Essays In International Finance

by

Ingomar Krohn

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Warwick Business School

September 2018

Contents

List of Tables

List of Figures

Acknowledgments

I am indebted to my supervisors Philippe Mueller and Paul Whelan, and Michael J. Moore for their guidance and mentorship during my PhD journey. I am grateful for the numerous meetings and discussions, and the opportunities to collaborate on interesting projects.

Also, I would like to thank my co-authors Alexander Mende, Vikas Raman, and Vladyslav Sushko for sharing their wisdom and for being always available to me as point of contact, for questions, and for discussions. I truly learnt a lot from every meeting and I am thankful for these collaborations.

I am extremely grateful to my family. The continuous and unlimited support of my parents, Sabine and Ingomar, and my sister, Niniane, during all of my studies has been invaluable. I could not have done this journey without them. Further, I cannot be thankful enough to my partner Wai Yee who stood by my side at all times. She has been supporting me in every possible way.

Lastly, I would like to mention my friends and PhD colleagues in Warwick, and abroad, and thank them for casual conversations about possible research ideas, countless joyful discussions, many critical debates, and their support. A very incomplete list includes: Andrew, Danilo, Iram, Johannes, Julius, Kolja, Miranda, Oscar, and Pedro.

Finally, I greatly acknowledge financial support from the Economic and Social Research Council and Doctoral Training Centre in Warwick [grant no. 1500668].

Declarations

This thesis is submitted to the University of Warwick in support of the requirements for the degree of Doctor of Philosophy. I confirm that I have not submitted the thesis for a degree at another university.

Chapter 2 is circulated as a working paper with the title ”Dealer information and macro fundamentals - new evidence on hybrid exchange rate models” and co-authored by Michael J. Moore. A version of this chapter is accepted for publication at the *Journal for International Money and Finance.*

Chapter 3 is circulated as a working paper with the title ”Performance, Persistence, and Pay: A New Perspective on CTAs”, and co-authored by Alexander Mende, Michael J. Moore, and Vikas Raman.

Chapter 4 is circulated as a working paper with the title ”FX Spot and Swap Market Liquidity and the Effects of Window Dressing”, and co-authored by Vladyslav Sushko.

Ingomar Krohn

September 2018

Abstract

This thesis consists of three chapters, in which I examine recent developments in the area of international finance. In Chapter 2, I introduce a new class of hybrid exchange rate models that combine macroeconomic variables with information from a foreign exchange interdealer trading platform. Building upon the work by Chinn and Moore (2011), I examine the power of hybrid models to explain and forecast exchange rate dynamics. I provide compelling evidence that hybrid models produce more accurate in-sample predictions than a conventional macroeconomic Taylor rule and well-established market microstructure models.

Chapter 3 provides an analysis of recent developments in the commodity trading advisor (CTA) industry, a growing alternative investment class, in which fund managers primarily take long and short positions in derivative markets. Based on the largest and cleanest cross-sectional CTA dataset employed in the literature to date, I assess the highly debated performance of CTAs (Bhardwaj et al. (2014), Gregoriou et al. (2010)) and discuss different characteristics of their return dynamics. Furthermore, I provide empirical evidence for a robust link between manager skill, manager compensation, and future fund performance. The results are consistent with a rational market where investors compete to invest with successful CTA managers who set their fee structure to signal their skills to investors.

In Chapter 4, I examine liquidity dynamics of FX spot and swap instruments. While trading volume of FX derivatives has been growing in the post-financial crisis period (BIS (2016)), knowledge about these instruments' liquidity dynamics is still limited. In this essay, I show that market and funding liquidity are inherently linked. In particular, I find that changes of dealers' quoting activity in the FX spot market has a significant impact on liquidity conditions in FX derivative markets. Further, FX dealers' quoting activity exhibits distinct seasonal up- and down-swings around quarter-end months. In these periods, small-volume and less-informed dealers appear to substitute large dealers as market makers. This change in the composition of the dealer ecosystem leads to a decline in market liquidity and to an increase in funding costs of FX derivative instruments. In line with Brunnermeier and Pedersen (2009), the increased funding costs seem to induce liquidity spirals at the end of the year, whereby a worsening of funding conditions is associated with lower levels of market liquidity in currency spot and swap markets.

Chapter 1

Introduction

In this thesis, I discuss three different topics in the area of international finance. Two chapters focus explicitly on developments in the foreign exchange market, the deepest asset class with trading volume exceeding US$ 5 trillion a day (BIS (2016)). In the other chapter, I analyse recent dynamics in the commodity trading advisor (CTA) industry. Managers of these alternative investment vehicles do not only actively trade currency options and futures but take long and short positions in various derivatives and across most major asset classes. As indicated by the industry's growing assets under management over the last twenty years or so, CTAs have established themselves as a popular alternative investment class and they have become an inherent component of today's financial market infrastructure.

While all three chapters are directly or indirectly related to international currency markets, an additional common characteristic is their distinct empirical component and the extensive data analysis which was conducted for each essay. In two chapters I employ high-frequency based data that was originally obtained in raw format at the tick-by-tick frequency from FX trading platforms. In the first essay, the data source is Thomson Reuters Dealing 3000 interdealer platform, which provides information on quotes at the top of the order book and on executed trades in the FX spot market. In the third essay, I use information about bid and ask prices of spot and derivative instruments, and about the name and location of active dealers obtained from Thomson Reuters Tick History Database. Both rich and very detailed datasets offer various angles to study recent dynamics in the foreign exchange market. The analysis in the second chapter which concerns dynamics in the CTA industry is primarily based on conventionally employed monthly return information for a large cross-section of funds. I use data from Barclay's Hedge Fund database and from proprietary datasets, which are all provided by the managed account specialist Risk and Portfolio Management SB. With the aim to contribute to the international finance literature, the structure of the thesis can be summarized as follows.

In Chapter 2, I introduce a new class of hybrid exchange rate models that combine macroeconomic variables and information from interdealer trading platforms. Building upon the work by Chinn and Moore (2011), I assess the power of hybrid models to explain and predict exchange rate dynamics. In contrast to earlier studies, I employ information from one of the largest foreign exchange interdealer datasets, which covers nineteen U.S. dollar and euro currency pairs for a period of more than ten years. In addition, I examine the impact of different measures of trading and quoting activity - namely market order flow (Evans and Lyons (2002)) and net limit order flow (Kozhan et al. (2015)) - on the dynamics of currency prices at the monthly frequency and as a component of a hybrid exchange rate model. The comprehensive analysis points towards largely unexplored benefits that emerge from combining information from conventional macroeconomic and market microstructure models. The results from the in-sample analysis show that the hybrid model produces almost always more accurate predictions than its individual model components. In contrast, a simple out-of-sample forecasting exercise produces mixed evidence across the currency cross-section. Depending on the currency pair, the hybrid model is outperformed by a conventional Taylor rule or simple market microstructure models.

Chapter 3 provides an extensive empirical analysis of recent developments in the commodity trading advisor industry (CTA), using the largest and cleanest cross-sectional dataset on fund characteristics, trading strategies, and return information analyzed so far. Contrary to recent studies (Bhardwaj et al. (2014), Gregoriou et al. (2010)), I show that CTA managers generate positive significant net-excess returns for investors. These net returns are highly positively skewed and move counter-cyclically to equity markets, offering investors an alternative investment opportunity with unique risk-return dynamics. In addition, I find that CTA managers produce significant annualized gross abnormal excess returns of more than 5% that cannot be explained by conventional factor models. Lastly, I document that the performance of managers is persistent. Following Berk and van Binsbergen (2015), I show that manager skill and manager compensation predict future fund performance for up to 36 months. These findings are consistent with a rational market where investors compete to invest with successful CTA managers who use fees to signal their skills to investors.

In Chapter 4, I return to the analysis of international currency markets and I examine liquidity dynamics of FX spot and swap instruments. While trading volume of FX derivatives has been growing in the post-financial crisis period (BIS (2016)), knowledge about these instruments' liquidity dynamics, and the link with FX spot market activity is still limited. Previous research examining FX liquidity so far (e.g. Mancini et al. (2013), Karnaukh et al. (2015)), has focused purely on spot market liquidity characteristics and largely ignored other currency instruments. In this essay, in contrast, I consider spot and swap markets at the same time, show that market and funding liquidity are closely related, and document that liquidity dynamics across these FX instruments are inherently linked. In particular, I find that quoting activity of dealers in the spot market has an impact on liquidity conditions in the swap market. Further, dealers' quoting activity exhibits seasonal up- and down-swings around quarter-end periods, during which small-volume and less informed dealers appear to substitute large dealers as market makers. This leads to a decline in liquidity and increase in funding costs of FX derivative trading. In line with Brunnermeier and Pedersen (2009) the increased funding costs, seem to induce liquidity spirals at year-ends, whereby a worsening of funding conditions is associated with lower levels of market liquidity in currency spot and swap markets.

Chapter 2

Dealer information and macro fundamentals - New evidence on hybrid exchange rate models

2.1 Introduction

This paper introduces a new class of hybrid exchange rate models that combine foreign exchange (FX) market microstructure-based order flow measures and macroeconomic fundamentals in one and the same model. The analysis is motivated by recent findings in Chinn and Moore (2011), which provide empirical evidence that models accounting for FX dealer information and macroeconomic measures improve the performance of economic models. This paper builds upon this novel result. Specifically, I exploit the link between a conventional Taylor rule and changes in currency prices to construct hybrid exchange rate models. I argue that a bridge between a conventional macroeconomic approach and FX trading dynamics can be built by proxying dealers' risk premia by the aggregated trading and quoting activity in the FX interdealer market. I evaluate the in-sample performance of such hybrid models using various criteria and conduct a stylized one-month ahead out-of-sample forecasting exercise to show the robustness of this finding and, subsequently, discuss the advantages and shortcomings of hybrid exchange rate models. Examining the benefits of combining models from the recent FX market microstructure literature (Evans and Lyons (2002)) with conventional Taylor rule fundamentals (Engel and West (2005), Engel et al. (2008)), this paper addresses various research questions: Do hybrid models fit exchange rate data more accurately than their individual nested model components? Is net limit order flow a significant driver of currency dynamics in a hybrid exchange rate model framework and at the comparably low monthly frequency? What is the impact of cross-currency interdependencies on the analysis? How does a hybrid model perform in a stylized out-of-sample forecasting exercise compared to conventional macroeconomic and market microstructure models?

Equipped with a new class of hybrid models, I discuss these questions and aim to contribute to the literature in various ways. First, this paper provides evidence that hybrid models have a more accurate model fit than its individual components separately. I use various model selection criteria and short-term in-sample predictions to illustrate the advantages of employing hybrid exchange rate models. The performance of hybrid models is superior to conventional models for U.S. dollar exchange rates and for the most frequently traded euro pairs, such as the EUR/USD and EUR/JPY. For some currency pairs, the increase in the hybrid model's goodness of fit measure is even larger than the aggregated $\bar{R}^2$ measure of the individual nested models. In particular these cases highlight the possible gains of using hybrid models to analyse exchange rate dynamics. The findings are robust to different specifications of order flow that account for the time-varying trading and quoting activity in the FX interdealer market.

Second, I build upon recent results by Kozhan et al. (2015) and show that net limit order flow is a significant driver of exchange rates at the monthly frequency. While the level of significance varies largely across currencies, net limit order flows' impact on prices is almost always smaller than the impact of market orders. These findings are in line with Bloomfield et al. (2005) who argue that informed traders use market orders and act as liquidity takers at the beginning of the trading period, before changing their quoting activity, increasingly quote limit orders and provide liquidity. In this study, this behaviour is empirically captured by the negative correlation coefficient between market and net limit orders and by the smaller coefficients associated with net limit orders that I obtain in the regression analysis. While the impact of net limit orders has been examined at the intraday and daily level, to the best of my knowledge this study is the first to document their significant impact on returns as part of a macroeconomic model and at the monthly frequency.

Third, I use the comprehensive cross-section of the underlying FX order flow dataset and conduct seemingly unrelated regression (SUR) analysis to highlight the importance of cross-currency interdependencies for the analysis of exchange rate. Accounting for the significant correlation across euro and U.S. dollar pairs, I employ an adapted measure for the system-wide model fit that explicitly penalizes the larger number of estimates in the system. This measure confirms that the hybrid approach produces a better system-wide goodness of fit than its individual model components. Further, I use the SUR model as well as panel fixed effects estimations to conduct a stylized one-month ahead out-of-sample forecasting exercise. In both regression approaches, I find that hybrid models outperform a conventional macroeconomic Taylor rule for several U.S. dollar pairs, but often produce higher mean square errors than market microstructure models. For euro pairs, market microstructure models produce the most accurate out-of-sample forecasts. While the focus of this paper is the in-sample analysis, these findings suggest that hybrid models can be potentially useful to forecast exchange rates out-of-sample.

Finally, all empirical evidence presented in this study is based on one of the largest foreign exchange interdealer order flow datasets analysed so far. I examine 10 U.S. dollar and nine euro currency pairs between January 2004 and February 2014. To the best of my knowledge this is the broadest cross-sectional coverage of foreign exchange order flow data, covering information not only about trades but also about limit order submissions and cancellations at the top of the order book. Accounting for these two measures of dealer activity, for different base currencies, and for a period of more than 10 years, the dataset provides new insights into the link between currency prices and order flows in the foreign exchange market.

The paper is organized as follows. In section 2, I review recent developments in the foreign exchange rate literature. The research methodology is described in section 3, and I introduce the data in section 4. The main findings are presented in section 5. Section 6 concludes the discussion.

2.2 Literature Review

The apparent disconnect between macroeconomic fundamentals and exchange rates is well documented in the literature (e.g. Obstfeld and Rogoff (2000)) and has received increasing attention since the seminal paper by Meese and Rogoff (1983). Despite numerous empirical assessments, however, researchers' success in explaining exchange rate movements remains limited if solely macroeconomic variables are used as explanatory factors.

An alternative angle to understand exchange rate patterns is provided by market microstructure models, which established themselves as a new stream in the exchange rate literature within the last 15 years or so. Instead of focusing on macroeconomic fundamentals, the portfolio shift model by Evans and Lyons (2002) determines exchange rate changes by investors' demand for a currency. Evans and Lyons show that accounting for data from FX trading platforms, measured by market order flow, increases the explanatory power of exchange rate models significantly. Several studies subsequently confirm this finding (e.g. Killeen et al. (2006)) and show that information incorporated in market microstructure variables is a crucial driver of exchange rate.

Surprisingly, macroeconomic and market microstructure models have evolved largely separately from each other, although both literature streams attempt to explain the same phenomenon: the dynamics of exchange rates. Possible benefits from combining different modelling approaches have not been systematically exploited yet, even though a synergy of models could help to improve academics' understanding of the development of currency prices.

One exception is Chinn and Moore (2011), who examine an ad-hoc hybrid version of the traditional money-income model that includes market order flow as an additional regressor. Estimating an error correction model for dollar/euro and dollar/yen currency pairs, the authors illustrate that a hybrid exchange rate model has a superior model fit compared to the individual nested model components. The improvement in explanatory power is driven by the fact that market order flow aggregates otherwise disperse private information and make it public to a larger group of FX dealers. Consequently, market participants benefit from a more comprehensive set of information on which they can base their decisions.

This paper builds upon these findings, but I extend the modelling approach by Chinn and Moore (2011) in several dimensions. First, I derive the hybrid macroeconomic model from the link between a Taylor rule and changes in currency prices. This step is motivated by the recently documented success of Taylor rule models to explain exchange rate patterns (Engel et al. (2008), Molodtsova and Papell (2009)). Generally, Taylor rules provide a link between short-term interest rates-the central banks' main direct policy tool—and dynamics of macroeconomic fundamentals. These are inherently influencing asset prices such that the bridge to exchange rate changes can be built using the uncovered interest rate parity condition.

Second, the class of hybrid models in this study encompass up to two measures from the FX interdealer markets that capture different dimensions of dealer activity. I use information about exchange rates that are incorporated in market order flow following Evans and Lyons (2002) portfolio shift model. The measure is defined as the sum of net buyer-initiated and seller-initiated trades and captures demand-shifts for a particular currency. Moreover, I use a recently introduced transaction flow measure, namely net limit order flow, as additional explanatory variable. As outlined in detail in the data description, net limit order flow is defined as the difference between executed and cancelled limit orders that lead to price changes at the top of the order book. In line with the original work by Kozhan et al. (2015), I postulate that dealers' submitted and cancelled limit orders have an impact on the price discovery process of exchange rates.

By incorporating this additional measure, the analysis not only considers a subset of dealers' activity in the form of market orders but accounts for a larger set of trading dynamics that take place in the FX interdealer market. The importance of this additional dimension of quoting activity at the intra-day and daily frequency has been recently documented by Kozhan et al. (2015). This paper is the first to assess the impact of net limit order flow within the framework of a macroeconomic model and at the monthly frequency. Finally, while most of the existing FX market microstructure literature focuses on the most liquid U.S. dollar pairs, the scope of the employed dataset allows for a more comprehensive empirical assessment. I look at a considerably larger cross-section of pairs that includes euro and U.S. dollar denominated exchange rates, fewer liquid pairs, currencies from emerging and developed economies, and from different exchange rate regimes that have received less attention so far.

2.3 Methodology

The starting point to construct the hybrid exchange rate models is a conventional Taylor rule that establishes a link between short-term interest rates and macroeconomic fundamentals. The relationship between the central bank policy tool and macroeconomic developments is derived following the approach by Molodtsova and

Papell (2009), who define a central bank's policy function as

$$i_t^* = \pi_t + \phi(\pi_t - \pi) + \gamma y_t + r^* + \kappa q_t \tag{2.1}$$

where $i*_t$ is the central bank's target for the interest rate, π^* is the target level of inflation, π_t refers to the inflation rate, y_t is the output gap, measured as the difference between potential and current output in period t, r^* is the equilibrium level of the real interest rate, and q_t refers to the real exchange rate. The last factor is included following the work of Clarida et al. (1998), who argue that central banks take into account the stability of currency prices when setting the interest rate. Next, combining the central bank's target for inflation and interest rate in one parameter ($\mu = r^* - \phi\pi^*$) and $\varphi = (1 + \phi)$, Equation (2.1) can be rewritten as

$$i_t^* = \mu + \varphi\pi_t + \gamma y_t + \kappa q_t \tag{2.2}$$

The intercept term (μ) measures deviations from the short-term interest rate and the target rate of inflation. Further, I allow the Taylor rule specification to account for the impact of interest rate inertia, defined as

$$i_t = (1 - \rho)i_t^* + \rho i_{t-1} + v_t \tag{2.3}$$

where i_t denotes the short-term interest rate and $\rho \in [0, 1]$ is a smoothing parameter that captures the gradual adjustment of interest rates to the target level. Substituting (2.2) in (2.3), I derive the following policy response function

$$\hat{i}_t = (1 - \rho)(\mu + \varphi\hat{pi}_t + \gamma\hat{y}_t + \kappa\hat{q}_t) + \rho\hat{i}_{t-1} + \hat{v}_t \tag{2.4}$$

where "^" denotes variables of the foreign country. The home country's policy response function has the same set-up, but the parameter of the real exchange rate is set to zero ($\kappa = 0$). It implies that the interest rate is determined only by its own lagged term, inflation, and output gap:

$$\imath_t = (1 - \rho)(\mu + \varphi\pi_t + \gamma y_t + \kappa q_t) + \rho i_{t-1} + v_t \tag{2.5}$$

Assuming both countries set their interest rates according to Equations (2.4) and (2.5), the interest rate differential can be written as

$$i_t - \hat{\imath}_t = \alpha + \beta_1(\pi_t - \hat{\pi}_t) + \beta_2(y_t - \hat{y}_t) + \beta_3(i_{t-1} - \hat{\imath}_{t-1}) + \beta_4 q_t + v_t - \hat{v}_t \tag{2.6}$$

where the parameters of the policy response functions are summarized as $\alpha = (1 - \rho)\mu, \beta_1 = (1 - \rho)\varphi, \beta_2 = (1 - \rho)\gamma, \beta_3 = \rho$, and $\beta_4 = (1 - \rho)\kappa$.

As the relationship between interest rate differential and macroeconomic fundamentals is established, I turn next to the link between currency prices, interest rates, and order flow dynamics. A crucial starting point is the uncovered interest parity (UIP) condition given by

$$s_t - s_{t-1} = i_{t-1} - \hat{\imath}_{t-1} + \epsilon_t \tag{2.7}$$

where the change in currency prices between period t and $t - 1$ determined by the lagged interest rate differential between home and foreign country ($i_{t-1} - \hat{\imath}_{t-1}$) and a residual term ϵ_t. If UIP holds, the link between currency prices and fundamentals can be built by substituting the interest rate spread from Equation (2.6) into the UIP condition. For the analysis in this paper, I depart from this assumption and account for the fact that UIP does not necessarily hold. I follow Breedon et al. (2016) and argue that fluctuations of currencies' prices are not only driven by the interest rate spread between two countries but also by an additional risk premium term, which I denote as δ_t. As the future path of exchange rates is unknown to market participants, currencies are risky assets, so that depending on their expectations about the future dynamics of the exchange rate and the likelihood of sudden shocks to currency prices, dealers require additional compensation for holding currencies as part of their portfolio. This additional risk component ultimately has an impact on the exchange rates, as it affects dealers' quoting activity.

To allow for the additional risk premium, I follow Breedon et al. (2016) and decompose the residual term ϵ in a risk premium term (δ_t) and an error term (u_t), so that Equation (2.7) can be re-written as

$$s_t - s_{t-1} = i_{t-1} - \hat{i}_{t-1} + \delta_t + u_t \tag{2.8}$$

While it is notoriously difficult to measure risk premia precisely, I assume dealers' trading decisions and quoting activity in the FX market can serve as a reasonable proxy for dealers' decisions to hold and trade a currency pair. Hence, holding a larger fraction of a currency as part of the portfolio, dealers increase their exposure to shocks affecting the underlying asset and, therefore, the risk premium for holding the asset must be larger. These shifts in demand for certain currencies in turn are captured by aggregated market and net limit order flow in the interdealer market and, therefore, can be considered as a reasonable proxy for the risk premium in Equation (2.8). Under this assumption, the hybrid model can be formulated as a combination of Equations (2.6) and (2.8), so that

$$s_t - s_{t-1} = \alpha + \beta_1 \tilde{\pi}_{t-1} + \beta_2 \tilde{y}_{t-1} + \beta_3 i_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t \tag{2.9}$$

where $\tilde{\pi}_{t-1} = \pi_{t-1} - \hat{\pi}_{t-1}$, $\tilde{y}_{t-1} = y_{t-1} - y_{t-1}$, $i_{t-2} = i_{t-2} - i_{t-2}$, denote the country differentials for inflation, output gap and lagged interest rates, respectively.

The hybrid model in Equation (2.9) captures the benefits of combining macroeconomic fundamentals and FX dealer information in one single approach. First, the hybrid model nests the link between a conventional macroeconomic Taylor rule and changes in currency prices if UIP holds ($\gamma_1 = \gamma_2 = 0$). In such a model, I expect estimates $\beta_1, \beta_2, \beta_3, \beta_4$ to be positive. An increase in the home country's inflation is associated with a contractionary monetary policy response. For example, as a response to an increase in price levels, central banks pursue a tighter monetary policy and increase the country's short-term interest rate. Such a policy decision translates into an appreciation of the exchange rate ($\beta_1 > 0$). As Molodtsova and Papell (2009) point out, an increase in the level of inflation leads not only to a contemporaneous rise in the home interest rate but also affects market participants' expectations about the long-lasting impact of the policy intervention. The revision of expectations can result in a further appreciation, which accelerates the initial impact of the change in currency prices.

Further, β_2 captures the impact of output gap differences between two countries. A positive spread implies an increase in economic activity that exceeds the potential output level. This divergence is likely to occur during economic booms and business cycle upswings and, in combination with inflationary pressure, may lead to an adjustment of the central bank's interest rate path. Following the same line of argument as before, the increase in prices is associated with a rise in the interest rate and the appreciation of the base currency compared to the foreign currency ($\beta_2 > 0$).

The impact of interest rate inertia is captured by β_3 and is expected to be positive as an increase in interest rate in the home country is associated with an appreciation of the domestic currency. Since interest rates enter the Taylor rule specification with a lag, I account for the fact that changes in monetary policy may not be immediately incorporated in asset prices ($\beta_3 > 0$).

Lastly, the impact of an increase in the real exchange rate is associated with increasing returns. Since real exchange rates are determined by the sum of log nominal exchange rate and the log price differential, an increase in the real exchange rate, ceteris paribus, is driven by lower price levels in the home country or higher price levels in the foreign country. The higher the domestic price level relative to the foreign price level, the more distinct the impact on exchange rate returns will be ($\beta_4 > 0$).

Second, Equation (2.9) nests the portfolio shift model ($\beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$), in which contemporaneous market order flow has a significant and positive impact on the price discovery process of currencies (Evans and Lyons (2002)). In line with the original model, I expect the estimate of γ_1 to be positive ($\gamma_1 > 0$) because an increase in order flow is associated with a higher net demand for the home currency. The shift in demand leads to an increase in the exchange rate.

Third, the hybrid model accounts for the portfolio shift model, which also includes net limit orders ($\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$). In this setup, I expect estimates of both order flow components to be positive, as the buying pressure from net limit order flow works in the same direction as that of market orders. Larger demand for the base currency leads to an appreciation of the exchange rate. Therefore, γ_2 is expected to be positive. The dealer's decision to submit either market orders or limit orders, however, depends on the signal and strength of private information. Following Kozhan et al. (2015) and Bloomfield et al. (2005) I suspect that dealers primarily employ market orders to exploit gains from private information and so price signals from limit orders should be smaller than signals from market orders. Therefore, I expect that the magnitude of γ_1 is larger than that of γ_2.

2.4 Data

The empirical analysis is conducted at the monthly frequency, since most macroeconomic data are not available at a higher level. I use exchange rate information for 19 currency pairs, which are obtained from the interdealer platform Reuters Dealing. The sample period starts in January 2004 and ends in February 2014 (122 observations) and covers more than 10 years of data. I split the data into two separate panels, where the euro (EUR) and the U.S. dollar (USD) are the base currencies, respectively. The euro serves as numeraire currency for the following nine currency pairs: Swiss franc (EUR/CHF), Czech koruna (EUR/CZK), British pound (EUR/GBP), Hungarian forint (EUR/HUF), Japanese yen (EUR/JPY), Norwegian krone (EUR/NOK), Polish zloty (EUR/PLN), Swedish krona (EUR/SEK) and U.S. dollar (EUR/USD). For currency pairs denominated in U.S. dollars, I consider the Canadian dollar (USD/CAD), Swiss franc (USD/CHF), British pound (USD/GBP), Israeli shekel (USD/ILS), Indian rupee (USD/INR), Japanese yen (USD/JPY), Mexican Peso (USD/MXN), Polish Zloty (USD/PLN), Singapore dollar (USD/SGD) and South African rand (USD/ZAR). For each pair, I construct a time-series of currency returns (Δs_t), which is defined as the difference of the end-of-month log spot exchange rate between month t and $t-1$.

The order flow data are obtained from Thomson Reuters Dealing 3000, which is one of the largest electronic trading platforms in the FX market. While the original dataset contains order flow information for 80 exchange rates, I only choose currency pairs that are available for the entire sample period and that are denominated by either the euro or U.S. dollar. The currency pairs included in the analysis account for approximately 73% of foreign exchange market global turnover in April 2013 (BIS (2013b)). In contrast to earlier empirical assessments, the sample covers two base currencies, and includes developing and emerging markets currencies as well as different exchange rate regimes. While most of the advanced economies are classified as free-floating currency regimes that follow inflation-targeting monetary policy frameworks, some of the smaller economies employ more restrictive exchange rate arrangements. These include Singapore, Switzerland, the Czech Republic, Hungary, Israel, South Africa, and India (IMF (2014)).

While the original transaction data from Thomson Reuters is obtained at the intraday frequency, I transform order flow measures to the monthly level by aggregating order flow data within each month. Following previous studies (e.g. Love and Payne (2008)), I measure market order flow as the number of buyer-initiated minus seller-initiated trades. Since I only have information on the number of trades but not on traded volume, the implicit assumption of the market order flow measure is that all trades are of equal size. As shown by previous research, this is not a concerning shortcoming and the significant impact of transaction flow data on exchange rate returns exists, even if order flow is measured by the number of trades instead of trading volume (Rime et al. (2010)). I denote the measure of market order flow in month t as (mo_t)

In addition to market order flow, I employ a new microstructure-based measure called net limit order flow. Following Kozhan et al. (2015), net limit order flow is based on submitted and cancelled limit orders that change the price at the top of the order book. The measure is constructed in the following way. I construct a time series of net order submissions, which accounts for the difference between the number of bid and offer submissions, and a second time series of net cancellations, which includes the difference between the number of cancellations of bid and offer orders. Equipped with these two series I define net limit order flow as the difference between net limit orders and net cancellations. The measure considers that private information is transmitted through limit order quoting activity and it therefore differs fundamentally from the conventional market order flow measure. Further, net limit order flow explicitly accounts for submitted and cancelled limit orders because we postulate that both dealer decisions can affect the price path of currencies. Since it accounts for two dimensions of dealer activity—limit order submissions and cancellations—the measure differs from limit order imbalances, which have been examined in earlier studies (e.g. Cheung and Rime (2014)). In line with market orders, I aggregate the net limit order flow measure to a monthly frequency and denote the variable in period t as (lo_t)

To illustrate the crucial role of net limit order flow in the interdealer market, Table 2.1 provides an overview of the monthly average trading and quoting activity.

As shown, the average number of trades and the number of submitted and

Table 2.1: Summary Statistics: Market And Limit Orders

This table reports the average number of trades executed through market orders (mo) and the number of net limit orders, defined as the sum of submitted and cancelled orders, (lo) for nine euro and ten U.S. dollar pairs.

Euro			U.S. dollar		
	mo	lo		mo	lo
USD	29,829	359,106	CAD	145,386	171,365
CHF	426	50,440	CHF	338	130,375
CZK	6,860	22,242	USD	189,749	322,253
GBP	92,423	356,597	ILS	4,435	18,474
HUF	11,516	34,150	INR	31,810	47,412
JPY	265	200,627	JPY	3,231	180,895
NOK	26,952	101,711	MXN	68,469	201,141
PLN	19,288	65,861	PLN	926	124,702
SEK	29,735	100,374	SGD	30,459	52,775
			ZAR	32,028	127,122

Notes: The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair. The data source is Thomson Reuters Dealing 3000 trading platform.

cancelled limit orders varies significantly across currency pairs. For example, the pair EUR/USD shows one of the highest number of trades and orders across euro currencies, reflecting that it is the most frequently traded currency pair in the foreign exchange market (BIS (2013b)). When the quoting currency is associated with an emerging market or small economy, dealer quoting activity is substantially lower.

The variation in market order flow and net limit orders across pairs illustrated in Table 2.1 can be explained by the fact that Reuters Dealing 3000 is not the only electronic trading platform in the foreign exchange market. For example, it is the main trading venue for commonwealth and many emerging market currencies, as reflected in the comparatively large number of trades and order submissions for EUR/GBP (356,597) and GBP/USD (322,253). In contrast, the average number of market orders involving the Japanese yen in the sample is comparatively low (EUR/JPY 265; USD/JPY 3,231), although the Japanese yen is one of the most frequently traded currencies. Yet, even though dealer activity for certain currency pairs is higher on alternative market venues, transaction dynamics between the main trading platforms are closely linked and highly correlated with each other at low frequencies (Breedon and Vitale (2010)). This should ensure that the results are representative of the overall dynamics in the FX market.

Lastly, even though the average monthly number of market orders is low for some currency pairs (for example when Japanese yen or Swiss franc is the respective quote currency), the dataset has the advantage that it covers submitted and cancelled limit orders so that an additional dimension of dealer activity is measured. The high quoting activity at the top of the order book (lo) may be indicative of the important role played by net limit orders in the foreign exchange market.

An indication of the significant impact of both transaction measures for the dynamics of currency prices is illustrated in Table 2.2. The table shows the correlation coefficients between market order and net limit order flow and the log spot exchange rate for all 19 currency pairs. In 17 out of 19 cases the correlation coefficient between market orders (mo_t) and the change in prices (Δs_t) is positive and statistically significant at least at the 10% level. Only for USD/MXN the coefficient is negative, but the magnitude of the contemporaneous relationship is much lower than for other pairs. This positive correlation is in line with the predictions of the portfolio shift model (Evans and Lyons (2002)). Positive co-movement between currency prices and order flow can be interpreted as net demand for the base currency, which leads to an appreciation of the euro or U.S. dollar, respectively. The correlation is not significant, although positive, when the Swiss franc is the quote currency. A possible explanation for the missing significant linear relation might be the low number of market order submissions on this trading platform, as indicated by Table 2.1.

While the role of market order flow is clear, the impact of net limit order flow (lo_t) on exchange rate changes is more ambiguous at the monthly frequency.

Table 2.2: Correlation Coefficients: Market And Net Limit Orders

This table reports the correlation coefficients between the change in the log spot exchange rate (Δs_t), market order flow (mo_t) and net limit order flow (lo_t), for nine euro and ten U.S. dollar pairs. Correlation coefficients marked in bold are different from zero at least at the 10% level of significance.

Panel A: Euro pairs

	USD		CHF		CZK		GBP		HUF	
	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t
mo_t	**0.47**		0.05		**0.36**		**0.37**		**0.32**	
lo_t	**0.24**	0.03	0.00	**-0.32**	**0.25**	-0.05	-0.06	**-0.45**	-0.02	**-0.21**

	JPY		NOK		PLN		SEK	
	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t
mo_t	**0.10**		**0.42**		**0.37**		**0.28**	
lo_t	**0.20**	-0.02	0.02	-0.03	-0.05	**-0.28**	-0.06	-0.02

Panel B: U.S. dollar pairs

	CAD		CHF		GBP		ISL		INR	
	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t
mo_t	**0.31**		**0.10**		**0.42**		**0.37**		**0.28**	
lo_t	-0.02	**-0.21**	**0.20**	-0.02	0.02	-0.03	-0.05	**-0.28**	-0.06	-0.02

	JPY		MXN		PLN		SGD		ZAR	
	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t	Δs_t	mo_t
mo_t	**0.25**		**-0.17**		**0.29**		**0.33**		**0.41**	
lo_t	**0.30**	**-0.24**	0.13	**-0.33**	0.04	-0.05	**0.31**	**-0.50**	0.09	-0.12

Notes: The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair.

The correlation coefficient is positive for 11 currency pairs, but its magnitude is almost always lower than that of market orders. Yet, the positive correlations can be interpreted as a first indication that limit orders convey information about the exchange rate path. The statistical significance, however, is lower than that for market order flow. As shown, the correlation coefficient between net limit order flow and the exchange rate is only significantly different from zero for six currency pairs. This contrasts with Kozhan et al. (2015), who examine this relationship at the daily frequency. Two explanations come to mind. First, the diverging results based on different data frequencies indicate that information transmitted through limit orders is lower than that transmitted through market orders. Second, as the correlation coefficients are largely not significantly different from zero, information transmitted via net limit order flow is transitory and short-lived. The impact of net limit order flow on the exchange rate vanishes for most currency pairs when the data is aggregated to the monthly frequency.

Yet, net limit order quoting activity is an important determinant for the analysis, as pointed out by the relationship between both order flow variables. It is negative for 18 currency pairs, indicating the changing trading behaviour of informed traders to generate gains from trades based on their private information. As Bloomfield et al. (2005) argue, informed traders use both market and limit orders but their trading behaviour varies over the course of the trading period. While informed traders primarily quote market orders and take liquidity to profit from their private information at the beginning of the trading period, they switch to limit order submissions once the value of information is exploited and currency prices have adjusted. By submitting limit orders during later periods of the trading day, informed traders then provide liquidity to other investors who use market orders to reduce inventory risk. This change in dealer activity is reflected by the negative correlation coefficient between the two contemporaneous flow variables. Further, the negative correlation represents the general trade-off faced by dealers to submit a either limit or a market order. While limit orders are cheaper than market orders, their execution is price contingent and not necessarily immediate. In contrast, market orders are executed at the prevalent market price. The significant correlation coefficients indicate that it is important to include limit order flow in the model even if the analysis is conducted at the monthly level.

In addition to the foreign exchange microstructure data, I construct a dataset of macroeconomic variables. The main source is the IMF's International Financial Statistics database accessed via Datastream. Price levels are measured by the monthly consumer price index (CPI) and inflation is constructed by calculating the monthly difference of CPI (IFS code: $64 - ZF$). Since gross domestic product is available only quarterly, I use the industrial production index (IFS code: $66 - ZF$) as proxy for output for most of the countries. The output gap is calculated using a one-sided Hodrick-Prescott filter. As a proxy for the short-term interest rate, I use money market rates (IFS code: $60 - B–ZF$).

2.5 Results

I begin the empirical analysis and estimate the hybrid model with both order flow measures for the euro and U.S. dollar-based exchange rates. Results are displayed in Table 2.3.

As illustrated, the performance of the hybrid model varies fundamentally across currency pairs. For the most frequently traded exchange rates, the adjusted goodness of fit measure is in the double-digit region. For euro currency pairs the maximum value is 0.32 (EUR/USD) and it increases up to 0.42 for U.S. dollar pairs (USD/CAD). In contrast, much less variation of returns is explained when the Swiss franc (USD/CHF, EUR/CHF) or most emerging markets serve as the quoting currency. For example, for the Mexican peso and the Indian rupee the adjusted R^2 is only 0.02 and 0.05, respectively. For these four currency pairs order flow data

Table 2.3: Regression Results: Hybrid Model

This table reports regression results of the hybrid model $s_t - s_{t-1} = \alpha + \beta_1 \tilde{\pi}_{t-1} + \beta_2 \tilde{y}_{t-1} + \beta_3 i_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t$, where $\tilde{\pi}_{t-1} = \pi_{t-1} - \hat{\pi}_{t-1}$, $\tilde{y}_{t-1} = y_{t-1} - \hat{y}_{t-1}$, $i_{t-2} = i_{t-2} - \hat{i}_{t-2}$, denote the differences of inflation, output gap and lagged interest rates between home and foreign country, respectively. Changes in prices are measured in basis points. The estimates of the intercept are omitted to save space. Numbers in parentheses refer to Newey-West adjusted standard errors. The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair. *,**,*** denote the 10%, 5%, and 1% level of significance, respectively.

Panel A: Euro pairs

	β_1	β_2	β_3	β_4	γ_1	γ_2	$\bar{R}^2$
USD	100.86***	-677.95	-30.85		169.69***	66.64***	0.32
	(35.85)	(1,406.35)	(34.88)		(30.31)	(20.16)	
CHF	10.52	-66.05	13.03	-328.45	226.74	12.13	0.00
	(23.40)	(167.60)	(29.69)	(303.65)	(462.84)	(29.15)	
CZK	10.10*	-343.42	-6.74	-171.49	275.94***	84.74***	0.28
	(21.94)	(203.24)	(21.46)	(265.13)	(60.33)	(26.24)	
GBP	-5.71	-100.78	18.06	-461.34	69.76***	27.95***	0.12
	(51.47)	(483.25)	(41.68)	(435.71)	(9.90)	(13.45)	
HUF	14.75	404.75	0.17	30.64	173.59***	22.54	0.07
	(14.49)	(427.21)	(17.10)	(470.63)	(50.31)	(39.75)	
JPY	-65.00**	822.14	-12	-1015.09*	2704.53***	112.06**	0.23
	(32.98)	(526.52)	(24.47)	(609.30)	(554.51)	(55.51)	
NOK	-16.48	121.07	5.58	-442.26	155.97***	3.3	0.16
	(14.96)	(220.04)	(44.60)	(306.14)	(34.06)	(4.88)	
PLN	-22	601.22	38.7	-333.48	150.88***	16.19	0.13
	(31.26)	(410.42)	(30.60)	(627.11)	(46.20)	(23.19)	
SEK	-35.94	192.33	5.54	-20.69	118.09***	-3.36	0.07
	(40.24)	(123.11)	(21.82)	(437.25)	(36.32)	(5.34)	

Table 2.3: Regression Results: Hybrid Model

This table reports regression results of the hybrid model $s_t - s_{t-1} = \alpha + \beta_1 \tilde{\pi}_{t-1} + \beta_2 \tilde{y}_{t-1} + \beta_3 i_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t$, where $\tilde{\pi}_{t-1} = \pi_{t-1} - \hat{\pi}_{t-1}$, $\tilde{y}_{t-1} = y_{t-1} - \hat{y}_{t-1}$, $i_{t-2} = i_{t-2} - \hat{i}_{t-2}$, denote the differences of inflation, output gap and lagged interest rates between home and foreign country, respectively. Changes in prices are measured in basis points. The estimates of the intercept are omitted to save space. Numbers in parentheses refer to Newey-West adjusted standard errors. The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair. *,**,*** denote the 10%, 5%, and 1% level of significance, respectively.

Panel B: U.S. dollar pairs

	β_1	β_2	β_3	β_4	γ_1	γ_2	$\bar{R}^2$
CAD	24.72	59.17	-75.94**	(453.55)	103.57***	58.17	0.42
	(30.36)	(745.08)	(32.08)	(310.20)	(7.94)	(58.28)	
CHF	87.71***	(406.00)	(43.48)	291.00	224.10	44.01	0.04
	(27.31)	(366.03)	(23.19)	(310.71)	(200.52)	(97.67)	
GBP	3.33	39.96	-92.59**	356.16	61.47***	50.61***	0.22
	(21.05)	(572.82)	(45.67)	(311.53)	(13.80)	(19.06)	
ILS	11.19	(80.37)	-46.61***	(97.81)	428.31***	24.74	0.11
	(12.84)	(396.90)	(16.40)	(255.53)	(139.95)	(36.15)	
INR	9.09	469.59	-18.74*	(339.45)	32.79	56.48	0,05
	(6.39)	(649.14)	(11.41)	(246.43)	(48.59)	(89.90)	
JPY	22.32	965.86*	(17.09)	(20.36)	300.67***	238.70***	0.22
	(15.89)	(472.72)	(15.50)	(209.16)	(80.84)	(43.92)	
MXN	23.27	-3103.32**	(5.50)	(154.66)	-45.54*	10.96	0.06
	(14.35)	(1,346.32)	(25.01)	(242.85)	(26.94)	(16.20)	
PLN	44.41**	734.94	(43.53)	(643.24)	1317.23***	21.49	0.10
	(19.22)	(817.81)	(30.92)	(428.61)	(435.76)	(67.68)	
SGD	(2.41)	(375.52)	(7.53)	56.35	161.14***	217.82***	0.38
	(6.66)	(276.42)	(10.80)	(153.69)	(25.20)	(42.03)	
ZAR	18.18	(781.91)	34.87	396.56	283.95***	24.84*	0,19
	(17.70)	(2,242.08)	(25.16)	(602.67)	(69.56)	(13.84)	

do not add any or only very little explanatory power. As the adjusted R^2 reaches double-digit values for 14 out of 19 currency pairs, I conclude that interdealer order flow information plays a relevant role for exchange rate dynamics at the monthly frequency.

Looking at the coefficients of macroeconomic fundamentals, only 10 out of 75 associated coefficients are significant at the 10% level or higher. For euro pairs, the inflation differential is the main macroeconomic driver (e.g. EUR/USD and EUR/JPY), while coefficients of the output gap are not significant at all.

Yet, even if coefficients associated with macroeconomic variables are significant, the signs of the estimated coefficients often contradict prior expectations. For example, while I expected the output gap differential to be positively related to changes in currency prices, all significant coefficients are negative for U.S. dollar exchange rates. These results are indicative of the disconnect between macroeconomic fundamentals and exchange rates (Obstfeld and Rogoff (2000)).

In contrast to macroeconomic fundamentals, net limit orders and market order flow significantly determine the price path of currencies. The coefficient of market order flow is significant at the 10% level for almost all euro currency pairs and is significant for seven out of 10 U.S. dollar pairs. These results are in line with previous studies that find market order flow is a significant driver of exchange rates even at the monthly frequency. Inferences about limit order flow are more ambiguous. While I find that 18 out 19 coefficients corresponding to net limit order flow (γ_2) are positive, for only half of these cases the coefficients are significantly different from zero. The positive coefficients of both order flow measures confirm that a higher demand for the base currency leads to an appreciation of the exchange rate.

Furthermore, I document that in 16 out of 19 pairs the magnitude of market order flow coefficients is larger than that of limit orders ($\gamma_1 > \gamma_2$). I formally test in Table 2.5 if the difference between the obtained coefficients is significantly different from zero and report the p-value of the t-test associated with the null hypothesis: $H_0 : \gamma_1 - \gamma_2 > 0$.

Table 2.5 shows that the null hypothesis is rejected at least at the 10% level for almost all euro currency pairs and for roughly half of the U.S. dollar pairs. The impact of net limit orders on the exchange rate is significantly smaller than that of market orders, confirming the results of Kozhan et al. (2015). The findings are also in line with Bloomfield et al. (2005), suggesting that market participants first employ market orders to exploit profits from their private information and subsequently increase limit order quoting activity

Table 2.4: Comparison: Market And Net Limit Order Coefficients

This table reports the p-value of an one-sided t-test with the null-hypothesis $H_0 : \gamma_1 - \gamma_2 > 0$ against the alternative hypothesis $H_A : \gamma_1 - \gamma_2 \leq 0$, to assess if the coefficient associated with market order flow is significantly larger than the estimated coefficient of net limit orders. The parameters γ_1 and γ_2 refer to the estimates of market and net limit order flow in a hybrid exchange rate model.

Panel A: Euro pairs					
	USD	CHF	CZK	GBP	HUF
p-value	0.01**	0.24	0.00***	0.04**	0.01**
	JPY	NOK	PLN	SEK	
p-value	0.00***	0.00***	0.00***	0.00***	
Panel B: U.S. dollar pairs					
	CAD	CHF	GBP	ISL	INR
p-value	0.08*	0.32	0.28	0.00***	0.73
	JPY	MXN	PLN	SGD	ZAR
p-value	0.22	0.98	0.00***	0.97	0.00***

Notes: The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair.

once prices adjust and information advantages become smaller.

In a next step, I examine the extent to which the hybrid model outperforms its nested components. Table 5 displays the adjusted R^2 measure for the original market microstructure portfolio shift ($R_1^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$), its extension by Kozhan et al. (2015), which includes net limit order flow ($R_2^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$), a conventional Taylor rule ($R_3^2 : \gamma_1 = \gamma_2 = 0$), a hybrid model only augmented with market order flow ($R_4^2 : \gamma_2 = 0$), and the hybrid model with both order flow components (R_5^2 : No restrictions imposed). The highest adjusted R^2 for each currency pair is marked in bold. In addition, I employ a series of F-tests as model selection criteria to establish whether a hybrid model can be considered the preferred modelling choice over its nested conventional models. Table 6 reports p-values for the respective F-tests.

Table 2.5: Goodness Of Fit: Hybrid Model And Its Nested Components

This table reports the adjusted goodness of fit of the hybrid model and of its nested models. The hybrid model is specified as form $s_t - s_{t-1} = \alpha + \beta_1 \tilde{\pi}_{t-1} + \beta_2 + \tilde{y}_{t-1} + \beta_3 \tilde{i}_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t$. The following restrictions are imposed to estimate the nested components of the hybrid model: $R_1^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$, $R_2^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, $R_3^2 : \gamma_1 = \gamma_2 = 0$, $R_4^2 : \gamma_2 = 0$, R_5^2 : No restrictions.

	Euro pairs						**U.S. dollar pairs**				
	$\bar{R}_1^2$	$\bar{R}_2^2$	$\bar{R}_3^2$	$\bar{R}_4^2$	$\bar{R}_5^2$		$\bar{R}_1^2$	$\bar{R}_2^2$	$\bar{R}_3^2$	$\bar{R}_4^2$	$\bar{R}_5^2$
USD	0.22	0.27	0.11	0.29	**0.32**	CAD	0.34	0.37	0.09	0.41	**0.42**
CHF	0.00	0.00	0.00	0.00	0.00	CHF	0.00	0.00	**0.05**	0.04	0.04
CZK	0.12	**0.18**	0.00	0.12	0.18	GBP	0.08	0.13	0.11	0.20	**0.22**
GBP	0.13	0.14	0.00	0.12	0.12	ILS	0.10	0.10	0.01	**0.12**	0.11
HUF	**0.09**	0.08	0.01	0.08	0.07	INR	0.02	0.01	0.05	0.05	**0.05**
JPY	0.15	0.19	0.05	0.19	**0.23**	JPY	0.05	0.19	0.01	0.07	**0.22**
NOK	**0.17**	0.16	0.00	0.17	0.16	MXN	0.02	0.02	0.04	0.06	**0.06**
PLN	0.13	0.13	0.00	0.13	**0.13**	PLN	0.08	0.07	0.05	**0.11**	0.10
SEK	0.07	0.06	0.00	0.07	**0.07**	SGD	0.01	0.39	0.00	0.11	**0.38**
						ZAR	0.16	0.17	0.03	0.18	**0.19**

Notes: The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair.

Beginning with the results in Table 5, I note that hybrid models produce the highest adjusted R^2 for four euro-denominated exchange rates (EUR/USD, EUR/JPY, EUR/PLN, EUR/SEK), while individual market microstructure models produce higher values for the remaining currencies. In these cases, macroeconomic fundamentals do not add explanatory power to the hybrid model. For U.S. dollar pairs, hybrid models produce the highest R^2 for all exchange rate pairs but the Swiss franc. Interestingly, including net limit order flow in the hybrid model regressions increases adjusted R^2 substantially in some cases. For example, for the Singapore dollar net limit order flow adds a significant proportion to the explanatory power of the hybrid model ($R^2 = 0.38$). For USD/JPY, explanatory power increases from 0.7 to 0.22 when net limit order flow is included in the hybrid model.

Next, I use a series of F-tests as model selection criteria and test for the joint

Table 2.6: F-Tests: Hybrid Model And Its Nested Components

This table reports the p-values for a series of F-tests, for which we impose restrictions on the coefficients of the hybrid model. The hybrid model is specified as form $s_t - s_{t-1} = \alpha + \beta_1\tilde{\pi}_{t-1} + \beta_2 + \tilde{y}_{t-1} + \beta_3\tilde{i}_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t$. We conduct four different F-tests for each currency pair, testing the following null hypotheses: $H_1^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$, $H_2^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, $H_3^2 : \gamma_1 = \gamma_2 = 0$, $H_4^2 : \gamma_2 = 0$.

	Euro pairs					**U.S. dollar pairs**			
	H_1^2	H_2^2	H_3^2	H_4^2		H_1^2	H_2^2	H_3^2	H_4^2
USD	0.00***	0.01***	0.00***	0.01**	CAD	0.00***	0.01**	0.00***	0.09*
CHF	0.87	0.77	0.76	0.82	CHF	0.08*	0.06*	0.68	0.52
CZK	0.03**	0.54	0.00***	0.00***	GBP	0.00***	0.00***	0.00***	0.05*
GBP	0.58	0.75	0.00***	0.24	ILS	0.24	0.2	0.00***	0.56
HUF	0.73	0.68	0.01**	0.58	INR	0.12	0.07*	0.43	0.32
JPY	0.01**	0.05*	0.00***	0.02**	JPY	0.00***	0.08*	0.00***	0.00***
NOK	0.56	0.42	0.00***	0.7	MXN	0.09*	0.07*	0.12	0.53
PLN	0.47	0.39	0.00***	0.51	PLN	0.16	0.11	0.02**	0.74
SEK	0.45	0.36	0.00***	0.61	SGD	0.00***	0.81	0.00***	0.00***
					ZAR	0.07*	0.12	0.00***	0.12

Notes: The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair.*,**,*** denote the 10%, 5%, and 1% level of significance, respectively.

significance of coefficients from four different nested models. Focusing on column H_0^3 in Table 2.6, the low p-value indicates that coefficients on transaction flow variables are jointly non-zero and that the hybrid models are the preferred choice over the conventional Taylor rule for almost all currency pairs in the sample. In 15 out of 19 cases I reject the null hypothesis H_0^3. Further, in seven out of those 15 currency pairs the hybrid model augmented with market and net limit orders should be selected to account for all significant regressors (H_0^4).

While adjusted R^2 values and F-tests already indicate the superior performance of hybrid models over their nested components, I also conduct a one-step ahead in-sample forecasting exercise as an additional model selection criterion. First, I estimate the hybrid models and their nested components and calculate the mean square error (MSE) for each of these models. Second, I use the conventional macroeconomic model as a base line comparison and calculate relative measures for all four models (MSE_1^{Rel} to MSE_4^{Rel}). Results are displayed in Table 2.7; a smaller value than 1 indicates a superior predictive power of the tested model compared to the macroeconomic Taylor rule. The closer the value to zero, the more accurate the model's prediction is. For ease of reading, I marked the lowest relative MSE for each currency in bold.

Table 2.7: In-Sample Predictions: Hybrid Model And Its Nested Components

This table reports the relative mean square errors (MSE^{Rel}) of a one-month ahead in-sample forecasting exercise of the hybrid model $s_{t+1} - s_t = \alpha + \beta_1 \tilde{\pi}_{t-1} + \beta_2 + \tilde{y}_{t-1} + \beta_3 \tilde{i}_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t$. We estimate the model in different specifications and impose the following restrictions for different model: MSE_1^{Rel} : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$, MSE_2^{Rel} : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, MSE_3^{Rel} : $\gamma_2 = 0$, MSE_4^{Rel} : No restrictions. Mean square errors are expressed relative to the conventional Taylor rule mode. Values smaller than one indicate more accurate in-sample performance of the tested model, compared to the conventional Taylor rule. Values in bold mark the smallest relative mean square error for each currency pair.

	Euro pairs					**U.S. dollar pairs**			
	MSE_1^{Rel}	MSE_2^{Rel}	MSE_3^{Rel}	MSE_4^{Rel}		MSE_1^{Rel}	MSE_2^{Rel}	MSE_3^{Rel}	MSE_4^{Rel}
USD	1.099	1.071	0.996	**0.977**	CAD	1.128	1.127	0.987	**0.986**
CHF	1.023	1.021	0.980	**0.979**	CHF	1.020	1.018	0.996	**0.996**
CZK	0.998	0.997	0.996	**0.995**	GBP	1.190	1.184	0.998	**0.998**
GBP	1.010	0.994	0.997	**0.987**	ILS	1.046	1.031	1.000	**0.979**
HUF	1.013	1.007	0.998	**0.993**	INR	1.045	1.045	0.993	**0.988**
JPY	1.070	1.068	0.982	**0.978**	JPY	1.008	0.974	0.998	**0.962**
NOK	1.014	1.012	1.000	**0.999**	MXN	1.033	1.025	1.000	**0.996**
PLN	0.999	0.955	0.993	**0.949**	PLN	1.057	1.052	1.000	**0.996**
SEK	1.010	1.009	0.997	**0.996**	SGD	1.053	1.050	0.999	**0.999**
					ZAR	1.061	1.054	0.984	**0.974**

Notes: The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair.

As illustrated in Table 2.7, the hybrid model shows a better predictive performance than the conventional macroeconomic model. While differences are small for several U.S. dollar pairs, for EUR/PLN or EUR/USD the predictive performance of the hybrid models is notably better. The hybrid models also outperform conventional market microstructure approaches, indicating that the improved performance is not only driven by order flow variables but by the combination with macroeconomic fundamentals.

Lastly, I show that these results are robust to different economic factors, such as the degree of trading and quoting activity. In Table 8 I employ the same series of F-tests as described earlier but instead of employing standard order flow measures, I use alternative specifications that are scaled by the market thickness and dealer activity in the market of each respective currency.

In the panel on the left of Table 2.8, I use normalized order flow measures (mo_t^N,lo_t^N) whereby I scale monthly market order flow by the number of trades and net limit order flow by the total number of submitted and cancelled orders. This alternative measure considers the time-varying degree of trading activity in each month in the interdealer market. In the middle panel, I construct trade- and order-weighted order flow measures as a second alternative measure of dealer activity (mo_t^W,lo_t^W). To this end, I scale daily order flow by the number of trades on each day relative to the total number of trades during each month. Similarly, order-weighted net limit order flow is constructed by multiplying net limit orders on each day by the number of submitted and cancelled orders during the day relative to the number of total limit order book activity over the entire month. Both measures are then aggregated to the monthly level. In contrast to normalized order flow measures, weighted transaction data allocate relatively more emphasis to order flow that is associated with higher dealer activity.

In the panel to the right of Table 8, I report results from a two-stage least square approach, to alleviate concerns about the contemporaneous relationship between order flow regressors and exchange rate changes. In these regressions, I use lagged order flows (mo_{t-1},lo_{t-1}) as instruments for the contemporaneous terms (mo_t,lo_t) that are originally included in the hybrid model in Equation (2.9). To be more concrete, in a first step I regress lagged transaction flow data on contemporaneous values and construct instruments from the first-stage residuals. These are then included in the second-stage estimation of the hybrid model. As lagged flow data are uncorrelated with the error term of the hybrid model but highly correlated with contemporaneous order flow, the approach fulfils the two most important criteria for the choice of an appropriate instrument.

As presented in Table 8, I find that results are qualitatively similar to the original model, that is, they are weakest for the normalized order flow, and almost identical for the weighted order flow and IV estimation. For all three specifications, I reject the null hypotheses of the conventional macroeconomic Taylor rule (H_0^3) in favour of a hybrid model with market order flow for at least seven out of nine euro pairs, and for five out of 10 U.S. dollar exchange rates. For EUR/USD, EUR/CZK, EUR/JPY, and EUR/NOK, and USD/CHF, USD/JPY, USD/SGD, USD/GBP, and USD/CAD, net limit order flow has a significant impact (H_0^4) and improves the explanatory power of the hybrid model. Lastly, both market microstructure models (H_0^1 and H_0^2), are rejected for various currency pairs, pointing toward a higher variation of exchange rate patterns that is jointly explained by macroeconomic variables and transaction flow measures.

Table 2.8: Alternative Order Flow Specifications

This table reports the p-values for a series of F-tests, for which I impose restrictions on the coefficients of the full hybrid model that employs alternative order flow specifications ($\tilde{mo}^{Alt}$,$\tilde{lo}^{Alt}$): $s_t - s_{t-1} = \alpha + \beta_1 \tilde{\pi}_{t-1} + \beta_2 \tilde{y}_{t-1} + \beta_3 \tilde{i}_{t-2} + \beta q_{t-1} + \gamma_1 mo_t^{Alt} + \gamma_2 lo_t^{Alt} + u_t$, where $Alt = N, W, IV$ refer to normalized (N), weighted order flow (W), and the IV-approach (IV), respectively. Normalized order flow refers to monthly order flow scaled by the total number of trades; weighted order flow is the sum of daily order flow weighted by the number of trades and quotes within each month. For the instrumental variable approach, I use lagged order flow as an instrument in a first-step regression and include the obtained residuals as a proxy of contemporaneous order flow dynamics in the second-stage regression. For each alternative order flow measure, I conduct four different F-tests, testing each of the following four null hypotheses: $H_1^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$, $H_2^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, $H_3^2 : \gamma_1 = \gamma_2 = 0$, $H_4^2 : \gamma_2 = 0$. The sample period is January 2004 to February 2014. *,**,*** denote the 10%, 5%, and 1% level of significance, respectively.

Panel A: Euro pairs

	Normalized order flow ($\tilde{mo}^N$,$\tilde{lo}^N$)				Weighted order flow ($\tilde{mo}^W$,$\tilde{lo}^W$)				IV-Approach ($\tilde{mo}^{IV}$,$\tilde{lo}^{IV}$)			
	H_0^1	H_0^2	H_0^3	H_0^4	H_0^1	H_0^2	H_0^3	H_0^4	H_0^1	H_0^2	H_0^3	H_0^4
USD	0.00***	0.00***	0.02**	0.45	0.01***	0.01***	0.00***	0.23	0.00***	0.00***	0.00***	0.01**
CHF	0.51	0.65	0.15	0.18	0.88	0.78	0.93	0.94	0.87	0.77	0.76	0.82
CZK	0.28	0.62	0.01**	0.04**	0.11	0.83	0.00***	0.01 **	0.03**	0.54	0.00***	0.00***
GBP	0.95	0.91	0.45	0.86	0.7	0.75	0.00***	0.41	0.59	0.76	0.00***	0.24
HUF	0.21	0.32	0.10*	0.08*	0.7	0.66	0.00***	0.49	0.62	0.62	0.00***	0.36
JPY	0.10*	0.06**	0.03**	0.35	0.02**	0.04**	0.00***	0.03 **	0.01***	0.06	0.00***	0.02**
NOK	0.13	0.59	0.06*	0.02**	0.63	0.5	0.00***	0.69	0.48	0.35	0.00***	0.69
PLN	0.72	0.63	0.02**	0.56	0.64	0.5	0.00***	0.96	0.48	0.38	0.00***	0.59
SEK	0.35	0.59	0.16	0.08*	0.37	0.32	0.00***	0.44	0.42	0.39	0.00***	0.47

Table 2.8: Alternative Order Flow Specifications

This table reports the p-values for a series of F-tests, for which I impose restrictions on the coefficients of the full hybrid model that employs alternative order flow specifications ($\tilde{mo}^{Alt}$,$\tilde{lo}^{Alt}$): $s_t - s_{t-1} = \alpha + \beta_1 \tilde{\pi}_{t-1} + \beta_2 \tilde{y}_{t-1} + \beta_3 \tilde{i}_{t-2} + \beta q_{t-1} + \gamma_1 mo_t^{Alt} + \gamma_2 lo_t^{Alt} + u_t$, where $Alt = N, W, IV$ refer to normalized (N), weighted order flow (W), and the IV-approach (IV), respectively. Normalized order flow refers to monthly order flow scaled by the total number of trades; weighted order flow is the sum of daily order flow weighted by the number of trades and quotes within each month. For the instrumental variable approach, I use lagged order flow as an instrument in a first-step regression and include the obtained residuals as a proxy of contemporaneous order flow dynamics in the second-stage regression. For each alternative order flow measure, I conduct four different F-tests, testing each of the following four null hypotheses: $H_1^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$, $H_2^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, $H_3^2 : \gamma_1 = \gamma_2 = 0$, $H_4^2 : \gamma_2 = 0$. The sample period is January 2004 to February 2014. *,**,*** denote the 10%, 5%, and 1% level of significance, respectively.

Panel B: U.S. dollar pairs

	Normalized order flow ($\tilde{mo}^N$,$\tilde{lo}^N$)				Weighted order flow ($\tilde{mo}^W$,$\tilde{lo}^W$)				IV-Approach ($\tilde{mo}^{IV}$,$\tilde{lo}^{IV}$)			
	H_0^1	H_0^2	H_0^3	H_0^4	H_0^1	H_0^2	H_0^3	H_0^4	H_0^1	H_0^2	H_0^3	H_0^4
CAD	0.01***	0.01***	0.27	0.86	0.00***	0.02**	0.00***	0.15	0.00***	0.01**	0.00***	0.09*
CHF	0.01**	0.06*	0.06*	0.04**	0.08*	0.06*	0.65	0.55	0.08*	0.06*	0.7	0.54
GBP	0.00***	0.0***	0.7	0.59	0.00***	0.00***	0.00***	0.11	0.00***	0.00***	0.00***	0.05*
ILS	0.19	0.41	0.00***	0.18	0.33	0.25	0.00***	0.8	0.23	0.19	0.00***	0.57
INR	0.1	0.06*	0.33	0.87	0.12	0.07*	0.25	0.33	0.12	0.07*	0.46	0.35
JPY	0.37	0.26	0.02**	0.94	0.00***	0.13	0.00***	0.00***	0.00***	0.08*	0.00***	0.00***
MXN	0.03**	0.06*	0.07*	0.12	0.12	0.07*	0.13	0.95	0.1	0.08*	0.11	0.49
PLN	0.37	0.31	0.00***	0.43	0.18	0.12	0.11	0.78	0.14	0.10*	0.03**	0.75
SGD	0.52	0.63	0.31	0.26	0.00***	0.84	0.00***	0.00***	0.00***	0.81	0.00***	0.00***
ZAR	0.18	0.13	0.32	0.49	0.09*	0.15	0.00***	0.13	0.08*	0.13	0.00***	0.12

2.5.1 Cross-currency Interdependence In The FX Market

The large cross-sectional dimension of the interdealer order flow dataset provides the additional opportunity to account for the cross-currency interdependencies of exchange rates and to assess their impact on the exchange rate analysis. Building upon results from individual time series regressions, I conduct seemingly unrelated regression (SUR) analysis in two sets of equations with EUR and USD as base currency, respectively (Zellner (1962)).

There are various justifications for estimating exchange rates in a system of equations as opposed to individual time series regressions. For example, changes in one exchange rate are likely to affect the value of other pairs that share the same base currency. Further, the dynamics of frequently traded currencies are closely related to each other. If exchange rates are analysed individually, interdependencies across currencies are not considered, affecting the efficiency of obtained coefficient estimates.

For example, EUR/GBP and EUR/JPY show large correlation coefficients of 0.51 and 0.47 with EUR/USD, respectively. Correlations between U.S. dollar-based currency pairs are even higher, with coefficients as high as 0.60. To account for the significant relation between currency prices, I set up systems of seemingly unrelated regressions for each base currency that take the form

$$S_t = BX_t + U_t \tag{2.10}$$

where S_t is a $k \times 1$ vector containing the dependent variable spot exchange rate (Δs_t^j) as dependent variable. For example, for euro currency pairs, let $j =$ *USD, CHF, CZK, GBP, HUF, JPY, NOK, PLN, SEK*, then $S_t = (\Delta s_t^{USD}, ..., \Delta s_t^{SEK})$, B is a matrix of coefficients and X_t contains the explanatory variables for the hybrid models and their individual components. I take the following approach. First, I estimate the hybrid model in a SUR framework and extract the variance-covariance matrices of the two systems of equations. Second, I formally assess the co-movement of the exchange rate pairs in each panel by testing if the off-diagonal entries of these matrices are zero. In this case particular case the SUR estimation would lead to identical results as ordinary least square regressions. I test for the diagonality condition by calculating the Breusch and Pagan (1980) test statistic. Under the null hypothesis of diagonality, the test statistic follows a $\chi^2_{(M(M-1)/2)}$ distribution, with $M_{EUR} = 9$ and $M_{USD} = 9$ denoting the number of equations in each system. The 5% critical value is $\chi^2_{36} = 50.998$.

Table 2.9: SUR Model: Cross-Equation Correlations

This table reports the Breusch and Pagan (1980) test statistic $\lambda_{LM} = \frac{1}{T}\sum_{i=2}^{M}\sum_{j=1}^{i-1} r_{ij}^2$ where M refers to the number of equations in the seemingly unrelated regression framework (SUR) and r_{ij}^2 are the sample correlation coefficients, which are calculated based on the sample standard covariances, such that $r_{ij}^2 = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}}$. I estimate a hybrid model of the form $s_t - s_{t-1} = \alpha + \beta_1\tilde{\pi}_{t-1} + \beta_2\tilde{y}_{t-1} + \beta_3\hat{\tilde{i}}_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t$ in a SUR framework for euro and U.S. dollar currencies separately and I assess the hybrid model's nested components by imposing the following restrictions: $\lambda_{LM}^1 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$, $\lambda_{LM}^2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, $\lambda_{LM}^3 : \gamma_1 = \gamma_2 = 0$, $\lambda_{LM}^4 : \gamma_2 = 0$, λ_{LM}^5 : *No restrictions.*

	λ_{LM}^1	λ_{LM}^2	λ_{LM}^3	λ_{LM}^4	λ_{LM}^5
Euro pairs	344.41	345.98	417.55	359.98	360.48
U.S. dollar pairs	1077.55	1037.84	1184.62	1081.6	1037.49

Notes: The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair. The SUR system for euro currencies includes EUR/USD, EUR/CHF, EUR.CZK, EUR/GBP, EUR/HUF, EUR/JPY, EUR/NOK, EUR/PLN, and EUR/SEK. For the U.S. dollar-based SUR estimation, the USD/CAD, USD/CHF, USD/ILS, USD/INR, USD/JPY, USD/MXN, USD/PLN, USD/SGD, and USD/ZAR are included.

As shown in Table 2.9, the obtained statistics range between 344.41 (λ_{LM}^1) and 417.55 (λ_{LM}^3) for euro and between 1037.49 (λ_{LM}^5) and 1184.62 (λ_{LM}^3) for U.S. dollar pairs. This large values suggest that the null hypothesis should be rejected for all five models at the 1% level. The simple test procedure highlights the significant cross-currency correlations and points toward the importance to conduct the exchange rate analysis in a system of equations.

To test if the hybrid model outperforms its nested components after considering the cross-currency correlation, I use McElroy (1977) system-wide goodness of fit measure as an evaluation criterion. I also modify the original measure to account for the total number of parameters estimated in each system. The results for all five models and both systems of equations are displayed in Table 2.10.

For euro pairs, adjusted system-wide R_{SUR}^2 exhibit the highest value for the hybrid model ($R_{SUR}^{2,5} = 0.093$) that incorporate both order flow variables and macroeconomic fundamentals. The model comprising solely macroeconomic regressors, produces an adjusted system-wide R_{SUR}^2 of only 0.005 ($R_{SUR}^{2,3} = 0.093$). Similarly, I find that the best fit for U.S.

Table 2.10: SUR Model: System-wide Explanatory Power

This table reports McElroy's system-wide measure of fit of a seemingly unrelated regression framework (SUR) $R^2_{SUR} = 1 - \frac{M}{tr\Sigma^{-1}S_{yy}}$, where M denotes the number of equation in the system, Σ is the residual cross product and S_yy is the mean deviation cross product matrix. The adjusted $\bar{R}^2_{SUR}$ is calculated as $\bar{R}^2_{SUR} = 1 - \frac{(1-R^2_{SUR})(N-1)}{N-p-1}$ and explicitly accounts for the total number of observations. We estimate a hybrid model of the form $s_t - s_{t-1} = \alpha + \beta_1\tilde{\pi}_{t-1} + \beta_2\tilde{y}_{t-1} + \beta_3 i_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t$ in a SUR framework for euro and U.S. dollar currencies separately and we estimate models based on its nested components by imposing the following restrictions: $R^{2,1}_{SUR} : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$, $R^{2,2}_{SUR} : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, $R^{2,3}_{SUR} : \gamma_1 = \gamma_2 = 0$, $R^{2,4}_{SUR} : \gamma_2 = 0$, $R^{2,5}_{SUR}$: *No restrictions.*

	$R^{2,1}_{SUR}$	$R^{2,2}_{SUR}$	$R^{2,3}_{SUR}$	$R^{2,4}_{SUR}$	$R^{2,5}_{SUR}$
Euro pairs	0.089	0.11	0.047	0.129	0.146
U.S. dollar pairs	0.051	0.1	0.048	0.107	0.155
	$\bar{R}^{2,1}_{SUR}$	$\bar{R}^{2,2}_{SUR}$	$\bar{R}^{2,3}_{SUR}$	$\bar{R}^{2,4}_{SUR}$	$\bar{R}^{2,5}_{SUR}$
Euro pairs	0.073	0.087	0.005	0.084	0.093
U.S. dollar pairs	0.035	0.077	0.006	0.06	0.102

Notes: The sample period is January 2004 to February 2014, comprising 122 monthly observations for each currency pair. The SUR system for euro currencies includes EUR/USD, EUR/CHF, EUR.CZK, EUR/GBP, EUR/HUF, EUR/JPY, EUR/NOK, EUR/PLN, and EUR/SEK. For the U.S. dollar-based SUR estimation, the USD/CAD, USD/CHF, USD/ILS, USD/INR, USD/JPY, USD/MXN, USD/PLN, USD/SGD, and USD/ZAR are included.

dollar exchange rates is obtained with the hybrid model, which accounts for both order flow variables. R^2_{SUR} rises to 0.101 ($R^{2,5}_{SUR}$), clearly outperforming all other models.

Overall, the SUR analysis supports the approach to combine variables from the microstructure and macroeconomic literature results in one and the same model. Moreover, the analysis highlights the significant impact of cross-currency correlations across exchange rates in the FX market, which should be taken into account in an out-of-sample forecasting exercise.

2.5.2 Out-of-sample Forecasting Performance

While the assessment so far points toward the superior explanatory performance of hybrid models over their nested components, in a last step I assess whether hybrid models can be potential tools to forecast exchange rates out-of-sample. To this end, I take the following approach. First, as I provided evidence for significant co-movement across exchange rates, I use the SUR regression framework to conduct a one-step ahead forecast for each individual exchange rate. Second, as the length of the time series is limited by the available order flow data, I also use a fixed effect panel framework to construct a one-step ahead out-of-sample forecast. This approach is motivated by recent studies (e.g. Engel et al. (2008)) that show that the forecasting performance of exchange rate models can be improved by reducing the number of coefficients across currency pairs. For both estimation techniques I use the conventional Taylor rule model as a benchmark and calculate Campbell and Thompson (2008)'s OOS-R2 statistic as an evaluation criterion

$$R^2_{OOS} = 1 - \frac{MSE_{\Delta \bar{s}_{t+1|t}}}{MSE_{\Delta \tilde{s}_{t+1|t}}} = \frac{\sum_{t=M+1}^{T-1} (\Delta s_{t+1} - \Delta \bar{s}_{t+1|t})^2}{\sum_{t=M+1}^{T-1} (\Delta s_{t+1} - \Delta \tilde{s}_{t+1|t})^2} \tag{2.11}$$

where $\bar{s}_{t+1|t}$ refers to the conditional one-step ahead forecast from the hybrid model and $\tilde{s}_{t+1|t}$ is the conditional forecast from the conventional Taylor rule. Positive values indicate a smaller mean square error of the hybrid model. The one-step ahead forecasts based on SUR regressions are constructed using a rolling window of 80 months and the OOS-R^2 statistic for all five models is displayed in Table 2.11.

The OOS-R^2 statistics indicate that the hybrid model that includes only market order flow as an additional regressor produces more accurate forecasts than the macroeconomic model for EUR/CZK, USD/CHF, USD/ILS, USD/JPY, and USD/ZAR. In contrast, net limit order flow does not seem to add much forecasting power. This confirms earlier results that the impact of net limit order flow on the exchange rate path is transitory and appears to vanish completely in an out-of-sample assessment.

Further, Table 2.12 summarizes the OOS-R^2 statistics for the fixed effects panel forecasts. In this estimation setting, I set the size of the rolling window to 40 months to increase the number of one-step ahead forecasts for each currency pair.

As illustrated by the statistics marked in bold, market microstructure models produce smaller mean square errors than the Taylor rule and the hybrid model. For euro pairs, the hybrid model generally performs worse than the macroeconomic Taylor rule, as indicated by the negative $R^{2,3}_{OOS,FE}$ and $R^{2,4}_{OOS,FE}$ For U.S. dollar currency pairs, the performance of the hybrid model is better and $R^{2,3}_{OOS,FE}$ is positive for CHF, ILS, JPY, and PLN. For these

Table 2.11: Out-Of-Sample Forecast: SUR Estimation

This table reports the $OOS-R^2$-statistic $R^2_{OOS} = 1 - \frac{MSE_{\Delta \bar{s}_{t+1|t}}}{MSE_{\Delta \tilde{s}_{t+1|t}}} = \frac{\sum_{t=M+1}^{T-1}(\Delta s_{t+1} - \Delta \bar{s}_{t+1|t})^2}{\sum_{t=M+1}^{T-1}(\Delta s_{t+1} - \Delta \tilde{s}_{t+1|t})^2}$ where $\bar{s}_{t+1|t}$ refers to the conditional one-step ahead forecast from the hybrid model and $\tilde{s}_{t+1|t}$ is the conditional forecast from the conventional Taylor rule. The hybrid model for the out-of-sample exercise is specified as $s_{t+1|t} = \alpha + \beta_1 \tilde{\pi}_{t-1} + \beta_2 \tilde{y}_{t-1} + \beta_3 \tilde{i}_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t$. $\tilde{s}_{t+1|t}$ is obtained from the same regression with $\gamma_1 = \gamma_2 = 0$. I impose the following restrictions to estimate the nested components of the hybrid model: $\bar{R}^{2,1}_{OOS,SUR} : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$, $\bar{R}^{2,2}_{OOS,SUR} : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, $\bar{R}^{2,3}_{OOS,SUR} : \gamma_2 = 0$, $\bar{R}^{2,4}_{OOS,SUR}$: *No restricions.*

	Euro pairs					**U.S. dollar pairs**			
	$\bar{R}^{2,1}_{OOS,SUR}$	$\bar{R}^{2,2}_{OOS,SUR}$	$\bar{R}^{2,3}_{OOS,SUR}$	$\bar{R}^{2,4}_{OOS,SUR}$		$\bar{R}^{2,1}_{OOS,SUR}$	$\bar{R}^{2,2}_{OOS,SUR}$	$\bar{R}^{2,3}_{OOS,SUR}$	$\bar{R}^{2,4}_{OOS,SUR}$
USD	-0.33	-0.34	-0.46	-0.5	CAD	-0.22	-0.34	-0.31	-0.35
CHF	-0.49	-1.5	-0.27	-0.84	CHF	**0.11**	**0.04**	**0.04**	-0.07
CZK	-0.01	-0.14	**0.04**	-0.10	ILS	**0.46**	**0.46**	**0.03**	-0.01
GBP	-1.27	-1.34	-1.36	-1.40	INR	-0.08	-0.04	-0.12	-0.17
HUF	-0.34	-0.48	-0.38	-0.61	JPY	**0.24**	**0.15**	**0.11**	**0.08**
JPY	-0.17	-0.33	-0.29	-0.51	MXN	**0.09**	**0.08**	-0.12	-0.15
NOK	-0.31	-0.31	-0.37	-0.33	PLN	-0.1	-0.05	-0.18	-0.13
PLN	**0.08**	-0.01	-0.21	-0.23	SGD	**0.00**	-0.71	-0.08	-0.89
SEK	-0.13	-0.22	-0.29	-0.41	ZAR	**0.08**	**0.04**	**0.03**	-0.02

Notes: The sample period is January 2004 to February 2014. The SUR system for euro includes EUR/USD, EUR/CHF, EUR.CZK, EUR/GBP, EUR/HUF, EUR/JPY, EUR/NOK, EUR/PLN, and EUR/SEK. For the U.S. dollar-based SUR estimation, the USD/CAD, USD/CHF, USD/ILS, USD/INR, USD/JPY, USD/MXN, USD/PLN, USD/SGD, and USD/ZAR are included.

pairs the combination of macroeconomic fundamentals and order flow data leads to superior out-of-sample forecasts than using only macroeconomic fundamentals. Compared to the in-sample evaluation, the hybrid model's success over the conventional Taylor rule in an out-of-sample forecasting exercise is limited to a few currency pairs. Yet, these cases point toward the possible benefits of combining macroeconomic and transaction flow information in one and the same model. Furthermore, as I do not use real-time data on macroeconomic fundamentals, the information used in this stylized forecasting exercise differs from

Table 2.12: Out-Of-Sample Forecast: Panel Fixed-Effect Estimation

This table reports the $OOS-R^2$-statistic $R^2_{OOS} = 1 - \frac{MSE_{\Delta \bar{s}_{t+1|t}}}{MSE_{\Delta \tilde{s}_{t+1|t}}} = \frac{\sum_{t=M+1}^{T-1}(\Delta s_{t+1} - \Delta \bar{s}_{t+1|t})^2}{\sum_{t=M+1}^{T-1}(\Delta s_{t+1} - \Delta \tilde{s}_{t+1|t})^2}$ where $\bar{s}_{t+1|t}$ refers to the conditional one-step ahead forecast from the hybrid model and $\tilde{s}_{t+1|t}$ is the conditional forecast from the conventional Taylor rule. The hybrid model for the out-of-sample exercise is specified as $s_{t+1|t} = \alpha + \beta_1 \tilde{\pi}_{t-1} + \beta_2 \tilde{y}_{t-1} + \beta_3 \tilde{i}_{t-2} + \beta q_{t-1} + \gamma_1 mo_t + \gamma_2 lo_t + u_t$. $\tilde{s}_{t+1|t}$ is obtained from the same regression with $\gamma_1 = \gamma_2 = 0$. I impose the following restrictions to estimate the nested components of the hybrid model: $\bar{R}^{2,1}_{OOS,FE}$: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_2 = 0$, $\bar{R}^{2,2}_{OOS,FE}$: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, $\bar{R}^{2,3}_{OOS,FE}$: $\gamma_2 = 0$, $\bar{R}^{2,4}_{OOS,FE}$: *No restrictions.*

	Euro pairs					**U.S. dollar pairs**			
	$\bar{R}^{2,1}_{OOS,FE}$	$\bar{R}^{2,2}_{OOS,FE}$	$\bar{R}^{2,3}_{OOS,FE}$	$\bar{R}^{2,4}_{OOS,FE}$		$\bar{R}^{2,1}_{OOS,FE}$	$\bar{R}^{2,2}_{OOS,FE}$	$\bar{R}^{2,3}_{OOS,FE}$	$\bar{R}^{2,4}_{OOS,FE}$
CHF	**0.152**	**0.146**	-0.249	-0.28	CAD	-0.737	-0.987	-0.681	-0.935
CZK	**0.41**	**0.378**	-0.18	-0.158	CHF	-0.005	**0.004**	**0.002**	**0.003**
GBP	-0.062	-0.221	-0.384	-0.436	ILS	**0.039**	-0.037	**0.002**	-0.007
HUF	**0.061**	**0.063**	-0.192	-0.149	INR	-0.189	-0.151	-0.141	-0.096
JPY	**0.082**	**0.058**	-0.13	-0.175	JPY	-0.056	-0.017	**0.001**	**0.048**
NOK	**0.049**	-0.154	-0.451	-0.674	MXN	-0.134	-0.28	-0.137	-0.4
PLN	-0.149	-0.276	-0.15	-0.295	PLN	**0.017**	**0.041**	**0.001**	**0.014**
SEK	**0.11**	**0.015**	-0.026	-0.216	SGD	-0.323	-0.346	-0.215	-0.265
					ZAR	-0.125	-0.539	-0.129	-0.556

Notes: The sample period is January 2004 to February 2014. The SUR system for euro includes EUR/USD, EUR/CHF, EUR.CZK, EUR/GBP, EUR/HUF, EUR/JPY, EUR/NOK, EUR/PLN, and EUR/SEK. For the U.S. dollar-based SUR estimation, the USD/CAD, USD/CHF, USD/ILS, USD/INR, USD/JPY, USD/MXN, USD/PLN, USD/SGD, and USD/ZAR are included.

the information available to market participants. This difference may not only affect the performance of the conventional Taylor rule, but also the forecast precision of hybrid models that explicitly account for the interactions between transaction flow measures and macroeconomic fundamentals. Another caveat of this approach is the small number of available observations in the forecasting exercise. The length of the time series is determined by the data availability of order flow variables. Once longer time series of transaction flow data become available, the forecasting performance of the hybrid models can be assessed in greater detail and over longer forecasting horizons. I leave this exercise to future research.

2.6 Conclusion

This paper introduces a new class of hybrid exchange rate models that combines foreign exchange dealer information with macroeconomic fundamentals embedded in a conventional Taylor rule. I argue that the link between macroeconomic models and interdealer trading dynamics can be established by interpreting transaction flow variables as a proxy for dealers' risk premiums and assuming that the uncovered interest rate parity condition does not hold. Equipped with a new class of hybrid models, I conduct an extensive empirical assessment and compare their performance with conventional market microstructure models and a macroeconomic Taylor rule. The findings can be summarized as follows.

First, I document a higher goodness-of-fit measure and superior in-sample predictive power of hybrid models compared to conventional market microstructure and macroeconomic models. I show that the adjusted goodness of fit measure increases up to 0.32 for euro and 0.42 for U.S. dollar exchange rates. For currencies traded against the U.S. dollar, the hybrid model generally shows a strong performance, while for currencies traded against the euro the hybrid model performs best for the most frequently traded currencies, such as EUR/USD and EUR/JPY. For more than half of the sample, the hybrid model is the preferred modelling choice over nested market microstructure models, as indicated by various F-tests and other model selection criteria. These findings are robust to different order flow specifications that explicitly consider the degree of dealer activity in the foreign exchange market.

Second, I show that a new microstructure measure, based purely on submitted and cancelled limit orders at the top of the order book, is a significant driver of monthly exchange rates. This measure has a large and significant impact for the most liquid currency pairs (EUR/USD, EUR/JPY, GBP/USD, USD/JPY) but also for a few emerging market currencies (e.g. USD/SGD). The estimated magnitude of this new measure, called net limit order flow, is lower than that of market orders and appears to be more transitory.

Lastly, I highlight the importance of cross-currency interdependencies and analyse a large cross-section of exchange rates in a seemingly unrelated regression framework. I use an augmented system-wide measure to account for the number of parameters and compare the hybrid model with its nested individual model components. For both base currencies, euro and U.S. dollar, the hybrid model explains the largest proportion of the system-wide variation. Further, I conduct a stylized one-month ahead out-of-sample forecasting exercise in a seemingly unrelated regression and panel fixed effects regression set-up. For U.S. dollar pairs, forecasts from both models outperform the conventional Taylor rule in several cases, while the most accurate forecasts for euro pairs are produced by market microstructure models.

Overall, the empirical evidence I present points toward the advantages of combining market microstructure approaches to analyse exchange rate dynamics. As more transaction flow data become available, future research can build upon these findings and focus further on the out-of-sample performance of hybrid exchange rate models.

Chapter 3

Performance, Persistence, and Pay: A New Perspective on CTAs

3.1 Introduction

A growing academic literature examines why investors continue to allocate their capital to seemingly unsuccessful active managers. While numerous studies focus on the performance of actively managed mutual funds (e.g. Gruber (1996); Cremers and Petajisto (2009); Berk and Green (2004)) and hedge funds (e.g. Ackermann et al. (1999); Agarwal et al. (2015a); Stulz (2007)), performance of commodity trading advisors (CTAs) has received less attention.[1] That said, the broad consensus emerging from extant studies is that the average CTA does not create value for its investors (e.g. Elton et al. (1987), Elton et al. (1989), Elton et al. (1990), Bhardwaj et al. (2014)). Yet, as indicated by their rapidly growing assets under management (AUM) from USD 24.9 billion to USD 339.7 billion between 1994 and 2016,[2] CTAs have become a popular investment vehicle for practitioners and a

[1] They are often excluded from hedge fund studies, such as Bollen (2013), Agarwal et al. (2009) or Titman and Tiu (2011).

[2] Information on the industry's AUM refers to Barclay's yearly estimates of the industry's overall assets under management. Accessed via `https://www.barclayhedge.com/research/indices/cta/mum/CTA_Fund_Industry.html` and `https://www.barclayhedge.com/research/indices/ghs/mum/`

fundamental component of today's financial markets.[3] I offer a new perspective on this puzzle. More specifically, I employ one of the largest and cleanest CTA datasets explored so far to analyse the performance of CTAs, discuss the cross-sectional variations within the category, assess CTA manager skill and performance persistence, and examine whether managerial compensation is justified by managerial performance.

Apart from their ever-growing presence, CTAs are also unique in that even while they are one of the more populous categories of alternative investments, their investment strategies are relatively undiversified and identifiable, making it is easier to benchmark and evaluate their performance (Fung and Hsieh (2001)). Such a unique advantage in modelling returns not only results in an accurate estimation of CTA performance but also enables me to circumvent the opaqueness of hedge fund investments and gain valuable insights into the operational efficiency of the alternative investments universe.

I find that CTA managers generate economic and statistically significant positive net excess returns. Further, I document that pre-fees alphas are positive and significant, indicating that CTA managers beat passive benchmark strategies. Following the rationale of Berk and van Binsbergen (2015), I also provide evidence that CTA managers add significant value to their customers and that CTA performances are persistent for a horizon of up to three years. Finally, I show that CTAs' compensation scheme predicts future performance, providing managers with an avenue to signal their skill to investors. While previous papers have invoked investor irrationality or severe information asymmetry to reconcile their findings with the continued growth of CTAs, the results of this study are indicative of a well-functioning, competitive marketplace with rational investors and fund managers. In fact, the results are perfectly in accordance with the main predictions of the Berk and Green (2004) model. One, there is significant and persistent cross-sectional variation in manager skill. Two, investors compete to invest with successful managers. Finally, managerial com-

---

`HF_Money_Under_Management.html`. In terms of AUM, CTAs became the third-largest hedge fund category in 2016, after Fixed Income (USD 556.2 billion) and Multi-Strategy (USD 360 billion) hedge funds.

[3]The growing popularity in the CTA industry, its increasing AUM, and associated risk factors are also evident in the recently increasing number of financial newspaper articles, for example, "Trend is your friend, say investors flocking to futures", (`https://www.ft.com/content/e367ca58-e72d-11e4-a01c-00144feab7de`) "Computer-driven trend hedge funds thrive despite falling markets," (`https://www.ft.com/content/dc33992c-beca-11e5-846f-79b0e3d20eaf`) or "Risk in new form of "portfolio insurance" sparks fear." (`https://www.ft.com/content/3eba3f56-08c6-11e7-97d1-5e720a26771b`)

pensation, functioning as a balancing mechanism, is set so that the ex-ante net alpha is zero.

The analysis is based on data derived from Barclay's Hedge Fund Database (BarclayHedge). The database covers on average 70% of the industry, in terms of AUM. Risk and Portfolio Management AB (RPM), a managed account industry specialist based in Stockholm, Sweden, provides me with the data. Since RPM has been downloading the entire BarclayHedge database daily since 2002, data entries are not rewritten and no return histories are deleted (Patton et al. (2015)). The dataset is therefore largely free of graveyard-bias and captures the wide cross-sectional dynamics of alive as well as defunct funds that stopped reporting during the sample period. First, equipped with this rich dataset, I construct equally and value-weighted portfolios of CTAs and show that funds generate on average 4.1% and 4.5% annualized net excess returns. These returns are net of all fees and are delivered to investors. Furthermore, I make use of an additional, small but representative proprietary dataset, provided by RPM, which contains solely realized returns and validates my findings. I show that portfolios constructed from BarclayHedge and from the proprietary dataset have the same distributional characteristics, highlighting the accuracy of the findings and verifying the economic and statistically significant performance of CTAs.

Second, to identify cross-sectional variations among CTAs, I use a novel trading strategy classification obtained from RPM. The classification is based on RPM's private information about a fund and on the fund's own trading strategy description, which it reports to BarclayHedge. All funds that start to report to BarclayHedge are categorized according to one of the four strategy groups: systematic and discretionary trend followers and their non-trend following counterpoints. In contrast to classifications that are available from commercial hedge fund databases, the fund categorization in this study allows me to distinguish between funds that use trend-following trading strategies (trend follower) and those that use a different trading approach (non-trend follower), for example, short-term or fundamental traders. The analysis shows that the differentiation between these groups is crucial and that return dynamics across these trading strategies are fundamentally different from each other. For example, systematic and discretionary trend followers generate 6.0% and 7.4% average annualized returns, compared to only 3.5% and 1.8% by non-trend following CTAS.

Third, I discuss additional attractive characteristics of CTA returns, such as their positive skewness and correlation with other assets classes. Approximately 64% of the funds have positively skewed returns, indicating that CTAs might be attractive to investors with preferences for skewed returns (Polkovnichenko (2005), Brunnermeier and Pedersen (2009),Mitton and Vorkink (2007)). Furthermore, I show that CTA returns move strongly counter cyclically to equity markets. In times of equity market turmoil (SP500 average: -10.1%), CTAs average monthly excess return is at least 1.7%; and when equity markets flourish (SP500 average: 8.4%), CTAs average excess return is -0.8%. Further, in extreme events, CTAs' return generating process is almost entirely uncorrelated with those of hedge funds—during months with the highest and lowest returns of the hedge fund research index, trend-following CTAs constantly produce returns between 2.4% and 3.5%. Even though CTAs are often classified as a subcategory of hedge funds, similar to Liang (2004), the analysis emphasizes substantial differences between these two asset classes and the possible diversification benefits from including CTAs in an investor's portfolio that cannot be obtained by investing in other active investment vehicles, such as hedge funds.

Fourth, I find that CTAs generate abnormal gross returns over and above benchmark trading strategies such as time series momentum (Moskowitz et al. (2012)) and option straddle factors (Fung and Hsieh (2001)). For the equally weighted and value-weighted CTA portfolios, the gross alpha is 8.4% and 6.4% on an annual basis, respectively. Furthermore, I document that CTAs, especially systematic trend followers, exploit price movements during periods of market turmoil. During these times, they produce an annual gross alpha of 27.4%, indicating that they successfully exploit price trends during crisis periods.

Next, I employ a recent approach by Berk and van Binsbergen (2015) to analyse the value added by CTA managers. In line with Berk and Green (2004) and Berk (2005), the authors argue that the amount of capital funds attract from investors is a better measure of managerial skill than the pre-fees regression alpha. Since the profitability of a fund depends on the return as well as the amount of capital managed by a CTA, the authors construct a proxy for a fund's added value that takes both dimensions into account. I find large cross-sectional variation in managerial skill. Specifically, I document that 41% of CTAs in the sample generate negative value, compared to standard time series momentum strategies. Moreover, I find that the average added value of a CTA is USD 0.49 million per month.

Finally, using the valued added measure I provide evidence that CTA performance is persistent for up to three years. Sorting funds into quintile portfolios, the top 20% of CTAs significantly outperforms the bottom 20% over various forecasting horizons. In line with the argument that funds with greater investment skills demand higher compensations, I also find that the costliest investments in CTAs outperform funds that demand less compensation. This analysis shows that funds with higher accruing fees are more successful and that managers can use their compensation as credible signal of their skill. These findings are indicative of an efficient CTA market.

The paper is most closely related to the existing literature examining the return dynamics of CTAs, most notably to Bhardwaj et al. (2014). In contrast to their results, my findings suggest that CTAs generate significant excess net returns to investors, that these net returns are positively skewed, and that CTAs generate significant pre-fees alpha, especially during periods of equity market stress. A likely explanation for the difference in results is that the analysis is based on a substantially larger set of CTAs, allowing for a wider representation of market dynamics. It covers on average 70% of the industry in terms of AUM, which is more than three times larger than the 21% industry coverage in Bhardwaj et al. (2014).

In addition, this study is the first to provide insight into the heterogeneity of CTA trading strategies, to show that there is significant and persistent cross-sectional variation in the skills of CTA managers, and that CTA manager pay is commensurate with manager performance. In this respect, this paper is also closely related to Berk and van Binsbergen (2015), who similarly examine managerial skill in the mutual fund universe. While I focus solely on CTAs, the results also contribute to the larger debate on the rationality of investors who place money with fund managers. For example, Griffin and Xu (2009), find no evidence for constant significant positive hedge fund alphas. In contrast, Agarwal and Naik (2000), Agarwal et al. (2009), and Ibbotson and Chen (2006) argue in favour of the hypothesis that hedge fund managers are skilled and generate abnormal returns beyond standard beta-risk factors. This paper's findings are consistent with this view that fund managers exhibit significant and persistent skill, and that "being able to pick good hedge funds can therefore be highly rewarding" (Stulz (2007)).

Finally, while Brown and Goetzmann (2003), Kazemi and Li (2009), and Gregoriou et al. (2010) use fund classifications available in commercial databases to identify performance differences among CTAs and hedge funds, I use a novel classification system to explicitly distinguish between trend- and non-trend-following CTAs. As I highlight in various exercises, their trading strategies and performances are substantially different from each other. In contrast to Arnold et al. (2013), who also distinguishes trend-followers and other CTA trading strategies, I do not analyse factors that determine the survival of funds but rather examine performance differences between these trading strategies. Earlier papers that have used the same fund classification (Elaut and Erdös (2016)) focus on only one trading strategy, but do not compare performance differences among CTAs.

The rest of this paper proceeds as follows. In the next section I introduce the datasets I use for the analysis and describe in detail the steps of data cleaning taken to alleviate the impact of possible biases. In section 3, I discuss CTA performance as well as dynamics of net-of-fee returns. Section 4 assesses the managerial skill of CTA managers and the persistence of CTA returns. Section 5 concludes.

3.2 Data

The main underlying database for the analysis is Barclay's Hedge Fund Database (BarclayHedge). Risk and Portfolio Management AB (RPM), a managed account industry specialist based in Stockholm, Sweden, provided me with the data. BarclayHedge is the single most comprehensive database for CTAs. Compared to other commercially available hedge fund databases, it has a low proportion of missing information and large coverage of defunct funds, which have stopped reporting to the data provider (Joenväärä et al. (2016)). For the analysis, I focus on funds' flagship programs, which refer to a fund's longest track record and highest assets under management. This leaves me with 3,017 individual CTAs and 208,959 fund-time observations for the period 1985 to December 2015. In order to allow for a comparison between these results and the previous literature, I follow the same cleaning procedure outlined in Bhardwaj et al. (2014). Table 3.1 summarizes each step and its impact on the dataset.

Since most commercial hedge fund databases begin to keep a track record of defunct funds in 1994, I restrict the analysis to the post-January 1994 period and drop returns associated with earlier reporting dates. This should reduce the impact of a potential survivorship bias in the database (Elaut et al. (2016)). Further, I only consider funds that report information denominated in U.S. dollars and exclude the records of 174 CTAs that

Table 3.1: Data Cleaning Steps

This table summarizes data cleaning steps of Barclay's Hedge Fund Database as of 15/03/2016. The number of funds for an equally weighted (EW) portfolio (no minimum reporting time) is 2,620. The value-weighted (VW) portfolio consists of 1,924 individual CTA flagship programs. AUM refers to assets under management

Data screening steps	# Funds removed	# Funds remaining
Starting Sample		3017
Stopped reporting before 1994	164	2853
Not reporting in USD	174	2679
Missing date of entry to database	9	2670
Not reporting "net all fees"	36	2634
Unrealistic return	1	2633
Funds created in 2016	13	2620
Funds with missing AUM	696	1924

use a different base currency. I also delete nine funds, for which I cannot identify an exact reporting start date, 36 entries that do not report returns "net all fees" and one entire fund history that reports unrealistic returns, such as −99.99%. Also, to allow for more than two months' reporting delay, I do not include funds added to BarclayHedge after December 2015.[4]. Lastly, to be able to construct a value-weighted index, I delete CTAs that do not report assets under management (AUM) for the first or last observation. For missing AUM observations within a fund's record, I proximate the AUM by linear interpolation between the first and last available non-zero entry.[5]

After applying these filters, I am left with a sample of 2,620 funds and 195,682 cross-sectional observations to construct an equally weighted (EW) portfolio CTA index. The value-weighted (VW) index is based on a cross-section of 1924 CTAs and 131,485. In terms of size, the underlying data for the analysis consists of approximately three times as many CTA flagship programs as previous studies on CTA performance. In terms of AUM, I cover on average 70% of the CTA industry over the entire sample period, which is

[4] I obtained the database from RPM in March 2016

[5] This approach closely follows Bhardwaj et al. (2014) even though the authors only delete funds with missing information about AUM for the first reported observation.

significantly larger than the industry coverage of 21% in Bhardwaj et al. (2014).[6]

3.2.1 Biases In Commercial Hedge Fund Databases

It is well documented in the academic literature (for a recent survey see e.g. Agarwal et al. (2015b)) that commercial hedge fund databases are subject to various biases. Concerning CTAs in particular, Fung and Hsieh (1997) find that the average annual return of surviving funds is 3.4% higher than the average annual return of their total sample of 901 CTAs in the period 1986–1996. Bhardwaj et al. (2014) show that EW and VW indices that include solely surviving funds generate 4.15% and 2.21% higher average annualized returns than portfolios of both alive and defunct funds. Including defunct funds in the analysis, therefore, takes into account the fact that worse performing funds may stop reporting and drop out of the database. In the sample of this study, the performance of EW and VW indices would be artificially inflated by 2.5% and 1.2%, respectively, if I considered only the 507 funds still alive at the end of the sample period and omit those that dropped out over time.

In addition to survivorship bias, I account for funds' tendency to report returns retrospectively after they have entered the database. This practice is termed "backfill bias" (Gregoriou et al. (2010)). Since CTAs use commercial databases to market their performance to investors, backfilled returns can lead to an artificial upward bias of the return structure. A common approach in the literature has been to exclude the first 12–24 months of the analysis to account for possibly retrospective reported return structures. However, as Bhardwaj et al. (2014) point out, a generic screen of the first "x-month" of reported returns does not clean the data sufficiently. They find that funds backfill on average 31 months in their sample. Instead of discarding a fixed number of first few months of each fund, the authors recommend using the fund's reporting start date as indicator and to exclude all reported returns prior to this date from the analysis.

In my version of BarclayHedge, I can follow the authors' suggested practice for most of the sample period and delete a fund's entire history prior to its entry in the database. I can infer the start of a fund's report history in BarclayHedge since RPM has downloaded the entire databases daily since February 2002 and flags the first entry of a fund to the

[6] I use BarclayHedge's estimate of the CTA industry size as benchmark. The annual data of the estimated industry size are accessible via: `https://www.barclayhedge.com/research/indices/cta/mum/CTA_Fund_Industry.html`.

database. I use this flag to minimize a potential upward bias in my analysis, caused by backfilled returns. For the first eight years, January 1994–January 2002, for which the reporting start date cannot be pinned down, I take a conservative approach and delete the first 36 reported months of a fund's track record.

Further, funds may revise their reports ex-post or even ask database vendors to delete the entire performance records after a fund stops reporting to the database (Patton et al. (2015)). If a fund has performed poorly in the past, it might have a greater incentive to delete its history, leading to an upward bias among defunct CTAs. Since the data for this study has been downloaded and stored daily by RPM, this BarclayHedge version is largely free of this "graveyard" bias.

3.2.2 CTA Trading Strategies

To understand and assess performance differences among CTAs, I supplement information on return dynamics from BarclayHedge with a trading strategy classification, which is obtained from RPM and allows me to distinguish between trend- and non-trend-following CTAs. Funds that enter BarclayHedge are categorized on a weekly basis into one of the categories shown in 3.2. The classification is based on comparing funds' past return dynamics with trend-following benchmarks, on the fund's own strategy description, on conversations with CTA managers and, if known, on observing funds' portfolio re-balancing and investment behaviour across different asset classes. Usually a fund is classified to one of the four categories only once when a fund starts to report its performance to BarclayHedge.[7]

Table 3.2 shows three different levels of classification. As shown in column (1) funds can be identified as discretionary or systematic trading CTAs. Systematic traders are characterized by their use of algorithmic trading models and an extensive quantitative analysis of financial data that forms the basis for funds' investment decisions. In contrast, for discretionary strategies managers' ability to exploit chart patterns or divine global supply/demand imbalances from fundamental data plays a much more fundamental role. Column (2) distinguishes between trend-following funds and non-trend followers. Trend-following funds take directional long and short positions in various asset classes and generate returns by exploiting persistent price trends (Kaminsky (2011)). In contrast, I consider non-trend followers as

[7]Based on anecdotal evidence, we assume that the strategy of funds do not change drastically over the course of their reporting time.

Table 3.2: Trading Classification

This tables shows the different strategy classifications provided by RPM. The different dimensions are divided into columns (1) and (2). For example, funds can be classified as either systematic or discretionary. Further, these classes can be differentiated between trend and non-trend follower. Funds that cannot be classified are grouped into "Others."

(1)	(2)
Systematic	Trend follower
	Non-trend follower
Discretionary	Trend follower
	Non-trend follower
Others	

fundamental, short-term, commodity and FX traders. This classification is a novel feature of this study, since I can distinguish between the following strategy classifications, which are not available in any commercially available hedge fund database: systematic trend follower, systematic non-trend follower, discretionary trend follower, and discretionary non-trend follower.[8] However, as my analysis shows, it is crucial to account for the heterogeneity among systematic and discretionary funds, since their return dynamics are fundamentally different from each other. Using RPM's strategy classification, I aim to reduce any "strategic self-misclassification" (Brown and Goetzmann (2015), p. 103) that may result from purely self-reported strategies.

3.2.3 Summary Statistics

For the analysis, I focus on CTAs with 24 months' reported information, which is a sufficiently long return history that is indicative of real return dynamics (Bhardwaj et al. (2014)). Table 3 summarizes the characteristics of the dataset.

As shown in Table 3.3, the EW and VW indices consist of 1,274 and 936 CTAs that report at least 24 months of returns. Two-thirds of these funds are systematic traders, while only 317 funds are categorized as discretionary. Less than 10% belong to the category

[8]While Elaut and Erdös (2016) use the same classification to analyze systematic trend followers, the aim of this paper is to provide an understanding of the overall industry dynamics and to show differences across all trading strategies.

Table 3.3: Summary Statistics

This table provides summary statistics of the composition of funds with at least twenty four months of reported return history. Average size is measured in million USD. Average age is measured in months. Funds are considered as "alive" if they report their information to BarclayHedge at the end of the sample period (December 2015). Information on individual trading strategies refers to funds that are considered for the strategies' EW index. Numbers in brackets refer to the proportion of funds in the overall sample.

			Systematic		Discretionary		
	EW Index	VW Index	Trend	Non-Trend	Trend	Non-Trend	Others
# of Funds	1274	926	487 (38%)	355 (28%)	29 (2%)	288 (23%)	115 (9%)
Avg. Size	234	USD 259	280	380	45	68	55
Avg. Age	70.8	69.2	81.2	64	62.7	65.4	63.6
Alive Funds	323	296	136	96	4	59	28

"Others." The average size of a CTA accounts for USD 234 million, measured by AUM of the last reported observation. However, there is a large variation in fund size across trading strategies. Systematic funds with an average size of USD 280 or USD 380 million assets under management for trend and non-trend followers are substantially larger than discretionary funds. Further, the long-lived CTAs with an average reporting time of 81.2 months tend to be systematic trend followers. The remaining sample average is approximately 64 months. Lastly, as indicated by the final row, most funds at the end of the sample are systematic funds.

3.3 CTA Performance

To evaluate the performance of CTAs, I start by examining the characteristics of funds' net excess returns—net returns in excess of the 3-Month Treasury Bill. To begin with, panel A of Table 3.4 shows the annualized average net excess return and volatility for the EW and VW indices for the period January 1994—December 2015. Over the entire sample period, the average annualized return accounts for 4.1% and 4.5% for the EW and VW index, respectively. Strikingly, both portfolios generate returns that are significantly different from zero at the 1% level, as indicated by the high t-statistics. The results are fundamentally different from earlier studies arguing that CTAs do not produce positive returns to investors. For example, Bhardwaj et al. (2014) find net excess returns are used up entirely by funds' high fee structure. Using a substantially larger cross-section of funds, representing on average 70% of the total CTA industry in terms of AUM, I show that CTAs' net-of-fee returns are economic and statistically significant. CTAs' profitability might be one simple explanation for the growing assets under management in the industry.

Table 3.4: CTA Performance

This table shows the performance analysis for EW and VW portfolios of funds with at least 24 reported return observations. Panels B and C display the results for the individual trading strategy, when funds are weighted on an equal or value basis, respectively. The column "T-Test" refers to the t-statistics of the null hypothesis that the average return equals zero. ***, **, * denote level of significance at the 1%, 5%, and 10% level, respectively.

	Jan 1994 - Dec 2015		
	Avg. Ann. Return	Avg. Ann. Volatility	T-Test
Panel A: All CTAs			
EW Index	4.1%	7.2%	2.65***
VW Index	4.5%	7.6%	2.81***
Panel B: By Trading Strategy (EW)			
Systematic Trend	5.1%	11.7%	2.03**
Systematic Non-Trend	3.1%	4.3%	3.37***
Discretionary Trend	4.7%	15.0%	1.46
Discretionary Non-Trend	2.8%	4.3%	3.10***
Panel C: By Trading Strategy (VW)			
Systematic Trend	6.0%	11.7%	2.41**
Systematic Non-Trend	3.5%	5.8%	2.84***
Discretionary Trend	7.4%	15.7%	2.21***
Discretionary Non-Trend	1.8%	6.5%	1.28

I also find CTAs' positive performance is largely driven by systematic traders, who generate significant positive returns of 5.1% and 3.1% for trend and non-trend followers, respectively. In contrast, the performance of discretionary funds is not necessarily significantly different from zero. Also, even though trend-following funds appear to generate higher returns, these benefits are associated with higher levels of risk. While the annualized average volatility of VW systematic and discretionary non-trend-following portfolios is 5.8% and 6.5%, respectively, it increases to 11.7% and 15.7% for trend-following counterparts. Even though various existing biases in all commercial databases have been identified by academic research, an issue for all studies so far has been that no source of validation is available to verify the process of data cleaning and analysis results. In this study, I am able to alleviate this major shortcoming by using a proprietary dataset of realized CTA returns as a validation mechanism. The data are provided by RPM and are based on realized returns from a set of 51 representative managers that report directly to RPM. While the cross-section of this dataset is smaller than the BarclayHedge coverage, it is worth highlighting that the returns from this database are realized rather than reported returns. Importantly, this implies that these data do not suffer from backfill or graveyard bias, or any form of retrospective window-dressing. Furthermore, since the set of CTAs has been actively managed by RPM, funds have been added to and dropped from the database. Therefore, the set of funds also consists of alive and defunct funds, circumventing concerns about survivorship bias. Even though the number of funds is small, the return dynamics are a representative sample of the overall CTA industry. For example, the correlation between a value-weighted index of the benchmark returns and BarclayHedge's CTA index is 0.83.

To alleviate concerns about remaining or undetected biases in the dataset, I compare the return dynamics of the EW and VW CTA portfolios from BarclayHedge with EW and VW indices based on realized returns from RPM's proprietary dataset. I conduct a t-test to assess if the indices based on reported return and realized return data are on average significantly different from each other. I postulate that if the results were driven by data biases or inadequate data cleaning, I would reject the null hypothesis that the reported return and realized return data have the same return dynamics. Also, I conduct the Kolmogorov-Smirnov (KS) test to check if the distribution of returns between the indices is significantly different. Failing to reject the null hypothesis, however, implies that the dataset of realized returns is representative of the overall industry, strengthening my line of argument. The results of these assessments are shown in Table 3.5.

To start with, Table 3.5 shows the average annual return of BarclayHedge and the set of funds that I use for validating my results. While the difference between indices is

Table 3.5: Benchmark Comparison

This table compares CTA returns from an equally- (EW) and value- (VW) weighted portfolio reported to BarclayHedge with EW and VW portfolios based on realized return data, provided by RPM. The reported t-statistic (”T-test”) refers to the test if the difference between the indices is, on average, different from zero. The column “KS-Test” refers to the Kolmogorov-Smirnov statistic, testing if both samples of funds are drawn from the same distribution. Note: EW index only covers the period April 2002–December 2015.

	Avg. Ann. Return				
	BarclayHedge	Realized Returns	Correlation	T-Test	KS-Test
EW Index	4.10%	2.70%	0.80	0.50	0.11
VW Index	4.50%	3.80%	0.82	0.30	0.09

slightly larger for the EW portfolios, it only accounts for 0.7% on an annual basis. Despite the performance differences, both indices largely follow the same dynamics. The correlation coefficient between EW and VW indices is 0.80 and 0.82, respectively. I interpret these values as a first indication that indices constructed from the proprietary data can be considered as a representation of the overall industry dynamics. Further, in column (5) I test the null hypothesis that both indices generate the same average return and in column (6) I test the null hypothesis that both return series are drawn from the same distribution.

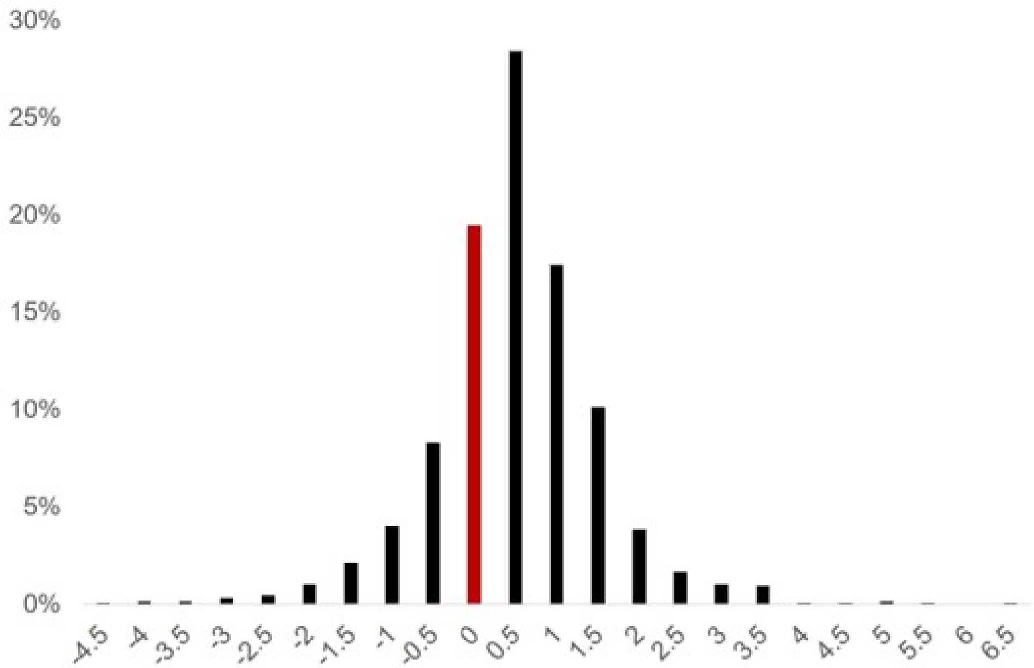
As shown in Table 3.5, columns (5) and (6), I am not able to reject the null hypothesis for either of the two tests. The t-statistics for the differences in mean returns are only 0.5 and 0.3 for the EW and VW index, respectively. Similarly, I am not able to reject the null hypothesis that returns are drawn for the same distribution, as seen from the small KS-statistics of 0.11 and 0.09 for the EW and VW index, respectively.

These results are crucial for this study as well as for papers examining the performance of hedge funds in general. First, they validate the steps of data cleaning, described in the previous section. They show that survivorship and backfill bias are the main forms of biases and that their impact can be significantly alleviated by including all defunct funds from the analysis and by deleting the entire return history prior to the first reporting date. Moreover, not being able to reject the null hypotheses suggests that the findings are not driven by artificially inflated return dynamics, but that they reflect accurately the level of profits generated by CTAs. This validation exercise provides further evidence that CTAs generate significantly positive net excess returns. Furthermore, the low values of the KS-test confirm the representative status of indices based on realized return data.

3.3.1 Characteristics Of CTA Returns

In this section, I analyse additional return characteristics that may further explain the growing popularity of CTAs among investors. I begin by assessing the skewness of returns at the individual fund level. Figure 3.1 shows the distribution of skewness for each fund's returns, where the red bar denotes funds whose returns have a skewness of zero.

Figure 3.1: Fund-level Skewness



Notes: This figure shows the skewness of net excess returns at the individual fund level. The x-axis refers to the return skewness of each fund and the y-axis measures the number of funds in each group relative to the entire sample. The red bar marks funds which returns have a skewness of zero.

As indicated by Figure 3.1, approximately 64% of funds have returns with positive skewness. In fact, for most funds the return skewness is 0.5. The maximum fund-level skewness is 6.18, resulting in a stretched right tail of the distribution. The mean and median are 0.27 and 0.25, respectively, highlighting the positively skewed distribution of returns at the fund level. The descriptive analysis suggests that investors, who prefer a larger upside risk and or have preferences for skewed returns, may allocate some of their capital to CTAs.

I confirm that CTAs serve as an alternative investment opportunity because they generate positive returns during times when equity markets perform particularly poorly. While this has been generally shown by previous studies (Kazemi and Li (2009); Bhardwaj et al. (2014)), in my analysis I contribute to the literature by assessing how CTAs perform in comparison to hedge funds and by pointing out performance differences across trading strategies.[9] Table 3.6 shows the monthly average excess return for the two CTA indices, the S&P 500 as proxy for equity markets and the Hedge Funds Research Index (HFRI).

As shown in panel A, CTAs generate average monthly net excess returns of 1.7% and 1.8% in bear markets when returns from equities are performing particularly poorly. In the worst 5% months of the S&P 500, its average monthly return accounts for -10.1% and hedge funds generate negative returns of -3.5%. The latter can be explained by the investment focus of most hedge funds on long-equity driven strategies. Conversely, during equity bull markets when the S&P 500 shows positive returns of 8.4%, CTA returns are negative. The same countercyclical dynamics appear when I assess the 5% best or worse months of the EW and VW indices in panels B and C, respectively.

In panel D, I depart from the existing literature and examine the tail correlation of CTA and hedge fund returns. Since CTAs are often considered a sub-category of hedge funds, I analyse the extent to which these two active investment classes show similar return dynamics during extreme market events. Interestingly, panel D clearly highlights how the timing of the return generating process of CTAs is fundamentally different during extreme events. The countercyclical correlation that I observe with equity markets does not exist. During the best and worst 5% months of the HFRI, returns of CTAs are essentially identical. While the HFRI index swings between -4.1% and 4.2%, the VW–CTA index generates 1.8% in both periods. This analysis shows that in extreme events, the two asset classes are largely uncorrelated with each other and indicates that the return generating process of CTAs cannot be replicated by either equity markets or hedge funds.

[9] I use Hedge Fund Research's value-weighted hedge fund index (HFRI) as a proxy for hedge fund returns. The data are obtained via Datastream.

Table 3.6: CTA Performance: Bull and Bear Markets

This table shows average monthly excess returns for the best and worst 5% months (13 months each) of the equally weighted (EW) and value-weighted (VW) CTA portfolio, S&P 500 value weighted index (excl. dividends) and HFRI Index (Hedge fund research index).

Panel A: S&P 500				
	Worst 5% S&P 500 months			
	CTA EW	CTA VW	S&P 500	HFRI
Monthly Average ER	1.7%	1.8%	-10.1%	-3.5%
	Best 5% S&P 500 months			
	CTA EW	CTA VW	S&P 500	HFRI
Monthly Average ER	-0.8%	-0.7%	8.4%	2.1%
Panel B: EW CTA Index				
	Worst 5% EW CTA Index months			
	CTA EW	CTA VW	S&P 500	HFRI
Monthly Average ER	-3.6%	-3.5%	2.1%	0.4%
	Best 5% EW CTA Index months			
	CTA EW	CTA VW	S&P 500	HFRI
Monthly Average ER	5.4%	5.0%	-2.1%	0.0%
Panel C: VW CTA Index				
	Worst 5% VW CTA Index months			
	CTA EW	CTA VW	S&P 500	HFRI
Monthly Average ER	-2.9%	-3.7%	1.2%	0.2%
	Best 5% VW CTA Index months			
	CTA EW	CTA VW	S&P 500	HFRI
Monthly Average ER	5.2%	5.2%	-2.4%	0.1%
Panel D: HFRI Index				
	Worst 5% HFRI Index months			
	CTA EW	CTA VW	S&P 500	HFRI
Monthly Average ER	1.8%	1.8%	-8.5%	-4.1%
	Best 5% HFRI Index months			
	CTA EW	CTA VW	S&P 500	HFRI
Monthly Average ER	1.9%	1.8%	-4.7%	4.2%

Table 3.7: CTA Performance: Bull and Bear Markets (By Trading Strategies)

This table shows average monthly excess returns for the best and worst 5% months (13 months each) of the equally weighted (EW) and value-weighted (VW) CTA portfolio, S&P 500 value weighted index (excl. dividends) and HFRI Index (Hedge fund research index)

Panel A: S&P 500

	Worst 5% S&P 500 months					
	Syst Trend	Syst Non-trend	Disc Trend	Disc Non-trend	S&P 500	HFRI
Monthly Average ER	3.1%	0.5%	2.7%%	0.3%	-10.1%	-3.5%

	Best 5% S&P 500 months					
	Syst Trend	Syst Non-trend	Disc Trend	Disc Non-trend	S&P 500	HFRI
Monthly Average ER	-2.0%	0.2%	-0.7%	0.2%	8.4%	2.1%

Panel B: HFRI Index

	Worst 5% HFRI Index months					
	Syst Trend	Syst Non-trend	Disc Trend	Disc Non-trend	S&P 500	HFRI
Monthly Average ER	3.3%	0.7%	3.5%	0.2%	-8.5%	-4.1%

	Best 5% HFRI Index months					
	Syst Trend	Syst Non-trend	Disc Trend	Disc Non-trend	S&P 500	HFRI
Monthly Average ER	2.4%	1.1%	3.4%	0.9%	4.7%	4.2%

Overall, Table 3.6 suggests that CTAs' countercyclical return movements are an additional benefit to investors while allocating capital to CTAs. Clearly, these benefits are not only reflected by smoothed returns across bear and bull markets, but also by lower return volatility achieved through risk diversification.

As shown in panel B, these benefits cannot be obtained by investing in hedge funds, since their returns differ from CTAs' return structure.

In Table 3.7 I repeat the assessment of assets' co-movements but I distinguish between the performances of individual trading strategies. In panel A, I document that trend-following CTAs are more sensitive to equity market swings than non-trend-following funds. For example, systematic trend followers fluctuate between 3.1% and -2.1% in the worst and best 5% months of the S&P 500 returns, while non-trend followers generate 0.5% and 0.2%, respectively. Similar dynamics can be observed for discretionary funds, for which returns fluctuate between 2.7% and -0.7% for trend followers and only between 0.3% and 0.2% for non-trend-following funds. Further, panel B reflects the disconnect between hedge fund and CTA returns. The thoroughly positive returns of all four trading strategies in HFRI's good and bad times point toward the fundamentally different investment approach taken by managers in each of the two active investment classes. In line with my earlier findings, this analysis suggests that not only average returns but also higher moments and the timing of return generation are crucial determinants for investors' decisions to allocate capital to CTAs.

3.4 Managerial Skill In The CTA Industry

The analysis of CTA performance has so far focused on the return generating process of net of fee excess returns. However, to make further statements about the skills of managers, I follow the literature and assess the gross returns of CTAs. Since most funds report net of fee returns to BarclayHedge, I follow the approach of French (2008) to obtain gross returns for each CTA in the database.

Table 3.8: CTA and Benchmark Returns: A Comparison

Panel A provides summary statistics for EW and VW CTA portfolios before and after fees. The t-statistics refer to the null hypothesis that funds generate on zero average returns ($H_0 : \mu_{CTA} = 0$). Panel B shows summary statistics for Moskowitz et al. (2012)'s time series momentum factors, available at `https://www.aqr.com/library/data-sets/timeseries-momentum-factors-monthly` and Fung and Hsieh (2001)'s portfolio straddle factors (PTFS), available at `https://faculty.fuqua.duke.edu/~dah7/HFRFData.htm`. The abbreviations BD, FX, COM, IR, STK, EQ refer to bonds, foreign exchange, commodities, interest rates, stocks, and equities, respectively. SP500 refers to the VW index including dividends and AGG denotes Barclay's Aggregate Bond Index.

Panel A: CTA returns - before and after fees

	Ann. Avg. Return	Ann. Volatility	Ann. Sharpe Ratio	T-Statistic
	EW Index			
Before fees	11.5%	8.2%	1.4	6.54
After fees	4.1%	7.2%	0.56	2.64
	VW Index			
Before fees	10.6%	8.5%	1.24	5.83
After fees	4.5%	7.6%	0.60	2.82

Panel B: Benchmark Strategies: Summary Statistics

	Ann. Avg. Return	Ann. Volatility	Ann. Sharpe Ratio
TSMOMCOM	0.121	0.153	0.79
TSMOMEQ	0.201	0.269	0.75
TSMOMBD	0.166	0.275	0.60
TSMOMFX	0.123	0.18	0.68
PTFSBD	-19.2%	0.53	-0.36
PTFSFX	-8.9%	0.674	-0.13
PTFSCOM	-4.5%	0.495	-0.09
PTFSIR	-11.9%	0.891	-0.13
PTFSSTK	-58.4%	0.488	-1.20
SP500	0.01	0.148	0.67
AGG	-0.11%	0.037	-0.03

For most funds, the fees consist of an annual management and a performance fee, which is charged only when the fund generates returns over a certain threshold. The management fee ranges from 0% to 20% with a mean of 1.8% and a standard deviation of 1%. The performance fee ranges from 0% to 50% and has an average of 20% and a standard deviation of 5%. Unfortunately, BarclayHedge does not have data on a fund's high-water mark or hurdle rate. Therefore, I take the most conservative approach and assume all funds have a high-water mark and for all CTAs I choose the 3-Month Treasury Bill as a hurdle rate.

Allowing for both features ensures that I do not overestimate gross returns artificially.[10]

As shown in Table 3.8, gross excess returns, defined as returns before fees but in excess of the risk-free rate, are approximately three times larger than net excess returns for the EW index, and roughly twice as large for the VW index. The impact of fees on the difference between net and gross excess returns is comparable to Bhardwaj et al. (2014) who construct gross returns using the same approach. In contrast to their paper, however, I find that gross and net excess returns are significantly different greater than zero, as indicated by the high t-statistics.

Next, equipped with EW and VW gross return indices, I assess whether funds can produce abnormal returns in excess of different alternative trading strategies. I use Fung and Hsieh (2001) portfolio straddle factors as a first benchmark strategy. The authors argue that trend-following strategies can be replicated by using option portfolio straddles and, therefore, are expected to explain a large proportion of the variation in gross CTA returns.[11].[12] Second, I use time series momentum factors (TSMOM) by Moskowitz et al. (2012) as simple normative benchmarks. Since CTAs generate returns by exploiting large consistent price trends, momentum trading is an alternative benchmark that replicates comparable return structure. Like the CTA gross indices, benchmark strategies do not include transaction costs, which makes using gross returns more accurate than using net returns. As shown

[10] I use different specifications and find that the impact of high water mark on CTA gross returns is small.

[11] The authors construct portfolio straddle factors for five different asset classes: bonds (PTFSBD), foreign exchange (PTFSFX), commodities (PTFSCOM), interest rates (PTFSIR) and stock markets (PTFSSTK).

[12] Time series momentum strategies are constructed for commodities (TSMOMCOM), equities (TSMOMEQ), bonds (TSMOMBD) and foreign exchange (TSMOMFX).

in panel B of Table 8, CTA gross returns outperform all the nine individual strategies, the S&P 500 and Barclay's Aggregate Bond Index (AGG) in terms of Sharpe Ratio.

While CTA returns appear to generate better Sharpe Ratio, I also test if managers can generate abnormal returns over and above these simple trading strategies. I postulate that a significant gross alpha would indicate that CTAs generate returns that beat passive trading strategies through their security selection skills and/or marketing timing ability. The results for the EW and VW indices are shown for both models in Table 3.9. In addition to the portfolio-straddle (PTFS) and time series momentum factors (TSMOM), I include returns from the S&P 500 and the AGG index as passive benchmarks (Bhardwaj et al. (2014)).

Tables 3.9 shows regression outcomes for different model specifications. As displayed, independent of the right-hand side variables, the intercept term is statistically significant at the 1% level. Furthermore, the intercept is also economically significant, highlighting the existence of managerial skills among CTAs. For example, as shown in column (4), when the VW index is the dependent variable and time series momentum factors are used as benchmark strategies, CTAs can generate 0.44% abnormal returns per month (5.3% annualized). Also, as shown in column (6), even adding PTFS and TSMOM factors in the same model (column (5)), leaves a significant abnormal gross excess return of 0.53% per month (6.4% annualized). Similar findings are seen in Table 3.10 with abnormal returns ranging from 0.37% (4.4% annualized) for systematic trend followers 1.26% (15.12% annualized) in Table 3.11 for discretionary trend followers.

Table 3.9: CTA Manager Skill and Gross Alpha

This table shows the regression results for two models, with Fung and Hsieh (2001)'s portfolio straddle factors (PTFS) and/or Moskowitz et al. (2012)'s time series momentum (TSMOM) factors used as explanatory variables. The abbreviations BD, FX, COM, IR, STK, EQ refer to bonds, foreign exchange, commodities, interest rates, stocks, and equities, respectively. Both models include the SP500 and Barclay's Aggregate Bond Index (AGG) as passive investment benchmark. The dependent variable is the VW CTA portfolio. The sample period is January 1994–December 2015. Coefficients are displayed in percentage terms. Numbers in parentheses refer to OLS t-statistics, based on Newey-West standard errors. ***, **, * indicate 1%, 5% and 10% level of significance, respectively.

	(1) EW	(2) VW	(3) EW	(4) VW	(5) EW	(6) VW
alpha	1.08***	1.00***	0.57***	0.44***	0.70***	0.53***
	(8.32)	(6.89)	(4.31)	(3.31)	(5.82)	(4.08)
TSMOMCOM			0.13***	0.12***	0.14***	0.12***
			(4.44)	(3.96)	(5.45)	(4.49)
TSMOMEQ			0.04**	0.06***	0.03**	0.05***
			(2.40)	(3.47)	(2.08)	(3.39)
TSMOMBD			0.09***	0.11***	0.07***	0.09***
			(5.38)	(6.23)	(4.54)	(5.28)
TSMOMFX			0.08***	0.09***	0.04	0.06**
			(3.04)	(3.36)	(1.58)	(2.33)
PTFSBD	0.02***	0.03***			0.02***	0.02***
	(2.82)	(2.65)			(2.62)	(2.61)
PTFSFX	0.04***	0.03***			0.04***	0.02***
	(6.02)	(4.19)			5.75	(3.43)
PTFSCOM	0.04***	0.03***			0.03***	0.02**
	(4.29)	(3.27)			(3.65)	(2.45)
PTFSIR	0.00	0.00			0.00	0.00
	(0.89)	(0.52)			(0.01)	(0.47)
PTFSSTK	0.01	0.01			0.01	0.00
	(1.45)	(1.13)			(1.04)	(0.47)
SP500	0.03	0.03	-0.02	-0.01	0.05*	0.05
	(0.90)	(0.87)	(-0.83)	(-0.33)	(1.68)	(1.54)
AGG	0.19	0.28**	0.15	0.18	0.08	0.13
	(1.62)	(2.23)	(1.25)	(1.42)	(0.73)	(1.12)
$\bar{R}^2$	0.31	0.21	0.3	0.34	0.48	0.43
N	264	264	264	264	264	264

Table 3.10: CTA Manager Skill and Gross Alpha: Systematic Traders

Table 10 shows the regression results for two models, with either Fung and Hsieh (2001)'s portfolio straddle factors (PTFS) or Moskowitz et al. (2012)'s time series momentum (TSMOM) factors used as explanatory variables. The abbreviations BD, FX, COM, IR, STK, EQ refer to bonds, foreign exchange, commodities, interest rates, stocks, and equities, respectively. Both models include the SP500 and Barclay's Aggregate Bond Index (AGG) as passive investment benchmark. The dependent variable is the VW CTA portfolio. The sample period is January 1994 to December 2015. Coefficients are displayed in percentage terms. Numbers in parentheses refer to OLS t-statistics, based on Newey-West standard errors. ***, **, * indicate 1%, 5% and 10% level of significance, respectively.

	Systematic Trend			Systematic Non-Trend		
	(1)	(2)	(3)	(4)	(5)	(6)
alpha	1.31***	0.37*	0.56***	0.69***	0.48***	0.46***
	(5.91)	(1.88)	(2.87)	(5.86)	(4.11)	(3.85)
TSMOMCOM		0.20***	0.21***		0.05**	0.06**
		(4.47)	(4.99)		(2.06)	(2.34)
TSMOMEQ		0.10***	0.09***		0.02	0.01
		(3.83)	(3.83)		(1.06)	(0.93)
TSMOMBD		0.17***	0.14***		0.07***	0.06***
		(6.49)	(5.45)		(4.55)	(4.07)
TSMOMFX		0.13***	0.09**		0.04*	0.02
		(3.44)	(2.54)		(1.78)	(0.97)
PTFSBD	0.04***		0.04***	0.01*		0.01
	(2.79)		(2.90)	(1.75)		(1.10)
PTFSFX	0.04***		0.03**	0.03***		0.02***
	(3.35)		(2.44)	(4.21)		(3.74)
PTFSCOM	0.06***		0.04***	0.01		0.00
	(3.71)		(2.94)	(0.65)		(0.01)
PTFSIR	0.00		0.00	0.00		0.00
	(- 0.54)		(0.51)	(- 0.22)		(0.43)
PTFSSTK	0.03*		0.02	- 0.01		- 0.01
	(1.81)		(1.26)	(- 0.96)		(- 1.45)
S & P500	0.02	- 0.04	0.05	0.07**	0.06**	0.08***
	(0.44)	(- 0.90)	(1.06)	(2.45)	(2.36)	(3.09)
AGG	0.49**	0.32*	0.26	0.11	0.02	- 0.01
	(2.50)	(1.72)	(1.47)	(1.08)	(0.17)	(- 0.10)
$\bar{R}$	0.22	0.38	0.47	0.10	0.14	0.19
N	264	264	264	264	264	264

Table 3.11: CTA Manager Skill and Gross Alpha: Discretionary Traders

Table 10 shows the regression results for two models, with either Fung and Hsieh (2001)'s portfolio straddle factors (PTFS) or Moskowitz et al. (2012)'s time series momentum (TSMOM) factors used as explanatory variables. The abbreviations BD, FX, COM, IR, STK, EQ refer to bonds, foreign exchange, commodities, interest rates, stocks, and equities, respectively. Both models include the SP500 and Barclay's Aggregate Bond Index (AGG) as passive investment benchmark. The dependent variable is the VW CTA portfolio. The sample period is January 1994 to December 2015. Coefficients are displayed in percentage terms. Numbers in parentheses refer to OLS t-statistics, based on Newey-West standard errors. ***, **, * indicate 1%, 5% and 10% level of significance, respectively.

	Discretionary Trend			Discretionary Non-Trend		
	(1)	(2)	(3)	(4)	(5)	(6)
alpha	1.57***	0.94***	1.16***	0.61***	0.50***	0.53***
	(5.21)	(2.92)	(3.72)	(4.76)	(3.71)	(3.91)
TSMOMCOM		0.25***	0.27***		0.03	0.04
		(3.49)	(4.05)		(1.15)	(1.34)
TSMOMEQ		0.01	0		0.03*	0.01
		(- 0.34)	(- 0.10)		(1.64)	(0.68)
TSMOMBD		0.11**	0.06		0.01	0.01
		(2.51)	(1.57)		(0.43)	(0.49)
TSMOMFX		0.1	0.02		0.02	- 0.00
		(1.58)	(0.30)		(0.78)	(- 0.09)
PTFSBD	0.04**		0.04**	- 0.01*		- 0.01
	(2.37)		(2.04)	(- 1.76)		(- 1.60)
PTFSFX	0.07***		0.07***	0.03***		0.03***
	(4.36)		(4.21)	(3.89)		(3.64)
PTFSCOM	0.08***		0.07***	0.02*		0.02*
	(3.78)		(3.31)	(1.91)		(1.64)
PTFSIR	0.00		0.01	0.00		0.00
	(0.03)		(0.53)	(0.30)		(0.44)
PTFSSTK	0.01		0.01	0.00		0.00
	(0.38)		(0.29)	(0.50)		(0.40)
S&P500	- 0.06	- 0.17**	- 0.03	0.03	0.00	0.03
	(- 0.89)	(- 2.41)	(- 0.44)	(0.95)	(- 0.01)	(0.99)
AGG	0.48*	0.57*	0.44	- 0.42***	- 0.37***	- 0.43***
	(1.80)	(1.90)	(1.59)	(- 3.73)	(- 2.95)	(- 3.51)
$\bar{R}$	0.24	0.14	0.29	0.11	0.04	0.10
N	264	264	264	264	264	264

Furthermore, the regression analysis shows that the PTFS and TSMOM factors explain a large proportion of the variance in CTAs' returns. For example, if solely PTFS factors are used as regressors, the adjusted R^2 accounts for at least 0.21 and for the TSMOM factors, adjusted R^2 increases to even 0.30 and 0.34 for EW and VW, respectively. Moreover, the combination of the two sets of factors results in an adjusted R^2 of up to 0.48, explaining nearly half of the variance of CTA returns. This significant increase, when combining the two sets of factors, highlights that PTFS and TSMOM factors capture different dynamics of CTAs' return generating process. Tables 3.10 and 3.11 show how the explanatory power of these factors varies between CTA trading strategies. Generally, TSMOM factors explain a larger degree of return variance than PTFS factors, pointing toward the similarities between time series momentum and CTAs' trend-following strategies. For systematic and discretionary trend followers the R^2 is 0.38 and 0.24 when solely the TSMOM factors are employed as regressors, while the R^2 remains comparably low for non-trend followers (0.14 and 0.06). Combining both sets of factors in one regression again leads to high explanatory power of up to 0.47, confirming the use of these factors as appropriate benchmark strategies.

3.4.1 Crisis Alpha

The diversity in trading strategies and managerial skill becomes even more apparent when looking at CTA returns in times of equity market turmoil. While the positive gross excess intercept term can be interpreted as an indicator of a manager's skill in general, I want to investigate further whether CTAs can make use of their skill during downturns in equity markets. CTAs generate positive excess returns of up to 3% during the worst 5% months of the S&P 500 (Table 3.12). Here I analyse whether these returns are subject to a trading strategy that cannot be replicated by the PTFS or TSMOM factors. I test for the existence of crisis alpha, by extending the previous regression by an additional intercept term and by estimating the following model:

$$R_t^G = \alpha_1 + \alpha_2 \mathbb{I} + \pi_t^{j,B} + \epsilon_t \tag{3.1}$$

where

$$\pi_t^{1,B} = \beta_1 + \beta_2 TSMOMCOM_t$$
$$+ \beta_3 TSMOMEQ_t + \beta_4 TSMOMBD_t + \beta_5 TSMOMFX_t$$
$$\pi_t^{2,B} = \beta_1 + \beta_2 PTFSBD_t$$
$$+ \beta_3 PTFSFX_t + \beta_4 PTFSIR_t + \beta_5 PTFSSTK_t$$

where R_t^G refers to the gross excess return of an EW or VW index α_1 is an intercept term and $\pi_t^{j,B}$ is the risk premium of a benchmark return strategy. Again, I use Fung and Hsieh (2001) (FH) portfolio straddle factors ($j = 1$) or Moskowitz et al. (2012)'s (2012) time series momentum factors as a benchmark ($j = 2$). To measure the skill of CTAs during crisis periods, I allow for α_2, where $\mathbb{I}$ refers to a dummy variable term, set equal to 1 during the 5% worse performing months of the S&P 500, such that it measures the skill of a CTA manager during market downswings. The intercept α_1 captures the average skill of managers during the remaining periods. The results are shown in Table 3.12. For brevity, I focus on the two intercept terms and their joint impact.

As seen in panel A, independent of the explanatory variables, the intercept term α_1 is positive and statistically significant. For the dummy variable intercept term (α_2) only EW indices and the VW index with FH factors as regressors show significant coefficients at the 10% level or higher. Concerning the joint impact (α_2), I find that both intercept terms are significant at least at the 10% level for all six specifications. For the time series momentum strategies, I find that the average monthly return in non-crisis times accounts for 0.49% and 0.37% for the EW and VW index, respectively. In crisis times this value triples to 1.70% abnormal monthly gross excess returns for the EW index and even 1.42% for the VW index ($\alpha_t otal$). Both values are not only economically but also statistically significant at the 5% and 10% level. Overall, CTAs appear to be particularly profitable investment opportunities during equity market downturns.

Panel B provides insights about what kind of trading strategy can generate abnormal gross excess returns during crisis periods. All but systematic trend funds generate significant

Table 3.12: Crisis Alpha

This table reports coefficient estimates of equation 3.1, focusing on the two intercept terms. In panel A, as indicated, either the equally (EW) or value- (VW) weighted index is used as dependent variable and either Fung and Hsieh (2001)'s portfolio straddle factors (FH), Moskowitz et al. (2012)'s time series momentum (TSMOM) factors, or both combined (All) are used as explanatory factors. In panel B, VW indices are used as dependent variable. Numbers in parentheses refer to OLS t-statistics, based on Newey-West standard errors. ***, **, * indicate 1%, 5% and 10% level of significance, respectively.

Panel A: Crisis Alpha - All CTAs

	FH-Factors		TSMOM- Factors		All Factors	
	EW	VW	EW	VW	EW	VW
α_1	0.90***	0.78***	0.49***	0.37***	0.61***	0.46***
	(6.30)	(4.98)	(3.52)	(2.65)	(4.73)	(3.28)
α_2	2.08***	2.34***	1.21*	1.06	1.24*	0.98
	(3.07)	(3.10)	(1.71)	(1.49)	(1.94)	(1.40)
$\alpha_{total} = \alpha_1 + \alpha_2$	2.98***	3.12***	1.70**	1.43*	1.85***	1.44*
	(3.96)	(3.74)	(2.23)	(1.86)	(2.65)	(1.89)
$\bar{R}^2$	0.33	0.24	0.30	0.35	0.48	0.43

Panel B: Crisis Alpha - By Trading Strategy

	TSMOM- Factors			
	Systematic Trend	Systematic Non-trend	Discretionary Trend	Discretionary Non-trend
α_1	0.24	0.47***	0.86**	0.45***
	(1.15)	(3.81)	(2.53)	(3.20)
α_2	2.04*	0.13	1.21	0.67
	(1.93)	(0.20)	(0.70)	(0.93)
$\alpha_{total} = \alpha_1 + \alpha_2$	2.28**	0.59	2.07	1.12
	(1.99)	(0.88)	(1.11)	(1.44)
$\bar{R}^2$	0.39	0.20	0.14	0.04

and positive monthly alphas during non-crisis periods (α_1). However, during times of market turmoil, only systematic trend followers can generate statistically significant alphas that account for more than 2% in each month. The crisis alpha (α_{Crisis}) is statistically significant at the 5% level. These findings are in line with Kazemi and Li (2009) who argue that systematic funds have a better market timing ability than discretionary traders, implying that systematic traders successfully adjust their portfolios just before equity turmoil and subsequently generate higher returns from directional investments with or against long-lasting price trends. Furthermore, the result can be linked to earlier studies (Kaminsky (2011); Kaminsky and Mende (2011)) that refer to crisis alpha as profits that are generated during crisis periods by exploiting large price trends. The analysis indicates that systematic trend followers are most adept at benefiting from distressed markets.

3.5 Managerial Skill and Performance Persistence

Having established CTA managers' skill through the analysis of gross returns, in this section I assess their performance using an alternate measure: the amount of capital that funds are able to extract from financial markets. To this end, I use an empirical procedure developed by Berk and van Binsbergen (2015) to estimate the value added by a fund as the gross excess return over a specific benchmark strategy, multiplied by its assets under management. Berk and van Binsbergen (2015) argue that this measure is more precise than net or gross abnormal returns obtained from standard regression models, as it takes into account the number of assets managed by a fund. For example, since the size of CTAs ranges between USD 10,000 and USD 5.3 billion,[13] the added value of two funds with the same abnormal return might vary greatly from each other because of the differences in the size of the funds' AUM. This dimension is not captured by the gross alpha. Therefore, calculating the added value of a CTA allows me to assess managerial skill from a new perspective that takes return dynamics and fund size into account.

According to Berk and van Binsbergen (2015), the value added by a fund between period $t-1$ and t is defined as:

$$V_i t = q_{i,t-1}(R^G_{i,t} - R^B_{i,t}) \tag{3.2}$$

where $q_{i,t-1}$ are fund i's assets under management in period $t-1$ measured in 2005 dollar terms,[14] $R^G_{i,t}$ is its gross return and $R^B_{i,t}$ is a return from an alternative benchmark investment that I calculate below. Once I have constructed the valued added for each

[13] Values refer to real AUM of the first reported entry to BarclayHedge.

[14] I transform nominal AUM to real AUM dividing it P_t/P_0, where P_t is the US-CPI index in period t and P_0 US CPI index in year 2005.

individual CTA, I calculate the average value $\hat{S}_i$ a fund generates over its entire lifetime as

$$\hat{S}_i = \sum_{t=1}^{T} \frac{V_i t}{T_i} \tag{3.3}$$

Similarly, the average value added across all funds is given by

$$\bar{S} = \frac{1}{N} \sum_{i=1}^{N} \hat{S}_i \tag{3.4}$$

where N refers to the total number of funds, represented in BarclayHedge. Lastly, I follow Berk and van Binsbergen (2015) and calculate a weighted measure of the average value added by taking into account the number of years a fund is actually reporting to BarclayHedge, that is

$$\bar{S}_W = \sum_{i=1}^{N} \frac{T_i \hat{S}_i}{\sum_{i=1}^{N} T_i} \tag{3.5}$$

Since more skilled managers stay alive for a longer period of time and, therefore, add more value, I would expect the weighted measure $\bar{S}_W$ to be larger than the simple cross-sectional average $\bar{S}$.

To construct the value added ($V_i t$) for each fund, I use Moskowitz et al. (2012) time series momentum factors as a benchmark trading strategy. More precisely, I estimate

$$\begin{aligned} R_t^G = {} & \beta_1 + \beta_2 TSMOMCOM_t \\ & + \beta_3 TSMOMEQ_t + \beta_4 TSMOMBD_t + \beta_5 TSMOMFX_t \end{aligned} \tag{3.6}$$

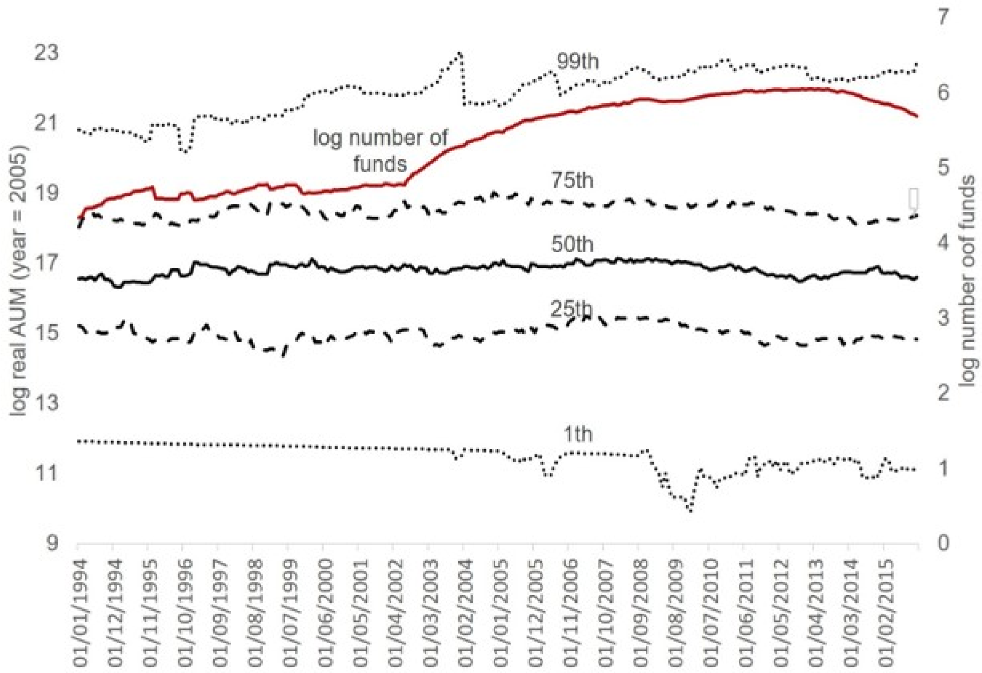
where β_i is the regression coefficient associated with one of the four time series momentum factors. Then, I reconstruct R_t^B from the regression's fitted values, so that the time series of benchmark returns obtained has the same level of risk implied by the four-TSMOM factor model. I choose these factors as a benchmark strategy for several reasons. First, benchmark factors should be tradeable portfolios that serve as simple passive strategy. This condition is clearly fulfilled by this benchmark since investors could simply reconstruct the TSMOM factors by investing into short and long portfolios, depending on an asset's prior returns. Second, previous research has emphasized CTAs' extensive use of time series momentum strategies (Baltas and Kosowski (2013); Elaut and Erdös (2016)). The high R^2 in the previous regression analysis of up 0.48 stresses the high explanatory powers of this trading style. Third, Moskowitz et al. (2012) argue that their time-series momentum factors are implementable strategies that generate the same payoff structure as Fung and Hsieh (2001)'s options straddle factors. Since time series momentum factors are easier to implement, I choose a passive strategy over the dynamic option straddle factors.

To alleviate concerns that results are driven by the growing size of the industry, I follow Berk and van Binsbergen (2015) and plot the log number of funds reporting to BarclayHedge as well as the log fund size of different percentiles over the entire sample period. Figure 3.2 illustrates that the median fund size (base year 2005) remains comparably stable over the entire period, while the number of reporting CTAs is growing, particularly since 2001. The growth of the industry's total AUM can therefore be attributed to an increasing number of CTAs, rather than an increase in the size of CTAs. These industry dynamics are comparable to those reported by Berk and van Binsbergen (2015).

As shown in Table 3.13, the average added value by a CTA is USD 0.49 million (base year 2005) and the reporting life time-weighted average is USD 1.27 million. Both numbers are significantly greater than zero at the 5% level using a one-sided t-test. Moreover, these values are substantially higher than the USD 0.27 and USD 0.14 million added value of mutual funds in the study by Berk and van Binsbergen (2015). In line with the authors, I argue that the differences in the cross-sectional means highlight that more talented managers have a longer lifespan. However, it is worth noting that, as demonstrated in the lower half of Table Berk and van Binsbergen (2015), there is a substantial cross-sectional variation in managerial skill. Value added by CTAs ranges from a loss of nearly USD 4 million in the bottom 1% to profits of USD 6.82 million in the very top. Furthermore, roughly two-fifths of the 926 funds do not add significant value. It is worth noting that Berk and van Binsbergen (2015) find that up to 59% of mutual funds are not able to outperform passive benchmark strategies. These results indicate that CTA managers are more skilled than mutual fund managers - a reassuring figure given that CTA managers' compensation is orders of magnitude greater than that of mutual fund managers.[15]

[15] As highlighted by Stulz (2007), the compensation schemes of mutual funds and hedge funds differ fundamentally from each other. Mutual fund managers' compensation is more strictly regulated, usually depends solely on the fund's assets under management and investors pay no additional

Figure 3.2: Development Of Real Assets Under Management in the CTA Industry



Notes: This figure shows the development of the growth of the assets under managements (AUM) in the commodity trading advisor industry. The 1th, 25th, 50th, 75th, and 99th percentile of the log real AUM are displayed against the left axis (black lines, base year = 2005), while the log number of funds (red line) refers to the axis on the right.

Next, I test for persistence in managerial skill. To this end, I employ a skills ratio as defined by Berk and van Binsbergen (2015):

$$SKR_i^\tau = \frac{\hat{S}_i}{\sigma_i^\tau} \tag{3.7}$$

where $\hat{S}_i = \frac{\sum_{t=1}^{\tau}}{V_{it}}$ and $\sigma_t^\tau = \sqrt{(\sum_{t=1}^{\tau}(V_{it} - \hat{S}_i^\tau)^2)/\tau}$. In line with the authors, I take the following approach. First, I split the sample into sorting and forecasting periods. In the sorting sample, funds are sorted in quintiles according to their level of skill. The minimum number of reported months for each fund i is 24 and I re-estimate the skills ratio for each point in time τ based on an extending window approach, including all the fund information from period 1 until τ. Second, for each τ I then estimate the value added for each fund in

performance fee. In contrast, the performance fee is a substantial component of a CTA manager's compensation and the results suggest that more skilful managers use a higher fee structure to signal their skill.

Table 3.13: Manager Skill and Added Value

This table shows the value added by CTA managers over Moskowitz et al. (2012) time series momentum factors. Values are expressed in million USD (base year = 2005). The null hypothesis tested is whether the cross-sectional weighted average or the cross-sectional mean is larger than zero (formally: $H_0 > 0$) ***, **, * indicate 1%, 5% and 10% level of significance, respectively.

Cross-sectional weighted average	1.27
Standard error of the weighted mean	0.02
p-value	0.04**
Cross-sectional mean	0.49
Standard error of the mean	6.81
p-value	0.01**
1st percentile	-3.94
5th percentile	-0.31
10th percentile	-0.21
50th percentile	0.006
90th percentile	0.57
95th percentile	1.47
99th percentile	6.82
Percent with less than zero	0.41
Number of funds	926

the periods $[V_{i,\tau+m}, ..., V_{i,\tau+m+h}]$, where h refers to the forecasting horizon and m to the minimum number of reported months after period τ. For each point in time I estimate the added value with information from the forecasting period only, not the sorting period. Since I estimate the benchmark return for each point in time with four time series momentum factors, I chose $m = 40$.[16] Concerning the forecasting horizon, I use different lengths for h, ranging between $h = 3$ and $h = 36$ months. The upper bound is chosen to account for the fact that the average lifetime of funds is approximately 70 months so that funds in the sorting period may not report any longer in the forecasting period. At the end, I obtain

[16] The minimum number of reporting months is chosen to be m=40 to allow for a sufficient number of degrees of freedom in each of the rolling regressions. Results are qualitatively similar to other specifications, such as e.g. $m = 30$ or $m = 50$.

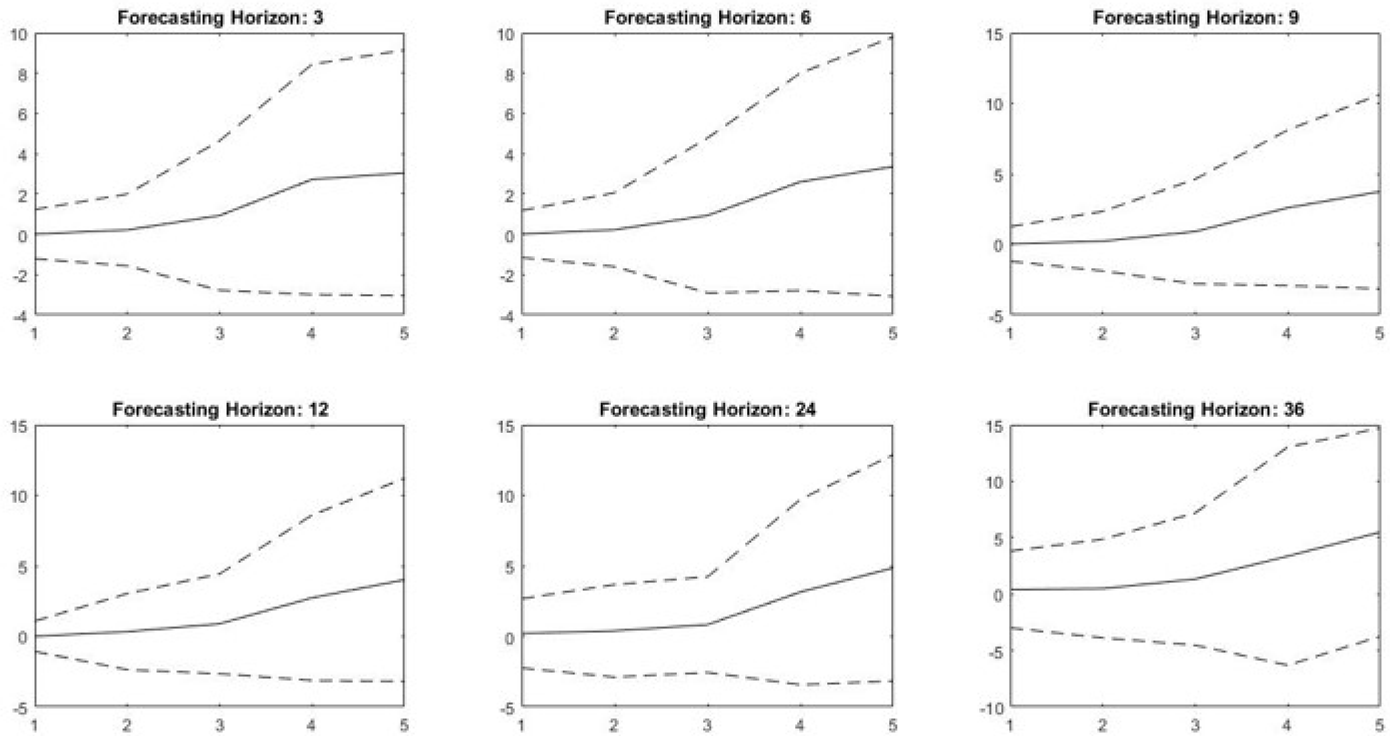
a time series of monthly average value added for each of the five portfolios. To evaluate persistence, I examine how often the valued added by the bottom 20% (Portfolio 1) is outperformed by the top 20% (Portfolio 5) and in how many months the latter outperforms the former. Results are shown in Table 3.14 and Figure 3.3.

Table 3.14: Manager Skill and Performance Persistence

This table shows the average value added by CTAs (in USD million; base year = 2005) sorted in the bottom and top portfolios for different forecasting horizons. The t-statistic refers to the test whether the average value added test in the bottom and top portfolios are the same. The table also shows the number of times the top quintile outperforms the bottom quintile. Returns from Moskowitz et al. (2012)'s time series momentum factors are used as benchmark trading strategy. Portfolios are sorted based on a manager's skill ratio.

Forecasting horizon	Bottom 20% Value Added	Top 20 % Value Added	T-statistic	Top 20% outperforms bottom 20%
3	0.01	3	13.2	0.96
6	0.02	3.4	13.6	0.96
9	0.04	3.7	14	0.97
12	0.01	4	14.3	0.99
24	0.21	4.9	14.9	0.94
36	0.41	5.5	14.1	0.84

Figure 3.3: Predictability of Fund Performance: Manager Skills Ratio



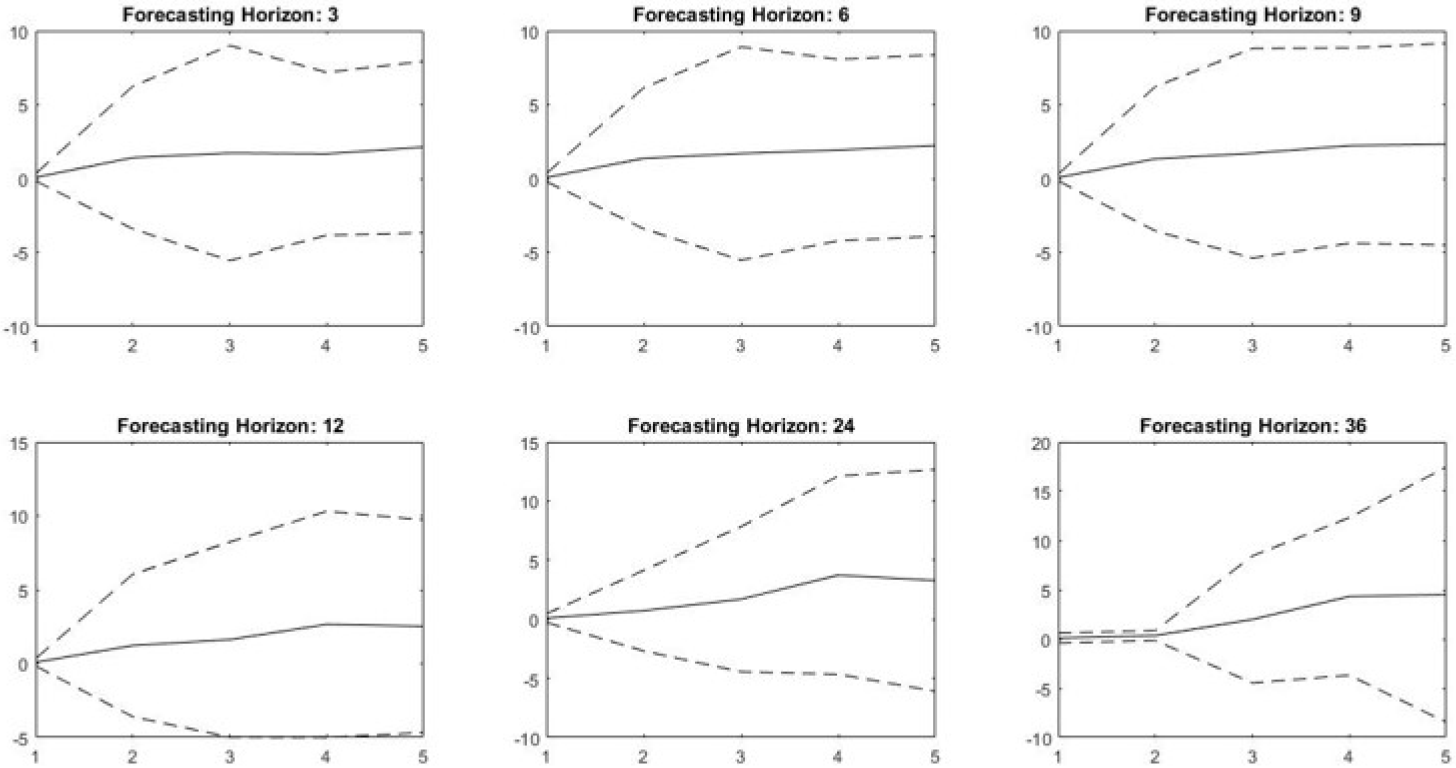
Notes: This figure shows the added value of CTAs sorted into portfolios over six different forecasting horizons (3, 6, 9, 12, 24, 36 months). The y-axis measures added value in USD million (base year = 2005) and the x-axis refers to the five portfolios. Portfolio 1 refers to CTAs with the lowest skill ratio and Portfolio 5 refers to CTAs with the highest skill ratio. The solid line refers to the average added value, while the dashed lines refers to the 95% confidence intervals.

Table 3.14 shows the value added by the funds in the top 20% and in the bottom 20% of the sample. As displayed, the values added by the two groups of CTAs differ significantly from each other. For example, the predicted added value of the bottom 20% is only USD 0.1 million, while the top 20% of CTAs' added value accounts for USD 3 million. As indicated by the large t-statistics, the added value between the two groups differs significantly across all forecasting horizons. This indicates that managers with more managerial skill persistently perform better than their less skilled peers. In addition, I find that in almost every month the top 20% outperforms the bottom 20%. The most skilled CTAs outperform the least skilled managers 96% of the time for the shortest forecasting horizon ($h = 3$).

Furthermore, in Figure 3.3 the solid line shows the average added value (y-axis) for each portfolio (x-axis) for all six forecasting horizons. Independent of the forecasting horizon, h, I find that more skilled funds (Portfolio 5) extract more value from capital markets than less skilled managers (Portfolio 1). I interpret this finding as evidence that better performance is not due to managers' luck but rather to their managerial skills.

Finally, I assess if investors can infer a priori whether some managers are more skilled than others. Since managerial skill is a scarce good and the cross-sectional variation is large, rational investors would prefer to allocate their capital to CTAs that provide the best performance. In line with Berk and van Binsbergen (2015), I assess whether investors can learn about managers' future performance based on their current compensation. If compensation predicts future performance, managers could use it as a credible and observable signal of their skill and attract more capital from investors. The existence of such a signalling mechanism would indicate an efficient and competitive CTA market (Akerlof (1970)). To control for the ex-ante predictability of future performance, I sort funds into quintile portfolios based on their compensation, which is defined as accrued fees multiplied by AUM. Using only the overall fee is problematic because the CTA fee structure is not very diverse. In my sample, 53% of funds report the typical 2/20 fee structure of management and performance fees. On the other hand, the amount of capital managed by a CTA varies greatly in the cross-section and is crucial for managerial compensation.

Figure 3.4: Predictability of Fund Performance: Manager Compensation



Notes: This figure shows the added value of CTAs sorted into portfolios over six different forecasting horizons (3, 6, 9, 12, 24, 36 months). The y-axis measures added value in USD million (base year = 2005) and the x-axis refers to the five portfolios. Portfolio 1 refers to the CTA with the lowest compensation and Portfolio 5 to the CTA with the highest compensation. The solid line refers to the average added value, while the dashed lines refers to the 95% confidence intervals.

Table 3.15: Compensation Scheme and Performance Persistence

This table shows the average value added by CTAs (in USD million; base year = 2005) sorted in the bottom and top portfolios for different forecasting horizons. The t-statistic refers to the test whether the average value added test in the bottom and top portfolios are the same. The table also shows the number of times the top quintile outperforms the bottom quintile. Returns from Moskowitz et al. (2012)'s time series momentum factors are used as benchmark trading strategy. Portfolios are sorted based on a manager's compensation scheme.

Forecasting horizon	Bottom 20% Value Added	Top 20 % Value Added	T-statistic	Top 20% outperforms bottom 20%
3	0.08	2.1	9.7	79%
6	0.07	2.2	9.5	83%
9	0.07	2.3	8.8	80%
12	0.09	2.5	9	82%
24	0.1	3.3	8.6	81%
36	0.1	4.5	8.2	72%

Table 3.15 provides support for the hypothesis that investors compete to allocate money to successful CTA managers. As indicated by the high t-statistics, funds that demand the highest compensation from investors outperform funds with the lowest compensation scheme. The value added by the costliest top 20% exceeds the performance of the bottom 20% in at least 72% of all months. Figure 3.4 supports this claim. In all cases, CTAs in Portfolio 5 add more value than funds in the lower ranked portfolios for up to nine months. For $h = 12$ and $h = 24$, Portfolio 4 slightly outperforms the most expensive funds leading to a slight kink in the solid line. In addition, 95% confidence bands increase with a larger forecasting horizon (indicated by the scale of the y-axis), adding greater uncertainty about a fund's future performance. However, in the short run, high compensation ex ante predicts future performance. Therefore, the results indicate that managers use their compensation to signal their skills to investors, who use this information while determining their fund allocations. Overall, I conclude that the value added provides additional evidence of managerial skill in the CTA industry and that CTA managerial pay increases commensurably with performance.

3.6 Conclusion

The CTA industry has grown rapidly over the past 20 years. However, extant empirical evidence indicates that CTA managers have generated statistically insignificant net excess returns and have passive benchmarks. If such is the case, why do professional, sophisticated investors continue to invest in these underperforming funds? Is this a consequence of investor irrationality? Or does the market thrive because it is too opaque to be aware of its own failing? Clearly, the puzzling growth of CTAs raises fundamental questions about academics' understanding of the operational efficiency of the CTA industry and the alternative investments market at large. I employ a large and representative dataset of CTAs to provide a new perspective on the performance of CTAs, the skill of their managers, and the relation between CTA managers' pay and performance.

The dataset is derived from the Barclay's Hedge Fund Database and data provided by Risk and Portfolio Management AB (RPM). The dataset has several advantages over those used in extant studies. First, it provides the most compressive of coverage of the CTA industry—70% of the total assets under CTA management, on average, between 1985 and 2015. Second, it is largely free of any graveyard bias as it has been downloaded by RPM on a daily basis for a large proportion of the sample period. Third, it enables me to classify CTAs into to four strategy groups: systematic and discretionary trend followers and their non-trend-following counterpoints. Additionally, I use a smaller proprietary dataset of realized CTA returns to validate my results obtained from the larger sample.

In contrast to earlier studies, I find that equally (EW) and value-weighted (VW) portfolios of CTAs generate on average 4.1% and 4.5% excess returns for investors on an annual basis. Notably, these returns are net of all fees. Despite high management and performance fees, CTAs are a profitable investment opportunity for investors. The results also show that CTA returns are positively skewed, countercyclical to equity markets and largely uncorrelated with hedge fund returns. I also document that CTA managers outperform normative benchmarks, such as time series momentum strategies, and produce up to 8.4% abnormal gross excess return on an annual basis. Testing formally in a regression framework for the existence of "crisis alpha", I find that systematic trend-following funds produce on average more than 27% annualized abnormal returns by exploiting large price trends during crisis times. Next, measuring managerial skill by the amount of capital that CTAs extract from financial markets, I show that CTAs, on average, add value of USD 1.27 million per month, with roughly 60% of the CTAs in the sample generating more value than passive benchmark trading strategies. Finally, I find that the cross-sectional differences in managers' skills are persistent up to three years, ruling out the possibility that the evidence relating to managerial skill is driven by luck. Moreover, I show that managerial fees predict future performance, indicating that investors are able to identify and reward skilled managers.

The results show the CTA market to be well-functioning, one in which rational investors compete to invest with skilled managers, whose compensation is set in equilibrium so that the expected net alpha is zero (Berk and Green (2004)).

Chapter 4

FX Spot and Swap Liquidity and the Effects of Window Dressing

4.1 Introduction

This paper assesses liquidity conditions in the foreign exchange (FX) market using intraday data for the post-financial crisis period. Employing information on interdealer quoting activity from the beginning of 2010 until mid-2017, I also account for FX swap dealers' response to recent regulatory changes. With average daily trading volumes exceeding $5 trillion, the FX market is the world's deepest financial market, yet FX liquidity conditions are notoriously difficult to assess. For one, unlike say equity markets, FX trading is fragmented across many venues and is primarily executed over-the-counter (OTC). Furthermore, trading volumes in FX derivatives are an important source of liquidity and price discovery in FX markets. Specifically, daily trading volume in foreign exchange swap markets has been increasing significantly in recent years and substantially exceeded spot market turnover in 2016 (BIS, 2016).[1] Hence, it is crucial to account for FX derivatives, in addition to spot trading, when assessing FX liquidity conditions.

[1] In April 2016, daily average turnover in the foreign exchange spot market was 1,652 billion US dollar-equivalents, compared to 2,380 billion US dollar-equivalent for FX swaps (BIS, 2016).

This paper contributes to the international finance literature in several ways. First, it adds to the study on liquidity dynamics in currency markets which only recently witnessed growing attention. Mancini-Griffoli and Ranaldo (2011) provide a systematic assessments of FX spot liquidity, highlighting the substantial variation of liquidity across currency pairs. Banti et al. (2012) combine data on returns and order flows across currencies to construct a measure of systematic FX liquidity risk. Karnaukh et al. (2015) provide further evidence for commonality in FX liquidity, using daily data covering a large cross-section of currency pairs for more than twenty years. Hasbrouck and Levich (2017) examine liquidity dynamics across a large number of currencies using one-month of settlement data, complemented with high-frequency data on quotes. In contrast to these studies, I do not limit the analysis to spot markets, but take into account liquidity conditions in the FX swap market as well.[2] This extension allows me to explicitly account for the joint behaviour of FX market liquidity and FX funding liquidity.

The theoretical framework for the interaction of these liquidity measures is grounded in Brunnermeier and Pedersen (2009). Whereas market liquidity broadly refers to the costs of trade execution and the ability to trade large volumes without generating an out-sized price impact, funding liquidity refers to the ease with which such trades and the associated market positions can be funded. Importantly, funding instruments are themselves traded, and their pricing can affect market liquidity conditions, which can then feedback to funding costs. While Banti and Phylaktis (2015) do assess this interaction of funding liquidity with FX market liquidity, they look at funding liquidity conditions in repo markets, whereas this analysis follows a novel approach by constructing all the funding liquidity measures from activity in FX markets themselves. I measure FX funding liquidity by the forward spread (e.g. forward discount computed from quotes of swap points) and, hence, I look at funding liquidity in the proximate market, rather than relying on more removed measures such as Libor-OIS or the TED spread. The FX forward point spread can be interpreted as the funding costs of a foreign exchange swap that is used to borrow (lend) US dollar while lending (borrowing) a local currency in the spot market.[3]

[2] BIS (2017) covered issues related to the liquidity of currency markets in the Americas, including FX derivatives.

[3] To the extent that measures of FX funding liquidity, particularly when adjusted by benchmark money market rates, are closely related to deviations from covered interest parity (CIP), this work is somewhat related the CIP literature. However, as my aim is neither to measure CIP arbitrage nor to explain CIP failure, I abstract from this literature. Links to a number of recent papers

Second, I examine intra-day liquidity conditions, using Thomson Reuters Tick History (TRTH) data obtained from Reuters Datascope, while aforementioned analyses are largely conducted at daily or lower frequencies. Huang and Masulis (1999) is another closely related study to have used TRTH data to assess liquidity conditions, but they focused on DM/USD, only on spot, and on one year of data between 1992 and 1993, whereas I cover a much longer and recent time-period, and corroborated the findings in both JPY/USD and EUR/USD.

Third, I estimate the effects of dealer competition on liquidity in spot and FX swaps markets building upon an early stream of literature, which examines the relationship between spot market liquidity, measured via bid-ask spreads and dealer competition (Huang and Masulis (1999)). I also add to studies relating FX price discovery and dealer informational advantages to dealer size (Rosenberg and Traub, 2009; Bjonnes et al., 2009; Phylaktis and Chen, 2010; Menkhoff et al., 2016). I contribute to the literature by assessing the impact of large and small dealers on liquidity conditions.

Fourth, in line with recent studies examining the market environment against the background of post-crisis regulatory frameworks (e.g. Adrian, Fleming, Or, and Vogt, 2017), I analyse current trends of liquidity dynamics and discuss the impact of the changing dealer behaviour. In particular, I draw a line along a major difference in the regulatory treatment of dealer balance sheets: I distinguish between globally systematically important banks (G-SIBs) and smaller banks, as this captures the different constraints faced by large and smaller dealers for providing liquidity in spot versus the swap markets. I document significant balance sheet window dressing by G-SIB dealers around regulatory reporting periods, which translates into their scaling back on their market-making activity in FX swaps. The dynamics have strengthened as banks began managing their balance sheets in order to adhere to the new leverage and liquidity regulations.[4] Hence, I provide further insights into the

examining the persistent failure of CIP in the period following the global financial crisis can be found here: https://www.bis.org/events/bissymposium0517/programme.htm.

[4]For example, the Basel standards required public disclosure of the Basel III leverage ratio by international banks as of January 2015, although the leverage ratio will become a mandatory part of the Basel III Pillar 1 requirements only in January 2018, after a period of monitoring and final calibration. Similarly, liquidity coverage ratios (LCR) was phased in at 60% in January 2015, and gradually rising to 100% over a period of four years. The former has implications for the contribution of off-balance sheet derivatives, such as FX swaps, to the total exposure calculation under the leverage ratio, while the latter affects how bank use FX swaps for their cross-currency liquidity management.

drivers of quarter-end anomalies in FX swap pricing documented by recent literature. Du and Verdelhan (2018) find that quarter-end spikes in short-term FX swap basis are in line with the maturity of the contracts relative to the reporting dates. Borio et al. (2016) show that quarter-end widening of FX swap basis is closely related to the quarter-end divergence in repo market spreads in respective currencies. Arai et al. (2016) take the fact that FX swap spreads and GC repo spreads widen at quarter-ends while Tri-party repo spreads do not widen as evidence of bank balance sheet management under the leverage ratio.[5] I contribute to this literature by showing that such quarter-end tightening in FX swap funding conditions is explained by the pull-back by G-SIB dealers from FX swap markets. I also document significant adverse spillovers to FX spot market arising from such dealer balance sheet window-dressing in FX derivatives.

The main data source for this study is Reuters Datascope. I obtain data for JPY/USD and EUR/USD spot market, 1-month swap points, and 1-month overnight index swap (OIS) rates for the period February 2010 to May 2017. The database documents entries at the milli-second frequency and comprises indicative quotes for the best bid and ask price. Further, the database stores the name and location of the dealers that are active in spot and swap markets and submit their quotes. The detailed track record allows me to conduct a comprehensive analysis of price and quantity dynamics of quote submissions in both, spot and derivative, markets. The main empirical analysis is conducted at the intra-day frequency leveraging price information with information on how many dealers were active, how many quotes were submitted, how quote submissions varied within a specific time horizon, and how all these metrics differed by dealer-type.

The main results are as follows. First, I find a robust relationship between FX funding and FX market liquidity. A deterioration in FX funding liquidity, measured by the widening of CIP deviations or simple FX swap spreads (forward discount) is associated with a widening of bid-ask spreads in both currency swap and spot markets.

Second, this link between FX market and FX funding liquidity conditions strengthened significantly since about mid-2014. In particular, while some tightening in FX swap market liquidity was always present around quarter-ends, these effects have become several

[5] This is because FX swap markets and GC repo markets rely on arbitrage-trading and market-making by banks, whereas the Tri-party repo market gets US dollar supply also from real money investors not subject to the leverage ratio.

times larger since 2014, with significant spillover to spot market liquidity also emerging in the latest period. The regime shift in the liquidity conditions in FX market appears related to the re-occurring liquidity droughts at quarter-ends. During each month corresponding to a quarter-end, controlling for higher variation in liquidity metrics, I find that FX funding and market liquidity exhibit stronger co-movement. Because the origins of quarter-end anomalies can be traced to core bank funding markets, such as unsecured overnight markets and repo markets, this implies stronger transmission of exogenous FX funding liquidity shocks to FX market liquidity. Statistical tests indeed point towards spillover of adverse FX funding liquidity shocks to market liquidity.

Third, I find that liquidity conditions and dealer activity are closely related. Specifically, the positive impact of dealer competition on FX market liquidity has decreased over time. While large dealers still dominate as market-makers in spot and their quoting intensity is associated with improved liquidity dynamics, they have also exhibited a tendency to pull-back from making markets in FX swaps around balance sheet reporting periods. As balance sheet window dressing by major dealers causes them to scale back their activity, funding costs significantly increase, market liquidity declines and volatility increases. For example, I find that funding costs at quarter-ends are approximately three times larger between July 2014 and May 2017 than during the European debt crisis.

Fourth, and related, I find that market-making activity in FX swaps can significantly diverge from that in spot. Whereas large dealers appear to dominate as principal market-makers in spot throughout most of the sample period, small dealers largely displace large dealers as market-makers in swaps in times when spreads are wide. Specifically, when liquidity conditions tighten due to the pull-back by G-SIBs from FX swap markets at quarter-ends, small dealers increase their quoting activity. In certain year-end periods, I also find evidence that smaller dealers are consistently quoting inside spreads, hence are the ones making-markets on average. Yet, small dealers stepping in during these times does not appear to lead to an improvement in liquidity conditions. I identify two reasons for this. One is that small dealers are low volume players, thus require wider bid-ask spreads and forward spreads for their market-making activity to be profitable. The second reason is that quoting activity by small dealers does not contribute to the same extent to price discovery as that by large dealer. Specifically, greater quoting intensity by small dealers does not suppress the dispersion of forward rate quotes in the same way that quoting intensity by large dealers does, indicating greater volatility of quotes around the "true" forward rate, constructed from the sum of spot rate and swap points, in an any given hour.

Finally, consistent with the price spillovers from forward points in FX derivative markets to bid-ask spreads in both swap and forward rates (eg. FX funding to market liquidity spillovers), starting July 2014 for JPY/USD (January 2015 for EUR/USD) I also find that heightened activity by smaller dealers in FX swap markets has a stronger effect on market liquidity (bid-ask spreads) in the *spot market* compared to smaller dealers in spot. This is noteworthy, especially because more than half of small dealers in FX swaps do not even participate in the spot market directly. In certain times, wider spreads quoted by small dealers in the swap market negatively impact spot market liquidity, and offset some of the positive effects on spot market liquidity from competition by large dealers.

Hence, the analysis suggests that funding liquidity has become a more important economic factor to understand bid-ask spreads in FX spot. I also show that dealer structure of FX markets has been changing over the span of the sample period, and that smaller banks appear to act more frequently as market makers in FX swaps as G-SIB dealers pull back. This adds nuance to the widely held view that FX liquidity provision is highly concentrated among a handful of largest dealers (King et al., 2011), while smaller dealers operate an agency model simply passing client flows into the wholesale FX market (Moore et al., 2016). While such concentration of liquidity provisions among the largest dealers appears to largely hold for spot markets, in the markets for FX swaps smaller dealers tend to turn to making markets and providing liquidity at times when spreads are wide enough to meet their hurdle rates. Since smaller dealers charge higher mark ups, their increased quoting activity does not necessarily lead to a narrowing of funding and market liquidity spreads, instead allowing the deterioration in FX liquidity conditions to persist until large dealers re-enter as market-makers as the quarter-end turn passes. Overall, this means that window dressing by large FX dealers in FX swaps has been disruptive not only to swap market liquidity but also to liquidity in spot.

This paper proceeds as follows. Section 2 describes the data and the measures of liquidity and dealer activity. Section 3 contains broad overview of liquidity measures at daily frequency. Section 4 contains the core intra-day analysis of FX liquidity dynamics. Section 5 concludes.

4.2 Data and Variable Definitions

I obtain data for JPY/USD and EUR/USD spot exchange rate and 1-month swap points from Reuters Datascope for the sample period 1st February 2010 to 31st May 2017. The dataset contains information on dealers' best bid and ask quote submissions, timed at the milli-second frequency. In addition, it documents the name and location of the dealer bank that submitted the quote. I also obtain information on 1-month overnight indexed swap rates for both countries. Table 4.1 shows the sample of tick history data for a two second window for spot JPY/USD.[6]

Table 4.1: Example: Two-second window for JPY/USD spot rate

RIC	Date	Time	Dealer	Bid	Ask
'JPY='	'5-May-17'	'14:47:29.348944'	'BKofNYMellon NYC'	112.620003	112.6399994
'JPY='	'5-May-17'	'14:47:29.381124'	'BARCLAYS LON'	112.610001	112.6399994
'JPY='	'5-May-17'	'14:47:29.640943'	'SOC GENERALE PAR'	112.599998	112.6399994
'JPY='	'5-May-17'	'14:47:30.065053'	'KASPI BANK ALA'	112.620003	112.6399994
'JPY='	'5-May-17'	'14:47:31.277082'	'SEB STO'	112.599998	112.6500015
'JPY='	'5-May-17'	'14:47:32.260157'	'RBS LON'	112.599998	112.6399994
'JPY='	'5-May-17'	'14:47:32.301189'	'RABOBANKGFM LON'	112.589996	112.6399994

Note: Data refers to a two-second window on 5th May 2017. The information is obtained from Thomson Reuters Tick History (TRTH) database and accessed via Reuters Datascope.

Since quotes are submitted to and documented by Reuters in irregular time intervals, I transform the raw data from the milli-second frequency to 1-min time series, using the last submitted ask and bid quote in each minute. I consider all submitted quotes, irrelevant if a certain dealer submitted more than one quote during each minute. Huang and Masulis (1999) refer to this methodology as quote-weighted price data. I keep a detailed record of

[6] While containing important information on quoting activity by FX dealer banks, the dataset is also subject to a number of limitations. First, it is primarily based on quote submissions on the Thomson Reuters Matching platform, which, together with EBS, only represent about 13% of global spot FX trading volume and 12% of global FX swaps trading volume, according to BIS (2016). Second, the data only has information on quotes and not traded prices or volumes, which precludes us from computing a number of popular measures of market liquidity based on the volume-return relationship. Third, the dataset does not contain information on the depth of the order-book, and the observed quotes are top-of-book quotes. Lastly, while Reuters is the main trading platform for commonwealth and emerging market currency pairs, for EUR/USD and JPY/USD it is EBS. As these are the two most frequently traded exchange rates, however, I believe it is pivotal to shed light on the link between liquidity dynamics and dealer activity in spot and swap market of these two currency pairs. Breedon and Vitale (2010) show that dynamics between EBS and Reuters are highly correlated and both markets are closely linked with each other.

the number of submitted quotes and I identify the number of unique dealers that are active in each 1-minute time-interval. In very few cases, in which no quote submission took place, I use the last available information to fill the gap. Next, while activity on FX markets is not restricted to specific trading hours, I clean the data in the spirit of earlier studies (e.g. Andersen, Bollerslev, Diebold, and Vega, 2003) and exclude certain trading hours and holidays. On weekends and in the occasion of a holiday, I delete data entries between 21:00:00 (GMT) of the previous day until 21:00:00 (GMT) of the holiday itself. For example, I drop information on weekends from Friday 21:00:00 until Sunday 21:00:00. We drop data on fixed holidays such as Christmas (24th - 26th December), New Year's (31st December - 2nd January) and July fourth (4th July).[7] In addition, I exclude flexible holidays, such as Good Friday, Easter Monday, Memorial Day, Labour Day, and Thanksgiving and the day after.

I obtain equally spaced time-series bid and ask prices for spot rate, 1-month swap points, and overnight index swap (OIS) rates, number of quote submissions, and number and names of active dealers. Lastly, I convert the series to the hourly frequency to reduce the impact of market microstructure noise. For the entire sample period, I obtain 44,088 observations.

4.2.1 Price Measures Of Market Liquidity

After these steps of data cleaning, I construct the following variables. Spot dealers quote spot bid and spot ask prices, FX swap dealers quote bid and ask swap points. Following Banti and Phylaktis (2015), I measure market liquidity at the hourly frequency in the foreign exchange spot and forward market by the bid-ask spread

$$Spread_h^S = \frac{S_h^{ask} - S_h^{bid}}{S_h^{mid}} \tag{4.1}$$

$$Spread_h^F = \frac{F_h^{ask} - F_h^{bid}}{F_h^{mid}} \tag{4.2}$$

where the mid-price is calculated as the arithmetic average between ask and bid price in each respective market segment. The bid and ask forward exchange rates are implied by the swap points quoted by dealers in FX swaps. We define the 1-month forward rate implied

[7] In 2015, the official holiday is 3rd July, since July 4th falls on a Saturday.

by swap points: $F = S + SP * 10^{-2}$ for USD/JPY and $F = S + SP * 10^{-4}$ for USD/EUR, where S denotes the spot rate and SP are 1-month swap points.

4.2.2 Price Measures Of Funding Liquidity

Swap point quotes from FX swap dealers contain another important piece of information. For example, if the reported swap points are negative, this indicates that USD is trading at a forward discount. Hence, the pricing of FX swaps reflects the costs of obtaining say USD today at the spot rate S in exchange for say JPY, and reversing this transaction in one month at the pre-agreed forward exchange rate F. This, effectively, represents the cost of term funding of one currency against another.

Hence, this main measure of FX funding liquidity is based on the forward spread, which I calculate as:

$$Fdiscount_h = \frac{F_h^{mid} - S_h^{mid}}{S_h^{mid}} \tag{4.3}$$

where F^{mid} and S^{mid} refer to the mid-price 1-month forward and spot rates, respectively.

As an alternative measure of funding liquidity, I adjust the forward spread (forward discount) by the level of benchmark interest rates, OIS rates, in the two currencies of the same maturity. This is because, over a longer horizon, the level of the forward-spot differential should change to reflect the relative interest rate differentials in the two currencies, as stipulated by the covered interest parity (CIP). Hence, an alternative measure of FX funding liquidity is based on annualising the implied 1-month interest in the raw forward discount, then adjusting it by the OIS rates in the two currencies. Effectively this comes down to computing deviations from CIP:[8]

$$CIPdev_h = (1 + \frac{r_h^{mid}}{100}) - (1 + \frac{r_h^{mid*}}{100}) \times \left(\frac{F_h^{mid}}{S_h^{mid}}\right)^{(360/30)} \tag{4.4}$$

[8] That said, adjustment of the forward discount by the OIS rates should not be considered as a measure of CIP arbitrage profits (see, for example, Rime, Schrimpf, and Syrstad, 2017), but is simply used to account for the relative cost of funding liquidity via FX swaps in the two currencies taking into account the level of benchmark interest rates.

where r_h^{mid} and r_h^{mid*} refer to the mid-price OIS rates of both currencies.

By now, it should be fairly obvious that the pricing of FX swaps is reflective of both market and funding liquidity.[9] First, the quotes for swap ask (bid) points are the quotes for the differential between ask (bid) spot and ask (bid) forward rate, thus implying a price for both. Second, the forward discount implicit in the swap points provides a measure of term funding of one currency against another.

4.2.3 Quantity Measures Of FX Liquidity

While these price-based measures are used to explore the relationship of liquidity dynamics, I use the following additional quantity-based measures that account for FX dealer structure and quoting activity. First, following Huang and Masulis (1999), I measure dealer competition by tracking the total number of quote submissions. I do this not only for spot, but also for forward points, with the number of quotes per hour denoted by Q_h^S and Q_h^F, respectively. In addition, I construct a measure of dealer competition at the extensive margin by counting the total number of active unique dealer banks within each hour. I denote this measure as N_h^S and N_h^F for spot and forward points, respectively. I treat all dealers from one bank, independently of their branch's location, as one market participant. It is worth noting the difference between these two measures. While former one accounts for the quoting activity of banks, the later only takes into account the actual number of banks active in the market. Lastly, I combine these two variables and measure quoting intensity as the ratio of submitted quotes and active banks $(\frac{Q_h^S}{N_h^S})$ and $(\frac{Q_h^F}{N_h^F})$. I interpret this measure as indicator of dealer competition at the intensive margin. An overview of the measures is provided in Table 4.2.

4.2.4 Large Versus Small Dealers

Since I know what dealer is active in the market, I distinguish between small and large dealers as an additional dimension of the analysis. An earlier study that looked at the size dimension is Phylaktis and Chen (2010). These authors relied on the ranking of the Annual Euromoney FX Survey (EMS) to make a dealer classification by size. An alternative approach is to distinguish between large and small dealers in view of recent policy implementations and rely on the classification of global systematically important banks (G-SIBs) by the Financial

[9]See Baba et al. (2008) for an exposition of cash flows in an FX swaps.

Table 4.2: Benchmark hourly and daily measures

Measure (hourly)	Definition	Daily
$Spread^S_h = \frac{S^{ask}_h - S^{bid}_h}{S^{mid}_h}$	market liquidity	mean
$Spread^F_h = \frac{F^{ask}_h - F^{bid}_h}{F^{mid}_h}$	market liquidity	mean
$Fdiscount_h = \frac{F^{mid}_h - S^{mid}_h}{S^{mid}_h}$	FX funding liquidity	mean
$CIPdev_h = (1 + \frac{r^{mid}_h}{100}) - (1 + \frac{r^{mid*}_h}{100}) \times \left(\frac{F^{mid}_h}{S^{mid}_h}\right)^{360/30}$		mean
$Q^P_h = \#Quotes^P_h$	dealer competition, *intensive margin*	sum
$N^P_h = \#Dealers^P_h$	dealer competition, *extensive margin*	sum
Q^P_h / N^P_h	dealer competition, *quoting intensity*	sum
$Disp^P_h = \sqrt{\sum_{i=1}^{h_i} \frac{q_i}{Q_h} \left(\frac{P_i - \bar{P}_h}{P_T}\right)}$	weighted quote dispersion; $P \equiv S, F$	mean
$Vol^P_h = \frac{\sum (r_P - \bar{r}_{P,h})^2}{n-1}$	variance; $r_P = ln(P_h) - ln(P_{h-1})$; $P \equiv S, F$	-

Stability Board, (BIS, 2011, 2013a). Table 4.3 shows the comparison between the thirty G-SIB bank dealers and the top FX dealers according to the 2016 Euromoney survey. While almost all 30 G-SIBs are ranked as top FX dealers in the Euromoney Survey (Table 4.3, left column). [10] 27 additional banks would be considered as large dealers by the Euromoney FX Survey but are not included in the list of G-SIBs.

Figure 4.1 shows daily time-series of the percentage share of all top-of-book quotes for the spot rate and forward points JPY/USD of large dealers, $Q^{S,L}_t / Q^S_t$ and $Q^{F,L}_t / Q^F_t$, classified according to the G-SIB designation versus dealers that are only part of the Euromoney Survey. During the entire sample period, large dealers categorised according to the G-SIB classification were responsible for 37.2% of daily spot quote submissions, on average, whereas using the broader Euromoney Survey for classification would raise that share to 49.8%. Figure 4.1 points towards an increased concentration of spot liquidity provision by tier-1 and tier-2 dealers. The share of spot quotes in JPY/USD submitted by large dealers approximately doubles between 2011 and 2014. Yet, as both figures show, the trends of both

[10]The only exception is Mizuho FG, which is classified as G-SIB but not listed in the Euromoney Survey. Note that non-bank dealers such as Citadel Securities, XTX Markets, Tower Research Capital, or Virtu Financial do not appear in the database by name. This is because their access to Reuters Matching trading platform is prime-brokered by major FX dealer banks. Therefore, the quotes of such non-bank market-makers on Reuters Matching appear under their prime-broker's name and not their own. Dealers that are not appear in the database are marked in grey in Table 4.3.

Table 4.3: G-SIB classification vs Euromoney FX Survey rankings

This table reports the comparison between large dealer categorisation based on G-SIB classification with that based on Euromoney FX Survey rankings in 2016.

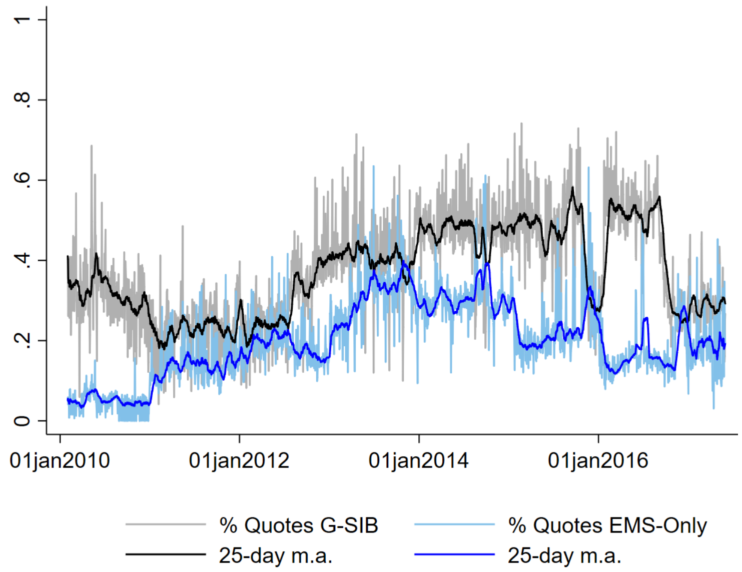
G-SIB Classification	Euromoney FX Survey (EMS)
Agricultural Bank of China	Alfa Bank
Bank of America Merrill Lynch	ANZ Banking Group
Bank of China	Bank of Montreal
Bank of New York Mellon	BBVA
Barclays	CIBC
BNP Paribas	Citadel Securities
China Construction Bank	Commerzbank
Citigroup	Commonwealth Bank of Australia
Credit Suisse	Danske Bank
Deutsche Bank	Den norske Bank
Goldman Sachs	Jump Trading
Groupe BPCE	Lloyds Banking Group
Groupe Credit Agricole	Lucid Markets
HSBC	National Australia Bank
ICBC	Natixis
ING Bank	Nomura
JP Morgan Chase	Rabobank
Mitsubishi UFJ FG	RBC Capital Markets
(Mizuho FG)	Saxo Bank
Morgan Stanley	Scotiabank
Nordea	SEB
Royal Bank of Scotland	TD Securities
Santander	Tower Research Capital
Societe Generale	Virtu Financial
Standard Chartered	Westpac Banking
State Street	XTX Markets
Sumitomo Mitsui FG	Zurich Cantonalbank
UBS	
Unicredit Group	
Wells Fargo	

Notes: Banks that are classified as large dealers according to G-SIB classification are also considered as large dealers according to the Euromoney FX Survey (EMS). Banks marked in grey are not available in our database and banks that are only part of the G-SIB classifaction but not listed in EMS are marked with parenthesis.

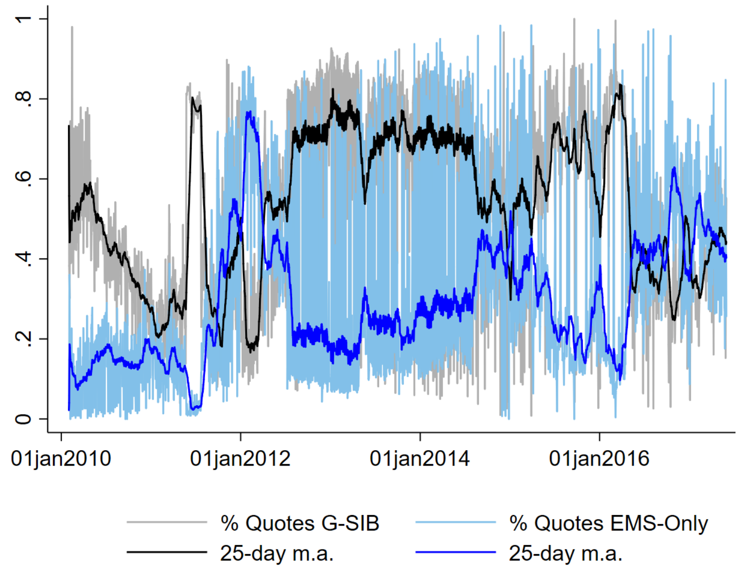
classifications are very different from each other. First, quoting activity by tier-1 dealers becomes increasingly volatile in the second half of the sample. Second, in periods when tier-1 dealers (G-SIBs) pull back and decrease their market making activity (e.g. December

Figure 4.1: Large and small dealer characteristics in JPY/USD

(a) Daily percentage quotes from large dealers in spot: G-SIB vs Euromoney classification



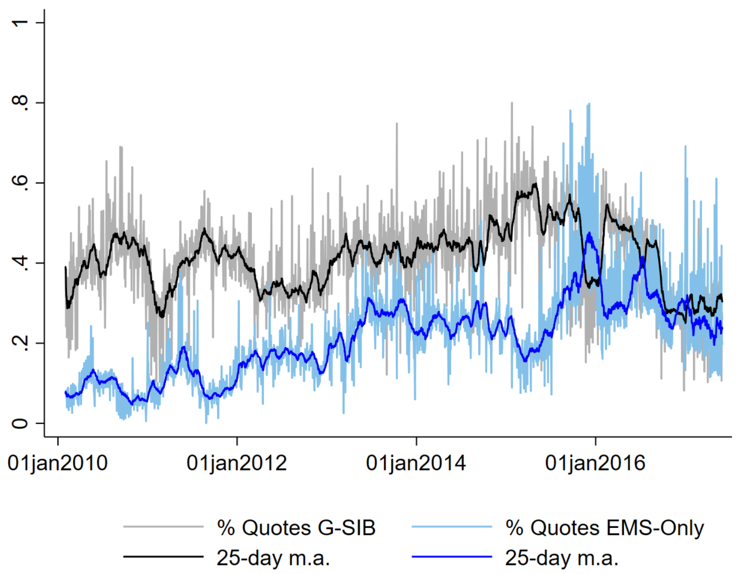
(b) Daily percentage quotes from large dealers in swap markets: G-SIB vs Euromoney classification



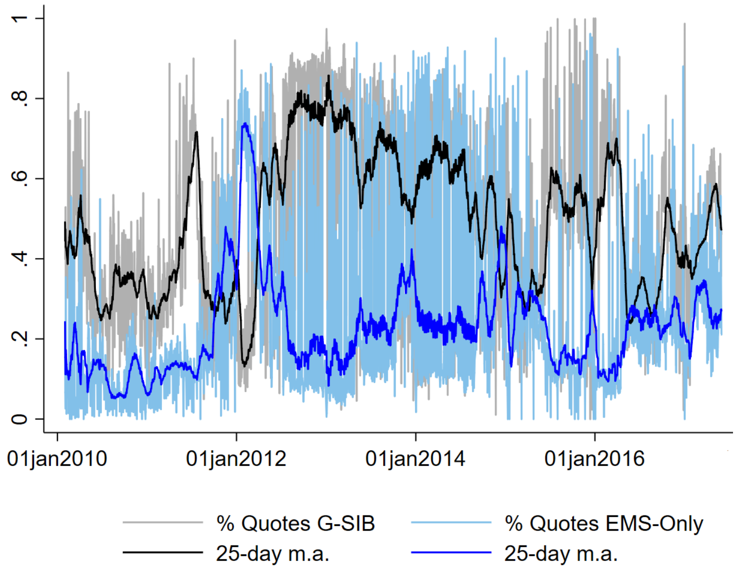
See Table 4.1 for the corresponding list of dealers. G-SIB (tier-1) dealers refer to banks that are classified as globally systematically important banks. EMS are banks that are listed as large dealers by the FX Euromoney Survey but that are not part of the G-SIB classification.

Figure 4.2: Large and small dealer characteristics in EUR/USD

(a) Daily percentage quotes from large dealers in spot: G-SIB vs Euromoney classification



(b) Daily percentage quotes from large dealers in swap markets: G-SIB vs Euromoney classification



See Table 4.1 for the corresponding list of dealers. G-SIB (tier-1) dealers refer to banks that are classified as globally systematically important banks. EMS are banks that are listed as large dealers by the FX Euromoney Survey but that are not part of the G-SIB classification.

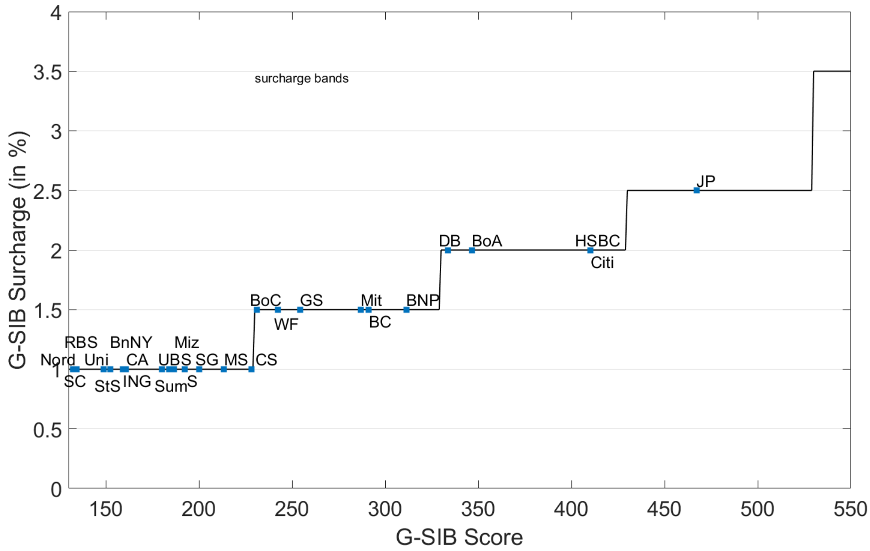
2016), tier-2 dealers increase the number of quote submissions.

Moving to FX swap markets, the substitution effect between tier-1 and tier-2 quotes is even stronger than in spot, as indicated by the counter cyclical quoting activity of tier-1 versus tier-2 dealer activity (Figure 4.1b). The decreasing quoting activity in FX swap points at general FX dealer aversion to making-markets in FX derivatives around regulatory reporting periods, because, unlike cash, derivatives exposures are costly in terms of capital and collateral requirements under Basel III, and national leverage and liquidity regulations. Tier-1 dealers are also subject to the annual G-SIB surcharge and are choosing to actively manage down their balance sheets to avoid crossing into the next G-SIB bucket.[11] The incentives to window-dress balance sheet would be strongest for banks whose end-of-year G-SIB score puts them closest to the next G-SIB bucket for the following year. Figure 4.3 illustrates this using end-2016 G-SIB scores for the banks that fall into this category in the JPY/USD sample, and what this implies about their proximity to the higher G-SIB bucket, and hence a higher capital surcharge beginning 2017. Similar dynamics for EUR/USD are shown in Figure 4.2.

During the entire sample period, large dealers categorised according to the G-SIB classification where responsible for 52.2% of daily quote submissions, on average. Whereas, using the broader Euromoney Survey for classification would raise that share to 73.5%. This implies that the majority of dealers classified as small in the swap market are still the type to make it into the Euromoney Survey, unlike most of the small dealers in the spot market. Hence, it may be more accurate to think of small dealers in FX swaps as tier-2 dealers, while many more small dealers in spot are better thought of as tier-3 dealers, because they would not feature in the Euromoney Survey.

[11] See, for example, J.P. Morgan "Making sense of Libor's mysterious rise", North American Fixed Income Strategy, 14 December, 2017.

Figure 4.3: G-SIB surcharge and bucket cut-off points (2016/2017)

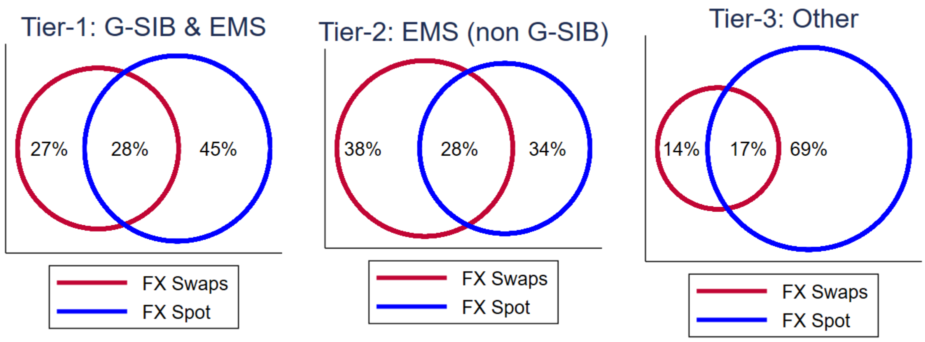


The figures show G-SIB banks included in the sample for JPY/USD as they are positioned according to their G-SIB score at end-2016 and their proximity to the next G-SIB bucket as of beginning 2017.

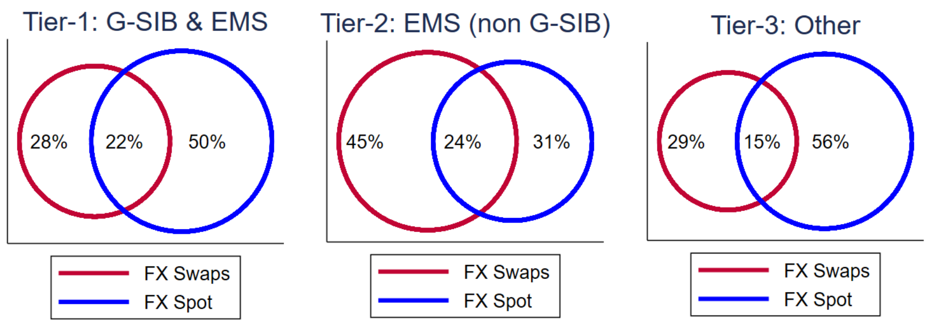
Indeed, the comparison of Figures 4.4a and 4.4b shows that most small (tier-3) dealers are only active in the spot market in both sub-sample periods, while tier-2 banks increase their activity in FX swaps markets since about mid-2014. In contrast, tier-1 dealer behaviour changes from quoting in both markets to greater liquidity provision in spot markets only. I interpret the Venn diagrams as further evidence that quoting activity between tier-1 and tier-2 banks differ significantly from each other in spot and swap markets, with tier-1 dealer shifting away from quoting FX swaps to only quoting spot as bank adopted to the new regulatory reporting templates as of January 2015. As Figure 4.4c shows, the tier-1 dealer shifting away from quoting both FX forward points and spot (e.g. FX swaps) to only quoting spot are particularly pronounced around the quarter-end and year-end regulatory reporting periods.

Figure 4.4: Activity of small and large dealers by market segment in JPY/USD

(a) Average share of dealers active in spot and derivative markets: Feb 2010 - Jun 2014



(b) Average share of dealers active in spot and derivative markets: Jul 2014 - May 2017



(c) Percentage of large dealer activity in spot and derivative markets

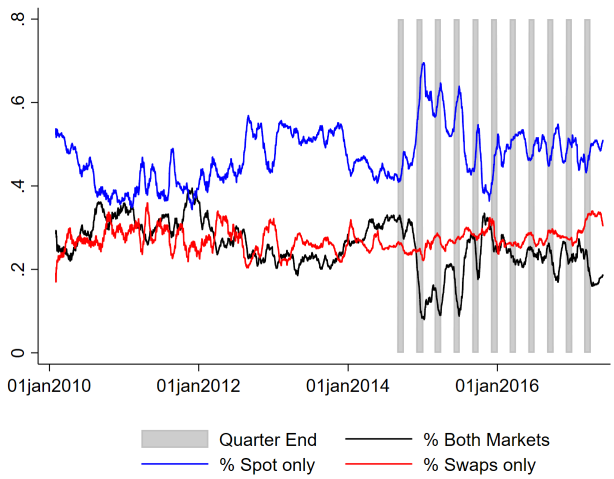
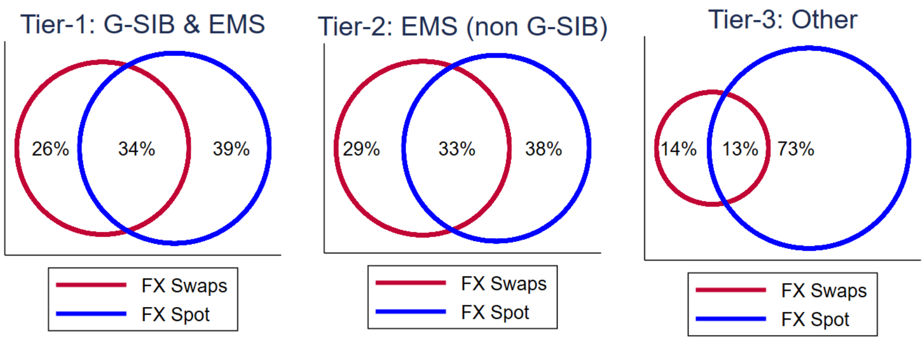


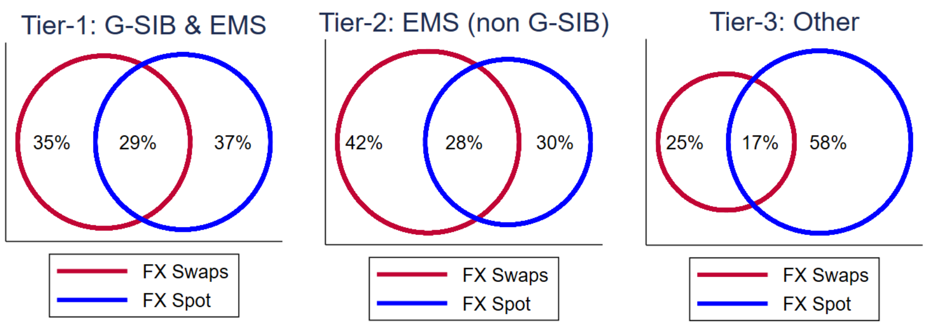
Figure 4.4a and 4.4b show the average share of dealers active in different market segments for the period February 2010 - June 2014 and July 2014 to May 2017, respectively. The red line refers to dealers that are only active in the swap market, the blue line refers to dealers active only in spot markets, and the intersection refers to dealers that are active in both markets. Figure 4.4c shows the 25-day moving average of large dealers (tier-1) that are only active in spot markets, only in swap markets, and in both markets.

Figure 4.5: Activity of small and large dealers by market segment in EUR/USD

(a) Average share of dealers active in spot and derivative markets: Feb 2010 - Jun 2014



(b) Average share of dealers active in spot and derivative markets: Jul 2014 - May 2017



(c) Percentage of large dealer activity in spot and derivative markets

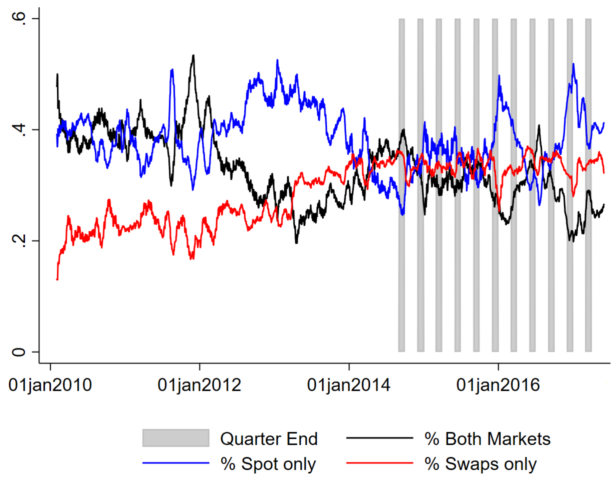


Figure 4.4a and 4.4b show the average share of dealers active in different market segments for the period February 2010 - June 2014 and July 2014 to May 2017, respectively. The red line refers to dealers that are only active in the swap market, the blue line refers to dealers active only in spot markets, and the intersection refers to dealers that are active in both markets. Figure 4.4c shows the 25-day moving average of large dealers (tier-1) that are only active in spot markets, only in swap markets, and in both markets.

4.3 Liquidity Measures In The Long-Run

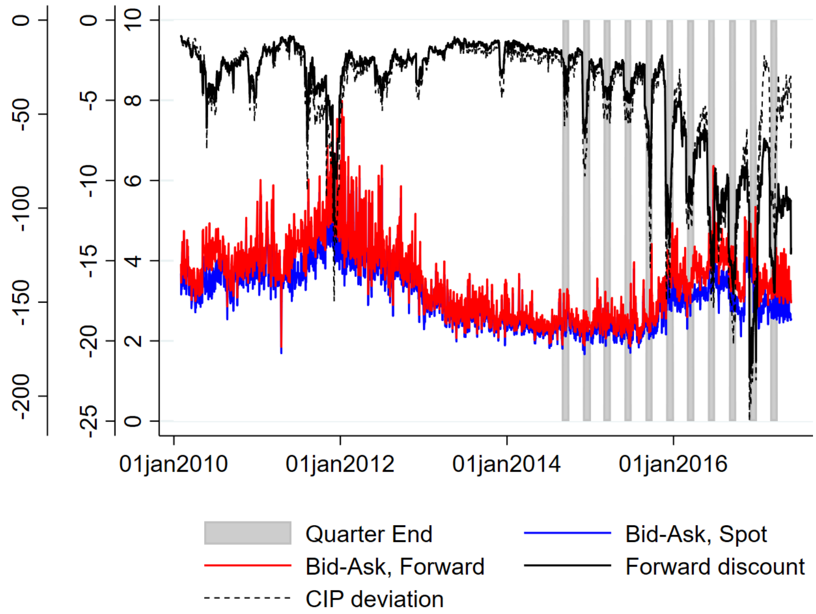
This section looks at daily trends in liquidity and dealer competition measures over the sample period February 2010 to May 2017. Figures 4.6a and 4.6b show the dynamics of the price-based liquidity measures for JPY/USD and EUR/USD, respectively. Market liquidity in the spot and swap markets move very closely (correlation of 0.97 for JPY/USD and 0.98 for EUR/USD) over most of the sample period. For both currency pairs, bid-ask spreads increase during the European debt crisis at the end of 2011 and beginning of 2012. They remain comparably low afterwards and start to increase gradually for both currencies from mid-2014 until the end of the sample. Since a higher bid-ask spread is associated with more illiquid market conditions, both Figures seem to suggest that market liquidity has declined towards the end of the sample.

For funding liquidity, I observe a similar pattern. As indicated by the black plot, there is an increase in (absolute) forward discount in the middle of 2011 and relatively stable funding costs in the period after the European debt crisis. From the third quarter in 2014 onward, funding liquidity drops on a re-occurring basis. These liquidity droughts are particularly prevalent during quarter-end periods, indicated by the grey areas. Clearly, the level of funding liquidity since about mid-2014 follows a different pattern compared to the 2010 to 2014 period.

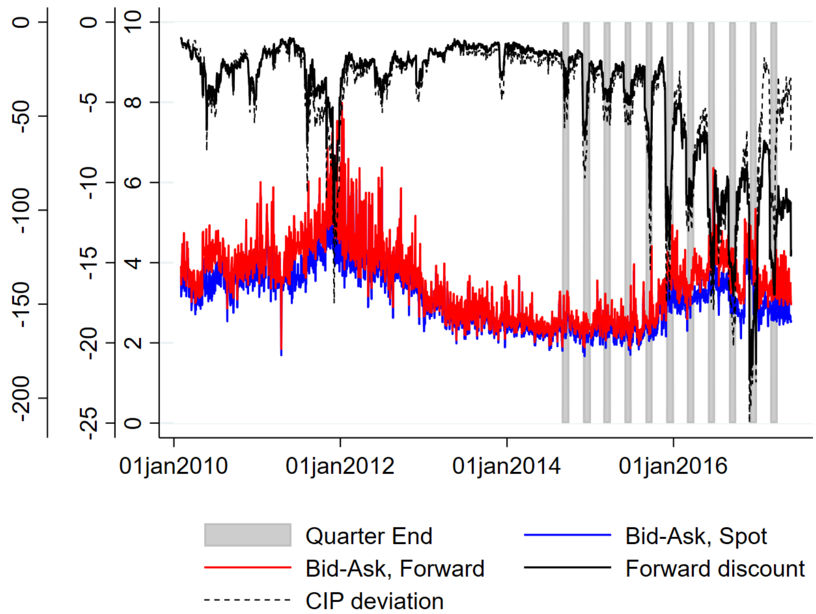
Table 4.4 shows summary statistics for price-based FX liquidity measures across the months falling on quarter-ends (QE), versus those on month before (BQE) and after (AQE). The sample is split according to the apparent regime change in FX liquidity conditions with the emergence of the quarter-end turn in forward points in September 2014 for JPY/USD and March 2015 for EUR/USD; hence I pick the second sub-sample cut-off two month prior to the QE month for each currency pair. For each liquidity measure, Panel A shows the average level as well as p-values of a one-sided t-test that the liquidity measures are significantly worse than in the rest of the months in each sub-sample. Panel B shows the volatility of each liquidity measure as well as the p-values of the variance ratio test that volatilities are significantly different compared to the rest of the months in each sub-sample.

Figure 4.6: Bid-ask spreads in spot and forward rate, forward discount, and CIP deviations

(a) JPY/USD



(b) EUR/USD



The outside y-axis shows OIS-based 1-month CIP deviations, in basis points; the middle y-axis shows 1-month forward discount, in basis points as a percentage of spot price; the inner y-axis shows bid-ask spreads, in basis points as a percentage of mid-price.

Table 4.4: Summary Statistics: Liquidity dynamics (price-based) at quarter end

This table reports the average and standard deviation of spot market and swap market liquidity, measured by the bid-ask spread of spot rate ($Spread^S$) and forward rate ($Spread^F$), and funding liquidity, measured by forward discount ($Fdiscount$) and CIP deviations ($CIPdev$) for JPY/USD and EUR/USD for two different sub-sample periods. QE refers to months at quarter-ends (March, June, September, December), BQE are months before quarter-end (February, May, August, November), and AQE refer to the first month after quarter-end (January, April, July, Octobers). In Panel A, numbers in parentheses refer to the p-value of a one-sided t-test that market and funding liquidity in the respective months is larger than in the rest of the months in each sub-sample. In Panel B, numbers in parentheses denote the p-value of a variance ratio test.

	JPY/USD						EUR/USD					
	02/2010 - 06/2014			07/2014 - 05/2017			02/2010 - 12/2015			01/2015 - 05/2017		
	QE	BQE	AQE	QE	BQE	AQE	QE	BQE	AQE	QE	BQE	AQE
Panel A: Average and t-test												
$Fdiscount$	-2.91***	-2.39***	-2.40***	-9.94***	-6.24***	-6.29***	-0.37***	0.12	0.31	-10.38***	-8.38	-8.47
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.83)	(1.00)	0.00	(1.00)	(0.99)
$CIPdev$	-29.83***	-23.66	-23.50	-91.86***	-45.24	-48.63***	-25.36***	-21.01	-20.97	-65.96***	-40.72	-45.47
	(0.00)	(1.00)	(1.00)	(0.00)	(1.00)	(0.00)	(0.00)	(1.00)	(1.00)	(0.00)	(1.00)	(0.99)
$Spread^S$	3.41	3.47	3.41	2.79*	2.80**	2.65	1.97	1.99	1.95	2.74	2.75	2.70
	(0.70)	0.12	(0.76)	(0.05)	(0.02)	(1.00)	(0.45)	(0.11)	(0.91)	(0.38)	(0.27)	(0.81)
$Spread^F$	3.72	3.73	3.67	3.21***	3.15	3.03	2.18	2.17	2.12	3.02	2.94	2.95
	(0.35)	(0.28)	(0.83)	(0.00)	(0.27)	(1.00)	(0.12)	(0.21)	(0.98)	(0.15)	(0.73)	(0.67)

Table 4.4: Summary Statistics: Liquidity dynamics (price-based) at quarter end

This table reports the average and standard deviation of spot market and swap market liquidity, measured by the bid-ask spread of spot rate ($Spread^S$) and forward rate ($Spread^F$), and funding liquidity, measured by forward discount ($Fdiscount$) and CIP deviations ($CIPdev$) for JPY/USD and EUR/USD for two different sub-sample periods. QE refers to months at quarter-ends (March, June, September, December), BQE are months before quarter-end (February, May, August, November), and AQE refer to the first month after quarter-end (January, April, July, Octobers). In Panel A, numbers in parentheses refer to the p-value of a one-sided t-test that market and funding liquidity in the respective months is larger than in the rest of the months in each sub-sample. In Panel B, numbers in parentheses denote the p-value of a variance ratio test.

	JPY/USD						EUR/USD					
	02/2010 - 06/2014			07/2014 - 05/2017			02/2010 - 12/2015			01/2015 - 05/2017		
	QE	BQE	AQE	QE	BQE	AQE	QE	BQE	AQE	QE	BQE	AQE
Panel B: Standard deviation and ratio test												
$Fdiscount$	1.76 ***	1.29	0.88	5.23 ***	3.89	3.89	2.89	2.73	3.21	5.37 ***	4.13	4.25
	(0.00)	(1.00)	(1.00)	(0.00)	(1.00)	(1.00)	(0.76)	(0.99)	(1.00)	(0.00)	(1.00)	(0.99)
$CIPdev$	21.41 ***	15.43	10.21	44.70 ***	29.75	28.22	0.39	14.09	0.37	0.69	23.33	0.76
	(0.00)	(1.00)	(1.00)	(0.00)	(1.00)	(1.00)	(0.11)	(1.00)	(0.80)	(0.68)	(1.00)	(0.01)
$Spread^S$	0.78	0.83 *	0.79	0.58 ***	0.55	0.46	0.47 *	0.37	0.44	0.84 *	0.66	0.85
	(0.87)	(0.06)	(0.67)	(0.00)	(0.21)	(1.00)	(0.09)	(0.65)	(0.83)	(0.07)	(0.66)	(0.03)
$Spread^F$	0.90	0.95	0.92	0.77 ***	0.69	0.64	21.86 ***	0.45	12.70	39.21 ***	0.69	25.25
	(0.81)	(0.13)	(0.59)	(0.00)	(0.75)	(1.00)	(0.00)	(0.67)	(1.00)	(0.00)	(1.00)	(1.00)

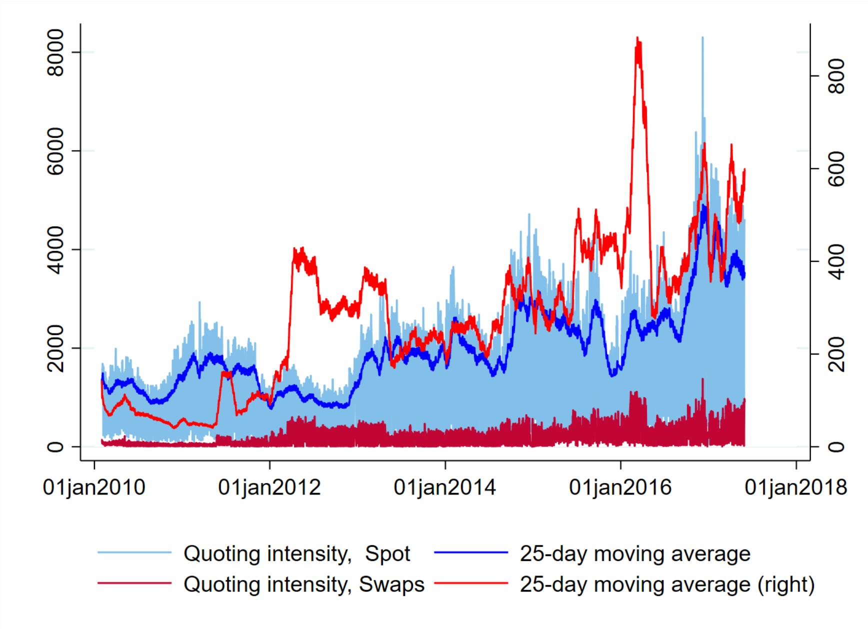
Quarter-end anomalies of the type picked-up in this paper are a recent phenomenon that has emerged since about September 2014 for JPY/USD (March 2015 for EUR/USD). Their origins are exogenous to the FX market as such, attributed to the window dressing by global banks, as some banks shrink their balance sheets so as to manage their regulatory costs associated with the new post-crisis capital and liquidity requirements. Such balance sheet window-dressing appears to have first-and-foremost affected short-term money markets and on balance sheet funding instruments, such as repurchase agreements (CGFS, 2017 and Aldasoro, Ehlers, and Eren, 2018). However, strong effects have also been documented for off-balance sheet instruments, such as FX swaps (see Arai, Makabe, Okawara, and Nagano, 2016 and Du and Verdelhan, 2018).

Takeaways from the results reported in Table 4.4 are as follows. First, FX funding liquidity, as measured by either swap points, $Fdiscount$, or swap points adjusted by the level of benchmark interest rates, $CIPdev$, deteriorates significantly at quarter-end months over the entire period (both wider spreads and higher spread volatility), but the magnitudes of the fall in liquidity at quarter-ends are several times larger in the most recent period. Second, market liquidity in JPY/USD has begun exhibiting significant deteriorations at quarter-ends in the most recent period, as indicated by wider level and volatility of bid-ask spreads in spot and swaps. Market liquidity in EUR/USD appears less affected, although the volatility of bid-ask spreads particularly in the swap market, but also to a lesser extent in spot, has risen.

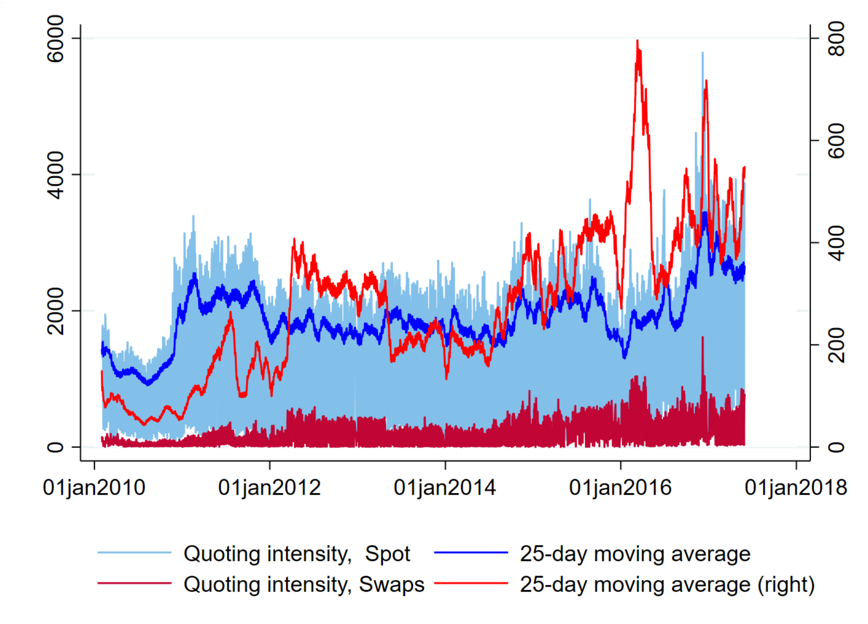
Further, Figures 4.7a and 4.7b show the dynamics of quantity-based measures of market activity. They show FX dealer quoting intensity in spot (blue) and swap market (red) as well as the moving average for both markets. Spot market trading intensity is up to five to ten times higher than in the swap market. In addition, in both markets, I observe an increase in market activity towards the end of the sample. The moving averages indicate a rise in quote submissions compared to the number of banks. The spiking of dealer quoting intensity is particularly pronounced for FX forwards for both currency pairs.

Figure 4.7: Measures of dealer competition in spot and swap markets

(a) JPY/USD



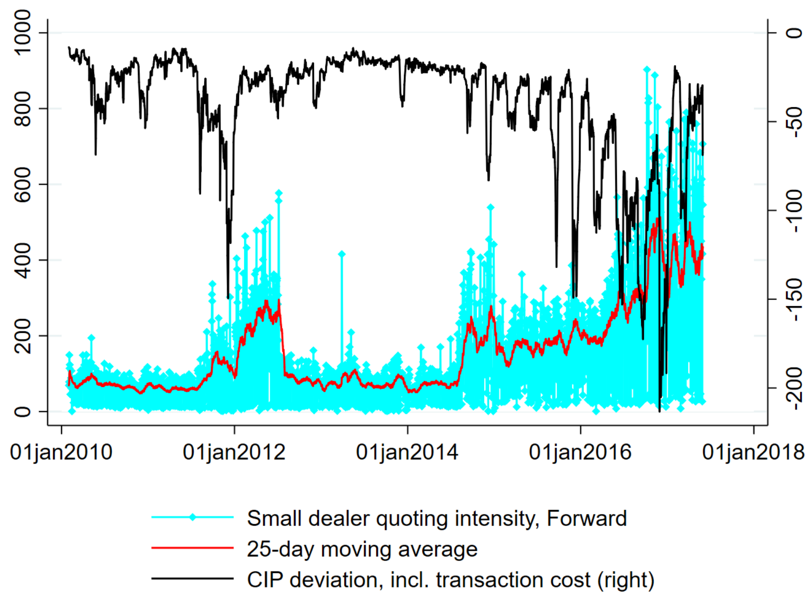
(b) EUR/USD



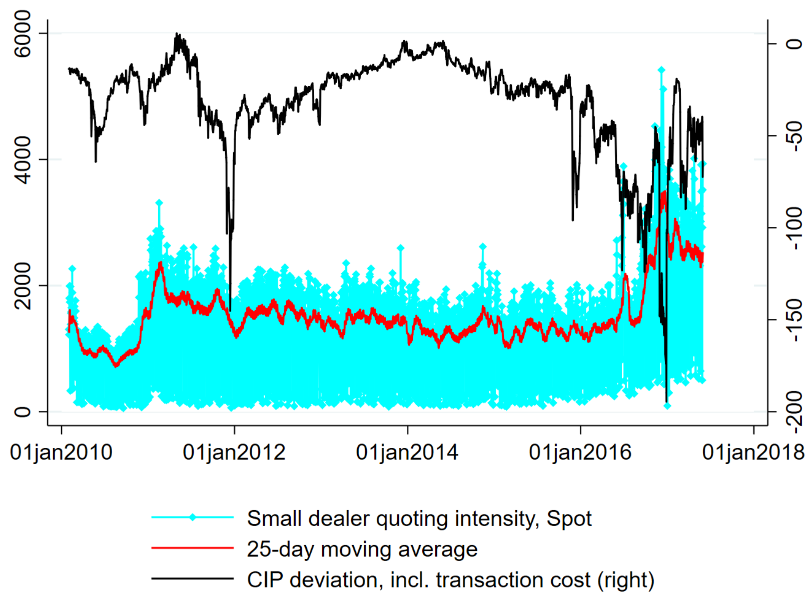
The figure shows daily dealer quoting intensity, defined as the total number of quotes divided by the total number of active dealers in a given day t, in spot and forwards, Q_t^S/N_t^S and Q_t^F/N_t^F; all measures based on the top of the order book.

Figure 4.8: Small dealer quoting intensity in the swap markets

(a) JPY/USD



(b) EUR/USD



The figure shows daily dealer quoting intensity of small dealers in forwards, defined as the total number of quotes divided by the total number of active dealers in a given day t, $Q_t^{F,SD}/N_t^{F,SD}$; all measures based on the top of the order book.

The steep increase in quoting intensity in FX swaps is driven by increased activity of smaller dealers. Figures 4.8a and 4.8b plot the quoting intensity of smaller dealers in the swap market against CIP deviations that take account of transaction costs (measured by bid-ask spreads). As the figures show, whenever bid-ask spreads and forwards spreads widen, as measured by the transaction cost-adjusted CIP deviations, small dealers tend to increase their quoting intensity. This was temporarily the case during the euro area sovereign debt crisis, but became more persistent since about mid-2014 as price-based measures of FX liquidity conditions deteriorated first in JPY/USD and then in EUR/USD. One possibility is that smaller institutions not constrained by G-SIB surcharge or that have greater balance sheet availability are trying to take advantage of the dislocations in FX swap markets. For example, smaller banks could borrow dollar cash in the Fed funds market and then lend it out in FX swaps.

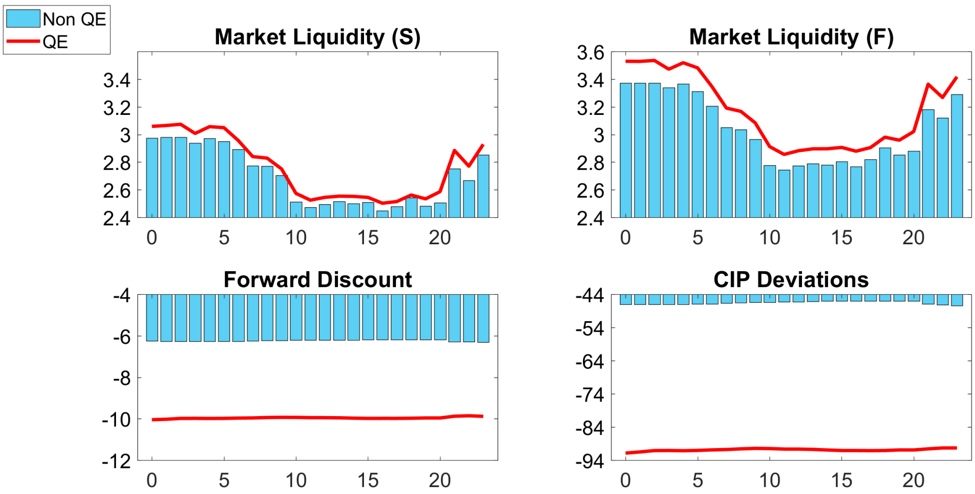
From Figures 4.6 through 4.8 I draw the following conclusions. First, liquidity dynamics, market participation, and trading activity vary to a great extent over the sample period. For example, while market liquidity in spot and swap markets appears to be greatly impacted by the European debt crisis during the first years of the sample, bid-ask spreads and funding costs decrease significantly between 2013 to mid-2014. From mid-2014 until the end of the sample liquidity dynamics are tightening again, despite an overall calm market environment compared to the crisis years. Second, liquidity droughts appear to emerge on a re-occurring basis and are stronger towards the end of a quarter. Third, as is also empirically established later, the rise in total quoting activity by FX dealers is not necessarily associated with an improvement in liquidity conditions.

4.4 Intraday FX liquidity Dynamics

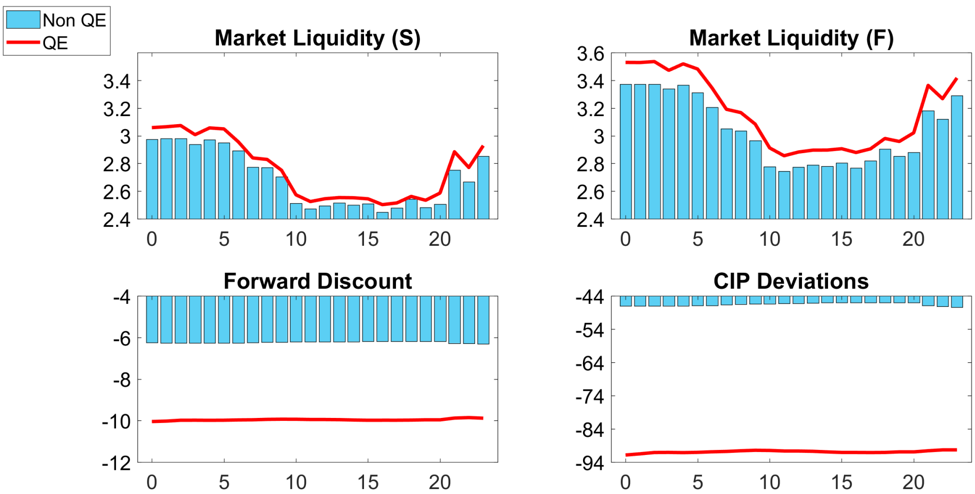
In this section, I move to the analysis at the hourly frequency, using measures constructed from tick-level data and described in Table 4.1. Figures 4.9 and 4.10 show the variation in FX market and funding liquidity in spot and swap markets during the trading hours for the two sub-sample periods for JPY/USD and EUR/USD, respectively. Market liquidity (measured by bid-ask spreads) tends to be lower during the beginning and end of the trading day, resembling a reversed J-shaped form of liquidity (blue bars).

Figure 4.9: Intraday liquidity dynamics: JPY/USD

(a) Feb 2010 - Jun 2014



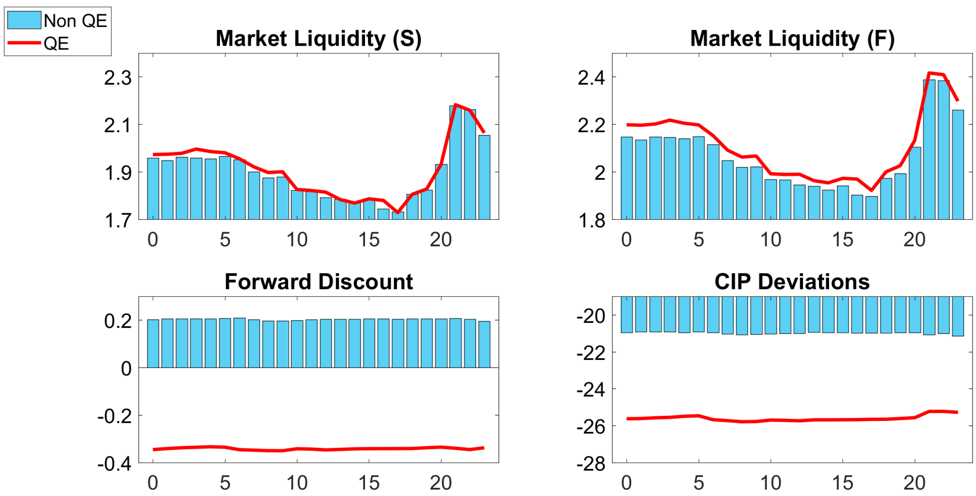
(b) Jul 2014 - May 2017



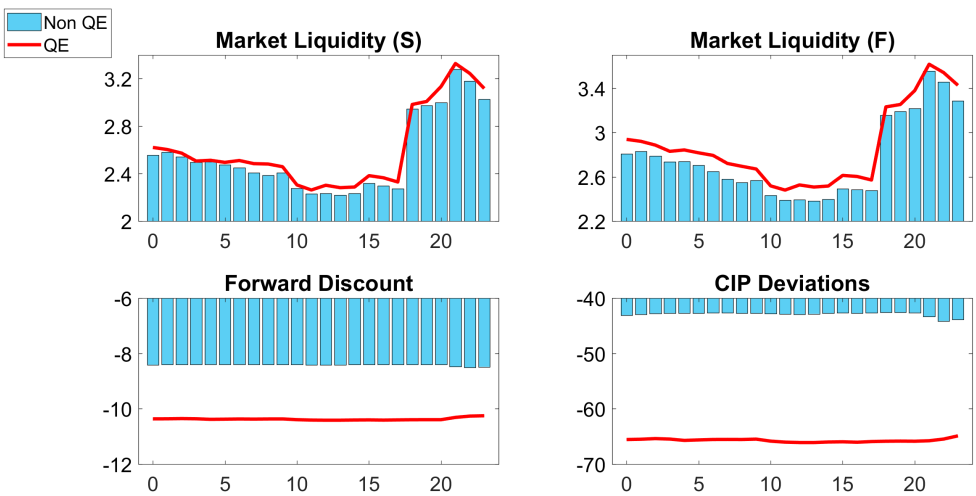
Figures (4.9a) and (4.9b) display average intraday levels of market liquidity in spot and swap markets, measured by spot rate bid-ask spread (S, top left), forward rate bid-ask spread (F, top right), funding liquidity ($Fdiscount$, bottom left) and CIP deviations ($CIPdev$, bottom right). Blue bars refer to non quarter-end months wand red lines refer quarter-end months; GMT time-stamps.

Figure 4.10: Intraday liquidity dynamics: EUR/USD

(a) Feb 2010 - Jun 2014



(b) Jul 2014 - May 2017



Figures (4.10a) and (4.10b) display average intraday levels of market liquidity in spot and swap markets, measured by spot rate bid-ask spread (S, top left), forward rate bid-ask spread (F, top right), funding liquidity ($Fdiscount$, bottom left) and CIP deviations ($CIPdev$, bottom right). Blue bars refer to non quarter-end months wand red lines refer quarter-end months; GMT time-stamps.

Measures of FX funding liquidity, both $Fdiscount$ and $CIPdev$, in turn exhibit an inverted J-shape, also indicating worse liquidity conditions when London and New York based traders are largely absent.

Red lines indicate averages for each trading hour during quarter-end months. As shown in Figure 4.9a, market liquidity changed little during quarter-ends in the February 2010 to June 2014 period, but FX funding liquidity conditions were usually worse. During the second sub-sample period, from July 2014 to May 2017, shown in Figure 4.9b, FX funding liquidity measures are considerably worse in levels (blue bars), and their deterioration at quarter-ends is much larger in relative terms, with spreads in both $Fdiscount$ and $CIPdev$ about two times wider (red lines). Furthermore, unlike the earlier period, bid-ask spreads exhibit widening at quarter ends for both spot at swap markets, indicating possible spillovers from FX funding to FX market liquidity at quarter-ends during the most recent period. Figures 4.10a and 4.10b show qualitatively similar results for EUR/USD.

Statistical tests confirm that intraday co-movement between FX market and funding liquidity has strengthened since the appearance of quarter-end anomalies in funding markets in mid-2014 for JPY/USD and early 2015 for EUR/USD. Table 4.5 shows that pairwise correlations as well as percentage of variation explained by a common factor has increased across all combinations of bid-ask spreads in spot and swaps with FX funding liquidity measures, for both currency pairs.

Table 4.5: Intraday conditional co-movement of liquidity measures

This table reports the average co-movement between spot market and swap market liquidity, measured by the bid-ask spread of spot rate ($Spread^S$) and forward rate ($Spread^F$), and funding liquidity, measured by forward discount ($Fdiscount$) and CIP deviations ($CIPdev$) for JPY/USD and EUR/USD. Funding liquidity is measured either by forward points or by CIP deviations. ρ refers to the average correlation coefficients across trading hours and PCA refers to the proportion of variation explained by the first principal component.

	JPY/USD			
	02/2010 00:00- 06/2014 23:00		07/2014 00:00 - 05/2017 23:00	
	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^F$
$\rho_{Fdiscount}$	-0.44	-0.50	-0.51	-0.63
ρ_{CIPdev}	-0.41	-0.46	-0.47	-0.57
$PCA_{Fdiscount}$	72.22	74.80	75.62	81.66
PCA_{CIPdev}	70.52	73.16	73.73	78.26
	EUR/USD			
	02/2010 00:00 - 12/2014 23:00		01/2015 00:00 - 05/2017 23:00	
	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^F$
$\rho_{Fdiscount}$	0.24	0.27	-0.28	-0.40
ρ_{CIPdev}	-0.34	-0.46	-0.32	-0.43
$PCA_{Fdiscount}$	62.00	63.54	64.16	69.81
PCA_{CIPdev}	67.15	72.85	65.89	71.74

Notes: Hourly sample; GMT time-stamps.

4.4.1 Short- and Long-run Liquidity Dynamics

The descriptive statistics point towards time-varying liquidity dynamics across sub-sample periods. They also indicate that the co-movement between FX market and FX funding liquidity conditions intensified in the last sub-sample period. While funding liquidity has tended to deteriorate at quarter-ends even in the pre-2014 period, these funding liquidity droughts have intensified since mid-2014. Furthermore, it is only in the latest sub-sample period that FX funding liquidity droughts appear to spillover to market liquidity conditions.

To formally examine the relationship between liquidity conditions in spot and swap markets, and the interaction between their market liquidity and funding liquidity components, I estimate a conditional error correction model (ECM), derived from an autoregressive distributed lag model specification for the two sub-sample periods. Following Pesaran et al. (2001) the specification allows to assess the long- and short-run specification between a set of variables independent of the order of integration of the variables in the system. As the dynamics of variables vary across the sample period, displaying mean-reversion in some months but high persistence in others, inferences about non-stationarity from standard unit root tests are highly dependent on the chosen time-period. Modelling the relationship between dealer activity and liquidity in an ARDL model, however, allows me to take an agnostic view about the order of integration, and to model long- and short-run dynamics without classifying variables as either stationary or non-stationary. I formulate the following two conditional ECMs as:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{h-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_h \quad (4.5)$$

where $\mathbf{z_h^S} = (Spread_h^P, |Fdiscount|_h, Q_{LD,h}^P/N_{LD,h}^P, Q_{SD,h}^P/N_{SD,h}^P, Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ is a vector of endogenous variables. LD, SD denote large and small dealers, for both spot and forward points, $P = S, F$. The vector contains bid-ask spread as a measure of market liquidity, forward points as a measure of funding liquidity, quoting intensity of large and smaller dealers, and realized volatility as control variables. α denotes an intercept and the term $\sum_{i=1}^{23} \delta_i H_i$ refers to hourly dummy variables and their associated coefficients. Long-run dynamics are captured by the lagged terms of the dependent and independent variables while short run dynamics are driven by the contemporaneous and lagged differenced terms. I test for the existence of a long-run relationship applying Pesaran et al. (2001) bound testing procedure. First, I test if all long-run coefficients are significantly different from zero using a F-test ($H_0 : \theta_i = 0$). Second, I test if the coefficient of the cointegrating relationship is smaller and significantly different from zero. I estimate the identical specification for every sub-sample period and only vary the number of lags p. Then I examine the significance of the long-run coefficients. If both null hypotheses are rejected, I conclude that there exists a long-run relationship between variables in vector $\boldsymbol{z^S}$ and $\boldsymbol{z^F}$.

Table 4.6: Intraday conditional co-movement of liquidity measures

This table is based on estimation results of the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{h-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_t$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |Fdiscount|_t, Q_{j,h}^P/N_{j,h}^P, Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and $j = LD, SD$, denote quoting activity of large and small dealers, respectively. The coefficients are scaled by the standard deviation of the explanatory variables in each sub-sample.

	JPY/USD					
	02/2010 00:00 - 06/2014 23:00			07/2014 00:00 - 05/2017 23:00		
Variable	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|Fdiscount\|$	0.380***	0.213***	0.193***	0.416***	0.301***	0.649***
Q_{LD}^F/N_{LD}^F	-0.287***			-0.100***		
Q_{SD}^F/N_{SD}^F	0.350***		0.202***	0.180***		0.311***
Vol^F	0.183***			0.427***		
Q_{LD}^S/N_{LD}^S		-0.180***	-0.148***		-0.064***	-0.205***
Q_{SD}^S/N_{SD}^S		0.124***			0.039	
Vol^S		0.143***	0.154***		0.382***	0.271***

	EUR/USD					
	02/2010 00:00 - 12/2014 23:00			01/2015 00:00 - 05/2017 23:00		
Variable	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|Fdiscount\|$	0.069***	0.085***	0.076***	0.265***	0.051	0.084***
Q_{LD}^F/N_{LD}^F	-0.147***			0.011		
Q_{SD}^F/N_{SD}^F	0.000		-0.026***	0.080*		0.043*
Vol^F	0.407***			0.295***		
Q_{LD}^S/N_{LD}^S		-0.161***	-0.179***		-0.374***	-0.359***
Q_{SD}^S/N_{SD}^S		-0.025*			0.062	
Vol^S		0.303***	0.356***		0.158***	0.156***

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). P-values assigned based on HAC robust standard errors: *** p<0.01, ** p<0.05, * p<0.1.

Table 4.7: Long-run liquidity dynamics in JPY/USD

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{t-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_t$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |Fdiscount|_t, Q_{j,h}^P/N_{j,h}^P, Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and $j = LD, SD$, denote the quoting intensity of large and small dealers, respectively. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects are omitted for brevity.

Sample:	2/01/2010 00:00-6/30/2014 23:00			7/01/2014 00:00-5/31/2017 23:00		
Variable:	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|Fdiscount\|$	0.273***	0.153***	0.138***	0.089***	0.064***	0.022***
	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)
Q_{LD}^F/N_{LD}^F	-0.006***			-0.002***		
	(0.00)			(0.00)		
Q_{SD}^F/N_{SD}^F	0.021***		0.012***	0.007***		0.012***
	(0.00)		(0.00)	(0.00)		(0.00)
Vol^F	1.734***			2.293***		
	(0.60)			(0.24)		
Q_{LD}^S/N_{LD}^S		-0.008***	-0.007***		-0.002***	-0.007***
		(0.00)	(0.00)		(0.00)	(0.00)
Q_{SD}^S/N_{SD}^S		0.003***			0.001	
		(0.00)			(0.00)	
Vol^S		1.353***	1.457***		2.053	1.457***
		(0.41)	(0.42)		(0.31)	(0.42)
θ	-0.07***	-0.12***	-0.12***	-0.09***	-0.11***	-0.12***
$F - Stat$	90.20	129.961	133.250	124.24	110.62	133.25
Hour dummies	yes	yes	yes	yes	yes	yes
Adj. R^2	0.273	0.280	0.281	0.261	0.274	0.281
Obs	26484	26484	26484	17592	17592	17592

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). HAC robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. F-statistic based on the Pesaran et al (2001) bounds test: 1% critical values 3.29 for I(0) and 4.37 for I(1).

Table 4.8: Long-run liquidity dynamics in EUR/USD

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{t-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_t$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |Fdiscount|_t, Q_{j,h}^P/N_{j,h}^P, Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and $j = LD, SD$, denote dealer quoting activity of large and small dealers, respectively. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects omitted for brevity.

Sample:	2/01/2010 00:00-6/30/2014 23:00			7/01/2014 00:00-5/31/2017 23:00		
Variable:	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|Fdiscount\|$	0.035***	0.043***	0.039***	0.057***	0.011	0.006***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Q_{LD}^F/N_{LD}^F	-0.003***			0.000		
	(0.00)			(0.00)		
Q_{SD}^F/N_{SD}^F	0.000		-0.001***	0.003**		0.002
	(0.00)		(0.00)	(0.00)		(0.00)
Vol^F	8.821***			3.787***		
	(0.90)			(0.66)		
Q_{LD}^S/N_{LD}^S		-0.002***	-0.002***		-0.004***	-0.004*
		(0.00)	(0.00)		(0.00)	(0.00)
Q_{SD}^S/N_{SD}^S		-0.001*			0.001	
		(0.00)			(0.00)	
Vol^S		6.573***	7.721***		2.030***	2.006***
		(0.55)	(0.65)		(0.35)	(0.34)
θ	-0.10***	-0.12***	-0.130***	-0.09***	-0.11***	-0.11***
$F-Stat$	91.76	106.04	113.64	51.77	59.92	60.79
Hour dummies	yes	yes	yes	yes	yes	yes
Adj. R^2	0.277	0.275	0.275	0.347	0.355	0.349
Obs	29457	29457	29457	14619	14619	14619

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). HAC robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. F-statistic based on the Pesaran et al (2001) bounds test: 1% critical values 3.29 for I(0) and 4.37 for I(1).

Table 4.6 shows the coefficients estimates of the long-run equations, expressed in terms of economic magnitudes by scaling by the standard deviations of the regressors. Table 4.7 and Table 4.8 show the complete test results for the long-run relationship among the variables for JPY/USD and EUR/USD, respectively. The reported F-statistics of the Pesaran et al. (2001) bounds test exceed I(1) critical values for all equation, indicating the presence of a statistically significant long-run relationship among the selected measures of liquidity and volatility.

Focusing on Table 4.6, the ECM-ARDL model estimation results point at several takeaways. First, there is a strong and robust relationship between FX market liquidity, as proxied by bid-ask spreads in both swap ($Spread^F$) and spot ($Spread^S$), with FX funding liquidity, as proxied by the absolute forward discount ($Fdiscount$). For example, a one standard deviation widening in $Fdiscount$ is associated with 41.6*bp* (26.5*bp*) wider bid-ask in JPY/USD (EUR/USD) swap, and a 64.9*bp* (8.4*bp*) wider bid-ask spread in JPY/USD (EUR/USD) spot. The link between funding and market liquidity strengthens in the second sub-sample period, particularly for JPY/USD where economic magnitude of the coefficient on $Fdiscount$ increases more than three-fold in the swap bid-ask spread equation. For EUR/USD I also observe a substantial strengthening of the liquidity relationship from 6.9*bp* to 26.5*bp* in the swap market, while the relationship in spot increases only slightly and remains comparably low.

Second, the positive net effect of dealer competition on market liquidity in FX swaps has all but disappeared. A one standard deviation increase in the quoting intensity by large dealers in the swap market, Q^F_{LD}/N^F_{LD} used to be associated with a 28.7*bp* (14.7*bp*) narrowing of bid-ask spreads on JPY/USD forward rate spread (EUR/USD forward rate spread) in the 2010 to mid-2014 (December 2015 for EUR/USD), but the effect becomes small and not statistically significant in the mid-2014 to 2017 period. In contrast, the negative association between quoting intensity of smaller dealers, Q^F_{SD}/N^F_{SD}, and market liquidity in FX swaps has persisted for both currency pairs, and even strengthened significantly in the case of EUR/USD. Thus, a one standard deviation increase in Q^F_{SD}/N^F_{SD} is associated with 0.31*bp* (0.04*bp*) wider bid-ask spread in FX swap market for JPY/USD (EUR/USD).

Third, in contrast to the swap market, dealer competition in the spot market has continued to contribute to significant narrowing of bid-ask spreads also in the post-2014 period. A one standard deviation increase in quoting intensity by large dealers in spot, Q^S_{LD}/N^S_{LD}, is associated with 20.5bp (35.9bp) narrower bid-ask spreads in JPY/USD (EUR/USD) spot. For neither of the two currency pairs do small dealers in spot contribute to a tighter bid-ask spread in the spot market in the second half of the sample. Both coefficients are small in magnitude (0.039 and 0.062 for JPY/USD and EUR/USD) and not significantly different from zero.

Fourth, rises in small dealer (primarily tier-2 bank) activity in swaps appears to have negative spillovers on spot market liquidity. Specifically, even though small dealer competition in spot markets does not seem to have a statistically significant effect on market liquidity in JPY/USD and EUR/USD, higher quoting intensity by small dealers in swaps is associated with wider bid-ask spreads in the spot (Table 4.7, last column). Similarly, when small dealer quoting intensity in EUR/USD spot is replaced with small dealer quoting intensity in swaps in the $Spread^S$ equation, the coefficient is two times larger in magnitude and takes on a positive sign (Table 4.8, last column). This is noteworthy especially because approximately half of small dealers in FX swaps do not even participate in spot market directly (see, Figure 4.5, above).

These results are obtained controlling for time-of-day effects with hourly dummies, as well as for intraday volatility in both spot and swap markets. The results are also robust to measuring FX funding liquidity using CIP deviations instead of the un-adjusted forward discount (see Tables A5 and A6).

4.4.2 Adverse Liquidity Effects Of Small Dealer Competition

What are the possible economic reasons behind the negative relationship between FX market liquidity and small dealer competition? The first reason is that small dealers charge higher spreads. This can be gleaned from Table 4.9, which shows simple average of the median hourly bid-ask spreads and forward discounts computed from forward quotes by large and small dealers. For both JPY/USD and EUR/USD, the bid-ask spreads of forward rates (expressed as a percentage of mid-forward rate, in basis points) are significantly higher for small dealers compared to large dealers. Similarly, the forward discount (forward spread, expressed as a percentage of mid-spot rate) is also somewhat wider for small dealers compared to large dealers. This is consistent with smaller dealers facing higher hurdle rates to enter as market-makers in the swap market, presumably due to being smaller volume players. Hence, their competition does not lead to the narrowing of the spreads to the levels that can be supported by large dealers.

Table 4.9: Forward rate bid-ask spreads and forward discounts quoted by large vs small dealers

	JPY/USD		EUR/USD	
Dealer category:	$Spread^F$	$Fdiscount$	$Spread^F$	$Fdiscount$
Large dealers	3.322bp	-4.471bp	2.218bp	-2.990bp
Small dealers	3.648bp	-4.579bp	2.247bp	-2.993bp

Notes: Average median hourly quotes. Large dealer are banks classified as G-SIBs and appearing in the Euromoney FX Survey rankings. 2/01/2010 00:00 to 5/31/2017 23:00 sample period.

The second reason for the negative relationship between FX market liquidity and small dealer competition in the swap market relates to the relative informational disadvantage of small dealers compared to large dealers. Bjonnes et al. (2009) find that order flow of large dealer banks is more informative than that of small banks, in terms of return predictability. Menkhoff et al. (2016) find evidence that informative order-flow of sophisticated investors affects foreign exchange rate via the intermediation of large dealers. My logic is consistent with this literature. Large dealers intermediate the lion share of customer flows inside their internal liquidity pools. As a result, their activity on anonymous primary interdealer venues, such as Reuters Matching, is largely driven by the hedging of any residual inventory imbalances reflective of aggregate informed trading of their client base. This would suggest that, on average, large dealers possess more precise information about the "true" market forward exchange rate at any point in time (again, because they observe the FX hedging activity of their diverse client base) compared to small dealers.

In order to test this, I follow recent studies which assess the distribution of quote submissions. For example, Corsetti et al. (2017) use information on both quotes and trades to construct a quote dispersion measure that accounts for market participants' reaction to new information based on the speed of trade execution. As I do not possess information on trades but only on quote submissions, the measure of dispersion follows Jankowitsch et al. (2011) and is applied to forward quotes within each hour:

$$Disp_h^F = \sqrt{\sum_{i=1}^{h_i} \frac{q_i^F}{Q_h^F} \left(\frac{F_i - \bar{F}_h}{\bar{F}_h} \right)} \tag{4.6}$$

where q_i^F accounts for the number of forward quote submissions within a minute, Q_h^F denotes the total number of submissions within the hour, F_i denotes the forward mid-price in minute i and $\bar{F}_h$ is the average forward price of each hour. In times of higher volatility and low liquidity, I expect the dispersion of quotes to be comparably larger and $Disp_h^F$ to increase.

I then, once again, formulate a conditional ECM, but for the system that includes $Disp_h^F$, quoting intensity by large and small dealers, and hourly volatility of the forward rate, Vol_h^P, as control:

$$\Delta Disp_h^F = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Disp_{h-1}^F + \boldsymbol{\theta x_{h-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_h \tag{4.7}$$

where $\mathbf{z_h^S} = (Disp_h^F, Q_{LD,t}^F/N_{LD,h}^F, Q_{SD,h}^F/N_{SD,h}^F, Vol^F) = (Disp_h^F, \mathbf{x_h^P})'$. Table 4.10 shows the results. Consistent with the hypothesis outlined above, large dealer quoting intensity is associated with a reduction in the dispersion of forward quotes in both JPY/USD and EUR/USD. In contrast, smaller dealer quoting intensity is not associated with a reduction in the forward quote dispersion in JPY/USD, while their marginal effect on dispersion is less than that of large dealers in the EUR/USD swap market.

To summarise, the results reported in Tables 4.9 and 4.10 indicate that two effects are at play in generating the negative relationship between liquidity in the FX swap market and competition by small dealers. The first one relates to their wider required intermediation spreads, both bid-ask spreads and the forward spread (forward discount). The second one relates to their informational disadvantage and hence greater uncertainty about the actual market mid-rate for pricing FX swaps, which leads to greater dispersion and volatility of the forward quotes.

Table 4.10: Forward quote dispersion of small and large dealers

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. In the top part of the table coefficients are scaled by the sample standard deviation, while raw regression estimates are displayed in the bottom part. Specifically, for the second sub-sample period (JPY/USD: July 2014 to May 2017; EUR/USD: January 2015 to May1 2017) we estimate:

$$\Delta Disp_h^F = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Disp_{h-1}^F + \boldsymbol{\theta x_{t-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^F} + \boldsymbol{\beta \Delta x_h^F} + u_t$$

where a vector $\mathbf{z_h^F} = (Disp_h^F, Q_{j,h}^F/N_{j,h}^F, Vol_h^F) = (Disp_h^F, \mathbf{x_h^F})'$ and $j = LD, SD$, denotes quoting intensity of large and small dealers, respectively. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects omitted for brevity.

	JPY/USD	EUR/USD
Q_{LD}^F/N_{LD}^F	-0.15 ***	-0.10 ***
Q_{SD}^F/N_{SD}^F	0.12 *	-0.09 ***
Vol^F	3.45 ***	3.24***
Q_{LD}^F/N_{LD}^F	-0.003 ***	-0.002***
	(0.00)	(0.00)
Q_{SD}^F/N_{SD}^F	0.005*	-0.004***
	(0.00)	(0.00)
Vol^F	18.548***	41.625***
	(4.28)	(2.79)
θ	-0.31***	-0.61***
$F-Stat$	111.93	624.79
Hourly dummies	yes	yes
Adj. R^2	0.60	0.68
Obs	17050	14247

Hourly sample: 1/06/2014 00:00 to 5/31/2017 23:00; GMT time-stamps. *** p<0.01, ** p<0.05, * p<0.1.

4.4.3 Three-tier Dealer Classification

To further understand the varying impact of quoting activity from different dealer segments, in a next step I apply an even more granular dealer classification to confirm the intuition of the earlier findings. I explicitly distinguish between dealers that are classified as G-SIB banks (tier-1), dealers that are listed in the *Euromoney Survey* but that are not classified as a G-SIB banks (tier-2), and other small dealers (tier-3). While I have separated tier-1 banks from the remaining dealer universe in the previous estimations, in this analysis I allow for an explicit split between small dealers (tier-2 and tier-3) and thereby differentiate between banks that vary in size but which are all not affected by recent regulatory changes. I re-examine the estimation of the ARDL-ECM model of Equation 4.5 and then discuss the impact of dealers' quoting intensity on the forward spread dispersion.

To begin with, Table 4.11 displays the economic impact of a change in dealer quoting intensity for each of the three dealer classifications (tier-1: $T1$, tier-2: $T2$, tier-3: $T3$). Coefficients are scaled by the sample standard deviation, while raw regression coefficients and cointegration parameters for both currency pairs are shown in Tables 4.12 and 4.13. Focusing on the second sub-sample, Table 4.11 shows how the positive impact of quoting intensity of the three different groups on market liquidity conditions declines almost monotonically in magnitude or statistical significance from tier-1 to tier-3 banks.

For example, for JPY/USD an increase in tier-1 quoting intensity leads to a decline in market liquidity in the swap market by $10.2bp$, tier-2 dealers lower the spread, though the effect is not significant, and a one standard deviation increase of tier-3 dealer quoting intensity significantly worsens market liquidity conditions by $66.2bp$. Further, I find that activity of tier-3 dealers ($40.5bp$) deteriorates liquidity conditions in the spot market by a factor more than six times as large as the quoting intensity of tier-2 dealers ($6.1bp$).

For EUR/USD, I document similar market characteristics though the impact of quoting intensity of tier-2 and tier-3 dealers on market liquidity dynamics is lower than for JPY/USD. Between January 2015 and May 2017, liquidity conditions in the swap market do not significantly improve neither because of an increase in tier-1 nor in tier-2 dealers quoting intensity, while tier-3 dealers' quoting widens the spread ($2.2bp$). I also note that tier-3 swap dealer activity affects liquidity conditions in the spot market, however, the magnitude of spread changes due to swap dealers is much lower ($1.5bp$) than the impact of spot dealer quoting intensity. Again, I show in the appendix (Table A7 and A8) that the results are qualitatively similar if funding costs are measured by CIP deviations instead of by the forward spread.

Table 4.11: Intraday conditional co-movement of liquidity measures

This table is based on estimation results of the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{h-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_t$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |Fdiscount|_t, Q_{j,h}^P/N_{j,h}^P, Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and $j = T1, T2, T3$, denotes Tier-1, Tier-2, and Tier-3 dealer quoting intensity, respectively. The coefficients are scaled by the standard deviation of the explanatory variables in each sub-sample.

	JPY/USD					
	02/2010 00:00 - 06/2014 23:00			07/2014 00:00 - 05/2017 23:00		
Variable	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|Fdiscount\|$	0.372***	0.196***	0.198***	0.265***	0.322***	0.194***
Q_{T1}^F/N_{T1}^F	-0.254***			-0.102***		
Q_{T2}^F/N_{T2}^F	0.372***		0.181***	-0.032		0.061**
Q_{T3}^F/N_{T3}^F	0.419***		0.171***	0.662***		0.405***
Vol^F	0.156***			0.393***		
Q_{T1}^S/N_{T1}^S		-0.471***	-0.481***		-0.241***	-0.164***
Q_{T2}^S/N_{T2}^S		-0.143***			0.105***	
Q_{T3}^S/N_{T3}^S		0.122*			-0.035	
Vol^S		0.139***	0.144***		0.353***	0.370***

	EUR/USD					
	02/2010 00:00 - 12/2014 23:00			01/2015 00:00 - 05/2017 23:00		
Variable	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|Fdiscount\|$	0.072***	0.074***	0.079***	0.277***	0.145***	0.097***
Q_{T1}^F/N_{T1}^F	-0.154***			0.012		
Q_{T2}^F/N_{T2}^F	0.088***		-0.014	-0.145		-0.208
Q_{T3}^F/N_{T3}^F	-0.064*		-0.029**	0.022**		0.015*
Vol^F	0.405***			0.294***		
Q_{T1}^S/N_{T1}^S		-0.048***	-0.100***		-0.228***	-0.292***
Q_{T2}^S/N_{T2}^S		-0.052***			0.211***	
Q_{T3}^S/N_{T3}^S		0.013			-0.265***	
Vol^S		0.266***	0.357***		0.111***	0.155

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). P-values assigned based on HAC robust standard errors: *** p<0.01, ** p<0.05, * p<0.1.

Table 4.12: Long-run liquidity dynamics in JPY/USD (By Bank Tiers)

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{t-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_t$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |Fdiscount|_t, Q_{j,h}^P/N_{j,h}^P, Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and $j = T1, T2, T3$, denotes Tier-1, Tier-2, and Tier-3 dealer quoting intensity, respectively. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects are omitted for brevity.

Sample:	2/01/2010 00:00-6/30/2014 23:00			7/01/2014 00:00-5/31/2017 23:00		
Variable:	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|Fdiscount\|$	0.268***	0.141***	0.143***	0.057***	0.069***	0.041***
	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Q_{T1}^F/N_{T1}^F	-0.005***			-0.002***		
	(0.00)			(0.00)		
Q_{T2}^F/N_{T2}^F	0.017***		0.008***	-0.001		0.002**
	(0.00)		(0.00)	(0.00)		(0.00)
Q_{T3}^F/N_{T3}^F	0.037***		0.015***	0.024***		0.015***
	(0.00)		(0.00)	(0.00)		(0.00)
Vol^F	1.480***			2.113***		
	(0.51)			(0.23)		
Q_{T1}^S/N_{T1}^S		-0.006***	-0.006***		-0.002***	-0.002***
		(0.00)	(0.00)		(0.00)	(0.00)
Q_{T2}^S/N_{T2}^S		-0.002***			0.002***	
		(0.00)			(0.00)	
Q_{T3}^S/N_{T3}^S		0.002*			0.000	
		(0.00)			(0.00)	
Vol^S		1.316	1.367		1.893	1.987
		(0.40)	(0.40)		(0.26)	(0.26)
θ	-0.08***	-0.12***	-0.12***	-0.10***	-0.12***	-0.12***
$F-Stat$	83.57	113.65	113.58	125.71	109.87	93.74
Hour dummies	yes	yes	yes	yes	yes	yes
Adj. R^2	0.28	0.28	0.28	0.27	0.28	0.27
Obs	26484	26484	26484	17592	17592	17592

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). HAC robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. F-statistic based on the Pesaran et al (2001) bounds test: 1% critical values 3.29 for I(0) and 4.37 for I(1).

Table 4.13: Long-run liquidity dynamics in EUR/USD (By Bank Tiers)

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{t-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_t$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |Fdiscount|_t, Q_{j,h}^P/N_{j,h}^P, Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and $j = T1, T2, T3$, denotes Tier-1, Tier-2, and Tier-3 dealer quoting intensity, respectively. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects omitted for brevity.

Sample:	2/01/2010 00:00-6/30/2014 23:00			7/01/2014 00:00-5/31/2017 23:00		
Variable:	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|Fdiscount\|$	0.037***	0.037***	0.040***	0.059***	0.031***	0.021***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Q_{T1}^F/N_{T1}^F	-0.004***			0.000		
	(0.00)			(0.00)		
Q_{T2}^F/N_{T2}^F	0.001***		0.000	-0.001		-0.002
	(0.00)		(0.00)	(0.00)		(0.00)
Q_{T3}^F/N_{T3}^F	-0.002*		-0.001**	0.001**		0.001*
	(0.00)		(0.00)	(0.00)		(0.00)
Vol^F	8.780***			3.768***		
	(0.90)			(0.66)		
Q_{T1}^S/N_{T1}^S		-0.001***	-0.002***		-0.003***	-0.004***
		(0.00)	(0.00)		(0.00)	(0.00)
Q_{T2}^S/N_{T2}^S		-0.003***			0.004***	
		(0.00)			(0.00)	
Q_{T3}^S/N_{T3}^S		0.000			-0.004***	
		(0.00)			(0.00)	
Vol^S		5.777***	7.739***		1.430***	1.985***
		(0.52)	(0.65)		(0.28)	(0.33)
θ	-0.10***	-0.14***	-0.13***	-0.09***	-0.14***	-0.11***
$F-Stat$	80.67	105.28	96.76	44.42	70.37	52.46
Hour dummies	yes	yes	yes	yes	yes	yes
Adj. R^2	0.28	0.28	0.28	0.35	0.38	0.35
Obs	26484	26484	26484	17592	17592	17592

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). HAC robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. F-statistic based on the Pesaran et al (2001) bounds test: 1% critical values 3.29 for I(0) and 4.37 for I(1).

Lastly, I document in Table 4.14 a gradually declining impact of dealer quoting intensity on the prevalent dispersion of quote submissions in the swap market. In the top panel, coefficients are scaled by the variables' standard deviation, while the bottom part of the table displays the raw long-run coefficients of Equation 4.7. I note that an increasing quoting intensity by tier-1 banks leads to a significant decline in the dispersion of forward quotes of 14.4*bp* for JPY/USD and 10.4*bp* for EUR/USD. Quoting intensity by tier-2 banks does not lower the quote dispersion for JPY/USD, while for EUR/USD it becomes smaller but to a lesser degree compared to the effect of tier-1 quoting activity (7.7*bp*). Notably, coefficients of tier-3 quoting intensity are positive for both currency pairs, providing further evidence that banks with a smaller customer base and those that are likely exposed to lower volumes of customer order flow contribute to a wider dispersion of forward quote submissions. This effect is particularly large for JPY/USD (33.3*bp*) while smaller and not significant for EUR/USD (1.4*bp*).

4.4.4 Contagion Versus Interdependence

In this subsection, I test for the presence of contagion from funding markets to market liquidity at quarter-end balance sheet reporting periods, when large dealers pull back and small dealers increase their quoting intensity in FX swaps.

I follow Forbes and Rigobon (2002) and calculate an adjusted correlation coefficient using hourly data. I then test for regime shifts between quarter-end and non-quarter end months.[12] Adjusting the correlation coefficient for heteroskedastic levels of volatility allows me to make further statements about contagions and spillovers, rather than simple co-movement. To this end, I estimate the following bivariate vector autoregressive model:

$$\Delta y_h = \phi(L)\Delta y_{h-1} + \eta_h \tag{4.8}$$

$$\Delta y_h = \left[\Delta Fdiscount_h, \Delta Spread_h^P\right] \text{ where } P = F, S \tag{4.9}$$

where Δy_h refers to the first differenced and de-seasonalized measures of funding and market liquidity. First, I de-seasonalise the FX liquidity metrics by regressing their changes on hourly dummies. Second, I estimate Equation (4.8) using a 200-hour rolling window

[12] Using a vector autoregression framework, Moinas et al. (2017) exploit regime shifts of volatility levels to examine liquidity dynamics in the European treasury bond markets.

Table 4.14: Forward quote dispersion (By Bank Tiers)

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. In the top part of the table coefficients are scaled by the sample standard deviation, while raw regression estimates are displayed in the bottom part. Specifically, for the second subsample period (JPY/USD: July 2014 to May 2017; EUR/USD: January 2015 to May 2017) we estimate:

$$\Delta Disp_h^F = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Disp_{h-1}^F + \boldsymbol{\theta x_{t-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^F} + \boldsymbol{\beta \Delta x_h^F} + u_t$$

where a vector $\mathbf{z_h^F} = (Disp_h^F, Q_{j,h}^F/N_{j,h}^F, Vol_h^F) = (Disp_h^F, \mathbf{x_h^F})'$ and $j = T1, T2, T3$, denotes Tier-1, Tier-2, and Tier-3 dealer quoting intensity, respectively. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects omitted for brevity.

	JPY/USD	EUR/USD
Q_{T1}^F/N_{T1}^F	-0.144 ***	-0.104***
Q_{T2}^F/N_{T2}^F	0.073	-0.77**
Q_{T3}^F/N_{T3}^F	0.333***	0.014
Vol^F	3.410***	3.239***
Q_{T1}^F/N_{T1}^F	-0.003***	-0.002***
	(0.00)	(0.00)
Q_{T2}^F/N_{T2}^F	0.002	-0.005**
	(0.00)	(0.00)
Q_{T3}^F/N_{T3}^F	0.012***	0.000
	(0.00)	(0.00)
Vol^F	18.314***	41.578***
	(4.20)	(2.78)
θ	-0.32***	-0.61***
$F-Stat$	97.08	518.69
Hour dummies	yes	yes
Adj. R^2	0.60	0.68
Obs	17050	14247

Hourly sample: 1/06/2014 00:00 to 5/31/2017 23:00; GMT time-stamps. *** p<0.01, ** p<0.05, * p<0.1.

and store the variance-covariance matrix for every single estimation.[13] Third, based on the obtained variance-covariance matrices from the hourly VAR regressions, I follow the ap-

[13]Every rolling estimation initially allows for 8 lags but I increase the lag length in a step-wise fashion until residuals are free of serial correlation.

proach in Forbes and Rigobon (2002) and construct an unconditional correlation coefficient as follows:

$$\rho = \frac{\rho^*}{\sqrt{(1 + \delta[1 - (\rho^*)^2])}} \tag{4.10}$$

$$\text{with } \delta = \frac{\sigma^{QE}_{Fdiscount}}{\sigma^{NQE}_{Fdiscount}} - 1$$

where ρ^* refers to the standard correlation coefficient between funding and market liquidity, and $\sigma^{QE}_{Fdiscount}$ and $\sigma^{NQE}_{Fdiscount}$ refer to the average variance of FX funding liquidity in quarter-end months (QE) and the two preceding non quarter-end months (NQE), respectively. Since the intra-day data allows us to construct the measure of co-movement for every rolling window estimation, I am able to obtain a time series of unconditional correlation coefficients for each QE and NQE period.

Having obtained the time-series of the adjusted correlation coefficients between funding and market liquidity measures in quarter-end months and the preceding two months, I then employ a one-sided t-test to examine for the following hypothesis:

$$H_0\text{: } \rho_{NQE} < \rho_{QE} \mid \rho_{QE} < 0 \qquad\qquad H_A\text{: } \rho_{NQE} > \rho_{QE} \mid \rho_{QE} < 0$$

where ρ_{NQE} and ρ_{QE} refer to the average of the adjusted correlation coefficients in the non-quarter-end and quarter-end months. Rejecting the null hypothesis indicates that the shocks to $Fdiscount$ in a quarter-end month lead to spillover to bid-ask spreads, even after adjusting for the higher level of volatility of funding conditions during these periods.[14]

Table 4.15 shows the results for JPY/USD. The average adjusted correlation coefficient is negative in a number of quarter-end as well as non-quarter-end months in both spot (upper panel) and forward (lower panel). Based on the t-test conducted on the adjusted correlation coefficients, I am able to reject the null of no spillovers in 3 out of 11 quarter-end months considered for spot, and 4 out of 11 quarter-end months for forwards.

[14]Qualitatively the same conclusions are drawn when shocks to CIP deviations at quarter-end months are considered. Results are summarized in Table A9 and A10 for JPY/USD and EUR/USD, respectively.

Table 4.15: Contagion from FX funding to market liquidity in JPY/USD

The table shows tests for contagion from FX funding liquidity to FX market liquidity in spot and swap markets. We follow Forbes and Rigobon (2002), and conduct a t-test of whether the correlations between $\Delta Fdiscount$ and $\Delta Spread^P$ is significantly more negative at quarter-ends, where $P = S, F$. The correlation coefficients are estimated using a 200-hour rolling window bi-variate VAR, and adjusted for heteroskedastic levels of volatility, thus allowing to make statements about contagions rather than a simple co-movement.

To spot market:	$\Delta Fdiscount \rightarrow \Delta Spread^S$					
	Q-end month		Prior 2 months		Contagion	
QEs:	ρ_{QE}	σ_{QE}	ρ_{NQE}	σ_{NQE}	t-stat	Reject H_0
09/14	0.004	0.009	-0.004	0.006	-14.13	
12/14	0.015	0.020	-0.019	0.015	-25.18	
03/15	0.075	0.022	0.018	0.046	-27.35	
06/15	-0.023	0.051	0.009	0.064	8.94	Yes
09/15	0.008	0.008	-0.011	0.020	-22.89	
12/15	0.020	0.057	0.013	0.066	-1.61	
03/16	0.067	0.042	-0.009	0.100	-17.41	
06/16	-0.078	0.062	-0.012	0.021	18.66	Yes
09/16	-0.007	0.053	-0.093	0.094	-19.36	
12/16	-0.044	0.046	0.014	0.058	17.36	Yes
03/17	0.024	0.081	-0.053	0.072	-15.36	
Avg. contagion	**-0.048**	**0.053**			**14.987**	**3/11 QEs**

To forward market:	$\Delta Fdiscount \rightarrow \Delta Spread^F$					
	Q-end month		Prior 2 months		Contagion	
QEs:	ρ_{QE}	σ_{QE}	ρ_{NQE}	σ_{NQE}	t-stat	Reject H_0
09/14	0.002	0.008	-0.005	0.007	-13.44	
12/14	0.012	0.019	-0.021	0.015	-25.1	
03/15	0.078	0.023	0.023	0.047	-25.66	
06/15	-0.028	0.052	0.011	0.061	10.71	Yes
09/15	0.016	0.014	-0.013	0.02	-26.63	
12/15	-0.011	0.025	0.03	0.067	14.61	Yes
03/16	0.053	0.057	-0.023	0.106	-15.23	
06/16	-0.078	0.061	-0.008	0.022	20.17	Yes
09/16	-0.008	0.076	-0.071	0.084	-12	
12/16	-0.06	0.05	-0.009	0.071	13.37	Yes
03/17	0.021	0.077	-0.059	0.066	-16.95	
Avg. contagion	**-0.044**	**0.047**			**14.715**	**4/11 QEs**

Hourly samples. Bivariate VAR, $\Delta y_h = \phi(L)\Delta y_{h-1} + \eta_h$ with $\Delta y_h = \left[\Delta Fdiscount_h, \Delta Spread_h^P\right]$, where $P = F, S$ and all endogenous variables de-seasonalised of hourly effects.

Table 4.16: Contagion from FX funding to market liquidity in EUR/USD

The table shows tests for contagion from FX funding liquidity to FX market liquidity in spot and swap markets. We follow Forbes and Rigobon (2002), and conduct a t-test of whether the correlations between $\Delta Fdiscount$ and $\Delta Spread^P$ is significantly more negative at quarter-ends, where $P = S, F$. The correlation coefficients are estimated using a 200-hour rolling window bi-variate VAR, and adjusted for heteroskedastic levels of volatility, thus allowing to make statements about contagions rather than a simple co-movement.

To spot market:	$\Delta Fdiscount \rightarrow \Delta Spread^S$					
	Q-end month		Prior 2 months		Contagion	
QEs:	ρ_{QE}	σ_{QE}	ρ_{NQE}	σ_{NQE}	t-stat	Reject H_0
03/15	0.029	0.052	-0.055	0.09	-19.33	
06/15	0.086	0.089	0.082	0.126	-0.65	
09/15	0.041	0.027	-0.012	0.048	-22.99	
12/15	-0.027	0.036	-0.006	0.069	6.44	Yes
03/16	0.123	0.062	0.053	0.049	-17.68	
06/16	0.011	0.007	-0.002	0.025	-13.6	
09/16	0.039	0.03	-0.082	0.064	-42.74	
12/16	-0.106	0.081	0.039	0.047	29.35	Yes
03/17	-0.031	0.096	-0.113	0.161	-10.77	
Avg. contagion	**-0.067**	**0.059**			**17.90**	**2/9 QEs**
To forward market:	$\Delta Fdiscount \rightarrow \Delta Spread^F$					
	Q-end month		Prior 2 months		Contagion	
QEs:	ρ_{QE}	σ_{QE}	ρ_{NQE}	σ_{NQE}	t-stat	Reject H_0
03/15	0.031	0.049	-0.075	0.1	-23.54	
06/15	0.077	0.096	0.005	0.102	-11.08	
09/15	0.04	0.028	-0.04	0.058	-31.32	
12/15	-0.026	0.034	-0.025	0.066	0.32	Yes
03/16	0.152	0.067	0.052	0.046	-23.83	
06/16	-0.01	0.029	-0.003	0.027	3.58	Yes
09/16	0.053	0.031	-0.078	0.058	-48.82	
12/16	-0.105	0.08	0.025	0.047	26.43	Yes
03/17	-0.017	0.098	-0.113	0.171	-12.04	
Avg. contagion	**-0.047**	**0.048**			**10.11**	**3/9 QEs**

Hourly samples. Bivariate VAR, $\Delta y_h = \phi(L)\Delta y_{h-1} + \eta_h$ with $\Delta y_h = \left[\Delta Fdiscount_h, \Delta Spread_h^P\right]$, where $P = F, S$ and all endogenous variables de-seasonalised of hourly effects.

Hence, the evidence in favour of contagion from FX funding to FX market liquidity in JPY/USD is strongest for December 2015, June 2016, and December 2016.

Table 4.16 shows the analogous test results for EUR/USD. Similar to JPY/USD, the number of months in which the null is rejected in favour of contagious spillovers is higher for forward bid-ask spreads than for spot bid-ask spreads. At the same time, the overall number of months, in which the results point towards contagion from FX funding liquidity to FX market liquidity, is slightly less that for JPY/USD. Still, both December 2015 and December 2016 turn out to be the quarter-end periods with the most robust evidence in favour of contagions, rather than simple co-movement. It is noteworthy that these months also fall on year-ends, when additional G-SIB surcharges apply to large dealer banks' balance sheets.

Overall, the empirical evidence suggests that a deterioration in funding liquidity at quarter-ends can spillover to market liquidity in spot and forward market. Taken together with the previous results on dealer activity, these findings suggest that the pull-back by G-SIBs from dealing in FX swaps at quarter- and year-ends can have a particularly contagious implications for spot market liquidity, as they are displaced by more expensive and less informed dealers.

4.4.5 Small Dealer Market-making: Case Study Of December 2016

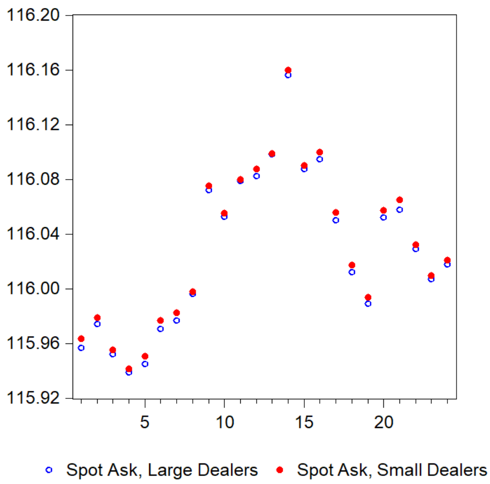
In this sub-section I provide evidence that small dealers displaced large dealers as market-makers in FX swaps for both currency pairs and also in spot markets for EUR/USD in December 2016. However, because these smaller volume players require higher hurdle rates, in terms of both bid-ask spreads on the forward points that they quote as well as wider forward discount, the increased competition by smaller dealers allows the low FX liquidity environment to persist. In contrast, large banks continued to dominate as market-makers in spot for JPY/USD and differences are very small for EUR/USD, indicating that it is likely their balance sheet constraints on the exposures to FX derivatives that explains their pull-back from quoting inside spreads in the swap market.

The left-hand panels of Figures 4.11 and 4.12 show the median hourly JPY/USD quotes of small dealers and large dealers during December 2016, for spot and forward points, respectively. The top graph displays ask and the bottom graph the associated bid quotes. The right-hand panels show the hypothetical location of small dealer quotes relative to large dealer quotes in the case that small dealers are actively making markets by quoting inside spreads. If the actual quotes correspond to the inside spread scenario, then this indicates that small dealers, not large dealers, would have been making markets on average during this month. The comparison of actual data (left) to the scenarios (right) in Figure 4.11 indicates that in December 2016, despite the pull-back by large dealers (G-SIBs) form the market in the aggregate, large dealers continued to make markets in spot. However, the comparison of actual data (left) to the scenarios (right) in Figure 4.12 of the forward quotes indicates that small dealers entered as market-makers in the swap market.

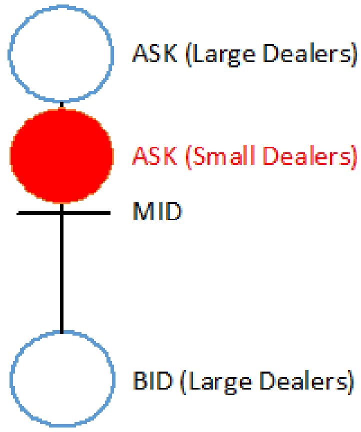
The results for EUR/USD, shown in Figure 4.13 and 4.14, are qualitatively similar for the swap market and small players appear to act as market-makers. In the spot market, differences in the submitted quotes of small and large dealers are low in magnitude and spreads in spot and swap markets by the different dealer segments are very similar. Yet, in contrast to JPY/EUR it appears that during most times of the day small dealers also act as market makers in this market segment.

A hypothesis that I so far reject is that smaller banks enter the market to source liquidity in one of the currencies. If this was the case, the test of inside versus outside spread by dealer category would have shown smaller dealers providing skewed quotes relative to large dealers. Data indicate that this is not the case.

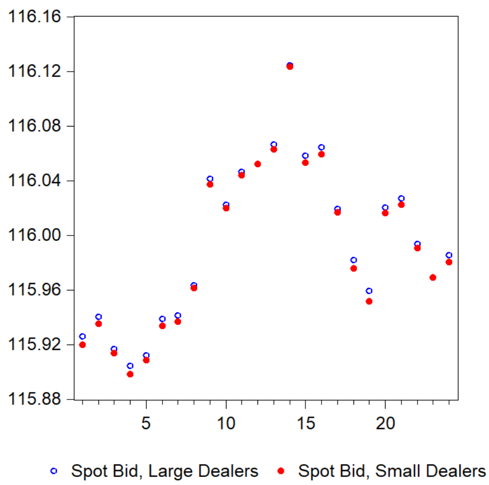
Such entry of small dealers in forwards rather than spot as market-makers is consistent with large dealers pulling back from trading in derivatives (e.g. forwards and FX swaps) but continuing to make markets in spot. Hence, the results indicate that small dealers can play an important role in market-making in FX swaps when funding conditions are tight and spreads are wide enough for smaller-volume players to profitably engage as market-makers. This adds nuance to the recently documented changes in FX market structure, whereby liquidity provision is bifurcated between the few large dealers making markets as principals and smaller dealers that operate an agency model simply passing client flows into the wholesale FX market. In this context, special periods, like quarter-ends, can be used for identification of funding liquidity effect on dealer competition and FX market activity.



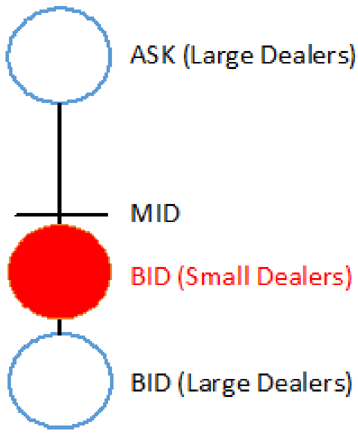
(a) Actual spot ASK



(b) Hypothetical: Smaller Dealers Short USD spot

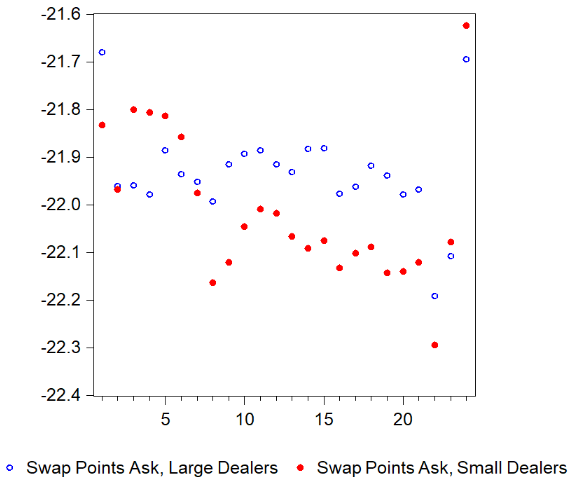


(c) Actual Spot BID

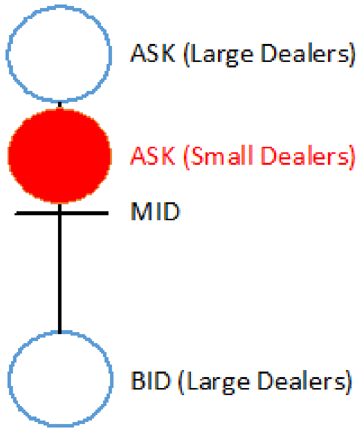


(d) Hypothetical: Smaller Dealers Long USD spot

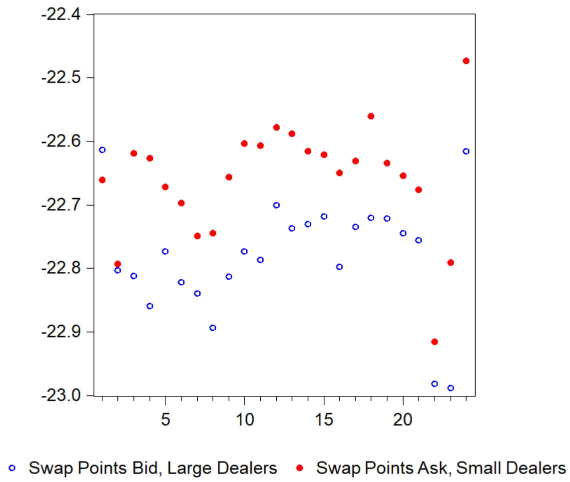
Figure 4.11: Median quote in JPY/USD spot (December 2016)
Spot quotes of small dealers compared to spot quotes of large dealers. Left-hand side figures indicate that small dealers did not act as market-makers in spot in December 2016. *Top-left:* Actual spot ask; *Top-right:* Hypothetical: Smaller Dealers Short USD spot scenario (ASK: SELL USD @ 116.160). *Bottom-left:* Actual spot bid; *Bottom-right:* Hypothetical: Smaller Dealers Long USD spot scenario (BID: BUY USD @ 116.124). GMT time-stamps.



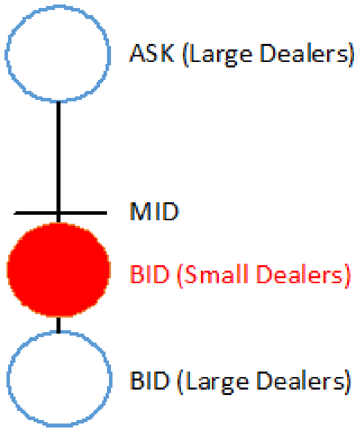
(a) Actual forward points ASK



(b) Hypothetical: Smaller Dealers Short USD forward

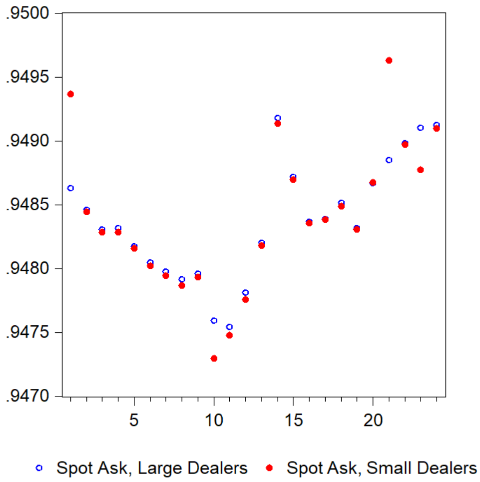


(c) Actual forward points BID

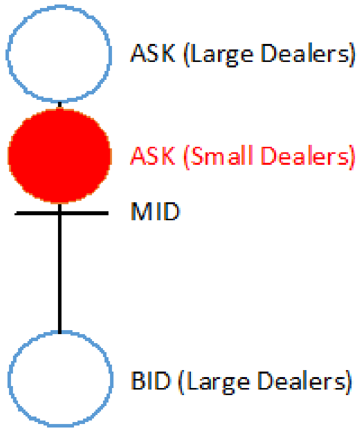


(d) Hypothetical: Smaller Dealers Long USD forward

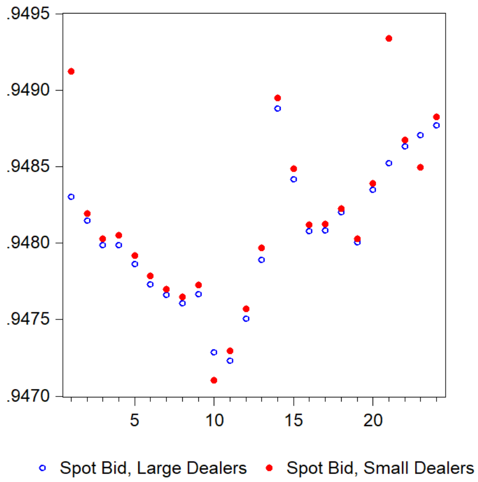
Figure 4.12: Median quote in JPY/USD forward points (December 2016)
Forward quotes of small dealers compared to forward quotes of large dealers. Indicates that small dealers acted as market-makers in swap markets in December 2016. *Top-left:* Actual forward points ask; *Top-right:* Hypothetical: Smaller Dealers Short USD forward scenario (ASK: SELL USD @ 116.160 - 0.221). *Bottom-left:* Actual forward points bid; *Bottom-right:* Hypothetical: Smaller Dealers Long USD forward scenario (BID: BUY USD @ 116.124 - 0.226). GMT time-stamps.



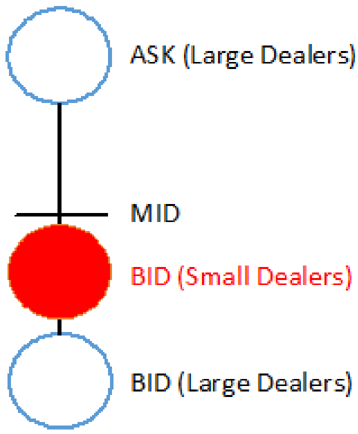
(a) Actual Spot ASK



(b) Hypothetical: Smaller Dealers Short USD spot

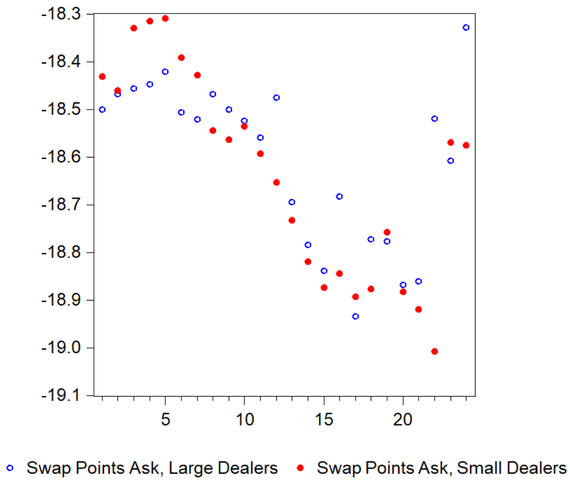


(c) Actual Spot BID

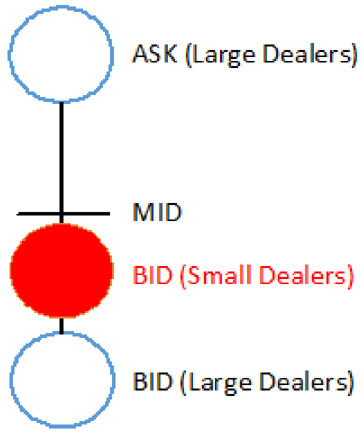


(d) Hypothetical: Smaller Dealers Long USD spot

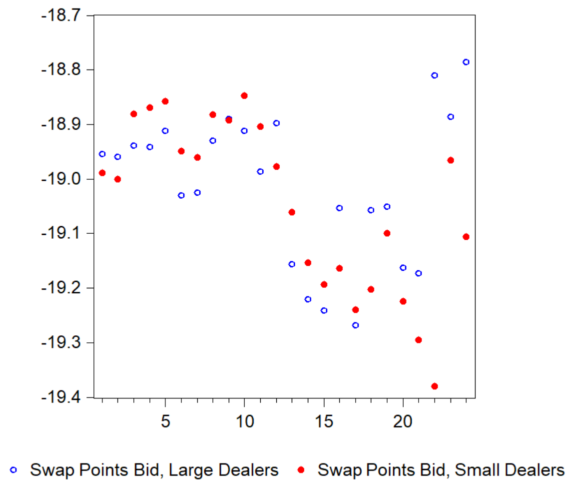
Figure 4.13: Median quote in EUR/USD spot (December 2016)
Spot quotes of small dealers compared to spot quotes of large dealers. Indicates that small and large dealers act as market-makers in spot in December 2016. *Top-left:* Actual spot ask; *Top-right:* Hypothetical: Smaller Dealers Short USD spot scenario (ASK: SELL USD @ 0.9491). *Bottom-left:* Actual spot bid; *Bottom-right:* Hypothetical: Smaller Dealers Long USD spot scenario (BID: BUY USD @ 0.9489). GMT time-stamps.



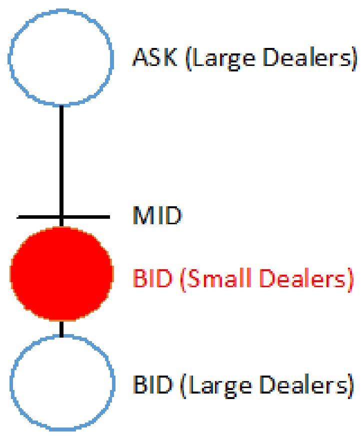
(a) Actual Forward Points ASK



(b) Hypothetical: Smaller Dealers Short USD spot



(c) Actual Forward Points BID



(d) Hypothetical: Smaller Dealers Long USD spot

Figure 4.14: Median quote in EUR/USD forward points (December 2016)
Forward quotes of small dealers compared to forward quotes of large dealers. Indicates that small dealers acted as market-makers in swap markets in December 2016. *Top-left:* Actual forward points ask; *Top-right:* Hypothetical: Smaller Dealers Short USD forward scenario (ASK: SELL USD @ 0.9491 - 0.00188). *Bottom-left:* Actual forward points bid; *Bottom-right:* Hypothetical: Smaller Dealers Long USD forward scenario (BID: BUY USD @ 0.9489 - 0.00192). GMT time-stamps.

4.5 Conclusion

Trading volumes in FX swaps are larger than spot, making these instruments crucial for price discovery in currency markets, yet there is hardly any literature on FX market liquidity taking the pricing of FX derivatives into account. In this paper, I measure the joint evolution of FX spot and forward market liquidity conditions. I draw on the pricing of both types of instruments to study the relationship between FX market liquidity and FX funding liquidity. The assessment of liquidity conditions also takes into account information on the number of dealers active at a given point in time, their quoting intensity, as well as dealer characteristics such as size. In particular, I account for the window dressing behaviour by large dealers in FX swap markets. The empirical strategy thus makes particular use of month-end and quarter-end dynamics in the FX swap market for the identification of exogenous FX funding liquidity shocks, and their impact on FX market liquidity.

The results based on intraday data for JPY/USD and EUR/USD show that FX spot and swap market liquidity is intimately linked. Furthermore, I find evidence for the presence of liquidity spirals in FX markets, a la Brunnermeier and Pedersen (2009). The co-movement between market and funding liquidity has increased in recent years, and the instances of extreme liquidity droughts have also risen. Statistical tests also point towards contagion of adverse FX funding liquidity shocks to market liquidity in both FX swap and spot markets in the most extreme periods, particularly at year-ends.

Competitive dynamics of FX dealers play an important role in these liquidity dynamics. Specifically, the positive impact of dealer competition on FX market liquidity has decreased over time. The structural break in the relationship between FX swap and spot market liquidity conditions, and with dealer activity, appears related to the window dressing behaviour by large FX swap dealers. While large dealers still dominate market-making in spot markets at all times, and their quoting intensity is associated with improved liquidity dynamics, they have exhibited a tendency to pull-back from market-making in FX derivatives, namely FX swaps, around balance sheet reporting periods. Yet, as large dealers are displaced by smaller, and as such more expensive and less informed, dealers in the forward markets, spot market liquidity appears to also suffer because liquidity conditions in spot and swap markets are tightly linked.

Hence, funding liquidity is now arguably a more important economic factor to understand bid-ask spreads in FX spot. As such, window dressing by large dealers in FX swaps has been disruptive not only to swap market liquidity but also to liquidity in spot. This does not mean that tighter regulation is harmful to liquidity conditions. Quite the opposite, liquidity in times of stress would have been enhanced with better capitalization and other regulatory reforms. However, the evidence does point at G-SIB dealer banks managing down their activity in FX swaps around regulatory reporting periods, which suggests that banks do not yet hold adequate capital to support the size of their desired FX swap business. Consistent with evidence in Gambacorta and Shin (2016) that higher bank capital is associated with greater amount of lending, the evidence can be interpreted to suggest that higher bank capitalisation would also translate into improved FX market liquidity.

A Additional Tables & Robustness Checks

Table A1: Intraday correlation coefficient of liquidity measures by trading hour

This table reports the average correlation coefficient by trading hour between spot market and funding liquidity ($\rho^{Fd}_{Sp^S}$) and swap market and funding liquidity ($\rho^{Fd}_{Sp^F}$) in two sub-sample periods for the JPY/USD and EUR/USD exchange rates.

	JPY/USD				EUR/USD			
GMT	Feb 2010 - Jun 2014		Jul 2014 - May 2017		Feb 2010 - Dec 2014		Jan 2015 - May 2017	
	$\rho^{Fd}_{Sp^S}$	$\rho^{Fd}_{Sp^F}$	$\rho^{Fd}_{Sp^S}$	$\rho^{Fd}_{Sp^F}$	$\rho^{Fd}_{Sp^S}$	$\rho^{Fd}_{Sp^F}$	$\rho^{Fd}_{Sp^S}$	$\rho^{Fd}_{Sp^F}$
0.00	-0.37	-0.43	-0.50	-0.59	-0.41	-0.50	-0.30	-0.42
1.00	-0.41	-0.47	-0.49	-0.58	-0.41	-0.50	-0.31	-0.44
2.00	-0.42	-0.47	-0.50	-0.59	-0.38	-0.48	-0.32	-0.45
3.00	-0.41	-0.46	-0.46	-0.57	-0.42	-0.51	-0.27	-0.41
4.00	-0.41	-0.46	-0.50	-0.59	-0.40	-0.49	-0.27	-0.42
5.00	-0.36	-0.41	-0.48	-0.57	-0.40	-0.51	-0.25	-0.41
6.00	-0.39	-0.44	-0.49	-0.58	-0.38	-0.50	-0.30	-0.46
7.00	-0.43	-0.47	-0.39	-0.53	-0.33	-0.46	-0.23	-0.40
8.00	-0.41	-0.45	-0.32	-0.48	-0.30	-0.42	-0.26	-0.41
9.00	-0.46	-0.49	-0.36	-0.50	-0.28	-0.41	-0.23	-0.38
10.00	-0.45	-0.48	-0.48	-0.56	-0.29	-0.41	-0.38	-0.49
11.00	-0.42	-0.45	-0.49	-0.56	-0.36	-0.46	-0.42	-0.52
12.00	-0.41	-0.46	-0.51	-0.58	-0.35	-0.47	-0.40	-0.51
13.00	-0.40	-0.46	-0.47	-0.55	-0.33	-0.46	-0.35	-0.46
14.00	-0.40	-0.47	-0.48	-0.54	-0.27	-0.40	-0.35	-0.46
15.00	-0.36	-0.43	-0.50	-0.56	-0.27	-0.41	-0.41	-0.50
16.00	-0.41	-0.47	-0.53	-0.59	-0.37	-0.49	-0.49	-0.56
17.00	-0.41	-0.48	-0.53	-0.59	-0.38	-0.50	-0.51	-0.57
18.00	-0.43	-0.49	-0.52	-0.58	-0.33	-0.47	-0.33	-0.41
19.00	-0.43	-0.49	-0.50	-0.56	-0.32	-0.46	-0.34	-0.41
20.00	-0.40	-0.46	-0.48	-0.56	-0.33	-0.45	-0.27	-0.36
21.00	-0.38	-0.44	-0.44	-0.55	-0.28	-0.39	-0.27	-0.35
22.00	-0.46	-0.51	-0.48	-0.58	-0.29	-0.39	-0.21	-0.32
23.00	-0.42	-0.46	-0.49	-0.59	-0.33	-0.43	-0.17	-0.30
Avg.	**-0.44**	**-0.50**	**-0.51**	**-0.63**	**0.24**	**0.27**	**-0.28**	**-0.40**

Hourly sample: 2/01/2010 00:00 to 5/31/2017 23:00; GMT time-stamps.

Table A2: Intraday correlation coefficient of liquidity measure, by trading hour incl. CIP deviations

This table reports the average correlation coefficient by trading hour between spot market and funding liquidity ($\rho^{CIPdev}_{Spread^S}$) and swap market and funding liquidity ($\rho^{CIPdev}_{Spread^F}$) in two sub-sample periods for the JPY/USD and EUR/USD exchange rates.

	JPY/USD				EUR/USD			
GMT	Feb 2010 - Jun 2014		Jul 2014 - May 2017		Feb 2010 - Dec 2014		Jan 2015 - May 2017	
	$\rho^{CIPdev}_{Spread^S}$	$\rho^{CIPdev}_{Spread^F}$	$\rho^{CIPdev}_{Spread^S}$	$\rho^{CIPdev}_{Spread^F}$	$\rho^{CIPdev}_{Spread^S}$	$\rho^{CIPdev}_{Spread^F}$	$\rho^{CIPdev}_{Spread^S}$	$\rho^{CIPdev}_{Spread^F}$
0	-0.37	-0.43	-0.50	-0.59	-0.41	-0.50	-0.30	-0.42
1	-0.41	-0.47	-0.49	-0.58	-0.41	-0.50	-0.31	-0.44
2	-0.42	-0.47	-0.50	-0.59	-0.38	-0.48	-0.32	-0.45
3	-0.41	-0.46	-0.46	-0.57	-0.42	-0.51	-0.27	-0.41
4	-0.41	-0.46	-0.50	-0.59	-0.40	-0.49	-0.27	-0.42
5	-0.36	-0.41	-0.48	-0.57	-0.40	-0.51	-0.25	-0.41
6	-0.39	-0.44	-0.49	-0.58	-0.38	-0.50	-0.30	-0.46
7	-0.43	-0.47	-0.39	-0.53	-0.33	-0.46	-0.23	-0.40
8	-0.41	-0.45	-0.32	-0.48	-0.30	-0.42	-0.26	-0.41
9	-0.46	-0.49	-0.36	-0.50	-0.28	-0.41	-0.23	-0.38
10	-0.45	-0.48	-0.48	-0.56	-0.29	-0.41	-0.38	-0.49
11	-0.42	-0.45	-0.49	-0.56	-0.36	-0.46	-0.42	-0.52
12	-0.41	-0.46	-0.51	-0.58	-0.35	-0.47	-0.40	-0.51
13	-0.40	-0.46	-0.47	-0.55	-0.33	-0.46	-0.35	-0.46
14	-0.40	-0.47	-0.48	-0.54	-0.27	-0.40	-0.35	-0.46
15	-0.36	-0.43	-0.50	-0.56	-0.27	-0.41	-0.41	-0.50
16	-0.41	-0.47	-0.53	-0.59	-0.37	-0.49	-0.49	-0.56
17	-0.41	-0.48	-0.53	-0.59	-0.38	-0.50	-0.51	-0.57
18	-0.43	-0.49	-0.52	-0.58	-0.33	-0.47	-0.33	-0.41
19	-0.43	-0.49	-0.50	-0.56	-0.32	-0.46	-0.34	-0.41
20	-0.40	-0.46	-0.48	-0.56	-0.33	-0.45	-0.27	-0.36
21	-0.38	-0.44	-0.44	-0.55	-0.28	-0.39	-0.27	-0.35
22	-0.46	-0.51	-0.48	-0.58	-0.29	-0.39	-0.21	-0.32
23	-0.42	-0.46	-0.49	-0.59	-0.33	-0.43	-0.17	-0.30
Avg.	**-0.41**	**-0.46**	**-0.47**	**-0.57**	**-0.34**	**-0.46**	**-0.32**	**-0.43**

Hourly sample: 2/01/2010 00:00 to 5/31/2017 23:00; GMT time-stamps.

Table A3: First principal component of liquidity measures by trading hour

This table reports the variation explained by the first principal component by trading hour between spot market and funding liquidity $({}^{Fd}_{Sp^S})$ and swap market and funding liquidity $({}^{Fd}_{Sp^F})$ in two sub-sample periods for the JPY/USD and EUR/USD exchange rates.

	JPY/USD				EUR/USD			
GMT	Feb 2010 - Jun 2014		Jul 2014 - May 2017		Feb 2010 - Dec 2014		Jan 2015 - May 2017	
	$\rho^{Fd}_{Sp^S}$	$\rho^{Fd}_{Sp^F}$	$\rho^{Fd}_{Sp^S}$	$\rho^{Fd}_{Sp^F}$	$\rho^{Fd}_{Sp^S}$	$\rho^{Fd}_{Sp^F}$	$\rho^{Fd}_{Sp^S}$	$\rho^{Fd}_{Sp^F}$
0	70.42	73.29	75.11	82.15	65.00	65.17	63.99	69.95
1	72.09	75.13	74.21	81.44	64.02	64.68	64.56	70.27
2	72.16	74.77	73.82	81.42	65.14	65.32	63.75	69.79
3	71.83	74.60	72.44	81.03	64.45	64.77	60.75	67.43
4	71.52	74.46	74.18	82.17	64.94	65.23	59.94	67.09
5	69.06	72.01	73.28	81.18	60.86	61.89	59.10	66.90
6	71.13	73.60	75.53	82.18	56.05	58.37	62.92	70.56
7	73.21	75.13	69.33	77.64	55.46	58.52	57.63	66.21
8	72.27	74.25	64.96	74.64	56.75	59.84	59.24	66.56
9	74.61	76.18	67.01	75.53	57.45	60.71	57.08	64.95
10	74.56	75.65	76.70	81.58	58.83	62.12	67.07	72.74
11	73.00	74.46	77.20	81.66	59.28	62.14	70.30	75.28
12	72.28	74.62	78.52	82.74	58.76	61.24	68.94	73.95
13	71.62	74.68	77.00	81.31	59.63	61.52	65.59	70.87
14	72.04	75.14	77.45	81.72	62.73	63.95	65.80	70.95
15	70.17	73.42	79.20	82.77	66.31	67.30	68.31	73.00
16	72.93	75.66	81.27	84.14	63.00	64.59	75.87	78.97
17	72.60	75.70	81.72	84.44	63.51	65.01	78.65	80.66
18	73.33	76.09	81.32	84.16	60.07	62.63	66.58	70.32
19	73.25	76.06	79.95	83.66	61.32	63.45	67.34	70.85
20	71.95	74.95	78.17	82.86	67.23	68.08	61.57	65.74
21	70.28	73.31	75.80	83.07	66.28	66.59	65.06	69.32
22	74.50	77.04	76.14	83.79	65.06	65.83	56.92	63.29
23	72.53	74.95	74.46	82.64	65.82	66.02	52.83	59.81
Avg.	**72.22**	**74.80**	**75.62**	**81.66**	**62.00**	**63.54**	**64.16**	**69.81**

Hourly sample: 2/01/2010 00:00 to 5/31/2017 23:00; GMT time-stamps.

Table A4: First principal component of liquidity measures, by trading hour incl. CIP deviations

This table reports the variation explained by the first principal component by trading hour between spot market and funding liquidity ($\frac{CIPdev}{Spread^S}$) and swap market and funding liquidity ($\frac{CIPdev}{Spread^F}$) in two sub-sample periods for the JPY/USD and EUR/USD exchange rates.

	JPY/USD				EUR/USD			
GMT	Feb 2010 - Jun 2014		Jul 2014 - May 2017		Feb 2010 - Dec 2014		Jan 2015 - May 2017	
	$\frac{CIPdev}{Spread^S}$	$\frac{CIPdev}{Spread^F}$	$\frac{CIPdev}{Spread^S}$	$\frac{CIPdev}{Spread^F}$	$\frac{CIPdev}{Spread^S}$	$\frac{CIPdev}{Spread^F}$	$\frac{CIPdev}{Spread^S}$	$\frac{CIPdev}{Spread^F}$
0	68.65	71.45	75.15	79.68	70.67	74.95	64.77	71.21
1	70.67	73.61	74.42	79.07	70.40	74.95	65.56	71.80
2	70.97	73.44	75.01	79.70	69.09	74.19	66.07	72.40
3	70.39	73.04	73.14	78.51	70.77	75.33	63.34	70.41
4	70.30	73.09	74.93	79.73	69.94	74.54	63.60	71.17
5	67.85	70.63	73.79	78.67	70.23	75.67	62.37	70.30
6	69.41	71.89	74.33	79.20	68.92	75.10	65.18	73.05
7	71.64	73.57	69.63	76.47	66.52	72.81	61.32	69.93
8	70.74	72.61	66.04	74.00	64.99	71.16	63.22	70.30
9	72.82	74.34	67.94	74.84	63.88	70.50	61.33	68.92
10	72.66	73.79	74.06	78.14	64.64	70.44	68.79	74.60
11	71.06	72.62	74.41	78.11	68.01	72.97	70.84	76.19
12	70.26	72.82	75.40	78.96	67.38	73.35	70.04	75.63
13	69.83	73.13	73.72	77.34	66.48	72.89	67.58	73.19
14	70.10	73.46	73.75	77.21	63.45	70.19	67.74	73.11
15	68.16	71.70	74.92	78.20	63.73	70.54	70.30	75.16
16	70.70	73.74	76.68	79.57	68.65	74.53	74.45	78.23
17	70.44	73.83	76.68	79.48	69.23	74.94	75.66	78.40
18	71.48	74.46	76.23	79.05	66.74	73.73	66.58	70.66
19	71.56	74.59	74.79	78.25	66.18	73.02	66.88	70.58
20	69.92	73.16	74.12	78.17	66.65	72.52	63.68	67.85
21	69.04	72.09	71.85	77.50	64.15	69.42	63.32	67.42
22	72.92	75.61	74.06	78.82	64.41	69.36	60.33	66.03
23	70.86	73.15	74.53	79.57	66.58	71.37	58.28	65.15
Avg.	**70.52**	**73.16**	**73.73**	**78.26**	**67.15**	**72.85**	**65.89**	**71.74**

Hourly sample: 2/01/2010 00:00 to 5/31/2017 23:00; GMT time-stamps.

Table A5: Long-run dynamics in JPY/USD incl. CIP Deviation

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{h-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_h$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |CIPdev|_h, Q_{LD,h}^P/N_{LD,h}^P, Q_{SD,h}^P/N_{SD,h}^P, Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and LD, SD denotes large and small dealers. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects omitted for brevity.

	2/01/2010 00:00 6/30/2014 23:00			7/01/2014 00:00 5/31/2017 23:00		
Variable:	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|CIPDev\|$	0.021***	0.012***	0.011***	0.009***	0.006***	0.001***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Q_{LD}^F/N_{LD}^F	-0.006***			-0.002***		
	(0.00)			(0.00)		
Q_{SD}^F/N_{SD}^F	0.022***		0.009***	0.011***		0.009***
	(0.00)		(0.00)	(0.00)		(0.00)
Vol^F	1.826***			2.537***		
	(0.63)			(0.27)		
Q_{LD}^S/N_{LD}^S		-0.008***	-0.007***		-0.002***	-0.001***
		(0.00)	(0.00)		(0.00)	(0.00)
Q_{SD}^S/N_{SD}^S		0.003***			0.002***	
		(0.00)			(0.00)	
Vol^S		1.384***	1.485***		2.098***	2.177***
		(0.42)	(0.43)		(0.33)	(0.23)
θ	-0.07***	-0.12***	-0.12***	-0.08***	-0.11***	-0.10***
$F-Stat$	88.01	129.99	132.67	120.10	107.98	92.29
Hour dummies	yes	yes	yes	yes	yes	yes
Adj. R^2	0.273	0.280	0.280	0.260	0.273	0.270
Obs	26484	26484	26484	17592	17592	17592

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). HAC robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. F-statistic based on the Pesaran et al (2001) bounds test: 1% critical values 3.29 for I(0) and 4.37 for I(1).

Table A6: Long-run dynamics in EUR/USD incl. CIP Deviation

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{h-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_h$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |CIPdev|_h, Q_{LD,h}^P/N_{LD,h}^P, Q_{SD,h}^P/N_{SD,h}^P, Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and LD, SD denotes large and small dealers. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects omitted for brevity.

Sub-sample period:	2/01/2010 00:00-12/31/2014 23:00			01/01/2015 00:00-5/31/2017 23:00		
Variable:	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|CIPDev\|$	0.009***	0.005***	0.006***	0.009***	0.003***	0.001***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Q_{LD}^F/N_{LD}^F	-0.003***			0.000		
	(0.00)			(0.00)		
Q_{SD}^F/N_{SD}^F	-0.001		-0.002***	0.003**		0.001***
	(0.00)		(0.00)	(0.00)		(0.00)
Vol^F	7.411***			3.484***		
	(0.78)			(0.60)		
Q_{LD}^S/N_{LD}^S		-0.001***	-0.001***		-0.004***	-0.004
		(0.00)	(0.00)		(0.00)	(0.00)
Q_{SD}^S/N_{SD}^S		0.000			0.001	
		(0.00)			(0.00)	
Vol^S		6.104***	6.022***		1.932***	1.908***
		(0.55)	(0.54)		(0.34)	(0.33)
θ	-0.12***	-0.12***	-0.13***	-0.09***	-0.11***	-0.11***
$F-Stat$	108.04	108.11	110.60	52.28	61.15	61.94
Hour dummies	yes	yes	yes	yes	yes	yes
Adj. R^2	0.279	0.276	0.276	0.348	0.355	0.349
Obs	29457	29457	29457	14619	14619	14619

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). HAC robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. F-statistic based on the Pesaran et al (2001) bounds test: 1% critical values 3.29 for I(0) and 4.37 for I(1).

Table A7: Long-run dynamics in JPY/USD incl. CIP Deviation by Bank Tiers

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{h-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_h$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |CIPdev|_h, Q_{T1,h}^P/N_{T1,h}^P, Q_{T2,h}^P/N_{T2,h}^P, Q_{T3,h}^P/N_{T3,h}^P Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and $T1$, $T2$, $T3$ denote the quoting intensity of Tier-1, Tier-2, and Tier-3 dealers, respectively. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects omitted for brevity.

	2/01/2010 00:00 6/30/2014 23:00			7/01/2014 00:00 5/31/2017 23:00		
Variable:	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|CIPDev\|$	0.023***	0.012***	0.012***	0.005***	0.006***	0.004***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Q_{T1}^F/N_{T1}^F	-0.005***			-0.002***		
	(0.00)			(0.00)		
Q_{T2}^F/N_{T2}^F	0.017***		0.008***	0.000		0.003***
	(0.00)		(0.00)	(0.00)		(0.00)
Q_{T3}^F/N_{T3}^F	0.043***		0.018***	0.030***		0.018***
	(0.00)		(0.00)	(0.00)		(0.00)
Vol^F	1.504***			2.146***		
	(0.51)			(0.24)		
Q_{T1}^S/N_{T1}^S		-0.006***	-0.006***		-0.003***	-0.002***
		(0.00)	(0.00)		(0.00)	(0.00)
Q_{T2}^S/N_{T2}^S		-0.003***			0.002***	
		(0.00)			(0.00)	
Q_{T3}^S/N_{T3}^S		0.002***			0.001***	
		(0.00)			(0.00)	
Vol^S		1.341***	1.369***		1.945***	1.998***
		(0.41)	(0.40)		(0.29)	(0.27)
θ	-0.08***	-0.12***	-0.12***	-0.10***	-0.11***	-0.12***
$F-Stat$	83.53	114.17	114.61	126.34	105.31	94.65
Hour dummies	yes	yes	yes	yes	yes	yes
Adj. R^2	0.28	0.282	0.28	0.27	0.28	0.28
Obs	26484	26484	26484	17592	17592	17592

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). HAC robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. F-statistic based on the Pesaran et al (2001) bounds test: 1% critical values 3.29 for I(0) and 4.37 for I(1).

Table A8: Long-run dynamics in EUR/USD incl. CIP Deviation

This table reports coefficients from the long-run equation of a conditional error correction model (ECM) derived from an autoregressive distributed lag (ARDL) model specification. Specifically, for the two sub-sample periods and for spot and swap market liquidity, measured by spot rate $Spread^S$ and forward rate $Spread^F$ bid ask spreads, $P = S, F$, we estimate:

$$\Delta Spread_h^P = \alpha + \sum_{i=1}^{23} \delta_i H_i + \theta_0 Spread_{h-1}^P + \boldsymbol{\theta x_{h-1}} + \sum_{i=1}^{p-1} \boldsymbol{\gamma_i \Delta z_{h-i}^P} + \boldsymbol{\beta \Delta x_h^P} + u_h$$

where a vector $\mathbf{z_h^S} = (Spread_h^P, |CIPdev|_h, Q_{T1,h}^P/N_{T1,h}^P, Q_{T2,h}^P/N_{T2,h}^P, Q_{T3,h}^P/N_{T3,h}^P Vol_h^P) = (Spread_h^P, \mathbf{x_h^P})'$ and $T1$, $T2$, $T3$ denote the quoting intensity of Tier-1, Tier-2, and Tier-3 dealers, respectively. F-statistics based on the results of the bound testing procedure for long-run relationship, robust to variables being in between I(0) and I(1). Constant and coefficients on short-run effects omitted for brevity.

Sub-sample period:	2/01/2010 00:00-12/31/2014 23:00			01/01/2015 00:00-5/31/2017 23:00		
Variable:	$Spread^F$	$Spread^S$	$Spread^S$	$Spread^F$	$Spread^S$	$Spread^S$
$\|CIPDev\|$	0.009***	0.004***	0.006***	0.009***	0.003***	0.004***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Q_{T1}^F/N_{T1}^F	-0.003***			0.000		
	(0.00)			(0.00)		
Q_{T2}^F/N_{T2}^F	0.000		-0.002	0.000		-0.002
	(0.00)		(0.00)	(0.00)		(0.00)
Q_{T3}^F/N_{T3}^F	-0.001***		0.000	0.001*		0.001*
	(0.00)		(0.00)	(0.00)		(0.00)
Vol^F	7.450***			3.482***		
	(0.78)			(0.61)		
Q_{T1}^S/N_{T1}^S		-0.001**	-0.001***		-0.003***	-0.004***
		(0.00)	(0.00)		(0.00)	(0.00)
Q_{T2}^S/N_{T2}^S		-0.002***			0.004***	
		(0.00)			(0.00)	
Q_{T3}^S/N_{T3}^S		0.000			-0.003***	
		(0.00)			(0.00)	
Vol^S		5.643***	5.992***		1.386***	1.887***
		(0.53)	(0.54)		(0.28)	(0.32)
θ	-0.12***	-0.13***	-0.13***	-0.09***	-0.14***	-0.11***
$F-Stat$	93.32	100.49	95.28	44.65	69.87	53.36
Hour dummies	yes	yes	yes	yes	yes	yes
Adj. R^2	0.279	0.278	0.276	0.348	0.377	0.350
Obs	26484	26484	26484	17592	17592	17592

Hourly sample; GMT time-stamps. ARDL lags chosen based on the Schwarz (Bayes) criterion (SC). HAC robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. F-statistic based on the Pesaran et al (2001) bounds test: 1% critical values 3.29 for I(0) and 4.37 for I(1).

Table A9: Contagion from FX funding to market liquidity in JPY/USD, incl. CIP Deviations

The table shows tests for contagion from FX funding liquidity to FX market liquidity in spot and swaps. We follow Forbes and Rigobon (2002), and conduct a t-test of whether the correlations between $\Delta CIPDev$ and $\Delta Spread^P$ is significantly more negative at quarter-ends, where $P = S, F$. The correlation coefficients are estimated using a 200-hour rolling window bi-variate VAR, and adjusted for heteroskedastic levels of volatility, thus allowing to make statements about contagions rather than a simple co-movement.

To spot market:	$\Delta CIPDev \rightarrow \Delta Spread^S$					
	Q-end month		Prior 2 months		Contagion	
QEs:	ρ_{QE}	σ_{QE}	ρ_{NQE}	σ_{NQE}	t-stat	Reject H_0
09/14	0.007	0.012	-0.006	0.011	-17	
12/14	0.019	0.021	-0.02	0.024	-24.53	
03/15	0.035	0.045	0.024	0.059	-3.16	
06/15	-0.01	0.054	0	0.058	2.63	Yes
09/15	0.005	0.008	-0.009	0.02	-17.2	
12/15	0.029	0.042	0.027	0.054	-0.55	
03/16	0.027	0.05	-0.008	0.097	-7.8	
06/16	-0.075	0.087	-0.008	0.021	13.74	Yes
09/16	0.005	0.052	-0.081	0.081	-20.75	
12/16	-0.077	0.047	0.004	0.049	24.97	Yes
03/17	-0.015	0.08	-0.051	0.068	-7.3	
Avg. contagion	**-0.054**	**0.063**			**13.78**	**3/11 QEs**

To forward market:	$\Delta CIPDev \rightarrow \Delta Spread^F$					
	Q-end month		Prior 2 months		Contagion	
QEs:	ρ_{QE}	σ_{QE}	ρ_{NQE}	σ_{NQE}	t-stat	Reject H_0
09/14	0.005	0.011	-0.007	0.011	-16.37	
12/14	0.015	0.02	-0.023	0.022	-25.04	
03/15	0.037	0.044	0.03	0.06	-2.04	
06/15	-0.015	0.056	0.002	0.055	4.39	Yes
09/15	0.014	0.014	-0.011	0.02	-23.29	
12/15	0.004	0.02	0.044	0.052	17.92	
03/16	0.029	0.064	-0.012	0.093	-8.24	
06/16	-0.059	0.099	-0.01	0.03	8.71	Yes
09/16	0.035	0.081	-0.029	0.086	-11.64	
12/16	-0.088	0.045	-0.013	0.059	22.45	Yes
03/17	-0.021	0.075	-0.054	0.065	-7.18	
Avg. contagion	**-0.054**	**0.067**			**11.85**	**3/11 QEs**

Hourly samples. Bivariate VAR, $\Delta y_h = \phi(L)\Delta y_{h-1} + \eta_h$ with $\Delta y_h = [\Delta CIPDev_h, \Delta Spread_h^P]$, where $P = F, S$ and all endogenous variables de-seasonalised of hourly effects.

Table A10: Contagion from FX funding to market liquidity in EUR/USD, incl. CIP Deviations

The table shows tests for contagion from FX funding liquidity to FX market liquidity in spot and swaps. We follow Forbes and Rigobon (2002), and conduct a t-test of whether the correlations between $\Delta CIPDev$ and $\Delta Spread^P$ is significantly more negative at quarter-ends, where $P = S, F$. The correlation coefficients are estimated using a 200-hour rolling window bi-variate VAR, and adjusted for heteroskedastic levels of volatility, thus allowing to make statements about contagions rather than a simple co-movement.

To spot market:	$\Delta CIPDev \rightarrow \Delta Spread^S$					
	Q-end month		Prior 2 months		Contagion	
QEs:	ρ_{QE}	σ_{QE}	ρ_{NQE}	σ_{NQE}	t-stat	Reject H_0
03/15	0.015	0.034	-0.064	0.108	-18.150	
06/15	0.112	0.051	0.043	0.123	-12.910	
09/15	0.021	0.043	-0.003	0.053	-7.860	
12/15	-0.033	0.044	-0.005	0.067	7.810	Yes
03/16	0.126	0.053	0.054	0.046	-21.090	
06/16	0.011	0.007	-0.002	0.023	-14.040	
09/16	0.044	0.037	-0.080	0.065	-40.060	
12/16	-0.123	0.072	0.040	0.058	35.020	
03/17	-0.039	0.041	-0.074	0.150	-6.120	Yes
Avg. contagion	**-0.078**	**0.058**			**21.42**	**2/9 QEs**

To forward market:	$\Delta CIPDev \rightarrow \Delta Spread^F$					
	Q-end month		Prior 2 months		Contagion	
QEs:	ρ_{QE}	σ_{QE}	ρ_{NQE}	σ_{NQE}	t-stat	Reject H_0
03/15	0.019	0.034	-0.073	0.11	-20.89	
06/15	0.101	0.049	0.009	0.114	-18.58	
09/15	0.021	0.042	-0.027	0.062	-15.2	
12/15	-0.04	0.04	-0.021	0.064	5.46	Yes
03/16	0.153	0.058	0.053	0.043	-27.12	
06/16	-0.01	0.029	-0.003	0.025	3.64	Yes
09/16	0.057	0.036	-0.07	0.063	-42.09	
12/16	-0.11	0.071	0.029	0.057	30.59	Yes
03/17	-0.043	0.039	-0.074	0.159	-5.14	
Avg. contagion	**-0.053**	**0.047**			**13.23**	**3/9 QEs**

Hourly samples. Bivariate VAR, $\Delta y_h = \phi(L)\Delta y_{h-1} + \eta_h$ with $\Delta y_h = \left[\Delta CIPDev_h, \Delta Spread_h^P\right]$, where $P = F, S$ and all endogenous variables de-seasonalised of hourly effects.

Bibliography

Ackermann, Carl, Richard McEnally, and David Ravenscraft, 1999, The performance of hedge funds: Risk, return, and incentives, *The Journal of Finance* 54, 833–874.

Adrian, Tobias, Vogt Fleming, Shachar Or, and Erik Vogt, 2017, Market liquidity after the financial crisis, Staff report no. 796, Federal Reserve Bank New York.

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik, 2009, Role of managerial incentives and discretion in hedge fund performance, *The Journal of Finance* 64, 2221–2256.

Agarwal, Vikas, Kevin Mullally, and Narayan Y. Naik, 2015a, Hedge funds: A survey of the academic literature, Working paper.

Agarwal, Vikas, Kevin Mullally, and Narayan Y. Naik, 2015b, Hedge funds: A survey of the academic literature, Working paper.

Agarwal, Vikas, and Narayan Y. Naik, 2000, Multi-period performance persistence analysis of hedge funds, *The Journal of Financial and Quantitative Analysis* 35, 327–342.

Akerlof, George A., 1970, The market for 'lemons': Quality uncertainty and the market mechanism*, *The Quarterly Journal of Economics* 84, 488.

Aldasoro, Inaki, Torsten Ehlers, and Egemen Eren, 2018, Frictions in Money Markets and Business Models of Global Banks: Evidence from the Japan Repo Premium, Technical report.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Clara Vega, 2003, Micro effects of macro announcements: Real-time price discovery in foreign exchange, *American Economic Review* 93, 38–62.

Arai, Fumihiko, Yoshibumi Makabe, Yasunori Okawara, and Teppei Nagano, 2016, Recent Trends in Cross-currency Basis, Bank of Japan Review Series, September 16-E-7, Bank of Japan.

Arnold, Julia, Robert Kosowski, and Paolo Zaffaroni, 2013, Survival of commodity trading advisors : Systematic vs . discretionary ctas.

Baba, Naohiko, Frank Packer, and Teppei Nagano, 2008, The spillover of money market turbulence to fx swap and cross-currency swap markets .

Baltas, Nick, and Robert Kosowski, 2013, Momentum strategies in futures markets and trend-following funds, Working paper.

Banti, Chiara, and Kate Phylaktis, 2015, Fx market liquidity, funding constraints and capital flows, *Journal of International Money and Finance* 56, 114 – 134.

Banti, Chiara, Kate Phylaktis, and Lucio Sarno, 2012, Global liquidity risk in the foreign exchange market, *Journal of International Money and Finance* 31, 267 – 291.

Berk, Jonathan B., 2005, Five myths of active portfolio management, *Journal of Portfolio Management* 31, 27 – 31.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Berk, Jonathan B., and Jules H. van Binsbergen, 2015, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1 – 20.

Bhardwaj, Geetesh, Gary B. Gorton, and K. Geert Rouwenhorst, 2014, Fooling some of the people all of the time: The inefficient performance and persistence of commodity trading advisors, *The Review of Financial Studies* 27, 3099–3132.

BIS, 2011, Globally systematically important banks: Assessment methodologyand the additional loss absorbency requirement - final document 1 – 32.

BIS, 2013a, Globally systematically important banks: updated assessment methodology and the higher loss absorbency requirement 1 – 23.

BIS, 2013b, Triennial central bank survey of foreign exchange and otc derivatives markets in 2013 1 – 24.

BIS, 2016, Triennial central bank survey of foreign exchange and otc derivatives markets in 2016 1 – 23.

BIS, 2017, Foreign exchange liquidity in the americas, *BIS Working Papers* 1 – 61.

Bjonnes, Geir H., Carol L. Osler, and Dagfinn Rime, 2009, Asymmetric information in the interbank foreign exchange market, Working Paper 2008/25, Norges Bank.

Bloomfield, Robert, Maureen O'Hara, and Gideon Saar, 2005, The "make or take" decision in an electronic market: Evidence on the evolution of liquidity, *Journal of Financial Economics* 75, 165 – 199.

Bollen, Nicolas P.B., 2013, Zero-r^2 hedge funds and market neutrality, *Journal of Financial and Quantitative Analysis* 48, 519–547.

Borio, Claudio, Robert McCauley, Patrick McGuire, and Vladyslav Sushko, 2016, Covered interest parity lost: understanding the cross-currency basis, *BIS Quarterly Review* 45–64.

Breedon, Francis, Dagfinn Rime, and Paolo Vitale, 2016, Carry trades, order flow, and the forward bias puzzle, *Journal of Money, Credit and Banking* 48, 1113–1134.

Breedon, Francis, and Paolo Vitale, 2010, An empirical study of portfolio-balance and information effects of order flow on exchange rates, *Journal of International Money and Finance* 29, 504 – 524.

Breusch, T. S., and A. R. Pagan, 1980, The lagrange multiplier test and its applications to model specification in econometrics, *The Review of Economic Studies* 47, 239–253.

Brown, Stephan J., and William N. Goetzmann, 2015, Hedge funds with style, *The Journal of Portfolio Management* 29, 101 –112.

Brown, Stephen J., and William N. Goetzmann, 2003, Hedge funds with style, *The Journal of Portfolio Management* 29, 101–112.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *The Review of Financial Studies* 22, 2201–2238.

Campbell, John, and Samuel B. Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average?, *Review of Financial Studies* 21, 1509–1531.

CGFS, 2017, Repo market functioning, *CGFS Publications* .

Cheung, Yin-Wong, and Dagfinn Rime, 2014, The offshore renminbi exchange rate: Microstructure and links to the onshore market, *Journal of International Money and Finance* 49, 170–189.

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

Clarida, Richard, Jordi Gali, and Mark Gertler, 1998, Monetary policy rules in practice: Some international evidence, *European Economic Review* 42, 1033 – 1067.

Corsetti, Giancarlo, Romain Lafarguette, and Arnaud Mehl, 2017, Fast trading, entropy and the foreign exchange market, Technical report.

Cremers, K. J. Martijn, and Antti Petajisto, 2009, How active is your fund manager? a new measure that predicts performance, *The Review of Financial Studies* 22, 3329.

Du, Tepper Alexander, Wenxin, and Adrien Verdelhan, 2018, Deviations from covered interest rate parity, *The Journal of Finance* 73, 915–957.

Elaut, Gert, and Peter Erdös, 2016, Adaptive time series momentum: An objective benchmark for systematic trend-following strategies, Working paper.

Elaut, Gert, Péter Erdös, and John Sjödin, 2016, An analysis of the risk-return characteristics of serially correlated managed futures, *Journal of Futures Markets* 36, 992–1013.

Elton, Edwin J., Martin J. Gruber, and Joel Rentzler, 1989, New public offerings, information, and investor rationality: The case of publicly offered commodity funds, *The Journal of Business* 62, 1–15.

Elton, Edwin J., Martin J. Gruber, and Joel Rentzler, 1990, The performance of publicly offered commodity funds, *Financial Analysts Journal* 46, 23–30.

Elton, Edwin J., Martin J. Gruber, and Joel C. Rentzler, 1987, Professionally managed, publicly traded commodity funds, *The Journal of Business* 60, 175–199.

Engel, Charles, Nelson C. Mark, and Kenneth D. West, 2008, Exchange rate models are not as bad as you think, in *NBER Macroeconomics Annual 2007, Volume 22*, NBER Chapters, 381–441 (National Bureau of Economic Research, Inc).

Engel, Charles, and Kenneth D. West, 2005, Exchange rates and fundamentals, *Journal of Political Economy* 113, 485–517.

Evans, Martin D. D., and Richard K. Lyons, 2002, Order flow and exchange rate dynamics, *Journal of Political Economy* 110, 170–180.

Forbes, Kristin J., and Roberto Rigobon, 2002, No contagion, only interdependence: Measuring stock market comovements, *The Journal of Finance* 57, 2223–2261.

French, Kenneth R., 2008, Presidential address: The cost of active investing, *The Journal of Finance* 63, 1537–1573.

Fung, William, and David A. Hsieh, 1997, Survivorship bias and investment style in the returns of ctas, *The Journal of Portfolio Management* 24, 30–41.

Fung, William, and David A. Hsieh, 2001, The risk in hedge fund strategies: Theory and evidence from trend followers, *The Review of Financial Studies* 14, 313.

Gambacorta, Leonardo, and Hyun Song Shin, 2016, Why bank capital matters for monetary policy, BIS Working Papers 558, Bank for International Settlements.

Gregoriou, Greg N., Georges Hübner, and Maher Kooli, 2010, Performance and persistence of commodity trading advisors: Further evidence, *Journal of Futures Markets* 30, 725–752.

Griffin, John M., and Jin Xu, 2009, How smart are the smart guys? a unique view from hedge fund stock holdings, *The Review of Financial Studies* 22, 2531–2570.

Gruber, Martin J., 1996, Another puzzle: The growth in actively managed mutual funds, *The Journal of Finance* 51, 783–810.

Hasbrouck, Joel, and Richard M. Levich, 2017, Fx market metrics: New findings based on cls bank settlement data, Working Paper 23206, National Bureau of Economic Research.

Huang, Roger D., and Ronald W. Masulis, 1999, Fx spreads and dealer competition across the 24-hour trading day, *The Review of Financial Studies* 12, 61–93.

Ibbotson, Roger G., and Peng Chen, 2006, The a,b,cs of hedge funds: Alphas, betas, and costs, Technical report.

IMF, 2014, Annual report on exchange rate arrangements and exchange restrictions 2014 1 – 109.

Jankowitsch, Rainer, Amrut Nashikkar, and Marti G. Subrahmanyam, 2011, Price dispersion in otc markets: A new measure of liquidity, *Journal of Banking and Finance* 35, 343 – 357.

Joenväärä, Juha, Robert Kosowski, and Pekka Tolonen, 2016, Hedge fund performance: What do we know, Working paper.

Kaminsky, Kathryn M., 2011, In search of crisis alpha: A short guide to investing in managed futures 1–8.

Kaminsky, Kathryn M., and Alexander Mende, 2011, Crisis alpha and risk in alternative investment strategies 1–8.

Karnaukh, Nina, Angelo Ranaldo, and Paul Söderlind, 2015, Understanding fx liquidity, *The Review of Financial Studies* 28, 3073.

Kazemi, Hossein, and Ying Li, 2009, Market timing of ctas: An examination of systematic ctas vs. discretionary ctas, *Journal of Futures Markets* 29, 1067–1099.

Killeen, William P., Richard K. Lyons, and Michael J. Moore, 2006, Fixed versus flexible: Lessons from {EMS} order flow, *Journal of International Money and Finance* 25, 551 – 579.

King, Michael R., Carol Osler, and Dagfinn Rime, 2011, Foreign exchange market structure, players and evolution, Working Paper 2011/10, Norges Bank.

Kozhan, Roman, Michael Moore, and Richard Payne, 2015, , market order flows, limit order flows and exchange rate dynamics, Working paper.

Liang, Bing, 2004, Alternative investmets: Ctas, hedge funds, and funds of funds, *Journal of Investment Management* 2, 76 – 93.

Love, Ryan, and Richard Payne, 2008, Macroeconomic news, order flows, and exchange rates, *The Journal of Financial and Quantitative Analysis* 43, 467–488.

Mancini, Loriano, Angelo Ranaldo, and Jan Wrampelmeyer, 2013, Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums, *The Journal of Finance* 68, 1805–1841.

Mancini-Griffoli, Thomas, and Angelo Ranaldo, 2011, Limits to arbitrage during the crisis: Funding liquidity constraints and covered interest parity, Working paper, Swiss National Bank.

McElroy, Marjorie B., 1977, Goodness of fit for seemingly unrelated regressions, *Journal of Econometrics* 6, 381 – 387.

Meese, Richard, and Kenneth Rogoff, 1983, Empirical exchange rate models of the seventies: Do they fit out of sample?, *Journal of International Economics* 14, 3–24, See also The Failure of Empirical Exchange Rate Models: No Longer New, But Still True, Economic Policy Web Essay, September 2001. © Copyright Elsevier Science. Posted with permission of Elsevier Science. One copy may be printed for individual use only.

Menkhoff, Lukas, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf, 2016, Information Flows in Foreign Exchange Markets: Dissecting Customer Currency Trades, *Journal of Finance* 71, 601–634.

Mitton, Todd, and Keith Vorkink, 2007, Equilibrium underdiversification and the preference for skewness, *The Review of Financial Studies* 20, 1255–1288.

Moinas, Sophie, Minh Nguyen, and Giorgio Valente, 2017, Funding constraints and market illiquidity in the european treasury bond market, Technical report.

Molodtsova, Tanya, and David H. Papell, 2009, Out-of-sample exchange rate predictability with taylor rule fundamentals, *Journal of International Economics* 77, 167 – 180.

Moore, Michael, Andreas Schrimpf, and Vladyslav Sushko, 2016, Downsized FX markets: causes and implications, *BIS Quarterly Review* 35–51.

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228 – 250, Special Issue on Investor Sentiment.

Obstfeld, Maurice, and Kenneth Rogoff, 2000, The six major puzzles in international macroeconomics: Is there a common cause?, NBER Working Papers 7777, National Bureau of Economic Research, Inc.

Patton, Andrew J., Tarun Ramadorai, and Michael Streatfield, 2015, Change you can believe in? hedge fund data revisions, *The Journal of Finance* 70, 963 –999.

Pesaran, M. Hashem, Yongcheol Shin, and Richard J. Smith, 2001, Bounds testing approaches to the analysis of level relationships, *Journal of Applied Econometrics* 16, 289–326.

Phylaktis, Kate, and Long Chen, 2010, Asymmetric information, price discovery and macroeconomic announcements in fx market: do top trading banks know more?, *International Journal of Finance and Economics* 15, 228–246.

Polkovnichenko, Valery, 2005, Household portfolio diversification: A case for rank-dependent preferences, *The Review of Financial Studies* 18, 1467–1502.

Rime, Dagfinn, Lucio Sarno, and Elvira Sojli, 2010, Exchange rate forecasting, order flow and macroeconomic information, *Journal of International Economics* 80, 72 – 88, Special Issue: {JIE} Special Issue on International Macro-Finance.

Rime, Dagfinn, Andreas Schrimpf, and Olav Syrstad, 2017, Segmented money markets and covered interest parity arbitrage, BIS Working Papers 651, Bank for International Settlements.

Rosenberg, Joshua V, and Leah G Traub, 2009, Price discovery in the foreign currency futures and spot, *The Journal of Derivatives* 17, 7–25.

Stulz, Rene M., 2007, Hedge funds: Past, present, and future, *Journal of Economic Perspectives* 21, 175–194.

Titman, Sheridan, and Cristian Tiu, 2011, Do the best hedge funds hedge?, *The Review of Financial Studies* 24, 123–168.

Zellner, Arnold, 1962, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of the American Statistical Association* 57, 348–368.