The Market Microstructure Approach to Foreign Exchange:

Looking Back and Looking Forward

This draft: 14 October 2012

Michael R. King, Carol Osler and Dagfinn Rime***

Abstract

Research on foreign exchange (FX) market microstructure stresses the importance of order flow, heterogeneity among agents, and private information as crucial determinants of short-run exchange rate dynamics. Microstructure researchers have produced empirically-driven models that fit the data surprisingly well. But currency markets are evolving rapidly in response to new electronic trading technologies. Transparency has risen, trading costs have tumbled, and transaction speed has accelerated as new players have entered the market and existing players have modified their behavior. These changes will have profound effects on exchange rate dynamics. Looking forward, we highlight fundamental yet unanswered questions on the nature of private information, the impact on market liquidity, and the changing process of price discovery. We also outline potential microstructure explanations for long-standing exchange rate puzzles.

JEL Classification: F31, G12, G15, C42, C82.

Keywords: exchange rates, market microstructure, information, liquidity, electronic trading.

***Michael King is at University of Western Ontario, Carol Osler is at Brandeis University and Dagfinn Rime is at the Norges Bank. The authors would like to thank Michael Melvin, {names} for helpful comments and suggestions. The views expressed in this paper are those of the authors and do not necessarily represent those of the Norges Bank. Please address all correspondence to Michael R King, Richard Ivey School of Business, UWO, mking@ivey.ca, tel: 519-661-3084.
The ancient and honorable field of international finance has grown furiously of late in activity, in content, and in scope.

Michael R. Darby

These opening words, written by the editor to introduce the inaugural issue of the Journal of International Money and Finance (JIMF) in 1982, could well have been written about the field of foreign exchange (FX) research today. Over the past thirty years, research on exchange rates has continued to grow in response to the puzzles and controversies that naturally arose following the move to floating rates after the breakdown of Bretton Woods.

This paper focuses on one facet of this research, FX market microstructure. Researchers in this field take a microeconomic approach to understanding the determination of exchange rates, which are, after all, just prices. Microstructure research in general analyzes the agents that trade in financial markets, the incentives and constraints that emerge from the institutional structure of trading, and the nature of equilibrium. FX microstructure builds on this general approach, but with models and empirical tests tailored to the study of currency markets.

This survey looks back at the key findings from the FX microstructure literature over the past 30 years, highlighting along the way the many important contributions published by JIMF.¹ We interpret the impact of microstructure research on the exchange rate literature using insights from Karl Popper, who stressed the interplay between empirical analysis and the development of theory. Most of the major findings in microstructure are empirical. This emphasis reflects the access of microstructure researchers to high-quality data and the focus on the decision-making of individual FX market participants. The growing availability of rich datasets has permitted

powerful tests of exchange-rate theories, allowing quick ‘falsification’ and providing strong pointers for how they might be improved. The survey also highlights ways in which microstructure research has contributed to powerful new explanations for long-standing puzzles, such as the forward bias puzzle, the profitability of technical analysis, and the greater explanatory power of purchasing power parity (PPP) over the long run.

Having highlighted stylized facts from the FX microstructure literature, this article draws attention to structural changes in the currency markets over the past decades, which have important implications for this field of study. Finally this article looks forward and highlights those research areas and questions that remain contentious or unanswered, which may provide fruitful areas for future research.

The JIMF, always receptive to the ‘facts first’ approach favored by microstructure researchers, has been the leading outlet for this field (see Appendix A). A review of the literature shows that the JIMF has published twice as many FX microstructure papers as the next leading outlet. The JIMF has published key microstructure papers even if they adopted methodologies not widely accepted in economics (e.g., surveys), even if they reached conclusions at odds with the rest of international economics (e.g., Evans and Lyons 2002a); and even if they dealt with microstructural nonlinearities orthogonal to standard models (e.g., Osler 2005). The JIMF has thus played an important role in establishing this line of inquiry as a respected part of international economics.

This paper has five sections. Section 1 looks back to the origins of FX microstructure. Section 2 reviews the most powerful finding from microstructure, namely the impact of order flow on exchange rate movements. Section 3 discusses liquidity provision and price discovery, and outlines the implications for exchange rate models. Section 4 raises open questions prompted
by the FX market’s rapid, on-going evolution. Section 5 outlines potential microstructure explanations for long-standing exchange rate puzzles. Section 6 concludes.

Section 1: The emergence of FX microstructure

A focus on the microstructure of FX markets was a natural development in the scientific analysis of floating exchange rates. This view is best appreciated in terms of Karl Popper’s depiction of the progress of science. For Popper, science is an evolutionary process in which theories are proposed, falsified by evidence, and then improved in light of the evidence. Such criticism is an essential activity and represents the only route through which science can achieve progress. FX microstructure research is primarily empirical, and can be characterized as adopting this falsification approach to knowledge.

When fixed exchange rates were abandoned in the early 1970s there was almost no evidence available to guide the development of exchange-rate models. Few countries had experimented with floating exchange rates and those few experiments were brief. In the absence of other evidence theorists postulated that PPP holds continuously – a theory that was falsified in the short run and, in some tests, in the long run (Rogoff 1996). The emerging evidence highlighted a specific shortcoming of continuous PPP, namely the exclusion of investor behavior (Kouri, 1976). Since international investing was discouraged under Bretton Woods, evidence on investor behavior was limited. Constrained by the lack of data, economists proceeded inductively and developed elegant money supply and portfolio balance models to explain the determination and behavior of exchange rates. Underpinning these parsimonious models was the assumption that uncovered interest rate parity (UIP) held continuously.
As more data became available, these next-generation exchange rate models were falsified, as all models inevitably are. The empirical challenges to these models were broad-based. UIP failed to hold at short horizons. Rather than depreciating as UIP predicted, high interest currencies were found to generally appreciate (Hodrick 1987; Engel, 1996; Bacchetta et al, 2009). Even covered interest rate parity (CIP) did not hold during turbulent periods (Taylor 1989). More broadly, these models were shown to forecast short-run exchange rate movements less well than the random-walk hypothesis (Meese and Rogoff 1983; Faust et al. 2003). With no consensus on how to address this failure of UIP, one strategy was to modify existing models by introducing an exogenous, time-varying risk premium. As noted by Burnside et al. (2007), this approach was ‘fraught with danger’ because it introduced an important source of model misspecification. Scientific progress, according to Popper, mandated the development of a third round of exchange-rate models.

Given the discouraging track record of the inductive approach, microstructure researchers adopted a deductive ‘facts first’ approach and began speaking directly to currency traders and other FX market participants (Taylor and Allen, 1992; Cheung and Chinn 2001, 2004; Gehrig and Menkhoff, 2004; Lui and Mole, 1998; MacDonald and Marsh, 1996; Menkhoff, 1998, 2001; Menkhoff and Gehrig, 2006). Researchers quickly established that the standard macroeconomic theories did not reflect the actual process through which dealers set exchange rates. Goodhart (1988) forcefully makes the case for this pragmatic approach: “economists cannot just rely on assumption and hypotheses about how speculators and other market agents may operate in theory, but should examine how they work in practice, by first-hand study of such markets.” This strategy paralleled the approach taken by researchers of equity market microstructure.
**Stylized facts on FX markets**

Any visitor to a currency dealing room in the 1980s saw dealers on multiple telephones noisily trading with other dealers. Similar activity could be observed in almost any city worldwide, since the FX markets are geographically decentralized. Then, as now, London captured about one-third of global FX trading, New York captured about one-fifth and trading in Asia was divided among Hong Kong, Tokyo, and Singapore. In the most liquid currencies the trading day is 24 hours long and trading floors are busiest when both London and New York are open. A currency’s liquidity tends to be deepest during local trading hours and there is a brief “overnight” lull in all FX trading activity between about 19:00 and 22:00 GMT (Lyons, 2001; Rime 2003; Osler, 2009). By 2010 the U.S. dollar (USD) was involved on one side of roughly three-quarters of all spot transactions, followed by the euro (EUR) at 46 percent, the Japanese yen (JPY) at 20 percent and the UK pound (GBP) at 14 percent.

[Enter Table 1 and Figure 1: Average daily interdealer trading activity by the hour]

Currency dealers are employed by the major commercial and investment banks. They intermediate trades between end-customers, and manage inventories or trade speculatively with other dealers in an interdealer market. The major end-customer groups include corporations engaged in international trade; asset managers such as hedge funds, mutual funds, endowment and pension funds, and insurance firms; smaller banks and central banks; and governments. Until the mid-1980s, trading by financial and corporate agents both represented roughly 20% percent of the market (with interdealer trading accounting for the rest). While the share of corporate trading has held steady over the years, the share of financial agents has risen dramatically. In  

---

2 Details on the composition of FX trading across currency pairs, countries, and instruments can be found in the 2010 Triennial Central Bank Survey of FX market activity (BIS, 2010; King and Mallo, 2010).
trading by financial institutions represented over 50 percent of daily average turnover (King and Rime, 2010). Currency markets are lightly regulated, if at all, with low transparency, little monitoring, and no official reporting requirements.

Any microeconomic investigation of a market begins by identifying the key agents’ objectives and constraints. Some FX market agents resemble the actors in standard exchange rate models. Hedge funds, for example, correspond well to the rational investors. They use currencies as a store of value and are motivated by profits, given the standard compensation scheme of 2 percent of assets under management and 20 percent of any profits. Like their modeling counterparts, hedge funds condition their trades on exchange-rate forecasts which they base on information gathered with costly effort. Their trading is limited by personal risk aversion, as in the models, but also by firm-based risk considerations and funding constraints.

Exporters and importers also have identifiable counterparts in standard exchange rate models. Such firms use foreign currency as a medium of exchange and therefore purchase more (less) of a currency once it has depreciated (appreciated). Most such firms do not permit speculative trading (Osler 2009) and do not condition their trades on exchange rate forecasts (Goodhart 1988, Bodnar et al. 1998). This choice is rational given the high costs of risk control and monitoring in firms where speculative trading is permitted (Osler 2009).

Other agents in currency markets do not have recognizable counterparts in standard models. Most international asset managers do not condition their trades on forecasts of future exchange rates (Taylor and Farstrup 2006), a choice that may be rational in light of the close correspondence between exchange-rate dynamics and a random walk. These asset managers also appear to be somewhat indifferent to execution costs, though those costs tend to be high (Osler et
Retail traders do condition their trades on exchange rate forecasts, but those forecasts appear uninformative because as a group they lose money (Heimer and Simon 2011).

Dealers are entirely absent from the standard macroeconomic models. They earn bonuses based on trading profits earned from liquidity provision and speculative position taking. Dealers are constrained in their risk-taking by position and loss limits. Given the volatility of exchange rates and the low costs of trading, dealers generally maintain inventories close to zero, especially at the end of the day. Inventory half-lives in currency markets are measured in minutes, even at the smaller banks (Lyons, 1998; Bjønnes and Rime 2005; Osler et al. 2011), whereas in equity and bond markets they are measured in days (Madhavan and Smidt 1993; Hansch et al. 1998).

The dealers themselves trade actively with each other with interdealer trading accounting for over 60 percent of spot FX trades during the 1980s and early 1990s (BIS, 2010). Interdealer trading is carried out either directly or indirectly via limit-order markets run by the electronic brokers EBS and Thomson Reuters (Lyons, 1995). In the interdealer limit-order markets, prices are ‘firm’ and brokers’ best bid and ask quotes provide a reliable signal of ‘the market. No agents are specifically tasked with providing liquidity. Every agent can either supply (‘make’) liquidity by placing a limit order, or demand (‘take’) liquidity by entering a market order.

Researchers have noted two key features of the market that constrain equilibrium prices. First, when quoting prices to end-customers dealers always begin with the prevailing interdealer quotes and adjust from there. Customer quotes are therefore tied to interdealer quotes. Second, interdealer quotes are constrained in turn by dealers’ preference for holding zero inventories overnight. This preference, which is rational given the volatility of exchange rates, implies that whatever inventory is accumulated by dealers though market making during a given day must,
by the end of that same day, be sold to other customers. And it is the exchange rate that moves to induce the necessary trading by customers.

[Figure 3: A typical FX dealer’s inventory]

Researchers visiting trading floors naturally sought to gain access to transactions and quote data, only to discover that such information was not yet captured electronically. The first high-frequency databases were thus assembled by hand. Goodhart and Figliuoli (1991), for example, analyze the behavior of the interdealer market as a whole using minute-by-minute exchange-rate quotes. Their pioneering work documented key stylized facts, such as the tendency for bid-ask spreads to cluster at just a few price levels and for exchange rate returns to be negatively autocorrelated.

Lyons (1995, 1998) studies the daily positions of a single active FX dealer during 1992. He documents that the dealer averaged $100,000 in profits per day on volume of $1 billion per day (or one basis point) and his positions had a half-life of only 10 minutes. By decomposing the dealer’s profits, Lyons (1998) finds that intermediation was more important than speculation, consistent with that dealer’s approach, known as ‘jobbing’, which focused almost exclusively on providing liquidity to other dealers. Jobbing was unusual even in 1992, and seems to be extinct as a strategy today. In a more recent study of four dealers at a major Scandinavian bank, Bjønnes and Rime (2005) identify a greater role for speculation with dealers actively trading across the two electronic broker platforms, EBS and Thomson Reuters. Mende and Menkhoff (2006) find that intermediation was the dominant source of profits for the dealing room of a small German bank. These studies highlight important differences between small and large dealing banks, and confirm the market-held view that large banks extract a substantial information advantage from observing more extensive trading flows.
Insights from the market

Many FX market insights were first documented through surveys of dealers, many – if not most – of which were published in JIMF: Taylor and Allen (1992), Lui and Mole (1998), Cheung and Chinn (2001, 2004), Gehrig and Menkhoff (2004), Menkhoff (1998, 2001) and Menkhoff and Gehrig (2006). Despite being a standard tool in most social science disciplines, surveys have not been widely used in economics. By publishing these surveys, the JIMF demonstrated intellectual independence and made a valuable contribution to the profession.

Such surveys help explain otherwise puzzling results uncovered through empirical research. For example, the clustering of bid-ask spreads documented by Goodhart and Figliuoli (1991) was explained by dealer reports that informal ‘market conventions’ was a strong determinant when quoting spreads (Cheung and Chinn, 2001). Since the market is intensely competitive, deviating from the competitive equilibrium is costly. These surveys also showed that market participants were heterogeneous in their trading styles, their views about other market participants, and their beliefs about the determinants of exchange rates.

A key market insight, documented by Menkhoff and Gehrig (2006), is the shared belief among FX dealers that exchange rates respond to currency-market flows. The importance of flows follows logically from dealers’ preference for zero overnight inventory holdings. Traders generally view the importance of trading flows to be self-evident; indeed, dealers base their trading strategies on this perspective (Osler 2006). This belief, however is profoundly inconsistent with the first and second generation exchange-rate theories with their inductively-derived focus on stock holdings of assets to the exclusion of asset flows. The view in these models that FX stocks could simply be first-differenced to create flows proved incorrect, as not all cross-border FX flows take place within currency markets, as discussed later. Similarly the
assumption that UIP and PPP hold continuously implies that flows have no role in maintaining these equilibrium relations. Likewise the assumption that all information is public and interpreted identically by all agents is inconsistent with traders’ views of price discovery, with private insights about exchange rates incorporated in prices through trading.

**Section 2: Order flow and exchange rate returns**

The shared belief among traders that currency flows are a critical driver of exchange rates could not be rigorously tested until researchers gained access to high-frequency transaction data. Such data was initially scarce during the 1980s when FX trades were predominantly executed over the telephone with back-office settlement relying on paper tickets and facsimiles. Pioneers such as Charles Goodhart, Rich Lyons, and Richard Olsen painstakingly assembled detailed datasets from printed records generously provided by dealer banks and FX brokers including EBS, OANDA, and Thomson Reuters.

**Interdealer order flow influences FX returns**

Lyons (1995) provides the first estimates of how order flow influences exchange rates. Lyons’ data comprise the complete trading record of a specific dealer during one week in 1992. Lyons found that this dealer would raise his quotes by 0.0001 DEM for incoming orders worth $10 million. As he recognized, however, one cannot extrapolate from a single dealer to the overall market.

Subsequent empirical work using comprehensive transactions data strongly confirmed the importance of interdealer trading flows for explaining exchange-rate dynamics. To arrive at this insight, however, researchers had to sort out a few methodological issues. One currency had to
be identified as the asset being traded and one as the medium of exchange. Market convention dictates that the traded asset is the base currency (or denominator) in the standard exchange rate quote. In EUR-USD, for example, a standard quote is 1.25 USD per EUR, so the base currency is the euro. To establish whether a currency was in net demand or net supply, it was also necessary to assign a direction to trades (either buys or sells), which was challenging as every FX trade involves both a demander and a supplier of liquidity. This ambiguity was resolved by recognizing that the product provided by a dealer is liquidity, namely the ability to trade a given quantity of FX quickly and inexpensively.

By viewing liquidity as the product, it became straightforward to assign trade direction: the aggressor in the trade is the agent purchasing (or taking) liquidity. For standard trades, the customer is the aggressor so the customer’s trade determines the trade direction. When the customer buys the base currency, the trade is categorized as a purchase. When dealers trade with each other, the dealer initiating the transaction is the liquidity demander. If the dealer demanding liquidity is purchasing the base currency, the trade is categorized as a purchase. In the microstructure literature, trading flows are thus calculated as the difference between buyer-initiated trades and seller-initiated trades, a measure called “order flow”. In other asset markets, this measure is called the “order imbalance”. Order flow corresponds to net liquidity demand for the base currency, with positive order flow associated with the base currency appreciating.

Evans and Lyons (2002b) first provided estimates of the exchange-rate’s response to interdealer order flow using transactions in USD-Deutschemark (DEM) and USD-JPY during four months of 1996. They regress the base currency’s daily return, \( r_t \), on order flow, \( \Delta x_t \), and fundamentals, \( F_t \):

\[
    r_t = \alpha + \beta \Delta x_t + \gamma F_t + \phi_t
\]  

(1)
where the fundamental variables are interest differentials, either lagged or in first difference. Subsequent researchers have sometimes included lagged exchange-rate returns. Consistent with the belief of traders, the estimated coefficient on order flow is positive and economically significant, indicating that net demand for the base currency raises its value in terms of the other currency. Specifically, an extra $1 billion in daily interdealer order flow was associated with a 0.5% appreciation of the USD vis-à-vis the DEM, with explanatory power on the order of 60 percent. Evans and Lyons (2002a) show that the explanatory power is higher, exceeding 70 percent, when returns are allowed to respond to order flow across additional currencies. When returns are regressed only on the interest differential or its first difference, the explanatory power is consistently below 1 percent.

Traders applauded this research as a sign that academics were getting better attuned to reality, with dealing banks creating teams to analyze their own order flow. Microstructure researchers noted that the Evans and Lyons’ (2002a,b) findings matched similar order flow evidence in equity and bond markets (Holthausen et al. 1990; Chordia, Roll, and Subrahmanyam 2002; Simon 1991, 1994; Brandt and Kavajecz 2005; Pasquariello and Vega 2007). While some economists saw the new results as highlighting a new direction for research, others remained skeptical and called for more evidence, particularly as the data were proprietary and not available to other researchers. This key finding has now been replicated with longer datasets, datasets that cover more currencies, datasets that are more recent, datasets from both large and small dealing banks, and datasets including brokered rather than direct interdealer trades. ³ Table 2 presents new evidence on this result for a broader set of currencies, and for longer samples, than ever before.

Explaining the influence of interdealer order flow

The evidence that order flow influences exchange rates, though striking, could not be fully credible without a rigorous explanation. An early criticism was that order flow was not a determinant of exchange rate at all, but rather reflected reverse causality with exchange rate returns causing interdealer order flow. Empirical support for the causal influence from order flow to returns was provided by Evans and Lyons (2005), Killeen et al (2006) and Daniélsson and Love (2006). The study by Daniélsson and Love (2006) reveals that the estimated influence from interdealer order flow to FX returns is stronger when one controls for feedback trading at the level of the individual transaction. The intuition is that there is nothing additional to learn from feedback trading than from prices.

The analysis of macroeconomic news events provided a further challenge to the idea, standard in inductively-derived models, that public information instantly affects exchange rates with no role for trading in this process. Econometric analysis of transactions data revealed that the impact of news operates primarily through order flow (Love and Payne 2003, Evans and Lyons 2002c, 2008; Rime et al. 2010). As ever, the JIMF published key results, including Evans and Lyons’ (2005) finding that these order-flow effects do not happen instantaneously, but persist for days. Similarly, Cai, Lee and Melvin (2001) identify an independent role for customer order flow that is distinct from the impact of macroeconomic announcements and central bank intervention in their study of the dramatic volatility of the Japanese yen in 1998.

Having established a separate role for order flow, researchers were able to explain its effect by drawing on three mutually consistent theories already well-established in the broader
microstructure literature. The first theory focuses on dealers’ inventory and operating costs, the second postulates a finite price elasticity of asset demand, and the third focuses on information asymmetry. All were originally derived in optimizing models with fully rational agents.

**Inventory effects**

Dealers are compensated for the costs of operations and the risk of holding inventories by charging a bid-ask spread. Order flow will naturally cause quotes to move between the bid and ask, consistent with Evans and Lyons (2002b). Buyer-initiated trades will move the quotes upwards, while seller-initiated trades move the quotes downwards. Inventory effects and the bid-ask bounce documented in many asset markets are not entirely consistent with the empirical findings, however, because inventory effects should only persist for a few minutes, given the market’s liquidity, whereas the effect of order flow on exchange rates persists much longer. Berger et al. (2008), for example, show that the price impact of interdealer order flow declines gradually over time but remains statistically and economically significant even at three months.

**Finite elasticity of demand**

A lasting effect of order flow on exchange rates emerges when the price elasticity of supply and demand are finite (Shleifer 1986). Evans and Lyons (2002b) outline a model of currency trading in which finite elasticity is center stage. Every trading day in the model includes three rounds of trading. In Round 1, dealers begin the day with zero inventory and are contacted by random customers to trade. Dealers quote prices, trade with end-customers, and accumulate inventory. In Round 2, the dealers trade with each other, effectively redistributing the aggregate inventory among themselves. In Round 3, dealers want to return to their preferred zero overnight inventory position so they set the quotes at a level such that a second set of customers willingly
purchase dealers’ aggregate inventory. This model captures so many important aspects of currency markets that it has become the intellectual workhorse of the microstructure field.

The first set of customers can be viewed as demanding instantaneous liquidity from dealers in response to exogenous shocks to their desired currency holdings. Since these customers permanently change their currency holdings, they effectively demand overnight liquidity from the market as a whole. Dealers willingly provide instantaneous liquidity, but are reluctant to provide overnight liquidity so move prices sufficiently that other customers are willing to do so. The theory is not specific on which end-customers demand liquidity in Round 1 and which provide overnight liquidity in Round 3. Round-1 customers might include corporations that face exogenous shocks to currency demand due a competitor raising prices, technology changes, or changes in barriers to trade. Or they may be financial customers whose demand is influenced by private information, noise trading mistaken for information (Black 1986), portfolio rebalancing by customers, or other random liquidity shocks. Corporate demand in Round 3 might arise as exchange rate movements change relative product prices. Similarly financial institutions may respond endogenously due to risk aversion, since a weaker currency promises a higher risk premium, other things equal. The heavy reliance of traders on technical analysis, documented by Taylor and Allen (1992), among others, also creates endogeneity in financial demand because technical analysis involves momentum and contrarian trading.

Since theory is agnostic on the respective roles of corporate and financial customers in the Evans and Lyons (2002a,b) framework, researchers have viewed the question as empirical. Round 1 customers can be distinguished from Round 3 customers according to the correlation between their order flow and returns. Round-1 customer order flow should be positively correlated with contemporaneous exchange rate returns, while Round-3 order flow is negatively
correlated. Researchers have also used the intertemporal properties of order flow to identify Round-1 from Round-3 customers. Round-3 trading should lag Round-1 trading, but not vice versa. Round-1 customer order flow should not respond to lagged FX returns while Round-3 customer order flow should.

Based on this identification strategy, studies using different time horizons and data sources consistently find that financial customers demand liquidity in Round 1 while corporate customers provide overnight liquidity in Round 3. Bjøtnes, Rime, and Solheim (2005) use comprehensive data on trading in Swedish krona to identify financial institutions as Round-1 customers and corporations as Round-3 customers. Marsh and O’Rourke (2005) use daily customer order flow from Royal Bank of Scotland and find that financial order flow does not respond to lagged returns while corporate order flow responds negatively. King et al. (2010) use eleven years of daily data collected for the Canadian dollar by the central bank to show that corporate order flow is negatively correlated with exchange rate returns, while financial order flow is positively related. This relationship is present in the response of order flow to macroeconomic surprises, to changes in macroeconomic expectations, and to changes in prices of commodity futures that influence the Canadian dollar.

The Evans and Lyons (2002a,b) 3-round dealer model has a number of implications for modeling exchange rates. First, it highlights the crucial role of corporate customers in determining exchange rates. Second, it highlights the relevance of finitely elastic currency demand, which contrasts with the assumption of infinite price elasticity under continuous UIP and PPP theories. Third, it shows that exchange rate models need not explicitly include dealers. Dealers may be involved in virtually every currency transaction, but because they prefer to hold zero inventory overnight they do not provide overnight liquidity and are thus of limited relevance
beyond the intraday horizon. This statement does not imply that one cannot learn low-frequency dynamics from studying interdealer order flow. Since they primarily intermediate end-customers, interdealer order flow can potentially mirror the customer market. One cannot, however, capture the heterogeneity of end-customers with interdealer trading since the interdealer flow is the sum of orders across different end-customer types.

Private information

Information provides a third reason why interdealer order flow could have a persistent impact on exchange rates. This possibility emerges directly from classic microstructure studies such as Glosten and Milgrom (1985) and Kyle (1985) that model the process through which private information influences prices via the order flow of informed agents. In the equity markets that inspired these models, the existence of private information about individual firms is not questioned. The existence of private information in currency markets is more controversial because most exchange rate fundamentals are publicly announced, such as interest rates and general price levels. But such public announcements are necessarily delayed relative to the realization of the fundamental variable itself, which provides time for agents to gather private information. This timing gap is also present in the disconnect puzzle between macro fundamentals and exchange rate variability, suggesting there is scope for disagreement around public information. Many hedge funds and other active traders devote extensive resources to gathering market intelligence, which would not be rational if there was no pay off to this activity. It is likewise noteworthy that dealers consistently stress the importance of private information in surveys. For example, Cheung and Chinn (2001) report that dealers view larger banks as having an informational advantage due to their larger customer base and market network.

---

4 Proprietary traders employed by commercial and investment banks are classified as financial customers, not dealers.
The microstructure evidence supports the market participants’ belief in private information relevant for future exchange rates. Evans and Lyons (2005) and Bjønnes, Osler, and Rime (2011) show that customer and interdealer order flow, respectively, have predictive power for future exchange rates. King et al. (2008) find that customer order flow has explanatory power over and above macroeconomic fundamentals and commodity prices when predicting movements in the Canadian dollar. Evans and Lyons (2009) and Evans (2010) provide evidence that Citibank customer order flow can be used to predict future GDP and inflation rates. Finally, Rime, Sarno, and Sojli (2010) show that interdealer order flow has predictive power for upcoming macro statistical releases.

If private information explains part of the influence of order flow on exchange rates, then agents must have access to different information sets or hold heterogeneous beliefs about public information. A predictable hypothesis would be that the contribution of different agents’ trades to FX returns will depend on the extent to which they are informed. Microstructure research suggests that some of this heterogeneity reflects imperfect rationality. Osler and Oberlechner (2011) show that currency dealers, as a group, tend to be overconfident and this tendency does not diminish over time. Myriad studies show that professional exchange rate forecasts are biased, inefficient, and inconsistent across time horizons (MacDonald 2000). This microstructure view of information is strikingly different from the inductively-derived perspective of standard exchange-rate models which assume all agents are homogenous and perfectly rational, all information is immediately announced to the public, and prices respond instantaneously to news.

There is abundant evidence supporting the heterogeneity of beliefs about exchange rates and the role of trading for revealing private information. MacDonald and Marsh (1996), for example, find that FX forecasters hold significant differences of opinion due to their
idiosyncratic interpretation of widely-available information. This heterogeneity translates into economically meaningful differences in forecast accuracy, with the extent of these disagreements determining market trading volume. Evans and Lyons (2005) and Carlson and Lo (2006) show that the information contained in macro news announcements takes days to become impounded in prices through the trades of dealers and end-customers. Bauwens, Omrane and Giot (2005) find that both scheduled and unscheduled macro news announcements have a significant impact on FX markets, with volatility increasing prior to these announcements reflecting increased uncertainty among market participants. Dominguez and Panthaki (2006) argue that the standard definition of news in macro models should be broadened to incorporate both non-fundamental news and order-flow. Dunne, Hau and Moore (2010) identify a strong influence from FX order flow to equity markets, suggesting that FX order flow captures changes in heterogeneous beliefs about fundamentals relevant to different asset classes.

Currency markets also display informational heterogeneity across customer locations and customer types. Menkhoff and Schmeling (2008) find that agents located in centers of political and financial decision-making are better informed than others. Among customer types, studies consistently conclude that financial customers are better informed than non-financial customers (Bjønnes et al. 2011; Carpenter and Wang 2003; Frömmel, Mende and Menkhoff 2008; Osler and Vandrovych 2009). Trades by corporate customers in liquid currencies appear to carry zero information, consistent with their role as liquidity providers. The empirical evidence indicates that the trades of retail investors carry no information; instead retail traders generally lose money, which suggests they (unintentionally) serve as overnight liquidity providers rather than rational speculators (Heimer and Simon 2011; Nolte and Nolte 2011). Finally, research indicates
that larger dealers themselves bring their own independent private information to the FX market (Bjønnes et al., 2011; Menkhoff and Schmeling, 2010; Moore and Payne, 2010).

Bacchetta and van Wincoop (2006), Evans (2011), Evans and Lyons (2005c, 2006), Frankel et al. (1996), Lyons (2001), and Sarno and Taylor (2001) have developed FX microstructure models that attribute the influence of order flow on exchange rates to private, heterogeneous information.

Section 3: Liquidity and price discovery in currency markets

Liquidity provision and price discovery are perhaps the two most important functions of financial markets. FX microstructure research has naturally focused on these topics, with implications for the modeling of exchange rates. Given that much of the existing microstructure research focuses on equity markets, it is important for exchange rate modelers to recognize that the conclusions of equity microstructure research “cannot be taken over into and applied to the FX market because the nature of the markets differ” (Booth 1994, p. 210).

Liquidity provision

Classic theories of liquidity provision indicate that bid-ask spreads should rise – and liquidity decline – with dealers’ risk aversion, volatility, the expected time between trades, and trade size, and information asymmetry (Ho and Stoll 1981; Glosten and Milgrom 1985). While bid-ask spreads in the interdealer FX market appear to conform to these predictions, recent FX research shows they do not hold in the customer markets.

In the interdealer market, Glassman (1987) confirms that the variation in interdealer bid-ask spreads reflect greater uncertainty, with market makers judging the probability of exchange
rate changes based on recent and long-term volatility. Hartmann (1998) and De Jong, Mahieu and Schotman (1998) confirm the importance of trading volumes and volatility for explaining bid-ask spreads. Focusing on a specific time period, Kaul and Sapp (2006) document that FX dealers widened bid-ask spreads from December 1999 to January 2000 as Y2K concerns led to increased safe-haven flows and rising dealer inventories in an environment of greater uncertainty and lower liquidity. Similarly, Mende (2006) studies FX interdealer spreads around September 11, 2001, and confirms they widened dramatically on the day of the 9/11 attacks then reverted to normal the next day, reflecting the spike in risk aversion and volatility from the event.

Hau, Killeen and Moore (2002) study one setting where the behavior of interdealer bid-ask spreads did not conform to standard theory. They find that bid-ask spreads on the newly created EUR were wider, not narrower, than the prior DEM spreads, despite the greater transaction volumes in the new currency. The authors argue this widening was paradoxically due to the higher transparency of order flow in the interdealer market. With only one currency in which to trade vis-à-vis the USD, dealers had fewer options for managing inventories without risking detection by other dealers.

The behavior of interdealer bid-ask spreads also conforms to the theoretical predictions from research on limit-order markets. Using minute-by-minute quotes from Reuters, Goodhart and Figliuoli (1991) and Goodhart and Payne (1996) confirm that trades are a major factor in spread determination and the frequency of quote revisions. Since these data were indicative quotes, like much of the high frequency data available in the 1990s, there was concern that they might not accurately represent firm prices or transaction prices. Danielsson and Payne (2002) compare indicative and firm quotes and find that indicative prices lagged when the market moved quickly, but indicative bid-ask quotes were generally quite close to firm quotes when
sampled at horizons of 5 minutes or longer. Lo and Sapp (2008) show that FX dealers’ decision whether to submit limit or market orders in interdealer limit order books is conditional on the previous type of order submitted as well as the recent volatility of the market, consistent with the theoretical predictions from Parlour (1998) and Foucault (1999). More market orders are used early in the trading day when information flows into the market, consistent with Bloomfield et al. (2005). Finally, Menkhoff et al. (2011) confirm that liquidity in the interdealer market responds to changes in volatility, bid-ask spreads, and other market conditions similarly to other markets, driven by informed traders.

The end-customer segment of the FX market does not behave consistently with classic microstructure theories regarding liquidity provision. Classical theories predict, for example, that dealers will “shade their prices”, which means they shift prices down when their inventory is excessive and vice versa (Ho and Stoll 1981). Such price shading has been documented in equity markets (Madhavan and Smidt 1993) and bond markets (Dunne et al. 2007). But studies of individual currency dealers provide no evidence of price shading (Bjønnes and Rime 2005; Osler et al. 2011), with the lone exception of Lyons’ FX jobber (Lyons, 1995). Dealers explain that quote shading would reveal information about their inventory position that could make them vulnerable. They prefer to unload inventory quickly in the liquid interdealer market.

The empirical facts on the bid-ask spreads quoted by dealers to their customers is equally at variance with classical microstructure theories. The orthodox view is that dealers widen spreads to protect themselves from adverse selection when trading against informed customers (Glosten and Milgrom 1985; Glosten 1989; Madhavan and Smidt 1993). However, Osler et al. (2011) show that FX bid-ask spreads for more informed customers – such as financial customers or customers making bigger trades – are narrower, not wider. This finding has been confirmed by
Ding (2009) for trade size and Rietz, Schmidt, and Taylor (2009) for customer type. This discrimination in favor of larger trades could reflect the lower per-unit operating costs, or the stronger bargaining power of financial customers (Green, et al. 2007; Rietz et al. 2009). Financial customers are also generally better informed and dealers may quote strategically to see their order flow, which may provide them with an information advantage when trading with other dealers (Bjønnes et al. 2011). Dealers thus have an incentive to maximize their trading with informed customers rather than to avoid such trading (Naik et al., 1999).

The contrasting findings for bid-ask spreads between the segments of the FX market highlights the importance of market structure in pricing behavior. Market makers lose when trading against better informed customers in any market. In the NYSE, which has just one tier, market makers (or specialists) have no one else with whom to trade after observing an informed customer trade. In consequence, they must offset those losses by charging wider spreads to informed customers. The FX market is a two-tier market, so dealers who trade with informed customers can turn around and exploit the information when trading with other dealers. They can quote narrower spreads to informed customers and make up the difference through more informed interdealer trading.

**Market structure and price discovery**

Market structure influences the price discovery process. Classical microstructure theories assume a one-tier market in which adverse selection is the dominant concern. While some exchange rate models have attempted incorporating these theories, the discussion in the previous section highlights two reasons why the classic theories cannot be directly applied in the FX context: (i) no influence from adverse selection to customer bid-ask spreads has been detectable
in currency markets; (ii) the currency market’s two-tier structure provides a rigorous, rational basis for the observed pattern of those bid-ask spreads.

Osler et al. (2011) propose a three-stage price discovery process relevant to the two-tier structure of FX markets. Following the literature, they assume that information originates with end-customers. In Stage 1, an informed customer trades with a dealer who receives a signal about the customer’s private information. This information, however, does not become embedded in prices at this stage because the informed customer pays a narrower spread than uninformed customers. In Stage 2 the dealer trades on customer’s private information in the interdealer market, leading other dealers to adjust their quotes in line with the customer’s original trade. The price continues moving in the direction implied by the customer information even after the first dealer has traded. In Stage 3, other customers contact dealers and the quotes they receive reflect this new information, completing the price discovery process. There is now substantial evidence consistent with this price discovery process for exchange rates.

A first key feature of Stage 2 is the dealers’ decision to mimic the trades of their informed customers. Using a probit analysis, Osler et al. (2011) confirm that dealers are more likely to trade aggressively after larger customer trades and trades with financial customers. Bjønnes et al. (2011) show that larger dealers (who are viewed as more informed) also tend to trade more aggressively. More persuasively, these authors show that a dealer’s tendency to trade aggressively rises with the volume of their business with informed and/or financial customers but is unaffected by the volume of business with corporate customers and governments.

A second key feature of Stage 2 is the reaction of other dealers to an interdealer trade. Goodhart and Payne (1996) and Menkhoff and Schmeling (2010) provide evidence that other dealers adjust their quotes in the direction of the most recent trade, thereby contributing to the
impact of any embedded information. After observing an aggressive interdealer purchase, for example, other dealers raise their quotes. Less informed dealers will also reverse the direction of their trades so that it matches the direction of dealers who are viewed as better informed.

**The nature and sources of private information**

Microstructure research certainly provides extensive evidence that interdealer and financial order flows carry information about upcoming returns. The exact nature of that information remains the subject of some debate. Is the information fundamental or transitory? The predictive power of customer and interdealer order flow for exchange-rate fundamentals, reviewed above, certainly suggests that the information may be linked to fundamentals.

Dealers are considered among the best-informed agents in FX markets. Not only does dealer order flow anticipate returns (Rime, Sarno, and Sojli, 2010; Bjønnes et al., 2011), it does so better than the trades of any individual group including leveraged investors (Osler and Vandrovych, 2009). But dealers typically hold positions for only a handful of minutes (Bjønnes and Rime 2005; Lyons 1995). This time horizon seems inconsistent with the notion that the information behind their trades is related to fundamentals. In surveys, dealers express the view that exchange rate fundamentals either do not exist or, if they do exist, do not matter (Menkhoff 1998). Instead dealers focus on customer order flow as one source of information on which to take speculative positions. There is no necessary inconsistency here, however, as dealers could focus on order flow and yet the information it contains may still be related to fundamentals.

If dealers are extracting fundamental information from customer order flow, what fundamental information does it reveal. And which end-customer segment has this information? It has already been noted that certain categories of hedge funds devote substantial resources to
macroeconomic analysis in an effort to anticipate macro statistical releases. Dealers could certainly acquire this information by observing the hedge fund trades. Dealers might also gather relevant information embedded in the trades of other agents but unknown to those agents themselves (Evans, 2010). Corporate traders, for example might reflect fundamental information about the state of domestic output, even though the corporations themselves do not trade on exchange rate forecasts (Goodhart 1988; Bodnar 1998). By observing sufficient corporate trades, dealers may detect a pattern that is dispersed and not known to individual corporate traders (Lyons 2001). Likewise, the trades of real money institutional investors such as mutual funds and pension funds might unwittingly reveal fundamental information on investor risk aversion or portfolio rebalancing (Breedon and Vitale, 2010; Killeen, Lyons and Moore, 2006). Finally, recent research suggests that dealers bring their own independent information to the market (Bjønnes, Osler, and Rime, 2011; Moore and Payne, 2011). The source and nature of private information revealed by customer and dealer order flow remains an open question.

**Section 4: Electronic trading and exchange rate dynamics**

Electronic trading has brought huge structural changes to FX markets over the past decade. In this section we briefly describe the most important institutional changes and highlight their importance for exchange rate dynamics. Greater detail on these changes is available in King and Rime (2010) and King et al. (2011).

*Trading among dealers and traditional FX customers*

When introduced on FX trading floors in the late 1980s, Thomson Reuters Dealing replaced the telephone with an electronic system for dealers to exchange messages, allowing for speedier and more efficient interdealer trading. The more important change occurred in the early
1990s when Reuters introduced the first electronic limit-order market for FX, now known as Thomson Reuters Matching, while a consortium of dealers launched a competing platform, Electronic Broking Service (EBS). These systems revolutionized the interdealer segment, but remained inaccessible to end-customers. The landscape changed dramatically in the late-1990s, however, when a number of multibank trading platforms were launched that targeted end-customers directly. These systems enhanced transparency, improved operating efficiency, and reduced trading costs at the expense of greater concentration among the top dealers who streamed quotes to these platforms. Over the next decade, massive investments in the IT infrastructure by dealers and market participants opened the door to algorithmic trading, with hedge funds and high-frequency traders gaining direct access to interdealer markets from 2005 onwards (King and Rime 2010). Starting in the early 2000s, the top banks launched proprietary single-bank trading platforms for their customers, allowing them to create pools of liquidity that are not visible to the market.

Individual dealing banks now stream executable prices to their FX customers across an array of public and proprietary trading platforms. As electronic brokers and dealers have segmented the market, different platforms have been launched targeting small corporate customers, sophisticated corporate customers, institutional asset managers, and leveraged investors such as hedge funds. These platforms are expensive to design, develop, and maintain, which means there are now substantial barriers to entry and economies of scale in FX dealing. Market making has, therefore, become a far more concentrated business with Euromoney reporting that the share of wholesale FX trading at the top three banks reached 40% in 2010, up from only 19% in 1998.

---

The high volume of customer transactions passing through the largest dealing rooms has, in turn, changed the economics of inventory management. It was always more profitable to lay inventory off on other customers, but dealers generally chose the interdealer market because it was faster. Now that advantage has been severely eroded, so dealers generally hold or “warehouse” inventory accumulated in customer trades. The change in the paradigm of inventory management has been extremely rapid. Less than one quarter of trades were crossed internally in 2007, but by 2010 that fraction reached a reported 80 percent at the largest dealers. Anecdotal reports suggest that liquidity provision and flow-based customer businesses have become the biggest source of FX revenues at the largest dealing banks following the 2007-2009 crisis.

The increased volume of customer business captured by the largest dealers has consolidated their information advantage relative to smaller banks, many of whom have chosen to curtail market making in the most liquid currencies. Small banks still have a comparative advantage in trading their local currency, so they choose to focus primarily on that niche business as well as relationship management and credit provision. Large banks, meanwhile, have aggressively developed new approaches to extracting information from customer flows. Dealers are most comfortable internalizing the trades of uninformed customers (“non-directional” flows). Dealers still actively court the trades of informed customers, however, they tend not to internalize such trades, consistent with our earlier economic analysis of dealer incentives. Instead they aggressively trade in interdealer markets to unwind any associated inventory position and possibly to take outright speculative positions. To optimize this effort, dealers have developed systems to screen out “predatory” flows, meaning trades in which a customer profits at the dealer’s expense due time latency between trade instruction and execution on the dealer’s own
platforms. Dealers prefer to exclude such customers from their platforms. The warehousing of dealer trades in this new environment has not yet been the focus of research.

The reduction in interdealer trading associated with the single-bank platforms has tended to reduce market transparency, but another development in the dealer-customer trading has had the opposite effect. On request-for-quote (RFQ) platforms customers looking for the best available price can now request quotes from several dealing banks simultaneously. By eliminating the need to search across banks sequentially, this system greatly improved customers’ negotiating power. Bid-ask spreads for customers that previously had low bargaining power, most notably corporate customers, tumbled accordingly (Goodhart et al., 2002; Bjønnes and Rime, 2005). There are as yet no studies of RFQ platforms in currency markets.

**Retail FX trading**

Electronic trading has enabled individuals of modest wealth, previously shut out of the market, to trade speculatively for their own account. This trading generally takes place over a new type of electronic trading platform known as the retail aggregator. By bundling many small retail trades into trades that meet the minimum $1 million size for interdealer trades, retail aggregators can provide narrow spreads on even tiny trades. Retail trading has grown rapidly and was estimated to have reached $125–150 billion per day by 2010, or 8 to 10 percent of the market (King and Rime 2010). Since retail customer order flow is generally uninformed (Heimer and Simon 2011), these customers are a profitable group to serve. Currently there is fierce competition for such business among the large banks, since they can effectively use these traders to provide liquidity for more informed customers. Evidence on retail trading remains quite limited, and represents a potentially fruitful area for future research.
Algorithmic traders and liquidity

As electronic trading systems progressed, it became possible for order-submission strategies to be programmed and executed entirely by computers (Chaboud et al. 2009). In some cases, computers simply automated the process of splitting larger trade into smaller transactions to reduce price impact. In other cases, computers provide a competitive advantage to traders whose strategies rely on speed such as triangular or covered-interest rate arbitrage.

Traders soon realized they could use the high execution speeds to take advantage of tiny discrepancies in the prices or timing of different trading platforms. These “high-frequency traders” typically provide substantial liquidity to the market via the hundreds or even thousands of tiny limit orders they submit each day. This strategy was initially profitable and spread rapidly in consequence, but it has compromised the profitability of traditional strategies and major banks have consequently pulled back from supplying on multibank platforms. The flash crash of 2010 raises important questions about the reliability of liquidity provision from high-frequency traders. Due to the need for liquidity, high-frequency trading is concentrated in the spot FX markets for the major currency pairs. While this segment remained relatively liquid in the turbulent period following Lehman Brothers’ collapse (Baba and Packer, 2009; Melvin and Taylor, 2009), there is no rigorous evidence on whether high-frequency traders provide true liquidity.

An important question facing FX markets is whether high-frequency traders are really increasing the liquidity of FX markets or merely creating a liquidity mirage that dries up when it is needed most. Future research could investigate whether high-frequency traders contribute to market liquidity primarily by submitting limit orders or by linking liquid pools across different trading platforms, thereby reducing market fragmentation.


*Liquidity aggregators*

With many electronic platforms competing to attract customers, there could conceivably be a fragmentation in market liquidity that could compromise market efficiency. A similar fragmentation has been observed in equity markets, where the introduction of Alternative Trading Systems (ATS) has eroded the market shares of the established exchanges. In FX markets the tendency towards fragmentation has been offset by the development of electronic tools that collect streaming price quotes from many competing platforms. By providing information on best prices these so-called “liquidity aggregators” perform a function similar to the National Best Bid and Offer (NBBO) system in the US, which also distributes information on the best bid and offer prices across all equity exchanges. The liquidity aggregators of currency markets, however, also allow customers to trade at the best prices, which is one reason they have proven especially popular with hedge funds. Despite the perceive threat to liquidity from the proliferation of exchanges, the evidence to date from equity markets is reassuring (O’Hara and Ye 2011; Degryse et al. 2011). FX liquidity aggregators, and their influence on liquidity and market dynamics, have yet to be studied.

*The global integration of FX markets*

Only a few researchers have studied how electronic trading may contribute to the transmission of information as well as shocks across global FX markets. Evans and Lyons (2002b) and Cai, Howorka and Wongswan (2008) document how electronic trading in the major currencies affects the returns and volatility of currencies in other regions. While most researchers view order flow as a mechanism for aggregating dispersed private information, Evans and Lyons (2002c) also conjecture that it is a proxy for liquidity as it affects the price impact of trades.
Banti et al. (2012) adopt this liquidity view and use order flow from real money investors across twenty currencies to construct a measure of global liquidity risk. Similar to the Pastor–Stambaugh liquidity measure for the US stock market, Banti et al. (2012) show that liquidity risk is priced in the cross-section of FX returns. Mancini et al. (forthcoming) confirm this finding using data for 12 currencies using high-frequency data from the EBS platform. They construct a wide range of liquidity measures and show that there are strong commonalities in liquidity across global FX markets. They also show that low interest rate currencies used to fund carry trades offer insurance against exposure to liquidity risk. Their data cover only the global financial crisis from 2007-2009, so it remains to be seen if their strong results also hold in calmer periods.

Greater financial integration can have adverse effects as highlighted by the dislocation in global FX markets during the 2007-2009 financial crisis. Despite originating in the US sub-prime mortgage markets, Melvin and Taylor (2009) document how the crisis spread globally across asset markets in 2007 and 2008, making it difficult to trade FX anywhere in any substantial size. FX volatility spiked to unseen levels, liquidity disappeared, and the cost of trading currencies skyrocketed. Thus greater integration of FX markets through electronic trading may prove to be a double-edged sword that increases liquidity and lowers transaction costs in good times, but transmits shocks in bad times. Again, more research on this possibility is warranted.

Section 5: Microstructure and exchange rate puzzles

The earlier discussion of order flow stressed its importance for exchange rate determination. In this section, we explore whether order flow can contributed to resolving some of the long-standing puzzles in FX markets, namely the forward bias puzzle, the profitability of technical analysis, and the greater power of PPP at longer time horizons.
Order flow and the forward bias puzzle

The importance of order flow for exchange rate determination provides the foundation for recent breakthroughs in our understanding of the failure of UIP and the associated profitability of the carry trade. The failure of UIP is so well known in international economics that we state it here only briefly. Under UIP, equilibrium expected exchange rate returns should compensate investors for the interest rate differential and risk premium:

\[ E\{s_{t+1} - s_t\} = (i_t^* - i_t) + rp_t \]  

(2)

where \( s_t \) is the (log) price of the home currency in terms of the foreign currency, \( i_t^* \) and \( i_t \) are domestic and foreign interest rates, and \( rp_t \) is a time-varying risk premium.

While UIP typically implies that high interest rate currencies should depreciate relative to low interest rate currencies, hundreds of studies find the opposite (Hodrick 1987; Engel 1996). Many investors exploit this regularity by borrowing in currencies with low interest rates and investing the proceeds in currencies with high interest rate, a strategy known as the “carry trade” (Burnside 2012). Figures 5 depicts a popular carry trade during the early 2000s. Investors funded in Japanese yen where interest rates were low and invested in the Australian dollar where the interest rates were high, earning positive returns as the Australian dollar generally appreciated. Economists have sought to explain this puzzle by introducing a risk premium. While earlier studies focused on the variance of FX returns, recent evidence points to the skewness of returns (Rafferty 2011; Brunnermeier et al. 2009). High interest rate currencies tend to exhibit negatively-skewed returns, implying that they appreciate slowly over long periods but then crash abruptly, as illustrated for the JPY/AUD in Figure 5.
Current widely-respected explanations for the failure of UIP rely on the influence of order flow on FX returns. Plantin and Shin (2011) assume that (i) the exchange rate always faces an exogenous probability of returning to its fundamental value and (ii) order flow has a positive, linear influence on exchange rates. The latter assumption implies that the order flow associated with carry trades will itself cause the high-interest rate currency to appreciate. When combined with the assumptions of (iii) ‘slow moving’ investment capital (Mitchell et al. 2007), (iv) positive feedback from carry-trade returns to funds invested in this strategy, and (v) a non-zero interest differential, the model produces negatively skewed returns to carry trades (Breedon et al, 2011). This model fits an extensive array of microstructure findings beyond the influence of order flow on returns, as outlined in Osler (2012).

Abreu and Brunnermeier (2003) develop a generic model of bubbles that relies on the response of prices to order flow. They show that coordination problems can arise under asymmetric information. Traders will rationally choose not to sell until everyone else is selling, but under asymmetric information they have difficulty identifying when others will sell. Traders therefore wait longer before selling, which allows bubbles to emerge and prices to rise above
fundamentals. A crash inevitably occurs, though it is delayed. The information asymmetries and 
coordination problems highlighted in this model can certainly be found in currency markets.

The JIMF has once again contributed to this line of research. Osler (2005) examines how 
price-contingent trading, specifically stop-loss orders, create currency crashes when order flow 
influences returns. A stop-loss buy order instructs a dealer to buy a specific quantity of a 
currency if and when the currency’s value rises to a pre-specified level. A stop-loss sell order is 
triggered when the currency’s value falls. These orders are common in currency markets. 
Customers can place them free of charge and every dealer monitors a book of them. Such orders 
generate order flow that tends to push FX prices further, automatically triggering yet more stop-
loss orders and associated order flow, leading to a price cascade.

The violent losses that disrupt otherwise positive carry-trade returns are typically 
concentrated on just a small number of days, pointing to the importance of stop-loss orders. 
Stop-loss induced price cascades are familiar to all currency traders, who describe the associated 
currency moves as extremely rapid, with the rate jumping over key price levels without trading.7 
Stop-loss orders can thus account for the exceptional swiftness of carry-trade unwinds.

**Order flow and the profitability of technical analysis**

Many technical trading strategies are profitable in FX markets (Menkhoff and Taylor 
2007). This empirical finding can be explained using order flow models. The documented 
profitability of trend-following technical strategies, including those based on moving averages 
and filter rules, is consistent with the same theories that explain the profitability of the carry 
trade. Technical analysts predict that exchange rate down-trends (up-trends) will tend to be

---

7 Stop-loss orders contribute to extreme returns in other markets, such as the May 2010 flash crash in US equity markets (CFTC-SEC 2010).
interrupted at specific support (or resistance) levels. If these levels are crossed, the existing trend will intensify. These predictions have been substantiated for support and resistance levels distributed to customers by major market players (Osler 2000). As outlined in Osler (2003), this highly non-linear behavior follows logically when one combines the influence of order flow on returns with the asymmetric clustering of stop-loss and take-profit orders near round numbers.

**Order flow and the PPP puzzle**

It is by now well-established that PPP holds in the long run for most currency pairs (Rogoff, 1996). The short-run failure of continuous PPP highlights a potentially important area in which the standard models could be improved. The short-run irrelevance and long-run relevance of PPP emerges naturally in Black (1985), Driskill (1981), and Osler (1995), all of whom independently develop and test models based on the insight that exchange rates are influenced by trading flows in currency markets. More recently, Fan and Lyons (2003) observe that financial order flow appears to be informative only at short horizons, while corporate order flow is informative at long horizons. Thus order flow from different end-customers may drive exchange rates away from PPP in the short-run but towards PPP over a longer horizon.

To understand this intuition, recall that exogenously-driven financial flows have a stronger influence on short run exchange-rate movements than exogenously-driven corporate flows. Over the longer horizon, however, the influence of financial flows should be approximately zero. A financial agent who opens a position today, buying currency and creating positive order flow, will eventually liquidate that position, selling currency and creating offsetting negative order flow. Over a longer horizon, the net impact should be minimal. Froot

---

8 Dahl et al. (2008) discuss the consistency between these models and the stylized facts of currency microstructure.
and Ramadorai’s (2005) study of State Street’s investor flows provide support for this hypothesis as they find that the positive influence of financial order flow disappears after about one year.

In contrast, corporate trades do not net to zero because their FX order flow represents only one side of a round-trip corporate transaction. Consider a US firm paying for imports from a European exporter. The US importer first buys euros through a currency dealer to pay for the transaction, generating positive order flow. If the European firm’s costs are denominated in euros, the opposite side of their transaction does not generate any order flow. For this reason, corporate shocks may have a permanent effect on exchange rates that only shows up in the long run after the transitory effects of financial order flows have dissipated. This timing suggests that PPP should be relatively easier to detect in the long run, consistent with the evidence.

**Section 6: Concluding remarks**

Our brief tour of FX microstructure research highlights how it emerged as a natural response to the empirical failure of early models of floating exchange rates. Due to the absence of historical experience, early macro models were designed inductively. As time went on and more data became available, however, some of their elegant assumptions and implications, though intellectually attractive, were ‘falsified’ by a growing body of evidence. As Popper highlights, falsifying existing theories is the only way empirical research can foster progress and is thus one of the primary responsibilities of good research.

Microstructure researchers adopted a deductive approach of surveying FX market participants and assembling detailed, high-frequency datasets. This effort produced the breakthrough insight about the forces that drive exchange rates, most notably order flow within the currency market. This insight, in turn, has led to other powerful insights. It is now recognized
that currency traders hold heterogeneous beliefs and have access to different information, some of which is private. While financial customers appear to be the best informed, their trades have only a transitory impact. Corporations are typically less informed, provide liquidity in overnight markets, and may contribute to the persistent impact of order flow on exchange rates. This interaction between informed and uninformed agents is key to modeling short-run exchange-rate dynamics. It is also recognized that currency market structure differs in important ways from the other financial markets, so researchers must be wise and selective in their reliance on the broader microstructure literature when developing rigorous exchange rate models.

Given the dramatic changes in electronic trading over the past decade, many established relationships may need to be revisited while the number of open questions has multiplied. No doubt FX researchers will continue to look to JIMF for leadership in publishing innovative yet controversial articles that move the field forward.
References


Table 1: Distribution of foreign exchange market turnover (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Geographical distribution (total)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>29.3</td>
<td>32.6</td>
<td>32.0</td>
<td>32.0</td>
<td>34.6</td>
<td>36.7</td>
</tr>
<tr>
<td>United States</td>
<td>16.3</td>
<td>18.3</td>
<td>16.1</td>
<td>19.1</td>
<td>17.4</td>
<td>17.9</td>
</tr>
<tr>
<td>Japan</td>
<td>10.3</td>
<td>7.0</td>
<td>9.0</td>
<td>8.0</td>
<td>5.8</td>
<td>6.2</td>
</tr>
<tr>
<td>Singapore</td>
<td>6.6</td>
<td>6.9</td>
<td>6.1</td>
<td>5.1</td>
<td>5.6</td>
<td>5.3</td>
</tr>
<tr>
<td>Switzerland</td>
<td>5.4</td>
<td>4.4</td>
<td>4.5</td>
<td>3.3</td>
<td>5.9</td>
<td>5.2</td>
</tr>
<tr>
<td>b) Currency distribution (Spot)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USD</td>
<td>71</td>
<td>78</td>
<td>84</td>
<td>85</td>
<td>79</td>
<td>80</td>
</tr>
<tr>
<td>EUR*</td>
<td>75</td>
<td>58</td>
<td>43</td>
<td>44</td>
<td>42</td>
<td>46</td>
</tr>
<tr>
<td>JPY</td>
<td>22</td>
<td>24</td>
<td>26</td>
<td>21</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>GBP</td>
<td>9</td>
<td>12</td>
<td>11</td>
<td>13</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>CHF</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>6</td>
</tr>
</tbody>
</table>

Note: Panel a: Country percentage shares of daily average global total in April. Country volumes are adjusted for local inter-dealer double-counting, but not cross-border (i.e., "net-gross" basis according to terminology of the BIS Triennial survey). Countries are sorted based on 2010 market share. Panel b: Total spot volume in a currency as percentage share of total global spot volume. EUR includes legacy currencies. Source: BIS Triennial FX Survey.
Figure 1: Average daily interdealer trading activity by the hour across different currencies

Note: The horizontal axis shows hour of day (GMT), and the vertical axis shows the average number of trades. The five lines are for 1997, the three 4-year average for 1998-01, 2002-05, and 2006-09, respectively, and finally for 2010. From e.g. GBP/USD (figure c) we see the growth in number of trades since 1997. The exchange rates EUR/USD and USD/JPY are now primarily traded on the competing platform EBS, hence the decrease in number of trades from 1997 to 2010 for these two exchange rates. Source: Thomson Reuters Matching.
Figure 2: FX spot market turnover by counterparty type

Note: Figure shows the share of financial customers (left axis) and non-financial customers (right axis, dot-symbols) out of total spot trading. Third group not shown in graph is dealers. G4-currencies (solid lines) are USD, EUR (DEM before 1999), JPY and GBP; Emerging market currencies (dashed lines) are here MXN, KRW, RUB, PLN, TRL, TWD, INR, HUF, ZAR and BRL.

Figure 3: Typical dealer inventory

(a) Lyons (1995)  
(b) Bjønnes and Rime (2005)

<table>
<thead>
<tr>
<th>Currency</th>
<th>Daily Price Impact</th>
<th>Intradaay Price Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OF-coeff</td>
<td>t-stat</td>
</tr>
<tr>
<td>AUD</td>
<td>0.016</td>
<td>27.89</td>
</tr>
<tr>
<td>CAD</td>
<td>0.016</td>
<td>27.40</td>
</tr>
<tr>
<td>CHF</td>
<td>0.300</td>
<td>2.15</td>
</tr>
<tr>
<td>EUR</td>
<td>0.026</td>
<td>9.84</td>
</tr>
<tr>
<td>GBP</td>
<td>0.012</td>
<td>29.96</td>
</tr>
<tr>
<td>HKD</td>
<td>0.003</td>
<td>16.36</td>
</tr>
<tr>
<td>JPY</td>
<td>0.084</td>
<td>13.76</td>
</tr>
<tr>
<td>MXN</td>
<td>0.021</td>
<td>16.84</td>
</tr>
<tr>
<td>NZD</td>
<td>0.056</td>
<td>20.52</td>
</tr>
<tr>
<td>SGD</td>
<td>0.022</td>
<td>18.84</td>
</tr>
<tr>
<td>THB</td>
<td>0.097</td>
<td>10.72</td>
</tr>
<tr>
<td>ZAR</td>
<td>0.063</td>
<td>20.08</td>
</tr>
<tr>
<td>EURCZK</td>
<td>0.066</td>
<td>24.65</td>
</tr>
<tr>
<td>EURDKK</td>
<td>0.004</td>
<td>20.05</td>
</tr>
<tr>
<td>EURGBP</td>
<td>0.013</td>
<td>22.77</td>
</tr>
<tr>
<td>EURHUF</td>
<td>0.063</td>
<td>23.15</td>
</tr>
<tr>
<td>EURJPY</td>
<td>0.509</td>
<td>4.40</td>
</tr>
<tr>
<td>EURNOK</td>
<td>0.032</td>
<td>25.51</td>
</tr>
<tr>
<td>EURPLN</td>
<td>0.043</td>
<td>19.30</td>
</tr>
<tr>
<td>EURREN</td>
<td>0.094</td>
<td>14.13</td>
</tr>
<tr>
<td>EURSEK</td>
<td>0.029</td>
<td>22.31</td>
</tr>
<tr>
<td>AUDNZD</td>
<td>0.051</td>
<td>14.67</td>
</tr>
<tr>
<td>NOKSEK</td>
<td>0.062</td>
<td>11.33</td>
</tr>
<tr>
<td>Average</td>
<td>0.072</td>
<td>18.11</td>
</tr>
</tbody>
</table>

Note: Table shows measures of daily and intraday price impact of order flow for several currencies. Order flow comes from the electronic broker Reuters D3000.2 is the number of buy-order minus the number of sell-orders. The regression for the daily price impact is $100\Delta \log(s_t) = \alpha + \beta OF_t/10 + \epsilon_t$, and the first columns report the $\beta$'s and their robust t-statistics. The interpretation for e.g. AUD is that a net imbalance of 10 trades move the AUD by 0.016% (for AUD the median imbalance is 50 trades). The next two columns report explanatory power (adjusted $R^2$) and number of observations. The longest sample start in 1996. All series end November 2011. The intraday-measure of price impact is the average of daily intra-day correlations between return and order flow, together with a t-test on the average of these daily correlations.
Figure 4: Market concentration. Number of banks covering x% market share.

Note: Dots, measured on right axis, represents number of banks covering 75% of the market according to the BIS Triennial Survey. The dots are weighted average of a selection of 14 countries, where share of the total volume of these 14 countries is used as weight. Lines, on left axis, measure the number of banks covering 60 and 75% of the market using the annual survey by the Euromoney.
Appendix A: Peer reviewed articles on FX microstructure in top journals, 1982-2012

Column 1 shows the number of peer-reviewed (scholarly) articles published between 1982 and 2012 in eleven top finance and economics journals. Column 2 shows the number of FX articles based on a search of the abstract for “exchange rate” or “foreign exchange”. Column 3 shows the number of FX microstructure articles from column 2 that include the words “microstructure” or “order flow” in the abstract. Column 4 shows the FX microstructure articles as a percentage of all articles published. The journals, shown in descending order based on column 3, are: Journal of International Money and Finance (JIMF), Journal of Banking and Finance (JBF), Journal of International Economics (JIE), Journal of Financial Markets (JFM), Journal of Financial Economics (JFE), Journal of Finance (JF), Journal of Financial and Quantitative Analysis (JFQA), American Economic Review (AER), Journal of Political Economy (JPE), Review of Financial Studies (RFS), and Quarterly Journal of Economics (QJE).

<table>
<thead>
<tr>
<th>Journal</th>
<th>(1) Number of scholarly articles</th>
<th>(2) (exchange W1 rate*) OR (foreign W1 exchange) in Abstract</th>
<th>(3) microstructure OR (order W1 flow) in Abstract</th>
<th>(4) Column 3 as % of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>JIMF</td>
<td>1,510</td>
<td>776</td>
<td>30</td>
<td>2.0%</td>
</tr>
<tr>
<td>JBF</td>
<td>3,807</td>
<td>175</td>
<td>7</td>
<td>0.2%</td>
</tr>
<tr>
<td>JIE</td>
<td>2,033</td>
<td>374</td>
<td>7</td>
<td>0.3%</td>
</tr>
<tr>
<td>JFM (1998-2012)</td>
<td>282</td>
<td>11</td>
<td>4</td>
<td>1.4%</td>
</tr>
<tr>
<td>JFE</td>
<td>2,092</td>
<td>23</td>
<td>3</td>
<td>0.1%</td>
</tr>
<tr>
<td>JF</td>
<td>3,399</td>
<td>75</td>
<td>2</td>
<td>0.1%</td>
</tr>
<tr>
<td>JFQA</td>
<td>1,226</td>
<td>30</td>
<td>1</td>
<td>0.1%</td>
</tr>
<tr>
<td>AER</td>
<td>6,174</td>
<td>172</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>JPE</td>
<td>1,736</td>
<td>37</td>
<td>1</td>
<td>0.1%</td>
</tr>
<tr>
<td>RFS (1988-2012)</td>
<td>1,401</td>
<td>23</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>QJE</td>
<td>1,421</td>
<td>35</td>
<td>0</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Source: Business Source Complete, SciVest Scopus.