Financial crime "hot spots" – Empirical evidence from the foreign exchange market ◊

Florian El Mouaaouy*

Institute for Capital Markets and Corporate Finance, Munich School of Management, Ludwig-Maximilians-Universität München, Schackstraße 4, 80539 München, Germany

This version: September 2015

Abstract. This paper uses a natural experiment to investigate the effects of collusive benchmark manipulation on foreign exchange (FX) trading and market behavior. Specifically, from January 2008 to October 2013, traders illegally coordinated trading strategies in order to obtain informational advantages and to influence FX benchmarks around the London close. Two manipulation measures are constructed which implement an approach to indicate the probability and intensity of potential misconduct during the FX benchmark window. The findings show anomalies and abnormal rate cluster behavior in prosecuted FX data. They contribute to the understanding of irregularities in a highly volatile environment and suggest a simple, practical, and useful approach to study other financial benchmarks, markets and time periods from a regulatory perspective.

JEL classification: D82; G12; G14; G15

Keywords: Capital market regulation; Exchange rate market microstructure; Financial market misconduct; Financial benchmarks; Information asymmetries

◊ I would like to thank Menachem Brenner, Matthias Bröcker, Ralf Elsas, Markus Glaser, Luzi Hail, Bryan Howieson, and seminar participants at the 9th Annual Meeting of the Italian Society of Law and Economics 2013, the 17th Annual Meeting of the Swiss Society for Financial Market Research 2014, the Behavioural Finance Working Group Conference 2015, the 8th International Risk Management Conference 2015, the Munich Finance Day 2015, the 32nd Annual Conference of the European Association of Law and Economics 2015, the Ludwig-Maximilians-Universität München, the Universität Zürich, and the Universität Bayern e.V. I gratefully acknowledge the support from the research scholarship BayEFG.

* Corresponding author. Tel.: +49 (0)89 2180 2211; fax: +49 (0)89 2180 2016.
E-mail address: fl.elmouaaouy@bwl.lmu.de (F. El Mouaaouy).

1. Introduction

Does collusive benchmark manipulation affect trading and market behavior? If so, how can this kind of financial market misconduct be detected? These questions are important as indicated by the substantial antitrust fines against leading financial institutions in recent years (CFTC 2014, FCA 2014, DoJ 2015). Essentially, benchmarks (also known as "fixings") represent an average price during a certain period, the so called benchmark window. Market participants may use these benchmarks as a standard reference price (WMCompany 2011). While much of the literature has focused on the manipulation of closing stock prices (Spatt 2014), an emerging strand of work investigates the manipulation of specific financial benchmarks, such as the London inter-bank offered rate (LIBOR) (Abrantes-Metz et al. 2012, Fouquau and Spieser 2015). However, benchmark manipulation in the FX market has received comparatively little academic attention. This is surprising, because the FX market represents the largest market in the financial system with an estimated average trading volume of 5.3 trillion US dollars (USD) per day (BIS 2013) and the FX benchmark manipulation has led to the largest set of antitrust fines imposed for a violation of the Sherman Act in the history of the Department of Justice (DoJ 2015). Despite the size and importance of the FX market, little is known about it when compared to other financial markets (Mancini, Ranaldo, and Wrampelmeyer 2013).

The manipulation of FX benchmarks is illegal and has potentially far-reaching implications for international financial markets. Vaughan and Finch (2013) and Vaughan, Finch, and Choudhury (2013) use interviews of market participants and experts to investigate so-called "market-impact maximizing trades". These trades cause suspicious patterns of rise and fall in exchange rates that might be attributable to manipulative and collusive practices by cartel members. In their 2013 articles, the authors provide anecdotal evidence that traders of several banks colluded and coordinated their trades for financial gain. Authorities in the

US, UK, Switzerland and in other countries investigated financial institutions and screened their electronic communication. In November 2014, they confirmed that banks' traders collusively pushed large orders in the market in order to maximize the market impact of their trades shortly before and during the benchmark window around the London close at 4 p.m. UK time[1]. As a consequence, the authorities imposed substantial fines on leading global financial institutions totaling several billions of USD (CFTC 2014, FCA 2014, DoJ 2015). However, further investigations are yet to come. Since even small distortions in FX rates and a small change in their digits may cause extensive damage to international financial markets, a question that arises is what can market data tell us about these financial crime "hot spots"? This paper tries to provide some insights into this question by investigating the effects of collusive benchmark manipulation on FX trading and market behavior. The use of behavioral-based screening tools may lead to more successful detection results and help support the long-run confidence in and the stability of financial markets.

The main contribution of this paper is to construct two manipulation measures that indicate the proportion and intensity of irregularities in FX rates and describe the FX trading and market behavior during the benchmark window. Specifically, the first measure refers to abnormal FX rates and calculates the difference between World Markets Company and Reuters (WMR) benchmarks and Thomson Reuters (TR) rates. The second measure decomposes FX rates into several digit positions and investigates to the proportion of digit differences between WMR benchmarks and TR rates. WMR benchmarks represent an average FX rate around 4 p.m., while TR rates are spot FX rates which are snapped at that particular moment in time. These measures are applied to prosecuted FX data and are compared between a collusive and a non-collusive period. Given the highly volatile environment in the FX market (Mancini, Ranaldo, and Wrampelmeyer 2013), the focus on the benchmark window avoids the normal intraday seasonal market behavior (Melvin and

Prins 2011) and accounts for the characteristics of the benchmark calculation. This simple procedure can help both the management board and regulators to screen financial data and has a more practical and useful attitude towards manipulation detection in the FX market than other approaches in the field (Harrington 2008). All two manipulation measures provide a regulatory application for academics, practitioners, and regulators to study different settings for which prosecution information is not readily available, such as other time periods, benchmarks and markets as well.

Importantly, the second manipulation measure addresses a concern of the Benford method. The Benford method refers to an empirical observation of digits in data sets, known as Benford's law (Benford 1938, Newcomb 1881). Many studies apply the Benford test to detect manipulation patterns in reported data (Carslaw 1988, Thomas 1989, Nigrini and Mittermaier 1997, Herrmann and Thomas 2005, Guan, He, and Yang 2006). However, a concern of the Benford approach is that the approximation based on Benford's expected digit frequencies might result in highly significant deviations despite no material manipulation (Mitchell 2001, Mitchell and Izan 2006). To capture this concern, the second manipulation measure provides an approach that is more sensitive to the FX trading and market behavior during the benchmark window. It tests whether and to what extent manipulative "market-impact maximizing trades" affect the digits of FX rates by comparing the digit distribution of WMR benchmarks with TR rates at different digit positions, rather than referring to the Benford distribution. Since even a one digit change in FX rates due to manipulative activities may cause extensive damage, the digit approach of the second manipulation measure represents a promising tool for conducting further research into the FX benchmark manipulation and potentially shedding light on the concerns raised by authorities in a realistic market setting.

The fact that FX benchmarks of the G10 currencies are exposed to misconduct during the collusive period from 1 January 2008 to 15 October 2013 (FCA 2014) provides a clear-cut natural experiment that allows a pre-post comparison. Benchmark manipulation is unlikely during the post-period, because the authorities launched an industry-wide remediation program to ensure a fair and orderly FX market. The FX data are reported by Thomson Reuters Datastream and include nine of the G10 currencies vis-à-vis the USD. The G10 currencies represent the most widely traded currencies in the FX market. The data set includes daily WMR benchmarks which represent one of the most important reference prices and reflect an average rate around the London close at 4 p.m. In addition, the data set includes TR spot FX rates, which reflect the exchange rate at that particular moment in time. If market participants try to maximize the market impact of their trades during the benchmark window, the manipulated WMR benchmark will more likely, and to a higher extent, deviate from the TR rate. All two manipulation measures are applied to the prosecuted FX data and the collusive period is compared to the non-collusive period. On one hand, one might expect to find anomalies and abnormal rate cluster behavior when traders illegally coordinated trading strategies around the London close during the collusive period. On the other hand, it is not expected that FX rates free from manipulation would reflect an empirical bias during the non-collusive period.

Relying on the presented manipulation measures, it is found that anomalies and abnormal cluster behavior are more pronounced when the FX market was exposed to manipulative and collusive practices. The first manipulation measure indicates that the average value of abnormal FX rates is more than twice as high during the collusive period than during the non-collusive period. The second manipulation measure indicates that the digit differences between WMR benchmarks and TR rates are more likely during the collusive period than during the non-collusive period. The GBP/USD, EUR/USD, USD/CHF, and USD/JPY are among the most suspicious currency pairs. In addition, differences are found between the collusive and non-collusive period regarding two liquidity measures, namely, volatility and bid-ask spreads, which supports predictions of microstructure theories. The paper provides evidence corroborating the experimental findings of Cason (2000), the theoretical findings of Hillion and Suominen (2004), and the empirical findings of Comerton-Forde and Putniņš (2011) that describe how manipulative and collusive activities affect trading characteristics and price accuracy. The evidence also mitigates concerns that biased WMR benchmarks stem merely from a highly volatile environment. Rather, it seems that large distortions in FX rates might reflect potential misconduct during the collusive period. This interpretation is consistent with anecdotal evidence reported in Vaughan and Finch (2013), Vaughan, Finch, and Choudhury (2013), and with the findings reported by several authorities (CFTC 2014, FCA 2014, DoJ 2015).

The remainder of this paper is organized as follows. The next section presents the related literature. Section 3 provides background information on the manipulation process of WMR benchmarks. Section 4 develops the hypotheses and describes the manipulation measures. Section 5 presents the data set and descriptive statistics. Section 6 reports the results, section 7 conducts several robustness checks, and the last section concludes.

2. Related literature

The paper is related to two strands of literature. A first line of research is attributable to manipulation patterns in the market microstructure literature. Work by Allen and Gorton (1992), Lyons (1997), and Vitale (2000) touches on market microstructure models that allow for manipulation in equilibrium. Lyons (2001), more broadly, points out that the FX market represents one of the least regulated markets with limited transparency, heterogeneous market participants, and decentralized dealership structure. In an empirical investigation, Comerton-Forde and Putniņš (2011) develop a measure of the probability of closing price manipulation of US and Canadian stocks based on trading characteristics and stock price accuracy. More specific, Abrantes-Metz et al. (2012) and Fouquau and Spieser (2015) use cluster analyses and different comparative statistical techniques to compare LIBOR dynamics with other benchmark indicators of borrowing costs. Yet, much of this research has yielded inconclusive results on trading and market behavior. In addition, there is a gap in research on FX benchmark manipulation that addresses the concerns raised by the authorities in a realistic market setting. This paper thus contributes to and expands this stream of literature as the first manipulation measure provides a benchmark specific approach to study the FX trading and market behavior using prosecuted FX data.

A second line of research is attributable to the literature on cluster analysis of digits. This literature investigates the tendency of prices to deviate from a uniform distribution in many different markets (Doucouliagos 2005). Kahn, Pennacchi, and Sopranzetti (1999), for instance, provide theory and empirical evidence of bank deposit interest rate clustering. Based on their theoretical explanation of behavioral causes, Ashton and Hudson (2008) develop a model of retail banking markets and compare the observed digit distribution of interest rates at different levels of financial involvement with a random digit distribution. In this way, they indicate that firms set interest rates in a manner consistent with maximizing returns from customers. Moreover, Osler (2003) and Osler (2005) conduct cluster analyses of digits in the FX market with regard to currency orders and exchange rate dynamics. In addition, a large and growing body of digit analysis applies the Benford method (Newcomb 1881, Benford 1938) to detect accounting and taxation fraud in reported data (Carslaw 1988, Thomas 1989, Nigrini and Mittermaier 1997, Herrmann and Thomas 2005, Guan, He, and Yang 2006). According to these studies, a deviation from Benford's digit distribution would indicate irregularities in observed data.[2] The second manipulation measure investigates the proportion of digit differences between WMR benchmarks and TR rates at several digit positions. A higher proportion indicates a higher level of suspicious patterns. The paper thus contributes to and expands this line of work as the second measure is more sensitive to the FX trading and market behavior during the benchmark window than the approximation based on Benford's expected digit frequencies.

3. Manipulation of financial benchmarks

In recent years, misconduct encompassing several financial benchmarks has led to substantial antitrust fines. At the beginning of the financial crisis in 2007, liquidity concerns and irregularities cast doubt on reference prices, such as the LIBOR, and challenged the integrity of the inter-bank market (Gilbert 2007). Previous research has found anomalies in market data supporting these suspicions (Mollenkamp and Whitehouse 2008, Abrantes-Metz et al. 2012). Different authorities from around the world started to investigate inter-bank offered rates in 2009 (HMTreasury 2012). They have screened telephone calls, e-mails, and other electronic communications of financial institutions and have detected cases of longstanding manipulation and collusion. However, the detection of manipulated interest rates only reflects the tip of the iceberg, since other benchmarks, such as benchmarks deriving from currency, commodity and other financial markets, are vulnerable to similar issues (HMTreasury 2012).

This paper focuses on WMR benchmarks from the global FX market, which represents one of the most commonly used reference prices. WMR represents a leading global electronic platform for interdealer spot FX trading, which has been active since the late 1990s (Berger et al. 2008). WMR benchmarks are based on actual, executed trades and include the exchange rates at which a currency was bought or sold around 4 p.m. WMR calculates an average price during the benchmark window of 30 seconds[3] before and after the London close. These benchmarks capture FX rates on each trading day and exist for different currencies. They provide market participants with transparency into the FX market. For instance, banks (including central banks), investment management firms, hedge funds, commercial companies, and retail investors rely on this reference rate and may measure FX rates against it. They use WMR benchmarks for valuing, transferring and rebalancing asset portfolios (WMCompany 2011).

Recent work points to suspicious currency spikes with quick reversals in individual currency pairs around the end of the trading day in London. Melvin and Prins (2011) illustrate that the benchmark setting of FX benchmarks creates a conflict of interest between traders at banks and their clients during the benchmark window. Vaughan and Finch (2013) and Vaughan, Finch, and Choudhury (2013) use interviews of market participants and experts to provide anecdotal evidence of illegal practices.[4] They point to traders' attempts to manipulate WMR benchmarks alone or in collusion with traders at other firms. Collusion allows these cartel members to reduce information asymmetries and thereby to improve the chance of influencing certain benchmarks. Recently, different authorities used electronic communication screens to confirm these allegations (CFTC 2014, FCA 2014, DoJ 2015). They noticed that traders were engaged in manipulative and collusive activities for financial benefit and to the potential detriment of banks' clients and other market participants over several years.[5] Although financial markets and their products represent a complex system, the manipulation of FX benchmarks was based on the simple fraud by individuals in terms of inappropriate sharing of confidential information ahead of and during the benchmark window via electronic messaging services, such as chat rooms. They were engaged in a widespread manipulative and collusive effort to influence WMR benchmarks encompassing the period from 1 January 2008 to 15 October 2013 (FCA 2014). The authorities launched an industry-wide remediation program to ensure a fair and orderly FX market in the aftermath of this collusive period. Ongoing FX benchmark manipulation is therefore unlikely during the post-period, which provides a clear-cut natural experiment that allows a pre-post comparison. Fig. 1 illustrates a hypothetical example of this kind of manipulation.

<Please insert Figure 1 about here>

To help visualize the WMR benchmark manipulation, Fig. 1 depicts the process according to the findings of the Financial Conduct Authority (FCA 2014). Banks' clients may place orders to buy or sell a specified volume of a currency against another currency at an uncertain benchmark rate prior to the benchmark time. This agreement exposes banks to movements in the spot FX market. The bank will make a profit if its traders succeed in buying the currency at a lower average rate or in selling the currency at a higher average rate in the market than the resulting WMR benchmark. The authorities notice that banks' traders pushed orders with as much market impact as possible shortly ahead of and during the benchmark window. The traders thereby attempted to influence benchmark rates artificially to an advantageous level and to increase the spread between the WMR benchmark and the banks' average rate. In addition, traders at different banks illegally disclosed confidential bank information between themselves, such as customer identities, order flows and rate spreads, in order to reduce the risk of adverse rate movements and to increase the impact on the rate level. Subsequently, the traders sell or buy the currency at the less favorable WMR benchmark to or from their clients, respectively, than the clients would otherwise have had. This phenomenon of "market-impact maximizing trades" causes an increased trading activity, volatility, and directional exchange rate changes during the benchmark window (Melvin and Prins 2011).

4. Hypotheses development

This section develops the hypotheses and presents the manipulation measures. To account for "market-impact maximizing trades" shortly ahead of and during the benchmark window, two different types of FX rates are investigated. First, WMR benchmarks represent an average price around the London close. Second, TR rates represent spot FX rates which are snapped at 4 p.m. In contrast to the WMR benchmark, the TR rate reflects the exchange rate at that particular moment in time. The focus on the benchmark window avoids the normal intraday seasonal market behavior (Melvin and Prins 2011), which is unrelated to misconduct, but would otherwise drive the results. Moreover, based on the findings of FCA (2014), a clear-cut natural experiment allows a pre-post comparison of a collusive period (from 1 April 2010 to 15 October 2013) with a non-collusive period (from 16 October 2013 to 31 March 2015). The before-after estimator for manipulated FX rates indicates how much larger the values of the manipulation measures are during the collusive period relative to the non-collusive period in the same FX rate. This allows to control for the effects of rate-specific characteristics and possible bias from manipulators selecting non-random FX rates (Comerton-Forde and Putniņš 2011). To compute statistical significance for the subsequent hypothesis tests and to account for suspicious currency spikes, the proportion or the mean value of the variable of interest is calculated, rather than the median value. Statistically significant differences between the two periods would point to potential misconduct. If collusive and manipulative activities are in place, the WMR benchmark will more likely, and to a higher extent, differ from the TR rate. Two indicators are constructed to measure irregularities in the following, namely, abnormal FX rates and digit differences between WMR benchmarks and TR rates.

First, abnormal FX rates use TR rates as closing prices and WMR benchmarks as average prices, i.e., they are conceptually related to Comerton-Forde and Putniņš (2011), who calculate the abnormal trade size based on the mean value per trade of the trades shortly before the close and the mean value per trade of all the values traded during the day. They provide evidence on abnormal day-end returns in the case of manipulation. Note that the measure of this study provides a benchmark specific approach to study the FX trading and market behavior during the benchmark window. This is important because of the highly volatile environment in the FX market (Mancini, Ranaldo, and Wrampelmeyer 2013). In addition, absolute values are computed here in order to account for positive and negative currency spikes on each trading day t, since the FX rate P_t was manipulated in an upward or downward direction (FCA 2014). Formally, the first manipulation measure refers to abnormal FX rates as

$$MM(1)_t \quad = \left|\left(\frac{P_t^{M,TR} - P_t^{M,WMR}}{P_t^{M,WMR}}\right)\right|, \tag{1}$$

where the superscript M represents the mid variable of WMR benchmarks and TR rates, respectively. A higher level of MM (1) indicates a higher intensity of irregularities. Hence, it is hypothesized that

Hypothesis H1. Abnormal FX rates are higher during the collusive period than during the non-collusive period.

The second manipulation measure decomposes FX rates into several digit positions and investigates whether and to what extent manipulative "market-impact maximizing trades" affect the digits of FX rates. This approach draws upon the manipulation detection process in digit analysis.[6] It is crucial to account for single digits in the detection process of FX manipulation, because even a one digit change in FX rates may cause extensive damage. That said, five digit positions are separately investigated since most of the exchange rates are rounded up to 5 decimal places. The counting of the digit position starts with the first non-zero digit from the left. In other words, the first digit represents the digit at the first non-zero position of numbers and so on and so forth. Given the highly volatile environment in the FX market, a manipulation measure is provided that allows for small differences between the WMR benchmark and the TR rate at higher digit positions, but at the same time accounts for suspicious currency spikes at lower digit positions. Intuitively, if market participants collude and push large orders in the market in order to maximize the market impact of their trades shortly before and during the benchmark window, the volatility will increase and manipulated WMR benchmarks will more likely deviate from TR rates. Formally, the second manipulation measure is defined for all first digits $d_1 \in \{1,2,\ldots,9\}$ and the digits $d_j \in \{0,1,2,\ldots,9\}$ at higher positions $j = 2,\ldots,5$, in that an exchange rate has a digit difference with

$$MM(2)_{p,T} = \frac{N_{p,T}^{diff}}{N_{p,T}} \tag{2}$$

where $N_{p,T}^{diff}$ represents the number of digit differences between WMR benchmarks and TR rates at the digit position $p = 1,2,\ldots,5$ and $N_{p,T}$ represents the total number of observations during the period under consideration, T, such as the collusive period, the non-collusive period, a specific year or quarter. A digit difference occurs, if the digit of the WMR benchmark, $d_{p,t}^{WMR}$, differs from the digit of the TR rate, $d_{p,t}^{TR}$, on the trading day t. A higher proportion indicates a higher level of suspicious patterns. The lower the digit position, the higher is the intensity of irregularities. Hence, it is hypothesized that

Hypothesis H2. Digit differences between WMR benchmarks and TR rates are more likely during the collusive period than during the non-collusive period.

Several robustness checks were conducted to validate the findings. The WMR-TR spread is calculated as the proportional quoted spread between the mid variable M of WMR benchmarks and TR rates, respectively, on each trading day t. The proportional quoted spread makes the FX rate P_t of different currencies comparable and is, for instance, used in previous work to calculate the bid-ask spread (Mancini, Ranaldo, and Wrampelmeyer 2013) or matched stock estimates (Comerton-Forde and Putniņš 2011). Here again, absolute values are computed. Formally, the WMR-TR spread represents

$$Spread_t^{(WMR-TR)} = \left| \left(\frac{P_t^{M,WMR} - P_t^{M,TR}}{\left(P_t^{M,WMR} + P_t^{M,TR}\right)/2} \right) \right| . \quad (3)$$

A higher level of WMR-TR spreads indicates a higher intensity of irregularities. If WMR-TR spreads are statistically significantly higher during the collusive period than during the non-collusive period, the null hypothesis will be rejected.

As an additional robustness check, the digit distribution is compared between WMR benchmarks and TR rates. To indicate that market participants attempted to maximize the market impact of their trades around the London close, the digit distribution of manipulated WMR benchmarks would need to deviate from the digit distribution of TR rates. In other words, it is assumed that the digit distribution free from misconduct is the same for WMR benchmarks and TR rates. The manipulative and collusive activities of banks' traders represent a behavioral act, which causes nonrandom patterns of human behavior in FX rates and a bias in the digits of affected rates.[7] For instance, if an exchange rate of 0.72572 increases to 0.72581 during the benchmark window due to manipulative practice, the digits change at the fourth and fifth digit position. Thus, biased digit frequencies of WMR benchmarks should deviate from digit frequencies of TR rates during the benchmark window. Since it would not be expected that WMR benchmarks free from manipulation reflect this behavioral bias in a pronounced way, the approach allows for identifying those currency pairs which are most likely affected by suspicious interventions of market participants. The classical Chi² test for independence is used in order to draw statistical inference.[8] For the number of different digits, d, the usual Pearson statistic is

$$\chi^2_{ind} = \sum_{s=1}^{2} \sum_{i=1}^{d} \frac{\left(O_{s,i} - E_{s,i}\right)^2}{E_{s,i}} . \quad (4)$$

This test compares the observed digit frequencies, $O_{s,i}$, to the frequencies expected, $E_{s,i}$, if the digit distribution were independent of the comparison groups, s, namely, WMR benchmarks and TR rates. The expected frequencies are based on the marginal totals. If there is a treatment effect in that the WMR proportions do not match the TR proportions in a statistically significant manner, the null hypothesis will be rejected.

This digit approach draws upon the Benford method. Over the last three decades, extensive work applied the Benford method to data sets in order to detect digit patterns of potential misconduct. Yet, the set-up of a Benford test was predominantly applied to accounting and taxation data (Carslaw 1988, Thomas 1989, Nigrini and Mittermaier 1997, Herrmann and Thomas 2005, Guan, He, and Yang 2006). In regards to financial markets, Ley (1996) documents that the first digits of the time series of daily returns on stock market indices reasonably agree with the Benford distribution. Subsequently, De Ceuster, Dhaene, and Schatteman (1998), Abrantes-Metz, Villas-Boas, and Judge (2011), and Carrera (2015) apply the Benford method to financial data sets. In contrast, empirical work has had little to say about manipulation patterns of digits in the FX market.[9]

The Benford method refers to an empirical observation of digits in data sets, known as Benford's law (Benford 1938, Newcomb 1881). As regarded in the empirical literature, first digits in many naturally occurring data sets follow a certain logarithmic distribution, instead of a more uniform distribution as is the case with the second or higher digit positions (Raimi 1976). The properties of this mathematical law and its robustness against mathematical transformations result in an unchanged distribution when, for instance, being converted by some fixed number (Pinkham 1961) and when being transformed through arithmetic calculations such as the addition of two Benford sets (Boyle 1994, Hamming 1970). The mathematical probability literature provides more theoretical insights into this significant-digit law and its scale and base invariance (Hill 1995). Basically, the Benford (1938) law implies that a number has the first digit $d_1 \in \{1,2,\dots,9\}$ with probability

$$P(D_1 = d_1) = log\big(1 + (1/d_1)\big)\ . \qquad (5)$$

The screening approaches of previous studies on manipulation detection use different statistical techniques in order to compare the observed empirical digit distribution with the theoretical Benford distribution (Hill 1995, Nigrini and Mittermaier 1997). According to these studies, a deviation from the Benford distribution would indicate irregularities in the observed data. Although it represents a well-known test in many different fields of manipulation detection, a deviation from the Benford distribution is not conclusive evidence of manipulation because it is difficult to prove that non-manipulated FX data should follow Benford's law. The approximation based on Benford's expected digit frequencies might therefore result in highly significant deviations despite no material manipulation. Interpretations linking these Benford deviations to misconduct might be misleading (Mitchell 2001, Mitchell and Izan 2006). Thus, the approximation based on Benford's expected digit frequencies might potentially yield imprecise results in FX markets and so should be interpreted with due caution.

Nonetheless, with this concern in mind, the investigation of statistical anomalies in data sets is likely to represent the most robust instruments for identifying those allegations worthy of closer scrutiny (Harrington 2008). Therefore, the paper uses the Benford test as an additional robustness check. Following previous studies (Cleary and Thibodeau 2005), the classical Pearson Chi² test statistic for goodness of fit is used in order to draw statistical inference:

$$\chi^2_{fit} = \sum_{i=1}^{9} \frac{(O_i - E_i)^2}{E_i} \tag{6}$$

with O_i as observed frequency, and E_i as the frequency expected by Benford. If the observed first digit proportions of FX returns deviate from the expected Benford proportions in a statistically significant manner, the null hypothesis will be rejected. Again, it is noteworthy, however, that statistical evidence of irregularities in time series cannot provide conclusive evidence of misconduct in the FX market. Therefore, caution should be taken when interpreting the results.

Importantly, to mitigate the concerns of the Benford approach described above, the second manipulation measure compares the digit distribution of WMR benchmarks and TR rates, rather than referring to the theoretical Benford distribution. In addition, the second manipulation measure is calculated at different digit positions. The general formula for Benford's law also applies to different (higher than the first) digit positions, but it implies that the digit of interest is dependent (conditional) on all those preceding it (Hill 1995, Mitchell 2001).[10] Any such relation is not in place in the approach adopted here, since the second manipulation measure compares the digits of WMR benchmarks to TR rates. Given the highly volatile environment, the presented measure has the advantage to allow for small changes between the WMR benchmark and the TR rate at higher digit positions, but to account for suspicious currency spikes at lower digit positions. Another advantage of the presented manipulation measures is that no additional data than the data under consideration or complex models are needed (Harrington 2008). Conventional methods are often based on inaccessible data, such as price-cost margins, quantity and demand, on complex models, and on the comparison of different data sets.[11] Thus, this simple procedure has a more practical and useful attitude towards manipulation detection in the FX market than other approaches in the field. It provides a regulatory application for academics, practitioners, and regulators to study potential manipulation of other financial benchmarks of which prosecution information is not readily accessible.

To further test for the robustness of the results, three additional liquidity measures are used, namely, price dispersion, trading cost and return reversal.[12] Previous research on the effects of manipulative and collusive activities on market outcome provides evidence on substantial increases in volatility, bid-ask spread, and return reversal of prosecuted data

(Cason 2000, Hillion and Suominen 2004, Comerton-Forde and Putniņš 2011). Furthermore, the inappropriate disclosure of confidential information with regard to the identity, size and direction of client orders at forthcoming fixes (FCA 2014) might also represent a critical transparency condition. Cartel members get informational advantage over non-members by colluding. The experimental findings of Bloomfield and O'Hara (2000), in contrast, point to marginally significantly narrower bid-ask spreads of those dealers who benefit from less transparency. However, their experimental setting to examine the effect of information asymmetries on market behavior is different in that pre-trade communication is not allowed. The empirical approach adopted here is thus closely related to the experimental setting in Cason (2000), to the theoretical model of Hillion and Suominen (2004), and to the empirical analysis of Comerton-Forde and Putniņš (2011), which explicitly accounts for the manipulation or collusion treatment. The consequences might also apply to the FX market.

In contrast to Ley (1996), returns are calculated as the change in the log WMR benchmark (TR rate) from 4:00 p.m. on day t-1 to 4:00 p.m. on day t (hereinafter returns) which is commonly regarded in the FX microstructure literature:[13]

$$r_t = log\left(\frac{P_t^M}{P_{t-1}^M}\right) . \tag{7}$$

The first liquidity measure is price dispersion which represents a proxy for illiquidity. Mancini, Ranaldo, and Wrampelmeyer (2013) point to an inverse relation between liquidity and volatility. Following Menkhoff et al. (2012) the volatility proxy is measured as the absolute daily log return

$$L_t^{(pd)} = |r_t| . \tag{8}$$

The second liquidity measure refers to the cost aspect of illiquidity. Based on Mancini, Ranaldo, and Wrampelmeyer (2013), a low proportional quoted bid-ask spread indicates a liquid market and is given by

$$L_t^{(ba)} = (P_t^A - P_t^B)/P_t^M \ , \tag{9}$$

where the superscripts M, A and B refer to the mid, ask and bid rates, respectively. The third liquidity measure is return reversal which includes the mid rate at 11 a.m. the following morning, $P_{t+1}^{M,11}$.[14] The formula from Comerton-Forde and Putniņš (2011) is adopted which is defined as

$$L_t^{(rr)} = log\left(\frac{P_t^M}{P_{t+1}^{M,11}}\right) . \tag{10}$$

In addition, following Mancini, Ranaldo, and Wrampelmeyer (2013), returns, bid-ask spreads, and return reversals as well as the manipulation measure MM (1) are multiplied by 10,000 to obtain basis points (bps).

5. Data set and descriptive statistics

The fact that a specific time period is exposed to misconduct (FCA 2014) provides a natural experiment to explore empirical regularities and irregularities in the FX market. The collusive period, which according to the authorities was exposed to illegal coordination of trading strategies, encompasses the period from 1 January 2008 to 15 October 2013 (FCA 2014). The period in the aftermath of 15 October 2013 represents the control group, since the authorities launched an industry-wide remediation program to ensure a fair and orderly FX market. The full sample encompasses the period from 1 April 2010 to 31 March 2015 mainly because of data availability regarding TR rates. Using this period avoids a change to the Authority's Decision Procedure and Penalties Manual (DEPP) for imposing a financial penalty, which was introduced on 6 March 2010 (FCA 2014). Additionally a one year period before and after 15 October 2013 is investigated to avoid changes in the benchmark calculation (untabulated). On 14 December 2014 the calculation window was expanded to a 5 minute benchmark window encompassing 2.5 minutes before and after the benchmark time (WMCompany 2011). The results remain qualitatively unchanged.

According to the findings of the authorities, traders at different banks manipulated currency pairs of the G10 spot FX market (CFTC 2014, FCA 2014), in particular the EUR/USD currency pair (DoJ 2015). The USD is the world's dominant vehicle currency as it is on one side of 87%[15] of all trades in FX deals with the EUR/USD as the most important currency pair representing 24.1% of 2013 global FX market turnover (BIS 2013) and the most liquid currency pair according to the perception of market participants and the liquidity estimates of Mancini, Ranaldo, and Wrampelmeyer (2013). As a consequence, nine of the G10 currencies vis-à-vis the USD are used in the following analyses, namely, Australian dollar (AUD), Canadian dollar (CAD), Swiss franc (CHF), Euro (EUR), British pound (GBP), Japanese yen (JPY), Norwegian krone (NOK), New Zealand dollar (NZD), and Swedish krona (SEK). Thus, currencies which show relatively sparse transaction flows on a daily basis are avoided (Froot and Ramadorai 2005). Table 1 reports descriptive statistics for the daily FX data used in this study.

<Please insert Table 1 about here>

The data set is reported by Thomson Reuters Datastream, representing a total of 92,792 observations. The data set includes daily WMR benchmarks, which are calculated by WMR and reflect the average rate around the London close at 4 p.m. In addition, the data set includes TR spot FX rates, which are taken from the TR proprietary FX rate real-time feed and are snapped (hourly) at that particular moment in time. The data set covers daily bid, ask, and mid rates at 4 p.m.[16] The bid rate is offered by the prospective currency purchaser and the ask rate is offered by the prospective currency vendor. The mid variable is the arithmetic average of the ask and bid variable. Most of the time series contain 1,304 observations over the sample period and are rounded up to 5 decimal places.

6. Results

6.1. Abnormal FX rates

Many market participants use WMR benchmarks as reference prices for the currency market. G10 currencies of these benchmarks were exposed to manipulation and collusion until 15 October 2013 (FCA 2014). This section tries to corroborate this notion from an empirical perspective and compares the collusive period (from 1 April 2010 to 15 October 2013) with the non-collusive period (from 16 October 2013 to 31 March 2015). This natural experiment might indicate irregularities attributable to the manipulative and collusive "market-impact maximizing trades" as described above.

The first manipulation measure, *MM* (1), refers to abnormal FX rates in equation (1). A higher level indicates a higher intensity of irregularities. Fig. 2 illustrates the average values of abnormal FX rates across the nine currency pairs on a daily basis. The graphical depiction indicates structural breaks and points to considerably higher average values of abnormal FX rates for the April to October period and for year-end days. This seasonal market behavior is known from "market-impact maximizing trades" which cause an increased trading activity and volatility around the London close, in particular on year-end days (Melvin and Prins 2011). A slightly different picture emerges with respect to FX returns or volatility in Fig. 3. Here, potential effects of "market-impact maximizing trades" are not revealed as clearly as in Fig. 2. In addition, the intensity with which WMR benchmarks differ from TR rates in Fig. 2 seems to be more pronounced during the collusive period than during the non-collusive period.

<Please insert Figure 2 and Figure 3 about here>

Hypothesis H1 posits that abnormal FX rates are higher during the collusive period than during the non-collusive period. To test this hypothesis, a one-tailed *t*-test is used based on mean values, rather than median values, in order to account for suspicious currency spikes. Differences between the collusive and non-collusive period would indicate potential misconduct.

<Please insert Table 2 about here>

The results in Table 2 indicate that on average the abnormal FX rates are higher during the collusive period (8.82 bps) than during the non-collusive period (3.80 bps). The average value of abnormal FX rates is more than twice as high during the collusive period. This difference of 5.02 bps is statistically significant at the 1% level, which rejects the null hypothesis that abnormal FX rates are not higher during the collusive period than during the non-collusive period. In addition, Table 2 provides a more detailed breakdown of the data of the comparison between the collusive and non-collusive period and reports the results of the mean comparison test for single currency pairs.

Table 2 indicates that on average the abnormal FX rates are higher during the collusive period than during the non-collusive period for all currency pairs. Since the differences are statistically significant at the 1% level, the null hypothesis that abnormal FX rates are not higher during the collusive period than during the non-collusive period is rejected. The USD/CHF displays the largest difference of 8.67 bps.

In sum, these results show empirical evidence on higher abnormal FX rates when the FX market was exposed to manipulative practices (FCA 2014). This evidence is interpreted as support for Hypothesis H1 that abnormal FX rates are higher during the collusive period than during the non-collusive period. These findings are consistent with the notion that traders attempted to influence FX benchmarks artificially to an advantageous level during the benchmark window. A possible explanation for the differences between the collusive and non-collusive period might refer to the suspicious “market-impact maximizing trades” shortly ahead of and during the benchmark window as described above.

6.2. Digit differences

The second manipulation measure, *MM* (2), in equation (2) is designed to decompose the digits of FX rates and to indicate the proportion of digit differences between WMR benchmarks and TR rates at several digit positions. To obtain proportions, this measure calculates the number of digit differences and divides this number by the total number of observations at each of the five digit positions, separately. A higher proportion indicates a higher level of suspicious patterns. The lower the digit position of digit differences, the higher is the intensity of irregularities. This allows for evaluating the intensity of suspicious patterns in FX rates encompassing several digit positions. The counting of the digit position starts with the first non-zero digit from the left. In other words, the first digit represents the digit at the first position of numbers and so on and so forth. For instance, if the WMR benchmark is 0.72572 and the TR rate is 0.72581 at 4 p.m., the digits differ at the fourth and fifth digit position. Fig. 4 provides an overview of the proportions of digit differences between WMR benchmarks and TR rates. The lines illustrate the average value of *MM* (2) across the nine currency pairs on a quarterly (upper graph) and annual (lower graph) basis. The first through fifth digit position are separately investigated.

<Please insert Figure 4 about here>

The upper graph in Fig. 4 indicates a higher average value of *MM* (2) at the third and fourth digit position in quarters two and three, while the average levels at the first, second, and fifth digit position stagnate throughout all quarters. The variation at the third and fourth digit position resembles the seasonal market behavior of the first manipulation measure in Fig. 2. and is known from “market-impact maximizing trades” which cause an increased trading activity and volatility around the London close (Melvin and Prins 2011). Similarly for the annual breakdown (untabulated), the lower graph in Fig. 4 depicts a stagnating average value of *MM* (2) at the first, second, and fifth digit position over time. The average value ranges on a low level between 0.00% and 0.51% at the first digit position and between 0.87% and 5.09% at the second digit position, while the average ranges on a higher level between 69.44% and 76.66% at the fifth digit position. In contrast to the lower and upmost digit positions, the average level ranges between 10.76% and 38.35% at the third digit position and between 47.57% and 71.74% at the fourth digit position. The high variation at the middle digit positions is attributable to the decrease in the proportion of digit differences in the aftermath of the 2013 allegations. It seems that the digit differences between WMR benchmarks and TR rates are more likely during the collusive period than during the non-collusive period.

Hypothesis H2 posits that digit differences between WMR benchmarks and TR rates are more likely during the collusive period than during the non-collusive period. To test this hypothesis, two comparison tests of proportions are conducted. One is based on Pearson's Chi² test statistic for independence in equation (4) in order to compare the distribution of digit differences across the nine currency pairs. Here, d represents the number of different currency pairs and s represents the comparison groups, namely, the collusive and non-collusive period. The other test refers to a one-tailed Z-test in order to compare the values of *MM* (2) for single currency pairs between the collusive and non-collusive period. Differences between the two periods would indicate potential misconduct.

<Please insert Table 3 about here>

The Chi² statistics point to differences between the collusive and non-collusive period at the first through fifth digit position in Table 3. All differences are statistically significant at the 1% level, which reject the null hypothesis that the proportions of digit differences during the collusive period match the proportions of digit differences during the non-collusive period.

In addition, the Z-statistics indicate large differences in the values of *MM* (2) for single currency pairs between the collusive and non-collusive period (Table 3). The USD/JPY and NZD/USD show statistically significant differences at conventional levels at the first through fifth digit position. Furthermore, the differences are statistically significant at conventional levels at the second through fifth digit position for the USD/CHF and EUR/USD as well as at the third through fifth digit position for the GBP/USD. That is, the null hypothesis that the proportion of digit differences is not higher during the collusive period than during the non-collusive period is rejected in most cases.

Taken together, the result for the first digit position is weaker than the results for all other digit positions because of the stickiness at the first digit position of all currency pairs (see Table 1) and a low value of *MM* (2) of 0.32% over the sample period (Table 3). The data show the most suspicious patterns at the second through fifth digit position. Given the large size of the currency market, it is not surprising that the value of *MM* (2) at the second digit position (3.83%) is very low over the sample period. Nonetheless, statistically significant differences between the collusive and non-collusive period is suspicious because "market-impact maximizing trades" would require much effort to affect the second digit position. The proportions of digit differences are higher at the third (27.84%) and fourth digit position (65.20) over the sample period. The highest proportion of digit difference is at the fifth digit position (72.12%) which appears plausible because of the highly volatile environment in the FX market.

In sum, the results provide strong evidence supporting Hypothesis H2 that digit differences between WMR benchmarks and TR rates are more likely during the collusive period than during the non-collusive period. The empirical evidence is in line with the notion that market participants colluded and attempted to maximize the market impact of their trades around the London close during the collusive period. The interpretation that certain currency pairs are most likely affected by suspicious interventions of traders is also consistent with previous results that rely on interviews (Vaughan and Finch 2013, Vaughan, Finch, and Choudhury 2013) or electronic communication screens (CFTC 2014, FCA 2014, DoJ 2015). Suspicious patterns ranging from the fifth to the second digit position would have potentially far-reaching implications and would therefore support the substantial antitrust fines. Thus, it is concluded that higher values of the manipulation measure could potentially reflect collusive behavior among traders after controlling for seasonal market behavior. The proposed screening approach provides a basis for identifying those benchmarks worthy of closer scrutiny. Academics, practitioners, and regulators could focus further investigations on these benchmarks.

However, a caveat of this type of analysis is that the Chi² statistic is sensitive to the number of observations and that a significant deviation is not conclusive evidence of a manipulative practice. To address this concern and to control for seasonal market behavior, further robustness checks are described in the following section.

7. Robustness checks

7.1. WMR-TR spreads

Several robustness checks were conducted to validate the findings. The proportional quoted WMR-TR spread is calculated to detect large differences between WMR benchmarks and TR rates. A higher level indicates a higher intensity of irregularities. Fig. 5 depicts the average values of WMR-TR spreads across the nine currency pairs on a daily basis. Similar to abnormal FX rates in Fig. 2, the graphical depiction indicates structural breaks and points to higher average values of WMR-TR spreads during the collusive period than during the non-collusive period.

<Please insert Figure 5 about here>

The null hypothesis that WMR-TR spreads are not higher during the collusive period than during the non-collusive period is tested by relying on a one-tailed *t*-test based on mean values. Differences between the collusive and non-collusive period would indicate potential misconduct.

<Please insert Table 4 about here>

Similar to the results for the first manipulation measure, Table 4 indicates that, on average, WMR-TR spreads are higher during the collusive period (8.83 bps) than during the non-collusive period (3.80 bps). The average value of the spreads is more than twice as high during the collusive period. This difference of 5.03 bps is statistically significant at the 1% level and allows for rejecting the null hypothesis that WMR-TR spreads are not higher during the collusive period than during the non-collusive period.

All currency pairs show a higher average spread during the collusive period in a highly significant manner (Table 4). Here again, the USD/CHF stands out as most suspicious with the highest average WMR-TR spread (12.25 bps) during the collusive period and the largest difference between the two periods (8.73 bps). Thus, the results hold for the first manipulation measure.

Taken together, these findings provide evidence on larger proportional quoted spreads between WMR benchmarks and TR rates when banks' traders were involved in collusive activities (FCA 2014). This interpretation seems consistent with the notion that WMR-TR spreads are higher during the collusive period than during the non-collusive period. To find more about potential FX manipulation, the following analysis further examines isolated digit positions of FX rates.

7.2. Digit distribution

To indicate a treatment effect in the digits of FX data, the digit distribution between WMR benchmarks and TR rates is compared at different digit positions based on equation (4). Here, d represents the number of different digits and s represents the comparison groups, namely, WMR benchmarks and TR rates. Differences in the digit distribution among the comparison groups would indicate suspicious patterns.

<Please insert Table 5 about here>

An interesting pattern emerges from Table 5. While the digit distribution of WMR benchmarks generally conforms to the digit distribution of TR rates fairly closely, the EUR/USD shows a statistically significant deviation at the second digit position during the collusive period. This deviation is marginally significant (p = 0.089). In addition, the USD/CHF (p = 0.000) and the USD/JPY (p = 0.000) show a highly and statistically significant deviation at the fifth digit position during the collusive period. This pattern is particularly suspicious, since the WMR and TR distribution of both currency pairs conform fairly closely during the non-collusive period.

Against expectations, the results indicate statistically significant deviations for the USD/SEK (significant at the 5% level) at the fourth digit position during the non-collusive period as well as for the AUD/USD (significant at the 1% level), the USD/CAD (significant at the 5% and 10% level, respectively), and the NZD/USD (significant at the 1% level) at the fifth digit position during the collusive and non-collusive period.

A potential explanation for these unexpected findings might be that the digits at the fourth and fifth digit position are highly volatile and generally differ between WMR benchmarks and TR rates if not necessarily manipulation is in place. This is consistent with the results in Table 3. The USD/SEK shows no statistically significant difference between the collusive and non-collusive period at the fourth digit position. A similar explanation might apply to the USD/CAD and NZD/USD at the fifth digit position, which show no or only marginally statistically significant differences between the collusive and non-collusive period. The AUD/USD shows no statistically significant differences at most digit positions in Table 3.

In sum, the results on biased digit distributions corroborate the findings for the second manipulation measure. They lend support to the notion that there is a treatment effect on the digits of FX rates during the collusive period.

7.3. Benford test

The traditional Benford test of financial time series would test the null hypothesis that the first digit distribution of observed returns equals the first digit distribution expected by Benford's law (Ley 1996). To test this null hypothesis, the first digit distribution of daily WMR returns is compared to the Benford distribution. The same test is conducted with TR returns with the expectation of finding similar results. Following previous studies (Nigrini and Mittermaier 1997, Cleary and Thibodeau 2005), the results stem from Pearson's Chi^2 test statistic for goodness of fit to test the overall fit based on equation (6) and the Z-statistic to test differences in individual digits.

Fig. 6 illustrates the first digit proportions of daily FX returns. Panel A includes the first digits of all nine WMR currency pairs and Panel B covers the first digits of all nine TR currency pairs. The left graph encompasses the collusive period, while the right graph encompasses the non-collusive period in both panels. Bars represent the first digit proportions of WMR and TR returns, respectively, while the solid line displays the first digit proportions expected by Benford's law. The capped spikes for each first digit indicate the upper and lower bounds of the 99% confidence interval as measured by the Z-statistics.

<Please insert Figure 6 about here>

Fig. 6 depicts similar patterns of the first digit distribution. All four graphs point to the stylized facts seen in other studies (Ley 1996) that the first digit distribution of returns is skewed towards lower first digits and approximates the monotonic decline known from the

Benford distribution. That is, the WMR and TR sample display more lower first digits, such as the digit one, than higher first digits, such as the digit 9. The left graph in Panel A indicates statistically significant differences at the 1% level in roughly all first digits of WMR returns, except for the digits 4 and 5 during the collusive period. With a Chi² statistic of 167.91, the overall first digit distribution of WMR returns deviates from the Benford distribution in a highly significant manner (significant at the 1% level). During the non-collusive period, in contrast, the right graph in Panel A indicates statistically significant differences in the digit 1 (significant at the 1% level), digits 5 and 6 (significant at the 5% level), and digit 2 (significant at the 10% level), while the proportions of all other first digits remain within the conventional confidence intervals. However, with a Chi² statistic of 24.35, the overall first digit distribution of WMR returns deviates from the Benford distribution in a highly significant manner (significant at the 1% level) as well. Similar digit patterns are found for the TR returns in Panel B.

Taken together, the graphical depiction suggests that the first digit distribution of FX returns converges to the Benford distribution. The results provide empirical evidence that the overall first digit distribution of FX returns deviates from the Benford distribution in a more pronounced manner during the collusive period. However, the null hypothesis that the first digit distribution of observed FX returns equals the Benford distribution during both the collusive and non-collusive period must be rejected. These findings are consistent with the notion that the approximation based on Benford's expected digit frequencies might result in highly significant deviations despite no material manipulation and potentially yield misleading results in financial markets (Mitchell 2001, Mitchell and Izan 2006). Whether this finding holds in a currency breakdown or whether the investigation of single currency pairs provides more robust manipulation patterns is analyzed below.

<Please insert Table 6 about here>

A slightly different picture emerges with respect to the test statistics for the overall fit of WMR and TR returns investigating single currency pairs in Table 6. The table indicates a statistically significant difference between the first digit distributions of WMR returns and the Benford distribution for most currency pairs during the collusive period. The null hypothesis that the first digit distribution of WMR returns equals the Benford distribution is rejected for the AUD/USD, USD/CHF, GBP/USD, USD/NOK, NZD/USD, and USD/SEK (significant at the 1% level) and for the EUR/USD and USD/JPY (significant at the 5% level).

The overall first digit distribution of WMR returns deviates from the Benford distribution in a less pronounced manner during the non-collusive period. However, the null hypothesis that the first digit distribution of observed FX returns equals the first digit distribution expected by Benford’s law has to be rejected for the AUD/USD (significant at the 1% level) as well as for the USD/CAD and USD/SEK (significant at the 5% level). These exceptions are similar to the unexpected findings in Table 5 in the section above.

In addition, Table 6 reports the test results of the comparison between the first digit distribution of TR returns and the Benford distribution for single currency pairs. The results are similar to those reported before. The Chi^2 statistics indicate more pronounced deviations during the collusive period than during the non-collusive period.

Overall, the findings for the currency breakdown show empirical evidence on more pronounced Benford deviations during the collusive period than during the non-collusive period. Although the approximation based on Benford’s expected digit frequencies results in statistically significant deviations during the non-collusive period in some cases, the results of the Benford test validate this study’s findings and support the notion that market participants attempted to manipulate WMR benchmarks (FCA 2014).

7.4. Liquidity measures

Table 7 reports the results for the liquidity measure. The first liquidity measure, price dispersion $L^{(pd)}$, is a volatility proxy and is defined as the absolute daily log return given in equation (8). The second liquidity measure, trading cost $L^{(ba)}$, is the proportional quoted bid-ask spread based on equation (9). The third liquidity measure, return reversal $L^{(rr)}$, refers to equation (10). The null hypothesis that the respective liquidity measure is not higher during the collusive period than during the non-collusive period is tested by relying on a one-tailed *t*-test.

<Please insert Table 7 about here>

The data in Table 7 show a higher volatility and a larger bid-ask spread during the collusive period than during the non-collusive period. The difference between the collusive and non-collusive period is statistically significant at conventional levels. In contrast, the results for return and return reversal in Table 7 report no significant differences between the collusive and non-collusive period. The measure of return reversal is taken from Comerton-Forde and Putniņš (2011) who analyze closing price manipulation of US and Canadian stocks. They investigate the return from the closing price to the price at 11 a.m. the following morning. The difference in results found between Comerton-Forde and Putniņš (2011) and the present study might be due to the large time lag between the closing and morning price, since, in contrast to stocks, FX trading continues after the London close. However, the results remain qualitatively unchanged if a narrower time lag that includes the 5 p.m. rate of the same day is used instead (untabulated). Given the highly volatile environment in the FX market, this one-hour lag might still be too large, since Mancini, Ranaldo, and Wrampelmeyer (2013), for instance, use one-minute data with a lag length of 5 to estimate the return reversal of FX rates. Their theoretical model for estimating liquidity dimensions of price impact and return reversal is conceptually related to Kyle (1985) and measures the relation between exchange rate changes and order flow. Further research on manipulation detection in the FX market could investigate these patterns using high-frequency data.

In sum, the results of this study for the first and second liquidity measure are supported by Hillion and Suominen (2004) who point to an increase in volatility and bid-ask spreads in the last minutes of trading in the case of manipulation. This finding is also in line with the evidence on wider spreads when dealers collude in asset market experiments (Cason 2000) and when market participants manipulate closing stock prices (Comerton-Forde and Putniņš 2011). The findings thus add to the evidence on the effects of the manipulation and collusion treatment on volatility and bid-ask spreads in the FX market.

That said, it should be noted that the highly volatile environment might also drive the effects on differences between WMR benchmarks and TR rates during the benchmark window. Melvin and Prins (2011) report a concentrated trade activity and volatility around the London close. It is therefore likely that WMR benchmarks differ from TR rates, if not necessarily manipulation is in place.

To alleviate concerns that the normal volatility behavior might drive the results, the manipulation measure *MM* (1) is normalized in dividing abnormal FX rates by the volatility proxy.[17] The results are reported in Table 8.

<Please insert Table 8 about here>

The results for the normalized manipulation measure in Table 8 show generally higher average values during the collusive period than during the non-collusive period. The two Scandinavian currencies, NOK and SEK, form an exception and show lower average values during the collusive period. The differences are statistically significant in a highly significant manner for the USD/CHF (significant at the 5% level), GBP/USD (significant at the 5% level), and USD/JPY (significant at the 1% level).

Overall, the results for the normalized manipulation measure alleviate concerns that volatility might drive the manipulation measure. Although higher volatility during the collusive period is generally found in Table 7, a possible explanation for the significant results for the USD/CHF, GBP/USD, and USD/JPY in Table 8 might be that the difference between WMR benchmarks and TR rates stem not merely from the normal volatility behavior, but rather from manipulative activities in certain exchange rates. Further research on manipulation detection could investigate volatility and other market proxies during the benchmark window, such as volume, or trading activity of individual financial institutions using high-frequency data.

Taken together, the results show empirical evidence on suspicious patterns for those currency pairs in which the literature on market microstructure suggests manipulative and collusive activities (Cason 2000, Hillion and Suominen 2004, Comerton-Forde and Putniņš 2011). This evidence is interpreted as support for the hypotheses that manipulated WMR benchmarks will more likely, and to a higher extent, deviate from TR rates. This should be particularly the case, if market participants collude and push large orders in the market to maximize the market impact of their trades shortly before and during the benchmark window.

8. Conclusion

Trading and market behavior is an important yet understudied aspect of collusive benchmark manipulation. This paper sheds light on controversial "market-impact maximizing trades" and tests predictions of market microstructure theories in the FX market. It addresses a gap in the literature by investigating how informational advantage through collusion and trade coordination affects FX benchmarks. In an experimental setting, two manipulation measures are constructed that contribute to understanding more about the probability and intensity of potential misconduct in prosecuted FX data. The manipulation measures *MM* (1) and *MM* (2) contribute to the literature on market microstructure and on cluster analysis of digits, respectively, as they provide a benchmark specific approach to detect manipulation patterns in a highly volatile environment. These measures can be used in an academic and practical application to study whether and to what extent manipulation of asset prices might be in place, although prosecution information is not readily accessible.

The findings show anomalies in and abnormal cluster behavior of certain benchmarks during the collusive period, in particular of the GBP/USD, EUR/USD, USD/CHF and USD/JPY, those currencies which were also exposed to the LIBOR rigging. A potential explanation consistent with the findings is that coordinated interventions between market participants may improve the chance of influencing the benchmark rate to their advantage. It appears that traders have to reduce the high risk of mismanagement by sharing order information and coordinating trades with other banks. In view of the immense size of one of the least regulated markets in the financial system, where four dominant banks combine a market share of over 50% among hundreds of other financial institutions (Euromoney 2013), collusive activities among the main market participants seem to be essential to influence FX benchmarks. In this way, collusive "market-impact maximizing trades" could have an effect on the digit distribution of FX benchmarks and potentially influence even the second digit position. This interpretation seems in line with anecdotal evidence in Vaughan and Finch (2013) and Vaughan, Finch, and Choudhury (2013) and seems consistent with the findings reported by several authorities (CFTC 2014, FCA 2014, DoJ 2015). It is thus concluded that higher values of the manipulation measures could potentially reflect collusive and manipulative activities after controlling for seasonal market behavior. Hence, the presented approach could detect those FX benchmarks which are most likely affected by these manipulative and collusive activities of FX market participants.

The findings are related to key policy concerns of financial regulators and highlight far-reaching implications for international financial markets. They contribute to the understanding of anomalies in certain currency pairs. These nonrandom patterns of human behavior indicate suspicious interventions of market participants. One possible explanation may be that this behavioral bias arises from disruptive practices of traders at banks influencing FX benchmarks, all to be taken with the caveat that a significant deviation is not conclusive evidence of misconduct in the FX market. Another explanation of these irregularities might be attributable to seasonal market behavior, such as abnormal trading activity. The study accounts for this alternative explanation, but further research is required in this area. Consequently, the study's conclusions should be interpreted with due caution.

Harrington (2008) suggests three steps in the process of the detection of misconduct including screening, verification, and prosecution. The presented approach is an initial step to screen moving periods or specific time periods of suspicious financial benchmarks and markets in order to detect irregularities. The approach can also be used to investigate tick data within the benchmark window. The detection of irregularities worthy of closer scrutiny plays an important role because investigations beyond G10 currencies are yet to come (FCA 2014) and other financial markets are vulnerable to similar issues (HMTreasury 2012). As second step, anomalies can then be further investigated using conventional methods, such as electronic communication screens, in order to verify misconduct in a more costly procedure. As third step, the final task of prosecution is to provide conclusive evidence of coordination. For these reasons, using the presented empirical design as a routine and universal tool for screening financial markets is recommended. This procedure can help both the management board and regulators to validate their surveillance system, detect potential financial crime "hot spots" and introduce further investigations in a timely manner, and to provide statistical evidence of manipulation in enforcement procedures.

Notes

[1] UK time is used throughout the paper.

[2] A deviation from the Benford distribution is not conclusive evidence of manipulation, but it represents a well-known test in many different fields of manipulation detection.

[3] On 14 December 2014 the calculation window was expanded to a 5 minute benchmark period encompassing 2.5 minutes before and after the benchmark time. For detailed information on the calculation method see WMCompany (2011). This change is controlled for in the robustness checks. The results remain qualitatively unchanged.

[4] For a definition on illegal price manipulation in financial markets see Kyle and Viswanathan (2008). Spatt (2014) reviews the existing literature and contemporary developments on financial market manipulation and artificial market pricing.

[5] European Central Bank (ECB) benchmarks are also exposed to misconduct. In contrast to WMR benchmarks, the ECB benchmark is based upon the exchange rate for several spot FX currency pairs at 1:15 p.m.

[6] For further information on the Benford method see Newcomb (1881), Benford (1938), Nigrini (1996), and Nigrini and Mittermaier (1997).

[7] For examples on how banks engaged in manipulation and made significant profits see FCA (2014).

[8] For further information on the standard Pearson Chi² test statistics for goodness of fit and independence see Rao and Scott (1981).

[9] For previous research and empirical evidence on rate clustering in the FX market see Mitchell and Izan (2006). Carrera (2015) apply the Benford method to Latin American exchange rates in order to track exchange rate management and to investigate policy interventions.

[10] Financial time series themselves generally do not conform to Benford’s law at the first digit position. Ley (1996), for instance, therefore investigates returns.

[11] For further screening methods see Harrington (2008).

[12] Mancini, Ranaldo, and Wrampelmeyer (2013) use intraday data and distinguish between three categories of liquidity measures. They compute price impact and return reversal, the proportional quoted bid-ask spread to obtain trading cost, and the two–scale realized volatility to obtain price dispersion.

[13] Further recent studies that use the change in the log exchange rate include Evans and Lyons (2002), Froot and Ramadorai (2005), Bacchetta and Van Wincoop (2006), and Mancini, Ranaldo, and Wrampelmeyer (2013). A survey of the existing research in FX market microstructure is given by Marsh and O'Rourke (2005) and Osler (2012).

[14] The results remain qualitatively unchanged, if a narrower time lag is used including the 5 p.m. rate of the same day in the formula instead of the 11 a.m. rate of the following day.

[15] Net-net basis, adjusted for local and cross-border interdealer double-counting. The sum of the percentage shares of average daily turnover in April 2013 of individual currencies totals 200% instead of 100%, since two currencies are involved in each transaction (BIS 2013).

[16] TR mid rates at 11 a.m. and 5 p.m. are used for additional robustness checks.

[17] See Zhou (1996) who calculates daily returns normalized by its volatility.

Figures

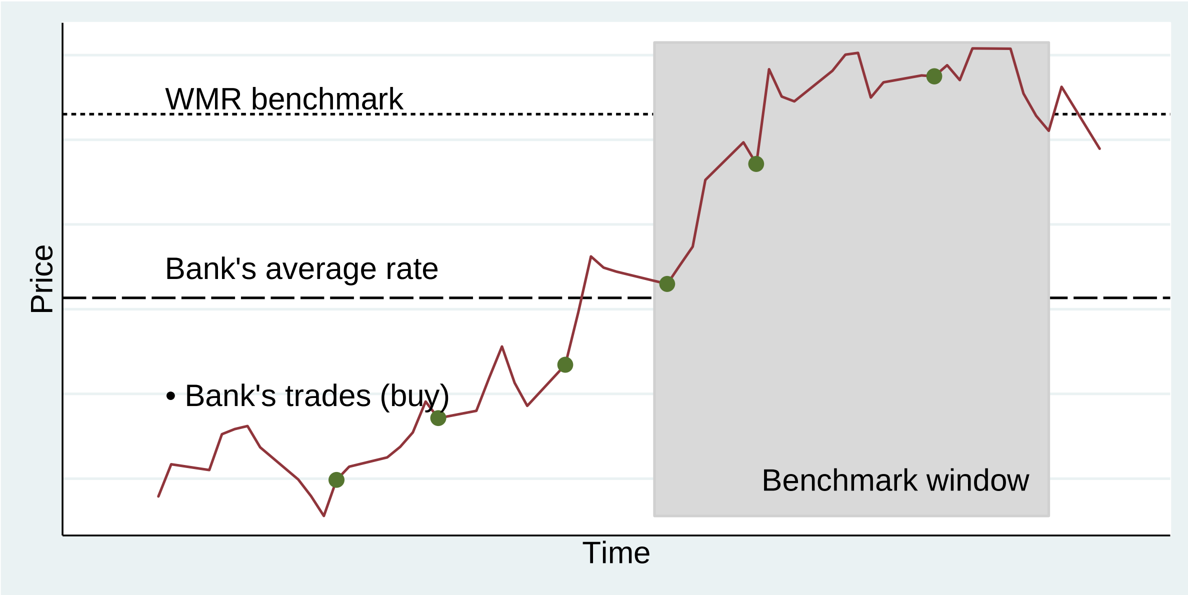


Figure 1. Hypothetical example of WMR benchmark manipulation based on FCA (2014). The figure illustrates the FX rate movement of a given currency shortly before and after the benchmark time at 4 p.m. The solid line indicates the FX rate. The dots indicate the exchange rates at which the bank trades and buys the currency. The long dashed line indicates the bank's average FX rate of these trades. The benchmark window encompasses 30 seconds before and after the benchmark time. The short dashed line indicates the WMR benchmark which represents the average value of the FX rates during the benchmark window.

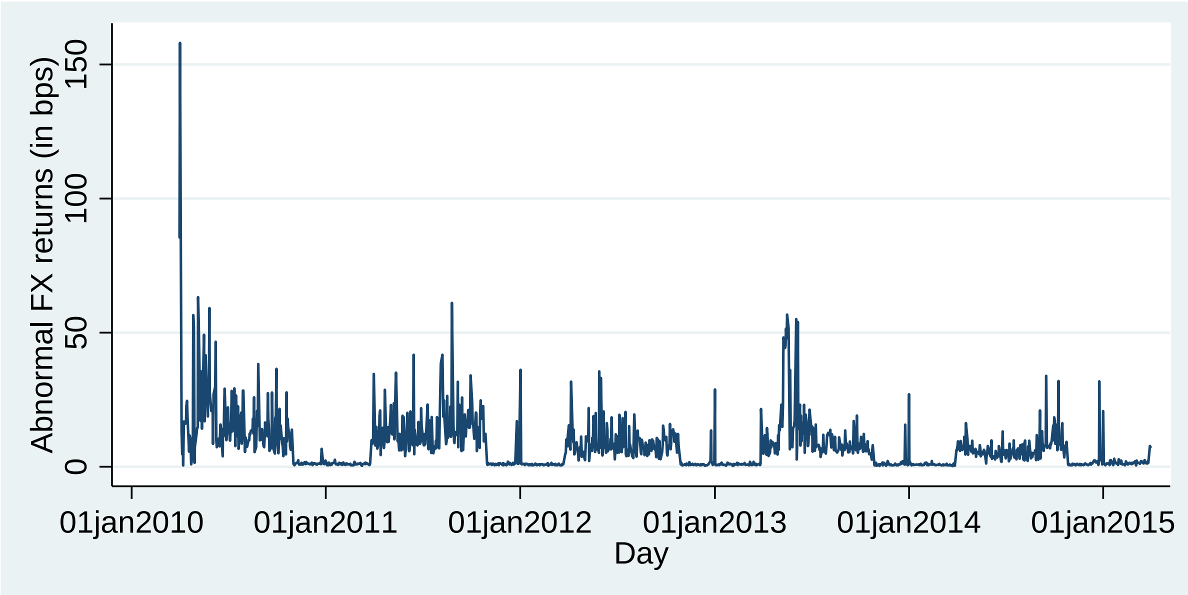
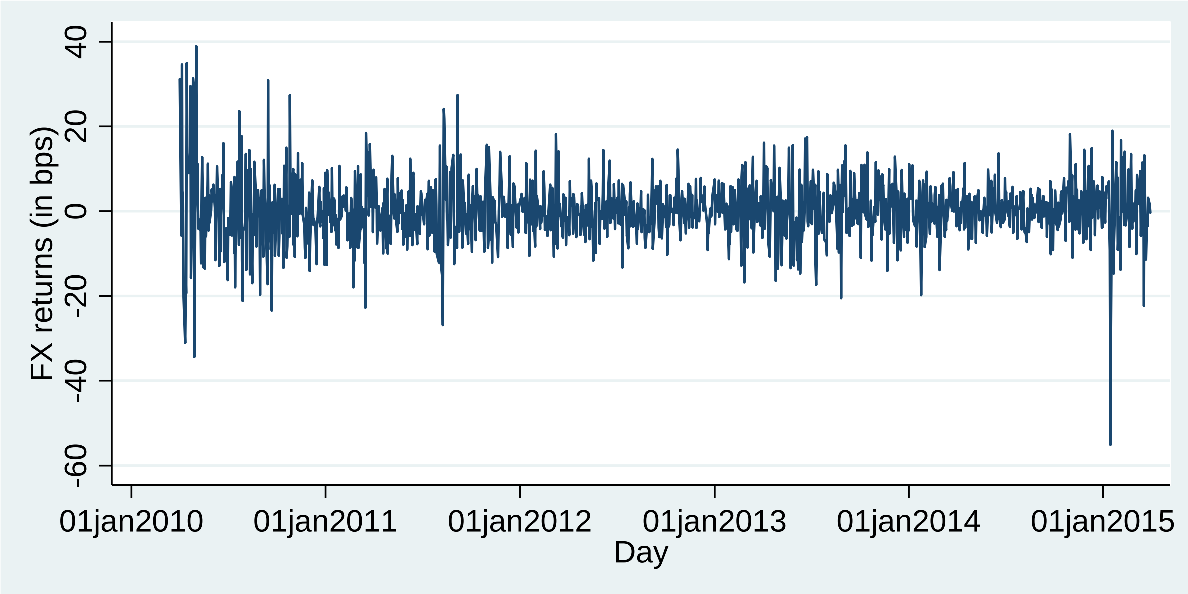


Figure 2. Trading and market behavior according to *MM* (1) in bps. The line illustrates the intensity that the WMR benchmarks differ from the TR rates based on equation (1). It represents the average of abnormal FX rates across the nine currency pairs on a daily basis. The collusive period is from 1 April 2010 to 15 October 2013, while the non-collusive period is from 16 October 2013 to 31 March 2015.

FX returns



FX volatility

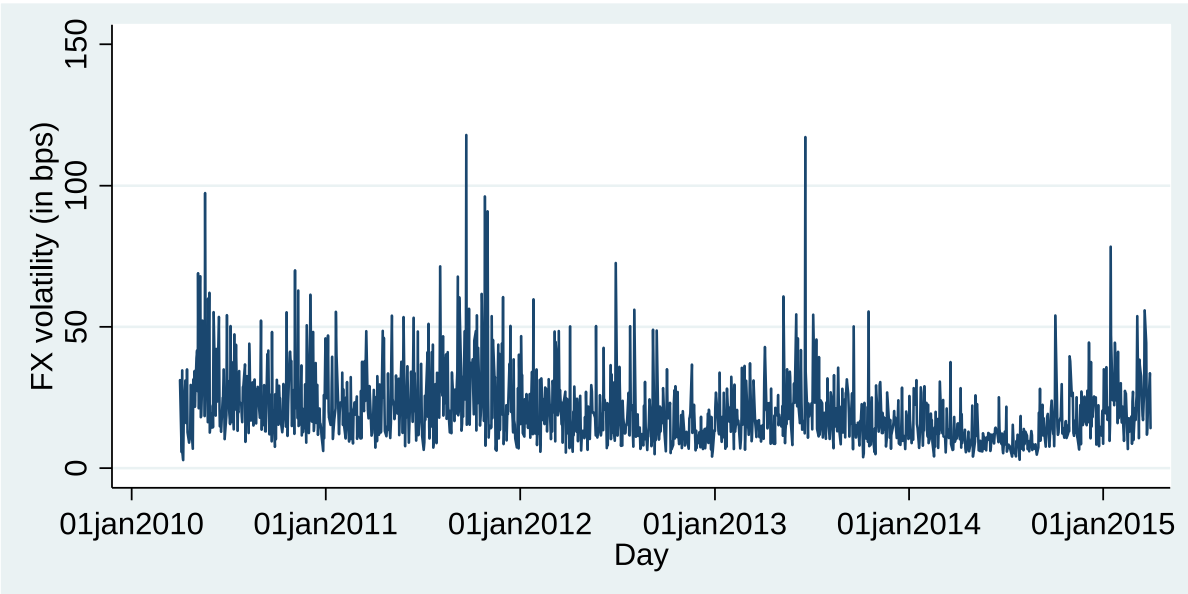
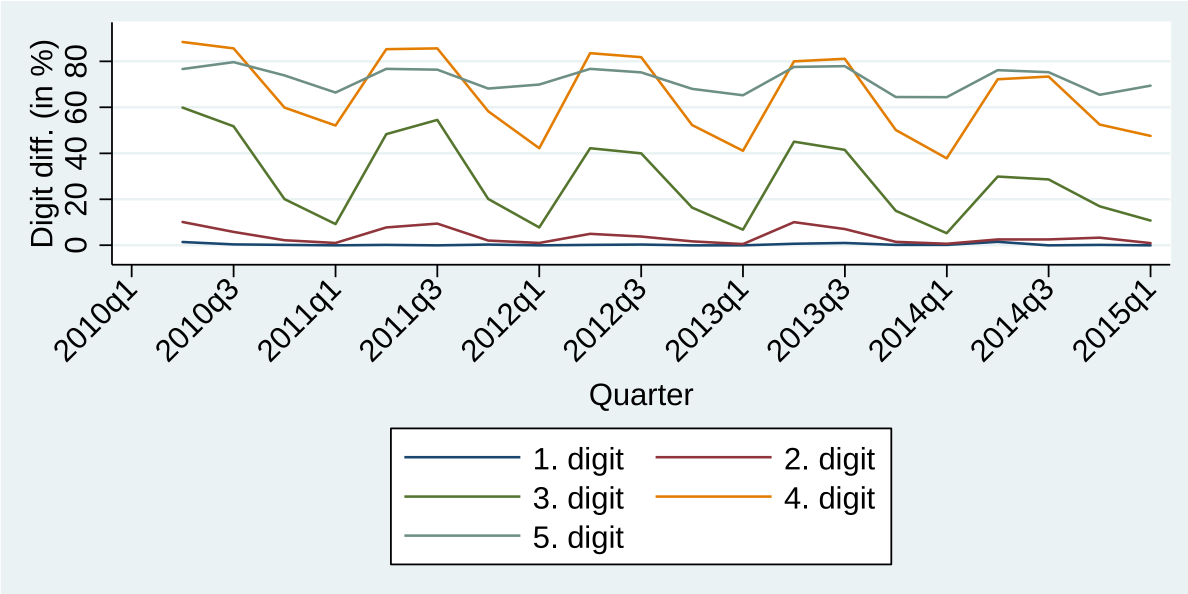


Figure 3. Daily FX returns (upper graph) and FX volatility (lower graph) based on TR rates. Return refers to the change in the log FX rates between two trading days in equation (7). Volatility is defined as the absolute daily log return given in equation (8). The lines illustrate the average across the nine currency pairs in bps. The collusive period is from 1 April 2010 to 15 October 2013, while the non-collusive period is from 16 October 2013 to 31 March 2015.

MM (2) on a quarterly basis



MM (2) on an annual basis

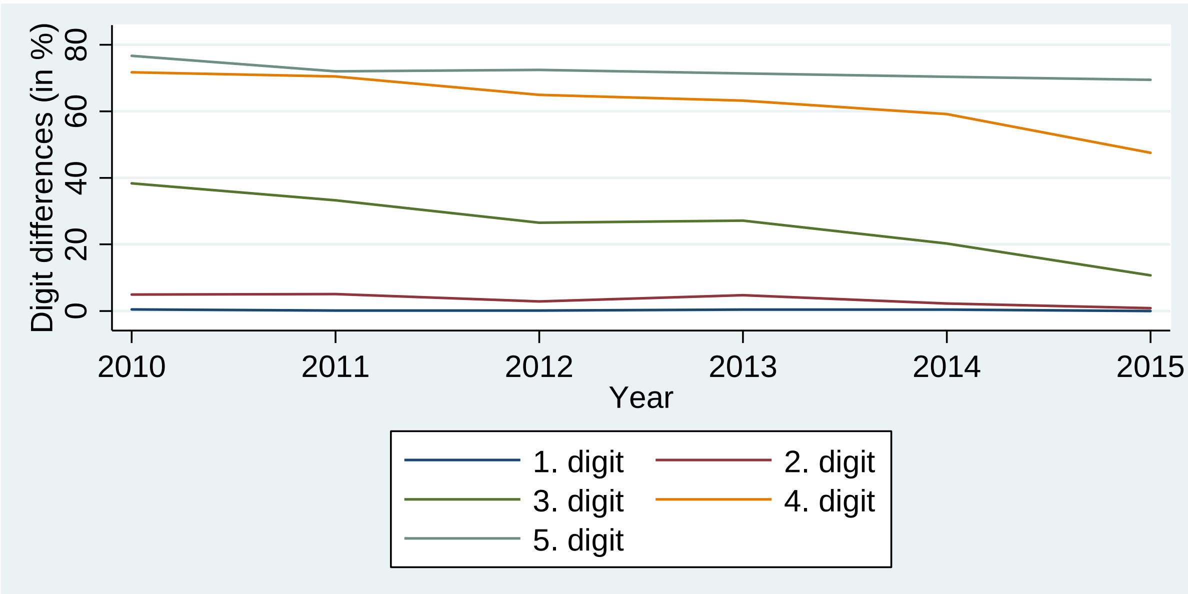


Figure 4. Trading and market behavior according to *MM* (2) in percent. The lines illustrate the proportion that the digits of WMR benchmarks differ from the TR rates at the first through fifth digit position based on the second manipulation measure in equation (2). They represent the average across the nine currency pairs on a quarterly basis (upper graph) and on an annual basis (lower graph). The collusive period is from 1 April 2010 to 15 October 2013, while the non-collusive period is from 16 October 2013 to 31 March 2015.

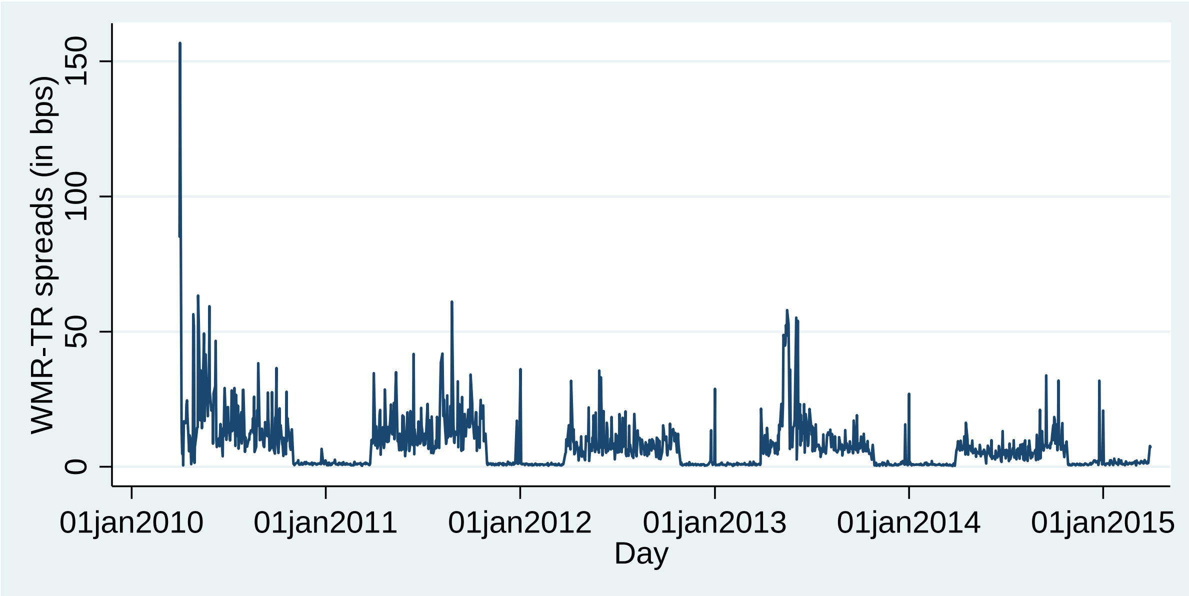
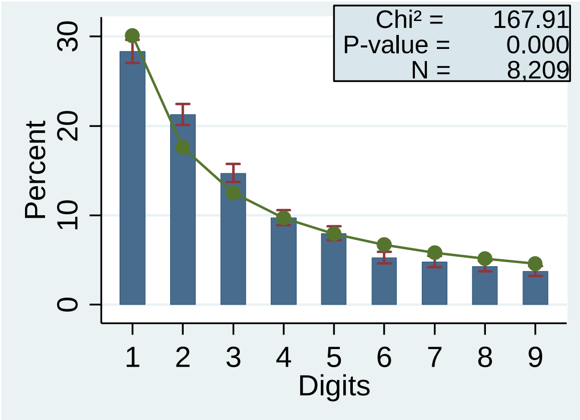


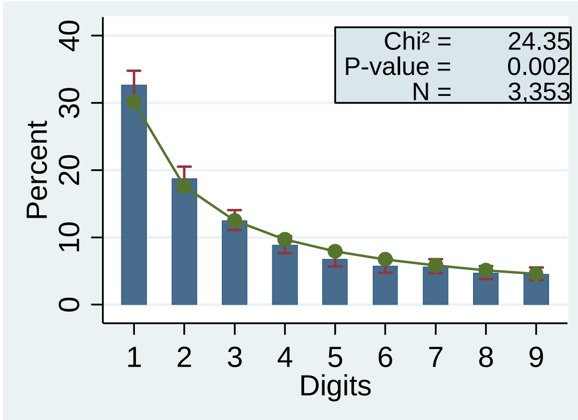
Figure 5. WMR-TR spreads in bps. The line illustrates the intensity that the WMR benchmarks differ from the TR rates based on equation (3). It represents the average across the nine currency pairs on a daily basis. The collusive period is from 1 April 2010 to 15 October 2013, while the non-collusive period is from 16 October 2013 to 31 March 2015.

Panel A. WMR returns

Collusive period

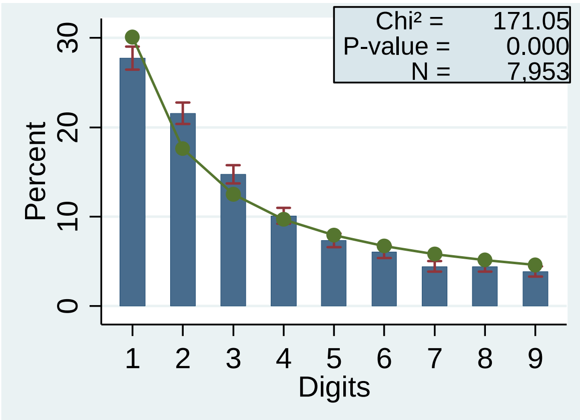


Non-collusive period



Panel B. TR returns

Collusive period



Non-collusive period

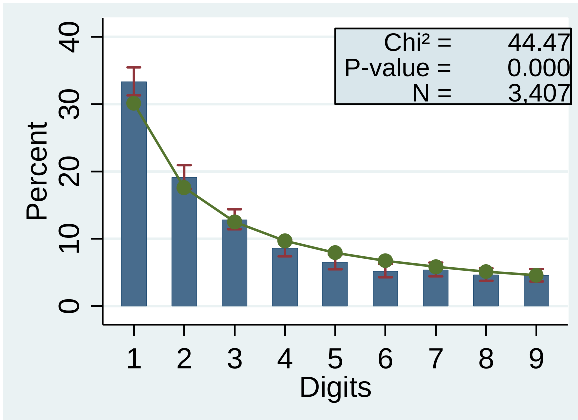


Figure 6. Distribution of first digits of the WMR and TR data set. The graphs depict the first digit proportions of daily FX returns of the nine currency pairs. Panel A includes the first digits of all nine WMR currency pairs and Panel B covers the first digits of all nine TR currency pairs. The collusive period encompasses the period from 1 April 2010 to 15 October 2013 (left). The non-collusive period encompasses the period from 16 October 2013 to 31 March 2015 (right). Bars represent the first digit proportions of FX returns, while the solid line indicates the first digit proportions expected by Benford's law (based on equation (5)). The capped spikes for each first digit represent the upper and lower bounds of the 99% confidence interval. The Chi² statistic is used to test the overall fit of the first digit distribution and the Z-statistic to test significant differences in individual digits. The null hypothesis is that the first digit distribution of FX returns equals the expected Benford distribution.

Tables

Table 1. Descriptive statistics for daily FX data.

Currency pair	WMR						TR					
	N	Mean	Std.dev.	Min	Median	Max	N	Mean	Std.dev.	Min	Median	Max
AUD/USD	1,304	0.97	0.08	0.76	0.98	1.10	1,282	0.97	0.08	0.76	0.98	1.10
USD/CAD	1,304	1.04	0.06	0.94	1.02	1.28	1,282	1.04	0.06	0.94	1.03	1.28
USD/CHF	1,304	0.94	0.06	0.73	0.93	1.16	1,304	0.94	0.06	0.73	0.93	1.16
EUR/USD	1,304	1.32	0.07	1.05	1.32	1.49	1,146	1.32	0.07	1.05	1.33	1.49
GBP/USD	1,304	1.59	0.05	1.43	1.59	1.72	1,282	1.59	0.05	1.43	1.59	1.72
USD/JPY	1,304	91.55	12.65	75.76	87.79	121.50	1,282	91.52	12.75	75.72	87.20	121.51
USD/NOK	1,304	6.03	0.52	5.22	5.93	8.36	1,282	6.03	0.53	5.23	5.93	8.36
NZD/USD	1,304	0.80	0.04	0.66	0.81	0.88	1,282	0.80	0.04	0.66	0.81	0.88
USD/SEK	1,304	6.83	0.50	5.99	6.68	8.75	1,282	6.82	0.50	6.00	6.67	8.75

Notes: This table reports descriptive statistics for daily currency pairs from 1 April 2010 to 31 March 2015 based on WMR benchmarks and TR rates at 4 p.m.

Table 2. Trading and market behavior according to *MM* (1).

	Collusive period (I)			Non-collusive period (II)			Difference (I-II)	
	N	Mean	Std.dev.	N	Mean	Std.dev.	Mean	P-value
ALL	9	8.82	2.05	9	3.80	0.83	5.02***	0.000
FX rate								
AUD/USD	902	8.77	11.67	380	4.04	6.61	4.73***	0.000
USD/CAD	902	8.14	10.92	380	3.94	5.68	4.20***	0.000
USD/CHF	924	12.19	35.47	380	3.52	5.61	8.67***	0.000
EUR/USD	766	6.32	8.80	380	2.81	4.85	3.51***	0.000
GBP/USD	902	6.65	8.78	380	2.83	4.93	3.81***	0.000
USD/JPY	902	6.61	9.87	380	2.95	4.72	3.67***	0.000
USD/NOK	902	10.17	12.82	380	5.14	7.73	5.03***	0.000
NZD/USD	902	10.16	12.82	380	4.50	7.25	5.66***	0.000
USD/SEK	902	10.36	13.45	380	4.45	5.58	5.91***	0.000

Notes: This table reports the mean comparison test results for the first manipulation measure in bps. *MM* (1) refers to abnormal FX rates using TR rates as closing prices and WMR benchmarks as average prices in equation (1). A higher level of *MM* (1) would indicate a higher intensity of irregularities. The collusive period encompasses the period from 1 April 2010 to 15 October 2013. The non-collusive period encompasses the period from 16 October 2013 to 31 March 2015. A statistically significant difference in the values of *MM* (1) between the two periods would indicate potential manipulation using a one-tailed *t*-test. The null hypothesis is that abnormal FX rates are not higher during the collusive period than during the non-collusive period.
*** Denote significance at the 1% level.

Table 3. Trading and market behavior according to *MM* (2).

	1. digit	Chi²	2. digit	Chi²	3. digit	Chi²	4. digit	Chi²	5. digit	Chi²
N^{diff}	36	17.29***	437	24.07***	3,180	120.00***	7,449	156.23***	8,239	23.15***
N	11,424		11,424		11,424		11,424		11,424	
MM (2) (ALL in %)	0.32		3.83		27.84		65.20		72.12	
	1. digit *MM* (2)		2. digit *MM* (2)		3. digit *MM* (2)		4. digit *MM* (2)		5. digit *MM* (2)	
	Diff (in %)	z	*Diff* (in %)	z	*Diff* (in %)	z	*Diff* (in %)	z	*Diff* (in %)	z
AUD/USD	0.44*	1.30	1.10	0.95	-0.97	-0.35	-14.07	-4.77	12.6***	4.23
USD/CAD	0.11	0.65	2.47***	2.71	17.19***	7.61	30.77***	10.11	-12.94	-4.52
USD/CHF	-0.20	-0.54	5.41***	3.84	15.15***	5.19	5.16**	2.10	4.08*	1.34
EUR/USD	0.00	NA	1.04**	2.00	4.26***	2.45	18.71***	6.06	7.52***	2.90
GBP/USD	0.00	NA	0.40	0.89	7.86***	4.63	25.56***	8.37	5.35**	2.20
USD/JPY	0.55*	1.45	3.95***	3.76	36.28***	13.19	49.75***	16.82	15.81***	6.61
USD/NOK	-1.92	-3.14	0.32	0.24	16.31***	5.47	2.87	1.11	3.98**	2.19
NZD/USD	0.55*	1.45	5.88***	3.97	20.48***	6.79	5.9***	2.45	3.93*	1.29
USD/SEK	0.33	1.13	5.05***	3.30	19.92***	6.59	2.77	1.09	0.13	0.07

Notes: This table reports the proportion test results for the second manipulation measure in percent. *MM* (2) refers to the proportion of digit differences between WMR benchmarks and TR rates in equation (2). A higher proportion would indicate a higher intensity of irregularities. Several digit positions are separately investigated. The collusive period encompasses the period from 1 April 2010 to 15 October 2013. The non-collusive period encompasses the period from 16 October 2013 to 31 March 2015. A statistically significant difference between the two periods would indicate potential manipulation using two test statistics. First, the distribution of digit differences across the nine currency pairs is compared between the two periods based on Pearson's Chi² test statistic for independence in equation (4), here with the number of different currency pairs d, the compared periods, s, and the joint number of digit differences during both periods N^{diff}. The null hypothesis is that the proportions of digit differences during the collusive period match the proportions of digit differences during the non-collusive period. Second, the values of *MM* (2) for single currency pairs are compared between the two periods based on a one-tailed Z-test. *Diff* represents the difference in *MM* (2) between the collusive and non-collusive period. The null hypothesis is that the proportion of digit differences is not higher during the collusive period than during the non-collusive period. "NA" refers to the case when no digit difference occurs.

*** Denote significance at the 1% level. ** Denote significance at the 5% level. * Denote significance at the 10% level.

Table 4. WMR-TR spread.

	Collusive period (I)			Non-collusive period (II)			Difference (I-II)	
	N	Mean	Std.dev.	N	Mean	Std.dev.	Mean	P-value
ALL	9	8.83	2.06	9	3.80	0.83	5.03***	0.000
FX rate								
AUD/USD	902	8.77	11.67	380	4.04	6.61	4.73***	0.000
USD/CAD	902	8.14	10.93	380	3.94	5.68	4.20***	0.000
USD/CHF	924	12.25	36.08	380	3.52	5.61	8.73***	0.000
EUR/USD	766	6.32	8.80	380	2.81	4.85	3.51***	0.000
GBP/USD	902	6.65	8.78	380	2.83	4.94	3.81***	0.000
USD/JPY	902	6.61	9.88	380	2.95	4.72	3.67***	0.000
USD/NOK	902	10.17	12.83	380	5.14	7.75	5.03***	0.000
NZD/USD	902	10.16	12.82	380	4.50	7.24	5.66***	0.000
USD/SEK	902	10.36	13.46	380	4.45	5.58	5.91***	0.000

Notes: This table reports the mean comparison test results for the WMR-TR spreads in bps. The proportional quoted spread between WMR benchmarks and TR rates is based on equation (3). A higher level of spreads would indicate a higher intensity of irregularities. The collusive period encompasses the period from 1 April 2010 to 15 October 2013. The non-collusive period encompasses the period from 16 October 2013 to 31 March 2015. A statistically significant difference between the two periods would indicate potential manipulation using a one-tailed *t*-test. The null hypothesis is that WMR-TR spreads are not higher during the collusive period than during the non-collusive period.
*** Denote significance at the 1% level.

Table 5. Digit distribution.

	Collusive period			Non-collusive period		
	N	Chi²	P-value	N	Chi²	P-value
1. digit						
AUD/USD	1,826	0.46	0.794	760	0.00	1.000
USD/CAD	1,826	0.04	0.842	760	NA	NA
USD/CHF	1,848	0.00	1.000	760	0.02	0.989
EUR/USD	1,690	NA	NA	760	NA	NA
GBP/USD	1,826	NA	NA	760	NA	NA
USD/JPY	1,826	0.98	0.807	760	0.00	1.000
USD/NOK	1,826	0.08	0.783	760	0.01	1.000
NZD/USD	1,826	0.47	0.792	760	0.00	1.000
USD/SEK	1,826	1.51	0.681	760	0.00	1.000
2. digit						
AUD/USD	1,826	4.05	0.908	760	0.33	1.000
USD/CAD	1,826	0.55	0.997	760	0.00	1.000
USD/CHF	1,848	1.38	0.998	760	0.20	1.000
EUR/USD	1,690	6.51*	0.089	760	0.00	1.000
GBP/USD	1,826	0.08	0.959	760	0.02	0.999
USD/JPY	1,826	2.40	0.983	760	0.03	1.000
USD/NOK	1,826	0.85	1.000	760	0.43	1.000
NZD/USD	1,826	2.87	0.969	760	0.29	1.000
USD/SEK	1,826	2.36	0.984	760	0.09	1.000
3. digit						
AUD/USD	1,826	1.62	0.996	760	1.23	0.999
USD/CAD	1,826	1.93	0.992	760	0.46	1.000
USD/CHF	1,848	2.67	0.976	760	2.46	0.982
EUR/USD	1,690	4.79	0.852	760	0.49	1.000
GBP/USD	1,826	2.12	0.989	760	0.89	1.000
USD/JPY	1,826	1.57	0.997	760	0.34	1.000
USD/NOK	1,826	3.45	0.944	760	5.33	0.804
NZD/USD	1,826	3.45	0.944	760	3.98	0.913
USD/SEK	1,826	6.26	0.714	760	1.57	0.997

(Continued)

Table 5. Continued.

	Collusive period			Non-collusive period		
	N	Chi²	P-value	N	Chi²	P-value
4. digit						
AUD/USD	1,826	12.17	0.204	760	3.32	0.950
USD/CAD	1,826	8.24	0.510	760	3.44	0.944
USD/CHF	1,848	3.92	0.917	760	13.17	0.155
EUR/USD	1,690	6.44	0.695	760	4.90	0.843
GBP/USD	1,826	10.31	0.326	760	1.50	0.997
USD/JPY	1,826	9.35	0.406	760	4.83	0.849
USD/NOK	1,826	5.55	0.784	760	2.36	0.984
NZD/USD	1,826	4.15	0.901	760	4.01	0.911
USD/SEK	1,826	6.86	0.651	760	19.86**	0.019
5. digit						
AUD/USD	1,826	40.04***	0.000	760	79.31***	0.000
USD/CAD	1,826	18.68**	0.028	760	16.76*	0.053
USD/CHF	1,848	327.20***	0.000	760	4.40	0.883
EUR/USD	1,690	6.11	0.729	760	1.35	0.998
GBP/USD	1,826	9.78	0.369	760	6.69	0.670
USD/JPY	1,826	919.41***	0.000	760	11.39	0.250
USD/NOK	1,826	9.00	0.437	760	11.60	0.237
NZD/USD	1,826	9.07***	0.003	760	80.54***	0.000
USD/SEK	1,826	8.88	0.448	760	2.68	0.976

Notes: This table reports the digit distribution test results for WMR benchmarks and TR rates. The digit distribution between WMR benchmarks and TR rates is compared based on Pearson's Chi² test statistic for independence in equation (4). *N* represents the joint number of observations. Several digit positions during the collusive and non-collusive period are separately investigated. The null hypothesis is that the WMR proportions match the TR proportions. "NA" refers to the case when only one bin is available, for instance, all FX rates display the digit one at the first digit position in both data sets. There is a variation in the numbers of bins and degrees of freedom.
*** Denote significance at the 1% level.
** Denote significance at the 5% level.
* Denote significance at the 10% level.

Table 6. Benford test.

Currency	WMR returns			TR returns		
	N	Chi²	P-value	N	Chi²	P-value
Collusive period						
AUD/USD	911	31.33***	0.000	896	42.16***	0.000
USD/CAD	908	11.58	0.171	894	23.38***	0.003
USD/CHF	913	26.27***	0.001	906	30.16***	0.000
EUR/USD	913	15.94**	0.043	762	14.89*	0.061
GBP/USD	912	36.40***	0.000	899	30.68***	0.000
USD/JPY	908	19.75**	0.011	897	11.92	0.155
USD/NOK	915	29.54***	0.000	901	27.29***	0.001
NZD/USD	914	51.00***	0.000	897	38.92***	0.000
USD/SEK	915	26.03***	0.001	901	35.70***	0.000
Non-collusive period						
AUD/USD	371	29.78***	0.000	379	25.80***	0.001
USD/CAD	373	19.47**	0.013	376	24.90***	0.002
USD/CHF	370	11.75	0.163	378	13.15	0.107
EUR/USD	370	4.01	0.856	378	6.02	0.645
GBP/USD	374	10.91	0.207	380	17.64**	0.024
USD/JPY	374	6.27	0.618	379	7.18	0.517
USD/NOK	375	9.90	0.272	380	8.40	0.396
NZD/USD	371	2.65	0.954	377	10.20	0.251
USD/SEK	375	16.96**	0.031	380	16.56**	0.035

Notes: This table reports the digit distribution test results for WMR and TR FX returns. The first digit distribution of WMR returns and TR returns, respectively, is compared with the Benford distribution based on Pearson's Chi² test statistic for goodness of fit in equation (6). The expected proportions are based on Benford's law based on equation (5). The null hypothesis is that the first digit distribution of FX returns equals the expected Benford distribution. The critical value at a significance level for Chi² of 10% is 13.36, of 5% is 15.51, and of 1% is 20.09 (8 degrees of freedom).

*** Denote significance at the 1% level.

** Denote significance at the 5% level.

* Denote significance at the 10% level.

Table 7. Liquidity measures.

	Collusive period (I)			Non-collusive period (II)			Difference (I-II)	
	N	Mean	Std.dev	N	Mean	Std.dev	Mean	P-value
Return								
WMR	9	0.06	0.41	9	0.50	2.39	-0.45	0.707
TR	9	0.00	0.41	9	0.50	2.38	-0.50	0.727
Volatility $L^{(pd)}$								
WMR	9	22.65	3.57	9	16.05	2.57	6.60***	0.000
TR	9	22.24	3.90	9	16.21	2.58	6.03***	0.001
Bid-ask spread $L^{(ba)}$								
WMR	9	5.09	1.68	9	4.34	1.78	0.75	0.186
TR	9	4.65	2.23	9	3.27	1.45	1.38*	0.069
Return reversal $L^{(rr)}$								
WMR	9	-0.05	0.75	9	-0.56	1.93	0.51	0.236
TR	9	-0.18	0.63	9	-0.59	1.90	0.41	0.272

Notes: This table reports the mean comparison test results for the liquidity measures in bps. Return refers to the change in the log FX rates between two trading days in equation (7). The first liquidity measure, price dispersion $L^{(pd)}$, is a volatility proxy defined as the absolute daily log return given in equation (8). The second liquidity measure, trading cost $L^{(ba)}$, is the proportional quoted bid-ask spread based on equation (9). The third liquidity measure, return reversal $L^{(rr)}$, refers to equation (10). The collusive period encompasses the period from 1 April 2010 to 15 October 2013. The non-collusive period encompasses the period from 16 October 2013 to 31 March 2015. Using a one-tailed *t*-test, the null hypothesis is that the respective liquidity measure is not higher during the collusive period than during the non-collusive period.

*** Denote significance at the 1% level.

* Denote significance at the 10% level.

Table 8. *MM* (1) normalized by volatility.

Currency	Collusive period (I)		Non-collusive period (II)		Difference (I-II)	
	Mean	Std.dev	Mean	Std.dev	Mean	P-value
AUD/USD	1.08	5.78	0.90	6.64	0.18	0.311
USD/CAD	1.49	5.24	1.18	4.14	0.31	0.154
USD/CHF	1.06	4.69	0.63	1.36	0.43**	0.040
EUR/USD	1.26	7.45	1.13	7.37	0.13	0.388
GBP/USD	1.29	5.43	0.63	1.64	0.66**	0.011
USD/JPY	1.07	3.50	0.59	2.05	0.48***	0.006
USD/NOK	0.98	2.62	1.40	5.67	-0.42	0.964
NZD/USD	1.09	3.43	0.83	4.18	0.26	0.124
USD/SEK	2.22	22.36	2.79	23.32	-0.58	0.661

Notes: This table reports the mean comparison test results for the first manipulation measure normalized by volatility. *MM* (1) is normalized in dividing abnormal FX rates in equation (1) by the volatility proxy in equation (8). The collusive period encompasses the period from 1 April 2010 to 15 October 2013. The non-collusive period encompasses the period from 16 October 2013 to 31 March 2015. A higher value would indicate a higher intensity of potential manipulation using a one-tailed *t*-test. The null hypothesis is that normalized abnormal FX rates are not higher during the collusive period than during the non-collusive period.

*** Denote significance at the 1% level.

** Denote significance at the 5% level.

References

Abrantes-Metz, Rosa M., Michael Kraten, Albert D. Metz, and Gim S. Seow. 2012. "Libor manipulation?" *Journal of Banking and Finance* 36 (1):136-150.

Abrantes-Metz, Rosa M., Sofia B. Villas-Boas, and George Judge. 2011. "Tracking the Libor rate." *Applied Economics Letters* 18 (10):893-899.

Allen, Franklin, and Gary Gorton. 1992. "Stock price manipulation, market microstructure and asymmetric information." *European Economic Review* 36 (2-3):624-630.

Ashton, John K., and Robert S. Hudson. 2008. "Interest rate clustering in UK financial services markets." *Journal of Banking and Finance* 32 (7):1393-1403.

Bacchetta, Philippe, and Eric Van Wincoop. 2006. "Can information heterogeneity explain the exchange rate determination puzzle?" *American Economic Review* 96 (3):552-576.

Benford, Frank. 1938. "The law of anomalous numbers." *Proceedings of the American Philosophical Society* 78 (4):551-572.

Berger, David W., Alain P. Chaboud, Sergey V. Chernenko, Edward Howorka, and Jonathan H. Wright. 2008. "Order flow and exchange rate dynamics in electronic brokerage system data." *Journal of International Economics* 75 (1):93-109.

BIS. 2013. Foreign exchange turnover in April 2013. Triennial central bank survey. Bank for International Settlements.

Bloomfield, Robert, and Maureen O'Hara. 2000. "Can transparent markets survive?" *Journal of Financial Economics* 55 (3):425-459.

Boyle, Jeff. 1994. "An application of Fourier Series to the most significant digit problem." *The American Mathematical Monthly* 101 (9):879-886.

Carrera, César. 2015. "Tracking exchange rate management in Latin America." *Review of Financial Economics* 25:35-41.

Carslaw, Charles A.P.N. 1988. "Anomalies in income numbers: Evidence of goal oriented behavior." *The Accounting Review* 63 (2):321-327.

Cason, Timothy N. 2000. "The opportunity for conspiracy in asset markets organized with dealer intermediaries." *Review of Financial Studies* 13 (2):385-416.

CFTC. 2014. CFTC orders five banks to pay over $1.4 billion in penalties for attempted manipulation of foreign exchange benchmark rates. U.S. Commodity Futures Trading Commission.

Cleary, Richard, and Jay C. Thibodeau. 2005. "Applying digital analysis using Benford's law to detect fraud: the dangers of type I errors." *Auditing: A Journal of Practice & Theory* 24 (1):77-81.

Comerton-Forde, Carole, and Tālis J. Putniņš. 2011. "Measuring closing price manipulation." *Journal of Financial Intermediation* 20 (2):135-158.

De Ceuster, Marc J.K., Geert Dhaene, and Tom Schatteman. 1998. "On the hypothesis of psychological barriers in stock markets and Benford's law." *Journal of Empirical Finance* 5 (3):263-279.

DoJ. 2015. Attorney General Lynch delivers remarks at a press conference on foreign exchange spot market manipulation. The United States Department of Justice.

Doucouliagos, Chris. 2005. "Price exhaustion and number preference: time and price confluence in Australian stock prices." *The European Journal of Finance* 11 (3):207-221.

Euromoney. 2013. Euromoney foreign exchange survey 2013. Euromoney Institutional Investor.

Evans, Martin D. D., and Richard K. Lyons. 2002. "Order flow and exchange rate dynamics." *Journal of Political Economy* 110 (1):170-180.

FCA. 2014. FCA fines five banks £1.1 billion for FX failings and announces industry-wide remediation programme. Financial Conduct Authority.

Fouquau, Julien, and Philippe K. Spieser. 2015. "Statistical evidence about LIBOR manipulation: A "Sherlock Holmes" investigation." *Journal of Banking and Finance* 50:632-643.

Froot, Kenneth A., and Tarun Ramadorai. 2005. "Currency returns, intrinsic value, and institutional-investor flows." *Journal of Finance* 60 (3):1535-1566.

Gilbert, Mark. 2007. "Barclays takes a money-market beating." *Bloomberg*, September 3.

Guan, Liming, Daoping He, and David Yang. 2006. "Auditing, integral approach to quarterly reporting, and cosmetic earnings management." *Managerial Auditing Journal* 21 (6):569-581.

Hamming, Richard W. 1970. "On the distribution of numbers." *Bell System Technical Journal* 49 (8):1609-1625.

Harrington, Joseph E. 2008. "Detecting cartels." In *Handbook of antitrust economics*, edited by Paolo Buccirossi, 213–258. Cambridge: MIT Press.

Herrmann, Don, and Wayne B. Thomas. 2005. "Rounding of analyst forecasts." *The Accounting Review* 80 (3):805-823.

Hill, Theodore P. 1995. "A statistical derivation of the significant-digit law." *Statistical Science* 10 (4):354-363.

Hillion, Pierre, and Matti Suominen. 2004. "The manipulation of closing prices." *Journal of Financial Markets* 7 (4):351-375.

HMTreasury. 2012. The Wheatley Review of LIBOR: Final report. Crown.

Kahn, Charles, George Pennacchi, and Ben Sopranzetti. 1999. "Bank deposit rate clustering: Theory and empirical evidence." *Journal of Finance* 54 (6):2185-2214.

Kyle, Albert S. 1985. "Continuous auctions and insider trading." *Econometrica* 53 (6):1315-1335.

Kyle, Albert S., and S. Viswanathan. 2008. "How to define illegal price manipulation." *American Economic Review* 98 (2):274-279.

Ley, Eduardo. 1996. "On the peculiar distribution of the US stock indexes' digits." *The American Statistician* 50 (4):311-313.

Lyons, Richard K. 1997. "A simultaneous trade model of the foreign exchange hot potato." *Journal of International Economics* 42 (3-4):275-298.

Lyons, Richard K. 2001. *The microstructure approach to exchange rates*. Cambridge, Mass.: MIT Press.

Mancini, Loriano, Angelo Ranaldo, and Jan Wrampelmeyer. 2013. "Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums." *Journal of Finance* 68 (5):1805-1841.

Marsh, Ian W., and Ceire O'Rourke. 2005. Customer order flow and exchange rate movements: Is there really information content?

Melvin, Michael, and John Prins. 2011. The equity hedging channel of exchange rate adjustment.

Menkhoff, Lukas, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf. 2012. "Carry trades and global foreign exchange volatility." *Journal of Finance* 67 (2):681-718.

Mitchell, Jason. 2001. "Clustering and psychological barriers: The importance of numbers." *Journal of Futures Markets* 21 (5):395-428.

Mitchell, Jason, and H.Y. Izan. 2006. "Clustering and psychological barriers in exchange rates." *Journal of International Financial Markets, Institutions and Money* 16 (4):318-344.

Mollenkamp, Carrick, and Mark Whitehouse. 2008. "Study casts doubt on key rate." *Wall Street Journal*, May 29.

Newcomb, Simon. 1881. "Note on the frequency of use of the different digits in natural numbers." *American Journal of Mathematics* 4 (1):39-40.

Nigrini, Mark J. 1996. "A taxpayer compliance application of Benford's law." *The Journal of the American Taxation Association* 18 (1):72-91.

Nigrini, Mark J., and Linda J. Mittermaier. 1997. "The use of Benford's law as an aid in analytical procedures." *Auditing: A Journal of Practice & Theory* 16 (2):52-67.

Osler, Carol L. 2003. "Currency orders and exchange rate dynamics: An explanation for the predictive success of technical analysis." *Journal of Finance* 58 (5):1791-1819.

Osler, Carol L. 2005. "Stop-loss orders and price cascades in currency markets." *Journal of International Money and Finance* 24 (2):219-241.

Osler, Carol L. 2012. "Market microstructure and the profitability of currency trading." *Annual Review of Financial Economics* 4 (1):469-495.

Pinkham, Roger S. 1961. "On the distribution of first significant digits." *The Annals of Mathematical Statistics* 32 (4):1223-1230.

Raimi, Ralph A. 1976. "The first digit problem." *The American Mathematical Monthly* 83 (7):521-538.

Rao, Jon N.K., and Alastair J. Scott. 1981. "The analysis of categorical data from complex sample surveys: Chi-squared tests for goodness of fit and independence in two-way tables." *Journal of the American Statistical Association* 76 (374):221-230.

Spatt, Chester. 2014. "Security market manipulation." *Annual Review of Financial Economics* 6 (1):405-418.

Thomas, Jacob K. 1989. "Unusual patterns in reported earnings." *The Accounting Review* 64 (4):773-787.

Vaughan, Liam, and Gavin Finch. 2013. "Currency spikes at 4 p.m. in London provide rigging clues." *Bloomberg*, August 28.

Vaughan, Liam, Gavin Finch, and Ambereen Choudhury. 2013. "Traders said to rig currency rates to profit off clients." *Bloomberg*, June 12.

Vitale, Paolo. 2000. "Speculative noise trading and manipulation in the foreign exchange market." *Journal of International Money and Finance* 19 (5):689-712.

WMCompany. 2011. WM/Reuters spot and forward rates methodology guide. The World Markets Company.

Zhou, Bin. 1996. "High-frequency data and volatility in foreign-exchange rates." *Journal of Business and Economic Statistics* 14 (1):45-52.