank versus outstanding gross notional positions



Note: To arrive at each institution's gross outstanding position in the inflation swap market, we first construct an unbalanced panel by tracking the trading activity of each institution across various execution dates of the trade, and then cumulatively construct a stock of their outstanding positions while taking into consideration older trades that expire. Source: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

4.3 Third identification strategy: heteroskedasticity across time

The third identification strategy exploits instead the fact that we have a time series of 879 days. The new assumption, to complement market segmentation and desk separation, is that fundamental news about inflation are lumpy. The underlying shocks may themselves be smooth over time, but it is around data release dates that traders learn the most about where inflation is heading.

In the data, at these dates, the volatility of inflation swap prices is noticeably higher. While liquidity shocks may also be higher during those dates, the major difference should be that actual fundamentals are revealed on these dates. The identifying assumption is that this spike in variance is driven by the emergence of fundamental shocks on these dates, which again seems reasonable. This allows for partial identification of the fundamental shocks alone. More formally:

Assumption 5. *(Heteroskedascity at known dates due to fundamentals.) If $\mathbf{\Sigma}_h$ is the variance-covariance matrix of the shocks $\boldsymbol{\varepsilon}$ at data release dates, and $\mathbf{\Sigma}_l$ the one at other dates, then the largest diagonal element of $\mathbf{\Sigma}_h'\mathbf{\Sigma}_l$ is the one associated with the variance of the fundamentals ε_π.*

In our sample, we have 48 monthly dates when the data on UK RPI inflation was released. As a first check, we indeed observe spikes in trading activity along with large price adjustments in the vicinity of these dates. In addition to these 48 dates, one more date in our sample had a large impact: 6^{th} September 2022 when ex-Prime Minister Truss announced a cap on energy prices that dramatically changed the properties of measured inflation.

To check the plausibility of assumption 5, we estimate a vector autoregression on those 49 dates, as well as a second VAR on the remaining dates. Calculating the product of the inverse of the variance-covariance matrix of residuals on the latter dates and the variance-covariance matrix of residuals on the former dates, the resulting matrix has a largest eigenvalue of 7.618. That is, there is clearly more variation in both prices and quantities during RPI release dates.

Finally, following Lütkepohl, Meitz, Netšunajev, and Saikkonen (2021), we implement a Wald-type test under the null hypothesis of homoskedasticity and so no identification. The null is rejected at a 0.1% significance level.

4.4 Dynamics and implementation

So far, we have discussed the model and identification in a purely static setting. If markets were efficient after one day, then this would suffice. However, there are good reasons to expect information to diffuse more slowly over several days, and that capacity constraints and regulations involve multiple successive days. The empirical model must be dynamic to allow for these features.

We estimate a Bayesian VAR with diffuse priors, a deterministic constant $\mathbf{c}$, and $L = 3$ lags:

$$\mathbf{Y}_t = \mathbf{c} + \sum_{\ell=1}^{L} \mathbf{\Phi}_\ell \mathbf{Y}_{t-\ell} + \mathbf{u}_t \quad \text{and} \quad \mathbf{u}_t = \mathbf{\Psi} \boldsymbol{\varepsilon}_t \tag{23}$$

Then, we implement the timing identification following Arias, Rubio-Ramírez, and Waggoner (2018), where sign restrictions on $\mathbf{\Psi}$ give set identification. For the granularity identification, we follow Stock and Watson (2018) and use the GIV as proxy instrumental variables. For the heteroskedasticity identification, we follow Brunnermeier, Palia, Sastry, and Sims (2021) and pick the shock that most closely satisfies the properties of a fundamental shock in our model.

To be able to compare the estimation results across the three identification strategies, we scale the impulse response functions in the following manner. The increase in the net notional position aggregated across all client institutions in the short horizon inflation market is always scaled to equal \$1bn in response to either the fundamental shock, the demand shock from hedge funds, or the supply shock from dealers. As for the demand shock from PFLDIs in the long horizon market, it is scaled such that it raises the net notional position aggregated across all client institutions in the long horizon inflation market by \$1bn.

4.5 Cross-verification of identification assumptions

The promise of having three identification strategies is not only that they can challenge the robustness of the findings, but also that they can internally cross-verify the identification assumptions that each strategy has independently imposed. Given any one of the identification strategies, we can let the data speak on whether the identification assumptions made by the other strategies seem to hold.

Starting with assessing the identification based on assuming differential reactiveness to fundamental shocks, the impulse response functions to a fundamental shock estimated under the other two identification strategies both have a rise in the net notional traded in the short-horizon market and a reduction in this quantity in the long-horizon market. Prices rise on impact in both markets, and the confidence bands are such that we do not reject the sign restrictions imposed by differential reactiveness on the estimates from the other two identification strategies.[18]

Next, turning to the identification based on the heteroskedasticity assumption, we find evidence that the fundamental shocks identified by sign restriction are more volatile on the RPI release dates. Across the 7560 importance sampler draws that give set identification, the ratio of the variance of the fundamental shock on these dates to that obtained from other dates remaining in our data sample has a median of 1.267, and is above one 91% of the times. For several draws, the variance

[18]These impulse responses are reported in Figure 12.

of the fundamental shock can be five times as large when compared to the non-inflation release and news dates (see Appendix E.6).

Third, we investigate the granularity assumption using the fundamental shocks $\hat{\varepsilon}_{\pi,t}$ estimated from the other two approaches. We compute the sample analogs of the exclusion restriction conditions by checking whether:

$$\frac{1}{T}\sum_{t=1}^{T} GIV_{\nu,t}\hat{\varepsilon}_{\pi,t} = 0, \quad \text{for} \quad \nu \in \{f, h, b\} \tag{24}$$

Using the draws of the estimated fundamental shocks from set identification, the median estimates are -0.006, 0.030, and 0.032 for $\nu = f$, h, and b, respectively.[19] Using instead the point estimates of the fundamental shock identified via heteroskedasticity, thesy are -0.071, -0.156 and -0.024, respectively. All are close to zero, supporting the validity of the granular instrumental variables.

Finally, we conclude this section by reporting the absolute value (since shocks are identified up to a sign) of the correlation of the estimated fundamental shocks obtained from each of the three strategies. Ordering the fundamentals from sign restriction strategy first, GIV strategy next, and heteroskedasticity strategy last, these correlations are:

$$\begin{bmatrix} 1 & 0.979 & 0.842 \\ \cdot & 1 & 0.783 \\ \cdot & \cdot & 1 \end{bmatrix} \tag{25}$$

Such a remarkably high degree of correlation across three estimates that come from completely different identification strategies, using completely different properties of the data, give confidence that the estimates that follow are robust.

5 Estimates and applications

This section presents our main estimates and findings from the UK RPI inflation swap market.

5.1 Impulse responses: the speed of adjustment to fundamentals

Figure 12 shows the impulse response to a fundamental shock for the three identification strategies. They are all qualitatively similar. Under our timing identification strategy, recall that the sign restricted responses of prices in both markets should increase, while those of the quantities traded in the long and short horizon markets should be negative and positive respectively. This is what

[19] Figure E.9 in Appendix E.6 shows the distribution of the sample analogs computed across all draws.

the differential reactiveness assumption implies, and as we noted earlier, the signs of the responses are verified by the other identification restrictions.

There are significant differences in the size of the price responses across strategies. They are much larger when identified with the granular strategy, and much smaller with the heteroskedasticity strategy, with the estimates from the timing strategy in between. However, the error bands are large and intersect across the three sets of price impulse responses, at both the short and long horizons. Because this pattern of very similar qualitative responses is true across other experiments, from now onwards we report only results with the timing identification strategy to conserve space. All the conclusions that we draw are robust to the identification strategy.

It is striking that this market completely adjusts within a handful of days. All responses stabilize within two to three days and then seemingly persist over time. This suggests that fundamental shocks are permanent and that information diffuses quite quickly. Given the high frequency of the data, perhaps the persistence is not too surprising: after news about inflation arrives, often in a lumpy way at a data release date, agents will adjust persistently their expectations of inflation until at least the next data release the following month. The quick adjustment is more surprising as it suggests that this market is close to being (weakly) informationally efficient.

5.2 Impulse responses: liquidity shocks and the slopes of the demand and supply functions

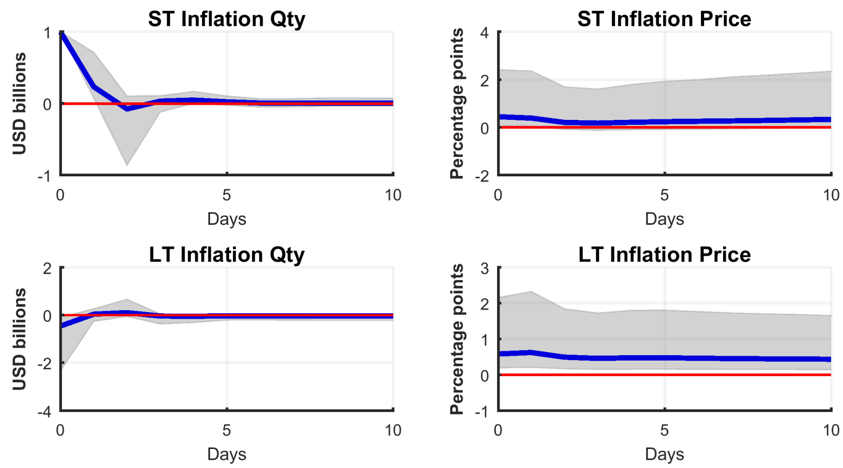
Figure 13 shows the estimated dynamic responses to the demand and supply liquidity shocks. These conform with the standard responses one would expect from shocks to supply and demand. Namely, swap prices rise and the quantity traded falls in both markets following a supply shock originating from dealers that would shift the supply function upwards, while both rise in response to a demand shock.[20]

Given the estimated impulse response functions, we use the relative responses to compute the slope of demand and supply functions in both markets. The intuition is as follows. For the case in which there is a shock to dealers that amounts to a reduction in supply, this would trace out the demand functions in both the long and short horizon markets *on impact*. In the long horizon market, for instance, we can subsequently compute the slope of the demand function locally by calculating dp/dq, where dp and dq are the impact responses of the swap breakeven rate and aggregate quantity traded. That is, the slope of the pension funds' demand function is locally identified here by a shift in dealer's supply function upon the realization of the shock. These provide insights into how sensitive prices and quantities will be to future shocks, as well as useful

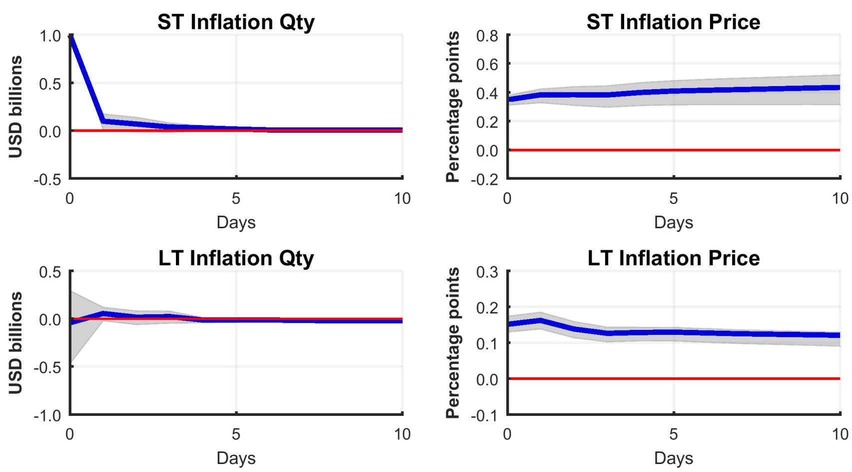
[20]Note that, on impact, we have left unrestricted the responses of price and quantity from the other market in the timing restriction strategy. So, this is a result, not an assumption.

Figure 12 Estimated impulse response functions to a fundamental shock

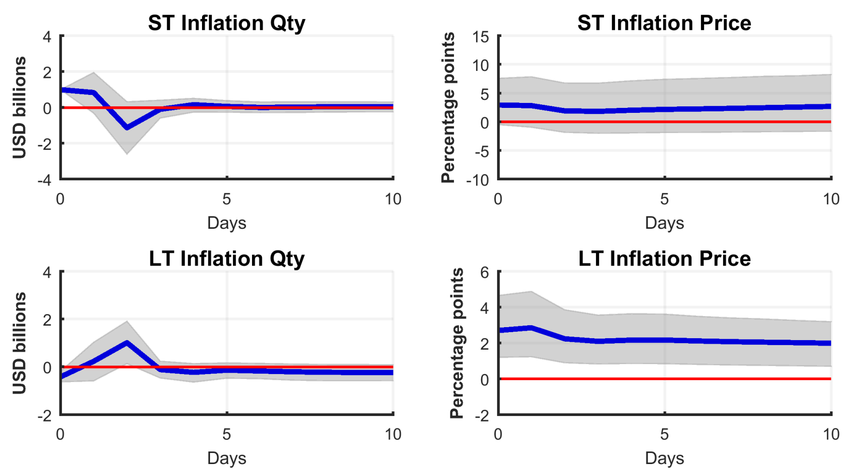
(i) With timing restriction strategy



(ii) With heteroskedasticity strategy



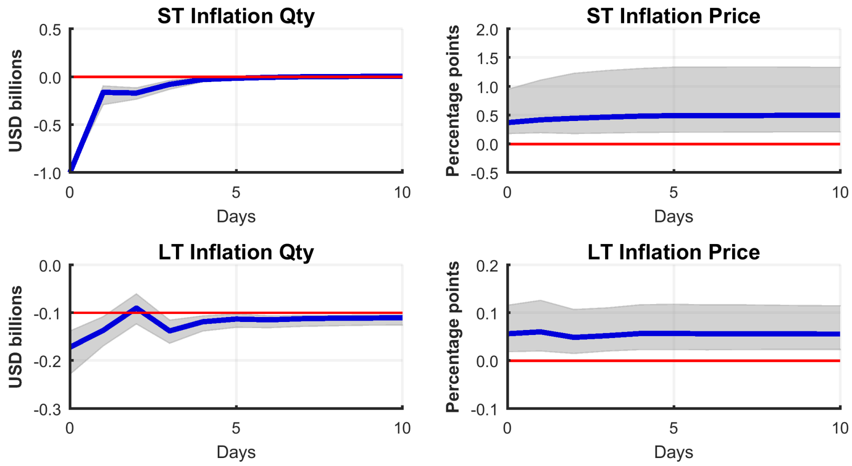
(iii) With granular identification strategy



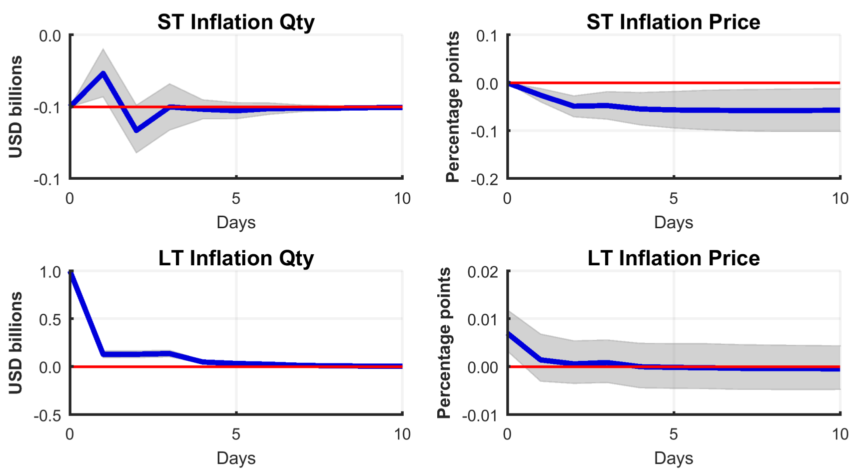
Note: Figure shows the median estimate with 68% confidence intervals.

Figure 13 Estimated impulse response functions to liquidity shocks

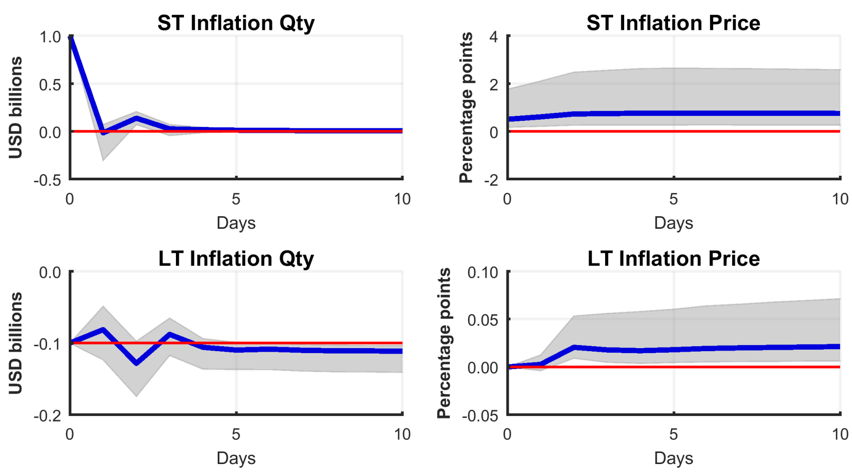
(i) Dealer Supply Shock ($\varepsilon_{b,t}$)



(ii) PFLDI Demand Shock ($\varepsilon_{f,t}$)



(iii) Hedge Fund Demand Shock ($\varepsilon_{h,t}$)

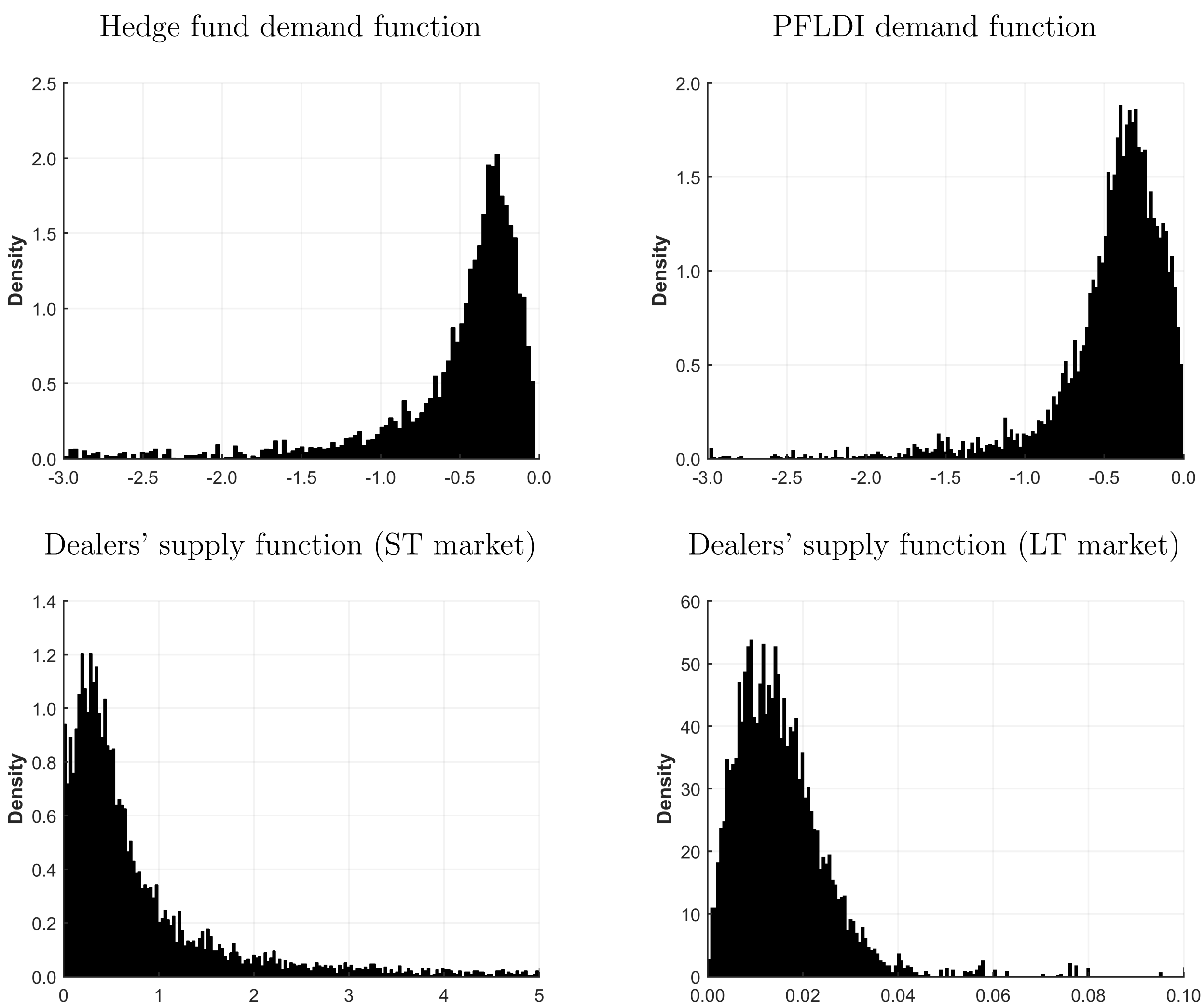


Note: Figure shows the median estimate with 68% confidence intervals.

statistics in writing more sophisticated models for these markets.

Figure 14 shows the estimated posterior distributions of the locally-estimated slopes of the demand and supply curves in both markets from the Bayesian SVAR. In the short-horizon market, dealer banks and hedge funds are similarly price-sensitive. Perhaps this is a reflection of sellers and buyers in this market being similarly sophisticated and have similar bargaining power when negotiating in this over the counter market. Instead, in the long horizon market, the slope of the supply function is close to zero. This almost-horizontal supply curve is close to a situation where dealers effectively set prices in the long-horizon market, with full bargaining power relative to their pension fund clients.

Figure 14 Distribution of estimated slopes of market demand and supply

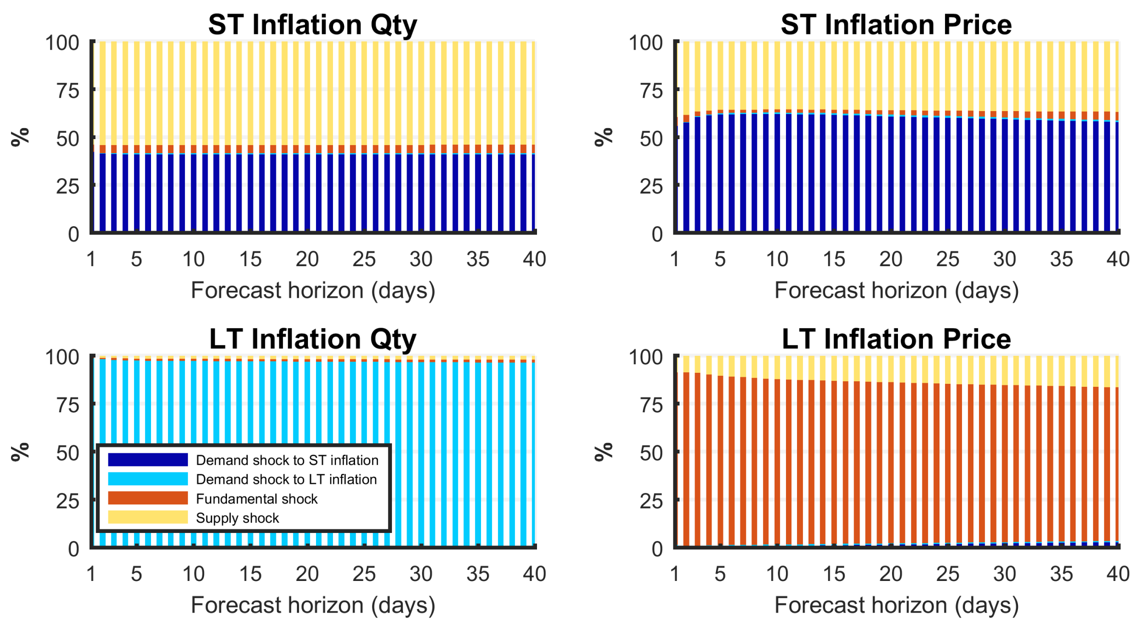


Note: Plotted slopes shown are truncated at the given axis values above.

5.3 Variance decompositions: the drivers of inflation swap prices at different horizons

Figure 15 shows the forecast error variance decompositions for the different shocks. We plot these at different forecasts horizons but, confirming the previous finding of quick response to shocks that have persistent effects, the decomposition is almost the same at all horizons.

Figure 15 FORECAST ERROR VARIANCE DECOMPOSITION



In the short-horizon, liquidity shocks drive nearly *all* the variation in inflation swap rates. Interestingly, while the focus in the literature on liquidity premia tends to be on the market makers, we find that supply shocks from dealer banks only account for about one third of the variance of prices. The remaining two thirds come instead from liquidity shocks affecting the demand from pension funds. When it comes to quantities, this ordering reverses, with liquidity shocks to dealer banks accounting for about 60% of the variance, while liquidity shocks to pensions funds account for the remaining 40%. Since regulatory constraints are probably not the main driver of hedge fund behavior, these results suggest putting more work into how background income risk (in the case of hedge funds, the flow of funds in and out) correlates with inflation, or what other capacity constraints hedge funds face in their ability to take inflation risk (perhaps linked to compensation and governance).

At long horizons, there is a stark contrast between quantities and prices. The former are almost entirely driven by liquidity frictions affecting pension funds. Changes in regulations, in constraints to taking risk, or in contributions and payouts into the funds come with large shifts in the demand curve form pension funds. Consistent with the flat supply curve from dealers that we found in the previous sub-section, this has a large effect on quantities, but very little effect on prices.

In such a market, prices can only move if the supply curve shifts, which could happen either because of shocks to fundamentals or liquidity shocks to dealers. This is precisely what the variance decomposition reveals. Fundamentals, that shift both supply and demand, account for roughly four fifths of the variance, with the remainder due to liquidity frictions affecting dealer banks that shift the supply curve.

Focusing on fundamentals, these variance decompositions suggest that changes in the 10-year inflation swap rate are a good indicator of fundamental expected inflation, but swap rates from the short horizon markets are not. For macroeconomic purposes, most movements in the inflation swap rates up to three years can be dismissed.

5.4 Historical decomposition: squeezing signal out of short-term prices

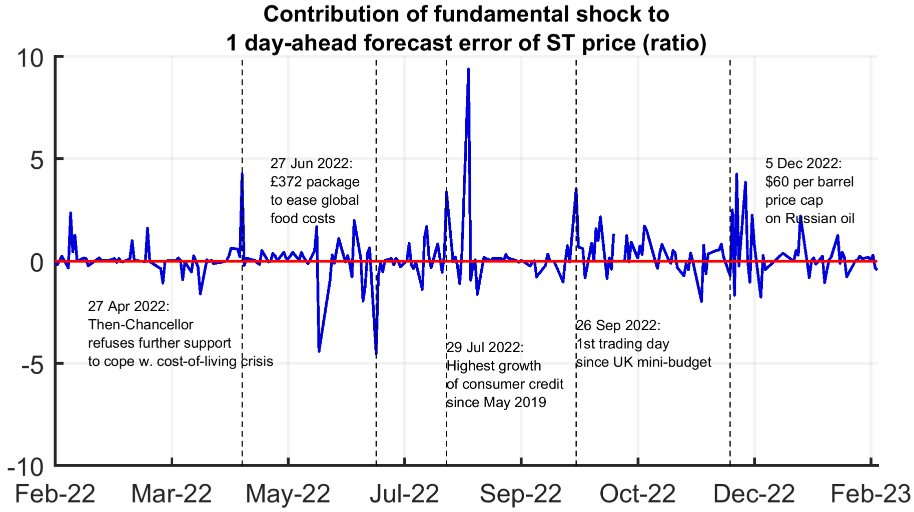
Before dismissing the inflation swap prices from the short horizon market entirely as a signal of fundamentals, we take a more microscopic approach by computing the contribution of the estimated fundamental shocks to one-day ahead forecast errors. This is reported in the top panel of Figure 16, which matches the local maxima in the swap prices to the most pertinent inflation news on that trading date. Even though liquidity shocks dominate, when large fundamental events happen, the inflation swap prices spike. So, at least for large events, they can carry some information about fundamental expectations of inflation.

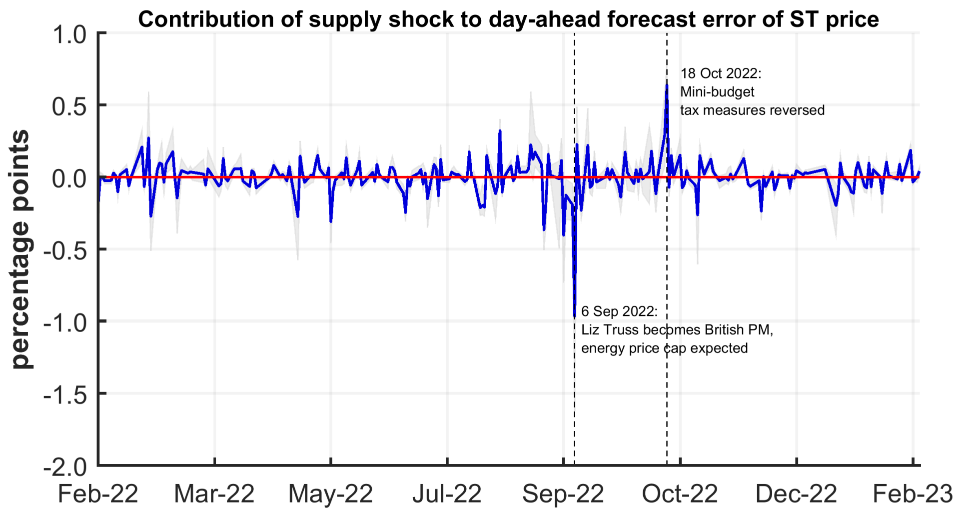
The bottom panel of Figure 16 focuses on the shocks to dealers' provision of inflation protection in the short horizon market. The variance decompositions suggested they account for 10-30% of the forecast error variance. There are two dates when they led to significant fluctuations in prices. The first is 6^{th} September 2022, when Liz Truss became British Prime Minister and was expected to unveil a package of energy price guarantees for households to cope with the rising energy cost. The second is 18^{th} October 2022, when the new government reversed nearly all the tax cuts earlier proposed by Truss, a shock in the opposite direction. We will discuss these in more detail below.

5.5 Historical decomposition: the Covid period.

Historical decompositions reveal the contributions of the unobserved shocks to the observed market prices during important macroeconomic episodes. One such episode was the initial Covid-19 virus period between February to September 2020 that coincided with the first cases of the virus and the periods following the announcement of the WHO declaring Covid-19 as a pandemic. The fall

Figure 16 CONTRIBUTION OF FUNDAMENTAL AND LIQUIDITY SHOCKS TO 1-DAY AHEAD FORECAST ERROR OF SHORT HORIZON INFLATION SWAP PRICES





in demand owing to repeated lockdowns led to fears of a depression and deflation.

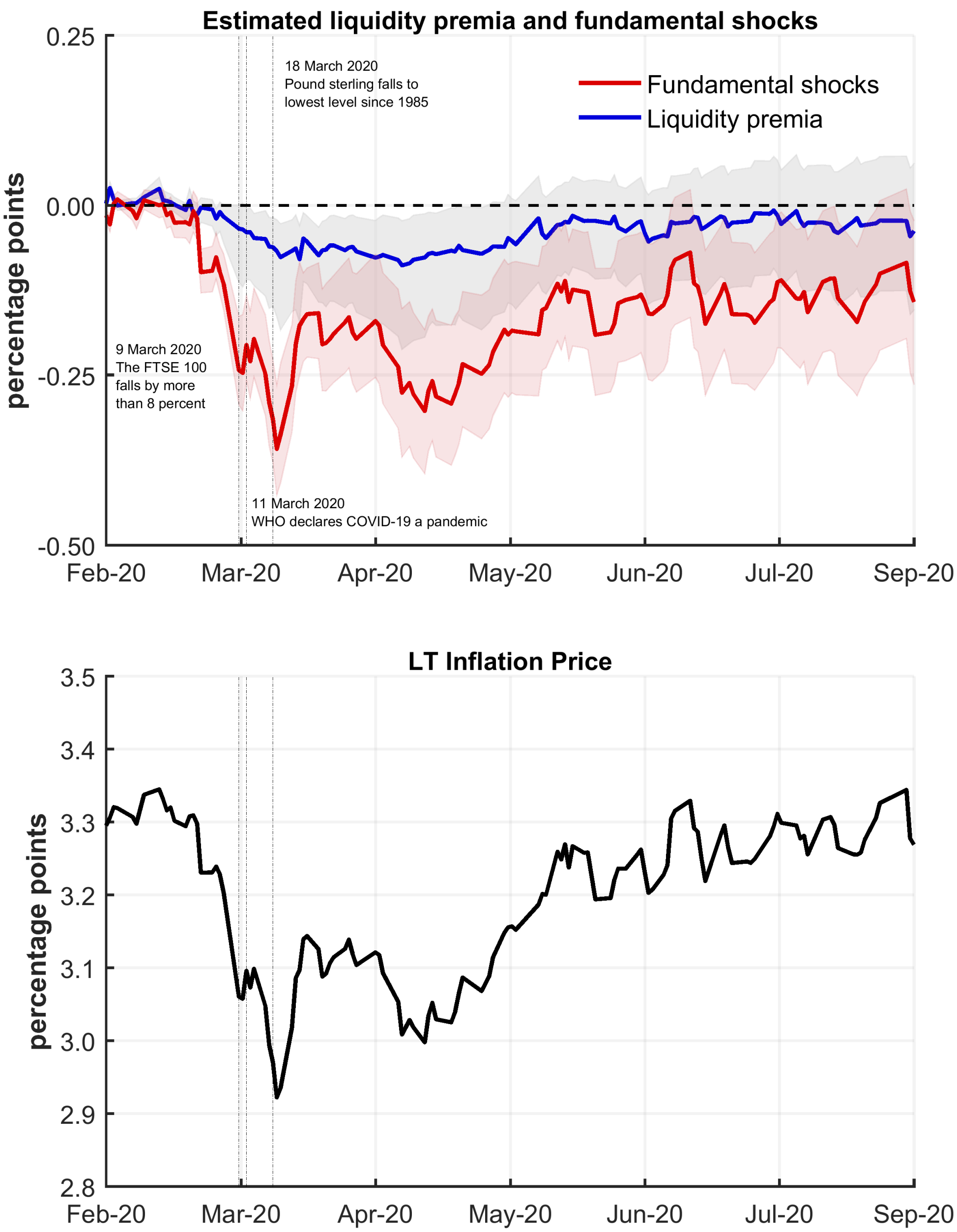
The top panel in Figure 17 shows the contribution of fundamentals and liquidity shocks to the inflation swap prices in the long-horizon market during this period, which are reported in the bottom panel.[21] The fall in swap prices was driven both by lower expected inflation and by liquidity shocks. Looking at swap prices alone would lead one to overstate the deflationary forces during the spring and summer of 2020.

Looking at a few dates validates our decomposition. Fundamental expected inflation started to dive on 9^{th} March 2020 when the FTSE 100 fell by more than 8 percent, before reaching a trough on 11^{th} March 2020 when the WHO officially declared Covid-19 to be a global pandemic. This

[21]The difference between the series in the bottom panel and the sum of the two series in the top panel is the contributions of the deterministic terms in the VAR.

movement reversed on 18^{th} March 2020 when the Pound depreciated to its lowest exchange rate since 1985, providing an inflationary shock. Overall, the fact that 10-year fundamental expected inflation fell on average by only 20bps during this period shows a comfortable degree of anchoring.

Figure 17 The Covid pandemic: swap rates, liquidity, and fundamentals



5.6 Historical decomposition: the star of the Ukraine war.

The second large shock to inflation during our sample was the start of the Ukraine War on the 24^{th} of February 2022. The months that followed came with large increases in the prices of crude oil and energy in the UK, and a re-evaluation of geopolitical constraints on trade. The price of inflation swaps rose sharply as we can see in Figure 18.

Again though, the rise in prices overstates the inflationary shock, because liquidity shocks were also elevating the prices by about 10bps during this period, as dealers constrained relative supply as demand grew. Our estimates are already able to pick up a rise in fundamental expected inflation by roughly 15bps in the lead-up to the war. Satellite information revealed heightened levels of Russian military activity at the border to Ukraine and financial markets were quick to incorporate the news. The date of the actual invasion provided a further increase of roughly the same magnitude.

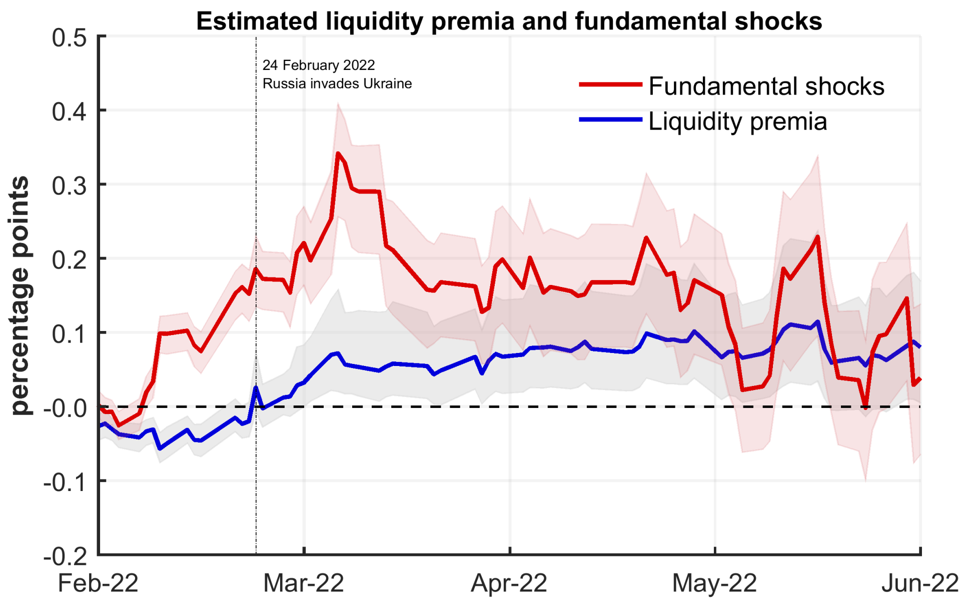
5.7 Historical decomposition: the energy price guarantees and the UK LDI crisis.

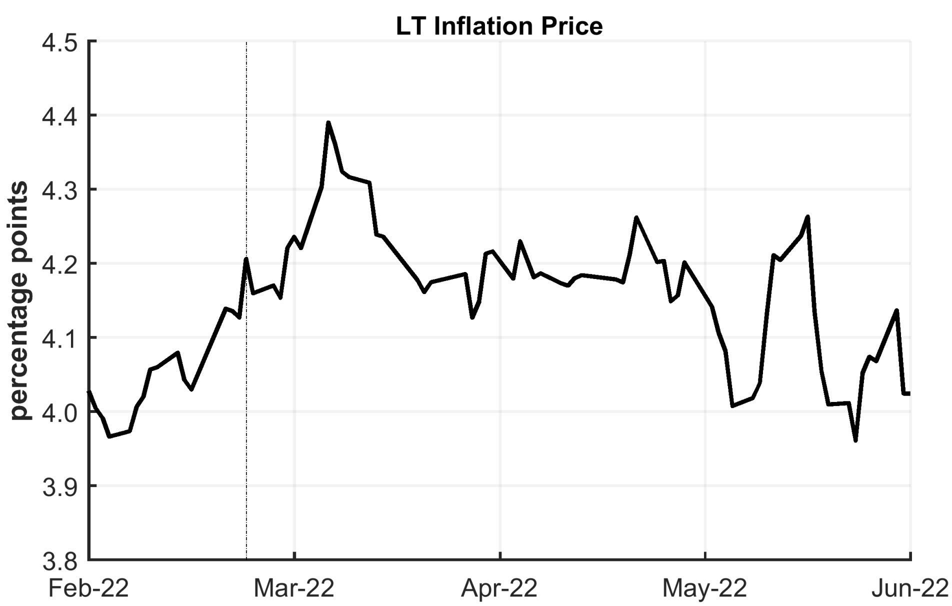
On the 5^{th} September 2022, Liz Truss became the UK's new Prime Minister having promised to tackle the UK's cost-of-living crisis brought about by the war in Ukraine. On the 6^{th} and 8^{th} September 2022, the government announced an "energy price guarantee", a price cap that substantially cut the effective prices that consumers would be paying for their household energy bills. This policy would have a large effect on measured headline RPI inflation in the following 12 to 24 months. On the 23^{rd} September 2022, the "Mini-Budget" was announced with large unfunded tax cuts. This fiscal expansion triggered a substantial fall in bond prices. This resulted in a sector-wide sell off of long-horizon bonds by liability-driven investment funds (the LDI in PFLDI) with further knock-on effects on wider financial stability. In order to stabilize the market, the Bank of England announced on 28^{th} September that it would temporarily buy a limited number of long-dated government bonds. One month later, on October 25^{th}, Rishi Sunak became Prime Minister and Jeremy Hunt the new Chancellor of the Exchequer. One of their first measures was to revert the tax cuts.

Figure 19 shows that in September and October of 2022, inflation swap prices were unusually volatile. The Truss energy caps reduced long-horizon inflation expectations, but the announcement of the fiscal expansion increased them sharply at first. This was quickly reversed as the Bank of England intervened, perhaps a reflection of successful communication with respect to the anchoring of expectations. One month later, when the tax cuts were reserved, expected inflation fell by approximately 20bps and stayed persistently lower. After the dust settled, this combination of shocks lowered long-horizon expected inflation.

Looking at the swap rates, liquidity shocks pushed those prices higher by more than 10bps, obscuring the fall in fundamentals. This contraction in supply by dealer banks may be a reflection of lasting concerns about the health of their counterparties in the PFLDI sector. Earlier, in

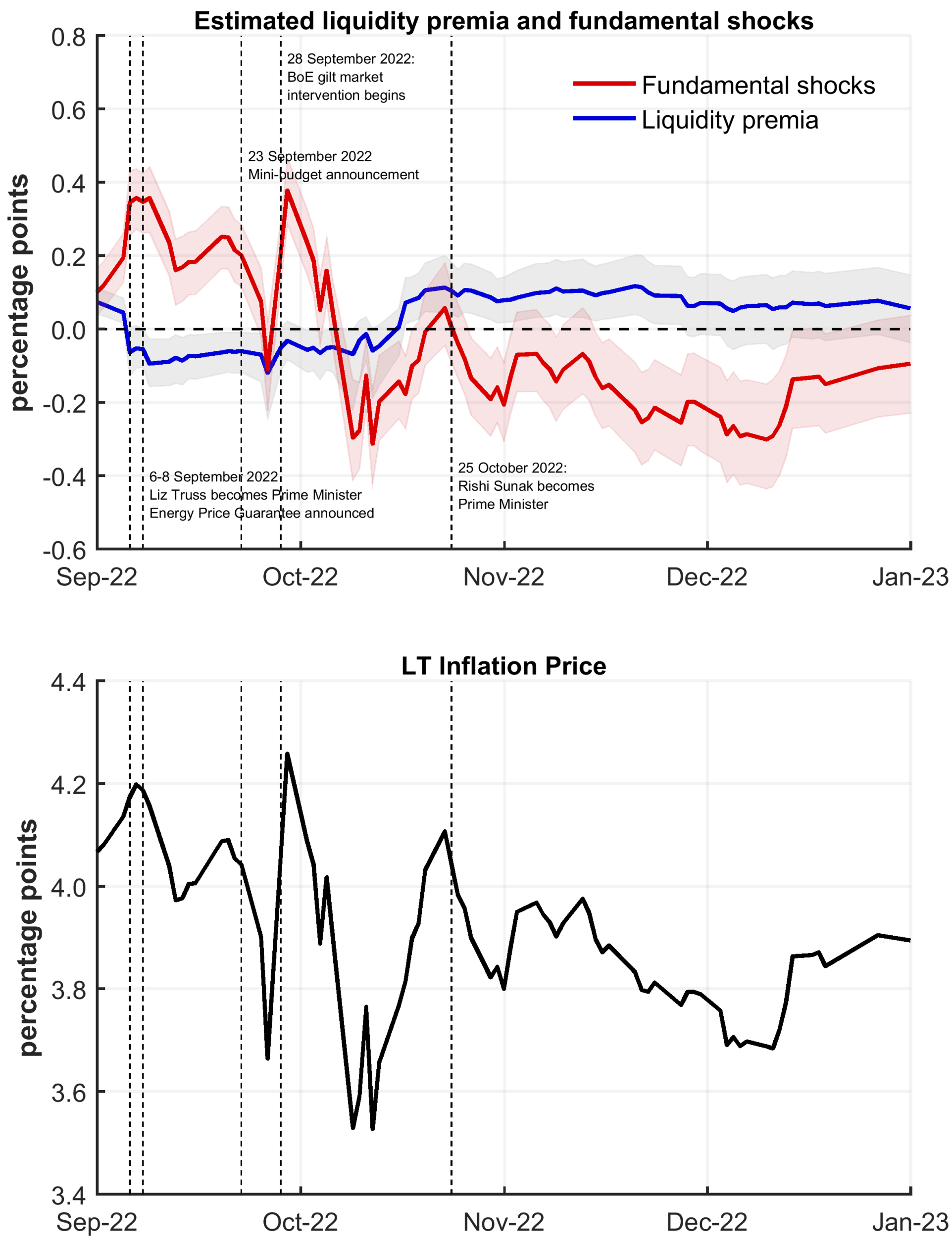
Figure 18 THE RUSSIA-UKRAINE WAR: SWAP RATES, LIQUIDITY, AND FUNDAMENTALS





September, the dominant liquidity shock came from the demand side, and pushed prices down. According to our estimates, PFLDIs were temporarily constrained in their ability to buy inflation swaps, which is consistent with the crisis in the sector during that month, and with the sector-wide deleveraging taking place in that month.

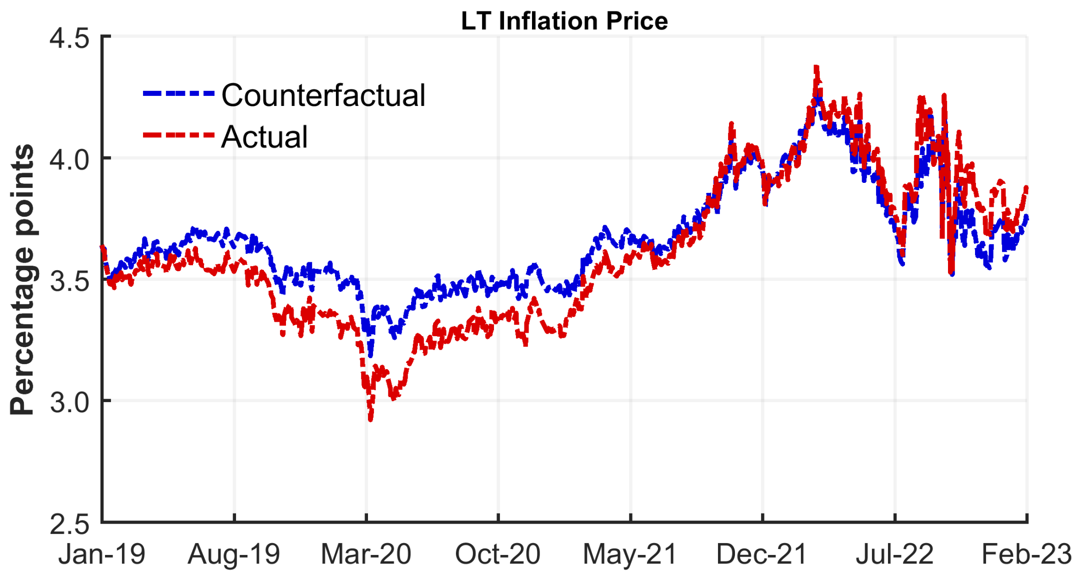
Figure 19 Autumn 2022: swap rates, liquidity, and fundamentals



5.8 Historical decomposition: a longer view on fundamental inflation

Zooming out from specific episodes, Figure 20 presents an historical decomposition for our entire sample. In red is the realized measure of p, a conventional measure of long term expected inflation. The blue line instead presents the price that deducts all the shocks to liquidity frictions from the realized price. As discussed before, our counterfactual measure points to higher expected inflation throughout 2020 — during the lockdowns of the Covid-19 pandemic — than the conventional measure. The period of financial disruption in Autumn 2022 generates a second significant deviation as conventional measures overstated expected inflation. This persists until February of 2023, and suggests a lower risk of inflation expectations being away from target.

Figure 20 FUNDAMENTAL EXPECTED INFLATION



6 Conclusion

This paper complements the prices of inflation swap contracts that are heavily used by researchers and policymakers with valuable data on the quantities behind the prices. We provide an analysis of the demand and supply functions in this large market in which inflation risk is traded. This leads to the following lessons about inflation risk:

First, there is a remarkable segmentation of the market across horizons. In short horizons, hedge funds and dealers are active, alternating between negative and positive net positions that average to zero. In long horizons, dealers steadily provide inflation protection to pension funds. Therefore, dealers provide the link between the two markets by supplying swaps to both.

Second, because we have high-frequency trading data at the institution-level, for four years with many news, we can identify the four supply and demand functions in this market. Three separate identification strategies exploit different dimensions of the data to do so. Reassuringly, the three yield similar empirical estimates, suggesting that our results are quite robust.

Third, in the market, prices seem to fully reflect information after two to three days. The markets are not fully efficient in that there is persistence from the day of impact to the next day. At the same time, this does not last beyond three days. Perhaps this is because investors in this market are very sophisticated, although subject to different constraints and operating with different beliefs.

Fourth, we find that the slope of the supply function of dealer banks is much smaller at long horizons than in the short horizon market. It remains unclear whether this is due to institutional characteristics, or due to the relative bargaining power of different clients, but future research can explore this intriguing result.

Fifth, fundamental shocks drive the long-horizon swap prices, while liquidity shocks drive the short-horizon prices. Therefore, even though one-year inflation swap prices are sometimes used (amongst other measures) to quantify expected inflation and to guide monetary policy, our estimates suggest that much of the variation in these prices can be explained by liquidity shocks to hedge funds and dealers. Comparing our estimates with key dates of market dysfunction confirms this view.

Sixth, and finally, we produce a measure of expected inflation at longer horizons that is cleaned of liquidity frictions. Over the four years of our sample, this measure did not fall as much during the Covid-19 pandemic, and did not rise as much in the second half of 2022. At the start of 2023, it painted a more optimistic picture of the anchoring of inflation expectations relative to conventional measures.

References

Abad, J., I. Aldasoro, C. Aymanns, M. D'Errico, L. F. Rousová, P. Hoffmann, S. Langfield, M. Neychev, and T. Roukny (2016): " Shedding light on dark markets: First insights from the new EU-wide OTC derivatives dataset," *ESRB: Occasional Paper Series No. 11.*

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2018): "Inference Based on Structural Vector Autoregressions Identified With Sign and Zero Restrictions: Theory and Applications," *Econometrica*, 86, 685–720.

Bai, J. (2009): "Panel Data Models With Interactive Fixed Effects," *Econometrica*, 77, 1229–1279.

Bai, J. and S. Ng (2002): "Determining the number of factors in approximate factor models." *Econometrica*, 70(1), 191–221.

Beechey, M. J., B. K. Johannsen, and A. T. Levin (2011): "Are Long-Run Inflation Expectations Anchored More Firmly in the Euro Area Than in the United States?" *American Economic Journal: Macroeconomics*, 3, 104–29.

Begenau, J., M. Piazzesi, and M. Schneider (2015): "Banks' risk exposures," *NBER Working Paper No. 21334.*

Benos, E., G. Ferrara, and A. Ranaldo (2022): "Collateral Cycles," *Bank of England Staff Working Paper No. 966.*

Boons, M., F. Duarte, F. de Roon, and M. Szymanowska (2020): "Time-varying inflation risk and stock returns," *Journal of Financial Economics*, 136, 444–470.

Brunnermeier, M., D. Palia, K. A. Sastry, and C. A. Sims (2021): "Feedbacks: Financial Markets and Economic Activity," *American Economic Review*, 111, 1845–79.

Campbell, J. Y., C. Pflueger, and L. M. Viceira (2020): "Macroeconomic Drivers of Bond and Equity Risks," *Journal of Political Economy*, 128, 3148–3185.

Cenedese, G., P. Della Corte, and T. Wang (2021): "Currency Mispricing and Dealer Balance Sheets," *The Journal of Finance*, 76, 2763–2803.

Cenedese, G., A. Ranaldo, and M. Vasios (2020): "OTC premia," *Journal of Financial Economics*, 136, 86–105.

Cieslak, A. and C. Pflueger (2023): "Inflation and Asset Returns," CEPR discussion paper 18003.

Covi, G., J. Brookes, and C. Raja (2022): "Measuring Capital at Risk in the UK banking sector: a microstructural network approach," *Bank of England Staff Working Paper No. 983.*

Czech, R., P. Della Corte, S. Huang, and T. Wang (2022): "FX Option Volume," *Bank of England Staff Working Paper No. 964.*

D'Amico, S. and T. B. King (2023): "Inflation hedging products," in *The Research Handbook of Financial Markets*, ed. by R. Gurkaynak and J. Wright, Edward-Elgar.

Fang, X., Y. Liu, and N. Roussanov (2022): "Getting to the Core: Inflation Risks Within and Across Asset Classes," Working Paper 30169, National Bureau of Economic Research.

Ferrara, G., P. Mueller, G. Viswanath-Natraj, and J. Wang (2022): "Central bank swap lines: micro-level evidence," *Bank of England Staff Working Paper No. 977.*

Gabaix, X. (2016): "Power laws in economics: An introduction," *Journal of Economic Perspectives*, 30, 185–206.

Gabaix, X. and R. S. Koijen (2020): "Granular instrumental variables," *National Bureau of Economic Research Working Paper Series.*

——— (2021): "In search of the origins of financial fluctuations: The inelastic markets hypothesis," *National Bureau of Economic Research Working Paper Series.*

Gabaix, X., R. S. Koijen, F. Mainardi, S. S. Oh, and M. Yogo (2022): "Asset Demand of U.S. Households," Harvard Umniversity manuscript.

Hanson, S., A. Malkhozov, and G. Venter (2022): "Demand and Supply Imbalance Risk and Long-Term Swap Spreads," *Unpublished working paper.*

Hau, H., P. Hoffmann, S. Langfield, and Y. Timmer (2021): "Discriminatory Pricing of Over-the-Counter Derivatives," *Management Science*, 67, 6660–6677.

Haubrich, J., G. Pennacchi, and P. Ritchken (2012): "Inflation Expectations, Real Rates, and Risk Premia: Evidence from Inflation Swaps," *The Review of Financial Studies*, 25, 1588–1629.

Jiang, E., G. Matvos, T. Piskorski, and A. Seru (2023): "Limited Hedging and Gambling for Resurrection by U.S. Banks During the 2022 Monetary Tightening?" *Unpublished Working Paper.*

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

Koijen, R. S., R. J. Richmond, and M. Yogo (2020): "Which investors matter for equity valuations and expected returns?" *National Bureau of Economic Research Working Paper Series.*

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

Koijen, R. S. J. and M. Yogo (2019): "A Demand System Approach to Asset Pricing," *Journal of Political Economy*, 127, 1475–1515.

Lütkepohl, H., M. Meitz, A. Netšunajev, and P. Saikkonen (2021): " Testing identification via heteroskedasticity in structural vector autoregressive models," *The Econometrics Journal*, 24, 1–2.

Mann, C. L. (2022): "Inflation expectations, inflation persistence, and monetary policy strategy," *Speech at the 53rd Annual Conference of the Money Macro and Finance Society, University of Kent.*

McPhail, L., P. Schnabl, and B. Tuckman (2023): "Do Banks Hedge Using Interest Rate Swaps?" *NBER Working Paper No. 31166.*

Ramsden, D. (2022): "Shocks, inflation, and the policy response," *Speech at the at Securities Industry Conference.*

Reis, R. (2020): "The People versus the Markets: A Parsimonious Model of Inflation Expectations," CEPR discussion paper XXX.

Stock, J. H. and M. W. Watson (2018): "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments," *The Economic Journal*, 128, 917–948.

Vayanos, D. and J.-L. Vila (2021): "A Preferred-Habitat Model of the Term Structure of Interest Rates," *Econometrica*, 89, 77–112.

Appendix

A Data: Additional Information

This section of the Appendix provides additional information on the data used in the empirical analysis.

A.1 Details on Cleaning Steps

In this section we briefly summarize the key steps that have been implemented to clean the trade-level data from the EMIR DTCC Trade Repository and to construct a daily time series of prices quantities in both the short-term (ST) and long-term (LT) inflation swap market.

We first define the preliminaries of our empirical work by restricting our data sample to daily DTCC OTC interest rate trade state files from the 2nd of January 2019 to the 10th of February 2023. Thus, our entire raw dataset consists of more than one thousand trade state files, each containing a stock of approximately 2.5-3.5 million outstanding trade-level reports at the time of the report.

Next, we describe the main steps taken to clean a single DTCC OTC interest rate trade state report:

1. We start by identifying inflation swap contracts from the raw report. We first extract the string that is associated with the floating leg fields of the derivative contract and use regular expressions to exhaustively capture strings that refer to inflation indexes. Subsequently, we check for the product classification type of the derivative contract and use its ISO 10962 6-character CFI code to verify whether it is an inflation swap contract. For example, a CFI code "SRGCSC" stands for: (S): Swaps; (R): Rates; (G): Underlying assets: inflation rate index; (C): Notional: constant; (S): Single or multi-currency: single; (C): Delivery: Cash. We only keep the raw swap trade reports that can be identified with a recognisable inflation index.

2. Next, we further shrink down the raw data by dropping non-effective trades at the time of report that will not go into effect even within the next ten years, as well as trades that have either been terminated or matured.

3. We then implement an extensive de-duplication of pairwise identical transaction reports at the counterparty level. Due to (UK) EMIR reporting requirements, both counterparties in a transaction must report the same trade to the trade repository. This leads to many trade reports that can be identified as pair-wise duplicates, one of which must be dropped to prevent double counting in the data. Such pair-wise reports are identified by their trade ID and the legal entity identifiers (LEI) of both counterparties, except that the reporting counterparty LEI and the LEI of the other counterparty are swapped depending on which counterparty reports the trade. For such pair-wise reports, we further validate their internal consistency by checking whether both reports contain the same contract-level information such as notional amount, date of maturity, intragroup status and whether they have the opposite counterparty sides (i.e., if one reporting counterparty to the pair reports the trade as a seller, the other reporting counterparty to the same pair must report the trade as a buyer). We drop a minority of such pairs where these contract-level information is inconsistent and proceed to keeping the latest report for the pairs that demonstrate such an internal consistency.

4. At this juncture, we have obtained a dataset consisting of all outstanding unique transactions of inflation swap contracts. We then proceed to drop the raw contract-level reports that have implausible notional amounts (less than $1,000 or more than $10bn), that are intragroup transactions, and those that are compression trades.

5. Finally, we obtain a cleaned trade state report by removing a tiny minority of trade reports that do not satisfy UK EMIR validation rules. We then allocate investors to an investor group using a best-endeavour sectoral classification, which is naturally subject to uncertainties (e.g. allocation of insurer with asset management arm).

Next, we append all cleaned DTCC OTC interest rate trade state files into a single dataset and restrict the data sample to dealer-client trades in UK RPI inflation swaps. This results in a total of more than 25 million observations over our sample period from January 2019 to February 2023.

A.2 Additional Summary Statistics

In this section, we provide additional summary statistics for a single trade state report, covering all outstanding trades in the UK RPI inflation swap market on 10^{th} February 2023. As shown in Table A.1, we observe the largest number of trades by pension funds, followed by insurers, LDI funds and hedge funds. Moreover, we find the largest average trades size for hedge funds with an average notional of $70m.

When analyzing the notionals across maturities, we find the largest outstanding notionals for the 1-year, 10-year and 3-year+ contracts. The average trade size seems to be decreasing with the maturity of the contract.

Table A.1 SUMMARY STATISTICS: DTCC DERIVATIVES REPOSITORY TRADE STATE REPORT

	Gross Notional	Mean	Std. Deviation	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile	Total Observations
By investor type									
PFLDIs	296,223.7	21.6	(43.9)	0.5	2.8	8.4	23.2	84.5	13,696
Hedge Funds	162,600.3	69.4	(111.3)	5.9	15.6	32.8	81.4	236.3	2,342
Non-dealer banks	44,491.2	21.3	(40.1)	0.3	1.8	8.6	24.7	82.7	2,085
Others	341,242.8	37.6	(91.5)	0.4	3.6	13.4	38.5	140.9	9,073
By initial maturity									
3-year or less	186,482.5	89.4	(169.2)	3.6	16.7	46.3	96.8	311.2	2,086
3 to 10-year, excl.	191,680.2	54.7	(88.9)	1.3	7.8	26.0	66.7	204.8	3,505
10-year or more	466,395.3	21.6	(43.9)	0.4	2.7	8.8	23.3	81.3	21,605
All	844,558.0	31.1	(71.7)	0.5	3.4	11.3	30.9	120.2	27,196

NOTE: All columns except the last are in units of USD millions. The category "Others" includes asset managers, central banks, insurers, non-financials, other financials, sovereign wealth funds, state entities, supranationals, trading services and proprietary trading firms. "All" also coincides with the statistics pertaining to the dealer bank sector since we report statistics on the dealer-client segment of the market. SOURCE: 10 February 2023 DTCC OTC interest rate trade state file.

B Proofs

B.1 The Asset Demand System of Pension Funds

In this section, we state the problem of pension funds (f) in the market for long horizon inflation swap contracts and formally derive their demand system.

Proof. Let each pension fund institution be denoted by (i). Each pension fund solves a portfolio allocation problem to maximise CARA utility given by:

$$\mathbb{E}_{f,i}\Big[-exp\big(-\widetilde{\gamma}_{f,i}a'_{f,i}\big)\Big], \ \ \widetilde{\gamma}_{f,i} = \frac{\gamma_{f,i}}{a_{f,i}} \tag{B.1}$$

where $a'_{f,i}$ is the terminal market value of pension fund i's chosen portfolio, and $\widetilde{\gamma}_{f,i}$ is its Arrow–Pratt measure of absolute risk aversion. Note that this is obtained from a scaling by the institution's initial position $a_{f,i}$, so that the degree of risk-aversion does not decrease mechanically with the size of the institution.

In any given period, the value of this portfolio is determined by the following budget constraint:

$$a'_{f,i} = a_{f,i} + (\pi - p)q_{f,i} + (d - s)e_{f,i} + y_{f,i} \tag{B.2}$$

This equation clarifies the environment and makes it transparent that the pension fund institution has access to *three* markets.

First, $q_{f,i}$ denotes its portfolio choice of inflation swap contracts that it can buy from the inflation market. In the data, this captures the net notional position. When a pension fund goes long on an inflation swap contract, it pays the swap breakeven rate p to its counterparty and receives the floating inflation rate π. In each payment period, $\pi - p$ can thus be understood as a marked-to-market return that is linked to realised inflation. We assume that π is a stochastic random variable, and each pension fund institution potentially disagrees on its first moment:

$$\mathbb{E}_{f,i}[\pi] = \mu_{f,i}\pi^e \tag{B.3}$$

where π^e denotes the fundamental expectation of inflation that is consistent with a rational expectations equilibrium. Second, $e_{f,i}$ denotes the total allocation of its portfolio over other market securities that can be thought of as a Lucas tree asset. This component of the portfolio yields a stochastic dividend d at the cost of s. Third, there exists a market for a risk-free asset with a risk-free return normalised to zero.[22] Lastly, $y_{f,i}$ denotes a source of mean-zero income risk that is idiosyncratic at the institutional-level. Taking market prices as given, the informational structure from the perspective of pension fund (f, i) is given by:

$$\boldsymbol{\nu} = \begin{pmatrix} \mathbb{E}_{f,i}[\pi] - p \\ \mathbb{E}_{f,i}[d] - s \\ \mathbb{E}_{f,i}[y_{f,i}] \end{pmatrix} = \begin{pmatrix} \mu_{f,i}\pi^e - p \\ \theta_d - s \\ 0 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2_\pi & \sigma_{\pi,d} & \sigma_{\pi,y_{f,i}} \\ \sigma_{d,\pi} & \sigma^2_d & \sigma_{d,y_{f,i}} \\ \sigma_{y_{f,i},\pi} & \sigma_{y_{f,i},d} & \sigma^2_{y_{f,i}} \end{pmatrix} \tag{B.4}$$

where $\boldsymbol{\nu}$ and $\boldsymbol{\Sigma}$ are the expected return vector and variance-covariance matrix respectively to a Gaussian distribution that characterises the stochasticity in this economy.

Next, we define an auxiliary term $\Delta a_{f,i}$ that equals to the change in the market value of the institution's portfolio overnight:

$$\Delta a_{f,i} = (\pi - p)q_{f,i} + (d - s)e_{f,i} + y_{f,i} \tag{B.5}$$

Since $\Delta a_{f,i}$ is a linear combination of random variables distributed by the Gaussian distribution, it also inherits the Gaussian distribution with the following moments:

$$\mathbb{E}_{f,i}[\Delta a_{f,i}] = (\mu_{f,i}\pi^e - p)q_{f,i} + (\theta_d - s)e_{f,i} \equiv \nu_{\Delta a_{f,i}} \tag{B.6}$$

$$\begin{aligned} Var_{f,i}[\Delta a_{f,i}] &= q^2_{f,i}Var[\pi - p] + e^2_{f,i}Var[d - s] + Var[y_{f,i}] \\ &+ 2q_{f,i}e_{f,i}Cov[(\pi - p)(d - s)] + 2q_{f,i}Cov[(\pi - p)y_{f,i}] \\ &+ 2e_{f,i}Cov[(d - s)y_{f,i}] \\ &\equiv \sigma^2_{\Delta a_{f,i}} \end{aligned} \tag{B.7}$$

[22]That is, the gross risk-free return equals 1 in this economy. This is without loss of generality.

Given so, the pension fund's objective is to maximise:

$$
\begin{aligned}
\mathbb{E}_{f,i}\Big[-exp\big(-\widetilde{\gamma}_{f,i}a'_{f,i}\big)\Big] &= \mathbb{E}_{f,i}\Big[-exp\big(-\widetilde{\gamma}_{f,i}(a_{f,i}+\Delta a_{f,i})\big)\Big] \\
&= -\frac{e^{-\gamma_{f,i}}}{\sqrt{2\pi\sigma^2_{\Delta a_{f,i}}}}\int_{-\infty}^{\infty} e^{-\left(\widetilde{\gamma}_{f,i}\Delta a_{f,i}+\frac{(\Delta a_{f,i}-\nu_{\Delta a_{f,i}})^2}{2\sigma^2_{\Delta a_{f,i}}}\right)} d\Delta a_{f,i} \\
&= -\frac{e^{-\gamma_{f,i}}}{\sqrt{2\pi\sigma^2_{\Delta a_{f,i}}}}\int_{-\infty}^{\infty} e^{-\left(\widetilde{\gamma}_{f,i}\left[\nu_{\Delta a_{f,i}}-\frac{\widetilde{\gamma}_{f,i}\sigma^2_{\Delta a_{f,i}}}{2}\right]+\frac{[\Delta a_{f,i}-(\nu_{\Delta a_{f,i}}-\widetilde{\gamma}_{f,i}\sigma^2_{\Delta a_{f,i}})]^2}{2\sigma^2_{\Delta a_{f,i}}}\right)} d\Delta a_{f,i}
\end{aligned}
\tag{B.8}
$$

This expression can be simplified considerably by noting that the probability density function of a random variable x with mean $\mathbb{E}[x]=\nu_{\Delta a_{f,i}}-\widetilde{\gamma}_{f,i}\sigma^2_{\Delta a_{f,i}}$ and variance $Var[x]=\sigma^2_{\Delta a_{f,i}}$ must satisfy:

$$
\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2_{\Delta a_{f,i}}}}e^{-\frac{[x-(\nu_{\Delta a_{f,i}}-\widetilde{\gamma}_{f,i}\sigma^2_{\Delta a_{f,i}})]^2}{2\sigma^2_{\Delta a_{f,i}}}}dx = 1 \tag{B.9}
$$

Substituting this equation above yields the following expression for the pension fund's objective:

$$
\mathbb{E}_{f,i}\Big[-exp\big(-\widetilde{\gamma}_{f,i}a'_{f,i}\big)\Big] = A_{f,i}e^{-\widetilde{\gamma}_{f,i}\left[\nu_{\Delta a_{f,i}}-\frac{\widetilde{\gamma}_{f,i}\sigma^2_{\Delta a_{f,i}}}{2}\right]} \tag{B.10}
$$

where $A_{f,i}=-exp(-\gamma_{f,i})$ is a constant.

It is straightforward to see that this is equivalent to a mean-variance maximisation problem. Using the auxiliary definitions of $\nu_{\Delta a_{f,i}}$ and $\sigma^2_{\Delta a_{f,i}}$, a formal statement of this problem is given by:

$$
\max_{\{q_{f,i},e_{f,i}\}}\Big[(\mu_{f,i}\pi^e-p)q_{f,i}+(\theta_d-s)e_{f,i}\Big]-\frac{\widetilde{\gamma}_{f,i}}{2}\Big[q^2_{f,i}\sigma^2_\pi+e^2_{f,i}\sigma^2_d+\sigma^2_{y_{f,i}}+2q_{f,i}e_{f,i}\sigma_{\pi,d}+2q_{f,i}\sigma_{\pi,y_{f,i}}+2e_{f,i}\sigma_{d,y_{f,i}}\Big]
$$

$$
s.t.\;\; G_f(q_{f,i},z_{f,i})\geq 0 \tag{B.11}
$$

Where $G_f(\cdot,\cdot)$ is a general constraint on the pension fund's ability to take positions in the inflation swap market. We assume that this function is continuous and differentiable. Taking market prices as given, the First-Order conditions necessary to attain a maximum are given by:

$$
q^*_{f,i}=\frac{\mu_{f,i}\pi^e-p^*}{\widetilde{\gamma}_{f,i}\sigma^2_\pi}-\frac{\sigma_{\pi,d}}{\sigma^2_\pi}e^*_{f,i}-\frac{\sigma_{\pi,y_{f,i}}}{\sigma^2_\pi}-\frac{\lambda_{f,i}G'_f(q^*_{f,i},z_{f,i})}{\widetilde{\gamma}_{f,i}\sigma^2_\pi} \tag{B.12}
$$

$$
e^*_{f,i}=\frac{\theta_d-s^*}{\widetilde{\gamma}_{f,i}\sigma^2_d}-\frac{\sigma_{\pi,d}}{\sigma^2_d}q^*_{f,i}-\frac{\sigma_{d,y_{f,i}}}{\sigma^2_d} \tag{B.13}
$$

Where $\lambda_{f,i}$ is the Lagrange multiplier associated with institution i's capacity constraint that satisfies the two-part Kuhn-Tucker conditions:

$$
\lambda_{f,i}G_f(q^*_{f,i},z_{f,i})=0,\;\; G_f(q^*_{f,i},z_{f,i})\geq 0,\;\; \lambda_{f,i}\geq 0 \tag{B.14}
$$

and $q^*_{f,i}$ denotes the institution's optimal portfolio allocation of inflation swap contracts in equilibrium. Combining the First-Order conditions above, and substituting the definition $\widetilde{\gamma}_{f,i}=\gamma_{f,i}/a_{f,i}$, the solution for $q^*_{f,i}$ can be purely obtained as a function of the model's primitives:

$$\frac{q^*_{f,i}}{a_{f,i}} = \frac{\mu_{f,i}\pi^e - p^*}{\gamma_{f,i}\sigma^2_\pi(1-\rho^2_{\pi,d})} - \frac{\sigma_d}{\sigma_\pi}\left[\frac{\theta_d - s^*}{\gamma_{f,i}\sigma^2_d(1-\rho^2_{\pi,d})}\right]\rho_{\pi,d} - \left[\frac{1}{(1-\rho^2_{\pi,d})\sigma^2_\pi}\right]\left(\frac{\sigma_{\pi,y_{f,i}}}{a_{f,i}} + \frac{\lambda_{f,i}g_f(q^*_{f,i}, z_{f,i})}{\gamma_{f,i}}\right) \quad \text{(B.15)}$$

where this solution assumes that the pension fund's idiosyncratic income risk is only correlated with π, that is, $\sigma_{d,y_{f,i}} = 0$, and $\rho^2_{\pi,d} = \sigma_{\pi,d}/\sigma_\pi\sigma_d$ is the correlation between π and d. This completes the proof. ■

B.2 The Asset Demand System of Dealer Banks

We now turn our attention to the dealer banks:

Proof. In the data, dealer banks are active in *both* the long horizon and short horizon markets by nature of being counterparties to both pension funds and hedge funds. Therefore, dealer banks differ from client institutions in that they have access to an additional market. The market value of their portfolio evolves according to:

$$a'_{b,i} = a_{b,i} + (\pi - p)q_{b,i} + (\Pi - P)Q_{b,i} + (d - s)e_{b,i} + y_{b,i} \quad \text{(B.16)}$$

where $q_{b,i}$ and $Q_{b,i}$ are dealer banks' demand of inflation swaps (measured in the data as a net notional position) in the long horizon and short horizon inflation market respectively. As with the client institutions, dealer banks also have general capacity constraints on their ability to take on risk in both inflation markets. However, given the high-frequency nature of our transaction-level data, we will assume *desk separation* within a period (i.e., one trading day), such that the ability for traders on one desk to take on risk in the short horizon market is independent of the risk taken by traders on the other desk which serves the long horizon market — at least within a day prior to books being balanced. Formally, this implies that dealer banks face two independent constraints given by:

$$G^S_b(Q_{b,i}, z_{b,i}) \geq 0, \quad \frac{\partial G^S_b(Q_{b,i}, z_{b,i})}{\partial Q} = g^S_b(Q_{b,i}, z_{b,i}) \quad \text{(B.17)}$$

$$G^L_b(q_{b,i}, z_{b,i}) \geq 0, \quad \frac{\partial G^L_b(q_{b,i}, z_{b,i})}{\partial q} = g^L_b(q_{b,i}, z_{b,i}) \quad \text{(B.18)}$$

where $G^S_b(\cdot,\cdot)$ and $G^L_b(\cdot,\cdot)$ are general functions that capture this constraint on risk-taking in both markets, and we assume that they are continuous and differentiable. Similar to before, $z_{b,i}$ is an exogenous shifter that governs the tightness of the constraints. Given that dealer banks also share the same CARA utility as the client institutions, it is straightforward to follow the same steps as before to obtain their optimal asset demand as a solution to a mean-variance maximisation problem given below:

$$\begin{aligned}\max_{\{q_{b,i}, Q_{b,i}, e_{b,i}\}} \quad & \Big[(\mu_{b,i}\pi^e - p)q_{b,i} + (\mu_{b,i}\Pi^e - P)Q_{b,i} + (\theta_d - s)e_{b,i}\Big] \\ & - \frac{\widetilde{\gamma}_{b,i}}{2}\Big[q^2_{b,i}\sigma^2_\pi + Q^2_{b,i}\sigma^2_\Pi + e^2_{b,i}\sigma^2_d + \sigma^2_{y_{b,i}} + 2q_{b,i}Q_{b,i}\sigma_{\pi,\Pi} + 2q_{b,i}e_{b,i}\sigma_{\pi,d} \\ & + 2q_{b,i}\sigma_{\pi,y_{b,i}} + 2Q_{b,i}e_{b,i}\sigma_{\Pi,d} + 2Q_{b,i}\sigma_{\Pi,y_{b,i}} + 2e_{b,i}\sigma_{d,y_{b,i}}\Big] \\ & s.t. \;\; G^S_b(Q_{b,i}, z_{b,i}) \geq 0, \;\; G^L_b(q_{b,i}, z_{b,i}) \geq 0 \quad \text{(B.19)}\end{aligned}$$

Taking market prices as given, the First-Order conditions necessary to attain a maximum are given by:

$$q^*_{b,i} = \frac{\mu_{b,i}\pi^e - p^*}{\widetilde{\gamma}_{b,i}\sigma^2_\pi} - \frac{\sigma_{\pi,\Pi}}{\sigma^2_\pi}Q^*_{b,i} - \frac{\sigma_{\pi,d}}{\sigma^2_\pi}e^*_{b,i} - \frac{\sigma_{\pi,y_{b,i}}}{\sigma^2_\pi} - \frac{\lambda^L_{b,i}g^L_b(q^*_{b,i}, z_{b,i})}{\widetilde{\gamma}_{b,i}\sigma^2_\pi} \quad \text{(B.20)}$$

$$Q^*_{b,i} = \frac{\mu_{b,i}\Pi^e - P^*}{\widetilde{\gamma}_{b,i}\sigma^2_\Pi} - \frac{\sigma_{\pi,\Pi}}{\sigma^2_\Pi}q^*_{b,i} - \frac{\sigma_{\Pi,d}}{\sigma^2_\Pi}e^*_{b,i} - \frac{\sigma_{\Pi,y_{b,i}}}{\sigma^2_\Pi} - \frac{\lambda^S_{b,i}g^S_b(Q^*_{b,i}, z_{b,i})}{\widetilde{\gamma}_{b,i}\sigma^2_\Pi} \tag{B.21}$$

$$e^*_{b,i} = \frac{\theta_d - s^*}{\widetilde{\gamma}_{b,i}\sigma^2_d} - \frac{\sigma_{\pi,d}}{\sigma^2_d}q^*_{b,i} - \frac{\sigma_{\Pi,d}}{\sigma^2_d}Q^*_{b,i} - \frac{\sigma_{d,y_{b,i}}}{\sigma^2_d} \tag{B.22}$$

Next, we eliminate demand for the Lucas tree asset $e^*_{b,i}$ from the system using substitution. Maintaining the assumption that $\sigma_{d,y_{b,i}} = 0$, and leaving out some algebra for brevity, this implies that $q^*_{b,i}$ and $Q^*_{b,i}$ can be expressed respectively as:

$$q^*_{b,i} = \frac{\mu_{b,i}\pi^e - p^*}{\widetilde{\gamma}_{b,i}\sigma^2_\pi(1-\rho^2_{\pi,d})} - \frac{\sigma_d}{\sigma_\pi}\left[\frac{\theta_d - s^*}{\widetilde{\gamma}_{b,i}\sigma^2_d(1-\rho^2_{\pi,d})}\right]\rho_{\pi,d} - \frac{\sigma_\Pi}{\sigma_\pi}\left[\frac{\rho_{\pi,\Pi} - \rho_{\pi,d}\rho_{\Pi,d}}{1-\rho^2_{\pi,d}}\right]Q^*_{b,i} - \left[\frac{1}{(1-\rho^2_{\pi,d})\sigma^2_\pi}\right]\left(\sigma_{\pi,y_{b,i}} + \frac{\lambda^L_{b,i}g^L_b(q^*_{b,i}, z_{b,i})}{\widetilde{\gamma}_{b,i}}\right) \tag{B.23}$$

$$Q^*_{b,i} = \frac{\mu_{b,i}\Pi^e - P^*}{\widetilde{\gamma}_{b,i}\sigma^2_\Pi(1-\rho^2_{\Pi,d})} - \frac{\sigma_d}{\sigma_\Pi}\left[\frac{\theta_d - s^*}{\widetilde{\gamma}_{b,i}\sigma^2_d(1-\rho^2_{\Pi,d})}\right]\rho_{\Pi,d} - \frac{\sigma_\pi}{\sigma_\Pi}\left[\frac{\rho_{\Pi,\pi} - \rho_{\Pi,d}\rho_{\pi,d}}{1-\rho^2_{\Pi,d}}\right]q^*_{b,i} - \left[\frac{1}{(1-\rho^2_{\Pi,d})\sigma^2_\Pi}\right]\left(\sigma_{\Pi,y_{b,i}} + \frac{\lambda^S_{b,i}g^S_b(Q^*_{b,i}, z_{b,i})}{\widetilde{\gamma}_{b,i}}\right) \tag{B.24}$$

From the expressions above, it becomes clear that for $Q^*_{b,i}$ (and its exogenous shifters) not to be a part of the solution for $q^*_{b,i}$ and vice versa, the variance-covariance matrix of stochastic returns in this economy must additionally be restricted such that:

$$\rho_{\pi,\Pi} = \rho_{\pi,d}\rho_{\Pi,d} \tag{B.25}$$

The expression for asset demand from dealer banks in the main text can subsequently be attained by using the definition for $\widetilde{\gamma}_{b,i} = \gamma_{b,i}/a_{b,i}$. ■

B.3 The Frictionless Equilibrium

Proof. The proof is standard and follows directly from the market clearing condition that holds in equilibrium. Noting that markets are complete, all idiosyncratic background risks can be fully insured away. All general capacity constraints on risk-taking are non-binding. Both imply that, $\sigma_{\pi,y_{\zeta,i}} = 0$ for $\zeta \in \{f, b\}$ and $\lambda^L_{b,i} = \lambda_{f,i} = 0$. Let Θ_f and Θ_b each denote a measurable set of pension funds and dealer banks respectively. In this long horizon market, market clearing requires:

$$\sum_{i\in\Theta_f}\widetilde{q}^*_{f,i} = -\sum_{i\in\Theta_b}\widetilde{q}^*_{b,i} \equiv \widetilde{q}^* > 0 \tag{B.26}$$

Substituting in the asset demand (supply) functions for each pair of counterparty institutions and grouping terms, it is straightforward to solve for the frictionless swap price as a function of fundamental expectation of inflation and a compensation for risk premium:

$$\widetilde{p}^* = \left[\frac{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i}\mu_{f,i}}{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i} + \sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i}} + \frac{\sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i}\mu_{b,i}}{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i} + \sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i}}\right]\pi^e - \frac{\theta_d - \widetilde{s}^*}{\sigma^2_d}\sigma_{\pi,d} \tag{B.27}$$

■

C Identification with Granular Instrumental Variables

In this section, we elaborate further on the identification strategy using granular instrumental variables. For exposition, consider the case of identifying the demand shock in the long horizon inflation swap market in which case we have to find a granular IV for ε_f.

Recall the demand system for PFLDIs and append to it a time dimension to allow for persistence in the data:

$$\frac{q^*_{f,i,t}}{a_{f,i,t}} = \frac{\mu_{f,i,t}\pi^e_t - p^*_t}{\gamma_{f,i}\sigma^2_\pi(1-\rho^2_{\pi,d})} - \left(\frac{\sigma_d}{\sigma_\pi}\right)\left[\frac{\theta_{d,t} - s^*_t}{\gamma_{f,i}\sigma^2_d(1-\rho^2_{\pi,d})}\right]\rho_{\pi,d} \underbrace{-\left[\frac{1}{(1-\rho^2_{\pi,d})\sigma^2_\pi}\right]\left(\frac{\sigma_{\pi,y_{f,i,t}}}{a_{f,i,t}} + \frac{\lambda_{f,i,t}g_{f,i,t}}{\gamma_{f,i}}\right)}_{+\varepsilon_{f,i,t}} \quad \text{(C.1)}$$

The fund-specific demand shock arising is denoted by $\varepsilon_{f,i,t}$. In equilibrium, the observed market price for the swap contract in the long market is a combination of the fundamental expectation of inflation, liquidity premia, minus a compensation for risk premia:

$$p^*_t = \Lambda_t\pi^e_t - rp^*_t + lp^*_t \quad \text{(C.2)}$$

where the components are defined as:

$$\Lambda_t = \frac{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i,t}\mu_{f,i,t}}{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i,t} + \sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i,t}} + \frac{\sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i,t}\mu_{b,i,t}}{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i,t} + \sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i,t}} \quad \text{(C.3)}$$

$$rp^*_t = \frac{\theta_d - s^*}{\sigma^2_d}\sigma_{\pi,d} \quad \text{(C.4)}$$

$$lp^*_t = -\frac{\sum_{i\in\Theta_b}\left\{\sigma_{\pi,y_{b,i}} + \frac{\lambda^L_{b,i,t}g^L_{b,i,t}}{\widetilde{\gamma}_{b,i,t}}\right\}}{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i,t} + \sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i,t}} - \frac{\sum_{i\in\Theta_f}\left\{\sigma_{\pi,y_{f,i}} + \frac{\lambda_{f,i,t}g_{f,i,t}}{\widetilde{\gamma}_{f,i,t}}\right\}}{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i,t} + \sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i,t}} \quad \text{(C.5)}$$

We make the following assumptions on $\mu_{f,i,t}$ and $\varepsilon_{f,i,t}$:

$$\mu_{f,i,t} = \kappa^\mu_{f,i}\mu_{f,t} \quad \text{(C.6)}$$

$$\varepsilon_{f,i,t} = \frac{\kappa^{lp}_{f,i}lp^*_t}{\gamma_{f,i}\sigma^2_\pi(1-\rho^2_{\pi,d})} + \tilde{\varepsilon}_{f,i,t} \quad \text{(C.7)}$$

Where (C.6) is an extension of (3), $\kappa^\mu_{f,i}$ measures the deviation of pension fund i's subjective expectation of fundamental inflation relative to the entire pension fund sector. $\tilde{\varepsilon}_{f,i,t}$ refers to the idiosyncratic component of fund-specific demand shock $\varepsilon_{f,i,t}$. These assumptions imply that the demand system above can be rewritten as:

$$\frac{q^*_{f,i,t}}{a_{f,i,t}} = \frac{(\kappa^\mu_{f,i}\mu_{f,t} - \Lambda_t)\pi^e_t + (\kappa^{lp}_{f,i} - 1)lp^*_t}{\gamma_{f,i}\sigma^2_\pi(1-\rho^2_{\pi,d})} + \widetilde{\varepsilon}_{f,i,t} = \boldsymbol{\omega}'_{f,i}\mathbf{F}_t + \widetilde{\varepsilon}_{f,i,t} \quad \text{(C.8)}$$

where $\mathbf{F}_t$ and $\boldsymbol{\omega}_{f,i}$ are respectively the common factors and the fund-specific factor loadings that map exactly into:

$$\mathbf{F}_t = \begin{pmatrix}\pi^e_t \\ lp^*_t\end{pmatrix}, \quad \boldsymbol{\omega}_{f,i} = \begin{pmatrix}\frac{(\kappa^\mu_{f,i}\mu_{f,t}-\Lambda_t)}{\gamma_{f,i}\sigma^2_\pi(1-\rho^2_{\pi,d})} \\ \frac{\kappa^{lp}_{f,i}-1}{\gamma_{f,i}\sigma^2_\pi(1-\rho^2_{\pi,d})}\end{pmatrix} \quad \text{(C.9)}$$

The granular IV thus isolates idiosyncratic shocks to institutional demand by removing all the common factors, $\mathbf{F}_t$. This equation can then be estimated using interactive fixed effects following the method by Bai (2009). Our empirical implementation estimates a modified regression equation that allows for persistence in institutional demand:

$$\frac{q_{f,i,t}}{a_{f,i,t}} = \sum_{j=1}^{J} \beta_j \frac{q_{f,i,t-j}}{a_{f,i,t-j}} + \boldsymbol{\omega}'_{f,i}\mathbf{F}_t + \widetilde{\varepsilon}_{f,i,t} \tag{C.10}$$

where we use $J = 3$ lags for this estimation to be consistent with the estimation of the Bayesian sign restrictions SVAR. Although our model implies a two-factor model, we allow for a larger number of factors to capture other sources of heterogeneity within the sector orthogonal to the model's components that could lead to disturbances in demand in the data, such as differences in fund structures (e.g., liability-driven investment, defined contributions). The number of factors are first chosen according to Bai and Ng (2002), and we adjust them appropriately such that they satisfy the identification of the fundamental shock. This procedure led to 11 factors being estimated for PFLDI demand, 11 for hedge fund demand and 8 for dealer bank supply. This implementation requires us to first construct a balanced panel using the DTCC trade repository data that allows us to observe the trading activity of each institution − identified by their legal entity identifiers − on a daily frequency. This allows us to measure e.g., $q_{f,i,t}$ if we were to construct the granular IV for the long horizon inflation market. We then measure $a_{f,i,t}$ using the gross notional amount of all outstanding inflation swap contracts traded by institution i from the PFLDI sector, and scale the quantities appropriately prior to estimating in this regression equation.[23] The granular IV for pension fund demand that emerges from this procedure is given by:

$$GIV_{f,t} = \sum_{i\in\Theta_f} a_{f,i,t}\widetilde{\varepsilon}_{f,i,t} \tag{C.11}$$

Recall from Equation (12) that the sector-wide demand shock from pension funds $\varepsilon_{f,t}$ is given by:

$$\varepsilon_{f,t} = -\frac{\sum_{i\in\Theta_f}\left\{\sigma_{\pi,y_{f,i}} + \frac{\lambda_{f,i}g_{f,i}}{\widetilde{\gamma}_{f,i,t}}\right\}}{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i,t} + \sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i,t}} \tag{C.12}$$

Using the assumption from (C.7) and the definition of $\varepsilon_{f,i,t}$, the equation above can be written instead as:

$$\varepsilon_{f,t} = \frac{\sum_{i\in\Theta_f}(\kappa^{lp}_{f,i}lp^*_t)\widetilde{\gamma}^{-1}_{f,i,t}}{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i,t} + \sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i,t}} + (1-\rho^2_{\pi,d})\sigma^2_\pi\frac{GIV_{f,t}}{\sum_{i\in\Theta_f}\widetilde{\gamma}^{-1}_{f,i,t} + \sum_{i\in\Theta_b}\widetilde{\gamma}^{-1}_{b,i,t}} \tag{C.13}$$

It is straightforward to see that $GIV_{f,t}$ qualifies as an instrumental variable for $\varepsilon_{f,t}$. Insofar as there is some granularity in $a_{f,i,t}$, $GIV_{f,t}$ will satisfy the relevance condition since $\mathbb{E}(GIV_{f,t}\varepsilon_{f,t}) \neq 0$.[24] For exclusion restriction, note that $\tilde{\varepsilon}_{fi,t} \perp \varepsilon_{x,t}, \varepsilon_{b,t}, \varepsilon_{\pi,t}$ by construction, since these three shocks are spanned by $F_{f,t}$. Hence, $\mathbb{E}(GIV_{f,t}\varepsilon_{x,t}) = 0$, $\mathbb{E}(GIV_{f,t}\varepsilon_{b,t}) = 0$ and $\mathbb{E}(GIV_{f,t}\varepsilon_{\pi,t}) = 0$. Thus, $GIV_{f,t}$ satisfies all required moments to qualify as a valid instrument for $\varepsilon_{f,t}$.

We subsequently apply the same procedure on data for hedge funds and dealer banks to obtain $GIV_{x,t}$ and $GIV_{b,t}$ as instruments for $\varepsilon_{x,t}$ and $\varepsilon_{b,t}$, which are the demand shocks in the short-horizon market and shocks to dealers' supply functions. We then use the simultaneous equations from the demand system to

[23]In constructing this measure, we carefully tracked the trading activity of each institution in our data sample and accumulated the stock of its outstanding positions by taking account of not only new inflation swap trades, but also trades entered into at the earlier part of the data sample that have eventually matured prior to the cessation of our data sample.

[24]No granularity means $GIV_{f,t} = 0$ as all the idiosyncratic shocks average out.

back out the granular IV for the fundamental shock. Recall that its static representation is given by:

$$\begin{pmatrix} Q_t \\ P_t \\ q_t \\ p_t \end{pmatrix} = constant + \underbrace{\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 & \mathbf{b}_4 \end{bmatrix}}_{} \underbrace{\begin{pmatrix} \varepsilon_{x,t} \\ \varepsilon_{f,t} \\ \varepsilon_{b,t} \\ \varepsilon_{\pi,t} \end{pmatrix}}_{\mathbf{B}} = constant + \begin{pmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \\ u_{4,t} \end{pmatrix} \tag{C.14}$$

Where each $\mathbf{b}_i$ for $i \in \{1, 2, 3, 4\}$ is a 4×1 column vector. We then use $GIV_{f,t}$, $GIV_{x,t}$ and $GIV_{b,t}$ as instruments to project $u_{4,t}$ on $u_{1,t}, u_{2,t}, u_{3,t}$. The residual that emerges from this procedure is then a valid instrument for $\varepsilon_{\pi,t}$, which we label as $GIV_{\pi,t}$. To see this, note that the matrix B by assumption can be inverted to obtain:

$$u_{4,t} = a_{4,1}u_{1,t} + a_{4,2}u_{2,t} + a_{4,3}u_{3,t} + \varepsilon_{\pi,t} \tag{C.15}$$

The residuals from the 2SLS regression exactly yields $\varepsilon_{\pi,t}$.[25] Next, we project u_t sequentially on $GIV_{f,t}$, $GIV_{x,t}$, $GIV_{b,t}$ and $IV_{\pi,t}$. This identifies the coefficients of the structural impact matrix b_1 to b_4 up to sign and scale.

D Robustness: Insurance Companies

In our baseline results, we have focused on the trading volumes of hedge funds in the short-horizon market, and on those of PFLDIs in the long-horizon market. However, as shown in Section 8, another notable player in the inflation swap market are insurance companies, in particular pension insurers. The latter have been heavily involved in the buy-in/buy-outs of pension fund liabilities in recent years—a trend that is set to continue in the coming years.

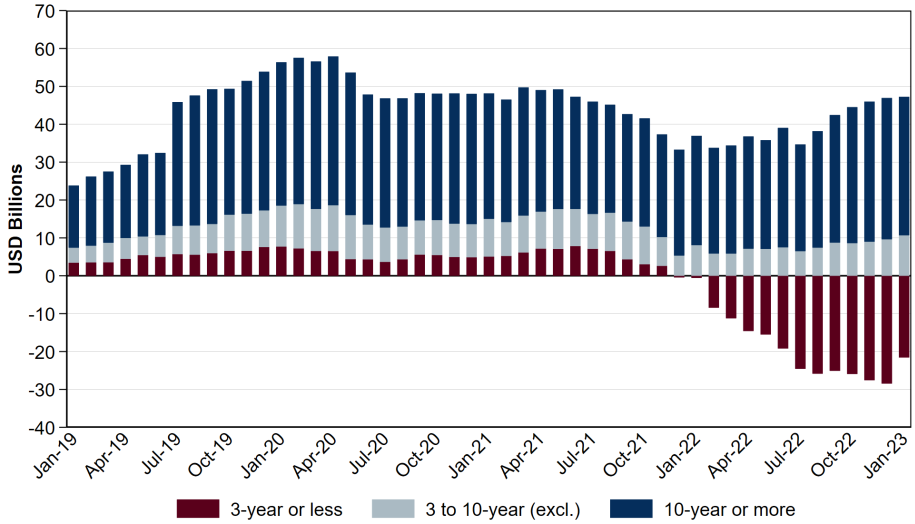
In contrast to pension funds, these insurers mainly use cash-flow driven investment strategies (CDI), which differ substantially from pension funds' usual LDI strategies. More precisely, the aim of the CDI strategy is to create an asset/derivative portfolio that closely matches the cash-flows on the liability side. Importantly, in terms of the inflation-indexation of pension liabilities, the most prevalent form sees inflation linkage floored at zero and capped at 5%, as measured by RPI.

For a CDI-investor that is fully inflation hedged, inflation moving above the 5% cap can be an issue, although usually a positive one: while the investor's inflation-linked asset rises in value, the liability stops tracking the higher inflation and effectively becomes a nominal liability for a period. The insurer will see its assets rise in value by more than its liabilities. But this creates a hedging mismatch: the fund now has too much inflation-linkage.

Given that RPI inflation has been above the 5% cap since late 2021, pension insurers have become net sellers of short-dated inflation swaps to reduce their over-hedged positions, see Figure D.1. Due to the mechanical nature of this re-balancing, we excluded insurers' trading volumes from our main analysis. However, as shown in the next sections, our baseline results remain robust to the inclusion of insurers' quantities.

[25] The numbering of the u's is arbitrary. One could also project $u_{1,t}$ on $u_{2,t}$, $u_{3,t}$ and $u_{4,t}$ instead and get a different $IV_{\pi,t}$ but this would be perfectly collinear with the instrument arising from the alternative projection.

Figure D.1 Net notional position of insurers in UK RPI inflation swaps



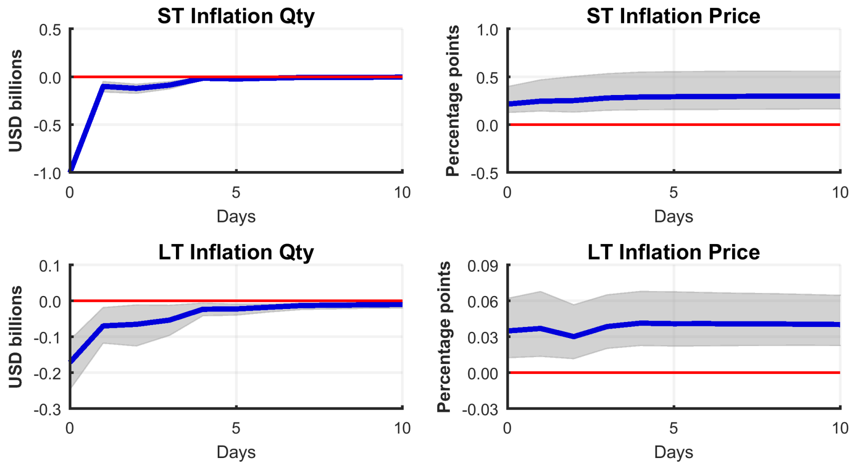
Source: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

D.1 Estimated impulse response functions

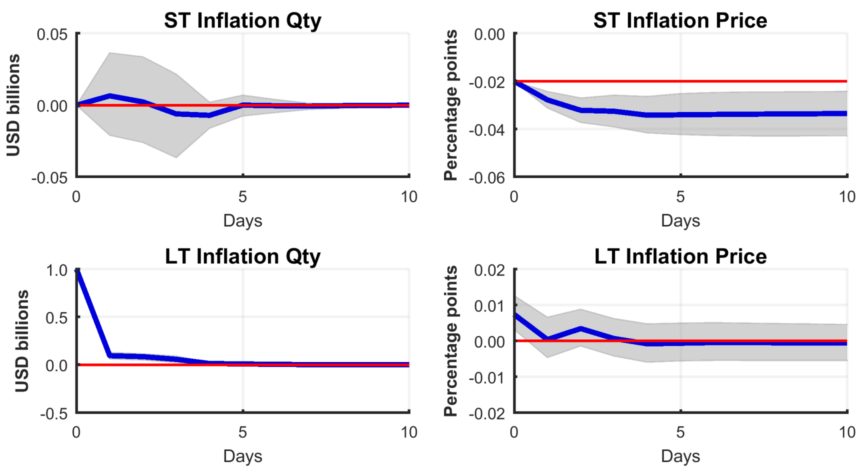
In this section, we report the estimated impulse responses obtained from imposing sign restrictions based on high-frequency timing when quantities of insurance companies are added to our baseline empirical specification. Comparing these estimates to Figure 13, we see that the results from the main text are preserved even when adding the variation that originates from insurance companies. Next, we juxtapose the estimated responses to a fundamental shock across each of the three identification strategies and contrast them with the results in the main text given by 12. Again, we find that the results remain remarkably stable.

Figure D.2 Estimated impulse response functions to liquidity shocks

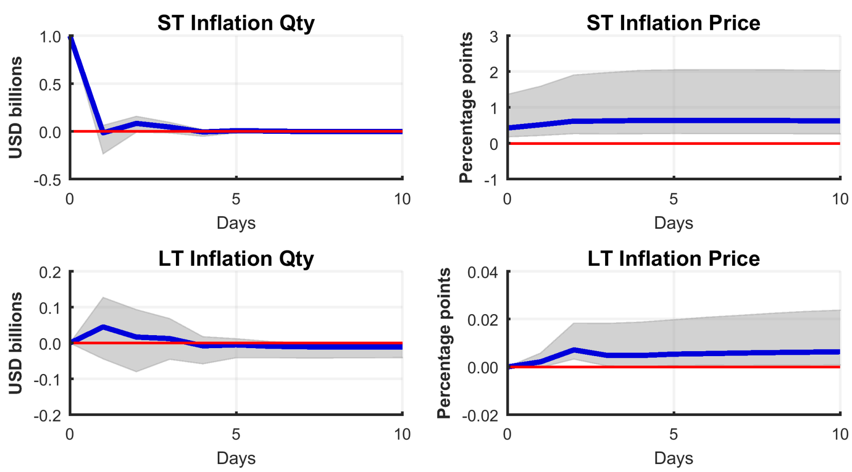
(i) Dealer Supply Shock ($\varepsilon_{b,t}$)



(ii) PFLDI Demand Shock ($\varepsilon_{f,t}$)



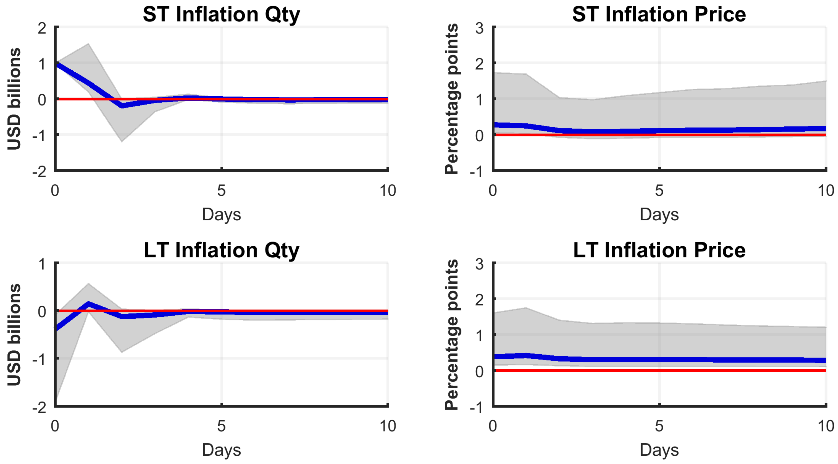
(iii) Hedge Fund Demand Shock ($\varepsilon_{h,t}$)



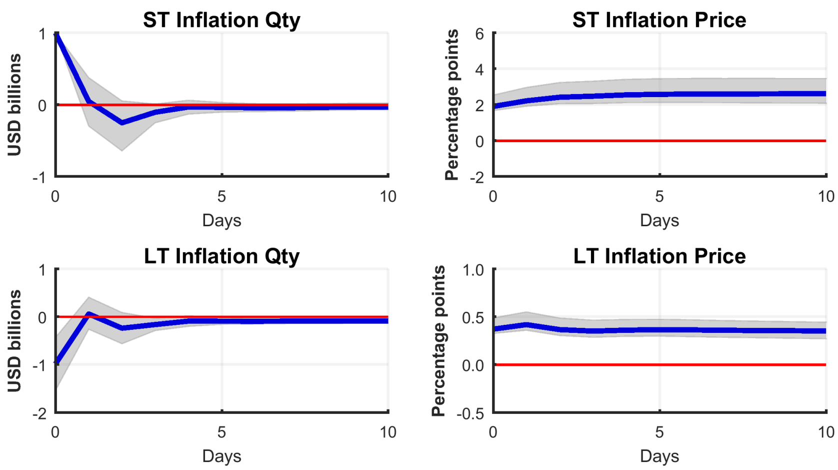
Note: Figure shows the median estimate with 68% confidence intervals.

Figure D.3 Estimated impulse response functions to a fundamental shock

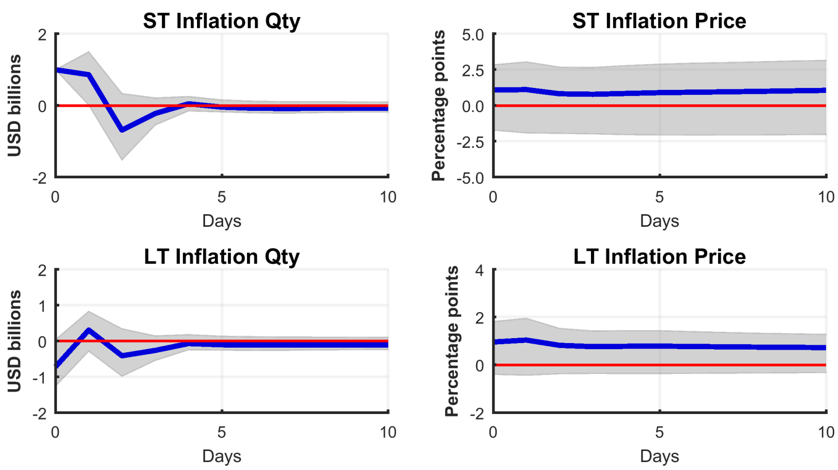
(i) With timing restriction strategy



(ii) With heteroskedasticity strategy



(iii) With granular identification strategy

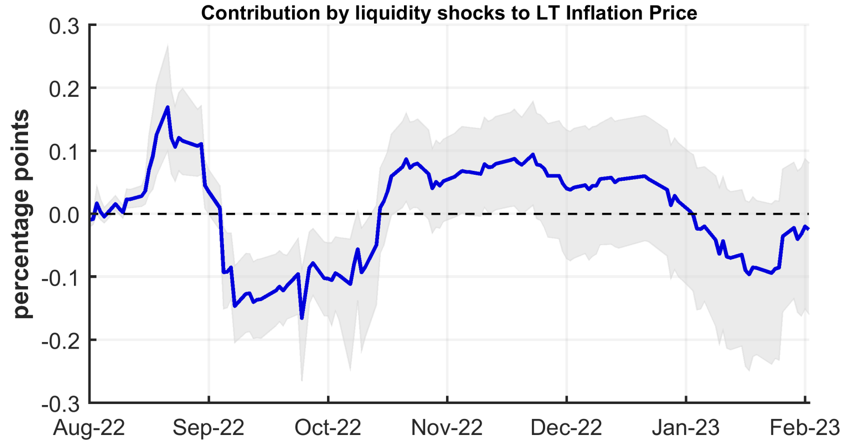


Note: Figure shows the median estimate with 68% confidence intervals.

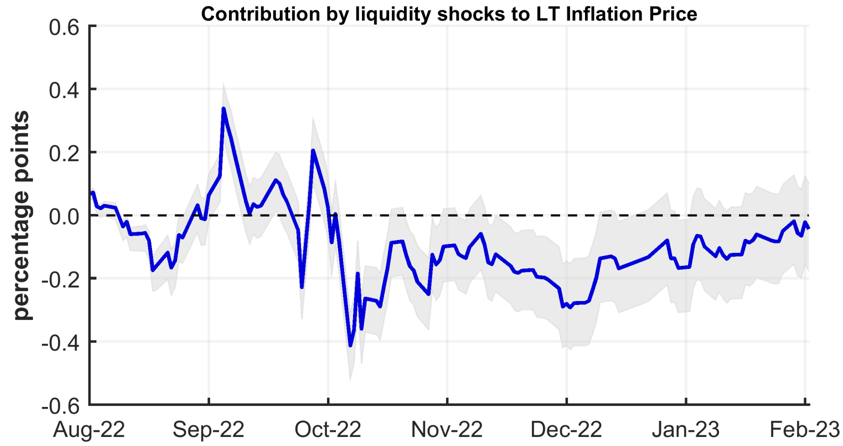
D.2 Estimated liquidity shocks across identification strategies

Figure D.4 ESTIMATED LIQUIDITY SHOCKS

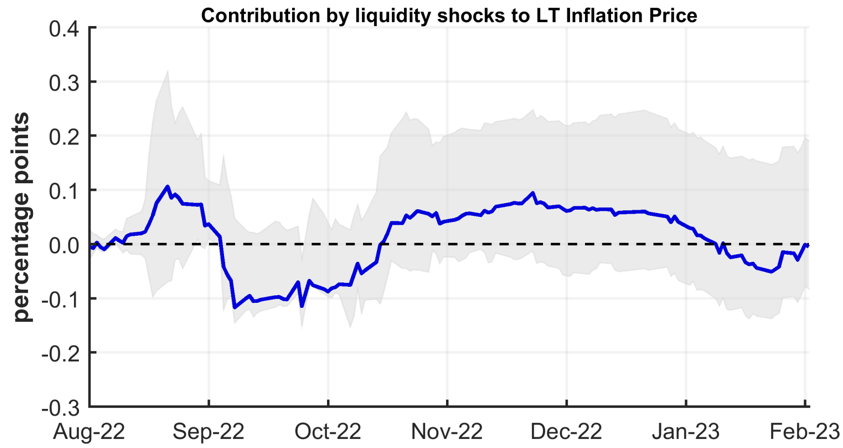
(i) With timing restriction strategy



(ii) With heteroskedasticity strategy



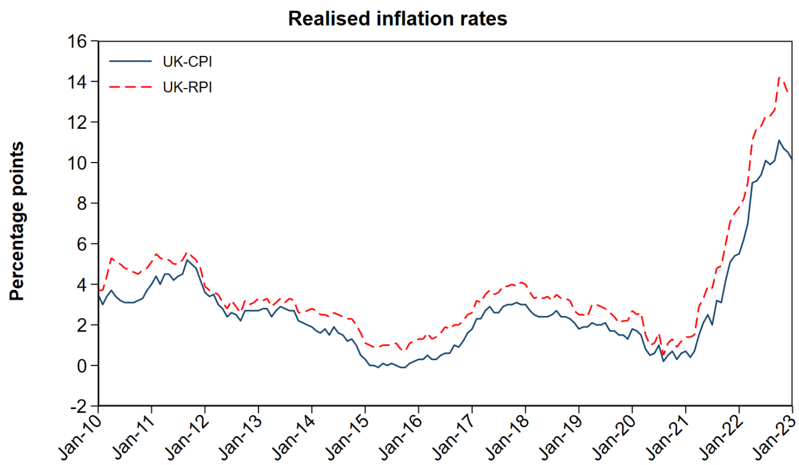
(iii) With granular identification strategy



E Additional Figures

E.1 RPI-CPI wedge

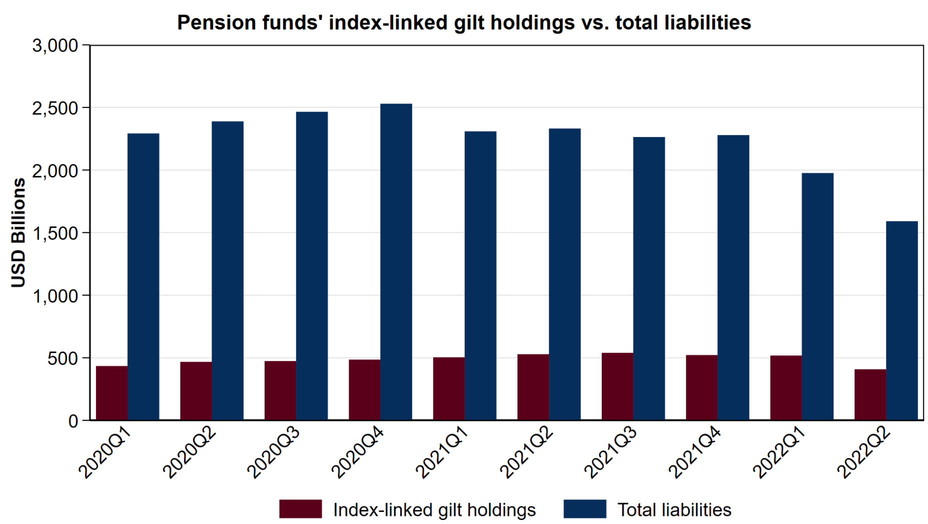
Figure E.1 Comparison of UK RPI and CPI



Note: Time-series comparison of the UK RPI and CPI. Source: Office for National Statistics.

E.2 Pension funds' index-linked gilt holdings

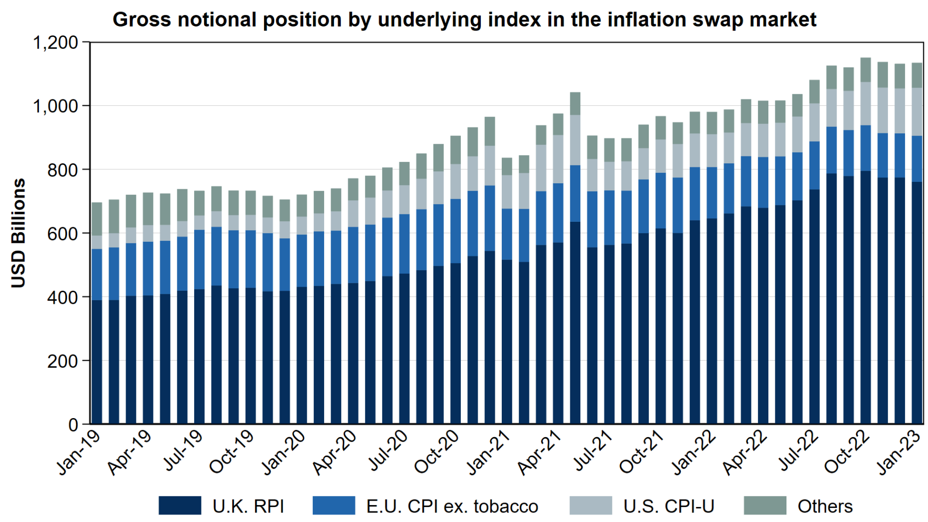
Figure E.2 Comparison of pension funds' index-linked gilt holdings and total liabilities



Note: Comparison of pension funds' index-linked gilt holdings and total liabilities in $ billions. Source: Pension Protection Fund PPF 7800 Data and Office for National Statistics.

E.3 Gross notional positions in the DTCC trade repository

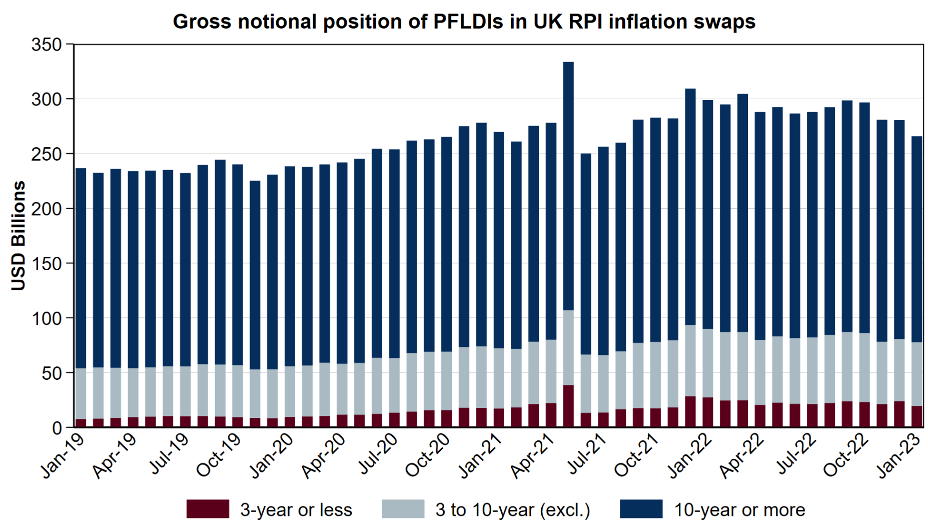
Figure E.3 Outstanding positions by underlying inflation index



Source: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

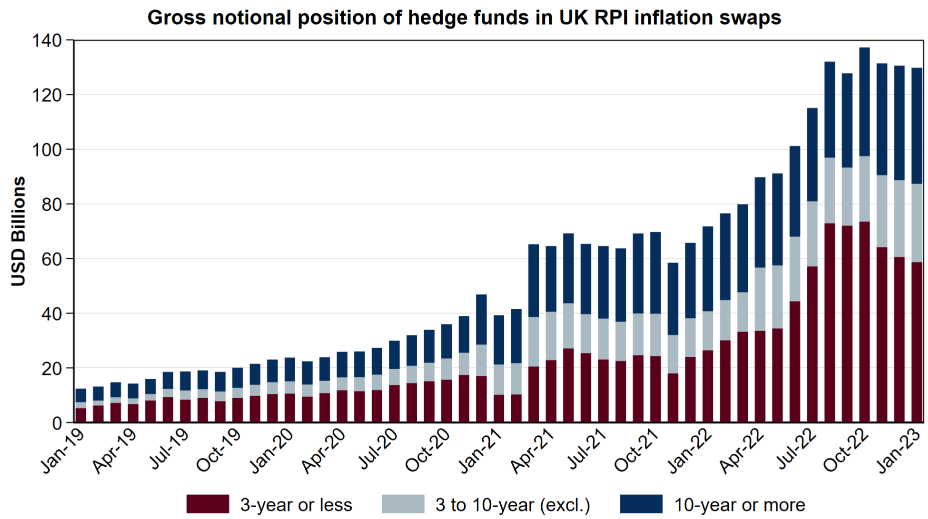
E.4 Gross notional positions in the UK RPI inflation swap market

Figure E.4 Pension fund and liability investment driven institutions



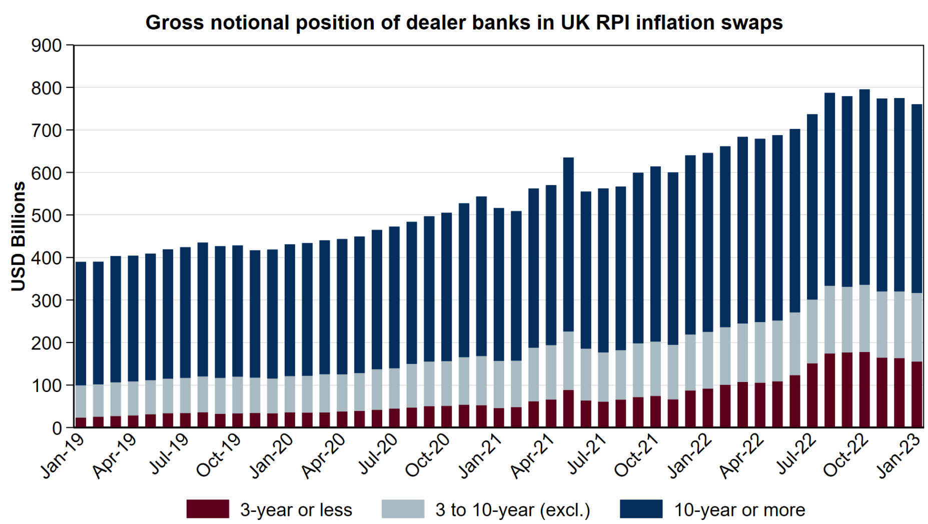
Source: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

Figure E.5 Hedge funds



Source: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

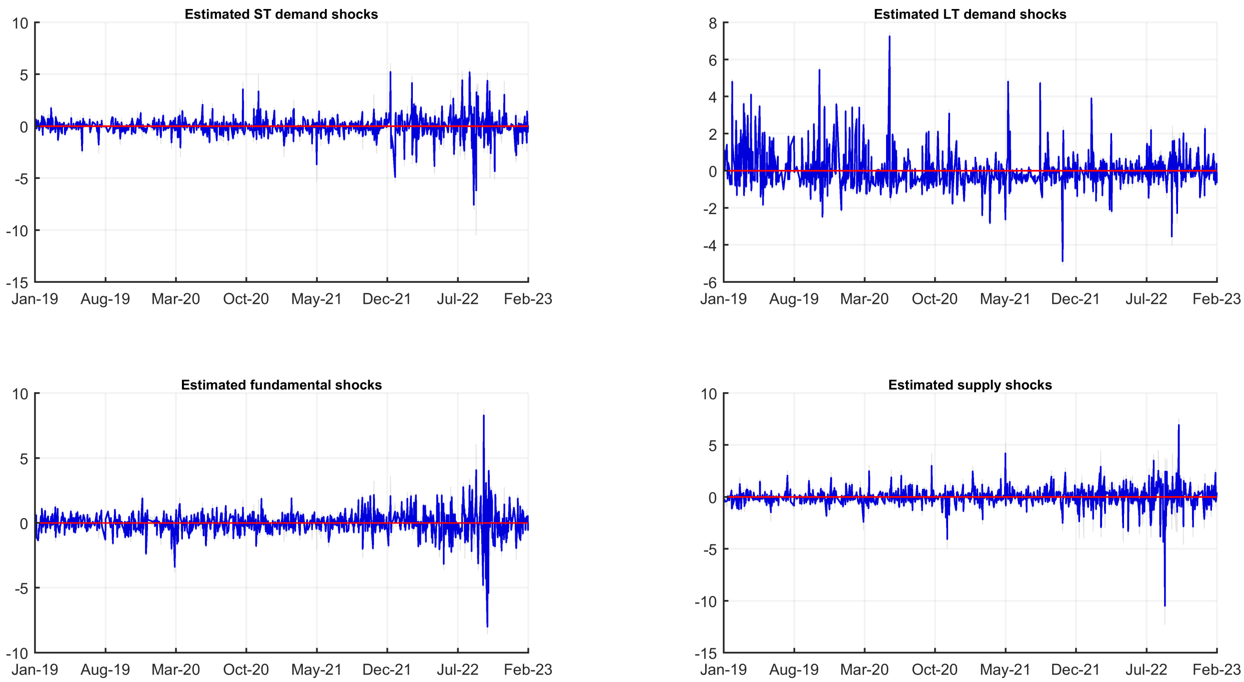
Figure E.6 Dealer banks



Source: DTCC Trade Repository OTC interest rate trade state files, from January 2019 to February 2023.

E.5 Estimated structural shocks

Figure E.7 Estimated fundamental and liquidity shocks



E.6 Cross-verification of identification strategies

Figure E.8 Distribution of ratio of variances

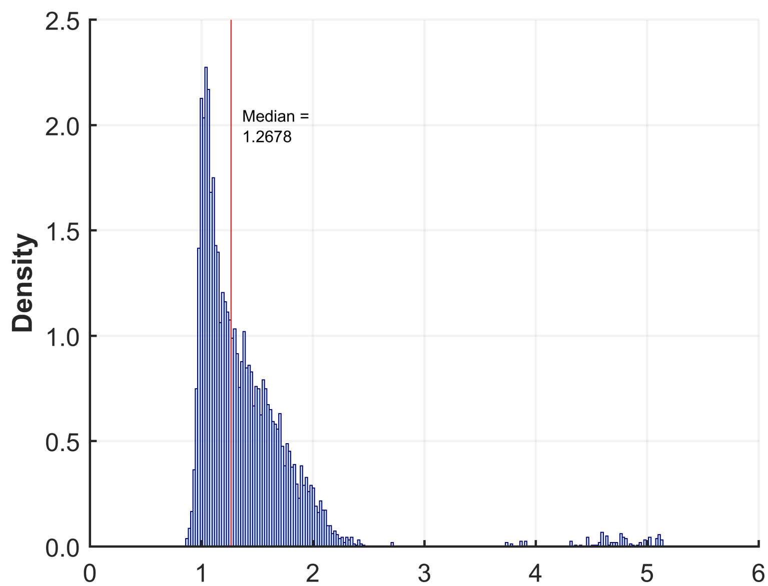
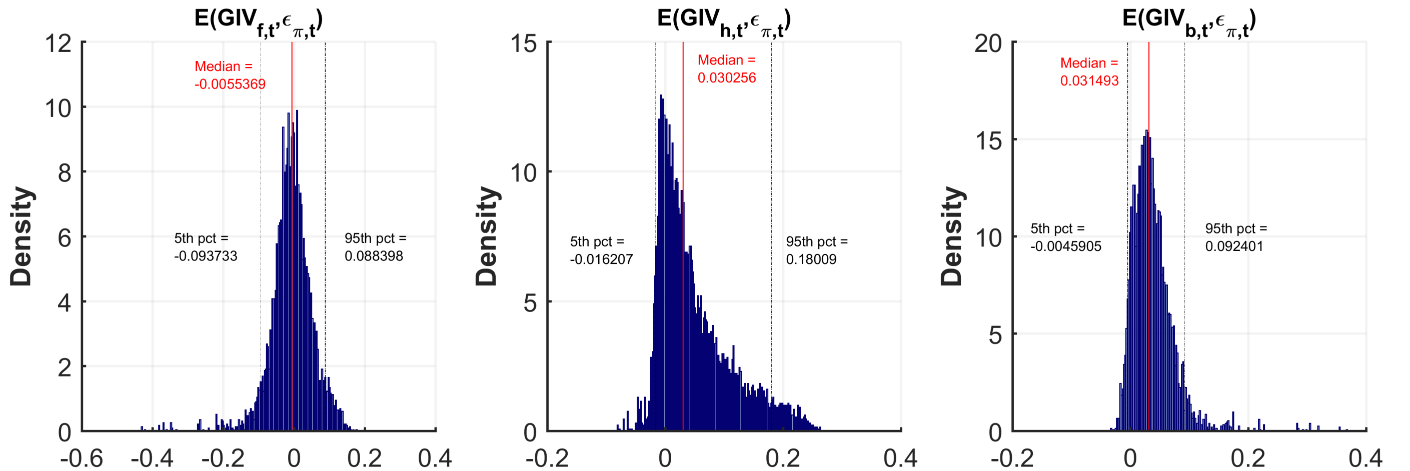


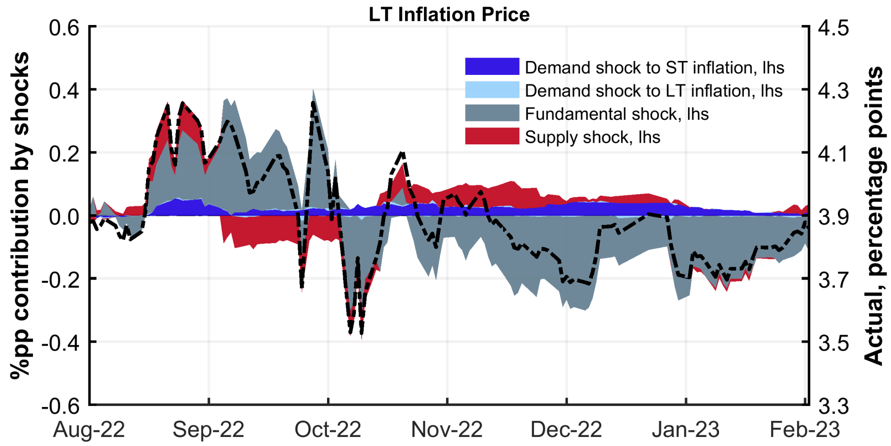
Figure E.9 Distribution of second moments



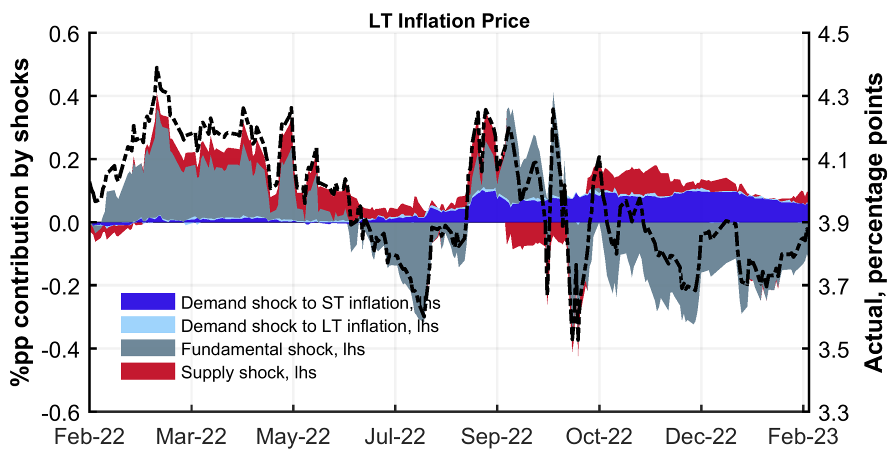
E.7 Historical Decompositions

Figure E.10 Inflation swap prices: long horizon markets

(i) From August 2022 (LDI crisis period)



(ii) From February 2022 (Ukraine war period)



(iii) From February 2020 (Pandemic period)

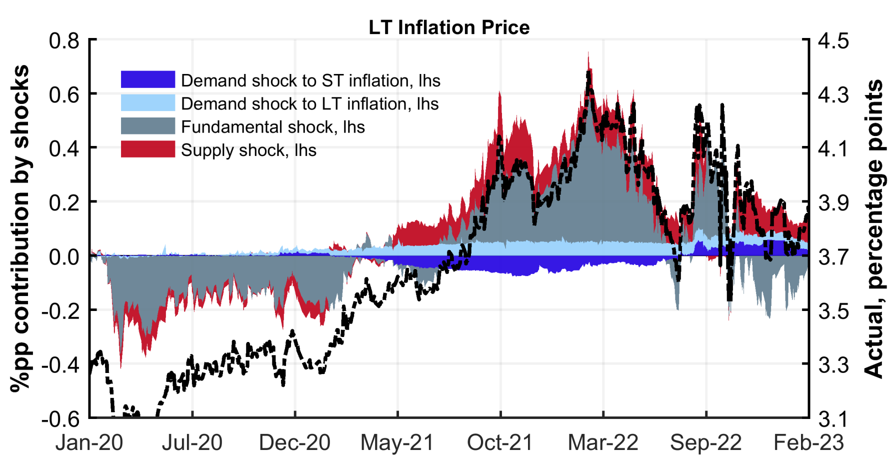
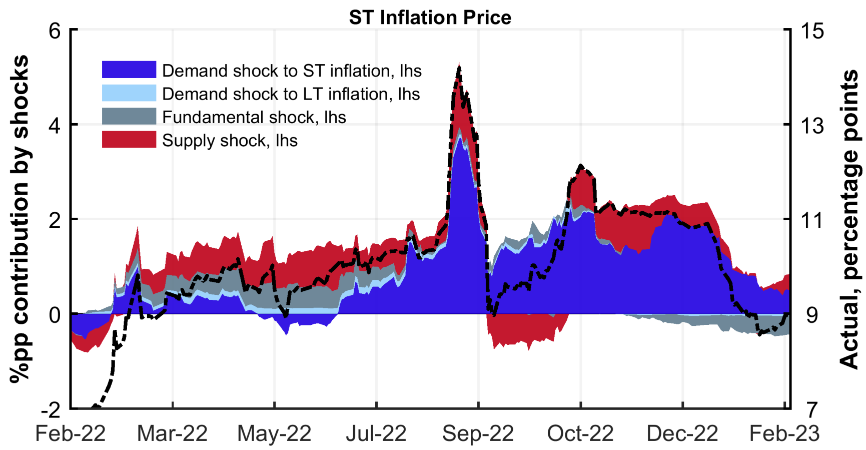
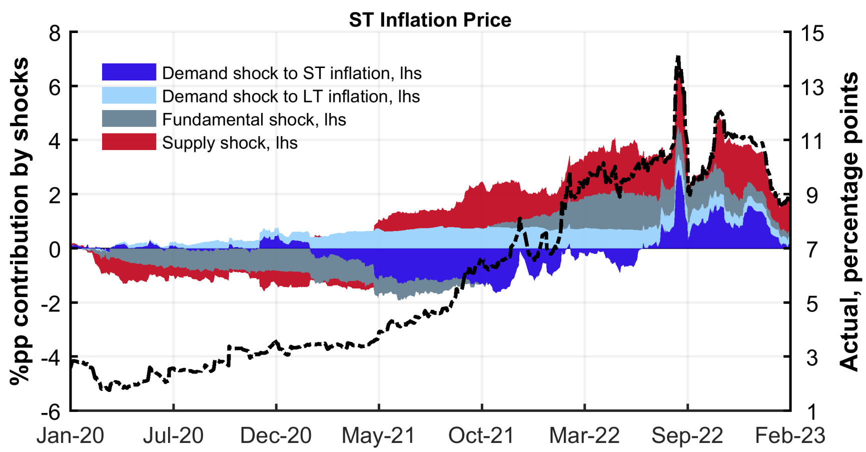


Figure E.11 Inflation swap prices: short horizon markets

(i) From February 2022 (Ukraine war period)



(ii) From February 2020 (Pandemic period)



(iii) Full sample