

## BIS Working Papers No 1094

# The foreign exchange market

by Alain Chaboud, Dagfinn Rime and Vladyslav Sushko

Monetary and Economic Department

April 2023

JEL classification: F31, G15.

Keywords: financial markets, foreign exchange, market microstructure, dealer intermediation, electronic trading, algorithmic trading.

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.
This publication is available on the DIS website (www.bis.org)
This publication is available on the BIS website (www.bis.org).
© Bank for International Settlements 2023. All rights reserved. Brief excerpts may be reproduced or translated provided the source is stated.
ISSN 1020-0959 (print) ISSN 1682-7678 (online)

## The Foreign Exchange Market\*

Alain Chaboud<sup>†</sup> Dagfinn Rime<sup>‡</sup> Vladyslav Sushko<sup>§</sup>
April 5, 2023

#### **Abstract**

This chapter discusses the structure and functioning of the spot foreign exchange (FX) market. The market structure, which has become far more complex over the past three decades, has mostly evolved endogenously as the global FX market is subject to notably less regulatory oversight than equity and bond markets in most countries. Major banks used to dominate liquidity provision but they have found their role challenged by High Frequency Trading firms in an increasingly fragmented electronic market. The information structure of the market has also changed. As such, high-frequency cross-asset correlations, especially with the futures market, have become more important. The chapter also discusses the important role of the official sector in the FX market, and it highlights a few special topics such as flash events and the FX fixing scandal. We conclude with some suggestions for future research.

**Keywords:** Financial markets, Foreign exchange, Market microstructure, Dealer intermediation, Electronic trading, Algorithmic trading.

JEL classification: F31, G15

<sup>\*</sup>This paper has been prepared for *The Research Handbook of Financial Markets*, edited by Refet Gürkaynak and Jonathan Wright. We are grateful to Refet Gürkaynak and Jonathan Wright, our editors, and to Angelo Ranaldo, our discussant, for providing valuable comments. We have also benefited from conversations and email exchanges with Mark Bruce, Alexei Jiltsov, Colin Lambert, Alexis Laming, Thomas Maag, Eugene Markman, Josh Matthews, James O'Connor, Roel Oomen, Dan Reichgott, Morten Salvesen, Merhrdad Samadi, James Sinclair, Nikola Tarashev and Clara Vega. We thank EBS, Refinitiv, and the BIS for kindly providing data. The views expressed in this chapter are solely those of the authors and should not be interpreted to represent the views of the Federal Reserve Board or of the Bank for International Settlements.

<sup>&</sup>lt;sup>†</sup>Board of Governors of the Federal Reserve System

<sup>&</sup>lt;sup>‡</sup>BI Norwegian Business School

<sup>§</sup>Bank for International Settlements

Foreign exchange spot is the simplest asset class one can trade, yet it has the most complex trading environment.

Quote from an executive at a major FX liquidity provider

#### 1 Introduction

The foreign exchange (FX) market, where the relative prices of the world's currencies are determined, is essential for international transactions in goods, services and financial assets. In addition, FX is often viewed as an asset class on its own. The end-users of the FX market are therefore comprised of a wide variety of financial and non-financial customers around the globe. The trading activity of these agents and their interaction with market intermediaries drives the process of exchange rate determination, which has an impact on virtually all international economic activity. As a result, the FX market is the largest financial market in the world. FX trading volumes are, for example, much larger than global equity market activity (King and Rime, 2010).

Currency trading takes place around the world and around the clock, with a weekly cycle beginning early Monday morning in the Asia/Pacific region and ending Friday afternoon in the Americas. Trading activity often peaks when London and New York daytime trading hours overlap and is relatively thin during the so-called "witching hour" period, the late afternoon in New York and early morning in Asia.

More than 50 currencies are regularly traded, but the US dollar (USD) has for a long time commanded the dominant status of a vehicle currency. The USD is on one side of almost 90% of all global FX transactions, with the euro (EUR) and Japanese yen (JPY) in distant second and third places (BIS, 2022). Similarly, while the number of market participants around the globe is very high, much of the liquidity provision in FX is concentrated among a relatively small number of global banks and non-bank liquidity providers. Thus, despite its global and dispersed nature, certain aspects of the FX market exhibit a high degree of concentration.

The market structure that supports this activity is constantly evolving. Broadly speaking, the FX market is an over-the-counter (OTC) market in which electronic trading has grown rapidly since the early 2000s. It used to be characterized by two very distinct segments, interdealer and dealer to customer. This distinction has become blurred over time, with a proliferation of trading venues, a growing variety of execution methods, and some non-bank actors emerging as liquidity providers alongside bank dealers.

This chapter discusses how the spot FX market is organized and how it functions, including the main participants, the structure of the market, and the role of the official sector. Knowing how the FX market functions is critical to understanding how it converges on equilibrium exchange rates, but this chapter does not focus on exchange rates or exchange rate determination per se, which is the subject of a vast literature.<sup>1</sup>

<sup>&</sup>lt;sup>1</sup>For an overview of the theoretical and empirical literature on exchange rates and exchange rate determi-

We note that the foreign exchange market also includes a large amount of trading in derivatives, including FX swaps, currency swaps, forwards, futures and options. FX swaps and currency swaps, important instruments primarily used for funding and hedging purposes, are covered in **Chapter Angelo**.

### 2 Market size and main players

Measuring global trading activity presents a challenge as the global FX market is obviously not under a single jurisdiction. However, a comprehensive and authoritative source of information, albeit infrequent, is the Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity (the "Triennial"). The Triennial provides a snapshot of daily FX trading activity every third year in the month of April.<sup>2</sup> The data for the Triennial are collected by central banks from bank-dealers in their jurisdictions and then aggregated, analyzed and published by the Bank for International Settlements (BIS). More frequent estimates can be obtained by using data from surveys conducted twice a year by foreign exchange committees (FXCs), industry groups sponsored by central banks in various countries.

#### 2.1 Daily trading volumes and the geography of trading

The average daily trading volume in the global FX spot market has long been on an upward trend (Figure 1).<sup>3</sup> According to the latest Triennial, daily trading volume in the global spot FX market averaged \$2.11 trillion in April 2022.<sup>4</sup> Illustrating the global role of the US dollar as the main vehicle currency, the top three most traded currency pairs were EURUSD, with 23% of the trading volume, followed by USDJPY with 14% and GBPUSD with 10%.<sup>5</sup>

Figure 1 shows that FX spot trading is concentrated in a few large financial centres. London alone accounted for 38% in 2022, while the combined share of the top four trading centres, which also included New York, Singapore and Hong Kong, amounted to 74% of global spot FX turnover (BIS, 2022).<sup>6</sup> The major FX trading hubs are not necessarily in countries that dominate global goods trade, as FX trading for financial motives, such as investments in foreign-denominated securities, far exceeds the transaction volume related to international trade. In addition, the main data centers that host the matching engines that power electronic FX trading venues form a critical part of the global FX market infrastructure. This has

nation, see, e.g., Maggiori (2022) and the chapters in James, Marsh, and Sarno (2012).

<sup>&</sup>lt;sup>2</sup>As of the writing of this chapter, the last Triennial survey was conducted in April 2019.

<sup>&</sup>lt;sup>3</sup>The notable spike in trading volume in 2014 was related to monetary easing by the Bank of Japan, while the contraction in 2015 owed importantly to the de-risking that followed the Swiss National Bank's (SNB) surprise removal of the floor of the Swiss franc against the euro (Moore, Schrimpf, and Sushko, 2016).

<sup>&</sup>lt;sup>4</sup>Average daily trading volume across all OTC FX instruments was \$7.5 trillion in April 2022. Exchange-traded FX futures and options added less than \$0.2 trillion to that total.

<sup>&</sup>lt;sup>5</sup>Each currency is assigned a three-letter code (the "ISO 4217" code). Exchange rates are then represented by a currency pair, with the base currency listed first. For instance, GBPUSD, "sterling-dollar," is quoted in dollars per British pound sterling, while USDJPY, "dollar-yen," is quoted in yens per dollar.

<sup>&</sup>lt;sup>6</sup>The latter two Asian financial centres have gradually overtaken Tokyo as major trading hubs.

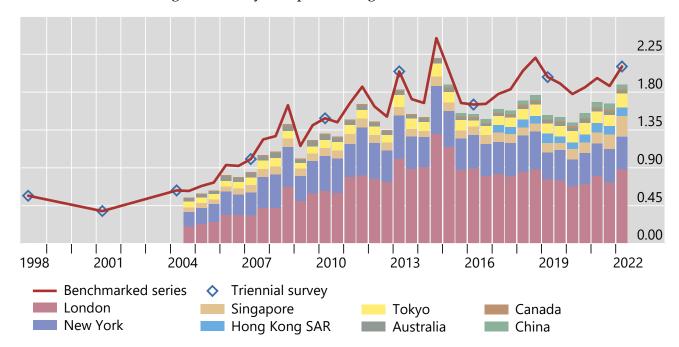


Figure 1: Daily FX spot trading volume, in \$ trillions

*Notes:* The semiannual FXC data used to create the benchmarked series are collected in April and October. Turnover from the China Foreign Exchange Trade System (CFETS) and from the Hong Kong Treasury Markets Association survey was added to the stacked bars beginning in April 2015 and April 2017, respectively.

Sources: BIS Triennial Central Bank Survey; Foreign Exchange Committee Surveys; CFETS.

incentivised some key players to co-locate their trading activity near these hubs.<sup>7</sup>

#### 2.2 Main types of counterparties

Figure 2 shows the share of trading volume of FX bank-dealers (the "reporting" dealers) with three broad counterparty categories covered by the central bank and BIS surveys: other bank-dealers, other financial institutions and non-financial customers. Non-financial customers, primarily corporations, use the FX market to support their core business activities, especially international trade. The broad category of "other financial institutions" has traditionally represented the financial customers of the reporting FX dealers. This includes institutional investors, asset managers, hedge funds, commodity trading advisors (CTAs), smaller banks that are not FX dealers, and central banks. More recently, principal trading firms (PTFs), often referred to as high-frequency traders (HFTs), have become important players within the category of "other financial institutions."

<sup>&</sup>lt;sup>7</sup>Among major data centers used for FX trading are London (LD4), New York (NY4), Tokyo (TY3), and, more recently, Singapore (SG1).

<sup>&</sup>lt;sup>8</sup>Note that many of the smaller banks trade FX directly with their own customers. Unlike the reporting dealers, however, they do not act as intermediaries in the global FX market.

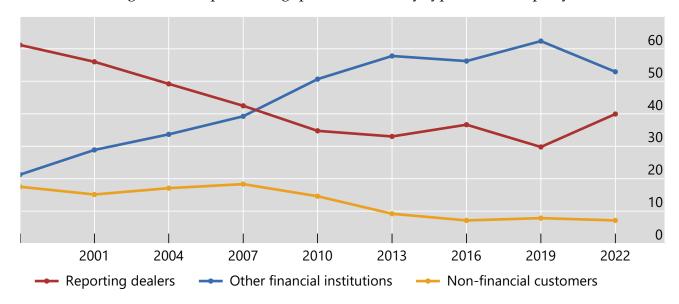


Figure 2: FX spot trading: percent shares by type of counterparty

Source: BIS Triennial Central Bank Survey.

Over the past 20 years, the share of global FX trading conducted with non-financial customers has declined from about 20% to less than 10%. In contrast, the share of financial customers, which used to be close to that of non-financial customers, has grown to more than 50%. This again reflects the fact that FX trading has become increasingly driven by the needs of financial customers as opposed to the needs arising directly from international trade.<sup>9</sup>

For a long time, bank-dealers constituted the sole sector that provided liquidity and warehoused risk for the rest of the market. The dealers were the archetypal liquidity providers (LPs) to financial and non-financial customers, which were the archetypal liquidity consumers (LCs). The dealers then traded among themselves to hedge their positions and rebalance their inventories, which generated much of the trading volume in the FX market. Figure 2 shows that in the late 1990s interdealer trading accounted for about two-thirds of all FX spot trading. The share of interdealer trading is now considerably lower. This is partly because dealers came to rely less on interdealer trading for inventory control and partly because the line between traditional LPs and LCs has become blurred. In particular, as PTFs have increased in importance in the FX market, some have taken on a dual role as both LPs and LCs, displacing some of the dealer-to-customer activity. For instance, according to a widely-followed yearly survey of the FX market by *Euromoney*, by 2022 PTFs accounted for almost a third of electronic FX trading with customers.<sup>10</sup>

<sup>&</sup>lt;sup>9</sup>Some retail investors also participate in the FX market, often trading with high leverage via retail margin brokers. But they generally represent a very small fraction of the trading activity of the "wholesale" market participants discussed earlier. Japan may be an exception, as FX margin trading by retail investors has become more substantial there (Mukoyama, Kikuta, and Washimi, 2018).

<sup>&</sup>lt;sup>10</sup>The results of these yearly surveys of the FX market, which *Euromoney* magazine has published for more

#### 3 The trading environment

The organisation of the FX market may best be understood in light of three key economic frictions: credit risk management, inventory risk management and asymmetric information. Along with advances in technology, these frictions have shaped the evolution of the FX market structure over time.

FX spot trades can involve very large sums and are settled two business days after the trade by the transfer of bank balances in the corresponding currencies ("T+2" settlement, discussed in Section 6.3). Hence, an FX trade has historically constituted a bilateral extension of credit, naturally leading to a market structure with banks at its core (Lyons, 2002). The introduction of FX prime brokerage (PB) in the late 1990s was a major advance in the management of credit. A service first offered to hedge funds by large FX dealers, PB enabled non-banks to conduct trades directly with a broad set of counterparties under the umbrella of its PB provider. This allowed a hedge fund to gain access to pricing and liquidity from a large number of bank-dealers solely on the basis of its credit relationship with its PB. Over time, prime brokerage has become widespread in the FX market. It allows various types of non-banks, including PTFs, to trade on a large number of venues, including some which were previously viewed as purely "interdealer." Prime brokerage has therefore contributed to the broadening of liquidity provision beyond bank-dealers and likely also to the proliferation of trading venues.

The management of inventory imbalances arising from trading with customers used to be a key driver of interdealer trading. The repeated passing of inventory imbalances between dealers in the market, dubbed "hot potato trading" (Lyons, 1996), contributed to the large interdealer share of total volume seen in the early years of Figure 2. However, the way bank-dealers managed their inventories began to change in the early to mid-2000s when the largest banks began to "internalise" some of their trades, waiting for an offsetting customer trade instead of immediately hedging the position in the interdealer market (Butz and Oomen, 2018). This was likely facilitated by technological advances as well as by the fairly high concentration of trading volume among the largest dealers at the time. The share of internalisation in major currency pairs is now estimated to have grown to 80% or more for the largest dealers (Moore et al., 2016; Schrimpf and Sushko, 2019). 11

Access to information shapes trading in all assets, and the structure of each market influences the speed of information aggregation. In the FX market, relevant information is dispersed among market participants, and the key intermediaries need to aggregate this information in order to set prices. Examples of relevant dispersed information are international portfolio allocation decisions by institutional investors, observations of the state of the economy in real time by firms (e.g., based on their knowledge of certain flows of imports and

than 40 years, are a very useful source of information on the global FX market, including about the relative importance of various liquidity providers. However, as many firms actively seek to be recognised in the survey, the results are best seen as indicative.

<sup>&</sup>lt;sup>11</sup>Butz and Oomen (2018) study how dealers adjust bid and ask quotes to influence the direction and arrival rate of customer orders, and they estimate that an average holding period before an offsetting trade (the "internalisation horizon") may be as short as one minute for a liquid pair like EURUSD. Barzykin, Bergault, and Guéant (2021) derive rules for dealer choices between internalisation and externalisation.

exports), or changes in risk preferences. Dealers then learn bits of dispersed information by observing the order flow of their various customers. Customers are not equally informed, and dealers profile them in order to better learn from the information content of their trades. Large banks that have a broad base of financial customers are better informed about FX developments than other banks (Bjønnes, Osler, and Rime, 2021; Menkhoff, Sarno, Schmeling, and Schrimpf, 2016; Ranaldo and Somogyi, 2021). It is also recognised that dealers bring their own independent information to the market (Moore and Payne, 2011).

The desire of market participants to manage the information revelation process in the market has likely been one of the factors driving the recent increase in the number of trading venues and available execution protocols. As we will discuss, these trading venues can come with important differences in the degree of disclosure of counterparty identities, the ability to assess a counterparty's price impact, or the option to tailor one's potential counterparties to a particular subset of market participants.

#### 3.1 The evolution of the FX trading environment over time

Figure 3 attempts to capture the evolution of the FX market structure over the last three decades. The upper panel, Figure 3a, depicts a stylized view of the market structure of the early to mid 1990s, Figure 3b the early to mid 2000s, and Figure 3c the most recent decade. The shaded area denotes the interdealer market, and the surrounding area denotes the dealer-to-customer market segment. Red arrows indicate voice execution, and blue arrows indicate electronic execution.

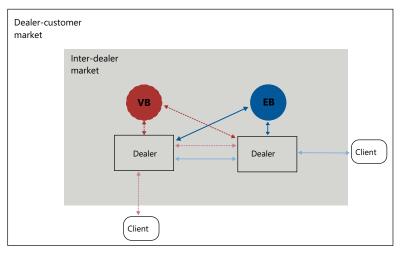
FX trading in the earlier period was often described as having a simple two-tier structure (Sager and Taylor, 2006). Customers traded directly with dealer banks, often by telephone or telex, in the "outer tier." Dealers were compensated in the form of bid-ask spreads and gained private information from the trades of their own customers (Lyons, 1996). Trading among dealers, either to manage inventory risk or for speculation, then constituted the "inner tier." Interdealer trading could take either a direct (bilateral) form or an indirect (brokered) form, initially only via voice brokers (VB). 12

The advent of two electronic brokers, by Reuters (now Refinitiv) and Electronic Broking Services (EBS), in the interdealer market in the early 1990s made the process of inventory risk management and price discovery far more centralized and efficient. The two electronic brokers, both organised as central limit order books (CLOBs), quickly became the main sources of price discovery and reference prices for the entire FX spot market, and therefore began to be referred to, jointly, as the "primary market" (we henceforth refer to them as "primary CLOBs"). <sup>13</sup>

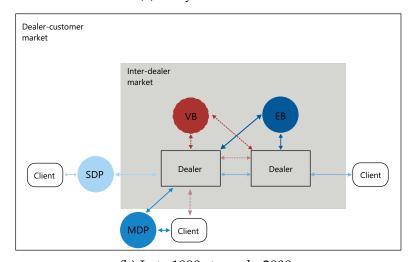
<sup>&</sup>lt;sup>12</sup>Besides telephone and telex, an electronic messaging platform called Reuters Direct Dealing System was also widely used for direct dealing. According to the 2022 Triennial, 23% of FX spot trading took place via a direct voice method, such as Bloomberg chat, and another 3% used a voice broker. Voice execution could be particularly advantageous to a customer seeking to execute a large trade.

<sup>&</sup>lt;sup>13</sup>EBS Market became the main venue for exchange rates such as EURUSD (and its predecessors), USDJPY, EURCHF, and, more recently, USDCNH, while Reuters Matching became the main venue for the "commonwealth" exchange rates, such as GBPUSD, USDCAD and AUDUSD, the scandinavian currencies, and most EME currencies.

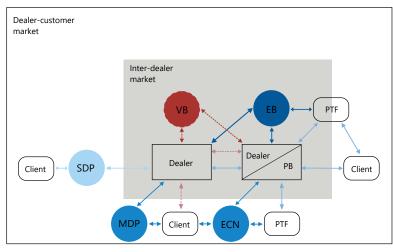
Figure 3: Evolution of the spot FX market



(a) Early to mid 1990s



(b) Late 1990s to early 2000s



(c) Recent decade

*Notes:* The figure builds on King, Osler, and Rime (2012); EB = electronic broker; MDP = multi-dealer platform; ECN = electronic communication network; PB = prime broker; PTF = principal trading firm; SDP = single-dealer platform; VB = voice broker.

The early 2000s offered new electronic trading opportunities in the dealer-to-customer segment, as shown in the middle panel (Figure 3b). Multi-dealer platforms (MDPs), likely the most important innovation at that time, enabled customers to submit a request for quote (RFQ) to multiple dealers simultaneously. MDPs significantly broadened choices for customers and introduced an important element of competition between dealers. In response, a number of individual FX dealer banks invested in proprietary single-dealer platforms (SDPs) for direct electronic trading. Over time, these SDPs came to replace telephone or telex in much of the bilateral trading. <sup>15</sup>

The lower panel, Figure 3c, provides a stylised view of the current market structure, which has continued to increase in complexity. The remains of the old two-tier structure can be seen, but the structure has been perturbed by two important innovations. First, the dealer-to-customer segment saw a proliferation of venues offering multiple types of execution protocols. This allowed an LP to provide liquidity to customers by streaming prices or by placing limit orders in an order book the latter resembling the trading environment of the interdealer market. We will refer to these venues collectively as electronic communications networks (ECNs), or "secondary ECNs." Second, PTFs have challenged the banks within the inner tier, both as LPs and LCs on the primary CLOBs. They have also, more recently, become active in the outer tier, where they stream prices to customers via ECNs. PTFs have thus established a firm foothold in the parts of the FX market that used to be exclusive domains of dealer-banks. At the same time, PTF trading in the FX market is enabled via prime brokerage arrangements with dealer-banks, including some of the very same banks they now compete with in the market.

#### 3.2 Market fragmentation and the declining role of the primary market

As shown in Figure 4, trading volume on the primary electronic brokers has declined substantially since the Global Financial Crisis, even as the overall volume in the FX spot market has increased. The emergence of a large number of alternative trading platforms and the aforementioned decline in interdealer trading of inventory imbalances have very likely contributed to this decline. Furthermore, some have cited the opening of the primary CLOBs to PTFs (via PB arrangements) as another factor. Some participants may indeed prefer to trade

<sup>&</sup>lt;sup>14</sup>An LC can also request a continuous stream of quotes for a given size and a certain amount of time (e.g., a trading day), known as a request for stream (RFS), effectively a continuous form of RFQ.

<sup>&</sup>lt;sup>15</sup>Although the number (and names) of platforms has changed over time, early examples of MDPs such as Currenex, Hostpot FX, and FXall continue to be important, and they have often diversified their offerings of execution methods. Banks have also responded to successful third-party MDPs by establishing FXSpotStream, which operates as a bank-owned consortium and provides multibank FX streaming and RFS service to customers. As for the SDPs, as of the writing of this chapter, examples include J.P. Morgan eXecute, Deutsche Bank AutobahnFX, UBS Neo, and Citi Velocity.

<sup>&</sup>lt;sup>16</sup>As of the writing of this chapter, examples of secondary ECNs include CboeFX, EuronextFX, LMAX and 360T, among others. For a comprehensive list of various trading venues, we refer the reader to https://www.marketfactory.com/venues/. For tractability, we will continue to refer to platforms with RFQ as the main execution protocol as MDPs.

<sup>&</sup>lt;sup>17</sup>According to the *Euromoney* FX survey, examples of customer-facing PTF liquidity providers are XTX Markets, Jump Trading, HC Technologies, and Citadel Securities.

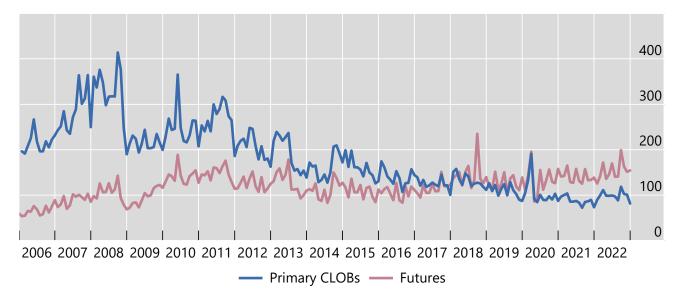


Figure 4: Primary CLOBs and currency futures exchanges, daily volumes \$ billions

*Notes:* Primary CLOBs refer to the wholesale electronic brokers EBS Market and Refinitiv Matching. In the latter part of the sample, the total volume may include other EBS platforms, such as EBS Direct. Exchange-traded futures volume mainly reflects that of the Chicago Mercantile Exchange, plus two smaller exchanges.

Sources: EBS; Refinitiv; BIS Exchange Traded Derivatives Statistics. January 2006 to December 2022.

on venues with less PTF participation, perhaps out of concern that their order flow may be quickly detected by the PTFs and that they may be subject to more adverse selection.

Despite their decline in trading volume, the primary CLOBs are still viewed by most market participants as a key locus of price discovery, and the market data generated by the EBS Market and Refinitv Matching platforms remain important to the market as a whole. Furthermore, when volatility spikes or market liquidity deteriorates on other venues, FX trading volume often tends to return to the EBS and Refinitiv CLOBs (Moore et al., 2016). This was the case in March 2020, for instance, when financial markets experienced turmoil at the beginning of the COVID pandemic.

Some LPs increasingly look to exchange-traded currency futures for reference prices and for hedging their spot activity. This appears to be the case particularly for PTFs, many of which entered the FX market with an already established business model in futures markets. But a growing number of market participants of all types now seems to consider currency futures traded on the Chicago Mercantile Exchange (CME) as at least a close cousin of the primary CLOBs. In contrast to the primary CLOBs, trading volume in currency futures has not declined. It now often exceed OTC FX spot trading volume in the primary market (Figure 4).

Figure 5 shows the total FX spot electronic trading volume over the past four Triennial surveys. Dark blue denotes multilateral platforms (CLOBs, ECNs, MDPs), while light blue denotes direct e-trading via SDPs. The declining trading volume of the primary CLOBs (which are part of the CLOBs/ECNs/MDPs category) is also shown. Recent growth in the dealer-to-customer segment has come from multilateral platforms, mostly at the expense of

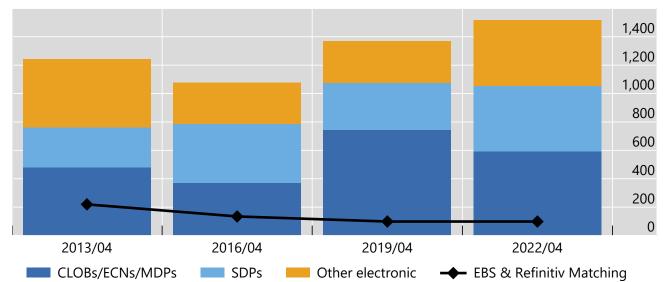


Figure 5: Total electronic spot trading volume, daily average \$ billions

*Notes*: Primary CLOBs refer to the wholesale electronic brokers EBS Market and Refinitiv Matching (in the latter part of the sample, the total may include other EBS platforms, such as EBS Direct). CLOB = central limit order book; ECN = electronic communication network; MDP = multi-dealer platform; SDP = single-dealer platform.

Sources: EBS; Refinitiv; BIS Triennial Central Bank Survey.

SDPs. This again reflects mainly the growth of ECNs offering a variety of trading protocols, as market participants can "shop around" in search of thebest execution. Illustrating the the degree of market fragmentation, Drehmann and Sushko (2022) show that customers can access liquidity via more than 30 MDPs and other ECNs, more than 20 SDPs, and via price streams from over a dozen PTFs. Faced with an increasing number of trading platforms and a greater variety of execution methods (such as RFQ, streaming and CLOBs), some LCs are also turning to liquidity aggregators. Liquidity aggregators are technology firms that assist their clients in accessing and comparing the various ECNs, MDPs and SDPs in real time, with the goals to obtain the best prices and maximise execution quality.

#### 3.3 Mapping the complex execution space

Figure 6 presents a stylized taxonomy of the variety of electronic trading venues currently used in the spot FX market. The classification of these venues and their associated trading mechanisms is based on two important dimensions: the level of pre-trade anonymity of the counterparties, and the level of "firmness" of the liquidity offered on the trading venues. This taxonomy oversimplifies what has become an increasingly complex market structure and likely misses some important exceptions, but it presents a useful way to consider the trading landscape.

As shown by the horizontal arrow, the level of pre-trade anonymity can range from fully disclosed to fully anonymous. Trading on an SDP, almost by definition, is fully disclosed, as the single LP knows the identity of all the potential LCs active on its platform. Similarly, on

an MDP where an LC submits an RFQ to several LPs, the identities of the potential counterparties are known before an actual trade takes place.

At the other end of the spectrum, the primary CLOBs are pre-trade fully anonymous, in the sense that displayed orders do not show a source and all participants can be matched based on the typical price and time priority rules of a CLOB. Note that the futures exchange, appearing to the side, offers the highest possible level of pre-trade (and post-trade) anonymity, basically by definition, as the exchange is the counterparty to each trade.

Secondary ECNs, in the middle, give their participants neither full pre-trade anonymity nor full pre-trade disclosure. This can be done, for instance, by assigning an alpha-numeric "tag" to each counterparty on the platform, a common practice on some ECNs. This may then allow a participant, having observed the execution quality of a trade against a counterparty with a particular tag, to decide whether to trade again with that unnamed counterparty. Furthermore, some other ECNs have created several separate CLOBs, each limited to a particular pool or tier of participants with a common trading style or set of characteristics. In that setting, while traders on each CLOB do not know ex-ante the identity of their counterparties, nor can follow them over time, they are assured to be matched only with a particular type of counterparty.

The second dimension in Figure 6 is the "firmness" of the liquidity offered on an electronic platform, i.e., the level of certainty of execution that follows a request to trade on a posted quote. This depends, importantly, on whether a practice known as "last look" is used by LPs or not. Last look is a process in which an LP, or more precisely the LP's computer, having received a request to trade on one of its posted quotes, takes a "last look" at the trade request before deciding whether to proceed with the transaction. The last look process occurs in a time period measured in milliseconds (Oomen, 2016; Cartea, Jaimungal, and Walton, 2019).

Last look was developed to address a number of potential issues arising from latencies within communication networks, the multiplicity of electronic platforms, and counterparties with different levels of technological sophistication. It has become broadly acceptable to use last look to conduct quasi-instantaneous "validity and price" checks, including verifying that sufficient bilateral credit is available and that the quoted price has not become stale. Other uses of last look, including the imposition of "additional hold time" before accepting or rejecting a trade, are much more controversial (see discussion in Section 5.1).

As shown in Figure 6, last look is not allowed on the primary CLOBs and liquidity there is therefore considered to be "firm." In contrast, last look is widely used on the secondary ECNs, with a few exceptions, and it is universally used on MDPs and SDPs, at least in the sense that LPs always reserve the option to use it. Liquidity on these platforms is therefore viewed as not completely firm: there is less certainty that a "hit" or request to trade on a quote will result in a trade.

Bringing the two dimensions of Figure 6 together, it is evident that a higher level of pretrade anonymity is associated with firmer liquidity. Anonymity and certainty of execution in an auction structure make the pre-trade environment on primary CLOBs closest to a full

<sup>&</sup>lt;sup>18</sup>Still, reflecting the OTC nature of the FX spot market, participants on the primary CLOBs can trade with each other only if they have established bilateral credit relationships, either directly or via their PBs. In practice most participants on the primary CLOBs can trade with each other.

Counterparties Tiered counterparty Counterparties **Futures** identified not identified pools or identification exchange before a trade using a "tag" before a trade **Fully Disclosed** Firm liquidity Primary (no last-look) **CLOBs** Secondary **ECNs** MDPs & SDPs Liquidity not firm (last-look allowed) Matching type: ← CLOB → ← CLOB/Streaming → RFQ/RFS —

Figure 6: A stylized taxonomy of main trading platforms

*Notes:* A stylized taxonomy of electronic trading mechanisms in the dimensions of pre-trade anonymity and the ability to use last look. Note that the ability to use last look may or may not be exercised by an LP. Primary CLOBs refer to the wholesale electronic brokers EBS Market and Refinitiv Matching. CLOB = central limit order book; ECN = electronic communication network; MDP = multi-dealer platform; SDP = single-dealer platform; RFQ = request for quote; RFS = request for stream.

exchange. This is likely one of the reasons why, despite their decline in trading volume, the prices shown on the primary CLOBs still tend to be viewed as reference prices for the entire FX market.

## 4 Algorithmic trading in FX

The proliferation of electronic execution in the spot FX market, now near 75% even for dealer-to-customer trading (BIS, 2022), has been accompanied by the growth of algorithmic (computer-driven) trading. A significant fraction of both LPs and LCs now use algorithms to drive at least some of their trading activity, with limited "human" intervention. The expansion of algorithmic trading in the FX market occurred first on the two primary CLOBs, but algorithmic trading has now expanded well beyond the primary market. <sup>19</sup>

Banks and non-banks deploy a variety of algorithms in the FX market. Execution al-

<sup>&</sup>lt;sup>19</sup>Because of the role that the primary CLOBs play in the market, algorithmic trading has had an important impact on price discovery in FX (Chaboud, Chiquoine, Hjalmarsson, and Vega, 2014).

gorithms attempt to minimize the price impact of large trades, market-making algorithms automatically post executable quotes, and opportunistic algorithms monitor the market in search of profit or hedging opportunities. An important distinction among these algorithms is whether they automate only the execution of the trade or whether the decision to trade is also left to the algorithm.

Figure 7 shows how the share of "manual trading" and "API-trading" on the EBS Market platform has evolved since 2004, illustrating the growth of algorithmic trading on one of the primary CLOBs, At the beginning of the sample, only banks were allowed on the platform and all trading was done manually by entering trading instructions on the specialized keyboard of an EBS terminal. After EBS allowed the direct interface of computers with its platform, banks began to trade algorithmically in 2004 (bank API). Non-banks began in 2005 (non-bank API), when they were first allowed on the EBS platform through PB arrangements. Note that, broadly speaking, the vast majority of the non-bank API trading shown currently represents the activity of PTFs, mostly using algorithms that automate both the decision to trade and the execution of the trade. In contrast, an important fraction of the Bank API activity likely reflects trades where the execution is automated but the initial decision to trade is taken by humans.<sup>20</sup>

As shown in Figure 7, by the end the sample, in 2022, algorithmic trading dominates. Bank API and non-bank API activity each account for a bit more than 40% of total trading volume, shares which have remained fairly even since 2015. The share of manual trading, in contrast, has continued to decline and now accounts for only about 15% of trading volume.<sup>21</sup>

Numerous types of algorithms are routinely used by banks and non-banks to provide and consume liquidity, as discussed earlier. Among those, "latency arbitrage" algorithms are a somewhat notorious sub-type of opportunistic algorithms which exploit a firm's speed advantage to generate profit opportunities.<sup>22</sup> Latency arbitrage algorithms are believed to be used almost exclusively by a subset of the PTFs active in the market. In an attempt to protect their clients from such strategies, particularly on platforms where manual traders and algorithmic traders coexist, several trading venues in the FX market have introduced various types of "speed bumps." ParFX, a platform launched in April 2013 by a group of large dealer banks, added randomised pauses (ranging from 20 to 80 milliseconds at the onset) to all trading instructions before processing them. EBS Market later in 2013 introduced a "latency floor" of a few milliseconds, short periods during which all trading messages are batched and randomized before being released to the order book. In 2014, Reuters/Refinitiv Matching, the other primary CLOB, introduced its own version of a speed bump.<sup>23</sup>

<sup>&</sup>lt;sup>20</sup>For example, a bank may decide to cover some of the order flow on its own SDP by trading on a primary CLOB using an execution algorithm.

<sup>&</sup>lt;sup>21</sup>The evolution of these shares over time has some notable features. For instance, manual trading drops in early 2020 as the COVID pandemic begins, probably reflecting the increase in working-from-home arrangements, bank algorithmic activity rises beginning in late 2014, likely due to the switch to algorithmic execution around fixings by a number of banks after the fixing scandal (see Section 6.2), and non-bank algorithmic activity peaks in 2011 after EBS reduces its tick size (Chaboud, Dao, and Vega, 2021).

<sup>&</sup>lt;sup>22</sup>For instance, a trader with faster technology than a certain LP may take advantage of the small time lag between when a market-moving trade takes place and when the LP updates its posted quotes in response, quickly trading on one of these now outdated quotes.

 $<sup>^{23}</sup>$ The speed bump introduced by Refinitiv was viewed as an evolution of the EBS speed bump. In particular,

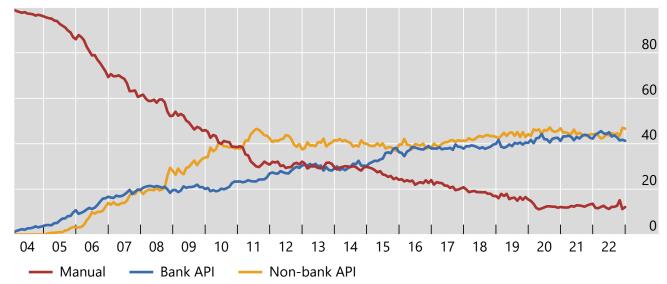


Figure 7: Manual vs. algorithmic execution on EBS Market

*Notes:* The Figure shows the proportion of trading volume on EBS Market executed manually versus algorithmically by banks and by non-banks.

Sources: EBS.

Finally, while execution algorithms have long been used by banks to optimise the execution of large trades, such algorithms are now also being used directly by some of the more sophisticated customers in the FX market. As described in a recent report on the use of execution algorithms in FX (Bank for International Settlements, 2020), sophisticated customers increasingly rely on smart order routing and execution algorithms to spread large orders over time and across multiple electronic venues. Customers in FX also increasingly use Transaction Cost Analysis (TCA) to monitor the execution quality of their trades.

#### 5 The role of the official sector

#### 5.1 Regulation, good practice, the FX Global Code

The level of regulatory oversight of the foreign exchange market varies greatly from country to country, depending importantly on the country's exchange rate regime and the restrictions on cross-border capital movements .<sup>24</sup> The regulatory authority can reside with the central bank, the ministry of finance, or another government agency. In the case of major floating exchange rates (often called the G-10 exchange rates), regulatory constraints are generally light compared to equity and bond markets. This is likely due in great part to the fact that an

the Refinitiv mechanism processes order cancellation messages before other types of messages when releasing each batch to the order book, providing additional protection against aggressive latency arbitrage strategies. That feature was later also adopted by EBS.

<sup>&</sup>lt;sup>24</sup>The IMF's *Annual Report on Exchange Arrangements and Exchange Restrictions* is a comprehensive source of information on that topic.

FX transaction does not entail the exchange of a security and that, by nature, the currencies of two different countries are involved. Even when they do not have formal regulatory authority over the FX market, many central banks sponsor Foreign Exchange Committees that provide a forum for a variety of market participants to meet regularly to discuss market functioning and encourage "good practice." In addition, the central banks of countries with major trading centres meet regularly under the auspices of the BIS to discuss the global foreign exchange market.

It is under this framework that a broad group of central banks and market participants cooperated to develop the FX Global Code, first published in 2017 and later updated in 2021. The FX Global Code, which is maintained by the Global Foreign Exchange Committee (GFXC) lists and discusses principles of good practice for the global FX market.<sup>25</sup> Specifically, the Code contains 55 principles of good practice in the areas of ethics, governance, execution, information sharing, risk and compliance, and confirmation and settlement. Market participants are encouraged to sign a "statement of commitment" to the Code to publicly signal their intention to adhere to its principles. The Code has gained broad international acceptance, particularly among large bank-dealers, and it is increasingly viewed as having an important influence on the behavior of FX market participants.

For instance, the 2021 update of the Code was accompanied by guidance on the proper use of last look. The guidance reaffirmed that last look should only be used for price and validity checks, and recommended that these checks should be applied "without delay." Market participants viewed that recommendation as advocating against a widely-used, but controversial, practice where some liquidity providers wait to observe additional price movements before accepting or rejecting a trade. While LPs argue that this "additional hold time" protects them against possible adverse selection, opponents point out that it allows LPs to only accept the most profitable trades. In the event, the guidance by the GFXC had a rapid and substantial impact on the behavior of market participants: numerous large bank-dealers announced an end to additional hold time, and some trading venues announced a reduction in the maximum length of their "last look window."

In the end, while still a fairly new experiment, the effect of the FX Global Code on the FX market is seen by some as evidence that a "soft" approach to regulation, which encourages good practice instead of imposing hard rules, can have a substantial impact on an important financial market.<sup>26</sup>

#### 5.2 FX intervention and exchange rate management

The official sector, mainly central banks and finance ministries, often has a critical influence on floating exchange rates.<sup>27</sup> Participants in FX markets are closely attuned to official macroeconomic data releases and central bank communications, and major exchange rates

<sup>&</sup>lt;sup>25</sup>The GFXC was created in 2017, and its members are the various FXCs from around the world. The Code is available on its website: http://www.globalfxc.org.

<sup>&</sup>lt;sup>26</sup>Skeptics may argue that the "soft" approach to regulation is only effective because deviating from accepted good practice could open a financial firm to civil legal liability.

<sup>&</sup>lt;sup>27</sup>The official sector obviously also plays a critical, but different, role in cases where countries choose to peg or closely manage the exchange value of their currency.

routinely react within milliseconds to the surprise component of these news. In addition, a more direct and willful way in which the official sector can affect exchange rates is via foreign exchange intervention, the purchase or sale of its currency in the FX market with the specific intent to influence exchange rates. Many major industrialized countries with floating exchange rates, including the United States, used FX intervention fairly frequently until the mid-1990s.<sup>28</sup> Since then, with a few notable exceptions, including Japan in the early 2000s and Switzerland in the years since the Global Financial Crisis, FX intervention by major industrialized countries has become rare.<sup>29</sup> FX intervention by developing and newly-industrialized countries has, in contrast, remained quite frequent.<sup>30</sup>

A distinction is often made between "sterilised" and "unsterilised" FX intervention, where sterilising an intervention operation means offsetting the associated effects on the domestic monetary base, typically by buying or selling bonds. Historically, most FX intervention operations have been sterilised. By design, unsterilised FX intervention is as much monetary policy as it is foreign exchange policy.<sup>31</sup>

A substantial academic literature has discussed the effects of FX intervention. Earlier work focused mainly on interventions by major industrialised countries (Dominguez and Frankel, 1993), but recent work has expanded to a study of the impact of FX intervention among a broader range of countries (Fratzscher, Gloede, Menkhoff, Sarno, and Stöhr, 2019), including countries which implement a "managed float" for their currencies (Frankel, 2019; Cavallino, 2019). Broadly speaking, research has traditionally discussed three channels through which FX intervention may affect exchange rates: a portfolio balance channel, when investors view domestic and foreign assets as imperfect substitutes; a signaling channel, where FX intervention is interpreted by the market as conveying information about future central bank policy; and a coordination channel, where intervention activity affects the trading behavior of FX traders in a way that reinforces the central bank's activity.

There is broad agreement that FX intervention, particularly if it is large, can have an immediate effect on exchange rates, but there is little consensus about the medium and long-term impact of FX intervention on exchange rates. In general, research has found that announcing an FX intervention, intervening in a direction consistent with monetary policy, and coordinating intervention among central banks can enhance the short-term impact of FX intervention. In addition, the impact of FX intervention in currencies outside of the few major currencies often seems to be larger and longer lasting. This suggests that FX interventions can have a larger impact in the presence of market imperfections, such as limited interme-

<sup>&</sup>lt;sup>28</sup>During that era, the intervention episodes that followed the 1985 Plaza Accord, an agreement among G5 countries to weaken the dollar, and the 1987 Louvre Accord, a G7 agreement to stabilise dollar exchange rates, are particularly well known. Several European countries also intervened heavily in the 1992 as the European Exchange Rate Mechanism (ERM) came under pressure, especially just before Great Britain withdrew from the ERM.

<sup>&</sup>lt;sup>29</sup>Japan has intervened in FX markets on several occasions since the 2000s, including in 2011 following the earthquake and nuclear accident and, most recently, in ate 2022.

<sup>&</sup>lt;sup>30</sup>An increasing number of central banks in emerging markets also conduct FX interventions using derivatives, allowing them to economise on FX reserves.

<sup>&</sup>lt;sup>31</sup>The recent intervention purchases of foreign currency by the SNB are an important example. In that case, the SNB specifically decided not to sterilise its intervention operations, resulting in a large increase in the size of its balance sheet, in essence conducting quantitative easing via the purchase of foreign currency.

diation capacity (Gabaix and Maggiori, 2015) and a number of other financial frictions that render domestic and foreign assets imperfect substitutes (Popper, 2023).

## 6 Special topics

#### 6.1 Flash events and other extreme events

Similar to several other financial markets with a large share of electronic trading, the FX market has in recent years experienced a number of flash events, with very sharp but often quickly retraced exchange rate movements accompanied by a temporary but almost complete disappearance of market liquidity. The analysis of these events can provide insights into the behavior of market participants and into the impact of specific aspects of the market structure. Among prominent recent events, on 7 October 2016 the GBP quickly depreciated almost 9% against the USD before recovering much of its losses within minutes (Bank for International Settlements, 2017); then on 3 January 2019, the JPY appreciated sharply against the USD and a number of other currencies before quickly retracing much of the move.

The two events share some interesting characteristics. They both occurred in the late afternoon in the United States (early morning in Asia), a time of day when the FX market typically exhibits its lowest trading volume and liquidity. In both cases, the linkages between the spot and the futures market seemed to be at play. Liquidity in the GBP fell to its lowest level precisely when the futures contracts at the CME experienced a trading halt in response to a large, but still fairly orderly, price movement. The JPY event occurred during the one hour of the day when there is no futures trading on the CME.<sup>32</sup> The two flash events also have some individual peculiarities. In particular, the JPY event involved the unwinding by retail FX investors of positions in higher-yielding currencies such as the Turkish lira and the Australian dollar.<sup>33</sup>

Finally, what is viewed as the most extreme event in the FX market in recent years was not a flash event per se, as it had a clearly-understood trigger. On 15 January 2015, the SNB removed the floor that it had imposed on the EURCHF exchange rate since 2011, a floor that had effectively prevented the appreciation of the CHF (Breedon, Chen, Ranaldo, and Vause, 2022). The SNB had maintained that floor by routinely intervening in the FX market, purchasing euros in an amount sufficient to keep the exchange rate at or above 1.20 Swiss francs per euro. The removal of the floor had not been expected by market participants (Mirkov, Pozdeev, and Söderlind, 2106), and the Swiss franc quickly appreciated up to 40% against the euro, forcing the SNB to resume intervention purchases of euros after about 20 minutes. The Swiss franc ended the day about 15 percent stronger against the euro. The impact of the associated volatility on leveraged FX positions led to a widespread review of risk-management practices in the FX industry.

 $<sup>^{32}</sup>$ FX futures at the CME trade 23 hours a day five days a week. Trading stops for one hour between 5 p.m. and 6 p.m. ET.

<sup>&</sup>lt;sup>33</sup>The retail FX sector in Japan is larger than in other developed countries, with a substantial number of retail FX margin traders. When the yen appreciates suddenly, currency moves can be amplified by triggering stop-loss orders (Tomohiro Niimi, 2016).

#### 6.2 Foreign exchange benchmark rates and the "fixing scandal"

Benchmark exchange rates, also commonly called "fixes," are calculated and published numerous times a day. They are widely used for a number of purposes, including valuing portfolios that contain foreign-currency denominated securities, constructing multi-country indices, such as the MSCI stock indices, and even settling some derivatives transactions. The best-known and most widely used of these benchmark rates is the one calculated at 4 p.m. in London by WMR.<sup>34</sup> Because some important FX market participants, such as asset managers, routinely request transactions executed at the 4 p.m WMR rates, the FX market experiences a substantial daily spike in trading volume around that time of the day. Especially large volumes occur at the end of each month, associated with the rebalancing of multi-country portfolios. There are other prominent daily fixes in the FX market, including the ones calculated by the ECB at 2:15 p.m. CET, and the Tokyo morning fixes calculated by various Japanese banks at 9:55 a.m. (Financial Stability Board, 2014; Ito and Yamada, 2017).

The role of these fixes, with very few exceptions (Melvin and Prins, 2015), attracted very little attention outside of the industry before 2013. This changed after accusations emerged in 2013 that some traders at large banks may be manipulating benchmark exchange rates (Vaughan, Finch, and Choudhury, 12 June 2013). In the end, what became known as the FX fixing scandal resulted in substantial fines being imposed on dealing banks in a number of countries, and reforms of the fixing process were implemented in 2015 after a study by the Financial Stability Board (Financial Stability Board, 2014, 2015). Among other reforms, the window over which the 4 p.m. fix is calculated was lengthened from one minute to five minutes with the goal to reduce the potential for manipulation.<sup>35</sup>

Despite the 2015 reforms, the issue of benchmark exchange rates continues to be a topic of concern. In particular, traders have pointed to ongoing unusual volatility patterns around the 4 p.m. fix, especially at month ends, leading to some calls for further reforms. The issue of FX benchmarks has also generated some academic research, with analysis focusing on the initial issue (Evans, 2018; Ito and Yamada, 2017; Osler and Turnbull, 2017) and on the impact of the fixing reform (Evans, O'Neill, Rime, and Saakvitne, 2018).

#### 6.3 Settlement risk

Transactions in the "spot" FX market, basically by definition, settle by the exchange of balances in one currency against balances in another. For the vast majority of currency pairs, settlement occurs two business days after the trading date, that is at "T+2."<sup>36</sup> FX settlement risk is the risk of loss incurred by a counterparty when it pays out the currency it sold but does not receive the currency it bought.

<sup>&</sup>lt;sup>34</sup>WM/Refinitiv (previously known as WM/Reuters) calculates benchmark rates every hour for a large number of spot, forward, and NDF exchange rates.

<sup>&</sup>lt;sup>35</sup>In late 2015, in an effort to discourage trading activity related to its daily FX reference rates, and thereby reduce the potential for manipulation, the ECB delayed their publication from 2:30 p.m. to 4:00 p.m. CET.

<sup>&</sup>lt;sup>36</sup>Transactions involving a handful of currency pairs settle at T+1, most prominently USDCAD. By convention, the "value date" for almost all FX trades around the world changes from one day to the next at 5 p.m. ET. Trades involving the New Zealand dollar are an exception.

The bankruptcy of Bankhaus Herstatt in 1974 demonstrated how FX settlement risk can undermine financial stability. Herstatt was a medium-sized German bank active in FX markets. On 26 June 1974, the German authorities closed the bank down. While Herstatt had already received Deutsche marks from its counterparties, it had not yet made the corresponding US dollar payments in New York. Herstatt's failure to pay led a number of banks to stop outgoing payments until their own payments due had been received. The consequences were systemic: the international payment system was severely affected for a period of time, and the erosion of trust caused lending rates to spike and credit to be curtailed.

The G-10 central banks launched in 1996 a comprehensive strategy to reduce FX settlement risk, resulting in the launch of Continuous Linked Settlement (CLS) in 2002. CLS is a specialized institution that settles FX transactions on a payment-versus-payment (PvP) basis, addressing settlement risk by ensuring that a payment in a currency occurs if and only if the payment in the other currency takes place.<sup>37</sup>

The establishment of CLS and other actions led to a substantial reduction in FX settlement risk. Kos and Levich (2016) estimate that the share of FX turnover settled by means of traditional correspondent bank arrangements declined from 85% to 13% between 1997 and 2013. The share of global FX transactions settled through CLS reached about 50% in 2013, but more recently the pace of adoption of PvP settlement appears to have slowed (Bech and Holden, 2019). Among likely causes are the growth in trading volume in emerging market currencies that are not CLS-eligible and the internalisation of customer flow by LPs, in which case the dealer is likely to settle trades across its own books.

There has been a renewed effort in recent years to reduce the remaining FX settlement risk in the system, including efforts to bring PvP settlement to additional currencies. In addition, the 2021 update of the FX Global Code now calls for market participants to use PvP settlement where available.

## 7 Topics for future research

The foreign exchange market has gone through major changes since the early 2000s, and this chapter concludes with a few open questions for future research grouped under three broad topics: Liquidity and market fragmentation; new market intermediaries and information; and the future of the FX market.

#### 7.1 Liquidity and market fragmentation

The decline in activity on the interdealer electronic brokers and the rapid growth in the number of electronic trading venues and protocols, particularly in the dealer-to-customer segment, have resulted in an increasingly fragmented trading environment. How has this affected liquidity in the market overall? Moreover, in a market where trading activity is widely dispersed, how does one properly measure market liquidity?<sup>38</sup>

<sup>&</sup>lt;sup>37</sup>Several regional PvP systems, such as Hong Kong CHATS, have also been introduced.

<sup>&</sup>lt;sup>38</sup>Research on FX market liquidity has frequently been limited by data availability. Using Reuters data, Evans and Lyons (2002) find that FX liquidity is time-varying and is inversely related to the amount of public

In such a fragmented market, market depth aggregated across all venues likely overstates the true depth of the market, resulting in a "liquidity mirage." The existence of last look further exacerbates this issue, because it allows LPs to more easily post the same liquidity simultaneously across many venues, knowing that they can reject some requests to trade. In such an environment, measures of liquidity besides bid-ask spreads and depth, such as estimated price impacts, may therefore be more useful.

Last look also allows LPs to post tighter bid-ask spreads, as these spreads no longer have to compensate them for the risk of an adverse price movement before a quote is hit. Importantly, this means that the FX market is now populated by a mixture of firm and non-firm liquidity, and the liquidity subject to last look is likely being offered at a better price. Customers trying to access that cheaper liquidity may however not be able to complete the trade. What is the "true" bid-ask spread in such a market?

Looking at the bigger picture, could the non-firm liquidity eventually drive out the firm liquidity, or is the coexistence of firm and non-firm liquidity a stable equilibrium? The answer to this question must depend importantly on the amount of information available to LCs, allowing them to evaluate the trade-off between the certainty of execution and the price. Finally, if liquidity is a priced risk factor, reflecting the risk of illiquidity when execution is most needed, could last look also impact the level of exchange rates?

#### 7.2 New intermediaries and price discovery

Changes in the composition of market intermediaries in the FX spot market also raise interesting research questions. As in several other important financial markets, the bank dealers now compete in the role of main liquidity providers with PTFs. The PTFs are lightly-capitalized non-banks that rely more heavily on speed and correlations with other asset classes to hedge their positions.

How has the risk-bearing capacity of FX market intermediaries, in total, changed as a result? Some observers argue that the market's risk-bearing capacity must have declined, even as average trading costs have likely declined as well. If so, is there evidence that this has affected market volatility or the frequency of flash events?

The evolution of market intermediaries also raises questions about information and price discovery in this market. The academic literature has long documented the explanatory power of certain customer order flows for exchange rate movements, information that only bank-dealers typically have had access to. Has the growing importance of PTFs reduced the explanatory power of customer order flow? Furthermore, has the presence of PTFs in this market lowered the incentives of bank-dealers to seek and study fundamental information that may affect the evolution of exchange rates, as in Weller (2018)?

There are also important questions about where price discovery occurs in this increasingly fragmented market. Do the primary CLOBs, despite their decreased trading volume,

information transmitted to the market. Mancini, Ranaldo, and Wrampelmeyer (2013) use EBS data to document that FX market liquidity co-moves with other financial markets, with common effects across currency pairs. More recently, Hasbrouck and Levich (2019) combine CLS data with commercial FX quotes data to design more comprehensive test of liquidity commonality.

still constitute the main locus of price discovery in the FX market? This chapter has discussed factors that argue in favor of that hypothesis. Alternatively, has price discovery, at least in part, moved to other sections of the market? The smaller size of the primary CLOBs, together with the increased role of PTFs who often rely on correlations with futures to price their spot liquidity, has led some market participants to argue that FX futures do, in some cases, lead price discovery in FX spot.<sup>39</sup> Can one envision in the future a highly fragmented FX spot market with price discovery occurring mainly in FX futures?<sup>40</sup>

#### 7.3 The future of the FX market

The structure of the global FX market has been in a constant state of flux for many years. An interesting aspect of the FX market is that it is generally under very light regulatory oversight, particularly as far as how trading venues operate. As a result, the structure and operation of the FX market have mostly developed endogenously, driven by commercial interests and the market participants' needs. For instance, platforms can easily set and change parameters such as the minimum trade size, minimum price increment, whether last look is allowed, the level of trader anonymity, the content and speed of data updates, etc..., as they compete to attract trading volume. The FX market should therefore provide a rich environment where both theoreticians and empirical economists can address questions such as what factors drive the evolution of a market and which platform characteristics are important to lead price discovery.

Finally, although much is speculative, the effect of new financial technologies (blockchain, cryptocurrencies, decentralized finance) on the FX market will be an important topic of research in coming years. This will particularly be the case if central bank digital currencies (CBDC) are issued by major jurisdictions around the world. This could affect not just the settlement of FX transactions, <sup>41</sup> but also impact price discovery and the role of the current intermediaries. That, in turn, could eventually lead to fundamental changes in the structure of the foreign exchange market.

#### References

Bank for International Settlements, 2017. The sterling "flash event" of 7 October 2016. Markets Committee Papers 9.

<sup>&</sup>lt;sup>39</sup>Chaboud et al. (2021) show some evidence of this occurring as early as 2011. Comparing price discovery in FX spot and futures markets used to be an active area of research (e.g Tse, Xiang, and Fung (2006), Rosenberg and Traub (2009), and Chen and Gau (2010).

<sup>&</sup>lt;sup>40</sup>Note that, to a large extent, this occurs in the US equity market (with E-mini S&P 500 futures) and the German sovereign bond market (with Bund futures).

<sup>&</sup>lt;sup>41</sup>For instance, the BIS Innovation Hub recently conducted a cross-border settlement of EUR vs. CHF FX transaction between a French and a Swiss commercial bank using wholesale CBDC issued by the Banque de France and the Swiss National Bank (Banque de France and Bank for International Settlements and Swiss National Bank, 2021).

- Bank for International Settlements, 2022. *Triennial Central Bank Survey of Foreign Exchange and Over-the-counter (OTC) Derivatives Markets in* 2022, October.
- Bank for International Settlements, 2020. FX execution algorithms and market functioning. Markets Committee Papers 13.
- Banque de France and Bank for International Settlements and Swiss National Bank, 2021. Cross-border settlement using wholesale CBDC. Project jura, BIS Innovation Hub.
- Barzykin, A., P. Bergault, O. Guéant, 2021. Market making by an FX dealer: tiers, pricing ladders and hedging rates for optimal risk control. Preprint arXiv:2112.02269, arXiv.
- Bech, M. L., H. Holden, 2019. FX settlement risk remains significant. *BIS Quarterly Review*: 48–49, December.
- Bjønnes, G. H., C. L. Osler, D. Rime, 2021. Price discovery in two-tier markets. *International Journal of Finance and Economics* 26:3109–3133.
- Breedon, F., L. Chen, A. Ranaldo, N. Vause, 2022. Judgement day: Algorithmic trading around the Swiss franc cap removal. *Journal of International Economics* 140.
- Butz, M., R. Oomen, 2018. Internalisation by electronic FX spot dealers. *Quantitative Finance* 19 1:35–56.
- Cartea, Á., S. Jaimungal, J. Walton, 2019. Foreign exchange markets with last look. *Mathematics and Financial Economics* 13 1:1–30.
- Cavallino, P., 2019. Capital flows and foreign exchange intervention. *American Economic Journal: Macroeconomics* 11 2:127–70.
- Chaboud, A. P., B. Chiquoine, E. Hjalmarsson, C. Vega, 2014. Rise of the machines: Algorithmic trading in the foreign exchange market. *Journal of Finance* 69 5:2045–2084.
- Chaboud, A. P., A. Dao, C. Vega, 2021. What makes HFT tick? Tick size changes and information advantage in a market with fast and slow traders. Available on SSRN.
- Chen, Y.-L., Y.-F. Gau, 2010. News announcements and price discovery in foreign exchange spot and futures markets. *Journal of Banking and Finance* 34 7:1628–1636.
- Dominguez, K. M., J. A. Frankel, 1993. *Does Foreign Exchange Intervention Work?* Columbia University Press.
- Derhmann, M., V. Sushko, 2022. The global foreign exchange market in a higher-volatility environment. *BIS Quarterly Review*: 33–48, December.
- Evans, M. D., 2018. Forex trading and the WMR fix. Journal of Banking and Finance 87:233–247.
- Evans, M. D., P. O'Neill, D. Rime, J. Saakvitne, 2018. Fixing the fix? assessing the effectiveness of the 4pm fix benchmark. Occasional Paper 46, FCA.

- Evans, M. D. D., R. K. Lyons, 2002. Time-varying liquidity in foreign exchange. *Journal of Monetary Economics* 49 5:1025–1051.
- Financial Stability Board, 2014. Final Report on Foreign Exchange Benchmarks, September.
- Financial Stability Board, 2015. Foreign exchange benchmarks: Report on progress in implementing the September 2014 recommendations, October.
- Frankel, J., 2019. Systematic managed floating. Open Economies Review 30 2:255–295.
- Fratzscher, M., O. Gloede, L. Menkhoff, L. Sarno, T. Stöhr, 2019. When is foreign exchange intervention effective? evidence from 33 countries. *American Economic Journal: Macroeconomics* 11 1:132–56.
- Gabaix, X., M. Maggiori, 2015. International liquidity and exchange rate dynamics. *The Quarterly Journal of Economics* 130 3:1369–1420.
- Hasbrouck, J., R. M. Levich, 2019. FX liquidity and market metrics: New results using CLS Bank settlement data. Typescript, Stern NYU.
- Ito, T., M. Yamada, 2017. Puzzles in the Tokyo fixing in the forex market: Order imbalances and bank pricing. *Journal of International Economics* 109:214–234.
- James, J., I. W. Marsh, L. Sarno (Eds.), 2012. The Handbook of Exchange Rates. Wiley.
- King, M. R., C. L. Osler, D. Rime, 2012. Foreign exchange market structure, players and evolution. In: James et al. (2012).
- King, M. R., D. Rime, 2010. The \$4 trillion question: What explains FX growth since the 2007 survey? *BIS Quarterly Review* 4:27–42.
- Kos, D., R. M. Levich, 2016. Settlement risk in the global FX market: How much remains? Typescript, Stern NYU.
- Lyons, R. K., 1996. Foreign exchange volume: Sound and fury signifying nothing? In: J. A. Frankel, G. Galli, A. Giovannini (Eds.), *The Microstructure of Foreign Exchange Markets*, University of Chicago Press, Chicago.
- Lyons, R. K., 2002. The future of the foreign exchange market. In: R. E. Litan, R. J. Herring (Eds.), *Brookings-Wharton Papers on Financial Services*, Brookings Institution.
- Maggiori, M., 2022. International macroeconomics with imperfect financial markets. In: K. R. Elhanan Helpman, G. Gopinath (Eds.), *Handbook of International Economics*, Elsevier, vol. 6 of *Handbook of International Economics*, Elsevier.
- Mancini, L., A. Ranaldo, J. Wrampelmeyer, 2013. Liquidity in the Foreign Exchange Market: Measurement, Commonality, and Risk Premiums. *Journal of Finance* 68 5:1805–1841.
- Melvin, M., J. Prins, 2015. Equity hedging and exchange rates at the london 4 p.m. fix. *Journal of Financial Markets* 22:50–72.

- Menkhoff, L., L. Sarno, M. Schmeling, A. Schrimpf, 2016. Information flows in foreign exchange markets: Dissecting customer currency trades. *Journal of Finance* 71 2:601–634.
- Mirkov, N., I. Pozdeev, P. Söderlind, 2106. Toward removal of Swiss franc cap: Market expectations and verbal interventions. Working Paper 2016-10, Swiss National Bank.
- Moore, M., A. Schrimpf, V. Sushko, 2016. Downsized FX markets: Causes and implications. *BIS Quarterly Review*: 35–51, December.
- Moore, M. J., R. Payne, 2011. On the sources of private information in FX markets. *Journal of Banking and Finance* 35 5:1250–1262.
- Mukoyama, Y., N. Kikuta, K. Washimi, 2018. Investment Patterns of Japanese Retail Investors in Foreign Exchange Margin Trading. Bank of Japan Review Series 18-E-3.
- Oomen, R., 2016. Last look. Quantitative Finance 17 7:1057-1070.
- Osler, C. L., A. Turnbull, 2017. Dealer trading at the fix. Working Paper 101R, Brandeis University.
- Popper H., 2023. Foreign Exchange Intervention. Working Paper, Santa Clara University.
- Ranaldo, A., F. Somogyi, 2021. Asymmetric information risk in FX markets. *Journal of Financial Economics* 140 2:391–411.
- Rosenberg, J. V., L. G. Traub, 2009. Price discovery in the foreign currency futures and spot market. *Journal of Derivatives* 17 2:7–25.
- Sager, M. J., M. P. Taylor, 2006. Under the microscope: the structure of the foreign exchange market. *International Journal of Finance & Economics* 11 1:81–95.
- Schrimpf, A., V. Sushko, 2019. FX trade execution: complex and highly fragmented. *BIS Quarterly Review*: 39–51, December.
- Tomohiro Niimi, 2016. Recent Trends in Foreign Exchange (FX) Margin Trading in Japan. Bank of Japan Review 2016-E-5.
- Tse, Y., J. Xiang, J. K. W. Fung, 2006. Price discovery in the foreign exchange futures market. *Journal of Futures Markets* 26 11:1131–1143.
- Vaughan, L., G. Finch, A. Choudhury, 12 June 2013. Traders said to rig currency rates to profit off clients. *Bloomberg News*.
- Weller, B. M., 2018. Does algorithmic trading reduce information acquisition? *Review of Financial Studies* 31 6:2184–2226.

#### Previous volumes in this series

1093 April 2023	Sovereign risk and bank lending: evidence from the 1999 Turkish earthquake	Yusuf Soner Başkaya, Bryan Hardy, Şebnem Kalemli-Özcan and Vivian Yue
1092 April 2023	Mobile Payments and Interoperability: insights from the academic literature	Milo Bianchi, Matthieu Bouvard, Renato Gomes, Andrew Rhodes and Vatsala Shreeti
1091 April 2023	FinTech, investor sophistication and financial portfolio choices	Leonardo Gambacorta, Romina Gambacorta and Roxana Mihet
1090 April 2023	Tackling the fiscal policy-financial stability nexus	Claudio Borio, Marc Farag and Fabrizio Zampolli
1089 April 2023	Intraday liquidity around the world	Biliana Alexandrova Kabadjova, Anton Badev, Saulo Benchimol Bastos, Evangelos Benos, Freddy Cepeda-Lopéz, James Chapman, Martin Diehl, Ioana Duca-Radu, Rodney Garratt, Ronald Heijmans, Anneke Kosse, Antoine Martin, Thomas Nellen, Thomas Nilsson, Jan Paulick, Andrei Pustelnikov, Francisco Rivadeneyra, Mario Rubem do Coutto Bastos and Sara Testi
1088 April 2023	Big techs and the credit channel of monetary policy	Fiorella De Fiore, Leonardo Gambacorta and Cristina Manea
1087 April 2023	Crypto carry	Maik Schmeling, Andreas Schrimpf and Karamfil Todorov
1086 April 2023	CBDC policies in open economies	Michael Kumhof, Marco Pinchetti, Phurichai Rungcharoenkitkul and Andrej Sokol
1085 March 2023	Supervisory policy stimulus: evidence from the Euro area dividend recommendation	Ernest Dautović, Leonardo Gambacorta and Alessio Reghezza
1084 March 2023	BigTech credit and monetary policy transmission: micro-level evidence from China	Yiping Huang, Xiang Li, Han Qiu and Changhua Yu
1083 March 2023	Commodity prices and the US Dollar	Daniel M Rees
1082 March 2023	Public debt and household inflation expectations	Francesco Grigoli and Damiano Sandri

All volumes are available on our website www.bis.org.