Heterogeneous Information Content of Global FX Trading*

Angelo Ranaldo† Fabricius Somogyi‡

First draft: November 14, 2018

Abstract

This paper studies the information content of trades in the world's largest over-the-counter market, the foreign exchange (FX) market. The results are derived from a comprehensive order flow dataset distinguishing between different groups of market participants and covering a broad cross-section of currency pairs. Our findings show that both the contemporary and permanent price impact are heterogeneous across agents, time, and currency pairs, supporting the asymmetric information theory. A trading strategy based on the permanent price impact capturing superior information generates high returns even after accounting for risk, transaction costs, and other common risk factors documented in the FX literature.

J.E.L. classification: G12, G15, F31

Keywords: Carry trade, Currency portfolios, Heterogeneity, Order flow, Price discovery

*We have benefited from comments by Karl Frauendorfer, Joel Hasbrouck, Lukas Menkhoff, Dagfinn Rime, Lucio Sarno, Maik Schmeling, Andreas Schrimpf, Paul Söderlind, Vladyslav Sushko, Abraham Thomas, Paolo Vitale, Florian Weigert. We also thank seminar and workshop participants at the University of St. Gallen, PiF Research Seminar, the CEU Cardenal Herrera University, Workshop on Microstructure and Asset Pricing, the Hong Kong Institute for Monetary Research, 14^{th} Central Bank Conference on The Microstructure of Financial Markets, respectively. We are also grateful to Tammer Kamel and his team at Quandl for providing access to their data. We are responsible for all remaining errors and omissions.

†University of St. Gallen, Switzerland. angelo.ranaldo@unisg.ch

‡University of St. Gallen, Switzerland. fabricius.somogyi@unisg.ch

1 Introduction

One of the most important questions in financial economics is how security prices are determined. This is especially true for the foreign exchange (FX) market that is the largest financial market in the world, with an average daily trading volume of $5.4 trillion (see BIS (2016)). Being almost entirely an over-the-counter (OTC) market, FX trading activity is relatively opaque and fragmented.[1] Without a centralized trading mechanism, information is dispersed across various types of market participants (e.g. large commercial banks, security dealers, investment management firms, multi-national corporations, hedge funds, and central banks). All of them possess distinct information sets and may contribute differently to FX determination.

This paper sheds new light on how different market participants determine currency values in the global FX market and on the information content of their trades. To do this, we utilize a novel and comprehensive dataset that has three main advantages: First, it is representative for the *global* FX market rather than a specific segment (e.g. interdealer) or source (e.g. customers' trades of a given bank). Second, it includes identity-based order flow data broken down into types of market participants such as corporates, funds, non-bank financial firms, and banks acting as price takers. The order flow represents the net of buy volume by price takers minus the sell volume by market maker FX transactions. Third, it provides *hourly* order-flow time series, which is the finest time granularity that has ever been studied for the global FX market. In this framework, we address two key questions: First, does order flow impact FX prices heterogeneously across market participants, time, and currency pairs? Second, does this heterogeneity provide significant economic value to be revealed in a profitable trading strategy?

Answering these questions is important for both regulators and academics who have sought to better understand how asset prices are determined and how (fundamental) information is processed in financial markets. The asymmetric information paradigm first formalised by Glosten and Milgrom (1985) and Kyle (1985) prescribes that when some agents have superior information about the fundamental value of an asset, their trades covey information to the market. This body of the literature outlines two main empirical predictions: First, asymmetric information is positively related to the price impact of the trade. Second, the price impact tends to be persistent given the information content. Another body of the literature shows that the non-informative 'frictions' such as trans-

[1] The FX market microstructure is explained in detail in Lyons (2006), King et al. (2012). More recent developments of the FX markets are discussed in e.g. Rime and Schrimpf (2013), Moore et al. (2016).

action costs (Roll (1984)), liquidity shocks (Grossman and Miller (1988)), and inventory effects (Stoll (1978)) generate temporary price impacts and reversals. Therefore, the information content of trades is an important issue that calls for further research. Our work provides novel insights into price formation and asymmetric information issues by dissecting order flow into end-user segments of the global FX market and investigating which of them contain superior information to predict future FX rates.

The FX market is the world's largest over-the-counter (OTC) market, in which notoriously dispersed and asymmetric information across market participants creates adverse selection, illiquidity, and other frictions, especially in distressed times e.g. during the financial crisis of 2008 (Duffie, 2012). As a consequence, regulators have implemented global regulatory reforms to increase transparency and market quality such as the Dodd-Frank Act (USA; 2010), EMIR (Europe; 2012), and MiFID II (Europe; 2014). Shedding light on dispersed information and fragmentation in global FX markets would support these regulations that have direct implications on financial stability, price effectiveness, and fairness. Furthermore, our study hopes to be relevant to global investors to gauge heterogeneous information contents of FX order flow data.

Our paper proceeds in two parts. In the first part, we empirically address the question whether global order flow impacts FX prices heterogeneously across market participants, time, and currency pairs. To do this, we use a novel and unique dataset from CLS Group (CLS) from 2012 to 2017. CLS Group operates the world's largest multi-currency cash settlement system, handling over 50% of global spot, swap, and forward FX transaction volume. This dataset includes hourly order flows divided into four types of market participants: Corporates, funds, non-bank financial firms, and banks acting as price takers as well as the aggregate buy and sell side for 30 currency pairs. This dataset has been recently introduced and made publicly accessible, thus allowing the replicability and extensions of our study. By dissecting order flow into customer segments we preserve the information diversity across market participants, which gets lost otherwise, when segments are aggregated.

Our analysis builds upon the empirical methodology that decomposes the order-flow price impact into transitory and permanent components. The transient component arises from *non-information* factors such as inventory control effects (see Hasbrouck (1988)), price discreteness, order fragmentation, price pressure and smoothing, etc. (see Hasbrouck (1991a)). On the other hand, a trade can also convey fundamental *information* bearing a persistent impact on the security price. Econometrically, this paradigm is implemented using a bivariate vector autoregressive (VAR) system in which price return and order flow both evolve endogenously and the latter is allowed to contemporaneously impact the former. This framework provides two key advantages: First, it captures the *persistent* price impact of the trade innovation (using impulse-response analyses), which is a more precise estimate of processing superior fundamental information than the *immediate* price impact because the latter is contaminated by transient (liquidity) effects. Second, it is a general setting encompassing serial dependence of trades and returns, delays in the effect of a trade on the price and non-linear trade-price relationships that can arise from inventory control, price pressure effects, order fragmentation as well as other 'frictions' such as price discreteness, non-competitive behaviours or transaction cost components (e.g. clearing fees).

We refine this VAR system by allowing for *heterogeneous* price impacts of different agents. More specifically, we estimate a bivariate VAR model that controls for order-size, short-term mean reversion, and hourly seasonalities in a rolling fashion to study the time variation of both contemporary and permanent price impact impounded by the order flow of different market participants. We find compelling evidence that order flow impacts FX spot prices heterogeneously across agents, time, and currency pairs, supporting the asymmetric information hypothesis. In particular, corporates have a significantly lower contemporaneous and permanent price impact than funds, non-bank financials or banks. Moreover, order flows by market participants follow divergent patterns. This is consistent with the idea that corporates reflect largely uninformed trading (see Menkhoff et al. (2016)) and two common practices applied by FX dealers: discount fees to beg customers' liquidity provision and offsetting informative order flow with the non-informative one to reduce their exposure to asymmetric information risk. On the other hand, funds, and non-bank financials have a positive price impact suggesting that they gain more access to superior information for instance trading all around the clock, as we show in this paper. Furthermore, we find the FX market to be fragmented in the sense that both the contemporary and permanent price impact vary heavily across currencies suggesting (time-varying) asymmetric information in the cross-section of FX rates.

In the second part of the paper, we analyse the economic value of order flow heterogeneity. To do this, we introduce a novel long-short trading strategy based on a simple idea that is consistent with the information asymmetry hypothesis: Order flows of agents and currencies impounding a persistent price impact convey superior information leading to better predictions of future evolutions of FX rates. The intuition behind this strategy relies on well-documented deviations from the Uncovered Interest rate Parity (*UIP*) condition and the forward premium puzzle: When regressing FX returns on interest rate differentials the slope coefficient is typically not equal to one but negative.[2] In other words, the forward premium points into the 'wrong' direction of the expected price movement. Given that our measure of persistent price impact captures order flows conveying superior information, it is naturally well suited to identify trading that correctly predicts currency values or conversely, that is more biased by *UIP* deviations. We provide empirical evidence that currency pairs with a larger permanent price impact are more likely to deviate from the *UIP* offering informational advantages. Specifically, the forward rate of a currency pair with a high (*low or negative*) aggregate permanent price impact is more likely to 'overshoot' (*'undershoot'*) the future spot rate.

To assess the economic value of heterogeneous information content of global FX trading, we take the perspective of a US-investor who timely transacts on this information. Our strategy, that we name ALP_{HML}, is an equally weighted long-short portfolio that it is rebalanced on a daily, weekly, and monthly basis. Trading signals are generated from estimating our bivariate VAR model in a twelve-month rolling window fashion. For every rolling window index the permanent price impact $\alpha_m^{j,k}$, i.e. the sum of the predicted quote revisions through lag m, is extracted for every agent j and currency pair k. For every currency pair $\alpha_m^{j,k}$ is summed up across agents to derive the (aggregate) informative price impact. The resulting α_m^ks are sorted across currency pairs by size in ascending order. The ALP_{HML} portfolio then consists of the 20% highest (*lowest*) α_m^k currency pairs in the long (*short*) leg. Within this context going long (*short*) means buying (*selling*) a foreign currency in the forward market and selling (*buying*) it in the spot market next period. Transaction costs are implemented using accurate quoted bid and ask rates for both forward contracts and spot transactions.[3] At monthly rebalancing ALP_{HML} generates a both economically and statistically significant annualised return of 9.82% (*7.00%*) and a Sharpe ratio of 1.22 (*0.89*) prior (*after*) transaction costs. Furthermore, we show that these returns cannot be explained by the main common FX risk factors such as momentum, carry, and real exchange rates (see Asness et al. (2013), Lustig et al. (2011), Menkhoff et al. (2012b, 2017)).

To sum up, two important findings emerge from our analysis: First, both the contemporary and permanent price impact systematically differ across agents, time, and currency pairs, supporting the asymmetric information hypothesis. Second, there is a

[2]For extensive surveys and regression results on *UIP* deviations, see e.g. Hansen and Hodrick (1980), Fama (1984), Hodrick (1988), Lewis (1995), Lustig and Verdelhan (2007).

[3]Transaction costs in FX spot and future markets are studied in e.g. Bessembinder (1994), Bollerslev and Melvin (1994), Ding (1999), Hartmann (1999), Huang and Masulis (1999), Hsieh and Kleidon (1996), Christiansen et al. (2011), Gilmore and Hayashi (2011), Mancini et al. (2013).

significant economic value in the permanent price impact estimates in the sense that it is possible to exploit their time-series and cross-sectional variation to build a simple, yet profitable long-short strategy.

Related literature. We contribute to the microstructure and FX asset pricing literature. First, our analysis of heterogeneous FX order flows provides empirical evidence of information asymmetry across market participants (e.g. Kyle (1985), Glosten and Milgrom (1985), Easley and O'Hara (1987, 1991), Holden and Subrahmanyam (1992)). Prior research has already shown indirect evidence of heterogeneous behaviours in FX markets by looking at intraday patterns of volatility (Engle et al. (1990)), FX rates (Ito et al. (1998)), or surrounding special moments such as central bank interventions (e.g. Peiers (1997)). Starting from the key contributions of Evans (2002), Evans and Lyons (2002, 2005)), several papers provide more direct evidence of information asymmetry by investigating how aggregate order flow determines FX rates[4]. The only few papers that study the order flow disaggregated by market participants focus on a specific market segment such as a single interdealer trading platform (e.g. Moore and Payne (2011), Chaboud et al. (2014), Breedon et al. (2018)) or on customers' order flow of a specific bank (e.g. Evans and Lyons (2006), Breedon and Vitale (2010), Osler et al. (2011), Breedon and Ranaldo (2013), Menkhoff et al. (2016)). We are the first analysing order flow data representative for the entire global FX spot market with a large cross-section of FX rates and relatively long sample period (compared to the previous microstructure literature). Building upon the seminal work by Hasbrouck (1988, 1991a,b), we propose a general model for heterogeneous price impacts across agents disentangling permanent (*informative*) and temporary (*uninformative*) effects. Thus, our findings provide empirical evidence of information asymmetry at a global scale and support prior theoretical research on FX determination with heterogeneous contributions from different agents by Bacchetta and van Wincoop (2005) and Evans and Lyons (2006).

Second, our paper contributes to the FX asset pricing literature. Lustig and Verdelhan (2007) are the first to build cross-sections of currency portfolios to show that consumption growth risk explains why *UIP* fails to hold. Lustig et al. (2011), Menkhoff et al. (2012a,b) and Asness et al. (2013) identify common risk factors in currency markets based on real exchange rate, global FX volatility, and momentum. Other factors ex-

[4]This vast literature on FX order flow includes e.g. Payne (2003), Bjønnes and Rime (2005), Berger et al. (2008), Evans and Lyons (2008), Love and Payne (2008), Phylaktis and Chen (2009), Breedon and Vitale (2010), Evans (2010), Menkhoff and Schmeling (2010), Rime et al. (2010), Chinn and Moore (2011), and Mancini et al. (2013).

plaining carry trade returns include macro variables such as global imbalances (e.g. Della Corte et al. (2016b)) or volatility risk premia (e.g. Della Corte et al. (2016a)). We add to this literature by investigating whether the heterogeneity in global FX order flows provides significant economic value to build a profitable trading strategy. Menkhoff et al. (2016) perform a thorough analysis that also studies FX order flow and dissect customer currency trades into end-user segments. However, we differ from this paper in two key aspects: a sounder methodology and more granular data. On the methodological side, Menkhoff et al. (2016) construct currency portfolios using lagged, *total* order flow, which confounds transient (*uninformative*) and persistent (*informative*) effects of trades. The methodology used in our paper is more accurate and consistent because it isolates the informative component of order flow, which genuinely generates superior predictions of FX evolutions and thus higher excess returns. On the data side, the dataset studied in our paper is representative for the global FX market and can be accessible to anyone, whereas their work relies on an anonymous bank-specific source. Furthermore, we use a comprehensive intraday (hourly) dataset encompassing 30 currency pairs (instead of 15 daily FX rates).

The remainder of this paper is structured as follows. Section 2 describes our dataset, Section 3 presents summary statistics and Section 4 outlines the theoretical foundations. Section 5 estimates a simple trade/ quote revision model and analyses price impact heterogeneity across agents, time, and currency pairs. Section 6 exploits the price impact heterogeneity by a profitable long-short trading strategy. Section 7 concludes.

2 Data

Our dataset on spot FX order flow by market participant comes from CLS Group (CLS), which is available from Quandl.com - a financial and economic data provider. CLS Group operates the world's largest multi-currency cash settlement system, handling over 50% of global spot, swap, and forward FX transaction volume. After each and every FX transaction, settlement members of CLS are entitled to submit the details of the order for authentication and matching by CLS. CLS *volume* data (rather than *order flow*) have been used in prior research by Fischer and Ranaldo (2011), Hasbrouck and Levich (2018), Gargano et al. (2018), and Ranaldo and Santucci de Magistris (2018). To the best of our knowledge, this is the first paper to study CLS order flow data.

2.1 Heterogeneous FX Information Content

Volume is recorded separately for buy and sell side market participants after instructions are received from both counterparties to the trade. Within the dataset, CLS records the time of the transaction as if it had occurred at the first instruction being received.

CLS receives confirmation on the majority of trade instructions from settlement members within two minutes of trade execution (see Hasbrouck and Levich (2018) for further details, who provide a similar description of CLS volume data). Most of the currently 66 settlement members are large multinational banks. Furthermore, there are over 20,000 'third party' clients of the settlement members, including other banks, funds, non-bank financial institutions, and corporations. On the settlement date, CLS mitigates counterparty risk by simultaneously settling both sides of the FX transaction.[5] The FX spot market works on a $t + 2$ settlement schedule, unless both parties are in North America (e.g. $t + 1$). That is, when a spot trade occurs in time t the settlement instructions are submitted to CLS specifying that the actual transfer should occur two days later (see Pojarliev and Levich (2012), Levich (2012, 2013) for further studies on CLS).

This dataset has several features that make it particularly suitable to investigate the information content of FX order flow and its statistical and economic value. First, CLS records the trading volume in the base currency as well as the number of transactions on an hourly basis from Sunday 9*pm* to Friday 9*pm* (London time, GMT with 7 months using BST) and thereby matches the whole FX trading week from the opening in Sydney on Monday morning to the closing in New York on Friday evening. Second, CLS sorts FX market participants into four distinct categories: corporates (CO), funds (FD), non-bank financial firms (NB), and banks (BA).[6] In other words, these labels refer to the identities of the entities who trade and not to the behaviour they exhibit.[7] The category fund includes pension funds, hedge funds, and sovereign wealth funds, whereas non-bank financial refers to insurance companies, brokers, and clearing houses. Corporate comprises any non-financial organization. Hence, there is substantial heterogeneity in the motives for market participation across the four end-user groups. These groups are likely to differ considerably in their sophistication and access to price relevant information.

Corporates, funds, and non-bank financial firms are always considered to be price

[5]See Galati (2002) and Lindley (2008) for details of the CLS settlement process and the systemic impact on settlement risk.

[6]It is important to note that we do not have data on individual customers but only on customer types.

[7]The reason being is that CLS is a payment-versus-payment platform that solely observes the executed trade price used for settlement and does not see the market behaviour of bids and offers that precede the execution or any other such details.

takers and are a subgroup of the total aggregate buy side. Banks acting as market makers are always reported on the sell side. In any given hour, CLS records the buy volume and buy trade count referring to how much of the buy currency was purchased by the price takers from the market makers identified in that record. The sell volume and sell trade count refer to how much of the sell currency was sold by the same price takers to the same market makers.

CLS uses two distinct methods to categorise market participants: identity-based and behaviour-based. For the first part, CLS classifies market participants into corporates, funds, non-bank financial firms, and banks based on static identity information. Assuming that all corporates, funds, and non-bank financials act as price takers leads to three possible transactor parings between price takers and market makers: corporate/ bank, fund/ bank, and non-bank financial/ bank.[8]

However, the above pairings only account for about 10-15% of the total activity in the FX market. The majority of activity in this market is bank/ bank. Therefore, CLS carries out a second analysis focusing on bank/ bank transactions to determine which banks are market makers and which banks are price takers. CLS maps all FX activity as a network. Market participants are nodes, while FX transactions are edges. Nodes are then separated into two groups based on their coreness in the network. Nodes that are mutually tightly interlinked and that maintain a consistently high coreness over time are considered market makers, while all other nodes are considered price takers. Thus, the total buy side activity takes into account the sum of the three categories above plus all trades between price taker banks and market maker banks, to get a total of "all buy side activity" versus "all sell side activity". Hence, by construction, the sell side includes only banks that were identified to be market makers. To avoid double counting, transactions between two market makers or between two price takers are excluded from this dataset.

Empirically, transactions between market makers make up most of the activity in the FX market. Typically, a price taker does an initial trade with one market maker, and that market maker hedges the resulting risk by trading with other market makers. A single initial trade can lead to a chain of downstream transactions where various market makers pass the "hot potato" around or slice up the risk in various ways. Consequently, the activity among market makers will be higher than between price takers and market makers. There are three further reasons why transactions between non-bank price takers

[8]Within this context, the term "price taker" is interchangeably used with the term "buy side", and the term "market maker" is used interchangeably with the term "sell side".

and market maker banks make up a relatively lower share of total FX turnover settled by CLS. First, many hedge funds and proprietary trading firms settle through prime brokers. CLS does not have look-through on these trades, and hence they appear as bank/ bank transactions. If those prime brokers are also market makers, the transactions would be excluded from the order flow dataset. Second, CLS has relatively low client penetration among corporates and real money funds who trade FX infrequently and do not need a dedicated third-party settlement service, since they are permitted to trade and settle directly with commercial banks. Third, there is an asymmetry in that market maker banks may engage in price-taking activity, but price taker banks are very unlikely to ever engage in market-making activity. However, excluding market maker banks from the buy side based on their overall trading behaviour discards actual price-taking originating from these banks.

Our full sample period spans from 2 September 2012 to 19 November 2017 and includes data for 16 major currencies and 30 currency pairs.[9] This large cross-section of FX rates is important for evaluating the economic value of order flow and deriving a profitable trading strategy from the inherent information fragmentation across currencies.

The order flow dataset is limited to spot transactions. Three characteristics of the dataset merit being discussed in more detail: First, it contains around five years of data, which is a relatively long compared to previous studies in FX microstructure but shorter than traditional asset pricing datasets. Using a high-frequency dataset raises the statistical value of order flow in a time series setting by mitigating endogeneity and reverse causality issues and allowing us to measure the price impact of order flow over short rolling windows (e.g. over quarterly time frames). Second, despite being the most comprehensive time series dataset on FX order flow, it does not cover the full FX (spot) market. The 2016 BIS triennial survey (see BIS (2016)) reports an average daily trading volume of $5.1 trillion. Conversely, CLS settles approximately $1.5 trillion or 30% of total FX volume. The reasons for this lack of coverage are manifold: First, FX options and non-deliverable forwards are not settled by CLS. Second, small banks with insufficient FX turnover avoid becoming a settlement member. Third, CLS does not settle some of the high-volume currencies, namely the Chinese renminbi and Russian

[9]The full dataset contains data for 18 major currencies and 33 currency pairs. To maintain a balanced panel we exclude the Hungarian forint (HUF), which enters the dataset later, on 07 November 2015. Moreover, we discard the USDKRW due to insufficient amount of trades per price taker category. The remaining 30 currency pairs are: AUDJPY, AUDNZD, AUDUSD, CADJPY, EURAUD, EURCAD, EURCHF, EURDKK, EURGBP, EURJPY, EURNOK, EURSEK, EURUSD, GBPAUD, GBPCAD, GBPCHF, GBPJPY, GBPUSD, NZDUSD, USDCAD, USDCHF, USDDKK, USDHKD, USDILS, USDJPY, USDMXN, USDNOK, USDSEK, USDSGD, and USDZAR.

rubel.

However, both Gargano et al. (2018) and Hasbrouck and Levich (2018) demonstrate that the CLS coverage is underestimated compared to the BIS survey, since a large fraction of the volume reported by BIS is related to interbank trading across desks and double-counts prime brokered "give-up" trades.[10] Adjusting for these facts shrinks total FX volume to $3.0 trillion per day, and thus CLS covers 50% of the market for FX.[11]

Third, this dataset does not cover all transactions originated by one of the three price taker categories. More precisely, if a corporation settles a trade via a prime broker who is member of CLS, then this trade would show up as a bank/ bank transaction. The reason being that CLS does not observe the originator of such a trade but only the settlement itself. Consequently, such a transaction would either be excluded from the dataset, if the prime broker is a market maker, or it would show up as a transaction originated by banks acting as price takers, if former is behaving as a price taker.

Following the standard approach in market microstructure, we measure order flow as net buying pressure z_t against the base currency, which we define as the buy volume by price takers in the base currency minus the sell volume by market maker trades of the counter currency against the base currency,

$$T_t = \begin{cases} +1 & \text{if } z_t > 0 \\ 0 & \text{if } z_t = 0 \\ -1 & \text{if } z_t < 0 \end{cases},$$

where a positive T_t indicates net buying pressure in the counter currency against the base currency. Therefore, order flow does not measure trading volume, but rather net buying (or selling) pressure.

[10] In the 2016 BIS report (cf. p. 9), "related party trades" and "prime brokers" generated $0.94 trillion and $0.89 trillion in turnover, respectively.

[11] In their online Appendix Gargano et al. (2018) further mitigate concerns about the representativeness of the sample by providing evidence that an almost perfect relationship exists between the share of currency-pair volume in the BIS Triennial Surveys and the CLS data.

2.2 Exchange Rate Returns

We pair FX volume data with hourly spot rates obtained from Olsen, a market-leading provider of high-frequency data and time-series management systems.[12] FX order flow and exchange return are both measured at hourly frequency. The exchange rate return is calculated as the log-difference in the exchange rate over a trading hour:

$$\Delta s_t = s_t - s_{t-1}, \tag{2.1}$$

where natural logarithms are denoted by lowercase letters. Returns are always calculated based on the base currency.

3 Summary Statistics

In this section we present summary statistics for our data on FX quotes and signed net volume. In Table 1 we report summary statistics for the quote in each currency pair. The first five rows report the sample mean, standard deviation of the mean, minimum, and maximum hourly return as well as the average relative spread over the full sample. The last row reports the first order autocorrelation.

There are three takeaways from the hourly spot returns summary statistics table: First, the average return over the hour is zero due to mean reversion (i.e. returns experience negative first order autocorrelation). Second, standard deviation of returns is reasonable in the range of 10-20 basis points (BPS). Third, the average relative spread varies substantially in the cross-section due to variations in transaction costs.

Table 2 reports detailed summary statistics for the hourly (absolute) net volume for the entire cross-section of currency pairs. The first five rows in Table 3 report the sample mean (*Mean*), standard deviation of the mean (*Std(Mean Vol)*), median (*Median*), 90^{th} percentile, and 10^{th} percentile of hourly (absolute) net volume in USD million. The last row displays the first order autocorrelation for aggregate volume *AC(1)*. Little surprisingly, the currency pairs with the highest hourly volume are AUDUSD, EURUSD, GBPUSD, USDCAD, and USDJPY, for which CLS settles on average USD133 mn., USD454 mn., USD202 mn., USD227 mn., and USD247 mn., respectively in spot trans-

[12] Olsen data are filtered in real time by getting a credibility tick assigned (ranging from 0 to 1). The filter looks at several factor i.e. the quality of the quoter, agreement with other quoters within the spread, the size of the movement with respect to the last tick, consistent lack of fluctuations, and many more. The number of ticks excluded from the supplied data due to credibility < 0.5 depends on the number of bad quoters for the rate but typically ranges from 0.5% to 3.0% per day.

actions per hour. Our ranking is largely in line with both the BIS Triennial Surveys and Gargano et al. (2018). Across currencies, we observe a mild first order autocorrelation of 8-17% that is consistent with prior research on stock markets e.g. Hasbrouck and Ho (1987). Funds and non-bank financials are the largest categories after banks acting as price takers, whilst corporates form the smallest group. There are at least two potential reasons why corporates show up less frequently in this dataset: First, corporates are known to use swap rather than spot transactions to hedge their currency exposure. Second, corporates access the FX spot market indirectly through prime brokers/ banks rather than directly via CLS. Moreover, large, multinational corporations are more likely to trade via CLS than their smaller, regional counterparts. Hence, we cannot rule out a potential bias towards large-sized corporations in the data.

Figure 1 fleshes out the idea that corporates trade at different times than funds or non-bank financials. For every market participant we report the average aggregate hourly volume for each hour of the trading day based on London time (no BST adjustment). Investigating at which hours market participants are most active helps to identify time-fixed effects in the trading behaviour of FX market participants. In the European morning, when only Asian markets are open, volume levels are relatively low. FX volume rises when European and London markets open at 6*am* and 7*am* London time, respectively. Around lunchtime, trading drops and picks up again when New York traders enter the market around 1*pm*. Volume is lowest during the night between 10*pm* and 11*pm*, when only the Australian market is open. This pattern persists across market participants. Banks, non-bank financials and funds all trade more around the clock. Banks are the largest subsection of the aggregate, with an average contribution of 30% to 50%. As expected, corporate trading is more concentrated in European working hours, i.e. 7*am* to 5*pm*.[13]

In line with the empirical evidence for equity markets (see Jain and Joh (1988), Gerety and Mulherin (1992)), we extend our analysis to FX order flow and find that (absolute) net volume is concentrated in the early and later parts of the trading day in London and New York suggesting heterogeneous information flows across (intraday) time.[14]

To conclude the descriptive analysis, we address three possible problematic issues on order flow data segregated by market participants groups stressed in Evans and

[13] We observe a similar pattern when using number of trades instead of volume per hour.

[14] Within these periods, more economically relevant news are released. See Foster and Viswanathan (1993) and Berry and Howe (1994). However, this pattern does not hold consistently among all currency pairs. For instance, the USDJPY experiences three peaks that include the opening hours in Tokyo and Sydney. EURGBP trading is concentrated between 7*am* and 5*pm* GMT.

Lyons (2006): (i) intra-temporal dependence, (ii) inter-temporal dependence, and (iii) representativeness. Rather than price impact parameters, the presence of these issues would force us to interpret the coefficients as a simple mapping of the variation in order flow segments into the flow of fundamental information that have yet to be fully assimilated by dealers across the market. First, both order flow and (signed) net volume exhibit low levels of intra-temporal correlation among different order flow/ (signed) net volume segments.

Second, Figure 2 plots the average correlation coefficients between customer order flows for horizons of 1,2,..., and 60 trading days (see the online Appendix for (signed) net volume). Average correlations between flows are based on the average correlation across all 30 currency pairs. A horizon of one day corresponds to non-overlapping hourly observations, whilst for longer horizons we sum over daily (overlapping) observations using the full sample. The shaded areas correspond to 95% confidence bands based on a moving-block bootstrap with 1,000 repetitions. We find global evidence that all correlations between financial (FD and BA) and non-financial customers (CO and NB) are significantly negative at all horizons, while there is hardly any significant correlation between flows of the non-financial customer groups. These results corroborate the risk-sharing hypothesis whereby financial players trade in the opposite direction of non-financial market participants. Therefore, our empirical analysis supports the idea that risk sharing takes place at a global scale and across customer segments rather than only in the interdealer segment or between customers of a given bank (Menkhoff et al. (2016)). Furthermore, these patterns indicate that across the entire cross-section of currencies, serial autocorrelation is none of an issue, since the difference between the Durbin-Watson test statistic and its critical value 2 is < 0.001 for all currency pairs. The Ljung-Box test for residual autocorrelation renders similar results: For more than 90% of currency pairs we do not reject the null hypothesis of no residual autocorrelation in order flow up to lag 24.

Third, instead of relying on a bank-specific source[15] or segment-specific trading platform [16], we use a comprehensive dataset, which covers around 50% of the global FX turnover compared to the BIS triennial survey (see BIS (2016), Hasbrouck and Levich (2018)). Despite this broad coverage, the possibility of an omitted-variable bias remains.

[15] See e.g. Evans and Lyons (2006), Breedon and Vitale (2010), Osler et al. (2011), Menkhoff et al. (2016).

[16] See e.g. Moore and Payne (2011), Chaboud et al. (2014).

4 Methodology

In this section we describe the methodology to investigate whether market participants exhibit a heterogeneous price impact in the FX spot market. It builds upon the framework of Hasbrouck (1988, 1991a), who introduces a vector autoregression (VAR) that makes almost no structural assumptions about the nature of information or order flow, but instead infers the nature of information and trading from the observed sequence of quotes and trades. Within the general setting of Hasbrouck (1991b) stock price movements are either related or unrelated to a recent trade. More specifically, trade-related price moves may convey superior private information, although the model itself does not make any structural assumptions on information. However, for the VAR model to be estimated consistently, it is necessary that the time series is covariance stationary with respect to the time index used. Jones et al. (1994) and Barclay and Hendershott (2003) argue that trades based on "public information" lead to price moves that are orthogonal to recent trade arrival.

Hasbrouck (1988) provides a useful model to separate permanent (information) effects and temporary (inventory) effects of a trade but suffers from the limitation that order flow is assumed to evolve exogenously. However, prices can feedback to order flow. To overcome this issue, Hasbrouck (1991a,b) proposes a bivariate VAR model that allows to decompose the price moves into trade-related and trade-unrelated components as well as to endogenise order flow. Consistent with this framework, we build an encompassing model that allows for heterogeneous order flows and controls for short-term mean reversion as well as hourly seasonalities. In particular, Eq. (4.1) describes the trade-by-trade evolution of the quote midpoint, whilst Eq. (4.2) refers to the persistence effect of order flow. We define T_t to be the buy-sell indicator ($+1$ for buys, -1 for sells) for trade t in a specific currency pair.[17] Furthermore, we define r_t to be the log FX-rate return based on the mid-quote. Easley and O'Hara (1987) present a theoretical asymmetric information model in which private information revealed by an order and the consequent change in quotes are positively related to order flow size. We account for these effects by introducing an order-size variable (cf. Hasbrouck (1988)) into the VAR specifications. Logarithms are taken to control for presumed non-linearities between order size and

[17] T_t^{CO} for corporate, T_t^{FD} for fund, T_t^{NB} for non-bank financial, and T_t^{BA} for banks acting as price takers, i.e. the orthogonalised volume that is the total buy side minus the aggregate (signed) net volume of every market participant.

quote revisions:

$$v_t = \begin{cases} +log(z_t) & \text{if } z_t > 0 \\ 0 & \text{if } z_t = 0 \\ -log(-z_t) & \text{if } z_t < 0 \end{cases}.$$

To support the interpretation of the regression coefficients, v_t is transformed by regressing it against current and lagged values of the trade indicator variable T_t. As proposed in Hasbrouck (1988), we extract the residuals from this regression, denoted by $\tilde{S}_t$, which are by construction uncorrelated with the indicator variable T_t.[18] Hourly dummies are included to control for daily seasonalities affecting FX rates and order flows.[19] More importantly, the VAR accommodates both lagged returns and order flow in both the return (Eq. (4.1)) and order flow equation (Eq. (4.2)), since many microstructure imperfections such as price discreteness, inventory effects, lagged adjustment to information, non-competitive behaviours, and order splitting are thought to cause lagged effects. The number of lags is selected to be five based on data-driven methods and theoretical foundations postulated by Hasbrouck (1991a,b). Our findings remain qualitatively unchanged when using more lags but become computationally expensive. In particular, the regression coefficients β_i^j at lags four/ five and beyond are mostly statistically insignificant.

$$r_t = \zeta_{1,l} D_{l,t} + \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{j \in C} \left(\sum_{i=0}^{5} \beta_i^j T_{t-i}^j + \sum_{i=0}^{5} \phi_i^j \tilde{S}_{t-1}^j \right) + \upsilon_1 \Delta s_{k,t;t-\tau} + \upsilon_2 \Delta s_{k,t;t-5\tau} + \epsilon_{r,t}, \quad (4.1)$$

$$T_t = \zeta_{2,l} D_{l,t} + \sum_{i=1}^{5} \gamma_i r_{t-i} + \sum_{j \in C} \left(\sum_{i=1}^{5} \delta_i^j T_{t-i}^j + \sum_{i=1}^{5} \omega_i^j \tilde{S}_{t-1}^j \right) + \epsilon_{T,t}, \quad (4.2)$$

where $D_{l,t}$ denotes a dummy variable matrix to account for time-fixed effects with $l = 24$ columns and $t = n$ rows, where element l, t is 1 if there was a trade in that hour, and $C = \{CO, FD, NB, BA\}$. Transactions are indexed by t. Moreover, the regression also takes into account the lagged exchange rate changes over the previous day $\Delta s_{k,t;t-\tau}$ and over the prior week $\Delta s_{k,t;t-5\tau}$. Here, $\tau = 24$ and t is measured at hourly frequency. The error terms $\epsilon_{r,t}$ and $\epsilon_{rT,t}$ can be interpreted as an expected (public information) and unexpected (private information) component, respectively.

Since we include contemporaneous T_t in Eq. (4.1) but not in Eq. (4.2), the system

[18] It is important to note that our main results remain qualitatively unchanged when excluding the order-size variable from our baseline VAR model.

[19] Prior research provides evidence of time-of-day seasonalities in FX returns (e.g. Ranaldo (2009)) and FX order flow (e.g. Breedon and Ranaldo (2013)).

is exactly identified and hence the error terms shall have zero mean and be jointly and serially uncorrelated:

$$
\begin{aligned}
&E(\epsilon_{T,t}) = E(\epsilon_{r,t}) = 0 \\
&E(\epsilon_{T,t}\epsilon_{T,s}) = E(\epsilon_{r,t}\epsilon_{r,s}) = E(\epsilon_{T,t}\epsilon_{r,s}) = 0, \text{ for } s \neq t.
\end{aligned} \tag{4.3}
$$

Next, in order to make Eqs. (4.1) and (4.2) more intuitive, the VAR shall be inverted to its vector moving average (VMA) representation. Hereby we follow the methodology in Hasbrouck (1991b), Hendershott et al. (2011) and thus derive:

$$
y_t = \begin{bmatrix} r_t \\ T_t \end{bmatrix} = \Theta(L)\epsilon_t = \begin{bmatrix} a_r D_t & b_r(L) & \vec{c_r}(L) & \vec{d_r}(L) & \vec{s_r} \\ a_T D_t & b_T(L) & \vec{c_T}(L) & \vec{d_T}(L) & 0 \end{bmatrix} \begin{bmatrix} \epsilon_{D,t} \\ \epsilon_{r,t} \\ \epsilon_{\vec{T},t} \\ \epsilon_{\vec{S},t} \\ \epsilon_{\vec{v},t} \end{bmatrix}, \tag{4.4}
$$

where $b_r(L), \vec{c_r}(L), \vec{d_r}(L), b_T(L), \vec{c_T}(L)$ and $\vec{d_T}(L)$ are lag polynomial operators. With $\vec{c_r}$, $\vec{c_T}$, $\vec{d_r}$, $\vec{d_T}$ being a row vector equal to $[\beta_i^{CO}\ \beta_i^{FD}\ \beta_i^{NB}\ \beta_i^{BA}]$, $[\delta_i^{CO}\ \delta_i^{FD}\ \delta_i^{NB}\ \delta_i^{BA}]$, $[\phi_i^{CO}\ \phi_i^{FD}\ \phi_i^{NB}\ \phi_i^{BA}]$, and $[\omega_i^{CO}\ \omega_i^{FD}\ \omega_i^{NB}\ \omega_i^{BA}]$, respectively. $\vec{s_r}$ refers to a row vector consisting of $[v_1\ v_2]$ from the return equation.

Permanent Price Impact. From Eqs. (4.1) and (4.2) we can derive the permanent price impact both on the individual agent level and aggregated across agents. Following Hasbrouck (1991a), the permanent price impact of agent $j \in C$, with $C = \{CO, FD, NB, BA\}$, can be calculated as follows:

$$
\alpha_m^j(\epsilon_{T^j,t}) = \sum_{t=0}^{m} E[r_t|\epsilon_{T^j,t}] = \sum_{t=0}^{m} \beta_t^j, \tag{4.5}
$$

where m indicates the number of lags, which is five in our case. Since α_m^j is cumulative over several hours (even weak effects can add up), VAR estimates of lower order ($t \leqslant 5$) are likely to overstate the long-run price impact. In other words, such a model would catch the initial positive impact of a trade on the quote but will miss the subsequent long-run reversion. Using the VMA representation, the cumulative impulse response (permanent price impact) aggregated across agents is given by:

$$\alpha_m(\epsilon_{T,t}) = \sum_{j \in C} \sum_{t=0}^{m} c_{\vec{r},t} = \sum_{j \in C} \alpha_m^j. \tag{4.6}$$

Within this framework, the cumulative impulse response function of the quoted price to a one-unit shock in the order flow equation is a measure of asymmetric information and adverse selection that accounts for the persistence in order flow as well as possible positive or negative feedback trading. Since α_m lies at the heart of the subsequent asset pricing analysis, it is important to underline that α_m possesses a natural interpretation as the information content of the innovation net of transient effects inherent in global FX trading.

To sum up, Hasbrouck (1991a,b) is the most suitable approach for capturing the permanent price impact conveyed by trade innovations robust to transitory effects such as price discreteness, inventory effects, information cascades and lagged adjustment to trades.

5 Heterogeneous Information Content of Global FX Trading

In this section, we analyse whether the price impact in the global FX spot market is heterogeneous across market participants, currency pairs, and time.

5.1 Estimating a Simple Trade/ Quote Revision Model

First, we estimate Eqs. (4.1) and (4.2) by standard ordinary least square (OLS) on the full sample controlling for seasonal time of the day effects, lagged returns, and order size.[20] Second, we apply a rolling window of twelve months to measure time variation of both the contemporary $c_0^{j,r}$ and permanent price impact α_m^j. The main advantage of the VAR approach lies in its potential for generalisation to gain a more nuanced view of the trade-quote interactions.[21] For the sake of clarity we will present results only for lagged

[20] To avoid misspecification in our regression analysis and check the validity of our assumptions in Eq. (4.3) we conduct a battery of diagnostic tests that are summarised in the online Appendix.

[21] As in Hasbrouck (1991a), T_t is defined as a limited dependent variable. As long as T_t and r_t are jointly covariance stationary and invertible, a VAR model as in Eq. (4.4) exists. However, even though the error terms are serially uncorrelated, they are not serially independent in general. The disturbance properties in Eq. (4.3) further ensure that the coefficients in Eq. (4.4) are estimated consistently by OLS. Nonetheless, estimation errors can lead to non-zero disturbance terms for certain out-of-sample data points (see Hasbrouck (1991b)).

return equation coefficients b_1^r (α_1) and b_1^T (γ_1), the contemporary price impact $c_0^{j,r}$ (β_0^j) and lagged order flow $c_1^{j,T}$ (δ_1^j), where $j \in C$ denotes one of the market participants. Table 4 shows the regression coefficients of the bivariate VAR estimated through five lags. The most important ones are those of $T_0^{j,r}$ that measure the contemporary price impact of a trade. Coefficients beyond lag three and four are seldom significant. To overcome the curse of heteroscedasticity and autocorrelation we apply heteroscedasticity and autocorrelation consistent standard errors (HAC errors) based on the Newey and West (1987) estimator of the covariance matrix (five lags).

For the vast majority of currency pairs regression coefficients bear the expected signs summarised in Table 4: b_1^r coefficients are negative and entail short-term mean reversion, while $c_0^{j,r}$ coefficients are positive and in line with the law of demand and supply. The true beauty of the log-level model in Table 4 is its interpretability: coefficients can be interpreted as percentage changes in the dependent variable for a one-unit-change of the independent variable.[22] The coefficients at longer lags alternate in sign and decay to zero after a few lags. From these results it stands out that all agents except for corporates have a significantly positive contemporary price impact.

For some currency pairs (e.g. EURGBP, EURNOK, EURUSD etc.) corporates experience significantly negative contemporary price impact parameters. The negative β_0^{CO} is consistent with Menkhoff et al. (2016) who analyse the customer order flow of a given bank. Our results point to the importance of this issue at a global scale. Rather than from informational motives, a negative relation between order flow and return arises from liquidity needs (Grossman and Miller (1988)) and dealers' inventory reasons (Stoll (1978)). Thus, corporate trading seem to be driven by risk sharing, hedging, and liquidity issues (i.e. implicitly paying an insurance premium), and additional costs unrelated to adverse selection. This idea squares well with the different timing in their trading behaviour (see Figure 1). Whereas banks and other financial institutions access a richer information set by trading around the clock, the trading activity of corporates is more segmented and limited within few hours. The negative β_0^{CO} is also consistent with two common practices applied by FX dealers time restriction when they need to fulfil their liquidity needs: First, dealers may apply discount fees to beg customers liquidity provision when they are in liquidity needs. Second, dealers offset order flows coming from

[22]Results are very similar when we use (signed) net volume (without order size variable $\tilde{S}_t^j$), calculated as the net of buy volume by price takers minus the sell volume by market maker transactions, broken down into types of market participants, instead of (binary) order flow as well as using transaction prices instead of mid-quotes for calculating r_t in Eq. (4.1). See the online Appendix for further results.

potentially more informed agents (e.g. other banks and financial firms) with the non-informative one from corporates to reduce or neutralise their exposure to asymmetric information risk. The negative correlations between corporates' order flow and that of other financial agents reported above are fully in line with this picture. The negative b_1^r coefficients entail negative autocorrelation in the quote revision. A negative relation between trades and lagged quote revision is consistent with inventory control hypothesis, since market makers reduce (increase) their quotes to stimulate more purchases (sales). The coefficients of the return over the previous day (υ_1) is negative and highly significant for all currency pairs, whilst the return over the prior week (υ_2) is negative, but insignificant for the majority of currency pairs.

Table 5 summarises the order flow equation coefficients, that too bear the expected signs: b_1^T is negative and highly significant, while $c_1^{j,T}$ coefficients are positively significant and reflect a strong positive autocorrelation in trades. For most currency pairs, $c_1^{j,r}$ is positive but statistically not always significant. This is consistent with the findings in the stock market literature e.g. Hasbrouck and Ho (1987), Hasbrouck (1988), Madhavan et al. (1997) and shows that purchases tend to follow purchases and similarly for sales. Rather than with inventory control mechanisms, the short-run predominance of positive autocorrelation can be reconciled with delayed price adjustments to new information. b_1^T again implies negative autocorrelation in the quote revisions. In the order flow equation estimation this implies Granger-Sims causality running from quote revisions to trades. This causality is in line with current microstructure theory, where a negative relation between trades and lagged quote revisions is consistent with inventory control effects and/ or the price experimentation hypothesis formulated by Leach and Madhavan (1992), in which the market maker sets quotes to extract information optimally from the traders.

For both the return and order flow equation, hourly dummies are mostly insignificant with a couple of exceptions at the opening/ closing of the Asian, European, and American market places. Order size coefficients are mostly positive and significant, but around a fraction of a BPS. Thus, larger trades subsequently lead to a larger price impact, increasing the spread and the level of asymmetric information.

So far we have centred our analysis on the contemporary price impact. We now turn to the permanent component. The sums of $c_t^{j,r}$ and $d_t^{j,r}$ measure the persistent effects of the trade indicator T_t^j and trade size variables $\tilde{S}_t^j$. A positive $d_t^{j,r}$ indicates that order size, in addition to trade direction, conveys information. Furthermore, in the model of Hasbrouck (1991a) α_m^j can be interpreted as the measure of asymmetric/ private information because trades are driven by a mixture of private (superior) information and liquidity needs rather than public information. Therefore, any persistent impact of a trade on price arises from asymmetric information signalled by that trade. This intuition is reflected in Eq. (4.4), which identifies all public information with the quote revision innovation ($\epsilon_{r,t}$) and all private information with the trade innovation ($\epsilon_{\vec{T},t}$). The dichotomy above ensures that $\epsilon_{\vec{T},t}$ reflects no public information and hence the impulse response function α_m^j can be interpreted as a measure of asymmetric/ private information. In Table 6 we summarise the estimates of permanent price impact for every agent and currency pair. In general, we find that the permanent (cumulative) price impact parameters are positive and significant across agents (with a handful of exceptions), except for corporates for the above-mentioned reasons. Hence, the positive (negative) α_m^j again reflects that order flows coming from financial firms (corporates) are informative (non-informative).

5.2 Heterogeneous Price Impact Across Agents

The question we address here is whether the estimates of price impact are significantly different across agents. In this and all subsequent sections we focus solely on the permanent price impact parameter α_m as it was defined in Section 4. The following results are de facto identical for both the contemporary and permanent price impact. Thus, the online Appendix collects output tables and all technical details.[23]

To assess if the permanent price impact parameter α_m^j significantly differs across agents, we test if all coefficients in Eq. (4.5) for a particular agent i are jointly significantly different from agent j's. The validity of this pairwise F-test is ensured by the general property of the log-level regression model, where the change in the independent variable by one unit can be approximately interpreted as an expected change by $10,000 \times$ regression coefficient $\times$ BPS in the dependent variable. For nearly every pairwise combination of agents we clearly reject the H_0 at a 5% global significance level. To overcome the curse of multiple testing, a Bonferroni correction is applied.

5.3 Fragmentation in the FX Spot Market Across Currencies

Another important question is whether the price impact varies across currency pairs. We find the global FX market to be fragmented in the sense that a particular agent i has a significantly different price impact parameter (both $c_0^{j,r}$/ α_m^j) across currency pairs. As before, we estimate Eq. (4.1) on the full sample and construct a pairwise F-test where

[23] See Tables B.7 and B.15 on page 12 and on page 22 in the online Appendix.

we test if all coefficients in Eq. (4.5) for a particular agent $i \in C = \{CO, FD, NB, BA\}$ are jointly significantly different in currency pair k than q. Again, the validity of this t-test is warranted by the general properties of a log-level model. For technical details and output tables see the online Appendix.[24]

The main result that emerges from this analysis is that corporates, funds, non-bank financials, and banks acting as price takers have a permanent price impact α_m^j which varies heavily across currencies. However, this type of fragmentation appears to be less pronounced for corporates than for funds or non-bank financials. The reasons for this are twofold: First, corporates are mostly active in the swap rather than spot market. Second, funds and non-bank financials are sophisticated and highly specialised investors who frequently trade on superior information that may well vary across currencies. All in all, our empirical analysis confirms earlier research on customer order flow (e.g. Evans and Lyons (2006), Osler et al. (2011), Menkhoff et al. (2016)). Furthermore, our results evidence that non-financial customer do not trade strategically but are rather hedging.

To sum up, two main results emerge: First, order flow impacts FX prices heterogeneously across agents. Second, the FX spot market suffers from fragmentation in the sense that the same agent has both a different contemporary and permanent price impact across currency pairs. Understanding the cross-sectional and time related difference across market participant and currencies constitutes an important and prosperous avenue for future research.

5.4 Time Varying Information Flows

In this section we introduce *time* as a third dimension of heterogeneity and study the time variation of both the contemporary and permanent price impact. Again we estimate Eq. (4.4) by OLS, but now in a rolling window fashion instead of using the full sample. We choose the rolling window to be one year but our results are robust to shorter horizons.

In Figure 3 we plot both the average $c_0^{j,r}$ and α_m^j across currency pairs over time. The plotted averages exclude any coefficients that are either heavy outliers with respect to the median or not significant at a 95% confidence level applying a simple two-sided t-test and the same joint F-test as in Table 6, respectively. Corporates seem to have the strongest time variation consistent with the idea that their trades are driven by uninformative reasons (market risk, hedging, or liquidity shocks) rather than a system-

[24] See Tables B.11 and B.19 on page 17 and on page 27 in the online Appendix.

atic processing of superior information. We use the Brown-Forsythe test to verify if corporates' price impact parameter exhibits a significantly higher variance than funds', non-bank financials' or banks'. For the vast majority of currency pairs we reject the null hypothesis of homoscedasticity across agents' price impact parameter at a 95% confidence level for all pairwise combinations. This is true for both the contemporary and permanent price impact.

The main difference is that the permanent price impact of sophisticated agents such as funds and banks is small on average across time, while financially less literate agents (corporates) experience stronger time variation in their permanent price impact. This is likely to reflect funds' and banks' superior financial sophistication to engage in strategic order submission behaviours such as order splitting and price impact smoothing.

6 Currency Portfolios

In the foregoing sections we have studied how order flow impacts FX spot prices heterogeneously. In the remaining part of this paper, we address the question whether this heterogeneity provides significant economic value. To do this, we introduce a simple, yet innovative trading strategy based on *UIP* deviations that exploits the (persistent) price impact heterogeneity.

6.1 Trading Strategy

Hasbrouck (1991a) demonstrates that any permanent price impact of a trade must arise from superior information about the future evolution of a security price. To capitalise on asymmetric information, a coherent trading strategy should apply this method to timely detect order flows conveying superior information across agents and assets. In the context of the global FX market, we consistently apply this idea by introducing a novel long-short trading strategy based on a simple idea: Order flows of agents and currencies impounding a persistent price impact convey superior information leading to better predictions of future evolutions of FX rates. The intuition behind this strategy relies on well-documented deviations from the Uncovered Interest rate Parity (*UIP*) condition and forward premium puzzle: When regressing FX returns on interest rate differentials the slope coefficient is typically not equal to one but negative.[25] In other

[25]See Hansen and Hodrick (1980), Fama (1984), Hodrick (1988), Lewis (1995), and Lustig and Verdelhan (2007) for extensive surveys and updated regression results on *UIP* deviations.

words, the forward premium points into the 'wrong' direction of the expected price movement.

Following the empirical work by Lustig and Verdelhan (2007) and Lustig et al. (2011) a straightforward interpretation of our trading strategy arises: Given that our measure of persistent price impact captures order flows conveying superior information (net of temporary liquidity effects), it is naturally well suited to identify trading that correctly predicts currency values or conversely, that is more biased by *UIP* deviations. As a consequence, currency pairs with a high (*low or negative*) aggregate permanent price impact are more likely to deviate from the *UIP* in the sense that $f_{t,t+1} \geqslant s_{t+1}$ ($f_{t,t+1} \leqslant s_{t+1}$), i.e. the log forward rate ($f_{t,t+1}$) 'overshoots' (*'undershoots'*) the future log spot rate (s_{t+1}). With this intuition in mind, the excess returns of this trading strategy are sourced by asymmetric (private) information and do not constitute compensation for risk nor are driven by contemporary liquidity effects.[26]

To be more precise, the long-short strategy (hereinafter ALP_{HML}) rests on five pillars: timing, weighting, rebalancing, signal extraction, and excess returns. Investment takes place *instantaneously* on the same day as the signal is extracted.[27] Throughout the entire investment period the strategy exhibits *equally weighted* long and short legs resulting into zero net exposure. To make our results comparable to other common FX risk factors (see Lustig et al. (2011), Menkhoff et al. (2017)) we form quintile portfolios ($Q_1, Q_2, ..., Q_5$) and build cross-sections of currency portfolios.[28] Portfolio rebalancing takes place at the *end* of every day, week, or month.

Trading signals are generated from estimating Eq. (4.1) in an twelve months *rolling window* fashion at daily frequency based on binary order flow and mid-quotes with the number of lags equal to five days.[29] The advantage of running this regression at daily rather than hourly frequency is two-fold: First, it is computationally less expensive and hence, easily replicable for global investors. Second, forward rates are usually not readily available at an hourly frequency and therefore using daily data ensures that signals are extracted at the same frequency as excess returns. Hence, investment starts in August 2013 after one year of formation period. This leaves us four years to test out-of-sample

[26]Due to the lack of order flow data being available in the FX market these returns were not arbitraged away in the past but are likely to vanish in the future.
[27]Results are robust to investing with a lag of one day up to a week.
[28]Lustig and Verdelhan (2007) were the first to build cross-sections of currency portfolios.
[29]The trading strategy is robust to our choice of model specification i.e. (signed) net volume instead of binary order flow and mid-quotes instead of transaction prices. In particular, it renders positive and significant returns for several different combinations of baseline VAR model, rolling window length, and number of lags.

performance. For every rolling window index and currency pair k we obtain the aggregate permanent price impact α_m^k (see Eq. (4)). All $\alpha_m^{j,k}$s that are insignificant at a 5% global significance level, i.e. individual test levels were Bonferroni corrected, are excluded. The intuition being that insignificant $\alpha_m^{j,k}$s do not convey any new, additional information.[30] Next, we sort α_m^k across currency pairs by size and in ascending order. The ALP_{HML} portfolio consists of the 20% highest (*lowest*) α_m^k currency pairs in the long (*short*) leg.

Following the FX asset pricing literature, going long or short in a specific currency pair involves forward positions. Therefore, the log *excess* return rx of buying a foreign currency in the forward market and selling it in the spot market after one day, week, or month is:

$$rx_{t+1} = f_{t,t+1} - s_{t+1} - \Delta s^*_{t,t+1}, \tag{6.1}$$

assuming that under normal conditions forward rates satisfy the Covered Interest rate Parity (*CIP*) condition.[31] $f_{t,t+1}$ denotes the log forward rate and s_t the log spot rate, both in units of the foreign currency per US dollar.[32] The last term is the change in the *USDXXX* spot rate. *XXX* being the base currency of a non-US currency pair such as the *EURGBP* (which is *EUR* in this case). Hence, $\Delta s^*_{t,t+1}$ applies only when investing in non-US currency pairs, because we assume that a US-investor has only US dollars. Since we have bid-ask quotes for spot and forward contracts[33] we can compute the investor's actual realised excess return net of transaction costs in the spirit of Lustig et al. (2011). The *net* log currency excess return for an investor who goes long in foreign currency x is:

$$rx_{t+1} = f^b_{t,t+1} - s^a_{t+1} - \mathbb{1}_{x \notin USD} \Delta s^{*,b,a}_{t,t+1}, \tag{6.2}$$

where $\mathbb{1}_{x \notin USD}$ is one if the currency pair is non-US and zero otherwise. x is the currency pair of interest and USD is the basket of all currency pairs, where the US dollar is in the base. The investor buys the foreign currency or equivalently sells the dollar forward at the bid price $f^b_{t,t+1}$ in period t, and sells the foreign currency or equivalently buys

[30] All our results remain qualitatively unchanged when we do not exclude any price impact coefficients.

[31] Akram et al. (2008) study high-frequency deviations from Covered Interest rate Parity (*CIP*). They conclude that *CIP* holds at daily and lower frequencies.

[32] Daily, weekly, and monthly forward bid and ask points are obtained from Bloomberg. Forward rates can be expressed as the forward discount/ premium (or forward points alternatively) plus the spot rate. Therefore, the simple (outright) forward bid and ask rates are $F^b_t = S^b_t + P^b_t$ and $F^a_t = S^a_t + P^a_t$, respectively, where P^b_t and P^a_t denote the bid and ask values of forward points.

[33] The bid-ask spread data are available for quoted spreads and not effective spreads. To be conservative, unlike earlier work (for example Goyal and Saretto (2009), Gilmore and Hayashi (2011), Menkhoff et al. (2016) and Gargano et al. (2018)) we do *not* employ 50% of the quoted bid-ask spread as the actual spread. The online Appendix shows the trivial increase in performance when applying the 50% rule to proxy effective spreads.

dollars at the ask price s^a_{t+1} in the spot market in period $t+1$. Similarly, for an investor who is long in the US dollar (and thus short the foreign currency) the net log currency excess return is given by:

$$rx_{t+1} = -f^a_{t,t+1} + s^b_{t+1} + \mathbb{1}_{x \notin USD} \Delta s^{*,a,b}_{t,t+1}, \tag{6.3}$$

and the (simple) portfolio return RX^p is given by:

$$RX^p_{t+1} = \sum_{k=1}^{K_t} w_{k,t+1} RX_{k,t+1}, \tag{6.4}$$

where $RX_{k,t+1}$ is a vector of simple returns based on Eq. (6.2) and Eq. (6.3), since log returns are not asset additive.

6.2 Trading Performance

In Tables 7 and 8 we present the annualised Sharpe ratio (SR), the annualised mean excess return (*Mean*), the maximum drawdown (MDD) and the Θ performance measure of Goetzmann et al. (2007) based on monthly rebalancing for both a US (USD) and European (EUR) investor perspective, respectively.[34] For both tables, Panel a) tabulates the five quintile portfolios, where the last column is a linear combination of going short the first and long the fifth quintile that takes into account the effect of compounding, whereas Panel b) benchmarks ALP_{HML} to common FX trading strategies. To overcome the curse of heteroscedasticity and autocorrelation we apply HAC errors using the plug-in procedure for automatic lag selection by Andrews and Monahan (1992) and Newey and West (1994).

First, we wish to compare performance to a pure order flow based strategy prior transaction costs. Hence, following the identical methodology as in Menkhoff et al. (2016), we construct a trading strategy (BMS) based on aggregate standardised *total* order flow and compare it to ALP_{HML}. We claim that the advantage of using a VAR model to measure the permanent price impact is its ability to separate temporary liquidity effects of order flow from persistent ones based on informational motives. Consequently, the predictive power of the permanent price impact on exchange rate changes should be higher and thus outperform a pure order flow based trading strategy such as BMS.

From Table 7 it is discernible that ALP_{HML} clearly outperforms BMS from a US/

[34]Prior transaction costs, trading performance remains similar for weekly and daily returns, but erodes significantly when transaction costs are taken into consideration.

European investor perspective. The main reason being that α_m measures the long-lived (ultimate) information effect of a trade net of temporary liquidity effects, while order flow itself can arise from both information and non-informational motives. Given that the dataset analysed in Menkhoff et al. (2016) is different and inaccessible, a direct one to one comparison is not permissible.[35]

In Table 8 we concentrate on after transaction cost performance of ALP_{HML}. From Panel a) and b), respectively, two main results emerge: First, a high performance of the ALP_{HML} strategy from both the US and European investor perspective. Second, our strategy clearly outperforms common FX risk factor strategies based on home base currency pairs i.e. DOL[36], the real exchange rate i.e. RER/ RER_{HML}[37], or momentum i.e. MOM_{HML}/ CAR_{HML}[38] (see Panel b)).

Note that both MOM_{HML} and CAR_{HML} are momentum based strategies, where latter is the 'classic' carry trade.[39] In line with, empirical research e.g. Lustig and Verdelhan (2007), CAR_{HML} generates negative excess returns and Sharpe ratios after transaction costs. This is presumably due to the negative relationship between exchange rate changes and the interest rate differential/ forward premium or discount.

Figure 5 depicts the cumulative (simple excess) returns of different rebalancing frequencies prior and after transaction costs. Gross returns are based on mid-quotes for both spot and forward rates. Daily and weekly rebalancing are substantially less profitable than monthly due to high transaction costs, but bear similar cumulative returns

[35] The differences are twofold: First, our data comes from a non-bank specific source that dissects the aggregate (signed) net volume into different customer end-user segments for a different, non-overlapping period in time. In particular, long-term and short-term demand-side investment managers comprise the largest players of their class, while funds and non-bank financials also incorporate small- and medium-sized investment managers. Second, Menkhoff et al. (2016) focus on daily rebalancing, while we perform monthly rebalancing due to significant transaction costs.

[36] The DOL portfolio consists of equally weighted long USD (or EUR) currency pairs. The portfolio return is given by Eq. (6.4).

[37] The RER and RER_{HML} are constructed based on Menkhoff et al. (2017), where currencies with a real exchanger rate above (*below*) the cross-sectional average are gone long (*short*). These weights are scaled such that their absolute sum equals unity. For RER_{HML} we rank the currency pairs based on the real exchange rate and form a "high minus low" portfolio in the sense that the top (*bottom*) quintile currencies receive a positive (*negative*) weight.

[38] The MOM_{HML} strategy involves a currency sorting based on the realised *CIP* deviation ($f^m_{t-1,t} - s^m_t$, where m stands for mid-quote) and goes long (*short*) currencies in the top (*bottom*) quintile, see Asness et al. (2013). For CAR_{HML} (see Lustig et al. (2011)) currency pairs are sorted based on the forward discount/ premium ($f^m_{t,t+1} - s^m_t$) and again a "high minus low" portfolio is formed where the long (*short*) leg consists of the 20% currency pairs with the highest forward premium (*discount*).

[39] MOM_{HML} assigns a positive (*negative*) weight to currencies that undershot (*overshot*) the future spot rate implied by *CIP* in the past period and bets that the direction of this deviation is going to persist over the next period. CAR_{HML} assigns a positive (*negative*) weight to currencies experiencing a forward premium (*discount*), hoping that they appreciate (*depreciate*) over the next period.

prior transaction costs. The investment period is the entire sample period (September 2012 to November 2017) minus twelve months of formation period to retrieve the first trading signal and therefore spans from August 2013 to November 2017. As discernible in Figure 5, cumulative returns seem to increase steadily over time and do not experience any regime switches.

In general, our results show that patient, less frequently rebalancing investors are rewarded both prior and after transaction costs by higher returns. In addition to the cumulative returns the maximum drawdown curves are constructed. This drawdown measure corresponds to the cumulative return of the ALP_{HML} portfolio relative to the last peak. With monthly rebalancing, the ALP_{HML} strategy is able to beat itself over extended periods of time and exhibits a maximum drawdown of 3.50% (*4.22%*) in a prior (*after*) transaction costs setting.

Finally, to overcome the statistical limitations of a relatively short out-of-sample period we use standard bootstrap techniques. Figure 7 presents bootstrapped p-values for ALP_{HML} prior and after transaction costs, respectively, for a US-investor perspective using $1,000$ bootstrap repetitions. Each of the four quarters displays one of the performance measures introduced at the beginning of this section. Taken together, these distributions underline the statistical significance of ALP_{HML} and confirm the appropriate use of asymptotic p-values. See the online Appendix for further distribution figures and outputs equivalent to Tables 7 and 8 but based on bootstrapped p-values.

6.3 Portfolio Weights

This subsection sheds light on the currency exposure of ALP_{HML} (US perspective) and analyses the decomposition of the long and short leg. As illustrated in Figure 4, this analysis delivers two main findings: First, our trading strategy exhibits a balanced exposure across currency pairs, where at least half of the currency pairs receive an average absolute weight of $\geqslant 2\%$ over time. The relatively dominant currency pairs are the AUDJPY, AUDNZD, GBPCHF, GBPJPY, and USDZAR.

Second, in Figure 9 we calculate the relative contribution of every agent's $\alpha_m^{j,k}$ to the aggregate permanent price impact α_m^k per currency pair and then take the average across all currencies for ALP_{HML} with monthly rebalancing. This figure clearly shows that all groups of market participants are represented providing further evidence of heterogeneous information across market participants. This picture is in line with our empirical finding that corporate trading is largely uninformative for predicting the evolution of FX rates but can indeed be informative for the short leg. Moreover, both the long and short leg appear to be equally balanced across agents. This result is not surprising if compared to Figure 3: Both sophisticated agents (funds, non-bank financials, and banks acting as price takers) and financially less literate agents (corporates) experience significant time variation of their permanent price impact, respectively.

6.4 Exposure Regression

In this section we address the question whether gross returns of ALP_{HML} (US perspective) can be explained by any of the common FX risk factors presented in Lustig et al. (2011), Asness et al. (2013) and Menkhoff et al. (2016, 2017). In Table 9, we regress monthly portfolio returns of the ALP_{HML} strategy on the monthly net excess returns associated with common risk factors: DOL, RER_{HML}, RER, MOM_{HML}, CAR_{HML}, and BMS as well as the return on the VIX index and the change in the iTraxx Europe CDS index. The regressions are based on simple gross excess returns i.e. Eq. (6.4).

The low R^2 is a clear indication for the low explanatory power of these common FX risk factors. In particular, the variation in excess returns of ALP_{HML} cannot be explained by traditional FX momentum (MOM_{HML}) and is negatively related to carry trade (CAR_{HML} à la Lustig et al. (2011)). The trading strategy generates a significant Jensen's alpha (α) of c. 60-80 BPS per month and information ratios (IR) of c. 30-37%, where the information ratio is defined as α divided by the residual standard deviation.

The relatively low correlations (see Table 10) between the monthly returns of ALP_{HML} and other common risk factors underpin our findings in Table 9. The correlation is negligible/ negative for momentum based ones (MOM_{HML} and CAR_{HML}). Consistent with the asymmetric information hypothesis, ALP_{HML} returns are more correlated to factors related to (currency) fundamental values, i.e. the real exchange rate (RER_{HML}, RER). As expected, ALP_{HML} is unrelated to standardised *total* order flow (BMS). In addition, ALP_{HML} relates positively to returns on the VIX index but negatively to changes in the CDS spread.

6.5 Transaction Costs

Table 11 summarises transaction costs associated with different rebalancing frequencies for the ALP_{HML} strategy. Transaction costs are defined as the difference in the excess return per annum over n successive days, weeks, or months with and without bid/ ask spreads. Both average and median transaction costs are substantially higher and more volatile when rebalancing occurs more frequently. This empirical observation is in line with the general property of daily data being more vulnerable to statistical noise or stale pricing that does not reflect economic value per se. However, the differences on a cost per transaction basis are negligible.

In Figure 10 we plot the empirical distribution of annual transaction costs for daily, weekly, and monthly rebalancing frequencies. The histograms emphasise that transaction costs are largely innocuous on a monthly basis, but severely kick in at higher rebalancing frequencies (daily/ weekly).

Rolling Over Forward Contracts. Gilmore and Hayashi (2011) introduce the concept of rolling over an open long/ short position instead of opening a new position and unwinding the old one. In other words, the investor opens a long position via a one-day, one-week or one-month forward outright contract in day, week, or month zero, maintains the position for n successive months via foreign exchange swaps, and then unwinds in day, week, or month n. Compared to the excess return prior transactions costs calculated from mid-quotes, the investor pays the difference between the bid and mid-rates (which equals half times the bid/ ask spread) when opening the position in day, week, or month 0, the difference between the offer and mid-rates when unwinding the position in day, week, or month n, and a daily, weekly, or monthly 'roll cost' in between. The roll cost is the difference between the bid and mid of the foreign exchange forward points of foreign exchange swaps. Typically, this will be far lower than the difference between the bid and mid of the forward (outright) rate.

In the spirit of Gilmore and Hayashi (2011), we derive the transaction costs for rolling over n successive periods for both a single currency and a portfolio of currency pairs. For technical details and derivations refer to the online Appendix, where we summarise the performance of the ALP_{HML} strategy as well as benchmark it to common FX strategies, when long/ short positions are rolled over instead of unwound and reopened. With monthly rebalancing performance improvement is small compared to daily or weekly rebalancing frequencies.

The reasons for this are twofold: First, on a monthly basis we have 51 rebalancing points over the entire investment period and therefore transaction costs are less weighty. Figure 5 underlines this point. Second, given that our portfolio is well diversified across a large cross-section of currency pairs the probability that the weights associated with currency pair in period $t-1$ and t coincide is low in general.

Similarly to Gargano et al. (2018), we find that transaction costs (i.e. roll costs) based on WM/Reuters data (see online appendix Gilmore and Hayashi (2011)) are higher than based on Olsen data (see Figure 11). Annual average roll costs for our cross-section of currency pairs are in the ballpark of a fraction of one percent, since the roll cost Z_t is effectively half of the relative forward points spread $(P_t^a - P_t^b)$ multiplied by the number of trading days, weeks, or months per investment period.

Beyond any doubt, the cost of rolling different currencies can vary considerably. Nevertheless, this does not impinge on the goal of this section to prove that it is typically cheaper to roll a position than to close it and then reopen it. Finally, in this section we have paid tribute to the importance of transaction costs and introduced a more sophisticated method of rolling over long and short positions. Especially at higher frequencies transaction costs can be reduced substantially.

6.6 Robustness Tests

We performed a number of additional analyses and robustness checks that we briefly summarise in this section. More detailed results are reported in the online Appendix. We focus on three of them: First, we test if cumulative returns are due to some periods performing extremely well, while others performing very poorly. Second, we explore the performance of the strategy using various sub-samples of currency pairs. Third, we check whether our results are sensible to including the contemporaneous price impact when calculating the permanent effect.

Rolling One Year Returns. Figure 12 demonstrates that our returns are robust to the length of investment period. Taking the *gross* returns from monthly rebalancing, we calculate the cumulative return over a period of 12 months in a rolling fashion. Cumulative annual returns are persistently positive. The numbers on the x-axis designate the starting month of the rolling window period, i.e. at tick 4, we measure the 12 months cumulative return for an investment starting in November 2013 and ending in October 2014 (month 16).[40]

Subsamples of Currencies. In Section 6.3 we have shown that the ALP_{HML} strategy exhibits a balanced exposure across currencies and over time. To alleviate any concerns that the economic profitability of ALP_{HML} is driven by just a few currency pairs, we twist our analysis by taking into account only a subset of the original 30 currency pairs. In Table 12 we report results for the annualised Sharpe ratio (SR), the annualised mean excess return (*Mean*), the maximum drawdown (MDD) and the Θ performance measure of Goetzmann et al. (2007) based on monthly rebalancing for both a US (USD)

[40] Note that our results remain robust when randomly splitting the sample into two to three non-overlapping periods.

and European (EUR) investor perspective.

There are five cases to be distinguished: First, USD currency pairs only (15 in total). Second, EUR and GBP currency pairs only (14 in total). Third, all pairs excluding emerging market currencies (i.e. USDILS, USDMXP, and USDZAR) and/ or fixed pairs (i.e. EURDKK, USDDKK, USDHKD, and USDSGD). Fourth, G10 currency pairs plus the most liquid EUR cross pairs (14 in total). Fifth, all currency pairs excluding the CHF crosses (27 in total). For each subsample the investment performance of ALP_{HML} remains economically and statistically significant at the 90% confidence level (either for US or EU perspective). Also, the development of their performance across time is similar. The annualised *gross* excess returns and Sharpe ratios range from 1.02% to 9.06% and 0.23 to 1.21, respectively. Note that the US/ EU differences can be traced back to additional transaction costs (i.e. $\Delta s^*_{t,t+1}$) that arise from trading in non USD/ EUR base currency pairs. Overall, the subsampling analysis alleviates three issues: First, our results are robust to the choice of currency pairs in the sample. Second, the ALP_{HML} performance does not seem to be driven by structural changes such as the implementation of new regulations (e.g. capital and liquidity requirements of Basel III). For instance, Du et al. (2018) show that *CIP* deviations concentrate at quarter-ends and in some currencies.[41] Moreover, the performance steadily increases over time and it is not originated by only a few currency pairs. Third, the ALP_{HML} performance is not affected by specific events such as the removal of the cap by the Swiss National Bank.

Excluding the Contemporary Price Impact. We next check whether our results are robust to including the contemporary price impact (c_0) in calculating the permanent price impact (α_m). To do this, we replicate our trading strategy ALP_{HML} based on signals $\alpha_n = \alpha_m - c_0$ ('residual price impact') and c_0, respectively, and compare both to trading based on α_m. In line with our prior, we find that both signals perform less well than trading on the permanent price impact. Therefore, we may conclude that trading on α_m, i.e. c_0 net of liquidity effects, maximises expected excess returns. However, excluding c_0 from α_m leaves our main results qualitatively unchanged.

To conclude, despite the relatively short sample period, our asset pricing analysis highlights the economic value of heterogeneous information contents of global FX trading. Our results are robust to our choice of currencies as well as the length of investment periods. As a result, global investors can effectively capitalise on information asymmetries across market participants.

[41] Du et al. (2018) document that high-interest-rate (low-interest-rate) currencies tend to exhibit a positive (negative) *CIP* basis i.e. the deviation from the CIP condition.

7 Conclusion

In this paper, we analyse the heterogeneous information content of global FX trading in an effort to improve our understanding of the world's largest financial market, the FX market. We address two main questions: First, given that the FX market hosts various types of market participants, does order flow impact FX prices heterogeneously? Second, do asymmetric information contents in global FX trading lead to a profitable trading strategy?

To answer these questions, we analyse a new and representative dataset of global FX order flows disaggregated by groups of market participants. We find compelling evidence that order flow impacts FX spot prices heterogeneously across agents, time, and currency pairs, supporting the asymmetric information hypothesis. In particular, corporates have a significantly lower contemporaneous and permanent price impact than funds, non-bank financials, or banks, suggesting (time-varying) asymmetric information across market participants and in the cross-section of FX rates.

To assess the economic value of order flow heterogeneity, we introduce a novel long-short trading strategy based on *UIP* deviations that exploits the persistent price impact. Given that our measure of permanent price impact captures the persistent effect of order flow net of temporary liquidity effects, it is naturally well suited to identify superior information that correctly predicts the evolution of FX rates and to puzzle out the forward premium bias. As a consequence, currency pairs with a high (*low or negative*) aggregate permanent price impact are more likely to deviate from the *UIP* in the sense that the forward rate 'overshoots' (*'undershoots'*) the future spot rate. Overall, the strategy generates an annualised excess return and a Sharpe ratio that are both economically and statistically significant, even after accounting for transaction costs. Furthermore, the returns generated by our strategy are unrelated to other common currency strategies and risk factors.

Our findings suggest that FX markets are still characterised by information asymmetries, heterogeneity, and fragmentation despite the ongoing efforts to redesign and regulate OTC markets including the Dodd-Frank Act, EMIR, and MiFID II. Future research should highlight whether the declared objectives such as an increase of transparency, price efficiency, and fairness have yet to be achieved or have produced the suited effects only on some market segments.

References

Akram, Q. F., Rime, D., and Sarno, L. (2008). Arbitrage in the foreign exchange market: Turning on the microscope. *Journal of International Economics*, 76(2):237–253.

Andrews, D. W. K. and Monahan, J. C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, 60(4):953.

Asness, C. S., Moskowitz, T. J., and Pedersen, L. H. (2013). Value and momentum everywhere. *The Journal of Finance*, 68(3):929–985.

Bacchetta, P. and van Wincoop, E. (2005). Can information heterogeneity explain the exchange rate determination puzzle? *American Economic Review*, 96(3).

Barclay, M. J. and Hendershott, T. (2003). Price discovery and trading after hours. *Review of Financial Studies*, 16(4):1041–1073.

Berger, D. W., Chaboud, A. P., Chernenko, S. V., Howorka, E., and Wright, J. H. (2008). Order flow and exchange rate dynamics in electronic brokerage system data. *Journal of International Economics*, 75(1):93–109.

Berry, T. D. and Howe, K. M. (1994). Public information arrival. *The Journal of Finance*, 49(4):1331–1346.

Bessembinder, H. (1994). Bid-ask spreads in the interbank foreign exchange markets. *Journal of Financial Economics*, 35(3):317–348.

BIS (2016). Triennial central bank survey - foreign exchange turnover in April 2016. *Bank for International Settlements (BIS) Triennial Survey*.

Bjønnes, G. H. and Rime, D. (2005). Dealer behavior and trading systems in foreign exchange markets. *Journal of Financial Economics*, 75(3):571–605.

Bollerslev, T. and Melvin, M. (1994). Bid—ask spreads and volatility in the foreign exchange market. *Journal of International Economics*, 36(3-4):355–372.

Breedon, F., Chen, L., Ranaldo, A., and Vause, N. (2018). Judgement day: Algorithmic trading around the swiss franc cap removal. *SSRN Electronic Journal*.

Breedon, F. and Ranaldo, A. (2013). Intraday patterns in FX returns and order flow. *Journal of Money, Credit and Banking*, 45(5):953–965.

Breedon, F. and Vitale, P. (2010). An empirical study of portfolio-balance and information effects of order flow on exchange rates. *Journal of International Money and Finance*, 29(3):504–524.

Chaboud, A. P., Chiquoine, B., Hjalmarsson, E., and Vega, C. (2014). Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance*, 69(5):2045–2084.

Chinn, M. D. and Moore, M. J. (2011). Order flow and the monetary model of exchange rates: Evidence from a novel data set. *Journal of Money, Credit and Banking*, 43(8):1599–1624.

Christiansen, C., Ranaldo, A., and Söderlind, P. (2011). The time-varying systematic risk of carry trade strategies. *Journal of Financial and Quantitative Analysis*, 46(04):1107–1125.

Della Corte, P., Ramadorai, T., and Sarno, L. (2016a). Volatility risk premia and exchange rate predictability. *Journal of Financial Economics*, 120(1):21–40.

Della Corte, P., Riddiough, S. J., and Sarno, L. (2016b). Currency premia and global imbalances. *Review of Financial Studies*, 29(8):2161–2193.

Ding, D. K. (1999). The determinants of bid-ask spreads in the foreign exchange futures market: A microstructure analysis. *Journal of Futures Markets*, 19(3):307–324.

Du, W., Tepper, A., and Verdelhan, A. (2018). Deviations from covered interest rate parity. *The Journal of Finance*, 73(3):915–957.

Duffie, D., editor (2012). *Dark Markets.* Asset Pricing and Information Transmission in Over-the-Counter Markets. Princeton University Press.

Easley, D. and O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19(1):69–90.

Easley, D. and O'Hara, M. (1991). Order form and information in securities markets. *The Journal of Finance*, 46(3):905–927.

Engle, R. F., Ito, T., and Lin, W.-L. (1990). Meteor showers or heat waves? heteroskedastic intra-daily volatility in the foreign exchange market. *Econometrica*, 58(3):525–542.

Evans, M. D. (2002). FX trading and exchange rate dynamics. *The Journal of Finance*, 57(6):2405–2447.

Evans, M. D. (2010). Order flows and the exchange rate disconnect puzzle. *Journal of International Economics*, 80(1):58–71.

Evans, M. D. and Lyons, R. K. (2002). Order flow and exchange rate dynamics. *Journal of Political Economy*, 110(1):247–290.

Evans, M. D. and Lyons, R. K. (2005). Do currency markets absorb news quickly? *Journal of International Money and Finance*, 24(2):197–217.

Evans, M. D. and Lyons, R. K. (2006). Understanding order flow. *International Journal of Finance & Economics*, 11(1):3–23.

Evans, M. D. and Lyons, R. K. (2008). How is macro news transmitted to exchange rates? *Journal of Financial Economics*, 88(1):26–50.

Fama, E. F. (1984). Forward and spot exchange rates. *Journal of Monetary Economics*, 14(3):319–338.

Fischer, A. M. and Ranaldo, A. (2011). Does fomc news increase global FX trading? *Journal of Banking & Finance*, 35(11):2965–2973.

Foster, F. D. and Viswanathan, S. (1993). Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models. *The Journal of Finance*, 48(1):187.

Galati, G. (2002). Settlement risk in foreign exchange markets and CLS Bank. *BIS Quarterly Review*.

Gargano, A., Riddiough, S. J., and Sarno, L. (2018). The value of volume in foreign exchange. *SSRN Electronic Journal*.

Gerety, M. S. and Mulherin, J. H. (1992). Trading halts and market activity: An analysis of volume at the open and the close. *The Journal of Finance*, 47(5):1765–1784.

Gilmore, S. and Hayashi, F. (2011). Emerging market currency excess returns. *American Economic Journal: Macroeconomics*, 3(4):85–111.

Glosten, L. R. and Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100.

Goetzmann, W., Ingersoll, J., Spiegel, M., and Welch, I. (2007). Portfolio performance manipulation and manipulation-proof performance measures. *Review of Financial Studies*, 20(5):1503–1546.

Goyal, A. and Saretto, A. (2009). Cross-section of option returns and volatility. *Journal of Financial Economics*, 94(2):310–326.

Grossman, S. J. and Miller, M. H. (1988). Liquidity and market structure. *The Journal of Finance*, 43(3):617–633.

Hansen, L. P. and Hodrick, R. J. (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy*, 88(5):829–853.

Hartmann, P. (1999). Trading volumes and transaction costs in the foreign exchange market. *Journal of Banking & Finance*, 23(5):801–824.

Hasbrouck, J. (1988). Trades, quotes, inventories, and information. *Journal of Financial Economics*, 22(2):229–252.

Hasbrouck, J. (1991a). Measuring the information content of stock trades. *The Journal of Finance*, 46(1):179–207.

Hasbrouck, J. (1991b). The summary informativeness of stock trades: An econometric analysis. *Review of Financial Studies*, 4(3):571–595.

Hasbrouck, J. and Ho, T. S. Y. (1987). Order arrival, quote behavior, and the return-generating process. *The Journal of Finance*, 42(4):1035–1048.

Hasbrouck, J. and Levich, R. M. (2018). FX market metrics: New findings based on CLS bank settlement data. *SSRN Electronic Journal*.

Hendershott, T., Jones, C. M., and Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *The Journal of Finance*, 66(1):1–33.

Hodrick, R. J. (1988). The empirical evidence on the efficiency of forward and futures foreign exchange markets. *Journal of International Economics*, 25(3-4):387–390.

Holden, C. W. and Subrahmanyam, A. (1992). Long-lived private information and imperfect competition. *The Journal of Finance*, 47(1):247–270.

Hsieh, D. A. and Kleidon, A. W. (1996). Bid-Ask Spreads in Foreign Exchange Markets: Implications for Models of Asymmetric Information. In *The Microstructure of Foreign Exchange Markets*, NBER Chapters, pages 41–72. National Bureau of Economic Research, Inc.

Huang, R. D. and Masulis, R. W. (1999). FX spreads and dealer competition across the 24-hour trading day. *Review of Financial Studies*, 12(1):61–93.

Ito, T., Lyons, R. K., and Melvin, M. T. (1998). Is there private information in the FX market? the tokyo experiment. *The Journal of Finance*, 53(3):1111–1130.

Jain, P. C. and Joh, G.-H. (1988). The dependence between hourly prices and trading volume. *The Journal of Financial and Quantitative Analysis*, 23(3):269.

Jones, C. M., Kaul, G., and Lipson, M. L. (1994). Information, trading, and volatility. *Journal of Financial Economics*, 36(1):127–154.

King, M. R., Osler, C., and Rime, D. (2012). *Foreign Exchange Market Structure, Players, and Evolution*. John Wiley & Sons, Inc.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6):1315.

Leach, J. C. and Madhavan, A. N. (1992). Intertemporal price discovery by market makers: Active versus passive learning. *Journal of Financial Intermediation*, 2(2):207–235.

Levich, R. (2013). Interest rate parity. In *The Evidence and Impact of Financial Globalization*, pages 417–427. Elsevier.

Levich, R. M. (2012). FX counterparty risk and trading activity in currency forward and futures markets. *Review of Financial Economics*, 21(3):102–110.

Lewis, K. K. (1995). Chapter 37, Puzzles in International Financial Markets. In *Handbook of International Economics*, pages 1913–1971. Elsevier.

Lindley, R. (2008). Reducing foreign exchange settlement risk. *BIS Quarterly Review*.

Love, R. and Payne, R. (2008). Macroeconomic news, order flows, and exchange rates. *Journal of Financial and Quantitative Analysis*, 43(02):467.

Lustig, H., Roussanov, N., and Verdelhan, A. (2011). Common risk factors in currency markets. *Review of Financial Studies*, 24(11):3731–3777.

Lustig, H. and Verdelhan, A. (2007). The cross section of foreign currency risk premia and consumption growth risk. *American Economic Review*, 97(1):89–117.

Lyons, R. K. (2006). *The Microstructure Approach to Exchange Rates (MIT Press)*. The MIT Press.

Madhavan, A., Richardson, M., and Roomans, M. (1997). Why do security prices change? a transaction-level analysis of NYSE stocks. *Review of Financial Studies*, 10(4):1035–1064.

Mancini, L., Ranaldo, A., and Wrampelmeyer, J. (2013). Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums. *The Journal of Finance*, 68(5):1805–1841.

Menkhoff, L., Sarno, L., Schmeling, M., and Schrimpf, A. (2012a). Carry trades and global foreign exchange volatility. *The Journal of Finance*, 67(2):681–718.

Menkhoff, L., Sarno, L., Schmeling, M., and Schrimpf, A. (2012b). Currency momentum strategies. *Journal of Financial Economics*, 106(3):660–684.

Menkhoff, L., Sarno, L., Schmeling, M., and Schrimpf, A. (2016). Information flows in foreign exchange markets: Dissecting customer currency trades. *The Journal of Finance*, 71(2):601–634.

Menkhoff, L., Sarno, L., Schmeling, M., and Schrimpf, A. (2017). Currency value. *The Review of Financial Studies*, 30(2):416–441.

Menkhoff, L. and Schmeling, M. (2010). Trader see, trader do: How do (small) FX traders react to large counterparties' trades? *Journal of International Money and Finance*, 29(7):1283–1302.

Moore, M., Schrimpf, A., and Sushko, V. (2016). Downsized FX markets: causes and implications. *BIS Quarterly Review*.

Moore, M. J. and Payne, R. (2011). On the sources of private information in FX markets. *Journal of Banking & Finance*, 35(5):1250–1262.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703.

Newey, W. K. and West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61(4):631–653.

Osler, C. L., Mende, A., and Menkhoff, L. (2011). Price discovery in currency markets. *Journal of International Money and Finance*, 30(8):1696–1718.

Payne, R. (2003). Informed trade in spot foreign exchange markets: an empirical investigation. *Journal of International Economics*, 61(2):307–329.

Peiers, B. (1997). Informed traders, intervention, and price leadership: A deeper view of the microstructure of the foreign exchange market. *The Journal of Finance*, 52(4):1589–1614.

Phylaktis, K. and Chen, L. (2009). Asymmetric information, price discovery and macroeconomic announcements in FX market: do top trading banks know more? *International Journal of Finance & Economics*, 15:228–246.

Pojarliev, M. and Levich, R. M. (2012). A new look at currency investing, CFA institute. *Financial Markets and Portfolio Management*, 28(3):305–306.

Ranaldo, A. (2009). Segmentation and time-of-day patterns in foreign exchange markets. *Journal of Banking & Finance*, 33(12):2199–2206.

Ranaldo, A. and Santucci de Magistris, P. (2018). Trading volume, illiquidity, and commonalities in FX markets. *Working Paper*.

Rime, D., Sarno, L., and Sojli, E. (2010). Exchange rate forecasting, order flow and macroeconomic information. *Journal of International Economics*, 80(1):72–88.

Rime, D. and Schrimpf, A. (2013). The anatomy of the global FX market through the lens of the 2013 Triennial Survey. *BIS Quarterly Review*.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, 39(4):1127.

Stoll, H. R. (1978). The supply of dealer services in securities markets. *The Journal of Finance*, 33(4):1133–1151.

8 Tables

Table 1: Summary Statistics for Hourly Spot Returns

in BPS	AUDJPY	AUDNZD	AUDUSD	CADJPY	EURAUD	EURCAD
Mean(Δ_r)	0.04	−0.04	−0.08	0.05	0.08	0.06
Std(Δ_r)	16.37	10.02	13.39	15.13	13.29	12.03
Min(Δ_r)	−540.61	−120.73	−228.41	−407.53	−140.71	−146.40
Max(Δ_r)	170.91	162.48	137.07	157.72	184.65	169.51
Avg. Spread	4.39	4.81	3.53	4.53	3.88	3.78
AC(1) in %	1.14	−3.37	0.35	1.21	2.12	1.04
in BPS	EURCHF	EURDKK	EURGBP	EURJPY	EURNOK	EURSEK
Mean(Δ_r)	0.00	0.00	0.04	0.10	0.09	0.05
Std(Δ_r)	11.10	0.54	11.15	13.77	11.23	8.77
Min(Δ_r)	−1,355.15	−10.03	−122.46	−502.24	−349.16	−101.96
Max(Δ_r)	248.53	11.37	434.97	203.05	282.01	184.08
Avg. Spread	2.87	2.73	3.36	3.39	6.53	5.67
AC(1) in %	−3.35	−18.22	−0.75	1.34	−0.07	−1.84
in BPS	EURUSD	GBPAUD	GBPCAD	GBPCHF	GBPJPY	GBPUSD
Mean(Δ_r)	−0.02	0.04	0.03	−0.03	0.07	−0.05
Std(Δ_r)	11.03	13.51	12.25	14.89	15.85	11.34
Min(Δ_r)	−183.95	−369.35	−503.66	−1,362.38	−895.73	−588.25
Max(Δ_r)	147.86	183.43	190.47	249.81	327.34	153.96
Avg. Spread	2.40	4.44	4.14	4.31	4.03	2.69
AC(1) in %	1.75	1.00	−0.69	−3.35	2.37	2.55
in BPS	NZDUSD	USDCAD	USDCHF	USDDKK	USDHKD	USDILS
Mean(Δ_r)	−0.04	0.09	0.02	0.03	0.00	−0.04
Std(Δ_r)	14.77	10.23	14.22	11.03	0.69	10.14
Min(Δ_r)	−204.26	−142.93	−1,377.04	−145.23	−20.41	−178.48
Max(Δ_r)	174.39	187.09	250.23	182.45	14.89	187.19
Avg. Spread	4.41	2.83	3.41	2.93	1.69	23.70
AC(1) in %	−2.24	−0.11	−4.22	1.45	−9.93	−12.84
in BPS	USDJPY	USDMXP	USDNOK	USDSEK	USDSGD	USDZAR
Mean(Δ_r)	0.13	0.13	0.12	0.08	0.03	0.19
Std(Δ_r)	12.54	15.77	14.72	13.19	6.79	20.35
Min(Δ_r)	−318.89	−356.76	−379.52	−164.75	−113.95	−249.15
Max(Δ_r)	156.68	572.61	367.60	300.75	108.06	558.23
Avg. Spread	2.75	6.11	7.50	6.44	3.91	11.32
AC(1) in %	1.67	3.16	0.35	0.08	−1.50	0.40

Note: This table presents summary statistics for average hourly returns of all currency pairs in our sample. The first five rows report the sample mean (Mean(Δ_r)), standard deviation (Std(Δ_r)), minimum (Min(Δ_r)), and maximum (Max(Δ_r)) of the returns as well as the average relative spread over the full sample in BPS (Avg. Spread). The last row reports the first order autocorrelation (AC(1)) for hourly returns in percent (%).

Table 2: Summary Statistics for Hourly (Net) Volume

in USD mn	CO	FD	NB	BA	in USD mn	CO	FD	NB	BA
AUDJPY	0.01	0.86	1.18	14.65	GBPCHF	0.01	1.40	0.60	5.76
AUDNZD	0.00	0.72	1.41	13.35	GBPJPY	0.07	1.58	1.96	16.45
AUDUSD	0.71	24.31	9.95	98.50	GBPUSD	3.14	41.88	13.51	143.16
CADJPY	0.00	0.28	0.54	4.65	NZDUSD	0.03	6.99	3.74	38.23
EURAUD	0.06	2.31	1.82	16.96	USDCAD	1.14	23.35	11.88	190.62
EURCAD	0.91	1.78	1.53	12.22	USDCHF	0.79	9.67	11.14	71.00
EURCHF	0.45	7.74	4.53	35.95	USDDKK	0.90	3.01	0.12	7.76
EURDKK	0.13	4.44	0.66	18.99	USDHKD	0.03	10.11	1.38	41.09
EURGBP	2.32	17.26	4.44	48.21	USDILS	0.02	0.96	0.21	10.88
EURJPY	0.20	6.19	6.70	37.60	USDJPY	2.32	47.05	17.56	179.90
EURNOK	1.02	4.67	2.23	18.68	USDMXP	0.26	9.13	2.07	31.90
EURSEK	1.81	7.81	2.36	23.61	USDNOK	0.17	4.20	1.42	19.66
EURUSD	17.30	123.73	26.28	286.30	USDSEK	0.30	6.63	1.62	23.77
GBPAUD	0.02	1.20	0.90	8.01	USDSGD	0.15	5.64	1.27	37.30
GBPCAD	0.07	0.84	0.73	6.40	USDZAR	0.06	4.01	1.41	22.50

Note: This table reports net (absolute value of buy side minus sell side) volume broken down by four categories of agents: corporates (CO), funds (FD), non-bank financials (NB), and banks acting as price takers (BA). All numbers are in USD million.

Table 3: Location Parameters for Hourly (Net) Volume

in USD mn	Corporate	Fund	Non-Bank Financial	Bank
Mean	1.15	12.66	4.50	49.47
Std(Mean Vol)	0.05	0.27	0.09	0.50
Median	0.00	1.45	0.99	20.81
90^{th}	1.19	27.68	9.63	122.11
10^{th}	0.00	0.00	0.04	2.31
AC(1) in %	8.21	12.82	15.77	17.26

Note: This table reports the sample mean (*Mean*), standard deviation of the mean (*Std(Mean Vol)*), median (*Median*), 90^{th} percentile, and 10^{th} percentile of hourly (absolute) net volume for the cross-section of all currency pairs in USD million. The last row displays the first order autocorrelation for aggregate volume (*AC(1)*) in percent (%).

Table 4: Return Equation Coefficients

The model is:

$$r_t = \zeta_{1,l} D_{l,t} + \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{j \in C} \left(\sum_{i=0}^{5} \beta_i^j T_{t-i}^j + \sum_{i=0}^{5} \phi_i^j \tilde{S}_{t-1}^j \right) + \upsilon_1 \Delta s_{k,t;t-\tau} + \upsilon_2 \Delta s_{k,t;t-5\tau} + \epsilon_{r,t},$$

where $D_{l,t}$ denotes a dummy variable matrix to account for time-fixed effects. $\Delta s_{k,t;t-\tau}$ and $\Delta s_{k,t-\tau;t-5\tau}$ account for the return over the prior day and week. Here, $\tau = 24$ and t is measured at hourly frequency and $C = \{CO, FD, NB, BA\}$. Transactions are indexed by t. r_t refers to the log-return in the mid-quote. $\tilde{S}_t^j$ controls for order size and refers to the residuals of regressing signed log-volume against current and lagged values of the trade indicator variable.

Eq. (4.1)	α_1	β_0^{CO}	β_0^{FD}	β_0^{NB}	β_0^{BA}	$\bar{R}^2$ in %	Eq. (4.1)	α_1	β_0^{CO}	β_0^{FD}	β_0^{NB}	β_0^{BA}	$\bar{R}^2$ in %
AUDJPY	***−5.430	0.019	*0.008	***0.010	***0.012	6.97	GBPCHF	***−9.060	*−0.065	−0.002	***0.015	***−0.005	6.79
	[3.800]	[0.280]	[1.781]	[4.986]	[12.513]			[3.032]	[1.876]	[0.927]	[3.291]	[5.341]	
AUDNZD	***−9.212	−0.016	−0.003	***−0.005	**−0.001	6.33	GBPJPY	−3.932	−0.014	0.000	**0.004	***0.008	7.16
	[12.982]	[0.538]	[1.048]	[4.999]	[2.235]			[1.506]	[0.910]	[0.043]	[2.257]	[6.806]	
AUDUSD	***−5.551	−0.012	***0.004	***0.012	0.001	7.02	GBPUSD	*−3.559	***−0.016	***0.003	***0.008	***0.004	7.31
	[7.147]	[1.640]	[4.320]	[15.489]	[1.148]			[1.817]	[4.243]	[3.856]	[11.042]	[6.309]	
CADJPY	***−4.577	0.075	−0.001	***0.009	***0.004	6.14	NZDUSD	***−7.462	*−0.044	***0.010	***0.007	***0.007	6.36
	[3.020]	[1.275]	[0.138]	[2.884]	[4.730]			[9.855]	[1.949]	[6.845]	[6.973]	[8.180]	
EURAUD	***−3.694	−0.013	0.002	0.001	***0.005	6.10	USDCAD	***−6.330	***−0.030	***0.004	***0.005	***0.003	6.88
	[3.109]	[0.775]	[1.599]	[0.878]	[6.807]			[6.820]	[5.332]	[4.284]	[7.526]	[4.976]	
EURCAD	***−5.093	***−0.036	0.002	***0.007	***−0.002	6.55	USDCHF	***−10.349	**−0.014	0.002	***0.013	0.001	7.30
	[3.911]	[5.944]	[0.955]	[4.691]	[2.784]			[3.043]	[2.546]	[1.470]	[13.146]	[1.599]	
EURCHF	**−8.790	−0.002	0.001	−0.002	***−0.006	6.64	USDDKK	***−4.316	***−0.045	−0.002	***0.017	***−0.003	6.26
	[2.365]	[0.498]	[0.920]	[1.122]	[7.575]			[4.064]	[5.375]	[1.325]	[3.423]	[3.784]	
EURDKK	***−24.854	0.000	***0.000	0.000	***0.000	11.25	USDHKD	***−17.325	0.001	***0.000	0.000	0.000	9.46
	[14.082]	[0.072]	[3.109]	[0.167]	[3.533]			[8.876]	[1.120]	[5.002]	[0.245]	[1.557]	
EURGBP	***−6.652	***−0.017	***0.003	0.001	***−0.004	6.24	USDILS	***−20.330	0.017	***0.004	***−0.011	***0.003	10.18
	[6.718]	[4.819]	[2.758]	[1.132]	[6.135]			[21.091]	[1.308]	[2.609]	[5.556]	[3.256]	
EURJPY	***−4.781	***−0.029	−0.002	***0.003	***−0.004	6.62	USDJPY	***−4.864	**−0.009	***0.006	***0.010	***0.003	7.39
	[2.906]	[2.780]	[1.401]	[3.160]	[4.482]			[3.920]	[2.379]	[6.434]	[13.681]	[3.966]	
EURNOK	***−5.893	***−0.019	***0.007	0.002	***0.002	6.54	USDMXP	−2.583	*−0.021	0.001	***−0.016	0.001	6.40
	[5.519]	[3.963]	[4.140]	[1.063]	[3.288]			[0.639]	[1.753]	[0.790]	[9.558]	[0.672]	
EURSEK	***−7.194	***−0.012	***0.005	***0.003	***0.002	6.49	USDNOK	***−5.992	***−0.039	***0.004	***0.007	***0.004	6.66
	[7.685]	[4.317]	[4.810]	[2.621]	[3.949]			[5.641]	[2.618]	[2.784]	[4.017]	[4.277]	
EURUSD	***−4.381	***−0.018	0.000	***0.007	**−0.001	7.24	USDSEK	***−5.702	−0.017	***0.005	***0.006	***0.004	6.25
	[3.989]	[9.794]	[0.364]	[10.237]	[2.041]			[5.776]	[1.500]	[3.701]	[3.997]	[4.365]	
GBPAUD	***−4.919	**0.040	***0.005	0.001	***0.003	6.25	USDSGD	***−8.103	***−0.013	***0.002	***0.003	***−0.001	6.90
	[4.545]	[2.398]	[2.728]	[0.708]	[3.541]			[10.225]	[3.080]	[4.078]	[3.628]	[2.705]	
GBPCAD	***−6.550	−0.036	−0.001	0.003	0.001	6.19	USDZAR	***−5.993	−0.025	**0.006	0.003	***0.004	6.82
	[5.409]	[0.910]	[0.504]	[1.351]	[1.086]			[5.310]	[1.128]	[2.538]	[1.226]	[2.984]	
Expected sign	-	+	+	+	+		Expected sign	-	+	+	+	+	
$D_{l,t}$						Yes	$D_{l,t}$						Yes
Lagged Ret.						Yes	Lagged Ret.						Yes
$\tilde{S}_t^j$						Yes	$\tilde{S}_t^j$						Yes

Note: The regression coefficients are estimated by OLS on the full sample. All coefficients are in %. T-stats in square brackets are based on HAC errors and stars (*/ **/ ***) denote significance at the 90%/ 95%/ 99% level, respectively.

Table 5: Order Flow Equation Coefficients

The model is:

$$T_t = \zeta_{2,l} D_{l,t} + \sum_{i=1}^{5} \gamma_i r_{t-i} + \sum_{j \in C} \left(\sum_{i=1}^{5} \delta_i^j T_{t-i}^j + \sum_{i=1}^{5} \omega_i^j \tilde{S}_{t-1}^j \right) + \epsilon_{T,t},$$

where $D_{l,t}$ denotes a dummy variable matrix to account for time-fixed effects, and $C = \{CO, FD, NB, BA\}$. Transactions are indexed by t. r_t refers to the log-return in the mid-quote. $\tilde{S}_t^j$ controls for order size and refers to the residuals of regressing signed log-volume against current and lagged values of the trade indicator variable.

Eq. (4.2)	γ_1	δ_1^{CO}	δ_1^{FD}	δ_1^{NB}	δ_1^{BA}	$\bar{R}^2$ in %	Eq. (4.2)	γ_1	δ_1^{CO}	δ_1^{FD}	δ_1^{NB}	δ_1^{BA}	$\bar{R}^2$ in %
AUDJPY	***32.245	−0.083	*0.034	0.007	***0.065	1.59	GBPCHF	***−24.533	***0.526	0.003	0.009	***0.020	0.31
	[7.708]	[0.344]	[1.657]	[0.699]	[11.206]			[3.734]	[3.835]	[0.212]	[0.604]	[3.413]	
AUDNZD	***−15.812	0.325	0.017	0.001	***0.051	0.50	GBPJPY	***29.055	0.011	**0.033	0.013	***0.045	0.82
	[2.756]	[1.542]	[0.743]	[0.070]	[8.975]			[4.648]	[0.117]	[2.358]	[1.548]	[7.794]	
AUDUSD	**−8.621	−0.005	0.010	***0.021	***0.034	0.40	GBPUSD	**−10.553	0.000	0.000	0.006	***0.043	0.57
	[2.017]	[0.132]	[1.449]	[3.621]	[5.912]			[2.210]	[0.018]	[0.016]	[0.978]	[7.556]	
CADJPY	0.926	0.395	0.032	*0.025	***0.026	0.17	NZDUSD	***−17.782	−0.006	0.011	0.005	***0.054	0.73
	[0.246]	[1.015]	[1.064]	[1.675]	[4.508]			[4.613]	[0.077]	[1.184]	[0.721]	[9.532]	
EURAUD	***−11.429	0.099	−0.001	0.003	***0.017	0.13	USDCAD	7.992	0.018	**0.016	0.006	***0.062	1.57
	[2.650]	[1.106]	[0.139]	[0.345]	[3.028]			[1.462]	[0.569]	[2.047]	[0.967]	[10.904]	
EURCAD	***−22.430	***0.122	−0.009	**0.021	***0.037	0.50	USDCHF	***−11.005	***0.135	***0.028	0.005	***0.034	0.41
	[4.777]	[3.353]	[0.806]	[2.163]	[6.383]			[3.159]	[3.623]	[3.424]	[0.786]	[5.848]	
EURCHF	−29.460	*0.063	***0.032	0.008	***0.066	1.65	USDDKK	−6.685	0.028	**0.020	0.018	***0.023	0.73
	[1.471]	[1.707]	[3.499]	[0.970]	[11.120]			[1.421]	[0.650]	[1.998]	[0.521]	[3.382]	
EURDKK	***374.927	*−0.138	**0.026	**0.087	***0.062	1.09	USDHKD	***−343.281	**0.284	0.011	0.024	***0.049	0.53
	[3.785]	[1.828]	[2.237]	[2.548]	[9.462]			[4.089]	[1.964]	[1.534]	[1.266]	[8.335]	
EURGBP	***−33.082	0.030	*0.015	−0.007	***0.041	0.88	USDILS	−3.552	0.185	0.016	−0.003	***0.076	1.34
	[6.397]	[1.377]	[1.862]	[0.973]	[7.114]			[0.715]	[1.150]	[1.211]	[0.140]	[11.322]	
EURJPY	−1.382	−0.031	**0.021	***0.017	***0.038	0.94	USDJPY	−6.166	0.017	***0.030	**0.014	***0.027	0.47
	[0.341]	[0.575]	[2.049]	[2.736]	[6.658]			[1.381]	[0.743]	[4.451]	[2.492]	[4.735]	
EURNOK	***−33.475	***0.059	***0.026	***0.042	***0.074	1.25	USDMXP	***−20.419	0.062	0.009	0.011	***0.046	0.41
	[5.964]	[2.661]	[2.579]	[4.495]	[12.543]			[4.420]	[1.375]	[1.064]	[1.304]	[7.804]	
EURSEK	***−35.644	**0.042	***0.036	***0.027	***0.073	1.10	USDNOK	*6.357	0.079	0.016	0.006	***0.061	0.76
	[5.655]	[2.328]	[3.989]	[2.868]	[12.539]			[1.705]	[1.263]	[1.645]	[0.648]	[10.207]	
EURUSD	***−21.832	**0.024	***0.028	0.000	***0.053	1.69	USDSEK	***−11.285	0.034	***0.025	0.001	***0.049	0.62
	[4.305]	[2.074]	[4.255]	[0.001]	[9.188]			[2.695]	[0.681]	[2.814]	[0.103]	[8.338]	
GBPAUD	−6.422	−0.160	0.015	**0.024	***0.020	0.12	USDSGD	***−75.327	−0.059	0.008	−0.006	***0.041	0.81
	[1.539]	[0.701]	[1.170]	[2.239]	[3.581]			[8.874]	[0.947]	[0.949]	[0.563]	[7.134]	
GBPCAD	**9.907	−0.081	0.008	***0.044	***0.026	0.23	USDZAR	***−22.957	0.013	0.007	*−0.016	***0.043	0.71
	[2.042]	[0.329]	[0.617]	[3.392]	[4.461]			[7.479]	[0.264]	[0.692]	[1.779]	[7.510]	
Expected sign	-	+	+	+	+		Expected sign	-	+	+	+	+	
$D_{l,t}$						Yes	$D_{l,t}$						Yes
Lagged Ret.						Yes	Lagged Ret.						Yes
$\tilde{S}_t^j$						Yes	$\tilde{S}_t^j$						Yes

Note: The linear regression coefficients are estimated by OLS on the full sample. T-stats in square brackets are based on HAC errors and stars (*/ **/ ***) denote significance at the 90%/ 95%/ 99% level, respectively.

Table 6: Permanent Price Impact Across Agents - Joint F-test

in BPS	α_m^{CO}	α_m^{FD}	α_m^{NB}	α_m^{BA}	in BPS	α_m^{CO}	α_m^{FD}	α_m^{NB}	α_m^{BA}
AUDJPY	2.636	−0.115	***1.073	***1.446	GBPCHF	−4.332	0.286	***1.955	***−0.107
	[0.251]	[2.174]	[7.765]	[32.220]		[1.499]	[1.202]	[8.635]	[6.218]
AUDNZD	0.952	1.049	***−0.597	*0.229	GBPJPY	−4.710	0.371	*−0.224	***1.096
	[0.460]	[2.246]	[6.285]	[2.668]		[0.482]	[0.790]	[2.753]	[13.858]
AUDUSD	0.265	***0.383	***1.036	0.405	GBPUSD	***−1.337	***0.533	***0.842	***0.909
	[1.829]	[3.949]	[43.294]	[1.480]		[8.330]	[3.944]	[24.585]	[10.873]
CADJPY	6.876	1.460	***0.593	***0.014	NZDUSD	***−5.160	***1.131	***0.915	***0.967
	[1.508]	[0.908]	[3.940]	[5.681]		[3.723]	[9.902]	[10.185]	[12.468]
EURAUD	−0.782	0.540	−0.034	***0.766	USDCAD	***−3.379	***0.560	***0.411	***0.316
	[0.289]	[1.340]	[1.093]	[8.882]		[16.296]	[7.564]	[11.882]	[4.688]
EURCAD	***−0.739	0.657	***0.398	**0.103	USDCHF	−1.353	0.541	***1.136	0.435
	[13.597]	[1.015]	[7.441]	[3.228]		[2.124]	[1.267]	[30.532]	[2.444]
EURCHF	−0.179	0.032	−0.229	***−0.245	USDDKK	***−4.548	*0.240	***1.549	***−0.186
	[0.427]	[0.486]	[0.687]	[13.336]		[14.978]	[2.742]	[3.834]	[4.544]
EURDKK	0.206	**0.046	0.032	***−0.003	USDHKD	0.153	***0.037	−0.005	0.010
	[1.106]	[2.898]	[0.960]	[4.251]		[0.343]	[5.478]	[0.350]	[2.104]
EURGBP	***−0.865	0.471	0.136	***0.221	USDILS	1.439	**1.177	***−0.869	***0.630
	[9.760]	[2.488]	[1.868]	[11.490]		[0.529]	[3.327]	[7.519]	[4.433]
EURJPY	***2.446	−1.153	**0.407	***0.227	USDJPY	**−0.896	***0.513	***0.443	***0.663
	[6.276]	[2.425]	[3.353]	[5.525]		[3.349]	[8.878]	[35.970]	[4.392]
EURNOK	***−1.993	***0.973	0.060	***0.495	USDMXP	***1.601	**−0.593	***−1.668	**0.663
	[11.413]	[7.980]	[2.098]	[3.658]		[4.798]	[3.056]	[27.065]	[3.304]
EURSEK	***−0.781	***0.875	***0.539	***0.434	USDNOK	***1.014	**0.657	***1.393	***0.068
	[11.287]	[9.837]	[4.627]	[4.168]		[4.203]	[3.354]	[7.387]	[5.625]
EURUSD	***−1.548	0.200	***0.476	***0.407	USDSEK	**−3.091	***1.608	***1.356	***0.278
	[34.875]	[1.129]	[20.045]	[4.211]		[2.986]	[7.565]	[7.253]	[4.265]
GBPAUD	10.471	0.440	0.636	***0.704	USDSGD	−0.057	***0.112	***0.605	*−0.007
	[1.619]	[1.561]	[2.386]	[3.901]		[2.117]	[3.856]	[4.714]	[2.705]
GBPCAD	−4.120	0.281	1.019	0.292	USDZAR	**−6.317	0.755	**0.502	***2.177
	[0.377]	[1.275]	[1.857]	[1.452]		[3.223]	[1.856]	[2.810]	[12.339]

Note: The numbers in brackets correspond to the test statistic for a joint F-test that the parameters in Eq. (4.5) are jointly different from zero. Stars (*/ **/ ***) denote significance at the global 90%/ 95%/ 99% level (α_g), respectively. For each individual test a Bonferroni correction is applied such that the local significance level is $\frac{\alpha_g}{m}$, where m is the number of multiple tests in the joint hypothesis. All regression coefficients are in BPS.

Table 7: Performance Benchmarking - *Gross* Returns - ALP_{HML}

Panel a)		Q_1	Q_2	Q_3	Q_4	Q_5	ALP_{HML}
SR	USD	−1.06	−0.62	−0.24	0.66	0.29	1.22
	EUR	−0.82	−0.34	−0.17	0.29	0.64	1.22
Mean in %	USD	**−7.51	−4.94	−2.19	5.24	2.13	***9.80
		[2.11]	[1.28]	[0.60]	[1.29]	[0.61]	[2.81]
	EUR	−4.22	−1.73	−0.90	2.24	5.38	***9.82
		[1.40]	[0.73]	[0.28]	[0.80]	[1.51]	[2.81]
MDD in %	USD	5.48	10.20	22.07	33.71	12.79	3.50
	EUR	12.02	9.52	18.38	14.65	5.06	3.50
Θ in %	USD	5.43	3.33	−3.37	−7.01	0.40	8.95
	EUR	2.59	0.72	0.06	−3.53	3.63	8.96

Panel b)		DOL	RER	MOM_{HML}	CAR_{HML}	BMS	ALP_{HML}
SR	USD	−0.43	0.73	−0.10	0.48	0.33	1.22
	EUR	−0.05	0.68	−0.14	0.43	0.25	1.22
Mean in %	USD	−2.79	*1.69	−0.77	3.12	1.08	***9.80
		[0.84]	[1.76]	[0.24]	[1.21]	[0.70]	[2.81]
	EUR	−0.23	*1.60	−1.00	2.82	0.82	***9.82
		[0.11]	[1.69]	[0.33]	[1.11]	[0.55]	[2.81]
MDD in %	USD	23.49	1.81	11.24	10.47	4.05	3.50
	EUR	5.79	1.79	11.83	10.61	4.11	3.50
Θ in %	USD	−4.06	0.31	−3.53	0.32	−1.33	8.95
	EUR	−1.03	0.50	−3.38	0.57	−1.07	8.96

Note: This table presents the out-of-sample economic performance of the ALP_{HML} strategy *before* transaction costs. Panel a) reports the annualised Sharpe ratio (SR), the annualised average (simple) *gross* excess return (*Mean*), the maximum drawdown (MDD) and the Θ performance measure of Goetzmann et al. (2007) for the quintile portfolios ($Q_1, Q_2, ..., Q_5$). Panel b) lists the same measures as Panel a) but for common FX trading strategies based on monthly rebalancing. In particular, DOL is based on USD (or EUR) currency pairs, RER on the real exchange rate (cf. Menkhoff et al. (2017)), MOM_{HML} on $f^m_{t-1,t} - s^m_t$ (cf. Asness et al. (2013)), CAR_{HML} on the forward discount/ premium ($f^m_{t,t+1} - s^m_t$, cf. Lustig et al. (2011)), and BMS is based on lagged standardised order flow (cf. Menkhoff et al. (2016)). Significance at the 90%/ 95%/ 99% level are represented by stars (*/ **/ ***), respectively. The numbers inside the brackets are the corresponding test statistics based on HAC errors correcting for serial correlation and small sample size (using the plug-in procedure for automatic lag selection by Andrews and Monahan (1992), Newey and West (1994)).

Table 8: Performance Benchmarking - *Net* Returns - ALP_{HML}

Panel a)		Q_1	Q_2	Q_3	Q_4	Q_5	ALP_{HML}
SR	USD	−0.84	−0.49	−0.35	0.84	0.13	0.86
	EUR	−0.55	−0.19	−0.07	0.44	0.50	0.89
Mean in %	USD	*−5.91	−3.76	−3.06	6.58	0.74	*6.69
		[1.68]	[1.00]	[0.89]	[1.63]	[0.27]	[1.91]
	EUR	−2.79	−0.87	−0.25	3.28	4.10	**7.00
		[0.93]	[0.40]	[0.11]	[1.25]	[1.18]	[2.00]
MDD in %	USD	6.42	11.01	24.77	40.35	14.53	4.22
	EUR	14.58	10.88	20.12	17.34	6.36	4.12
Θ in %	USD	5.43	3.33	−3.37	−7.01	0.40	6.10
	EUR	2.59	0.72	0.06	−3.53	3.63	6.39

Panel b)		DOL	RER_{HML}	RER	MOM_{HML}	CAR_{HML}	ALP_{HML}
SR	USD	−0.61	0.21	0.16	−0.54	0.11	0.86
	EUR	−0.27	0.31	0.24	−0.51	0.15	0.89
Mean in %	USD	−3.85	0.80	0.34	−3.35	0.56	*6.69
		[1.19]	[0.56]	[0.39]	[1.29]	[0.29]	[1.91]
	EUR	−0.98	1.21	0.53	−3.20	0.82	**7.00
		[0.60]	[0.80]	[0.59]	[1.21]	[0.38]	[2.00]
MDD in %	USD	26.54	3.48	2.77	17.03	12.87	4.22
	EUR	7.78	3.37	2.58	16.74	12.41	4.12
Θ in %	USD	−4.06	0.70	0.31	−3.53	0.32	6.10
	EUR	−1.03	1.10	0.50	−3.38	0.57	6.39

Note: This table presents the out-of-sample economic performance of the ALP_{HML} strategy *after* transaction costs. Panel a) reports the annualised Sharpe ratio (SR), the annualised average (simple) *net* excess return (*Mean*), the maximum drawdown (MDD) and the Θ performance measure of Goetzmann et al. (2007) for the quintile portfolios ($Q_1, Q_2, ..., Q_5$). Panel b) lists the same measures as Panel a) but for common FX trading strategies based on monthly rebalancing. In particular, DOL is based on USD (or EUR) currency pairs, RER/ RER_{HML} on the real exchange rate (cf. Menkhoff et al. (2017)), MOM_{HML} on $f^m_{t-1,t} - s^m_t$ (cf. Asness et al. (2013)), CAR_{HML} on the forward discount/ premium ($f^m_{t,t+1} - s^m_t$, cf. Lustig et al. (2011)). Significance at the 90%/ 95%/ 99% level are represented by stars (*/ **/ ***), respectively. The numbers inside the brackets are the corresponding test statistics based on HAC errors correcting for serial correlation and small sample size (using the plug-in procedure for automatic lag selection by Andrews and Monahan (1992), Newey and West (1994)).

Table 9: Exposure Regression Based on Monthly *Gross* USD-Excess Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
α	***0.008	***0.008	**0.007	**0.006	**0.008	***0.008	***0.008	**0.008	**0.008	**0.007
	[2.832]	[2.718]	[2.574]	[2.265]	[2.419]	[2.637]	[2.717]	[2.571]	[2.513]	[2.465]
DOL		−0.138	−0.009	−0.050	−0.100	−0.070	−0.153	−0.124	−0.136	0.240
		[0.644]	[0.033]	[0.257]	[0.464]	[0.368]	[0.644]	[0.572]	[0.609]	[1.137]
RER_{HML}			0.459							
			[1.134]							
RER				0.983						**1.835
				[1.543]						[2.189]
MOM_{HML}					0.308					
					[1.296]					
CAR_{HML}						−0.167				**−0.527
						[0.978]				[2.201]
BMS							−0.120			
							[0.350]			
ΔVIX								0.000		
								[0.976]		
ΔCDS									0.000	0.000
									[0.436]	[0.516]
R^2 in %	N/A	1.12	5.92	9.02	6.47	2.99	1.38	2.03	1.35	22.01
IR	0.35	0.34	0.30	0.30	0.36	0.37	0.34	0.33	0.34	0.37
#Obs	51	51	51	51	51	51	51	51	51	51

Note: In this table, we regress monthly *gross* excess returns by ALP_{HML} on monthly excess returns associated with common risk factors, where DOL is based on USD (or EUR) currency pairs; RER/ RER_{HML} are based on the real exchange rate (cf. Menkhoff et al. (2017)); MOM_{HML} is based on $f^m_{t-1,t} - s^m_t$ (cf. Asness et al. (2013)), CAR_{HML} is based on the forward discount/ premium ($f^m_{t,t+1} - s^m_t$, cf. Lustig et al. (2011)), and BMS is based on lagged standardised order flow (cf. Menkhoff et al. (2016)). ΔVIX is the return on the VIX index and ΔCDS the change in the iTraxx Europe CDS index. The information ratio (IR) is defined as α divided by residual standard deviation. Significance at the 90%/ 95%/ 99% level are represented by stars (*/ **/ ***), respectively. The numbers inside the brackets are the corresponding test statistics based on HAC errors correcting for serial correlation and small sample size (using the plug-in procedure for automatic lag selection by Andrews and Monahan (1992), Newey and West (1994)).

Table 10: Correlation with Common FX Risk Factors in %

	ΔVIX	ΔCDS	DOL	RER_{HML}	RER	MOM_{HML}	CAR_{HML}	BMS
ΔCDS	25.41							
DOL	7.58	14.11						
RER_{HML}	4.03	−3.88	−41.43					
RER	1.23	1.61	−23.57	88.02				
MOM_{HML}	17.18	14.68	−12.05	2.60	3.46			
CAR_{HML}	21.89	13.94	35.48	31.13	43.28	−15.79		
BMS	−4.21	−14.62	−21.57	−8.70	−2.21	−4.46	−4.39	
ALP_{HML}	1.87	−10.12	−10.22	24.51	29.77	23.66	−15.62	−1.40

Note: This table shows the time series cross-correlation at lag 0 between the *gross* excess return of HML_α (US perspective) and those associated with different FX risk factors, where DOL is based on USD (or EUR) currency pairs; RER/ RER_{HML} are based on the real exchange rate (cf. Menkhoff et al. (2017)); MOM_{HML} is based on $f^m_{t-1,t} - s^m_t$ (cf. Asness et al. (2013)), CAR_{HML} is based on the forward discount/ premium ($f^m_{t,t+1} - s^m_t$, cf. Lustig et al. (2011)), and BMS is based on lagged standardised order flow (cf. Menkhoff et al. (2016)). ΔVIX is the return on the VIX index and ΔCDS the change in the iTraxx Europe CDS index.

Table 11: Summary of Annual Trading Costs and Cost Per Trade

Panel a)	USD			EUR		
in %	daily	weekly	monthly	daily	weekly	monthly
Mean	44.89	16.00	2.85	40.22	14.52	2.57
Median	42.12	15.73	2.64	36.66	14.20	2.34
Std	17.95	4.65	1.11	17.47	5.26	1.10
Panel b)	**USD**			**EUR**		
in %	daily	weekly	monthly	daily	weekly	monthly
Mean	0.18	0.32	0.24	0.16	0.29	0.21
Median	0.17	0.31	0.22	0.15	0.28	0.20
Std	0.07	0.09	0.09	0.07	0.11	0.09

Note: Panel a) shows the average, median and standard deviation of annualised trading costs in percent for different rebalancing frequencies (daily, weekly, and monthly). Annualised transaction costs are approximated by the cost per trade times the number of trading days, weeks, and months per year. Panel b) lists the same measures as Panel a) but for the cost per trade associated with daily, weekly, and monthly rebalancing. Across both panels USD refers to a US-investor perspective, while EUR to a European (EU-centric) investor. Transaction costs are calculated as the difference in log returns between excess returns based on mid-quote and bid and ask quotes.

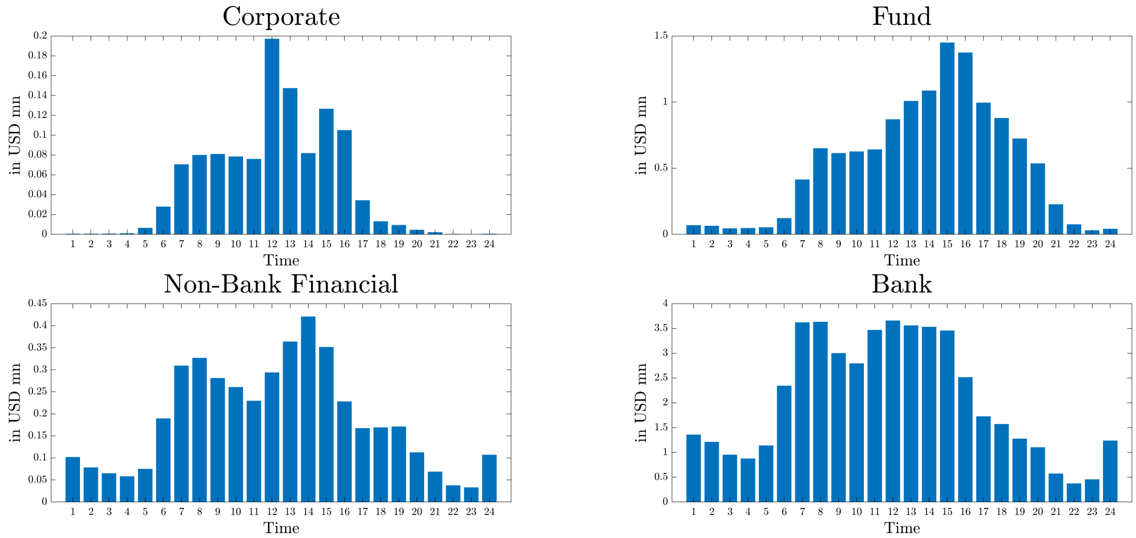
Table 12: Subsample Performance Benchmarking - *Gross* Returns

Panel a) US perspective	USD	EUR/ GBP	no EM	G10	no CHF	ALP_{HML}
SR	0.23	1.08	0.94	0.80	1.21	1.22
Mean in %	1.02	**6.07	**7.63	4.60	***9.06	***9.80
	[0.46]	[2.47]	[2.07]	[1.60]	[2.59]	[2.76]
MDD in %	7.92	3.47	4.97	4.49	3.96	3.50
MPPM in %	0.89	5.68	6.96	4.28	8.32	8.95
Panel b) EU perspective	**USD**	**EUR/ GBP**	**no EM**	**G10**	**no CHF**	ALP_{HML}
SR	0.95	0.85	0.85	1.20	1.01	1.22
Mean in %	**3.81	*4.81	*6.76	***5.79	**7.33	***9.82
	[1.97]	[1.80]	[1.96]	[2.74]	[2.28]	[2.77]
MDD in %	3.01	3.92	6.18	2.14	3.93	3.50
MPPM in %	3.62	4.49	6.16	5.45	6.74	8.96

Note: In this table, we report *gross* performance measures of ALP_{HML} based on five subsamples of currency pairs: (i) USD currency pairs only (15 in total), (ii) EUR and GBP currency pairs only (14 in total), (iii) all pairs excluding emerging market currencies (i.e. USDILS, USDMXP, and USDZAR) and/ or fixed pairs (i.e. EURDKK, USDDKK, USDHKD, and USDSGD), (iv) G10 currency pairs plus the most liquid EUR cross pairs (14 in total), (v) all pairs excluding the CHF crosses (27 in total). Significance at the 90%/ 95%/ 99% level are represented by stars (*/ **/ ***), respectively. The numbers inside the brackets are the corresponding test statistics based on HAC errors correcting for serial correlation and small sample size (using the plug-in procedure for automatic lag selection by Andrews and Monahan (1992), Newey and West (1994)).

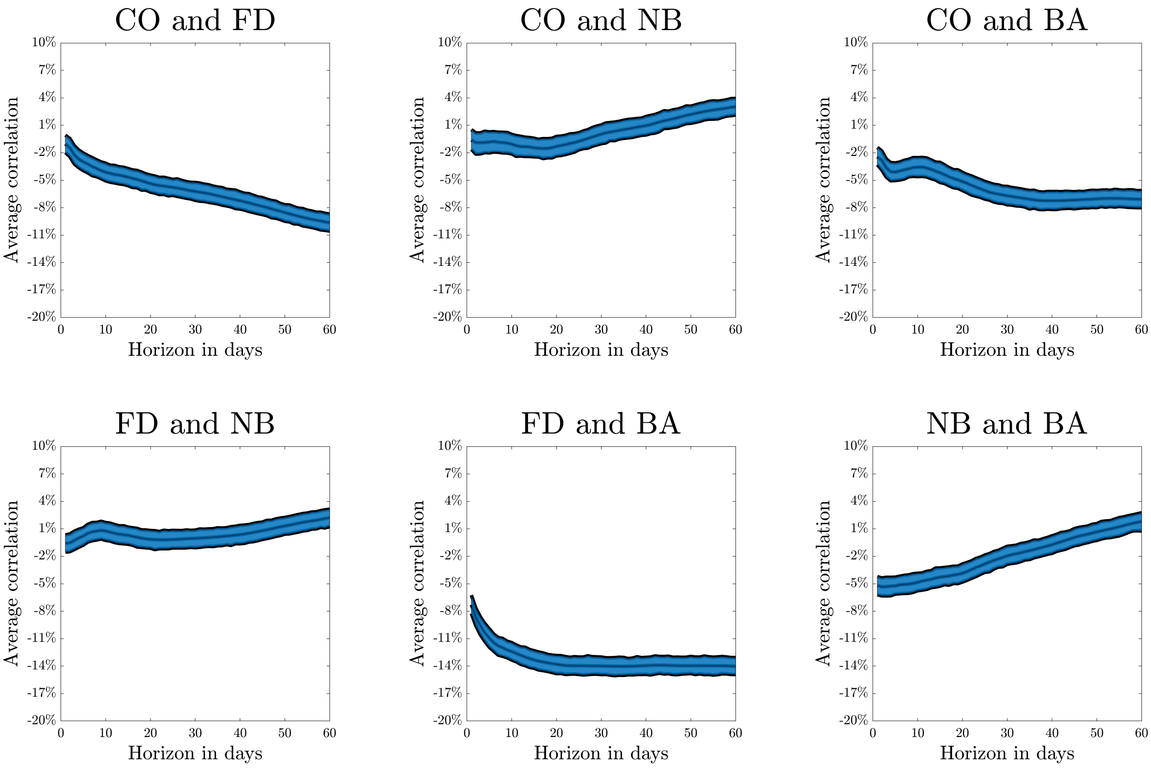
9 Figures

Figure 1: Distribution of Trading Volume Over a Day



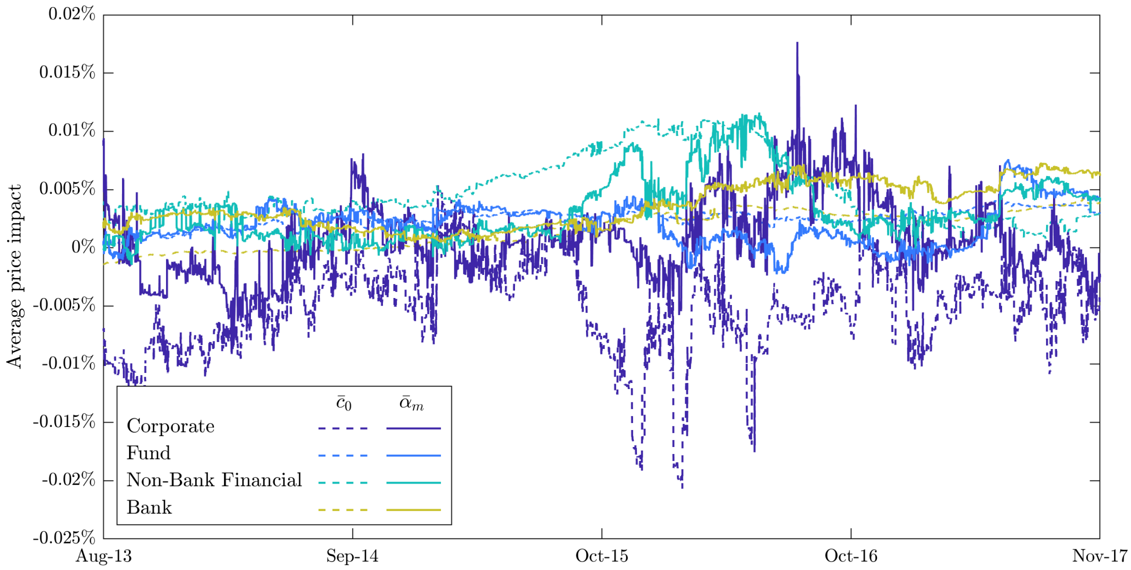
Note: This figure shows the average hourly volume (in USD million) during the entire trading day. The average is computed across all trading days and currency pairs using the entire sample period from Sep 2012 to Nov 2017. The numbers on the horizontal axis denote the closing time, e.g. the bar denoted 17 refers to volume between *4pm* and *5pm* (GMT, no BST adjustment).

Figure 2: Correlation of Customer Order Flows Over Longer Horizons



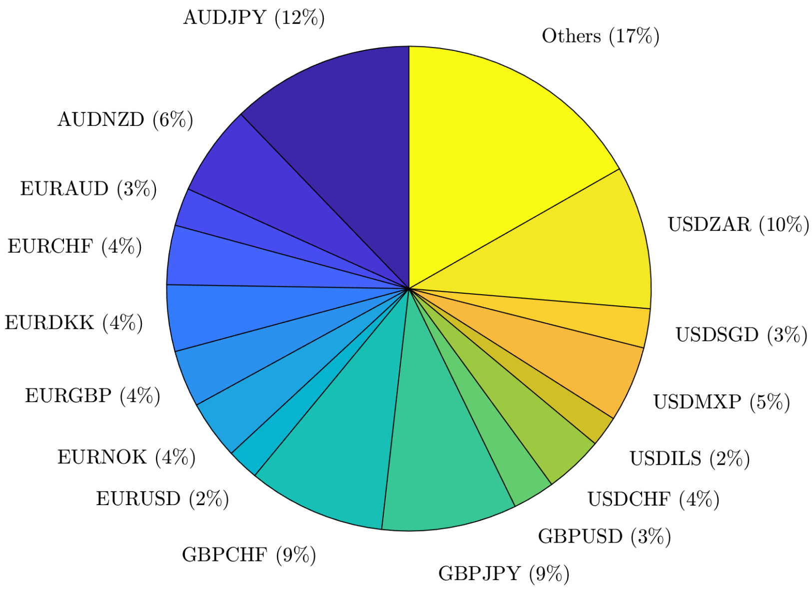
Note: Correlations are based on the average correlation across all currency pairs. A one day horizon corresponds to non-overlapping hourly observations. For horizons greater than one day we sum up order flow over n days in an overlapping fashion and calculate correlations based on the sum of n day order flow. The shaded areas correspond to 95% confidence bands based on a moving-block bootstrap with 1,000 repetitions.

Figure 3: Twelve Months Rolling Window Regression for $c_0^{r,k}$/ α_m^k



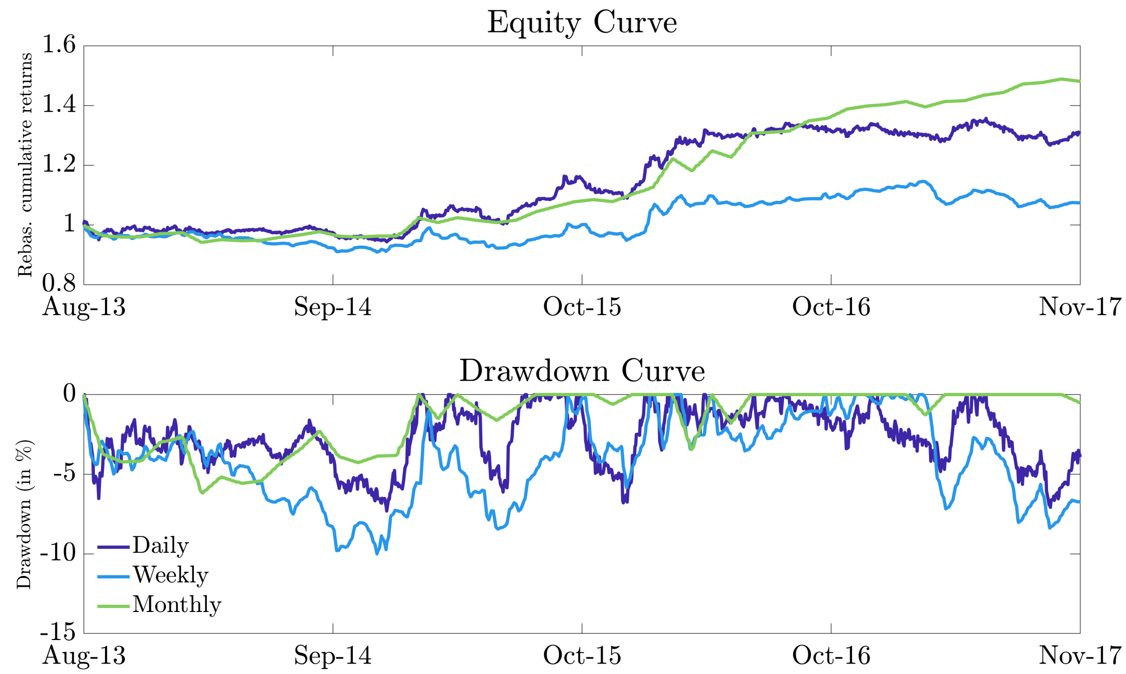
Note: The cross-sectional average contemporary ($\bar{c}_0^r$) and permanent ($\bar{\alpha}_m$) price impact are calculated after removing any coefficients that are either heavy outliers with respect to the median or not significant at a 95% confidence level applying a simple two-sided t-test and joint F-test, respectively.

Figure 4: Distribution of Absolute Currency Exposure (US-Investor)

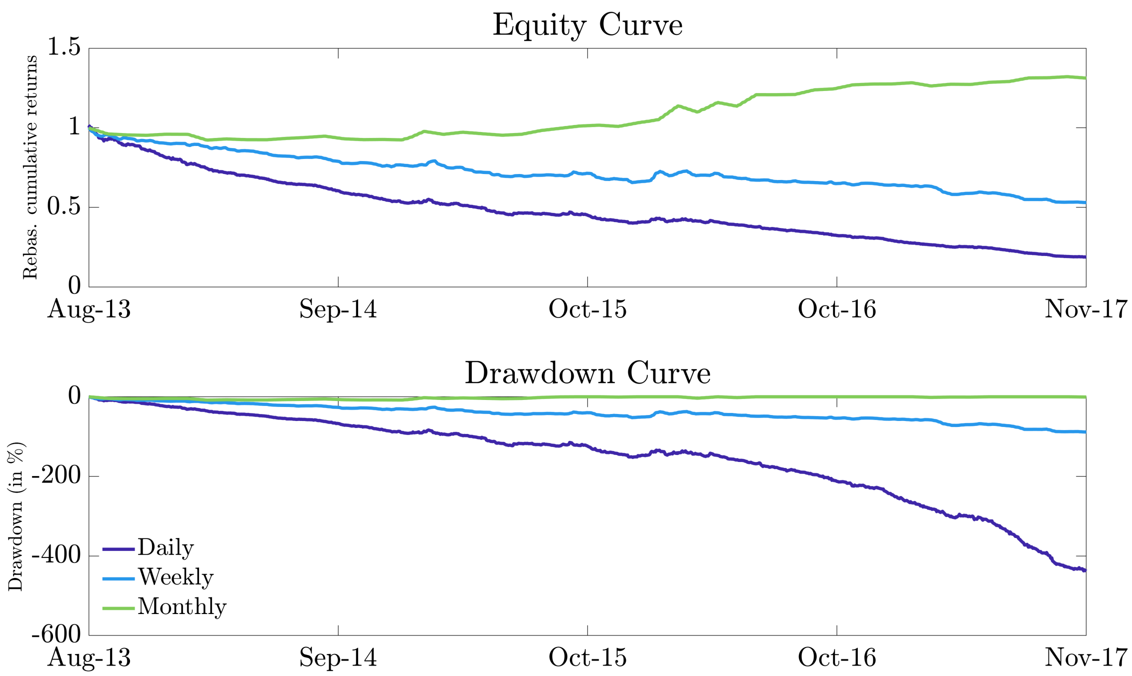


Note: Sum up absolute exposure to each currency pair over time and then normalise to one. 'Others' comprise currency pairs with a relative share $\leqslant 2\%$: AUDUSD, CADJPY, EURCAD, EURJPY, EURSEK, GBPAUD, GBPCAD, NZDUSD, USDCAD, USDDKK, USDHKD, USDJPY, USDNOK, and USDSEK.

Figure 5: Equity and Drawdown Curves (US Investor)



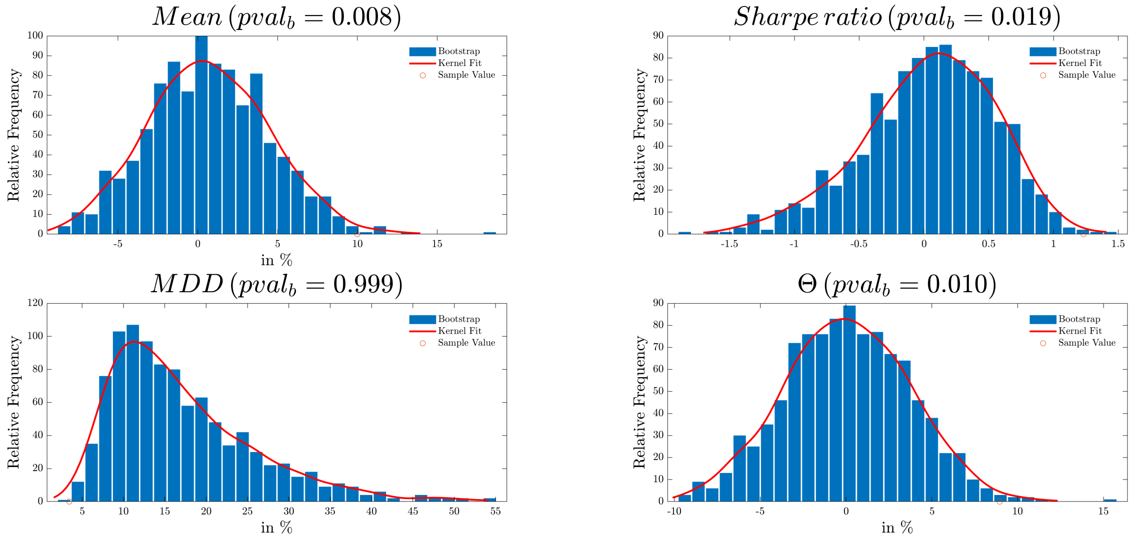
(a) Prior Transaction Costs



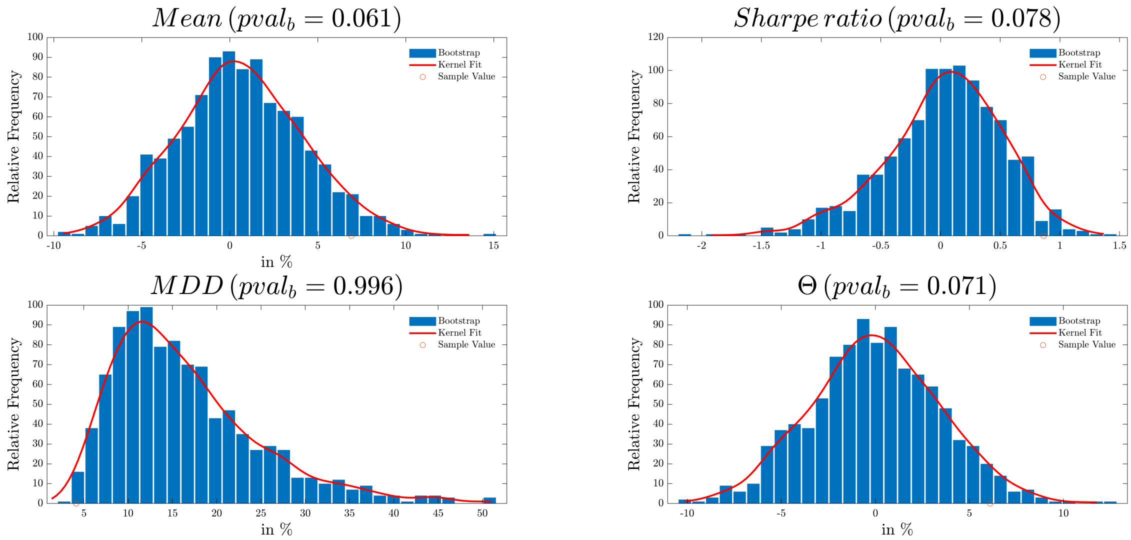
(b) After Transaction Costs

Note: For non-daily rebalancing frequencies missing data points are interpolated linearly.

Figure 7: Bootstrapped Economic Performance of ALP_{HML} (US-Investor)



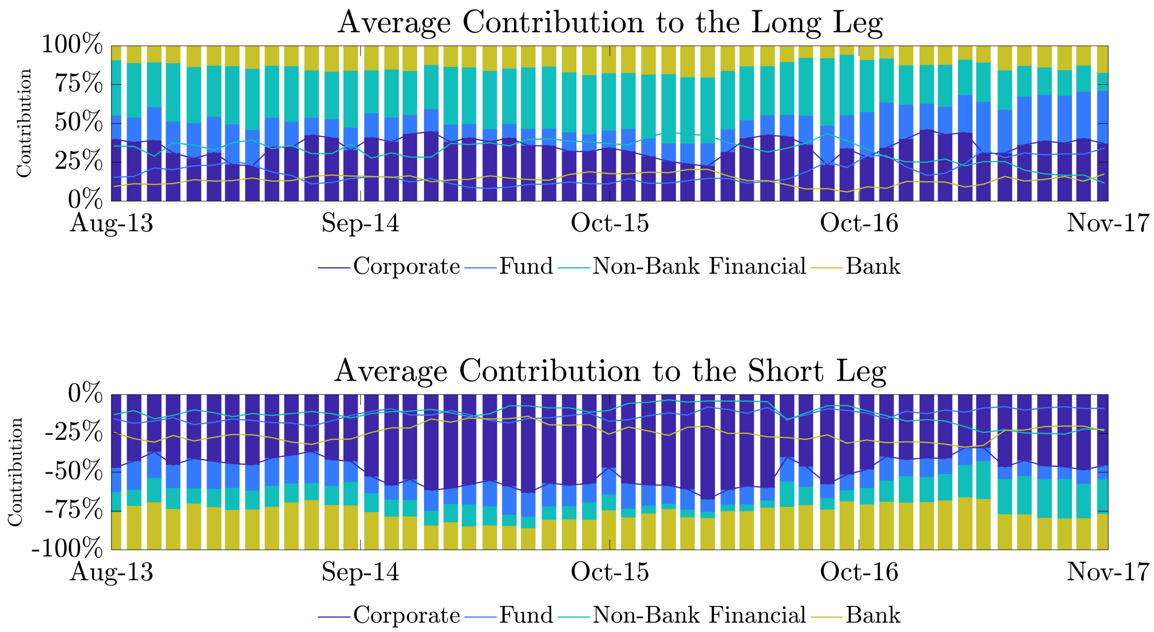
(a) Prior Transaction Costs



(b) After Transaction Costs

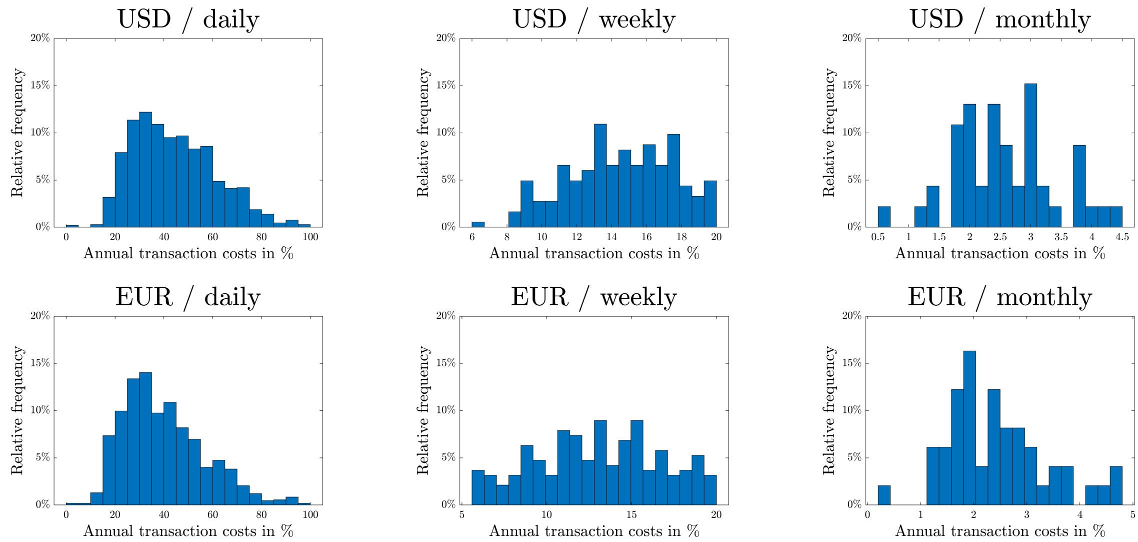
Note: Panel a) and b) depict bootstrapped p-values using 1,000 bootstrap repetitions for ALP_{HML} prior and after transaction costs, respectively, for a US-investor perspective. The upper-left plot displays the annualised mean excess return ($Mean$), the upper-right plot displays the annualised Sharpe ratio, the lower-left plot displays the maximum drawdown (MDD), and the lower-right plot displays the Θ performance measure of Goetzmann et al. (2007) based on monthly rebalancing. The bootstrapped p-values ($pval_b$) are reported in parenthesis in the titles.

Figure 9: Average Contribution to the Long and Short Leg



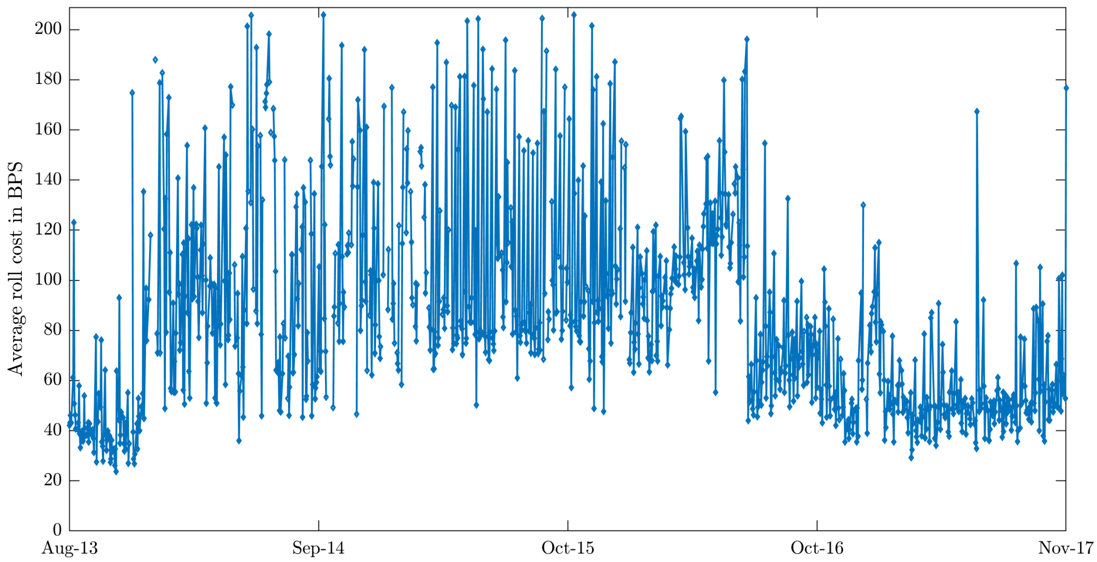
Note: Compute the relative share of each agent's $\alpha_m^{j,k}$ to the aggregate α_m^k and calculate the mean across all currency pairs.

Figure 10: Distribution of Annual Trading Costs



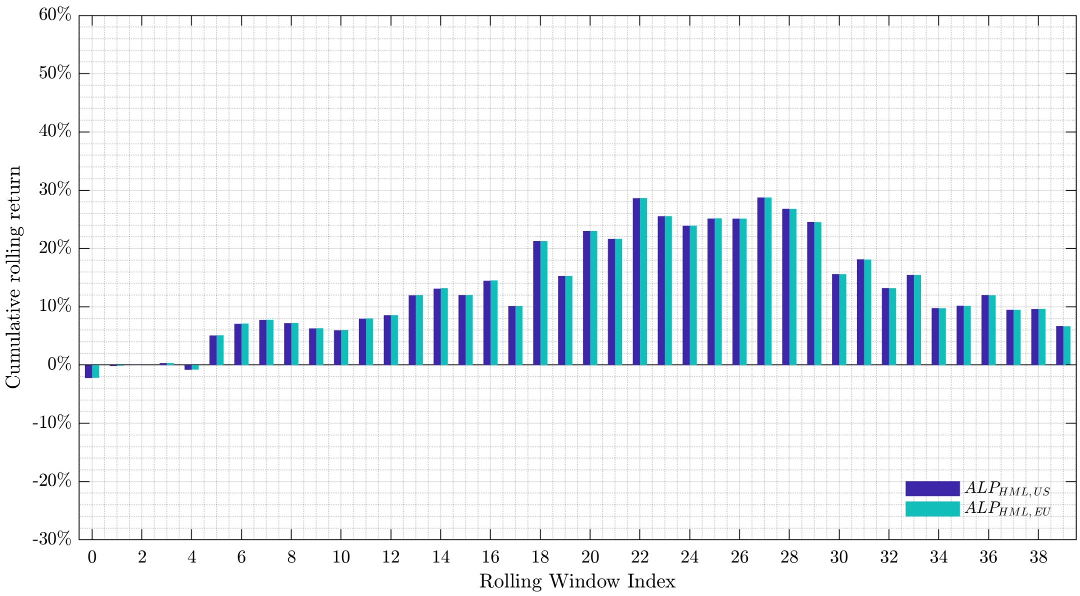
Note: This figure shows the empirical distribution of annual transaction costs for different rebalancing frequencies: daily, weekly, and monthly. Annualised transaction costs are approximated by the cost per trade times number of trading days, weeks, and months per year.

Figure 11: Annualised Roll Costs in BPS



Note: The average is calculated after removing outliers relative to the median.

Figure 12: Cumulative Rolling *Gross* Returns



Note: Rolling window *gross* returns for monthly rebalancing and one year investment horizon.