Journal of Financial Economics

Constrained liquidity provision in currency markets☆

Wenqian Huang [a], Angelo Ranaldo [b], Andreas Schrimpf [a,c], Fabricius Somogyi [d,*]

[a] *Bank for International Settlements, Centralbahnplatz 2, 4052, Basel, Switzerland*
[b] *University of Basel and Swiss Finance Institute, Peter Merian-Weg 6, Basel, 4052, Switzerland*
[c] *CEPR, 33 Great Sutton Street, London, EC1V 0DX, United Kingdom*
[d] *D'Amore-McKim School of Business, Northeastern University, 360 Huntington Avenue, Boston, 02115, MA, United States*

ARTICLE INFO

JEL classification:
F31
G12
G15

Keywords:
Currency markets
Dealer constraints
Market liquidity
FX
Liquidity provision

ABSTRACT

We devise a simple model of liquidity demand and supply to study dealers' liquidity provision in currency markets. Drawing on a globally representative data set of currency trading volumes, we show that at times when dealers' intermediation capacity is constrained the cost of liquidity provision increases disproportionately relative to dealer-intermediated volume. Consequently, the otherwise strong and positive relation between liquidity costs and trading volume diminishes significantly when dealers face tighter Value-at-Risk limits or higher funding costs. Using various econometric approaches, we show that this nonlinear effect of dealer constraints on market liquidity primarily stems from a reduction in the elasticity of liquidity supply, rather than changes in liquidity demand.

1. Introduction

Financial intermediaries play a crucial role in maintaining the functioning of modern financial markets. This is especially true for the foreign exchange market (FX), where dealer banks are the primary providers of market liquidity.[1] However, dealer banks' ability to provide liquidity in over-the-counter (OTC) markets heavily depends on their balance sheet capacity to absorb and fund trading positions. Constraints on dealers' intermediation capacity can in turn reduce their incentives to intermediate trades, increase liquidity costs, and generate violations of no-arbitrage conditions.[2]

Against this backdrop, the key contribution of this paper is to shed light on the link between the determinants of currency market liquidity and dealer intermediation constraints. To this end, we build a simple model of liquidity demand and supply and test its predictions by drawing on a globally representative data set on FX trading volumes. We find that the cost of providing FX liquidity increases disproportionately more relative to trading volume when dealers face tighter constraints on their intermediation capacity. More specifically, when the dealer sector is more constrained due to higher funding costs and more restrictive Value-at-Risk (VaR) constraints the otherwise strong and positive

☆ Nikolai Roussanov was the editor for this article. We thank him and two anonymous referees for their thoughtful recommendations. We are grateful to Matteo Aquilina, Briana Chang (discussant), Stijn Claessens, Stefania D'Amico (discussant), Tobias Dieler (discussant), Wenxin Du, Darrell Duffie, Ester Félez-Viñas (discussant), Semyon Malamud, Dagfinn Rime (discussant), Paul Söderlind, Frederik Simon (discussant), and Jonathan Wallen for helpful comments. We also thank seminar and conference participants at the BIS, 11th Workshop on Exchange Rates (SNB), King's College London, 11th Annual 2022 Stern/Salomon Center Microstructure Conference, 2022 SFI Research Days in Gerzensee, CLS Group, Credit Suisse, University of St.Gallen, VfS Annual Conference 2022, Boston Fed, University College Dublin, Reichman University (IDC Herzliya), Northeastern University, University of Lugano, ECB, 17th Central Bank Conference on the Microstructure of Financial Markets, 2022 New Zealand Finance Meeting, 2023 IBEFA Meeting at ASSA, QCGBF at King's Business School 2023 Annual Conference, and the 2024 European Finance Association Annual Meeting. Angelo Ranaldo acknowledges financial support from the Swiss National Science Foundation (SNSF grant 182303). All errors are our own.

[1] To be clear, we focus on the role of FX dealer banks as liquidity providers rather than cross-market arbitrageurs. This is consistent with the role that these institutions have played after the clampdown on proprietary trading in the aftermath of the Global Financial Crisis.

[2] See "Holistic Review of the March Market Turmoil", Financial Stability Board, November 2020.

Received 10 January 2024; Received in revised form 9 February 2025; Accepted 13 February 2025

Available online 24 February 2025
0304-405X/

correlation between the *cost* and the *quantity* of FX liquidity provision weakens substantially. Guided by our theoretical framework and by employing various econometric techniques, we show that this nonlinear effect of dealer constraints on market liquidity primarily stems from a reduction in the elasticity of liquidity supply, rather than changes in liquidity demand.

In the context of the FX spot market, we first identify two liquidity cost measures based on the well-known triangular no-arbitrage condition that ties together triplets of exchange rates (e.g., EURCAD, USDEUR, and USDCAD). The first measure captures **v**iolations in the **l**aw **o**f **o**ne **p**rice, which we label as VLOOP. Conceptually, VLOOP quantifies the divergence of the mid-quote prices from the triangular no-arbitrage relation. In line with the literature on intermediary asset pricing, VLOOP can be interpreted as the shadow cost of intermediary constraints.[3] The second component captures the round-trip **t**ransaction **cost** of performing a triangular arbitrage trade, which we label as TCOST. It reflects the dealer's realised compensation to take on inventory risks stemming from imbalanced customers order flows. Against this backdrop, VLOOP does not aim to measure executable arbitrage opportunities or potential arbitrage profits. This is because in an OTC market individual traders may face a different degree of trading frictions in the form of, for instance, transaction costs, price impact, and funding costs. Put differently, TCOST larger than VLOOP implies that the differentials in midquotes are within the no-arbitrage bounds set forth by bid–ask spreads.

Our model builds on the premise that tighter constraints reduce dealers' short-run flexibility to intermediate and provide liquidity to customers (see, e.g., Duffie, 2010, 2023). The dealer sector faces two types of constraints when intermediating customers' order flows: (i) debt funding costs η to finance inventory positions that stem from absorbing directional customer order flows; and (ii) VaR limits that arise from both regulatory and internal risk management practices (Adrian and Shin, 2010). The first one represents a cost factor in the "production" of market liquidity by dealers, while the second one is a hard constraint that can directly restrict dealers' intermediation capacity. In particular, VaR limits can become binding for a ω fraction of dealers. Taken together, dealers are less willing to intermediate in FX spot markets when their debt funding costs are higher and/ or VaR constraints are stricter. Thus, when dealers face tighter constraints, the liquidity supply curve becomes steeper, reflecting a decrease in the elasticity of their liquidity supply. This leads to the first empirical implication: no-arbitrage deviations (i.e., VLOOP) and transaction costs (i.e., TCOST) are positively related to measures of dealer constraints.

Going a step further, our model sheds light on how dealer constraints shape the relation between the price and the quantity of currency market liquidity. When price-sensitive customers demand more liquidity, both the price (i.e., VLOOP and TCOST) and the quantity (i.e., dealer-intermediated volume) of liquidity increase in equilibrium. When dealer constraints remain unchanged, the price and quantity increase proportionately. In contrast, when dealer constraints tighten, the price of liquidity increases disproportionately more relative to the quantity. This nonlinear effect has two main drivers. On the one hand, the more binding VaR constraints (captured by higher ω) imply that the unconstrained dealers need to accommodate a larger share of customer order flows. On the other hand, the unconstrained dealers are left to absorb the order imbalance at higher debt funding costs (captured by higher η). As a result, an increase in customers' demand for liquidity leads to a more pronounced increase in the equilibrium price compared to volume. This underpins our second, and most novel, empirical implication: while liquidity costs and dealer-intermediated volumes exhibit a positive correlation in normal times, this connection weakens significantly during times when dealer constraints are more stringent.

We test these two predictions using a unique data set on global FX spot trading activity from CLS Group. To measure dealer constraints in line with our model, we construct empirical measures capturing the two sources of dealer constraints. First, we employ dealers' debt funding costs to proxy for η and second, we use dealers' realised VaR and the number of VaR breaches in a given quarter to capture the tightness of the VaR constraint ω. We study these measures separately and, for conciseness, combined as a single metric of intermediary constraints that we dub "DCM" (referring to **D**ealer **C**onstraint **M**easure). To construct DCM, we first create (cross-sectionally averaged) time-series of debt funding costs, portfolio VaR, and the number of VaR breaches for 10 major FX dealer banks. We define the first principal component of these time-series as our DCM measure of dealer constraints.[4]

The following three core findings arise from our empirical analysis: First, the two liquidity cost measures VLOOP and TCOST tend to co-move over time, albeit their correlation is only about 54% on average. Second, when the dealer sector is largely unconstrained, the correlation between liquidity costs and dealer-intermediated volume is overall positive and ranges from 9 to 25%. This observation is consistent with dealers requiring a higher compensation when providing more immediacy to clients. Third, and most strikingly, when dealer constraints tighten, both liquidity cost measures increase disproportionately more relative to dealer-intermediated volumes. In times when the dealer sector is constrained, the conditional correlation between liquidity costs and the intermediated quantities drops by at least 50%. To establish this novel result, we estimate smooth transition regression (LSTAR) models,[5] which are well-suited for our analysis because constrained regimes are determined endogenously and may vary over time.

To broaden the scope of our proposed mechanism regarding the nonlinear effect of dealer constraints on market liquidity, we extend our empirical analysis in two ways. First, we connect our findings to Duffie et al. (2023), showing that volatility and liquidity costs in the US Treasury market co-move positively in normal times, but less so during times when dealer balance sheet utilisation is high. Our model sheds light on the origins of this positive correlation and can explain why it diminishes when dealers are more constrained. Moreover, we provide empirical evidence supporting this prediction in the context of the currency market. Second, we expand our analysis to FX forwards and swaps, where constraints on dealer intermediation capacity are even more consequential for liquidity provision than in FX spot. Specifically, we show that our mechanism, which operates through constraints on dealers' intermediation activity, has even more explanatory power when analysing market liquidity in currency forwards and swaps. Taken together, our results suggest that our economic mechanism is not only operational in the FX spot market but might also be at play in other OTC markets such as FX derivatives and US Treasuries.

Through the prism of our model, the drop in the correlation between liquidity costs and volume stems from a more inelastic (i.e., steeper) supply curve. However, when it comes to the empirical estimation, one might be concerned that our dealer constraint measure is correlated with factors affecting liquidity demand. Put differently, we can only attribute the weakening of the correlation between liquidity cost and intermediated volume in constrained periods to a drop in the elasticity

[3] In accord with this strand of literature (e.g., Adrian et al., 2014; Kisin and Manela, 2016; Duffie, 2018; Fleckenstein and Longstaff, 2018; Du et al., 2022), one may also refer to these shadow costs as "balance sheet costs" associated with FX spot liquidity provision.

[4] For robustness, we also consider other measures proposed in the related literature to capture the balance sheet capacity of financial intermediaries. Specifically, we consider the He et al. (2017) leverage ratio, credit default swap (CDS) premia (Andersen et al., 2019), and deviations from the covered interest rate parity (CIP) condition (Du et al., 2018; Rime et al., 2022).

[5] Studies using alternative forms of smooth transition regressions to perform exchange rate or carry trade predictability include, for instance, Kilian and Taylor (2003), Christiansen et al. (2011), Tenreyro and Thwaites (2016), and Jeanneret and Sokolovski (2019).

of liquidity supply once we are appropriately controlling for shifts in liquidity demand. To account for changes in liquidity demand, we rely on two approaches: First, we use a rich set of fixed effects (both time-series and cross-sectional) and control variables such as FX volatility and measures of price impact to account for observable and unobservable factors related to liquidity demand. Second, we employ a structural vector autoregression (SVAR) with sign restrictions to more explicitly disentangle liquidity demand and supply dynamics. Our setup to estimate liquidity demand and supply shocks closely follows Goldberg (2020) and Goldberg and Nozawa (2020), respectively, and employs the same set of sign restrictions. In a next step, we use both liquidity supply and demand shocks as alternative measures of tightening dealer constraints. In line with our model, it turns out that only liquidity supply (rather than demand) shocks are economically and statistically significant determinants of the variation in the correlation between liquidity costs (i.e., VLOOP and TCOST) and dealer-intermediated trading volume.

2. Related literature

Our paper contributes to three strands of literature. First, we contribute to the literature on currency market liquidity. Prior research in this field provides empirical evidence on the correlation between funding liquidity and market liquidity (Mancini et al., 2013; Karnaukh et al., 2015) but does not explore the fundamental drivers of these links. Our work goes significantly beyond these empirically focused papers by elucidating (both theoretically and empirically) the economic mechanism through which this connection arises. In particular, we show that studying the *joint behaviour of quantities and prices* allows us to better isolate the impact of dealer intermediation constraints on market liquidity conditions.

To do so, we leverage data on dealer-intermediated FX trading volumes from CLS group. The literature on trading volume is relatively scarce due to the lack of comprehensive data. Earlier research has instead focused on order flows (e.g., Evans, 2002; Evans and Lyons, 2002, 2005) primarily analysing the inter-dealer segment, which is dominated by two platforms: Reuters (e.g., Evans, 2002; Payne, 2003; Foucault et al., 2016) and EBS (e.g., Chaboud et al., 2008; Mancini et al., 2013; Chaboud et al., 2014). Other sources of FX spot volume are proprietary data sets from specific dealer banks.[6] The recent public access to CLS data has enabled researchers to study customer-dealer volume at a global scale (Hasbrouck and Levich, 2018, 2021; Cespa et al., 2021; Ranaldo and Somogyi, 2021; Ranaldo and Santucci de Magistris, 2022). We contribute to this strand of literature by investigating the impact of dealer constraints on both the cost and quantity dimensions of FX liquidity. We primarily focus on currency markets, but we also examine the similarities between the economic mechanism described in our model and that in other OTC markets, such as fixed income markets (Duffie, 2023; Duffie et al., 2023). Consequently, our findings shed light on how the functionality of OTC markets, more generally, is shaped by dealer intermediation constraints.

Second, our work relates to the broad literature that emphasises the role of intermediary frictions in affecting asset prices and financial market conditions.[7] Our main contribution is to show in depth how constrained dealers charge higher liquidity costs and decrease their elasticity of liquidity provision in the FX spot market. This finding is in accord with the evidence documented for other markets, in particular, stocks (Comerton-Forde et al., 2010; Çötelioğlu et al., 2020) and corporate bonds (Bao et al., 2018). Our research expands the literature by conceptualising and empirically examining how constraints on dealers, like debt financing costs and VaR limits, affect liquidity costs and trading volume in currency markets. Our results remain consistent across a suite of measures that are coherently tied to our theoretical framework, as well as when utilising broader proxies for dealers' balance sheet capacity such as the equity capital ratio of financial intermediaries (He et al., 2017), credit default swap spreads (Andersen et al., 2019), and deviations from CIP (Du et al., 2018, 2022; Rime et al., 2022; Du et al., 2023), respectively. Our findings are also in line with Nagel (2012) who shows that market makers' liquidity supply is increasing in their intermediation capacity but decreasing in the level of risk. Moreover, our paper corroborates the idea that market-wide liquidity conditions depend on intermediary constraints (e.g., Adrian and Shin, 2010) and that intermediary leverage and banks' risk management practices (e.g., following Value-at-Risk methodologies) tend to be procyclical (Adrian and Shin, 2013). Lastly, our findings suggesting that dealers' balance sheet space affects both the cost and quantity of liquidity provision are consistent with slow-moving intermediary capital being a key factor behind distortions in asset pricing relations (Duffie, 2010).

Finally, we add to the literature on limits to arbitrage along two dimensions. First, while prior research has mostly focused on constrained *arbitrageurs* (e.g., Shleifer and Vishny, 1997, Gromb and Vayanos, 2002, Hombert and Thesmar, 2014 and more recently Du et al., 2022 and Siriwardane et al., 2025), our main angle is to study constrained *dealers*. This emphasis on constrained dealers, rather than cross-market arbitrageurs (e.g., principal trading firms) is driven by several factors. To begin with, regulatory changes since the Global Financial Crisis period have significantly influenced the role of dealer banks. In particular, these regulations have incentivised banks to shift their business models from proprietary trading to market making. In addition, dealers' intermediation capacity is also affected by their risk management practices, such as VaR constraints, and their funding costs. Second, we propose to draw on the triangular no-arbitrage identity to derive two liquidity cost components with an economically meaningful interpretation. Thus, our key contribution is to elucidate the relation between liquidity costs and volumes using arbitrage conditions and to show how this relation critically depends on the intermediation capacity of dealers. In addition, a large body of prior research has studied limits to arbitrage in equities (see Gromb and Vayanos, 2010). However, many of the frictions considered in that literature, such as short sale constraints (e.g., Chu et al., 2020), do not apply to currency markets. Related to the stock market literature, recent studies document widespread mispricings in stressed times (Pasquariello, 2014), commonality in arbitrage deviations (e.g., Rösch et al., 2016; Du et al., 2022), and limits to arbitrage impacting market liquidity (Rösch, 2021). We add to this branch of the literature by identifying constrained dealers as the main driving force behind such commonalities and by showing that dealer constraints have a nonlinear effect on currency market liquidity.

3. A simple model of constrained liquidity supply

1. Both liquidity cost measures (i.e., VLOOP and TCOST) increase when FX dealers are more constrained in their intermediation capacity (see Proposition 1 below).
2. When dealer constraints tighten, the liquidity supply curve shifts inward and thereby increases the cost of liquidity provision but decreases dealer-intermediated volume (relative to the counterfactual). Thus, the correlation between liquidity costs and dealer-intermediated volume drops with tighter dealer constraints (see Proposition 2 below).

[6] See, for instance, Bjønnes and Rime (2005), Menkhoff et al. (2016), Gallien et al. (2018).

[7] See, for example, Gârleanu and Pedersen (2011), He and Krishnamurthy (2011), He and Krishnamurthy (2013), Adrian and Boyarchenko (2012), Adrian et al. (2014), He et al. (2017), Chen et al. (2018), Gospodinov and Robotti (2021), Baron and Muir (2021), Haddad and Muir (2021), Kargar (2021), and He et al. (2022).

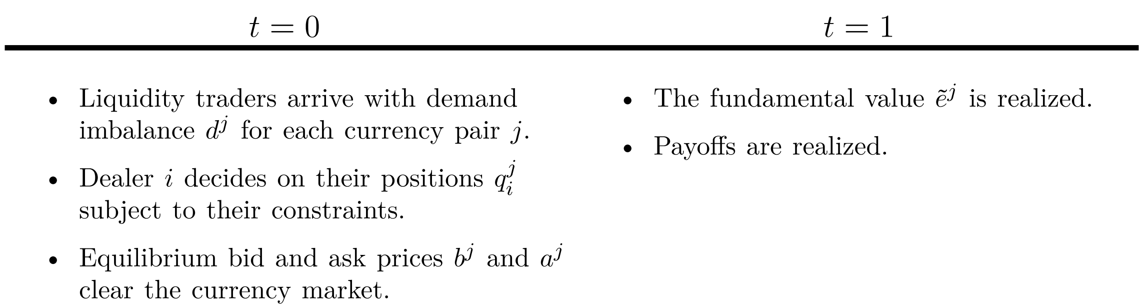

Fig. 1. Timeline.

Liquidity cost measures. Consider a trader exchanging one euro (EUR) to some amount of US dollar (USD), exchanging the amount of US dollar to some amount of Canadian dollar (CAD) and exchanging back the amount of Canadian dollar to euro instantaneously. Let m^j denote the midquotes of the three currency pairs, where $j \in \{x, y, z\}$, $x = USDEUR$, $y = EURCAD$, and $z = USDCAD$. The trader has identified a violation of the law of one price (VLOOP) if $m^z/(m^x m^y)$ is different from unity. More formally,

$$\text{VLOOP} = \frac{m^z}{m^x m^y}, \tag{1}$$

which is dimensionless with respect to the choice of base currency. Note that dimensionless here refers to the economic magnitude, rather than the sign, of VLOOP. Regardless of our choice of base and quote currency for pairs x, y, and z, the absolute value of VLOOP will always be the same. The only requirement is that the denominator is a synthetic replication of the currency pair in the numerator (or the other way around).

Clearly, such law of one price deviations are not necessarily profitable arbitrage opportunities due to the presence of transaction costs. We denote the bid–ask spread as $s^j = a^j - b^j$, where a^j and b^j are the ask and the bid price of currency pair j. Replacing the midquotes with the bid and ask prices, the pay-off from a triangular arbitrage trade is $b^z/(a^x a^y)$ and can be decomposed into VLOOP and round-trip transaction costs (TCOST)[8]:

$$\frac{b^z}{a^x a^y} = \frac{m^z\left(1 - \frac{s^z}{2m^z}\right)}{m^x\left(1 + \frac{s^x}{2m^x}\right) m^y\left(1 + \frac{s^y}{2m^y}\right)} = \underbrace{\frac{m^z}{m^x m^y}}_{\text{VLOOP}} \Bigg/ \underbrace{\frac{\left(1 + \frac{s^x}{2m^x}\right) \times \left(1 + \frac{s^y}{2m^y}\right)}{\left(1 - \frac{s^z}{2m^z}\right)}}_{\text{TCOST}}. \tag{2}$$

In equilibrium, VLOOP and TCOST are determined by the demand for, and supply of, liquidity. At t=0, liquidity traders arrive with demand imbalance d^j. They trade with dealers at price p^j, which could be either a^j or b^j depending on the direction of their trade. The fundamental value of currency pair j is stochastic, and denoted as $\tilde{e}^j$, with mean e^j. The three fundamental values are intimately linked via $e^z = e^x e^y$. We assume that the three currency pairs are i.i.d. and have the same volatility denoted as σ. At $t = 1$, the uncertainty is resolved and traders receive the fundamental value of each currency pair. Fig. 1 summarises the timeline of the model.

Traders. We model liquidity demand following the classic market microstructure literature (see, e.g., Grossman and Miller, 1988; Hendershott and Menkveld, 2014). Liquidity traders are price sensitive and arrive at $t = 0$. Their aggregate liquidity demand decreases in the bid–ask spread quoted by the dealers. Furthermore, the demand is higher when the currency pairs are more volatile, reflecting higher disagreement about fundamental values and associated portfolio rebalancing. Specifically, the demand for currency pair j is given by $\sigma(1 - s^j)$, which is increasing in volatility σ and decreasing in the bid–ask spread s^j.

Trading demand is imbalanced across the three currency pairs due to diverging liquidity needs among traders. For simplicity, we assume that a $\pi > 1/2$ fraction of traders in currency pair x are buyers and the rest are sellers. Conversely, for currency pair y, a $(1 - \pi)$ fraction of traders are buyers and the rest are sellers. For currency pair z, half of the traders are buyers, whereas the other half are sellers. Thus, traders impose net buying pressure $(2\pi - 1)$ in currency pair x and net selling pressure $(1 - 2\pi)$ in pair y. The net buying pressure is simply the buy orders minus the sell orders. Hence, for currency pair x, the net buying pressure is $\pi - (1 - \pi) = 2\pi - 1 > 0$, and it is $(1 - \pi) - \pi = 1 - 2\pi < 0$ for currency pair y. Thus, the aggregate (net) trading demand of the traders is to swap EUR for USD (i.e., being long currency pair $x = USDEUR$) and to swap EUR for CAD (i.e., being short currency pair $y = EURCAD$). As a result, the traders' demand imbalance that needs to be absorbed by dealers in pair j is given as follows[9]:

$$d^x = \sigma(1 - s^x)(2\pi - 1), \tag{3}$$

$$d^y = \sigma(1 - s^y)(1 - 2\pi), \tag{4}$$

$$d^z = 0. \tag{5}$$

Dealers. There is a unit mass of competitive dealers that intermediate buy and sell orders in the currency market (see Foucault et al., 2013, Sec. 3.5). The dealers start with zero inventory and are subject to Value-at-Risk (VaR) constraints that might be due to regulatory and/ or internal risk management practices. A proportion ω of the dealers have tight VaR constraints with low thresholds T_L and the rest (i.e., $1 - \omega$) have loose VaR constraints with high thresholds T_H. These differences between the VaR constraints reflect the differences in dealers' balance sheet capacity. The VaR of a dealer $i \in \{L, H\}$ is $VaR_i^j \equiv \sigma \times q_i^j$, where q_i^j denotes the dealer's net position in a given currency pair j. Note that because dealers start with zero inventory there are no benefits of netting across currencies.[10] Specifically, $q_i^j > 0$ indicates

[8] See the Online Appendix for numerical examples.

[9] While VLOOP captures violations of the law of one price, it is important to note that these violations do not directly imply the presence of profitable triangular arbitrage opportunities. This is especially the case when transaction costs are appropriately taken into account. The empirical evidence presented in Section 5 shows that *TCOST* is typically larger than VLOOP during our sample period. Thus, our model focuses on the situation where *TCOST* is larger than VLOOP. Moreover, to keep the model concise, we have chosen not to introduce cross-market arbitrageurs in the model. This is in line with how dealer banks operate in today's financial markets, that is, as liquidity providers rather than arbitrageurs trading on their own account (see CGFS, 2014; Lu and Wallen, 2024).

[10] As a result of this assumption, the analysis is the same at both the currency-level and the currency-pair-level. To lighten notations, we choose to model dealers' optimisation problem at the currency-pair-level, instead of at the currency-level. This modelling choice is consistent with common risk management practices incentivising dealers to keep the order book flat across

that the dealer holds a long position in the base currency and a short position in the quote currency. Hence, the VaR constraint is given by $\sigma \times q_i^j \leq T_i$ (see,e.g., Duffie and Pan, 1997; Adrian and Shin, 2013).

Dealer i finances their net position q_i^j in each of the three currency pairs by issuing debt (e.g., Scott, 1976; van Binsbergen et al., 2010). Specifically, the dealer faces a trade-off between their convex debt funding cost $\eta(q_i^j)^2$ and the spread between the prices they quote and the fundamental value of each currency.[11] Thus, the utility of dealer i is given as follows:

$$U_i^D = \underbrace{E\Big((p^x - \tilde{e}^x)q_i^x + (p^y - \tilde{e}^y)q_i^y + (p^z - \tilde{e}^z)q_i^z\Big)}_{\text{Gain from trade}} - \underbrace{\frac{\eta}{2}\Big((q_i^x)^2 + (q_i^y)^2 + (q_i^z)^2\Big)}_{\text{Debt funding cost}} \tag{6}$$

$$\text{subject to:} \quad \underbrace{\sigma \times q_i^x \leq T_i, \quad \sigma \times q_i^y \leq T_i, \quad \sigma \times q_i^z \leq T_i,}_{\text{VaR constraints}}$$
$$\text{where } i \in \{L, H\}. \tag{7}$$

All dealers are competitive. At $t = 0$, they take prices p^j as given and choose their net positions q_i^j (i.e., the quantity they are intermediating) subject to their VaR constraint. The sign of q_i^j is determined by the direction of customer flows. Therefore, the supply function of a dealer with a nonbinding VaR limit is simply pinned down by the first order conditions:

$$\frac{\partial U_i^D}{\partial q_i^j} = 0 = \begin{cases} a_i^j - e^j - \underbrace{\eta |q_i^j|}_{\text{Debt funding cost}}, & \text{if } q_i^j > 0, \\ b_i^j - e^j + \underbrace{\eta |q_i^j|}_{\text{Debt funding cost}}, & \text{if } q_i^j < 0 \end{cases} \tag{8}$$

The first order conditions in Eq. (8) suggest that there are two components in the dealer's supply function. The first one is related to the marginal value of buying and selling and reflects the spread between the quoted prices and the fundamental value (i.e., $a_i^j - e^j$ and $b_i^j - e^j$). The second component $\eta|q_i^j|$ is the debt funding cost, which depends on both the size and the direction of the incoming customer order flow. If the VaR constraint is binding, dealer i's net position is equal to T_i/σ. Thus, the dealer's net position (induced by customers' order flows) for currency pair j is bounded as follows:

$$q_i^j = \min\left\{\frac{p_i^j - e^j}{\eta}, \frac{T_i}{\sigma}\right\}. \tag{9}$$

Market clearing. At $t = 0$, traders' demand must be equal to dealers' liquidity supply:

$$d^j = \int_i q_i^j. \tag{10}$$

Equilibrium outcomes. In our context, the parameter space of interest is the situation where the dealer sector is only partially constrained, that is, when a nonzero proportion ω of dealers (with T_L) face binding constraints while the rest (i.e., $1 - \omega$) are not constrained by the VaR thresholds. In the following analysis, we suppress the subscript of H and L when discussing the VaR thresholds and use T to denote T_L. In this case, the dealer sector faces two sources of constraints: one stemming from the dealers' debt funding cost, and the other from their VaR limits. The proposition below outlines the impact of the two sources of constraints on liquidity costs:

Proposition 1. *Both VLOOP and TCOST are higher conditional on the dealer sector being more constrained, which corresponds to periods when*

i) dealers' debt funding costs are higher (i.e., higher η) and/ or;
ii) VaR limits are binding for a larger share of dealers (i.e., higher ω).

In addition, VLOOP and TCOST also increase in currency volatility (i.e., σ).

We delegate the proofs of the model to Appendix A and instead focus on the economic intuition here. Both VLOOP and TCOST increase in the bid–ask spread. It is evident from Eq. (8) that higher funding costs η lead to a larger balance sheet cost for a given order size. Thus, the spread increases in η, and so do VLOOP and TCOST.

A higher ω indicates that a larger proportion of dealers is constrained by the VaR limit, reducing the pool of unconstrained dealers available to balance the order flow. As a consequence, a smaller group of dealers is left to absorb the order imbalance, which intuitively leads to an increase in spreads and, as a result, higher VLOOP and TCOST. A surge in volatility σ affects VLOOP and TCOST along two channels. On the one hand, the demand for liquidity and hence, the order flow imbalance increases in volatility and thereby scales up the bid–ask spread. On the other hand, when VaR limits become more binding, the remainder of unconstrained dealers is only willing to absorb a larger amount of order imbalance if they are sufficiently compensated by a larger spread. Thus, both channels scale up the spread and hence, increase VLOOP as well as TCOST.

Liquidity supply and demand. Next, we examine the supply and demand curves of liquidity in the spot FX market. We use currency pair x with unbalanced order flows as an example. For simplicity, we suppress the superscript x for the rest of this section. First, we rewrite Eq. (3) such that the price of liquidity p^D (i.e., the bid–ask spread s) is a function of the demanded quantity q^D (i.e., net buying pressure d that needs to be warehoused by dealers):

$$p^D = 1 - \frac{q^D}{\sigma(2\pi - 1)}. \tag{11}$$

The demand curve is downward-sloping because liquidity traders are price sensitive. Moreover, the slope of the demand curve steepens with volatility σ since liquidity demand increases in volatility. The demand slope is also steeper when customer order flows are more unbalanced (i.e., when π is higher).

On the supply side, constrained dealers can only provide a fixed quantity (i.e., $\omega \times T/\sigma$) as dictated by the binding VaR constraint. Therefore, market clearing requires that the remaining trading demand is intermediated by dealers that are unconstrained. Thus, in equilibrium, the supply curve is determined by the first order condition Eq. (8) of the unconstrained dealers' maximisation problem. The price of liquidity p^S (i.e., the bid–ask spread s) is a function of the total supplied quantity q^S (i.e., the net buying quantity from the traders' perspective) minus the net position taken by the constrained dealers (i.e., $\omega \times T/\sigma$), that is,

$$p^S = \frac{\eta}{1-\omega}\left(q^S - \frac{\omega T}{\sigma}\right). \tag{12}$$

The supply curve is upward-sloping and its slope increases with debt financing cost η because the unconstrained dealers require a higher compensation for each additional marginal unit of currency that they intermediate. Furthermore, the slope is steeper when ω is higher, because a smaller amount of unconstrained dealers need to absorb the entire customer order imbalance. Overall, the dealer sector is more constrained when η and ω are large. Collectively, these variables can be

trading periods (Evans and Lyons, 2002). Hence, the net positions outlined above correspond to trading demands reflecting the need to exchange one currency for another.

[11] To simplify the notation, we assume that both the VaR threshold T_i and debt funding cost η are the same across the three currency pairs. Relaxing this constraint will not qualitatively affect any of our main results.

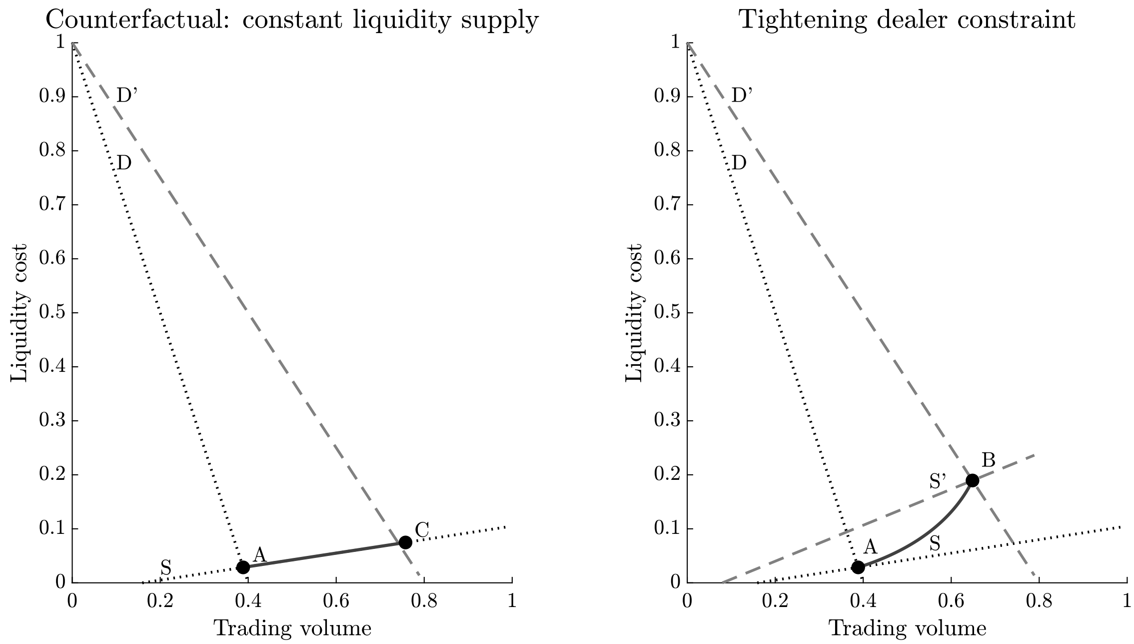

Fig. 2. Liquidity supply and demand
Note: This figure plots liquidity costs against dealer-intermediated volumes. The baseline parameters are $\pi = 0.7$, $e^x = 1.2$, $e^y = 1.1$, $e^z = 1.32$, where π denotes the fraction of traders that are buyers (sellers) in currency pair x (y), e^x, e^y, and e^z denote the fundamental values of currency pairs x, y, and z, respectively. When the dealer is unconstrained (i.e., S), $\eta = 0.05$ and $\omega = 0.2$, whereas $\eta = 0.1$ and $\omega = 0.4$ when the dealer is constrained (i.e., S'). The solid lines indicate the equilibrium outcomes when varying the volatility of the exchange rates σ from 0.5 to 0.7. The parameter space for the left and right panel are identical. The only difference is the assumption that the liquidity supply curve does not shift in the counterfactual, that is, holding σ, η, and ω constant (left panel). Both liquidity costs and dealer-intermediated volume are normalised to unity.

interpreted as a measure of the (shadow) cost of providing immediacy in the currency market.

Fig. 2 illustrates how a more constrained dealer sector affects both trading volume and liquidity costs in equilibrium. The left panel in Fig. 2 illustrates the counterfactual in which liquidity supply is kept constant while liquidity demand shifts outwards (due to an increase in volatility σ). As a result, liquidity costs and trading volume increase at the same rate (equal to the slope of the supply curve) from point A to C. The right panel in Fig. 2 in turn illustrates supply and demand in the constrained dealer sector. Specifically, it shows how the outward shift in liquidity demand is counterbalanced by the inward shift in liquidity supply due to a higher η and ω. Importantly, the inward shift in liquidity supply leads to an increase in liquidity cost but *lower* volume compared to the counterfactual shown in the left panel. Consequently, the co-movement between liquidity cost and volume *weakens* as the dealer sector becomes more constrained. Proposition 2 summarises these results.

Proposition 2. *Higher debt funding costs and/ or more stringent VaR limits give rise to a more constrained dealer sector and cause an inward shift in liquidity supply. As a result, the cost of liquidity provision increases, while dealer-intermediated volume falls relative to the counterfactual. Consequently, as dealer constraints tighten, the co-movement between liquidity costs and volume weakens.*

Following Proposition 2, the correlation between liquidity costs and the volume intermediated by dealers falls with the degree of dealer constrainedness. The economic intuition is that the increase in liquidity costs outpaces the increase in trading volume because dealers are constrained by either the debt funding cost and/ or by the binding VaR constraint.

Equipped with these theoretical propositions, the goal of our empirical analysis in the subsequent sections of the paper is threefold: First, to provide a thorough empirical examination of these theoretical predictions. Second, to document how the correlation between liquidity cost and volume changes conditional on dealer constraints tightening. Third, to use various econometric techniques (i.e., panel regressions with fixed effects and structural VaRs) to tease out to what extent the variation in this conditional correlation is driven by changes in the elasticity of liquidity supply.

4. Measuring the cost of FX spot liquidity provision

4.1. Data sources

Our empirical analysis employs trade and quote data from two main sources. The FX spot volume data come directly from CLS Group (CLS), which is the world's largest payment-vs-payment settlement system.[12] The data set features trading activity that passes through the main FX intermediaries (i.e., dealer banks) that are either trading with each other or with their customers, reflecting the market structure of FX markets where dealers play a central role (Schrimpf and Sushko, 2019). Given this market structure, the data set does not include any direct trading activity between two customers (e.g., corporates and funds).[13]

We obtain CLS data directly from CLS Group. The same data set has been used in prior research, among others, by Cespa et al.

[12] At settlement, CLS mitigates principal and operational risk by settling both sides of the trade at once. The comprehensiveness of CLS' coverage of global FX transactions is unmatched, as it handles more than half of global FX trading volumes. Cespa et al. (2021) show that there is an almost perfect overlap between the share of volume across currency pairs in the BIS Triennial Surveys and the CLS data.

[13] On electronic trading platforms, two non-dealer participants can trade with each other via prime brokerage. These trades – mostly between proprietary trading firms or hedge funds – will be classified as dealer-to-dealer trades in our sample. While this activity can be sizeable (Schrimpf and Sushko, 2019), it typically entails intraday trades that do not significantly affect end-of-day order imbalances (Huang et al., 2023).

(2021), Ranaldo and Somogyi (2021), and Ranaldo and Santucci de Magistris (2022). Note that Hasbrouck and Levich (2018, 2021) also analyse CLS data but using proprietary and transaction-level data. The aforementioned authors have also comprehensively described the data. The CLS volume data are available to us at the hourly frequency. The sample period spans from November 2011 to September 2022 and includes data for 18 major currencies and 33 currency pairs.

For testing the predictions of the model, we also need to construct empirical measures capturing the cost of dealers' liquidity provision. We derive these measures from the triangular no-arbitrage relation that ties together a triplet of currency pairs involving one non-dollar currency pair (e.g., AUDJPY) and two dollar legs (i.e., USDAUD and USDJPY). Based on the 33 currency pairs in the CLS data, we are able to construct a maximum of 15 such triplets of currency pairs,[14] involving 15 non-dollar currency pairs (i.e., AUDJPY, AUDNZD, CADJPY, EURAUD, EURCAD, EURCHF, EURDKK, EURGBP, EURJPY, EURNOK, EURSEK, GBPAUD, GBPCAD, GBPCHF, and GBPJPY) and 10 dollar pairs (i.e., USDAUD, USDCAD, USDCHF, USDDKK, USDEUR, USDGBP, USDJPY, USDNOK, USDNZD, and USDSEK). These 25 currency pairs cover at least 75% of global FX spot trading volume according to the Bank for International Settlements (see "Triennial central bank survey — global foreign exchange market turnover in 2022", September 2022).

Next, we combine the hourly FX volume data with intraday spot bid and ask quotes from Olsen, a well-known provider of high-frequency data. Olsen compiles historical tick-by-tick data from various electronic trading platforms, both from the inter-dealer and dealer-customer segments. A key advantage of the Olsen data are that it accurately matches both the cross-sectional and also the time-series dimension of the CLS volume data. A possible downside is that the bid and ask quotes are indicative and hence, do not correspond to actually executable prices. This means that choosing between Olsen data and inter-dealer prices (e.g., from EBS or Reuters) requires balancing the trade-off between comprehensive coverage (across currency pairs and time periods) and the tradeability of the quotes. We are convinced that for our empirical analysis the advantage of having a sufficiently large sample across both the time-series and cross-sectional dimension compensates for the indicative nature of the quotes. This is because our primary goal is not to pinpoint any specific arbitrage opportunities on a particular trading platform, but to develop a measure of trading costs that accurately represents the global currency market.[15]

4.2. Key variables

Liquidity cost measures. Our model implies two measures of liquidity costs in the FX spot market: (i) violations of the law of one price (VLOOP), and (ii) round-trip transaction costs (TCOST). VLOOP captures the price dislocations for two assets or trading positions with the same intrinsic value, while TCOST refers to the round-trip trading cost to take advantage of such dislocations. The VLOOP component of the triangular arbitrage trade is computed with midquote prices and reflects the difference between exchanging a currency pair directly or indirectly, that is, by using another currency (e.g., the US dollar) as a vehicle. The TCOST part is computed from the bid and ask quotes (depending on the base and quote currency) involved in the currency pair triplet. Clearly, whether VLOOP constitutes an actual arbitrage opportunity will depend on the degree of trading frictions or limits to arbitrage faced by an individual trader. TCOST captures some of these trading frictions in the form of bid–ask spreads.

We compute VLOOP and TCOST for $k = 1, 2, \ldots, 15$ triplets of currency pairs (see the Online Appendix for further details). A triplet is defined as one non-dollar pair (e.g., EURCAD) plus the two USD legs (e.g., USDEUR and USDCAD). At every point in time we take the perspective of an arbitrageur by, first, identifying the seemingly profitable direction of the trade (i.e., by conditioning on VLOOP being positive), and second, by computing the associated trading cost TCOST. To mitigate the effect of outliers, we remove observations at the top and bottom 1.5 percentiles of the hourly VLOOP and TCOST series. For our main analysis we rely on daily measures of VLOOP and TCOST that we obtain by summing up hourly observations for each day.

Fig. 3 shows the time-series and cross-sectional variation of hourly no-arbitrage violations VLOOP (left y-axis) and round-trip transaction costs TCOST (right y-axis), respectively. Economically, a higher reading of VLOOP coincides with a larger shadow cost of intermediary constraints, whereas TCOST captures the realised compensations for providing immediacy. Both measures of dealers' liquidity costs exhibit intuitive properties in the sense that they surge during market stress and mean-revert during calm periods. The large spike during the Covid-19 market turmoil in March and April 2020 is particularly well pronounced across all 15 triplets of currency pairs and is indicative of the global nature of the stress. The correlation of VLOOP and TCOST is positive for the entire cross-section and ranges from 15–40%. We interpret this as evidence of commonality in no-arbitrage violations (e.g., Rösch et al., 2016; Du et al., 2022) and market liquidity in the broader sense (Rösch, 2021).

Summary statistics. Table 1 reports the time-series average of hourly no-arbitrage deviations (VLOOP) and round-trip trading costs (TCOST). In addition, it tabulates hourly averages of direct trading volume in non-dollar currency pairs (e.g., AUDJPY) and synthetic trading volume in dollar currency pairs. By "synthetic" we refer to the sum of trading volume in two dollar pairs (e.g., USDAUD and USDJPY) within a triplet of currency pairs. Each row corresponds to one currency pair triplet, which we abbreviate as, for instance, AUD-USD-JPY.

This simple summary table conveys three main insights: First, deviations from fundamentals (as measured by VLOOP) are an order of magnitude smaller than round-trip transaction costs (as measured by TCOST). We interpret this result as suggestive evidence that dealers recharge their intermediation costs on the bid and ask prices offered to their customers. Another implication is that seemingly profitable violations of triangular no-arbitrage are most of the time not exploitable by the average trader as transaction costs are prohibitively high (i.e., there is no free lunch) as well as FX quantity conventions on major trading platforms (i.e., there are minimum required trading amounts in each currency pair). Second, trading volume in non-dollar currency pairs is considerably smaller relative to the synthetic volume in dollar pairs. This is essentially the case for all 15 currency pair triplets but the effect is less pronounced for those involving the NOK and SEK, where the euro crosses play a bigger role. Finally, the synthetic relative bid–ask spread is somewhat larger than the direct spread in non-dollar currency pairs.

Measures of dealer constraints. Our model suggests two main sources of dealer constraints that can have a bearing on liquidity provision: dealers' debt funding costs (η) and VaR constraints (ω). In our empirical implementation we seek to capture these intermediary constraints in a single metric that we dub "DCM", which stands for dealer constraint measure. We construct this measure in two steps.

As a first step, we create three time-series based on cross-sectional averages of the top 10 FX dealer banks' (i) debt funding costs (daily), (ii) realised VaR measure of their overall trading book (quarterly), and (iii) the number of VaR breaches (quarterly).[16] The first measure captures

[14] Note that to maintain a balanced panel, we also remove all currency pairs involving the Hungarian forint (HUF), which enters the data set later, on 7 November 2015.

[15] In the Online Appendix we conduct a comprehensive comparison of Olsen and EBS data for the full-year of 2016. The key takeaway is that EBS and Olsen quotes (and also VLOOP and TCOST) are positively correlated and the mean absolute difference is especially low for currency pairs that are mainly traded on EBS.

[16] We compute cross-sectional averages because the CLS volume data do not contain any information about traders' identities. See the Online Appendix for details on how we retrieve and compute each of these variables.

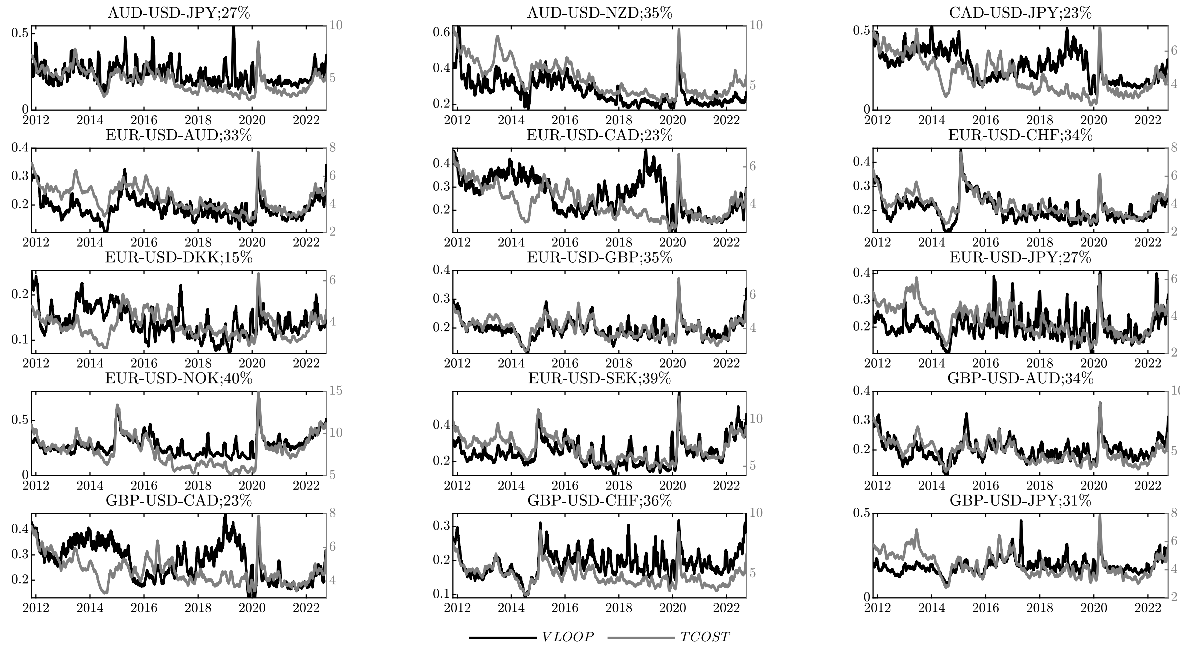

Fig. 3. No-arbitrage violations and round-trip transaction costs
Note: This figure plots the 22-day moving averages of hourly triangular no-arbitrage deviations VLOOP (left y-axis) and round-trip trading costs TCOST (right y-axis), respectively, for 15 triplets of currency pairs. Both variables are measured in basis points. The numbers in the titles refer to the correlation coefficient of VLOOP and TCOST. The sample covers the period from 1 November 2011 to 30 September 2022.

Table 1
Summary statistics.

	Liquidity cost in bps		Volume in $bn		Bid–ask spread in bps		*Volatility in bps*
	VLOOP	TCOST	Direct	Synthetic	Direct	Synthetic	*Direct*
AUD-USD-JPY	0.23	4.73	0.18	4.93	4.00	5.71	14.13
AUD-USD-NZD	0.27	5.61	0.09	1.93	4.17	7.25	8.96
CAD-USD-JPY	0.28	4.50	0.03	5.27	4.10	5.06	12.36
EUR-USD-AUD	0.19	4.40	0.13	7.47	3.40	5.52	11.30
EUR-USD-CAD	0.27	4.10	0.08	7.81	3.37	4.88	9.88
EUR-USD-CHF	0.21	3.91	0.36	6.56	2.60	5.27	6.49
EUR-USD-DKK	0.14	3.84	0.09	5.98	2.45	5.27	1.79
EUR-USD-GBP	0.20	4.01	0.59	7.94	3.11	4.89	9.33
EUR-USD-JPY	0.20	3.78	0.61	9.35	3.03	4.69	11.06
EUR-USD-NOK	0.27	7.92	0.24	6.06	6.43	9.55	11.66
EUR-USD-SEK	0.25	6.91	0.27	6.08	5.41	8.45	9.39
GBP-USD-AUD	0.20	4.95	0.04	3.51	4.03	5.91	12.14
GBP-USD-CAD	0.27	4.58	0.03	3.85	3.84	5.27	10.59
GBP-USD-CHF	0.19	4.84	0.03	2.60	3.99	5.66	10.55
GBP-USD-JPY	0.19	4.35	0.20	5.39	3.71	5.08	12.47

Note: This table reports the time-series average of hourly triangular no-arbitrage deviations $VLOOP$ in basis points (bps), round-trip trading costs $TCOST$ in bps, direct trading volume in non-dollar pairs (e.g., AUDJPY) in $bn, synthetic volume in dollar pairs in $bn, direct and synthetic relative bid–ask spreads, and realised volatility in non-dollar pairs in bps. By "synthetic" we refer to the sum of trading volumes and relative bid–ask spreads in two dollar pairs (e.g., USDAUD and USDJPY) within a currency pair triplet. Each row corresponds to a triplet of currency pairs, for example, AUDJPY, USDAUD, and USDJPY that we abbreviate as AUD-USD-JPY. The sample covers the period from 1 November 2011 to 30 September 2022.

the funding cost constraint η, while the last two measures both proxy the VaR constraint ω. To determine the top FX dealers we rely on the well-known Euromoney FX surveys. In every given year we assign an equal weight to each of the top FX dealers. Note that for certain variables (i.e., Value-at-Risk and number of VaR breaches) we were only able to collect data for a subset of banks. In such situations, we compute equally weighted averages based on the available set of dealer bank observations (see the Online Appendix for further details). For the VaR breaches we exploit the fact that US banks as well as foreign banks with US subsidiaries that are subject to the "Market Risk Capital Rule FFIEC 102" are required to report the number of VaR breaches in any given quarter since January 2015.

As a second step, we distil the information in the individual measures of dealer constraints to derive the composite dealer constraint measure DCM. The key advantage of DCM is that it encompasses all the model-based factors that can impact dealers' short-run flexibility to intermediate in currency markets. It is simply constructed by extracting the first principal component of the three individual dealer constraint series. The first principal component explains around 53% of the total variance of the individual dealer constraint time-series.

Fig. 4 depicts how the three model-derived dealer constraint measures and the composite measure vary over time. The three series exhibit a notable co-movement (cf. Table 2). Hence, the common component is well reflected by the composite dealer constraint measure. The decline in DCM from 2012 up to the Covid-19 pandemic is consistent with the drop in bank credit spreads after the European sovereign debt crisis (Berndt et al., 2023).

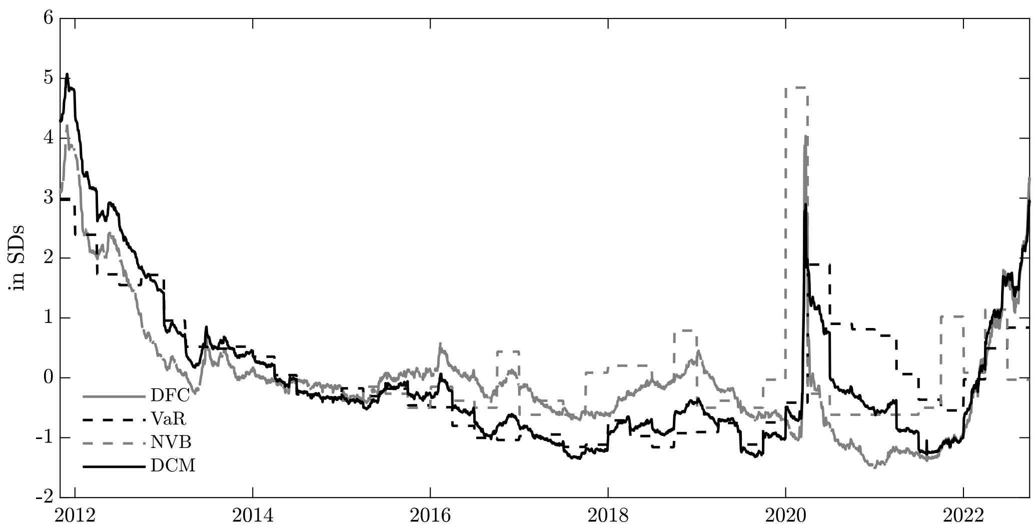

Fig. 4. State variables: Dealer constraint measure (DCM) and its components
Note: This figure plots different state variables that we observe at the daily and quarterly frequencies. Observations have been standardised by subtracting the sample mean and dividing by the standard deviation of every variable. The three state variables are primary FX dealer banks' (i) daily debt funding cost yield (DFC, solid grey line), (ii) quarterly Value-at-Risk measure (VaR, dashed black line), and (iii) the number of VaR breached in a given quarter (NVB, dashed grey line). We define our dealer constraint measure (DCM, black solid line) as the first principal component across these four variables. The sample spans from 1 November 2011 to 30 September 2022.

Besides the factors analysed explicitly in our theoretical framework, we also consider four additional dealer constraint measures proposed in the related literature. First, He et al. (2017) show that negative shocks to intermediary capital (i.e., dealers' quarterly leverage ratio) reduce their risk bearing capacity across many asset classes including FX. Second, an increase in dealers' credit default swap (CDS) premia and valuation adjustments (XVA) (Andersen et al., 2019) can hamper their willingness to make balance sheet space available when facing customer order flow imbalances. Third, CIP deviations reflect the relative tightness of dealer funding conditions (Rime et al., 2022) and balance sheet capacity in a broader sense (Du et al., 2018). We compute the average CIP deviations across our set of ten US dollar pairs.[17] Fourth, we follow Andersen et al. (2019) to devise an alternative measure for ω. Specifically, we compute the fraction of dealers that have 5-year CDS spreads above the 5-year USDJPY covered interest rate parity (CIP) basis. The intuition is that dealers can only arbitrage CIP violations if the deviations are larger than their credit spreads. We discuss robustness results based on these additional four measures in Section 6.

5. Liquidity provision and dealer constraints

In this section, we test the two main implications of our model. We start by exploring whether our liquidity cost measures (i.e., VLOOP and TCOST) are indeed positively related to various measures of dealer constraints (see Proposition 1). We then assess whether liquidity costs increase disproportionately more relative to dealer-intermediated volumes when dealer constraints tighten, in turn leading to a falling correlation between liquidity costs and trading volumes (see Proposition 2).

The analysis is split into three main parts. The first part presents motivating evidence using two simple correlation tables to support the empirical implications of the model. The first table (Table 2) relates our empirical liquidity cost measures to our model-derived measures of dealer constraints and provides evidence in favour of a positive association between the two. The second table (Table 3) shows that the correlation between the cost and the quantity of liquidity provision (i.e., dealer-intermediated volumes) depends on dealers' intermediation constraints. In the second part of the empirical analysis, we assess Proposition 2 of the model more formally. Specifically we rely on state-dependent regression analysis to quantify the change in the correlation between liquidity costs and trading volume, while controlling for factors influencing liquidity demand. Eventually, in the third part, we investigate the validity of Proposition 2 using structural vector autoregressions with sign restrictions that allow us to disentangle liquidity demand and supply shocks.

5.1. Motivating evidence

Table 2 provides motivating evidence in favour of the first prediction of our model. It illustrates how the two liquidity cost measures are contemporaneously positively related to dealers' debt funding costs η as well as VaR constraints ω. As outlined above, we employ two proxies for ω: First, we compute the average portfolio VaR across the top FX dealer banks and second, we identify the number of VaR breaches NVB in a given quarter for the same set of banks. The economic magnitude of the correlations is comparable across both VLOOP and TCOST, albeit TCOST seems to be more correlated with VaR as well as debt funding costs. In addition, VLOOP and TCOST are positively correlated with FX volatility σ, which is in line with our model. Hence, we will control for realised variance in all regression-based analyses.

The columns labelled $VLOOP\ FW$, $TCOST\ FW$, $VLOOP\ SW$, $TCOST\ SW$ show that these results are not confined to the FX spot market (represented by $TCOST$ and $VLOOP$) but are also reflected in the FX forward (FW) and swap (SW) markets, respectively. Specifically, we focus on the most frequently traded forward and swap contracts, which is the 1-week maturity. We follow Kloks et al. (2023) to compute forward and swap based measures of VLOOP and TCOST and find that these liquidity cost metrics are indeed positively correlated with measures of dealer constraints (i.e., η, VaR, and NVB). In the Online Appendix we provide additional details on how we construct these measures and show that the results in Table 3 carry over to currency forwards and swaps across various maturities (i.e., 1-month).

[17] Clearly, CIP deviations are only the symptoms of dealer balance sheet constraints and not their cause. Put differently, one can think of the CIP basis as a broad measure of limits to FX arbitrage.

Table 2
Correlations of key variables.

	$VLOOP$	$TCOST$	$VLOOPFW$	$TCOSTFW$	$VLOOPSW$	$TCOSTSW$	σ	η	VaR
TCOST	*** 59.28								
VLOOP FW	*** 19.74	* 5.99							
TCOST FW	*** 55.10	*** 94.76	*** 21.47						
VLOOP SW	7.19	1.54	*** 89.54	*** 15.06					
TCOST SW	*** 57.21	*** 97.90	*** 15.06	*** 98.94	** 9.07				
σ	*** 45.02	*** 45.43	−0.03	*** 38.79	−1.07	*** 42.41			
η	*** 30.90	*** 41.35	2.41	*** 36.52	3.56	*** 38.70	***13.47		
VaR	*** 26.86	*** 49.43	*** 14.70	*** 50.80	** 10.10	*** 50.23	−7.01	***41.07	
NVB	*** 25.06	*** 10.05	*** 15.51	4.46	*** 23.48	* 6.21	***29.88	***31.92	−3.76

Note: This table reports the pairwise Pearson correlation coefficient (in percent, %) of (log) changes in quarterly triangular no-arbitrage deviations $VLOOP$, round-trip trading costs $TCOST$, realised volatility σ, debt funding costs η, dealers' VaR measure VaR, and the number of VaR breaches NVB. $VLOOP\ FW$ and $TCOST\ FW$ are based on 1-week forward rates, whereas $VLOOP\ SW$ and $TCOST\ SW$ are the corresponding liquidity cost measures for FX swaps. Significant correlations at the 90%, 95%, and 99% levels are represented by asterisks *, **, and ***, respectively. The sample covers the period from 1 November 2011 to 30 September 2022 with the exception of NVB due to the fact that US and foreign dealer banks with US subsidiaries were not required to report VaR breaches prior to 2015.

Table 3
Liquidity provision cost characteristics across DCM percentiles..

DCM percentile		$\overline{VLOOP}$ in %	$\overline{TCOST}$ in %	$\overline{VLM}$ in $bn	$cor(VLOOP, VLM)$	$cor(TCOST, VLM)$	#Obs
Full sample	0.0	0.05	1.10	141.81	0.09	0.25	2801
Least constrained	0.1	0.05	1.12	144.41	0.10	0.25	2521
	0.2	0.05	1.15	147.85	0.10	0.24	2241
	0.3	0.05	1.17	150.57	0.10	*0.26	1961
	0.4	0.05	1.20	154.19	0.10	***0.28	1681
	0.5	0.05	1.22	156.58	0.10	***0.29	1401
	0.6	0.06	1.25	154.67	0.09	**0.27	1121
	0.7	0.06	1.30	159.25	***0.07	***0.21	841
	0.8	0.06	1.31	156.58	***0.05	***0.17	561
Most constrained	0.9	0.06	1.36	165.93	***−0.01	***0.08	281

Note: The first three columns in this table report the within-decile average of $VLOOP_{k,t}$ ($\overline{VLOOP}$) in %, $TCOST_{k,t}$ ($\overline{TCOST}$) in %, and the average $VLM_{k,t}$ ($\overline{VLM}$) in $bn across percentiles of the dealer constraint measure DCM. The underlying data are based on a panel of 15 currency pair triplets. Columns 4 and 5 tabulate the conditional Pearson correlation coefficient of (log) changes in the two liquidity cost measures (i.e., $VLOOP$ and $TCOST$) and total trading volume VLM across the percentiles of the dealer constraint measure. The last column shows the average number of observations for each DCM percentile. The asterisks *, **, and *** indicate that the correlation is significantly different from the full sample estimate (in the first row) at the 90%, 95%, and 99% levels. The corresponding test statistic for the conditional correlation cor^{τ} being equal to the full sample correlation $cor^{\tau=1.00}$, where $\tau \in 0.1, 0.2, \ldots, 0.9$ refers to DCM_t deciles, are based on the Fisher z-transformation. The sample covers the period from 1 November 2011 to 30 September 2022.

Table 3 presents empirical support for the second prediction of our model, namely that changes in dealer capacity have a nonlinear effect on market liquidity. The first three columns show how the average liquidity cost and dealer-intermediated volume (i.e., VLM) increase across the percentiles of our dealer constraint measure DCM. The monotonic increase in both liquidity cost measures and trading volume across the DCM percentiles suggests that the dealer sector as a whole accommodates the rise in trading demands even at times when intermediation constraints tighten. However, the increase in liquidity costs outpaces the increase in trading volume when the dealer sector is more constrained. For instance, TCOST increases by 23% compared to normal times, whereas volume (i.e., VLM) increases by only 17%. Eventually, columns 4 and 5 show the conditional correlation of (log) changes in each of our two liquidity cost measures and total trading volume.[18] Consistent with our model's predictions, we find that the correlation of volume with each of the two liquidity cost measures weakens substantially as DCM increases. For instance, the conditional correlation based on the highest DCM decile (i.e., when dealers are most constrained) is a mere 8% for TCOST, and hence economically and statistically significantly lower than the full-sample correlation of 25%.

These patterns are in line with the mechanism illustrated in Fig. 2. Specifically, they show that liquidity costs increase disproportionately more relative to intermediated volumes given the inward shift in liquidity supply. Note that this is *not* a mechanical effect as the unconditional correlation between changes in VLOOP or TCOST and DCM is less than 1%.

These initial results go beyond prior research exploring the nexus between FX market liquidity and funding liquidity (Mancini et al., 2013; Karnaukh et al., 2015). They do so by shedding light on the key mechanisms that lead to a deterioration of FX market liquidity when dealer constraints tighten. Moreover, they also relate to recent work by Duffie et al. (2023) showing that the relation between US Treasury market liquidity and volatility deteriorates during stressed periods (e.g., Covid-19 market turmoil). When we reiterate the conditional correlation analysis in Table 3 with volatility instead of trading volume, we find that also the liquidity cost–volatility relation tends to weaken when dealers are more constrained. This result aligns well with Duffie et al. (2023), suggesting that volatility and liquidity costs co-move positively in normal times, but less so when dealers exhaust their balance sheet capacity. Through the prism of our model, higher volatility is associated with more imbalanced customer demand as well as more constrained liquidity supply. In equilibrium, liquidity costs increase more than volatility, which reduces the correlation between the two variables. We document these additional findings in the Online Appendix.

Based on the theoretical framework in Section 3, the explanation for these empirical results hinges on two economic forces: For one, the price of market liquidity (i.e., VLOOP and TCOST) increases as dealers pass-through higher marginal funding costs (reflected by the surge of DCM) to their customers. For another, higher liquidity costs discourage customers' trading activity (given they are price-sensitive

[18] Note that the CLS volume data include the FX trading activity of all top dealer banks listed in the Euromoney FX surveys. In particular, the banks that show up in the Euromoney FX surveys are also the most dominant players on the CLS settlement system.

and have downward-sloping demand curves), thereby curbing trading demands. Consequently, when dealer constraints tighten, there is a marked imbalance: equilibrium trading volume expands less compared to the surge in the equilibrium price of liquidity. Therefore, our motivating evidence also highlights that combining information on both prices and quantities is a pivotal step in isolating the nonlinear impact of dealers' intermediation constraints on market liquidity conditions.

5.2. Regression analysis

To formally underpin the above reasoning that dealer constraints have a nonlinear effect on market liquidity, we employ smooth transition regression (LSTAR) models (e.g., van Dijk et al., 2002; Christiansen et al., 2011). These nonlinear LSTAR models are particularly well-suited for our analysis as constrained and unconstrained regimes are determined endogenously (i.e., the econometrician is not choosing a particular cutoff value) and may vary smoothly over time. In particular, the constrained and unconstrained periods (governed by γ and c) are determined by estimating a nonlinear regression model based on the generalised method of moments (GMM).[19]

For the LSTAR model, let $G(z_{t-1})$ be a logistic function depending on the 1-day lagged regime variable z_{t-1}:

$$G(z_{t-1}) = (1 + \exp(-\gamma'(z_{t-1} - c)))^{-1}, \tag{13}$$

where the parameter c is the central location and the vector γ determines the steepness of $G(z_{t-1})$. We use the 1-day lagged value of DCM in all our state-dependent regression analyses to rule out any contemporaneous relation between our dealer constraint measure and the amount of intermediated volume.[20] Hence, the LSTAR model is of the form

$$y_{k,t} = \lambda_t + \alpha_k + [1 - G(z_{t-1})]\beta_1' f_{k,t} + G(z_{t-1})\beta_2' f_{k,t} + \beta_3' \mathbf{w}_{k,t} + \varepsilon_{k,t}, \tag{14}$$

where the dependent variable $y_{k,t}$ is one of our two liquidity cost measures (i.e., VLOOP or TCOST), $f_{k,t}$ is the total aggregate trading volume (i.e., $VLM_{k,t}$) that is defined as the sum of trading volume in one non-dollar as well as two dollar pairs within a particular currency pair triplet k. The state-independent control variable $\mathbf{w}_{k,t}$ includes either the realised variance $RV_{k,t}$ or the price impact $Amihud_{k,t}$ in the non-dollar currency pair within each triplet k. We estimate $RV_{k,t}$ following Barndorff-Nielsen and Shephard (2002) as the sum of squared intraday midquote returns. Following Ranaldo and Santucci de Magistris (2022), we estimate the enhanced version of the Amihud (2002) measure as the ratio of daily realised volatility to aggregate daily trading volume. To limit the detrimental effect of outliers, we winsorize $Amihud_{k,t}$ at the 0.5% level.

The slope coefficients in (14) vary smoothly with the regime variable z_{t-1} from β_1 at low values of $\gamma' z_{t-1}$ to β_2 at high values of $\gamma' z_{t-1}$. There are two interesting boundary cases: First, if $\beta_1 = \beta_2$ we effectively have a linear regression. Second, the limit case where $\gamma \rightarrow \infty$ is equivalent to a linear regression with a dummy.

The key coefficient of interest in Eq. (14) is the difference between β_2 and β_1. It captures the change in the correlation between liquidity costs and dealer-intermediated volume across unconstrained and constrained regimes. To estimate all parameters (including γ and c) in Eq. (14), we use GMM and conduct inference based on Driscoll and Kraay's (1998) covariance matrix which allows for random clustering and serial correlation up to 8 lags. We choose the optimal number of lags (i.e., "bandwidth") using the plug-in procedure for automatic lag selection by Andrews and Monahan (1992) and Newey and West (1994), respectively. Across all regression specifications, both the dependent and independent variables are taken in logs and first differences. The obvious advantage of this is twofold: First, regression coefficients can be interpreted as elasticities. Second, FX volume in levels is non-stationary and persistent (see Ranaldo and Santucci de Magistris, 2022), hence taking first-differences is an effective way to render the time-series stationary.

Note that we include both cross-sectional α_k and time-series λ_t fixed effects to control for any unobservable heterogeneity that is constant across triplets of currency pairs k and days t, respectively. As a result, all reported R^2 are "within" rather than "overall" coefficients of determination. The inclusion of time-series fixed effects implicitly assumes that (lagged) dealer constraints have no direct bearing on liquidity costs. In the Online Appendix, we show that this assumption is reasonable given that the lagged dealer constraint measure is statistically insignificant in a regression without time-series fixed effects.

Table 4 presents our baseline results. We observe a consistent picture across all three specifications: the difference between the slope coefficient on trading volume in constrained and unconstrained periods (i.e., $\beta_2 - \beta_1$) is negative and statistically significant for both VLOOP and TCOST. Moreover, the estimated slope coefficients are at least 50% (e.g., $-0.07/0.12 = 58\%$) lower when dealer banks are constrained. Note that we control for several factors affecting the demand for liquidity. The day fixed effects λ_t control for any global market factors such as global volatility (Menkhoff et al., 2012) or global illiquidity (Karnaukh et al., 2015). In addition, following the intuition of our model, we also include the realised exchange rate variance $RV_{k,t}$ as a state-independent control variable to account for any differences in trading demands related to volatility. Moreover, currency volatility also controls for differences in dealer competition across currency pairs because it is highly correlated with the number of active dealers in the market (Huang and Masulis, 1999).

Related to our efforts to control for currency demand, one might wonder how much our results are driven by market-wide factors that are not dealer specific and which are also more related to liquidity demand rather than supply. To address this question, we conduct a placebo exercise where we explore a set of non-dealer specific regime variables that are presumably more exposed to liquidity demand as well as broad market conditions. In particular, we consider the VIX, the TED spread, the price of gold, and the LIBOR-OIS spread as alternative regime variables. We find that these state variables do not appropriately capture dealer constraints because the relation between liquidity costs and volume is not state-dependent. We document these additional results in the Online Appendix.

Thus far, the empirical results in this section provide two key takeaways that lend support to our model. First, in line with Proposition 1, an increase in measures of dealer constraints (i.e., DCM) is associated with a surge in liquidity costs (i.e., VLOOP and TCOST). Second, in line with Proposition 2, the correlation between liquidity cost measures and dealer-intermediated trading volume is significantly smaller during times when dealers are more constrained (i.e., the difference between the slope coefficients with respect to trading volume across unconstrained and constrained states is negative).

5.3. Disentangling liquidity demand and supply

Our theoretical framework in Section 3 suggests that the decline in the correlation between liquidity costs and trading volumes stems from a more inelastic (i.e., steeper) supply curve. However, empirically, one might be concerned that our dealer constraint measure (DCM), which we have used as a state variable in the previous subsection, is correlated

[19] Following Granger and Teräsvirta (1997), γ and c are free parameters that are bounded to avoid any corner solution (e.g., where all dealers are constrained at all times). Specifically, we allow γ to vary from 1 to 12, whereas c is bounded between -0.5 and $+0.5$. Our results are robust to varying these bounds.

[20] In the Online Appendix we show that our findings are robust to using up to 90 lags and are hence not driven by the fact that some of the DCM constituents are measured at the quarterly frequency (i.e., Value-at-Risk and number of VaR breaches). This exercise also provides evidence in favour of the idea that dealer constraints have a lasting (i.e., persistent) adverse effect on FX liquidity provision.

Table 4
Smooth transition regression with DCM as state variable..

	VLOOP			TCOST		
	(1)	(2)	(3)	(4)	(5)	(6)
γ	***12.00	***12.00	***12.00	***12.00	***12.00	***12.00
c	−0.03	−0.03	−0.03	***0.14	***0.14	***0.14
Unconstr. volume	***0.10	***0.10	***0.08	***0.12	***0.12	***0.10
	[4.21]	[4.10]	[3.24]	[16.90]	[16.85]	[12.92]
Constr. volume	−0.03	−0.03	−0.05	***0.05	***0.06	**0.03
	[0.91]	[0.97]	[1.43]	[4.11]	[4.17]	[2.54]
Amihud (2002)		−0.00			**0.00	
		[0.73]			[2.07]	
Realised variance			***0.02			***0.02
			[3.20]			[8.84]
Constr.-Unconstr.	***−0.13	***−0.13	***−0.12	***−0.07	***−0.07	***−0.06
	[3.16]	[3.16]	[3.06]	[4.45]	[4.46]	[3.96]
R^2 in %	0.09	0.09	0.14	2.38	2.40	3.44
Avg. #Time periods	2796	2796	2796	2801	2801	2800
#Currency triplets	15	15	15	15	15	15
Currency triplet FE	yes	yes	yes	yes	yes	yes
Time-series FE	yes	yes	yes	yes	yes	yes

Note: This table reports results from daily fixed effects LSTAR panel regressions of the form $y_{k,t} = \lambda_t + \alpha_k + [1 - G(z_{t-1})]\beta_1' f_{k,t} + G(z_{t-1})\beta_2' f_{k,t} + \beta_3' w_{k,t} + \epsilon_{k,t}$, where the dependent variable $y_{k,t}$ is a liquidity cost measure (i.e., $VLOOP$ or $TCOST$), $f_{k,t}$ ($w_{k,t}$) are state-dependent (*state-independent*) regressors, and $G(z_{t-1})$ is a logistic function depending on state variable z_{t-1}. The regime variable is the 1-day lagged value of the dealer constraint measure DCM_t. The optimal parameters γ and c are determined by nonlinear least squares minimising the concentrated sum of squared errors. Both dependent and independent variables are taken in logs and changes. The sample covers the period from 1 November 2011 to 30 September 2022. The test statistics based on Driscoll and Kraay (1998) robust standard errors allowing for random clustering and serial correlation (using the plug-in procedure for automatic lag selection by Andrews and Monahan (1992) and Newey and West (1994)) are reported in brackets. Asterisks *, **, and *** denote significance at the 90%, 95%, and 99% levels.

with factors simultaneously affecting both liquidity supply and demand.

To address this issue more conclusively, we now turn to a structural vector autoregression setup with sign restrictions. This econometric approach allows us to explicitly disentangle liquidity demand and supply dynamics. Specifically, we build on the approach by Uhlig (2005) and others (e.g., Canova and De Nicoló, 2002; Rubio-Ramírez et al., 2010), which has become widely used in economics and finance to estimate models with sign restrictions. Our empirical analysis proceeds in two steps.

In a first step, we estimate a structural (bivariate) vector autoregression (SVAR) model of liquidity cost measures (i.e., VLOOP or TCOST) and dealer-intermediated volume (VLM). To identify liquidity supply and demand shifts, we estimate the SVAR imposing sign restrictions in the spirit of Cohen et al. (2007), Goldberg (2020), and Goldberg and Nozawa (2020), respectively, using Bayesian methods (see the Online Appendix for further details). Let $Y_{k,t} = [X_{k,t}\ VLM_{k,t}]^T$ be a 2×1 vector containing $X \in \{VLOOP, TCOST\}$ and VLM in currency pair triplet k and day t. The bivariate panel SVAR for $Y_{k,t}$ is:

$$Y_{k,t} = \alpha_k + \sum_{i=1}^{l} B_{k,i} Y_{k,t-i} + \xi_{k,t}, \tag{15}$$

where $B_{k,i}$ is a 2×2 matrix of coefficients, l the lag length, $\xi_{k,t} = [\xi_{X;k,t}\ \xi_{VLM;k,t}]^T$ the reduced form error, and α_k is a 2×1 vector of currency triplet fixed effects. The vector of residuals $\xi_{k,t}$ can be mapped to the structural liquidity supply $\delta^s_{k,t}$ and demand $\delta^d_{k,t}$ shocks using the following relation:

$$\begin{bmatrix} \xi_{X;k,t} \\ \xi_{VLM;k,t} \end{bmatrix} = A_k \begin{bmatrix} \delta^s_{k,t} \\ \delta^d_{k,t} \end{bmatrix}, \tag{16}$$

where A_k is a 2×2 matrix and $\delta_{k,t} = [\delta^s_{k,t}\ \delta^d_{k,t}]^T$ is a 2×1 vector. Based on (15) and (16), the first column of A_k corresponds to changes in liquidity provision costs (i.e., VLOOP or TCOST) and dealer-intermediated volume associated with an increase in $\delta^s_{k,t}$. The second column in turn corresponds to changes in liquidity costs and intermediated volumes associated with an increase in $\delta^d_{k,t}$. Following Goldberg (2020), if A_k satisfies the following sign restrictions:

$$sign(A_k) = \begin{pmatrix} + & + \\ - & + \end{pmatrix}, \tag{17}$$

then $\delta^s_{k,t}$ can be interpreted as an inward shift in liquidity supply reflecting a tightening of dealer constraints, whereas $\delta^d_{k,t}$ corresponds to an outward shift in liquidity demand.

The sign restrictions in Eq. (17) assume that supply shifts lead to changes in liquidity costs and trading volume that have opposite signs. In other words, a shock to liquidity supply will lead to a rise in liquidity costs but at the same time a fall in dealer-intermediated volume. Demand shocks, by contrast, are assumed to lead to changes in liquidity costs and volume in the same direction. That is, in the case of demand shocks, the increase in liquidity costs goes in hand with a rise in dealer-intermediated volume.

These sign restrictions are fully consistent with our model (see Section 3), which rationalises how dealer constraints (i.e., ω and η) affect both the level and the slope of the liquidity supply curve. In particular, the SVAR model embraces the economic intuition in our model, namely, that an inward (*outward*) shift of the supply (*demand*) curve corresponds to a higher equilibrium price (i.e., cost of liquidity) when holding demand (*supply*) constant.

For illustrative purposes, Fig. 5 (Fig. 6) shows estimates of the impulse responses of $VLOOP$ ($TCOST$) and VLM to liquidity supply and demand shifts for the EUR-USD-JPY currency pair triplet. The impulse response functions for the other 14 triplets exhibit qualitatively similar patterns. In line with the above reasoning, concurrently with a supply shift, VLOOP (TCOST) rises and VLM declines. As shown in Fig. 5 (Fig. 6), contemporaneous with a supply shift, VLOOP (TCOST) rises by 32% (5%) and VLM declines by 16% (18%), according to the posterior mean. Contrarily, a demand shock is associated with an increase in VLOOP (TCOST) as well as an increase in VLM by 26% and 25% (12% and 20%), respectively.

In a second step, we estimate the correlation between the cost of liquidity provision (i.e., VLOOP or TCOST) and dealer-intermediated trading volume (i.e., VLM) in a 30-day rolling window fashion and estimate the following panel regression model:

$$\rho_{k,t} = \alpha_k + \phi_1 DCM_t + \phi_2 RV_{k,t} + \phi_3 Amihud_{k,t} + \epsilon_{k,t}, \tag{18}$$

where the dependent variable is the 30-day rolling window correlation of a liquidity cost measure (i.e., VLOOP or TCOST) and trading volume, α_k denotes currency triplet fixed effects, $RV_{k,t}$ the realised variance, $Amihud_{k,t}$ the price impact in the non-dollar currency pair within each

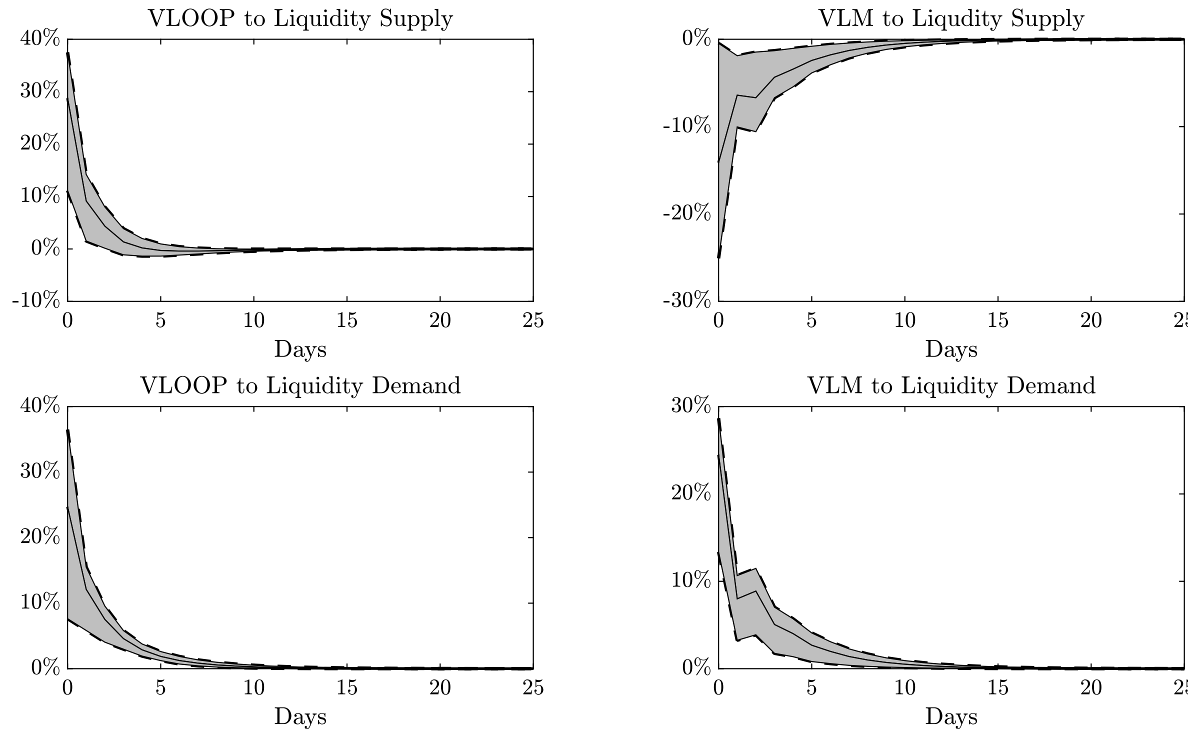

Fig. 5. Dynamic impulse response function for EUR-USD-JPY; VLOOP
Note: This figure plots the estimated dynamic impulse response of the shadow cost of intermediary constraints (VLOOP) and dealer-intermediated volume (VLM) associated with liquidity supply and demand shifts. The median response is shown by the black solid line. The grey shaded area marks a 95% confidence interval around the median. The sample covers the period from 1 November 2011 to 30 September 2022.

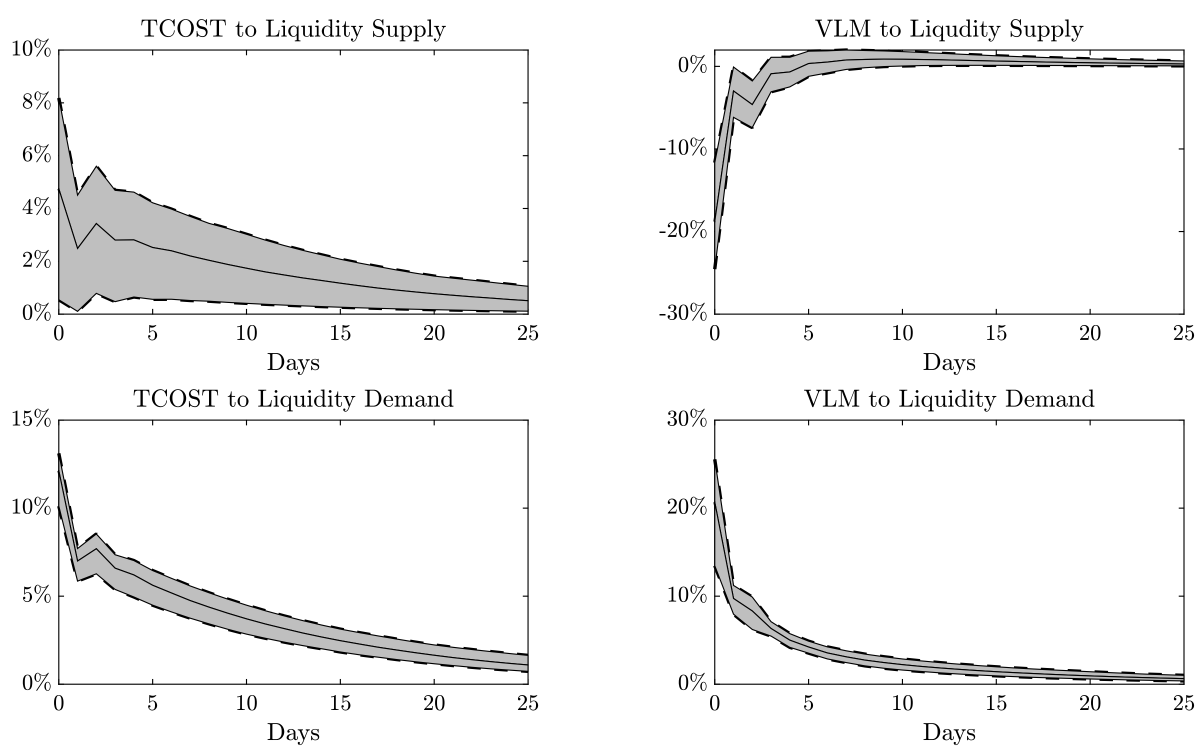

Fig. 6. Dynamic impulse response function for EUR-USD-JPY; TCOST
Note: This figure plots the estimated dynamic impulse response of dealers' compensation for enduring inventory imbalances (TCOST) and dealer-intermediated volume (VLM) associated with liquidity supply and demand shifts. The median response is shown by the black solid line. The grey shaded area marks a 95% confidence interval around the median. The sample covers the period from 1 November 2011 to 30 September 2022.

Table 5
Disentangling liquidity demand and supply..

	cor(VLOOP,VLM)			cor(TCOST,VLM)		
Panel A	(1)	(2)	(3)	(4)	(5)	(6)
DCM	***−0.09	***−0.09	***−0.10	*−0.08	*−0.08	*−0.07
	[2.62]	[2.64]	[2.73]	[1.92]	[1.90]	[1.75]
Realised variance		*0.02			−0.03	
		[1.72]			[1.56]	
Amihud (2002)			***0.07			***−0.14
			[3.93]			[5.39]
R^2 in %	0.84	0.89	1.06	0.63	0.73	1.40
Adj. R^2 in %	0.80	0.86	1.02	0.59	0.69	1.37
Avg. #Time periods	2773	2773	2773	2773	2773	2773
#Currency triplets	15	15	15	15	15	15
Currency triplet FE	yes	yes	yes	yes	yes	yes
Time series FE	no	no	no	no	no	no
Panel B						
δ^s	***−0.05	***−0.05	***−0.06	***−0.06	***−0.06	***−0.06
	[5.44]	[5.41]	[5.63]	[5.59]	[5.80]	[4.94]
δ^d	−0.00	−0.00	0.00	0.01	0.01	0.01
	[0.19]	[0.22]	[0.10]	[1.55]	[1.56]	[0.97]
Realised variance		**0.00			**−0.02	
		[2.20]			[2.07]	
Amihud (2002)			***0.04			***−0.06
			[2.60]			[4.63]
R^2 in %	0.24	0.24	0.29	0.24	0.28	0.51
Adj. R^2 in %	0.20	0.20	0.25	0.20	0.24	0.47
Avg. #Time periods	2773	2773	2773	2773	2773	2773
#Currency triplets	15	15	15	15	15	15
Currency triplet FE	yes	yes	yes	yes	yes	yes
Time series FE	yes	yes	yes	yes	yes	yes

Note: This table reports results from daily fixed effects panel regressions of the form $\rho_{k,t} = \alpha_k + \phi_1 DCM_t + \phi_2 RV_{k,t} + \phi_3 Amihud_{k,t} + \epsilon_{k,t}$, where the dependent variable is the 30-day rolling window correlation of a liquidity cost measure (i.e., $VLOOP$, or $TCOST$) and trading volume (i.e., VLM), α_k denotes cross-sectional fixed effects, $RV_{k,t}$ ($Amihud_{k,t}$) the realised variance (Amihud (2002) price impact) in the non-dollar currency pair within each triplet k, and DCM_t is our dealer constraint measure. Panel A shows the estimates of (18) using DCM, whereas Panel B uses both liquidity supply $\delta^s_{k,t}$ and demand shocks $\delta^d_{k,t}$ from the SVAR as alternative measures of tightening dealer constraints. All regressors have been normalised to have unit standard deviation. Hence, the regression coefficients measure the increase in ρ associated with a one standard deviation increase in DCM, δ^s, and δ^d, respectively. The sample covers the period from 1 September 2012 to 30 September 2022. The test statistics based on Driscoll and Kraay (1998) robust standard errors allowing for random clustering and serial correlation (using the plug-in procedure for automatic lag selection by Newey and West, 1994) are reported in brackets. Asterisks *, **, and *** denote significance at the 90%, 95%, and 99% levels.

triplet k, and DCM_t is our dealer constraint measure.

The regression in (18) may suffer from endogeneity of DCM_t due to a missing factor simultaneously affecting the correlation between liquidity costs and volumes $\rho_{k,t}$. In other words, DCM_t may not (fully) capture the dealer constraints η and ω in our model. To address this issue, we use the supply shocks that we extract from the SVAR as an alternative measure for tightening dealer constraints. Additionally, to account for potential shifts in liquidity demand, we include demand shocks as an additional control variable in (18).

Table 5 documents the results of estimating (18) using various measures of dealer constraints. In particular, Panel A shows the estimates of (18) using DCM, whereas Panel B uses both liquidity supply and demand shocks as alternative measures of tightening dealer constraints. We estimate demand and supply shocks from a panel SVAR with currency triplet fixed effects. The key takeaway from Table 5 is consistent with the LSTAR analysis (see Table 4) and corroborates the idea that more binding dealer constraints are associated with dealers' liquidity provision becoming less elastic (i.e., smaller $\rho_{k,t}$). It turns out that, in line with the model, liquidity supply (rather than demand) shocks are the pivotal determinant of the variation in the correlation between liquidity costs and trading volume.

6. Robustness tests

To investigate the robustness of our findings we run two additional empirical tests: (i) decompose the dealer constraint measure into its constituents and use alternative measures of dealer constraints (leverage ratio, CIP basis, CDS spreads) and (ii) capture the share of constrained dealers based on differences in CDS spreads and the CIP basis. We relegate detailed results on all additional robustness checks to the Online Appendix but provide a succinct summary below in this section.

Different components of dealer constraints. We consider the same LSTAR specification as in (14) but instead of our dealer constraint measure DCM we use its three constituents. In particular, we use the lagged value of primary FX dealer banks' quarterly Value-at-Risk measure (VaR), daily funding cost yield (DFC), and the number of VaR breaches (NVB) in a given quarter as regime variables. In addition, we follow the related literature and use three broad measures of dealer balance sheet capacity: (i) He et al. (2017) leverage ratio (i.e., 1−*capital ratio*) (quarterly), (ii) credit default swap (CDS) premia (daily), and (iii) the average CIP deviation (daily) across our set of ten US dollar currency pairs.

Table 6 reports the estimates of using each of the six aforementioned measures as a state variable. The difference between the constrained and unconstrained coefficient is negative and significant across all specifications for both VLOOP and TCOST. The only exception is column 9 (NVB for TCOST) where the difference is statistically insignificant. These estimates are in line with our baseline specification based on DCM in terms of economic magnitudes.

Share of constrained dealers. Here we investigate an alternative measure for the share of constrained dealers that is based on the related literature on funding liquidity and, in particular, Andersen et al. (2019). The key intuition is that arbitraging CIP violations is only beneficial to a dealer if the deviations exceed the dealer's credit spread. Put

Table 6
Smooth transition regression with different state variables..

	VLOOP						TCOST					
	DFC	VaR	NVB	HKM	CDS	CIP	DFC	VaR	NVB	HKM	CDS	CIP
γ	***12.00	*12.00	***12.00	***8.63	***12.00	***1.19	***1.00	***12.00	12.00	***12.00	***4.89	***12.00
c	***−0.25	0.06	***0.10	−0.01	***0.24	***0.50	***−0.37	***0.26	***0.50	−0.03	***0.03	***0.50
Unconstr. volume	***0.12	***0.08	***0.12	***0.10	***0.07	***0.14	***0.17	***0.10	***0.09	***0.11	***0.11	***0.09
	[3.90]	[3.09]	[4.52]	[3.67]	[3.07]	[2.94]	[6.50]	[12.51]	[10.68]	[11.66]	[12.81]	[11.58]
Constr. volume	−0.02	−0.03	−0.03	−0.02	*−0.07	*−0.15	0.01	***0.04	***0.12	***0.05	0.02	***0.05
	[0.68]	[1.00]	[0.78]	[0.76]	[1.87]	[1.84]	[0.52]	[3.62]	[5.68]	[5.68]	[1.05]	[3.05]
Realised variance	***0.02	***0.02	***0.02	***0.02	***0.02	***0.02	***0.03	***0.02	***0.03	***0.03	***0.03	***0.03
	[3.31]	[3.17]	[3.06]	[3.25]	[3.23]	[3.22]	[8.96]	[8.84]	[7.39]	[8.98]	[8.90]	[8.94]
Constr.-Unconstr.	***−0.14	***−0.11	***−0.16	***−0.12	***−0.14	**−0.28	***−0.16	***−0.05	0.03	***−0.06	***−0.10	**−0.04
	[3.37]	[2.73]	[3.03]	[3.00]	[3.16]	[2.37]	[3.44]	[3.57]	[1.40]	[4.45]	[5.18]	[2.31]
R^2 in %	0.15	0.13	0.28	0.14	0.14	0.12	3.42	3.41	3.57	3.45	3.52	3.33
BIC	94.08	94.08	88.67	94.08	94.08	94.08	52.57	52.57	49.43	52.57	52.55	52.58
Avg. #Time periods	2796	2796	1985	2796	2796	2796	2800	2800	1988	2800	2800	2800
#Currency triplets	15	15	15	15	15	15	15	15	15	15	15	15
Currency triplet FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Time-series FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Note: This table reports results from daily fixed effects LSTAR panel regressions of the form $y_{k,t} = \lambda_t + \alpha_k + [1 - G(z_{t-1})]\beta_1' f_{k,t} + G(z_{t-1})\beta_2' f_{k,t} + \beta_3' w_{k,t} + \varepsilon_{k,t}$, where the dependent variable $y_{k,t}$ is a liquidity cost measure (i.e., $VLOOP$ or $TCOST$), $f_{k,t}$ ($w_{k,t}$) are state-dependent (*state-independent*) regressors and $G(z_{t-1})$ is a logistic function depending on the regime variable z_{t-1}. The regime variables are the 1-day lagged values of primary FX dealer banks': daily funding cost yield (DFC, columns 1 and 7), quarterly Value-at-Risk measure (VaR, columns 2 and 8), number of VaR breaches per quarter (NVB, columns 3 and 9), quarterly He et al. (2017) leverage ratio (HKM, columns 4 and 10), daily credit default spread (CDS, columns 5 and 11), and daily average CIP basis in US dollar currency pairs (CIP, columns 6 and 12). Note that we assign an equal weight to each top 10 FX dealer bank (based on the Euromoney FX survey) when computing a cross-sectional average. The optimal parameters γ and c are determined by nonlinear least squares minimising the concentrated sum of squared errors. The row 'Constr. - Unconstr.' reports the difference between the slope coefficient on constrained and unconstrained volume, respectively. Both dependent and independent variables are taken in logs and changes. The sample covers the period from 1 November 2011 to 30 September 2022. The test statistics based on Driscoll and Kraay (1998) robust standard errors allowing for random clustering and serial correlation (using the plug-in procedure for automatic lag selection by Andrews and Monahan, 1992; Newey and West, 1994) are reported in brackets. Asterisks *, **, and *** denote significance at the 90%, 95%, and 99% levels.

differently, dealers that have credit spreads above the CIP basis are "constrained" in the sense that they are unable to perform the arbitrage trade. Specifically, we compare the 5-year basis in the USDJPY to the 5-year CDS spread of each top dealer bank and then compute the fraction of banks that have long-term CDS spreads below the long-term CIP basis. We focus on the USDJPY basis because it is the largest and most persistent for our sample period. Next, we use this fraction as an alternative measure for ω. Table 7 provides quantitative support in favour of this funding liquidity-based measure for the share of constrained dealers. The difference between the constrained and unconstrained coefficient is negative and significant across all specifications at the 10% significance level.

Additional analyses. In the Online Appendix we document eight additional robustness checks for our baseline result (see Table 4). First, we estimate the LSTAR currency pair triplet by triplet to shed light on the cross-sectional differences across currency pair triplets. In line with the panel regression, we find that the difference between the slope coefficient on trading volume in constrained and unconstrained periods is negative and significant for several currency pair triplets. Second, we split volume into inter-bank and customer-bank trades and find that large dealer banks mainly curtail their liquidity provision in trades with other banks (rather than customers). Third, we perform a subsample analysis to account for the rise of non-bank liquidity providers since 2016. In line with the hypothesis that non-bank liquidity providers are more flexible in their liquidity provision than traditional dealer banks, we find that the constrained minus unconstrained coefficient with respect to trading volume is almost twice as large in terms of economic magnitude for the first half than for the second half of our sample. Fourth, we relax time-series fixed effects and find that the lagged dealer constraint measure is insignificant in such a regression. Fifth, in a placebo exercise, we explore a different set of regime variables that are not dealer specific and which are thus more directly exposed to liquidity demand and general market conditions (e.g., the VIX index or the TED spread). We find no significant drop in the correlation between liquidity costs and trading volume for any of these alternative state variables. Sixth, we vary the number of lags in the LSTAR model and find that dealer constraints have a lasting adverse effect on FX liquidity provision. Seventh, we focus on the main London stock market trading hours to rule out that our results are driven by more illiquid trading hours. Lastly, we employ euro-based currency triplets to show that our results are not driven by the dominant role of the US dollar in FX trading.

To summarise, these additional robustness tests corroborate our previous results and support the main mechanisms of our theoretical framework. Dealers promote FX market liquidity in normal times through elastic liquidity provision. As such, dealer intermediation contributes to better market liquidity, that is, narrower spreads and more informative prices (i.e., lower transactions costs and tighter no-arbitrage conditions). However, when FX dealers are constrained they increase liquidity costs disproportionately more relative to their market-making activities (i.e., dealer-intermediated trading volumes).

7. Conclusion

In this paper, we have studied how constraints on dealers' intermediation capacity affect currency market liquidity. Using a simple model and a unique data set on global FX spot trading activity, we provide a novel analytical method to identify and measure how dealer constraints affect not only the price of market liquidity, but also its relation to the quantity of liquidity (i.e., trading volume). We show that during times when the dealer sector is more constrained, for instance, due to higher funding costs and/ or stricter Value-at-Risk limits, liquidity cost measures increase disproportionately more relative to equilibrium trading volumes. As a result, the otherwise strong and positive relation between liquidity costs and trading volume weakens by at least 50% relative to times when dealers are largely unconstrained. To account for changes in both liquidity demand and supply we employ a structural vector autoregression with sign restrictions and show that this result is mainly driven by a drop in the elasticity of liquidity supply rather than an increase in demand.

Our paper has implications for policymakers and academics alike. After the Global Financial Crisis erupted in 2008, policymakers have largely focused on making OTC derivatives markets more stable. However, the world's largest financial market measured by daily turnover,

Table 7
Smooth transition regression with constrained dealer share as state variable..

	VLOOP			TCOST		
	(1)	(2)	(3)	(4)	(5)	(6)
γ	***1.45	***1.43	***1.35	12.00	12.00	12.00
c	***0.50	***0.50	***0.50	***0.50	***0.50	***0.50
Unconstr. volume	***0.15	***0.15	***0.14	***0.11	***0.12	***0.09
	[4.25]	[4.11]	[3.68]	[14.77]	[14.85]	[11.07]
Constr. volume	**−0.10	**−0.11	**−0.12	***0.08	***0.09	***0.06
	[2.11]	[2.22]	[2.40]	[7.90]	[7.90]	[5.84]
Amihud (2002)		−0.01			**0.00	
		[1.59]			[2.27]	
Realised variance			**0.02			***0.03
			[2.47]			[9.16]
Constr.-Unconstr.	***−0.25	***−0.26	***−0.26	**−0.03	**−0.03	*−0.02
	[3.41]	[3.41]	[3.33]	[2.23]	[2.24]	[1.73]
R^2 in %	0.11	0.12	0.14	2.21	2.25	3.39
Avg. #Time periods	2604	2604	2604	2609	2609	2608
#Currency triplets	15	15	15	15	15	15
Currency triplet FE	yes	yes	yes	yes	yes	yes
Time-series FE	yes	yes	yes	yes	yes	yes

Note: This table reports results from daily fixed effects LSTAR panel regressions of the form $y_{k,t} = \lambda_t + \alpha_k + [1 - G(z_{t-1})]\beta_1' f_{k,t} + G(z_{t-1})\beta_2' f_{k,t} + \beta_3' w_{k,t} + \varepsilon_{k,t}$, where the dependent variable $y_{k,t}$ is a liquidity cost measure (i.e., $VLOOP$ or $TCOST$), $f_{k,t}$ ($w_{k,t}$) are state-dependent (*state-independent*) regressors, and $G(z_{t-1})$ is a logistic function depending on the state variable z_{t-1}. The regime variable is the 1-day lagged value of the share of constrained dealers (i.e., ω) that is captured by the fraction of top 10 dealer banks with 5-year CDS spreads exceeding the 5-year USDJPY CIP basis. The optimal parameters γ and c are determined by nonlinear least squares minimising the concentrated sum of squared errors. The row 'Constr. - Unconstr.' reports the difference between the slope coefficient on constrained and unconstrained volume, respectively. Both dependent and independent variables are taken in logs and changes. The sample covers the period from 1 September 2012 to 30 September 2022. The test statistics based on Driscoll and Kraay (1998) robust standard errors allowing for random clustering and serial correlation (using the plug-in procedure for automatic lag selection by Andrews and Monahan, 1992; Newey and West, 1994) are reported in brackets. Asterisks *, **, and *** denote significance at the 90%, 95%, and 99% levels.

the FX market, has been largely excluded from the post-crisis regulatory reforms. Specifically, policymakers have merely observed the ongoing changes such as the proliferation of multiple trading venues that have led to a surge in fragmentation of market liquidity in currency markets.[21] Our study shows that this type of fragmentation becomes amplified when dealer constraints tighten, that is, exactly when high market resilience would be desirable. With respect to the academic literature, our study covers the FX market, which is commonly regarded as one of the most liquid financial markets in the world. We leave the study of the role of dealer constraints on the liquidity provision in other OTC markets (e.g., government bonds and OTC derivatives) to future research.

CRediT authorship contribution statement

Wenqian Huang: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Angelo Ranaldo:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Andreas Schrimpf:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Fabricius Somogyi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

[21] See "FX execution algorithms and market functioning", Bank for International Settlements, Report submitted by a Study Group established by the Markets Committee, October 2020.

Declaration of competing interest

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

Appendix A. Proofs for the model

Proposition 1.

Proof. For currency pair z, which is assumed to have balanced order flows, Eq. (10) implies that all dealers are unconstrained since $q_i^z = 0$. However, for currency pairs x and y with unbalanced order flows the constrained dealers' *total* net positions are bounded by the following VaR constraints:

$$q_L^x = \frac{\omega T}{\sigma}, \qquad q_L^y = -\frac{\omega T}{\sigma}. \tag{19}$$

Thus, in equilibrium, the unconstrained dealers need to absorb the remaining order flows:

$$q_H^x = \underbrace{\sigma(1 - s^x)(2\pi - 1)}_{d^x} - \frac{\omega T}{\sigma}, \qquad q_H^y = \underbrace{\sigma(1 - s^y)(1 - 2\pi)}_{d^y} + \frac{\omega T}{\sigma}, \tag{20}$$

and the market clearing prices for currency pairs x and y are determined by the unconstrained dealers' supply functions that arise from Eq. (8). We assume that dealers have equal probabilities in meeting the traders and hence, take their orders "pro rata". Thus, the selling and buying amounts of the unconstrained dealers are

$$q_{H,S}^x = q_H^x \pi, \quad q_{H,B}^x = q_H^x(\pi - 1), \quad q_{H,S}^y = q_H^y(1 - \pi), \quad q_{H,B}^y = -q_H^y \pi. \tag{21}$$

There are two market clearing conditions for each currency pair: one for the case when the unconstrained dealers are buying (i.e., bid price) and another reflecting the situation when they are selling (i.e., ask price):

$$q^x_{H,B} = (1-\omega)\frac{b^x - e^x}{\eta}, \qquad q^x_{H,S} = (1-\omega)\frac{a^x - e^x}{\eta}; \tag{22}$$

$$q^y_{H,B} = (1-\omega)\frac{b^y - e^y}{\eta}, \qquad q^y_{H,S} = (1-\omega)\frac{a^y - e^y}{\eta}. \tag{23}$$

Taking currency pair x as an example (i.e., Eq. (22)), the left-hand side of the first (respectively second) equation is the amount that needs to be bought (respectively sold) by the unconstrained dealers (to clear the market), and the right-hand side is their supply function from Eq. (8). To solve this system of equations, we subtract both sides of the equations on the left from those on the right:

$$\left(\sigma(1-s^x)(2\pi-1) - \frac{\omega T}{\sigma}\right)(\pi - (\pi-1)) = \frac{(1-\omega)}{\eta}s^x \tag{24}$$

from which we can derive s^x. Similarly, s^y can be derived from Eq. (23). The bid–ask spread for currency pairs x and y turns out to be the same (due to the simplifying assumptions of symmetric directional order flows, as well as due to having the same σ, η, and T across currencies):

$$s^x = s^y = \frac{\eta((2\pi-1)\sigma^2 - \omega T)}{\sigma((2\pi-1)\eta\sigma + 1 - \omega)} \equiv s. \tag{25}$$

Similarly, the market clearing conditions for the balanced currency pair z are:

$$-\frac{1}{2}\sigma^z(1-s^z) = \frac{b^z - e^z}{\eta}, \qquad \frac{1}{2}\sigma^z(1-s^z) = \frac{a^z - e^z}{\eta}, \tag{26}$$

which leads to the following bid–ask spread:

$$s^z = \frac{\eta\sigma}{1+\eta\sigma}. \tag{27}$$

To derive the midquotes in currency pair x and y, we substitute $s = s^x = s^y$ into the market clearing conditions (22) and (23), respectively. Take currency pair x as an example. We first add both sides of the equations on the left in Eq. (22) to those on the right, and use $m^x = \frac{a^x+b^x}{2}$:

$$\left(\sigma(1-s)(2\pi-1) - \frac{\omega T}{\sigma}\right)(\pi + (\pi-1)) = \frac{(1-\omega)}{\eta}2(m^x - e^x), \tag{28}$$

$$m^x = e^x + (\pi - \frac{1}{2})\underbrace{\frac{\eta}{1-\omega}\left(\sigma(1-s)(2\pi-1) - \frac{\omega T}{\sigma}\right)}_{=s \text{ from Eq. (24)}}. \tag{29}$$

Using similar steps, we derive the midquotes for the other currency pairs. To summarise, the midquotes are:

$$\begin{aligned} m^x &= e^x + \left(\pi - \frac{1}{2}\right)\underbrace{\frac{\eta((2\pi-1)\sigma^2 - \omega T)}{\sigma((2\pi-1)\eta\sigma + 1 - \omega)}}_{s}, \\ m^y &= e^y + \left(\frac{1}{2} - \pi\right)\underbrace{\frac{\eta((2\pi-1)\sigma^2 - \omega T)}{\sigma((2\pi-1)\eta\sigma + 1 - \omega)}}_{s}, \quad m^z = e^z. \end{aligned} \tag{30}$$

The midquotes of currency pairs x and y in (30) deviate from their fundamental values e^x and e^y, respectively, if $\pi \neq \frac{1}{2}$ or $\eta \neq 0$. Contrarily, the midquote for currency pair z is equal to its fundamental value (i.e., $m^z = e^z$) because order flows are balanced. Substituting Eq. (30) into Eq. (1), deviations of the midquotes (set by the market clearing conditions and unconstrained dealers' supply function) from the fundamental values represent violations of the law of one price:

$$\text{VLOOP} = \frac{e^z}{(e^x + (\pi - \frac{1}{2})s)(e^y + (\frac{1}{2} - \pi)s)}, \tag{31}$$

and substituting Eq. (30) into Eq. (2), we can express TCOST as follows:

$$\text{TCOST} = \frac{\left(1 + \frac{s}{2(e^x + (\pi - \frac{1}{2})s)}\right)\left(1 + \frac{s}{2(e^y + (\frac{1}{2} - \pi)s)}\right)}{1 - \frac{s^z}{2e^z}}. \tag{32}$$

Taking the first order derivatives of s^z with respect to η and σ we have that s^z increases in both η and σ as follows:

$$\frac{ds^z}{d\eta} = \frac{\sigma}{(1+\eta\sigma)^2} > 0, \qquad \frac{ds^z}{d\sigma} = \frac{\eta}{(1+\eta\sigma)^2} > 0. \tag{33}$$

For currency pairs x and y we have that their bid–ask spreads s increase in η, σ, and ω as follows:

$$\frac{ds}{d\eta} = \frac{\sigma(1-\omega)((2\pi-1)\sigma^2 - \omega T)}{\sigma^2((2\pi-1)\eta\sigma + 1 - \omega)^2} > 0; \tag{34}$$

$$\frac{ds}{d\sigma} = \frac{\eta((1-\omega)(\sigma^2(2\pi-1) + \omega T) + 2\eta\sigma\omega T(2\pi-1))}{\sigma^2((2\pi-1)\eta\sigma + 1 - \omega)^2} > 0; \tag{35}$$

$$\frac{ds}{d\omega} = \frac{\sigma\eta((2\pi-1)\sigma^2 - (1+\eta\sigma(2\pi-1))T)}{\sigma^2((2\pi-1)\eta\sigma + 1 - \omega)^2} > 0, \tag{36}$$

where the first and third inequality come from the condition that the VaR thresholds are binding for the constrained dealers: $T/\sigma < \sigma(2\pi-1)(1-s) < \sigma(2\pi-1)$. To see the third inequality, note that the bid–ask spread can be re-written as

$$s = \frac{\eta\sigma(2\pi-1)\left(1 - \frac{\omega T}{\sigma^2(2\pi-1)}\right)}{1 + \eta\sigma(2\pi-1) - \omega}. \tag{37}$$

Replacing T with $T < \sigma^2(2\pi-1)(1-s)$, we have that

$$s > \frac{\eta\sigma(2\pi-1)(1-\omega(1-s))}{1+\eta\sigma(2\pi-1) - \omega}, \tag{38}$$

$$(1 + \eta\sigma(2\pi-1) - \omega)s > \eta\sigma(2\pi-1)(1-\omega) + \eta\sigma\omega s(2\pi-1), \tag{39}$$

$$s > \frac{\eta\sigma(2\pi-1)}{1+\eta\sigma(2\pi-1)} = 1 - \frac{1}{1+\eta\sigma(2\pi-1)}, \tag{40}$$

$$\frac{(2\pi-1)\sigma^2}{T} > \frac{1}{1-s} > (1+\eta\sigma(2\pi-1)), \tag{41}$$

$$(2\pi-1)\sigma^2 > (1+\eta\sigma(2\pi-1))T. \tag{42}$$

TCOST increases in s and s^z as evidenced by Eq. (32). In addition, all the spreads increase in η, σ, and ω as shown by Eq. (33) to Eq. (36). Hence, TCOST increases in η, σ, and ω.

As to VLOOP, we first note that VLOOP is above unity if $m^z > m^x m^y$, which is

$$\begin{aligned} &(e^x + (\pi - \frac{1}{2})s)(e^y + (\frac{1}{2} - \pi)s) < e^x e^y \\ \Leftrightarrow\ &(\pi - \frac{1}{2})(e^y - e^x)s - (\pi - \frac{1}{2})^2 s^2 < 0 \\ \Leftrightarrow\ &e^y < e^x + (\pi - \frac{1}{2})s \end{aligned} \tag{43}$$

where we use $e^z = e^x e^y$. Thus, we focus on the parameter space when the above inequality holds. Taking the derivative of VLOOP with respect to s and using Eq. (43) (which also holds with inequality when replacing $\pi - \frac{1}{2}$ with $2(\pi - \frac{1}{2}) = 2\pi - 1$), we have that

$$\begin{aligned} \frac{dVLOOP}{ds} &= \frac{8e^x e^y(2\pi-1)(e^x + (2\pi-1)s - e^y)}{(2e^x + (2\pi-1)s)^2(2e^y + (2\pi-1)s)^2} \\ &> \frac{8e^x e^y(2\pi-1)(\pi - \frac{1}{2})s}{(2e^x + (2\pi-1)s)^2(2e^y + (2\pi-1)s)^2} > 0 \end{aligned} \tag{44}$$

Because VLOOP increases in s, it also increases in η, σ, and ω. □

Proposition 2.

Proof. From Eq. (12), we can see that for a given quantity q^S, a higher η and/ or ω leads to a higher p^S. Thus, all else equal, increasing η or ω shift the supply curve inwards. By equating the demand curve Eq. (11) and the supply curve Eq. (12), the equilibrium price (which turns out to be equal to the relative bid–ask spread s in (25)) and quantity, which we denote as p^* and q^* are

$$q^* = \frac{(2\pi-1)(\sigma(1-\omega) + \eta\omega T)}{\eta\sigma(2\pi-1) + 1 - \omega}, \tag{45}$$

$$p^* = \frac{\eta((2\pi-1)\sigma^2 - \omega T)}{\sigma(\eta\sigma(2\pi-1) + 1 - \omega)} = s. \tag{46}$$

To see that the equilibrium quantity increases less than the increase of equilibrium price when dealers are more constrained, we show that q^*/p^* decreases in η and ω:

$$\frac{q^*}{p^*} = \frac{\sigma(2\pi-1)(\sigma(1-\omega)+\eta\omega T)}{\eta((2\pi-1)\sigma^2-\omega T)}, \tag{47}$$

$$\frac{d\frac{q^*}{p^*}}{d\eta} = -\frac{(2\pi-1)\sigma^2(1-\omega)}{\eta^2((2\pi-1)\sigma^2-\omega T)} < 0, \tag{48}$$

$$\frac{d\frac{q^*}{p^*}}{d\omega} = -\frac{(2\pi-1)\sigma^2((2\pi-1)\sigma^2-(1+\eta\sigma(2\pi-1))T)}{\eta((2\pi-1)\sigma^2-\omega T)^2} < 0 \tag{49}$$

where the last inequality comes from Eq. (42).

Furthermore, we show that the ratio of dealer-intermediated volume (i.e., VLM) and our liquidity cost measures (i.e., VLOOP and TCOST) also decreases when dealers are more constrained:

$$\frac{d\frac{VLM}{VLOOP}}{d\eta} < 0, \frac{d\frac{VLM}{VLOOP}}{d\omega} < 0, \frac{d\frac{VLM}{TCOST}}{d\eta} < 0, \frac{d\frac{VLM}{TCOST}}{d\omega} < 0. \tag{50}$$

Since the trading demand for currency pair j is $\sigma(1-s^j)$, the dealer-intermediated volume across the three currency pairs is

$$VLM = \sigma(1-s^x)+\sigma(1-s^y)+\sigma(1-s^z) = 2\sigma(1-s)+\sigma(1-s^z). \tag{51}$$

Hence, using Eq. (1) and $m^z = e^z$ and $s = s^x = s^y$, the ratio of dealer-intermediated volume and VLOOP is

$$\frac{VLM}{VLOOP} = \frac{\sigma}{e^z}(m^x m^y)(2(1-s)+(1-s^z)) \tag{52}$$

From Proposition 1, we have that $ds/d\eta > 0$ and $ds^z/d\eta > 0$. In addition, we can show that $d(m^x m^y)/d\eta$ is negative using the expression for m^x and m^y in Eq. (30):

$$\frac{dm^x m^y}{d\eta} = \left(\pi-\frac{1}{2}\right)(e^y-e^x-(2\pi-1)s)\frac{ds}{d\eta} < 0 \tag{53}$$

where the inequality comes from Eq. (43). Thus, as η increases, $VLM/VLOOP$ decreases.

Similarly, we can show that $VLM/VLOOP$ decreases in ω because we have $ds/d\omega > 0$, $ds^z/d\omega > 0$, and $d(m^x m^y)/d\omega < 0$:

$$\frac{dm^x m^y}{d\omega} = \left(\pi-\frac{1}{2}\right)(e^y-e^x-(2\pi-1)s)\frac{ds}{d\omega} < 0 \tag{54}$$

where the inequality comes from Eq. (43).

Using Eq. (2) and $m^z = e^z$, the ratio of dealer-intermediated volume and TCOST is given as follows

$$\frac{VLM}{TCOST} = \frac{\sigma(2(1-s)+(1-s^z))\left(1-\frac{s^z}{2e^z}\right)}{\left(1+\frac{1}{2(\frac{e^x}{s}+(\pi-\frac{1}{2}))}\right)\left(1+\frac{1}{2(\frac{e^y}{s}+(\frac{1}{2}-\pi))}\right)}. \tag{55}$$

When η or ω increases, we know from Proposition 1 that s and s^z increase. In addition, it is evident that the above ratio depends on η and ω only through s and s^z; and the ratio decreases in s and s^z because the numerator is decreasing in s and s^z, whereas the denominator is increasing in s and s^z, respectively. Hence, this ratio decreases in η and ω. □

References

Adrian, T., Boyarchenko, N., 2012. Intermediary leverage cycles and financial stability. SSRN Electron. J. http://dx.doi.org/10.2139/ssrn.2133385.

Adrian, T., Etula, E., Muir, T., 2014. Financial intermediaries and the cross-section of asset returns. J. Financ. 69 (6), 2557–2596. http://dx.doi.org/10.1111/jofi.12189.

Adrian, T., Shin, H.S., 2010. Liquidity and leverage. J. Financ. Intermediation 19 (3), 418–437. http://dx.doi.org/10.1016/j.jfi.2008.12.002.

Adrian, T., Shin, H.S., 2013. Procyclical leverage and value-at-risk. Rev. Financ. Stud. 27 (2), 373–403. http://dx.doi.org/10.1093/rfs/hht068.

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. J. Financ. Mark. 5 (1), 31–56. http://dx.doi.org/10.1016/s1386-4181(01)00024-6.

Andersen, L., Duffie, D., Song, Y., 2019. Funding value adjustments. J. Financ. 74 (1), 145–192. http://dx.doi.org/10.1111/jofi.12739.

Andrews, D.W.K., Monahan, J.C., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. Econometrica 60 (4), 953–966. http://dx.doi.org/10.2307/2951574.

Bao, J., O'Hara, M., Zhou, X.A., 2018. The Volcker rule and corporate bond market making in times of stress. J. Financ. Econ. 130 (1), 95–113. http://dx.doi.org/10.1016/j.jfineco.2018.06.001.

Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. R. Stat. Soc. Ser. B Stat. Methodol. 64 (2), 253–280. http://dx.doi.org/10.1111/1467-9868.00336.

Baron, M., Muir, T., 2021. Intermediaries and asset prices: International evidence since 1870. In: Nieuwerburgh, S.V. (Ed.), Rev. Financ. Stud. 35 (5), 2144–2189. http://dx.doi.org/10.1093/rfs/hhab077.

Berndt, A., Duffie, D., Zhu, Y., 2023. Across-the-curve credit spread indices. Financ. Mark. Inst. Instrum. 32 (3), 115–130. http://dx.doi.org/10.1111/fmii.12172.

van Binsbergen, J.H., Graham, J.R., Yang, J., 2010. The cost of debt. J. Financ. 65 (6), 2089–2136. http://dx.doi.org/10.1111/j.1540-6261.2010.01611.x.

Bjønnes, G.H., Rime, D., 2005. Dealer behavior and trading systems in foreign exchange markets. J. Financ. Econ. 75 (3), 571–605. http://dx.doi.org/10.1016/j.jfineco.2004.08.001.

Canova, F., De Nicoló, G., 2002. Monetary disturbances matter for business fluctuations in the G7. J. Monet. Econ. 49 (6), 1131–1159. http://dx.doi.org/10.1016/s0304-3932(02)00145-9.

Cespa, G., Gargano, A., Riddiough, S.J., Sarno, L., 2021. Foreign exchange volume. In: Nieuwerburgh, S.V. (Ed.), Rev. Financ. Stud. 35 (5), 2386–2427. http://dx.doi.org/10.1093/rfs/hhab095.

CGFS, 2014. Market-making and proprietary trading: industry trends, drivers and policy implications. BIS CGFS Papers No. 52.

Chaboud, A.P., Chernenko, S.V., Wright, J.H., 2008. Trading activity and macroeconomic announcements in high-frequency exchange rate data. J. Eur. Econ. Assoc. 6 (2–3), 589–596. http://dx.doi.org/10.1162/jeea.2008.6.2-3.589.

Chaboud, A.P., Chiquoine, B., Hjalmarsson, E., Vega, C., 2014. Rise of the machines: Algorithmic trading in the foreign exchange market. J. Financ. 69 (5), 2045–2084. http://dx.doi.org/10.1111/jofi.12186.

Chen, H., Joslin, S., Ni, S.X., 2018. Demand for crash insurance, intermediary constraints, and risk premia in financial markets. Rev. Financ. Stud. 32 (1), 228–265. http://dx.doi.org/10.1093/rfs/hhy004.

Christiansen, C., Ranaldo, A., Söderlind, P., 2011. The time-varying systematic risk of carry trade strategies. J. Financ. Quant. Anal. 46 (04), 1107–1125. http://dx.doi.org/10.1017/s0022109011000263.

Chu, Y., Hirshleifer, D., Ma, L., 2020. The causal effect of limits to arbitrage on asset pricing anomalies. J. Financ. 75 (5), 2631–2672. http://dx.doi.org/10.1111/jofi.12947.

Cohen, L., Diether, K.B., Malloy, C.J., 2007. Supply and demand shifts in the shorting market. J. Financ. 62 (5), 2061–2096. http://dx.doi.org/10.1111/j.1540-6261.2007.01269.x.

Comerton-Forde, C., Hendershott, T., Jones, C.M., Moulton, P.C., Seasholes, M.S., 2010. Time variation in liquidity: The role of market-maker inventories and revenues. J. Financ. 65 (1), 295–331. http://dx.doi.org/10.1111/j.1540-6261.2009.01530.x.

Çötelioğlu, E., Franzoni, F., Plazzi, A., 2020. What constrains liquidity provision? Evidence from institutional trades. Rev. Financ. 25 (2), 485–517. http://dx.doi.org/10.1093/rof/rfaa016.

van Dijk, D., Teräsvirta, T., Franses, P.H., 2002. Smooth transition autoregressive models — A survey of recent developments. Econometric Rev. 21 (1), 1–47. http://dx.doi.org/10.1081/etc-120008723.

Driscoll, J.C., Kraay, A.C., 1998. Consistent covariance matrix estimation with spatially dependent panel data. Rev. Econ. Stat. 80 (4), 549–560. http://dx.doi.org/10.1162/003465398557825.

Du, W., Hébert, B., Huber, A.W., 2022. Are intermediary constraints priced? In: Giglio, S. (Ed.), Rev. Financ. Stud. http://dx.doi.org/10.1093/rfs/hhac050.

Du, W., Hébert, B., Li, W., 2023. Intermediary balance sheets and the treasury yield curve. J. Financ. Econ. 150 (3), 103722. http://dx.doi.org/10.1016/j.jfineco.2023.103722.

Du, W., Tepper, A., Verdelhan, A., 2018. Deviations from covered interest rate parity. J. Financ. 73 (3), 915–957. http://dx.doi.org/10.1111/jofi.12620.

Duffie, D., 2010. Presidential address: Asset price dynamics with slow-moving capital. J. Financ. 65 (4), 1237–1267. http://dx.doi.org/10.1111/j.1540-6261.2010.01569.x.

Duffie, D., 2018. Post-crisis bank regulations and financial market liquidity. Lezioni Paolo Baffi di Moneta e Finanza.
Duffie, J.D., 2023. Resilience redux in the U.S. treasury market. In: Stanford University Graduate School of Business Research Paper No. 4552735, Jackson Hole Symposium, Federal Reserve Bank of Kansas City, August, 2023. http://dx.doi.org/10.2139/ssrn.4552735.
Duffie, D., Fleming, M.J., Keane, F.M., Nelson, C., Shachar, O., Van Tassel, P., 2023. Dealer capacity and U.S. treasury market functionality. Staff Reports (Federal Reserve Bank of New York), http://dx.doi.org/10.59576/sr.1070,
Duffie, D., Pan, J., 1997. An overview of value at risk. J. Deriv. 4 (3), 7–49. http://dx.doi.org/10.3905/jod.1997.407971.
Evans, M.D., 2002. FX trading and exchange rate dynamics. J. Financ. 57 (6), 2405–2447. http://dx.doi.org/10.1111/1540-6261.00501.
Evans, M.D., Lyons, R.K., 2002. Order flow and exchange rate dynamics. J. Political Econ. 110 (1), 247–290. http://dx.doi.org/10.1142/9789813148543_0006.
Evans, M.D., Lyons, R.K., 2005. Do currency markets absorb news quickly? J. Int. Money Financ. 24 (2), 197–217. http://dx.doi.org/10.1016/j.jimonfin.2004.12.004.
Fleckenstein, M., Longstaff, F., 2018. Shadow funding costs: Measuring the cost of balance sheet constraints. http://dx.doi.org/10.3386/w24224, Unpublished working paper. NBER.
Foucault, T., Kozhan, R., Tham, W.W., 2016. Toxic arbitrage. Rev. Financ. Stud. 30 (4), 1053–1094. http://dx.doi.org/10.1093/rfs/hhw103.
Foucault, T., Pagano, M., Roell, A., 2013. Market Liquidity. Oxford University Press, http://dx.doi.org/10.1093/acprof:oso/9780199936243.001.0001.
Gallien, F., Kassibrakis, S., Klimenko, N., Malamud, S., Teguia, A., 2018. Liquidity provision in the foreign exchange market. In: Swiss Finance Institute Research Paper No. 18–56. (18–56), http://dx.doi.org/10.2139/ssrn.3234406.
Gârleanu, N., Pedersen, L.H., 2011. Margin-based asset pricing and deviations from the law of one price. Rev. Financ. Stud. 24 (6), 1980–2022. http://dx.doi.org/10.1093/rfs/hhr027.
Goldberg, J., 2020. Liquidity supply by broker-dealers and real activity. J. Financ. Econ. 136 (3), 806–827. http://dx.doi.org/10.1016/j.jfineco.2019.11.006.
Goldberg, J., Nozawa, Y., 2020. Liquidity supply in the corporate bond market. J. Financ. 76 (2), 755–796. http://dx.doi.org/10.1111/jofi.12991.
Gospodinov, N., Robotti, C., 2021. Common pricing across asset classes: Empirical evidence revisited. J. Financ. Econ. 140 (1), 292–324. http://dx.doi.org/10.1016/j.jfineco.2020.12.001.
Granger, C.W.J., Teräsvirta, T., 1997. Modelling Nonlinear Economic Relationships. Oxford University Press.
Gromb, D., Vayanos, D., 2002. Equilibrium and welfare in markets with financially constrained arbitrageurs. J. Financ. Econ. 66 (2–3), 361–407. http://dx.doi.org/10.1016/s0304-405x(02)00228-3.
Gromb, D., Vayanos, D., 2010. Limits of arbitrage. Annu. Rev. Financ. Econ. 2 (1), 251–275. http://dx.doi.org/10.1146/annurev-financial-073009-104107.
Grossman, S.J., Miller, M.H., 1988. Liquidity and market structure. J. Financ. 43 (3), 617–633. http://dx.doi.org/10.1111/j.1540-6261.1988.tb04594.x.
Haddad, V., Muir, T., 2021. Do intermediaries matter for aggregate asset prices? J. Financ. 76 (6), 2719–2761. http://dx.doi.org/10.1111/jofi.13086.
Hasbrouck, J., Levich, R.M., 2018. FX market metrics: New findings based on CLS bank settlement data. SSRN Electron. J. http://dx.doi.org/10.2139/ssrn.2912976.
Hasbrouck, J., Levich, R.M., 2021. Network structure and pricing in the FX market. J. Financ. Econ. 141 (2), 705–729. http://dx.doi.org/10.1016/j.jfineco.2021.04.013.
He, Z., Kelly, B., Manela, A., 2017. Intermediary asset pricing: New evidence from many asset classes. J. Financ. Econ. 126 (1), 1–35. http://dx.doi.org/10.1016/j.jfineco.2017.08.002.
He, Z., Krishnamurthy, A., 2011. A model of capital and crises. Rev. Econ. Stud. 79 (2), 735–777. http://dx.doi.org/10.1093/restud/rdr036.
He, Z., Krishnamurthy, A., 2013. Intermediary asset pricing. Am. Econ. Rev. 103 (2), 732–770. http://dx.doi.org/10.1257/aer.103.2.732.
He, Z., Nagel, S., Song, Z., 2022. Treasury inconvenience yields during the COVID-19 crisis. J. Financ. Econ. 143 (1), 57–79. http://dx.doi.org/10.1016/j.jfineco.2021.05.044.
Hendershott, T., Menkveld, A.J., 2014. Price pressures. J. Financ. Econ. 114 (3), 405–423. http://dx.doi.org/10.1016/j.jfineco.2014.08.001.
Hombert, J., Thesmar, D., 2014. Overcoming limits of arbitrage: Theory and evidence. J. Financ. Econ. 111 (1), 26–44. http://dx.doi.org/10.1016/j.jfineco.2013.09.003.
Huang, R.D., Masulis, R.W., 1999. FX spreads and dealer competition across the 24-hour trading day. Rev. Financ. Stud. 12 (1), 61–93. http://dx.doi.org/10.1093/rfs/12.1.61.
Huang, W., O'Neill, P., Ranaldo, A., Yu, S., 2023. HFTs and dealer banks: Liquidity and price discovery in FX trading. In: Swiss Finance Institute Research Paper No. 23-48. http://dx.doi.org/10.2139/ssrn.4349184.
Jeanneret, A., Sokolovski, V., 2019. Commodity prices and currencies. SSRN Electron. J. http://dx.doi.org/10.2139/ssrn.3379090.
Kargar, M., 2021. Heterogeneous intermediary asset pricing. J. Financ. Econ. 141 (2), 505–532. http://dx.doi.org/10.1016/j.jfineco.2021.04.012.
Karnaukh, N., Ranaldo, A., Söderlind, P., 2015. Understanding FX liquidity. Rev. Financ. Stud. 28 (11), 3073–3108. http://dx.doi.org/10.1093/rfs/hhv029.
Kilian, L., Taylor, M.P., 2003. Why is it so difficult to beat the random walk forecast of exchange rates? J. Int. Econ. 60 (1), 85–107. http://dx.doi.org/10.1016/s0022-1996(02)00060-0.
Kisin, R., Manela, A., 2016. The shadow cost of bank capital requirements. Rev. Financ. Stud. 29 (7), 1780–1820. http://dx.doi.org/10.1093/rfs/hhw022.
Kloks, P., Mattille, E., Ranaldo, A., 2023. Foreign exchange swap liquidity. In: Swiss Finance Institute Research Paper No. 23-22. http://dx.doi.org/10.2139/ssrn.4398052.
Lu, L., Wallen, J., 2024. What do bank trading desks do? http://dx.doi.org/10.2139/ssrn.4898830, Harvard Business School Working Paper.
Mancini, L., Ranaldo, A., Wrampelmeyer, J., 2013. Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums. J. Financ. 68 (5), 1805–1841. http://dx.doi.org/10.1111/jofi.12053.
Menkhoff, L., Sarno, L., Schmeling, M., Schrimpf, A., 2012. Carry trades and global foreign exchange volatility. J. Financ. 67 (2), 681–718. http://dx.doi.org/10.1111/j.1540-6261.2012.01728.x.
Menkhoff, L., Sarno, L., Schmeling, M., Schrimpf, A., 2016. Information flows in foreign exchange markets: Dissecting customer currency trades. J. Financ. 71 (2), 601–634. http://dx.doi.org/10.1111/jofi.12378.
Nagel, S., 2012. Evaporating liquidity. Rev. Financ. Stud. 25 (7), 2005–2039. http://dx.doi.org/10.1093/rfs/hhs066.
Newey, W.K., West, K.D., 1994. Automatic lag selection in covariance matrix estimation. Rev. Econ. Stud. 61 (4), 631–653. http://dx.doi.org/10.2307/2297912.
Pasquariello, P., 2014. Financial market dislocations. Rev. Financ. Stud. 27 (6), 1868–1914. http://dx.doi.org/10.1093/rfs/hhu007.
Payne, R., 2003. Informed trade in spot foreign exchange markets: an empirical investigation. J. Int. Econ. 61 (2), 307–329. http://dx.doi.org/10.1016/s0022-1996(03)00003-5.
Ranaldo, A., Santucci de Magistris, P., 2022. Liquidity in the global currency market. J. Financ. Econ. 146 (3), 859–883. http://dx.doi.org/10.1016/j.jfineco.2022.09.004.
Ranaldo, A., Somogyi, F., 2021. Asymmetric information risk in FX markets. J. Financ. Econ. 140 (2), 391–411. http://dx.doi.org/10.1016/j.jfineco.2020.12.007.
Rime, D., Schrimpf, A., Syrstad, O., 2022. Covered interest parity arbitrage. Rev. Financ. Stud. 35 (11), 5185–5227. http://dx.doi.org/10.1093/rfs/hhac026.
Rösch, D., 2021. The impact of arbitrage on market liquidity. J. Financ. Econ. 142 (1), 195–213. http://dx.doi.org/10.1016/j.jfineco.2021.04.034.
Rösch, D.M., Subrahmanyam, A., van Dijk, M.A., 2016. The dynamics of market efficiency. Rev. Financ. Stud. 30 (4), 1151–1187. http://dx.doi.org/10.1093/rfs/hhw085.
Rubio-Ramírez, J.F., Waggoner, D.F., Zha, T., 2010. Structural vector autoregressions: Theory of identification and algorithms for inference. Rev. Econ. Stud. 77 (2), 665–696. http://dx.doi.org/10.1111/j.1467-937x.2009.00578.x.
Schrimpf, A., Sushko, V., 2019. FX trade execution: Complex and highly fragmented. BIS Quarterly Review.
Scott, J.H., 1976. A theory of optimal capital structure. Bell J. Econ. 7 (1), 33. http://dx.doi.org/10.2307/3003189.
Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. J. Financ. 52 (1), 35–55. http://dx.doi.org/10.1111/j.1540-6261.1997.tb03807.x.
Siriwardane, E., Sunderam, A., Wallen, J., 2025. Segmented arbitrage. J. Financ. http://dx.doi.org/10.2139/ssrn.3960980, Forthcoming.
Tenreyro, S., Thwaites, G., 2016. Pushing on a string: US monetary policy is less powerful in recessions. Am. Econ. J.: Macroecon. 8 (4), 43–74. http://dx.doi.org/10.1257/mac.20150016.
Uhlig, H., 2005. What are the effects of monetary policy on output? Results from an agnostic identification procedure. J. Monet. Econ. 52 (2), 381–419. http://dx.doi.org/10.1016/j.jmoneco.2004.05.007.