Customer Trading in the Foreign Exchange Market
Empirical Evidence from an Internet Trading Platform

Sandra Lechner∗  Ingmar Nolte†
University of Konstanz, University of Konstanz,
CoFE

∗Department of Economics, Box D124, University of Konstanz, 78457 Konstanz, Germany.
Phone +49-7531-88-3214, Fax -4450, email: Sandra.Lechner@uni-konstanz.de
†Department of Economics, Box D124, University of Konstanz, 78457 Konstanz, Germany.
Phone +49-7531-88-3753, Fax -4450, email: Ingmar.Nolte@uni-konstanz.de. The work is supported in part by the European Community’s Human Potential Program under contract HPRN-CT-2002-00232, Microstructure of Financial Markets in Europe; and by the Fritz Thyssen Foundation through the project ‘Dealer-Behavior and Price-Dynamics on the Foreign Exchange Market’. The authors want to thank Richard Olsen for helpful comments and for providing us with the data.
Abstract

This paper focuses on the relationship between customer trading’s order flow and price changes. Customers can either trade with interdealer banks or on internet trading platforms that have rapidly been emerging during the last decade. Our analysis is based on a customer dataset from Olsen and Associates’ FXTrade, which is one of the most well known internet currency exchanges. This dataset contains information on several order types. Furthermore it allows us to infer whether the order is submitted to close a certain previously established position or whether the order is submitted to open a new position. This information enables us to define a refined aggregated net order flow measure, which we compare to the standard net order flow measure (difference between the number of buyer and seller initiated trades). For the EUR/USD exchange rate we analyze the dynamic behavior and the interaction with the price change series for both order flow measures. The analysis is carried out on 11 different frequencies (between 1 min to 1 day). We find that the explanatory power of the refined order flow measure in comparison to the standard one is significantly higher. Furthermore, we find that in upwards (downwards) trending periods the order flow is subject to seasonality effects, since in these periods the order flow measure is positively (negatively) skewed. We specify structural VAR models to capture the bivariate dynamic structure of order flow and price changes.

JEL classification: G10, F31, C32

Keywords: Customer Dataset, Order Flow, Price Changes, Foreign Exchange Market, Aggregation Levels
1 Introduction

This paper focuses on the key determinant to explain exchange rate movements: order flow. We use information stemming from a customer dataset to analyze the interaction of order flow and price changes on different aggregation levels. Our customer dataset contains detailed information on traders’ characteristics and currency positions. Relating order flow to traders’ characteristics provides valuable insights into the dynamics of order flow and price changes. The way how information is aggregated by order flow is central in understanding the microstructure of the foreign exchange market. Indeed, the foreign exchange market is a highly decentralized market with low transparency. Information on the interpretation of specific news events, risk preferences, hedging demands, or central bank interventions is therefore widely dispersed and disaggregated among agents. The literature on the measurement of order flow in the foreign exchange market can be traced back to Lyons (1995), who suggests to aggregate the dispersed information into one single measure: the difference between the number of buyer- and seller initiated trades, throughout denoted as standard net order flow. In the portfolio allocation model of Evans & Lyons (2002a,b) order flow is the cornerstone to explain price movements. They propose a simple model, which outperforms the random walk forecast using additional information on order flow and refute the Meese & Rogoff (1983a,b) findings.

To our knowledge most of the existing studies on order flow, except the ones of Osler (2002) and Marsh & O’Rourke (2004), consider order flow obtained from a dataset on direct (e.g. Reuters Dealing 2000-1) or brokeraged (e.g. Reuters Dealing 2000-2, EBS) interdealer trading. Osler (2002) and Marsh & O’Rourke (2004) use a dataset on customer trades collected by the Royal Bank of Scotland. For a dealer-bank itself customer-trading-order-flow is the primary source of dispersed information, whereas interdealer-bank-order-flow is in part only a consequence of customer-order-flow. Customer trading amounts to 51% of the FX spot market trading and can be divided into trading with financial institutions without access to the interdealer market and non-financial customers. Customers have in general two ways of participating in the FX market: either by trading with interdealer banks, or by trading on internet trading platforms which have been rapidly emerging during the last decade. These platforms try to attract customers offering tight spreads
(equal or close to the spread in the interdealer market) regardless of the trade size.

In our analysis we use a dataset (1st October 2003 to 14th May 2004) on customer trades which stems from Olsen and Associates’ FXTrade, one of the most prominent internet currency trading platforms. One of the features of this dataset is that the trading mechanism on this platform allows us to distinguish between a broad set of order types (market orders and limit orders to open a new position, market orders and limit orders to close a previous established position, stop-loss, take-profit orders, and various further types). Given the information on these particular transaction types, we are able to construct a refined aggregated net order flow measure.

For the EUR/USD rate we compare the dynamic behavior, as well as the interaction with the corresponding price change series of our refined order flow measure with the behavior of the standard order flow measure, which is constructed by taking the difference between the number of buyer and seller initiated trades. This analysis is carried out on 11 different frequencies from 1 min to 1 day, and we find that the explanatory power of the refined order flow measure in comparison to the standard order flow measure is significantly higher. Furthermore, we find that in upwards (downwards) trending exchange rate periods order flow is subject to seasonality effects, since in these periods the order flow measure is positively (negatively) skewed.

We try to capture the bivariate dynamic structure of order flow and price changes by applying structural VAR models and we show that they are flexible enough to describe the dynamics on 10 min to 1 day aggregation levels. However, for higher frequencies these models seem to be too inflexible to describe the true nature of the process.

The paper is organized as follows. In Section 2, we briefly describe the foreign exchange market and explain in detail the trading mechanism and the different order types on the OANDA FXTrade platform. This is essential to understand the construction of our refined order flow. In section 3, we describe the dataset and explain how we construct our refined net order flow. Section 4 presents our empirical results. Section 5 concludes and gives an outlook on possible extensions.
2 The Foreign Exchange Market and OANDA

The Foreign Exchange Market

The FX-market is generally characterized by a high degree of decentralization, low-transparency, and 24h trading. According to the Triennial Bank for International Settlements’ (BIS) Report (2004), the nine most active trading centers in 2004 in the FX spot market are London (31.3%), New York (19.2%), Tokyo (8.3%), Singapore (5.2%), Frankfurt (4.9%), Hong Kong (4.2%), Sydney (3.4%), Zürich (3.3%) and Paris (2.7%) accounting for a total turnover of 82.5%. The three most actively traded currency pairs are USD/EUR (28%), USD/JPY (17%), and USD/GBP (14%). In 2004 the total average daily turnover amounts to 1773 bn$, which is proportioned into spot (35%), forward (12%), and swap (53%) market. In the FX spot market 48% of the turnover is generated by dealers (large commercial banks, investment banks and security houses) which trade very actively and participate in the interdealer market or trade with large corporate firms, 34% is generated by other financial institutions (smaller commercial banks, investment banks, security houses, mutual funds, pension funds, hedge funds, insurance companies, etc.), and the remaining 17% are generated by non-financial customers such as corporates and governments. Since our data-set stems from OANDA FXTrade (a platform for mainly non-financial customer trading), we will briefly describe the structure of the FX spot market, the classification of OANDA FXTrade within the market, and the trading rules on OANDA FXTrade.¹

There are several groups of agents who trade in the FX market. First of all there is the group of non-financial customers (17% of the FX spot market trading), which mainly consists of corporations (exporters, importers), and also, since the establishment of internet trading platforms, of individual investors. One important characteristic of this group is they do not have access to the interbank (interdealer) market and they have, if at all, only very limited information on each other. The second group of agents in the FX market consists of financial institutions (34% of the FX spot market trading), such as smaller commercial banks, mutual funds, hedge funds, insurance companies, etc. which similar to the first group have no access to the interbank market. Both groups have basically two channels to settle a transaction: i) directly via dealer-banks that offer a bid-ask spread, which is mainly driven by order handling costs (the smaller and the more unconventional the order (size), the higher the bid-ask spread) and ii) they can trade with each other via internet trading ¹For a more detailed description of the FX market we refer to Lyons (2001) and Rime (2003).
platforms, which try to offer independently of the trade size a small (close or equal to interbank) bid-ask spread to attract customers. The market of internet trading platforms itself is divided into two groups: a) platforms which are established by banks or a consortium of banks, such as FXConnect or Currenex and b) non-bank trading platforms such as Deal4Free or OANDA FXTrade, which is the source of our data-set. Usually these internet trading platforms are at least partially organized as so called crossing networks, since there is too little trading to have an (arbitrage free) price discovery. Crossing networks obtain bid and ask quotes completely or in addition to their own limit order book from other trading channels, e.g. electronic brokers like Reuters Dealing 3000-2 or EBS. By channeling a certain amount of (not necessarily local) customers to one trading platform, the decentralization of the customers market has been reduced by the establishment of internet trading platforms. Furthermore, depending on the platform, customers may have (limited) access to the limit order book and the history of trades and quotes. Therefore, the transparency is higher than in the direct customer-to-dealer-bank trading.

The third group of agents consists of dealer-banks (48% of the FX spot market trading), which trade in general as the counterparty with members of the first and second group or with each other in the interdealer market. Trading in the interdealer market is usually done in two different ways: either directly (by telephone and via Reuters Dealing 3000-1), or indirectly via brokers (voice brokers and electronic brokers). Telephone trading is the traditional form of direct (bilateral) trading, thereby dealer 1 (aggressor, taker, initiator) calls dealer 2 (non-aggressor, market maker, non-initiator) and requests quotes for a standardized trade size. After the announcement of the quotes, dealer 1 can buy, sell or decide not to trade given the quotes. After the trading decision the identity of dealer 1 is revealed to dealer 2. Since the introduction of Reuters Dealing 3000-1 (1st version in 1981) this type of bilateral communication is mainly done on a computer screen with a kind of messenger program, which is considered to be quicker and perhaps more efficient. Indirect trading in the interdealer market is done via brokers. In general, a dealer can act as a non-aggressor who submits a limit order to a broker, or he can act as an aggressor who submits a market order to a broker. We distinguish between voice brokers and electronic brokers: voice brokers are usually local (London based, New York based, etc.) companies who establish a (closed) telephone network and intercoms in the interested dealer-banks. They collect limit orders and announce the best bid and ask quotes (for standardized trade sizes) via the intercoms. The dealer can react by picking up the phone and responding whether he wants to sell or to buy. Electronic
brokers (basically Reuters Dealing 3000-2 and EBS) act world wide. The dealers are connected via trading stations; they can submit limit orders and market orders in electronic form. In contrast to trading via voice brokers, they get information on the recent trading-history. Therefore, trading via electronic brokers is more transparent, and the matching of limit and market orders is considered to be more efficient, since on a world-wide basis more participants (with different sources of information) trade via electronic brokers than via local voice brokers.

However, as pointed out above, the dataset that we use in our analysis contains customer trades from an internet trading platform (OANDA FXTrade). There are several reasons why this is an interesting dataset per se: i) customers have basically two possibilities to trade: either by trading with a dealer-bank or by trading on such a platform. As pointed out by Lyons (2002), there is recently a shift in the interdealer market from direct trading towards electronic brokerage trading. One argument to explain this shift is that there is more transparency on electronic brokerage systems. In the customer market one can expect the same shift from dealer-bank trading towards internet platform trading, since these platforms are also more transparent and moreover try to offer small (interbank) spreads to all customers. ii) customers orders reflect dispersed information and the flow (the dynamics or the interaction) of customer orders aggregates this information to a dealer-bank, or as Rime (2003) puts this “The trading that takes place with customers is private information for banks,...”. On internet trading platforms, customers can act as takers (submit market orders), as market makers (submit standard limit orders), or they can depending on their positions submit stop-loss, take-profit (special limit orders), and several further types. Since the information content and the aggregation of dispersed information probably differ between order types, this dataset allows us to analyze order flow, conditional on certain orders’ characteristics. iii) most of the research on order flow focuses on the interdealer market (e.g. Bjønnes & Rime (2003), Evans & Lyons (2002a,b), Payne (2003)), whereas the papers by Marsh & O’Rourke (2004) and Osler (2002) deal with customer orders observed by the Royal Bank of Scotland. However, to our knowledge there has been no analysis of customer data obtained by an internet trading platform. Based on Evans & Lyons (2002a,b) portfolio allocation model, customer orders are the primary source to identify dispersed information, whereas interdealer orders are only a consequence (e.g. “hot potato” trading, inventory control). Following this argument, the way dispersed information is aggregated can be analyzed with a customer dataset in a more reliable way than
with an interdealer dataset. iv) without using the portfolio allocation argument or the argument of Sarno & Taylor (2001) that order flow is a proxy for macroeconomic fundamentals but referring e.g. to Osler (2002), order flow contains information even if this information is not linked to interpretations of macroeconomic fundamentals and new announcements. This so called non-informative order flow, that may be linked to special limit-price setting strategies or technical analysis patterns and induces a specific clustering or herding behavior, is also private information to banks. Such a non-informative order flow can also be analyzed with the customer dataset from OANDA FXTrade.

In contrast to orders in direct interdealer trading and even in brokeraged interdealer trading, customer orders on internet trading platforms are much more multi-faceted. The standard definition of an aggregated net orderflow measure\(^2\) as the difference between buyer initiated and seller initiated trades, can be refined incorporating information on the order type and the status of the position. In our empirical analysis we will compare the standard orderflow measure with a refined orderflow measure. Therefore, it is indispensable to understand the trading mechanism and the order types on OANDA FXTrade.

**OANDA FXTrade in Detail**

OANDA stands for Olsen and Associates, and was established in 1995 by Richard Olsen. In early 2001, the company established a FXTrade platform, allowing customers to participate directly in the foreign exchange market in a very convenient way. This platform is a fully virtual marketplace for trading currencies via the internet, without limits on the trade size, and with 24 hours trading time, 7 days per week. The OANDA FXTrade platform is a market making system that executes orders using the exchange rate prevalent in the market (determined either by their limit order book or by a crossing network). OANDA FXTrade offers immediate settlement of trades and tight spreads as low as 2 to 3 pips on all transaction sizes. Given various boundary conditions, as for example sufficient margin requirements, orders are always executed. The OANDA FXTrade platform is based on the concept of margin trading, this means that the trader can enter into positions larger than his funds. The platform requires a minimum initial margin of 2% on positions in the major currency pairs and 4% in all other currency pairs, which correspond to

---

\(^2\)See e.g. Lyons (2001).
a leverage\textsuperscript{3} of 50:1 and 25:1 respectively. In other words, for each dollar margin available the trader can make a 50 dollar trade. The advantage (disadvantage) of margin trading is that the trader can leverage his capital and generate perhaps large profits (losses) relative to the amount invested. The margin available, a margin calculator, as well as the resulting units available to trade and the account summary, are provided on the interface. The trader receives a margin call when the net asset value (i.e. the current value of all open positions plus the value of the remaining deposited funds) becomes half the margin requirement. In other words, if the trader has not sufficient margin to cover twice the losses on an open position, a margin call order is used to close automatically all open positions using the prevalent market rates at this time.

Market orders (buy or sell) are executed immediately and affect existing open positions. The exchange rate used for the trade will not necessarily correspond to the exchange rate displayed in the buy/sell order window, but to the current exchange rate maintained at the server. In order to avoid a large fluctuation of the price between the submission and the execution of the order, the trader can specify lower and upper bounds while entering the order. An order will then be executed only if the exchange rate to be applied is located inside the bounds, otherwise the market order will be cancelled due to bound violations.

Limit orders are maintained in the system for up to one month. The server manages the limit order book, the current exchange rates, and the current market orders to match existing limit orders. The limit order can therefore be matched either against a market order or against a bid or an ask price obtained from the crossing network, which means that OANDA itself is at least for a while the counterparty. Stop-loss orders and take-profit orders are special limit orders in the sense that they can be set for existing open positions. They can be specified directly while entering a market or a limit order, but they can also be specified latter for existing open positions. Stop-loss and take-profit orders are automatically erased from the system whenever a position is closed due to further trading activity. Take-profit (TP) orders are typically set to close an existing position after a certain profit has been realized. Stop-loss (SL) orders, in contrast, specify that the position should be closed after the realization of a certain loss to avoid further losses.

\textsuperscript{3}A leverage of 50:1 is the maximum offered by OANDA FXTrade.
On the OANDA FXTrade platform, buying EUR/USD means that you are buying the base currency (EUR) and selling the quote currency (USD), whereas selling EUR/USD means that you are selling the base currency (EUR) and buying the quote currency (USD). Recorded units always refer to the base currency. The OANDA FXTrade record contains the following information on the transactions:

- **Buy/Sell market open (close)**: immediately executed to open or close a position in a specific currency pair.

- **Buy/Sell limit order**: the trader posted a buy or sell limit order to the system, which is then pending.

- **Buy/Sell limit order executed open (close)**: pending limit order is executed to open or close a certain position.

- **Buy/Sell take-profit close**: closes an open position by buying or selling the currency pair when the exchange rate reaches a predetermined level, in order to make a profit.

- **Buy/Sell stop-loss close**: closes an open position by buying or selling the currency pair when the exchange rate reaches a predetermined level in order to avoid further losses.

- **Buy/Sell margin call close**: closes automatically all open positions using the prevalent market rates at the closing time. This happens if the trader has not sufficient margin to cover two times the losses of all open positions.

- **Change order**: change of a pending limit order (limits for take-profit or stop-loss, the value of the upper or lower bounds, the quote as well as the number of units).

- **Change stop-loss or take profit on open trade**: change stop-loss or take-profit limit on an open position.

- **Cancel order by hand**: cancel a pending limit order by hand.

- **Cancel order: insufficient funds**: automatically recorded when the trader has not enough funds to open a new position.

- **Cancel order: bound violation**: market order or limit order is cancelled because the applied exchange rate is not located inside the specified bounds.

- **Order expired**: a pending limit order is expired.

The trading options on OANDA FXTrade are displayed in the following figure.
3 Description of the Dataset

Our data set is the activity record of OANDA FXTrade from 1st October 2003 to 14th May 2004 (227 days). This record contains for 30 currency pairs all trading activity (on a second by second basis) within this period and allows us to distinguish between the transaction types listed in Table 1 and described above. Furthermore, we have – depending on the order type – information, on the transaction prices (market orders, limit orders executed, stop-loss, take profit, margin call), on the bid and ask quotes (limit orders pending), on the transaction volume and, on the limits of stop-loss and take-profit orders. In our analysis we focus only on the most actively traded currency pair EUR/USD, which accounted for nearly 39% of all records with an average interrecord-duration of 8.5 seconds. For EUR/USD 13.5% of all transactions have a transaction-volume between 1€ (min.) and 100€, whereas only 1% of all transactions have a transaction-volume, which ranges between 50,000€ and 1,000,000€ (max.), where the average transaction-volume per trade is 26,546 €. The average number of different traders per day is 744. These numbers suggest that a certain part of our customers are “small” private investors (speculators). Table 1 also shows how the transaction-records are proportioned in the particular order types. We observe that approximately 40% of all records are (buy/sell) market orders to open or close a position, whereas only approximately 10% are initial (pending) limited orders, of which 70% are finally executed (7% of all orders). 12% of all records are stop-loss, take-profit or margin call orders, and 6% are cancellations due to various reasons. A fairly large part of 22% are changes of stop-loss and/or take-profit limits.

We construct from this data set, using only executed orders (market orders, limit orders executed, stop-loss, take profit, margin call), an equidistant EUR/USD price series for 11 frequencies (1 min, 2 min, 5 min, 10 min, 20 min, 30 min, 1 hour, 2 hour, 4 hour, 8 hour, 1 day) applying a previous tick rule. Moreover, we compute on these frequencies, in addition to the standard net order flow measure, our refined net order flow measure. The focus in the definition of our refined net order flow measure is the causal relationship between the price change in a given period to the order flow in this period, i.e. we favor with the construction of the refined order flow measure the causality from price changes to order flow and not vice versa.
<table>
<thead>
<tr>
<th>Transaction Record</th>
<th>Percentages</th>
<th>Std. Order Flow Signs</th>
<th>Ref. Order Flow Signs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buy margin call (close)</td>
<td>0.12</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Buy take-profit (close)</td>
<td>3.14</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Buy stop-loss (close)</td>
<td>2.18</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Sell market (close)</td>
<td>10.27</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Sell market (open)</td>
<td>10.61</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sell limitorder executed (close)</td>
<td>0.46</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Sell limitorder executed (open)</td>
<td>2.92</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Limit order: Sell</td>
<td>4.76</td>
<td>not used</td>
<td>not used</td>
</tr>
<tr>
<td>Limit order: buy</td>
<td>5.41</td>
<td>not used</td>
<td>not used</td>
</tr>
<tr>
<td>Buy limitorder executed (open)</td>
<td>3.22</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Buy limitorder executed (close)</td>
<td>0.46</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Buy market (open)</td>
<td>13.10</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Buy market (close)</td>
<td>8.27</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Sell stop-loss (close)</td>
<td>2.55</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Sell take-profit (close)</td>
<td>3.49</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Sell margin call (close)</td>
<td>0.17</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change order</td>
<td>3.01</td>
<td>not used</td>
<td>not used</td>
</tr>
<tr>
<td>Change stop-loss or take-profit</td>
<td>22.36</td>
<td>not used</td>
<td>not used</td>
</tr>
<tr>
<td>Cancel order by hand</td>
<td>2.41</td>
<td>not used</td>
<td>not used</td>
</tr>
<tr>
<td>Cancel order: insufficient funds</td>
<td>0.28</td>
<td>not used</td>
<td>not used</td>
</tr>
<tr>
<td>Cancel order: bound violation</td>
<td>0.20</td>
<td>not used</td>
<td>not used</td>
</tr>
<tr>
<td>Order expired</td>
<td>0.65</td>
<td>not used</td>
<td>not used</td>
</tr>
</tbody>
</table>

Table 1: Col. 1 states the record entries, col. 2 gives the corresponding percentages, col. 3 contains the signs for the construction of the standard net order flow measure and col. 4 contains the signs for the construction of the refined net order flow measure.

**Construction of the Standard and the Refined Order Flow**

Order flow measures incorporate and aggregate information contained in different order types into one single figure. The standard net order flow measure is constructed by taking the difference between the number of buyer initiated trades and all seller initiated trades (within a given period), or stated differently the cumulative sum of signed order flow, where buyer (seller) initiated orders get positive (negative) signs. Given the information in our data set, we consider buy (sell) market orders as buyer (seller) initiated trades, buy (sell) limit orders executed as seller (buyer) initiated trades, since the submitter of the limit order is not the initiating party, but the market maker, buy (sell) stop-loss and take profit orders as seller (buyer) initiated trades, since these are special limit orders, and buy (sell) margin call orders

---

4Compare Lyons (1995).
as buyer (seller) initiated trades, since these are market orders.

Our refined net order flow measure is constructed to explain the instantaneous causal relationship from price changes to order flow motivated by the following arguments. Let us first of all distinguish whether one has an open position or not. If one possesses an open position, this position can either generate losses or profits. According to Locke & Mann (2000) profits are usually realized much quicker than losses. This means one will close an existing open position with an unrealized profit more often than an existing open position with an unrealized corresponding loss (assuming the price itself is a martingale). Assume for a while that one has built an open position by buying EUR/USD, i.e. buying EUR, selling USD and there is a positive price change, then the trader probably wants to realize this profit by submitting a sell market order or a sell limit order (with a tight limit) to the system, which would result in a positive correlation between price changes and the flow of sell market (close) or sell limit order executed (close) orders. Now assume that one has established an open position by selling EUR/USD, i.e. selling EUR, buying USD and there is a negative price change. One realizes this profit by submitting a buy market order or a buy limit order to the system, which would result in a negative correlation between price changes and the flow of buy market (close) or buy limit order executed (close) orders.

Let us now consider the case when one has no open position, and remember that a certain share of the customers are private investors or speculators\(^5\). Abstracting from macroeconomic news interpretations, this group of investors has to rely on other sources to obtain information on the possible future development of the exchange rate: for example chart analysis. There are several survey studies among foreign exchange market dealers, which provide evidence that at least dealers rely on chart analysis quite frequently. The survey study of Taylor & Allen (1992) for example shows that at least 90% of the London based dealers rely on chart analysis information in addition to fundamental information. Moreover, it shows that dealers try to infer especially short-term exchange rate behavior with the help of chart analysis. In the foggy world of chart analysis an easy strategy is to follow the current trend.\(^6\) Assuming at least on an intraday basis, if there are no news events that at least some investors are trend followers, whereas the remaining investors act randomly. One should then observe the following contemporaneous correlations: a positive correlation between the flow of buy market (open) or buy limit orders exe-

\(^5\)This has also been confirmed by Richard Olsen.

\(^6\)Compare for example the book by Covel (2004).
cuted (open) orders and price changes and a negative correlation between the flow of sell market (open) or sell limit orders executed (open) orders and price changes. Now we consider stop-loss and take-profit orders: a buy/sell stop-loss (close) order is executed if one has an open position obtained by selling/buying EUR/USD, which generates losses. This means a buy/sell stop-loss (close) order is executed after a (sequence of) positive/negative price change(s), i.e. a positive/negative correlation between the flow of buy/sell stop-loss (close) orders and price changes. A buy/sell take-profit (close) order is executed if one has an open position obtained by selling/buying EUR/USD, which generates profits. This means a buy/sell take-profit (close) order is executed after a (sequence of) negative/positive price change(s), i.e. a negative/positive correlation between the flow of buy/sell take-profit (close) orders and price changes.

We can treat margin call orders in a similar way: a buy/sell margin call (close) order is executed if one has an open position obtained by selling/buying EUR/USD, which generates large losses. This results similar to stop-loss orders in a positive/negative correlation between the flow of buy/sell margin call (close) orders and price changes. The signs for the construction of the aggregated order flow measures are stated in Table 1 as well. Please keep in mind, by constructing the order flow measures as cumulative sums, we completely ignore correlations between the flows of different order types. The postulated signs for the construction of our refined order flow measure are supported by the empirical correlations stated in Table 2. It is shown in Table 2 (last two rows) that the contemporaneous correlation between order flow and price changes is much higher for the refined net order flow measure than for the standard one on all frequencies.

Furthermore, by postulating a contemporaneous causality direction from price changes to order flow, according to the explanations given above, we are in line with Osler (2002), who also postulates this causality direction, but we are opposed to Evans & Lyons (2002a,b) portfolio allocation model, which favors the causality direction from order flow to price changes. Of course, the truth is somewhere in between and the explanations above are more suited to explain order flow induced by short term price fluctuations, whereas longer term price reactions due to order flow is perhaps better explained by a portfolio allocation model. Stated differently: on higher frequencies we expect the causality direction: price changes to order flow and on lower frequencies we expect the opposite. In the empirical part we therefore analyze the dynamic behavior of the order flow measures and price changes on several
frequencies. However, our analysis focuses more on higher frequencies, since we do not have sufficient observations on lower frequencies (e.g. daily or less). On a daily level for example we have only 227 (respectively 163 without weekends) observations.
Table 2: Contemporaneous correlations between price changes and order flows for each transaction type as well as between price change and standard net order flow (second last row) and price changes and refined net order flow.
4 Empirical Findings

Table 3 contains some descriptive statistics of the price changes, the standard net order flow and the refined net order flow measure on different frequencies. We observe for all frequencies that price changes have a positive mean, which can be explained by the fact that in the period that we consider, 1st October 2003 to 14th May 2004, the EUR/USD exchange rate is increasing. Furthermore, the means of both order flow measures are also positive, which confirm the positive price changes. Nevertheless, the refined order flow measure has not only a higher mean than the standard one, but also a higher standard deviation. An important point that we want to emphasize is that the standard order flow measure has for all frequencies a negative skewness, whereas the refined order flow measure is always positively skewed. This positive skewness indicates that there is more mass on the right hand side of the distribution than on the left hand side – compare Figure 7 in the Appendix, too. The positive skewness is quite natural since on average we have positive price changes, which means that one should expect more positive net order flows. Furthermore, the skewness of the refined order flow measure is decreasing in the frequency, which is in line with the means of the price changes. The skewnesses of the standard order flow measure are therefore contrary to the means of the price changes. We also computed the refined net order flow measure for a period where the exchange rate is on average decreasing. In this period we find a left-skewed refined order flow measure.

In the same way as volatility (e.g. absolute price changes, absolute returns, realized volatility) is subject to diurnally seasonality patterns (compare e.g. Andersen & Bollerslev (1997) or Dacorogna, Gençay, Müller, Olsen & Pictet (2001)), we expect absolute order flow (volatility of order flow) to be subject to diurnally seasonality. Due to the fact that our refined order flow measure is positively skewed, we also expect a diurnally seasonality pattern in our refined order flow measure. In Figure 1 we plotted the diurnally seasonality function, computed by a Nadaraya-Watson kernel regression with a Gaussian kernel and optimal bandwidth selection according to Silverman (1986) on a 10 min aggregation level. The seasonality functions are plotted for both net order flow measures as well as for positive order flows and negative order flows. The time scale is measured in Eastern Standard Time (EST).
<table>
<thead>
<tr>
<th>Frequency</th>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 min</td>
<td>price change</td>
<td>0.0009</td>
<td>2.7420</td>
<td>-0.12</td>
<td>68.31</td>
<td>-115</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>0.0800</td>
<td>4.3605</td>
<td>-0.30</td>
<td>128.24</td>
<td>-192</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>0.4544</td>
<td>7.4346</td>
<td>2.09</td>
<td>225.70</td>
<td>-318</td>
<td>306</td>
</tr>
<tr>
<td>2 min</td>
<td>price change</td>
<td>0.0017</td>
<td>3.6351</td>
<td>-0.10</td>
<td>45.34</td>
<td>-126</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>0.1594</td>
<td>6.5507</td>
<td>-0.09</td>
<td>91.94</td>
<td>-224</td>
<td>162</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>0.9068</td>
<td>12.3057</td>
<td>2.03</td>
<td>201.81</td>
<td>-508</td>
<td>487</td>
</tr>
<tr>
<td>5 min</td>
<td>price change</td>
<td>0.0043</td>
<td>5.3610</td>
<td>-0.13</td>
<td>37.64</td>
<td>-151</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>0.3984</td>
<td>11.2235</td>
<td>-0.55</td>
<td>70.88</td>
<td>-346</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>2.2677</td>
<td>23.7098</td>
<td>1.04</td>
<td>154.11</td>
<td>-975</td>
<td>615</td>
</tr>
<tr>
<td>10 min</td>
<td>price change</td>
<td>0.0087</td>
<td>7.3330</td>
<td>0.03</td>
<td>28.21</td>
<td>-154</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>0.7997</td>
<td>16.9357</td>
<td>-0.53</td>
<td>55.11</td>
<td>-432</td>
<td>327</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>4.5446</td>
<td>38.8332</td>
<td>1.31</td>
<td>105.95</td>
<td>-1215</td>
<td>873</td>
</tr>
<tr>
<td>20 min</td>
<td>price change</td>
<td>0.0176</td>
<td>10.0299</td>
<td>-0.03</td>
<td>14.11</td>
<td>-105</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>1.5964</td>
<td>25.7609</td>
<td>-0.51</td>
<td>41.43</td>
<td>-510</td>
<td>424</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>9.0816</td>
<td>61.5974</td>
<td>1.79</td>
<td>38.49</td>
<td>-802</td>
<td>939</td>
</tr>
<tr>
<td>30 min</td>
<td>price change</td>
<td>0.0264</td>
<td>12.4035</td>
<td>-0.11</td>
<td>18.10</td>
<td>-143</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>2.3934</td>
<td>33.1342</td>
<td>-0.28</td>
<td>28.43</td>
<td>-471</td>
<td>492</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>13.6326</td>
<td>85.4083</td>
<td>2.36</td>
<td>47.98</td>
<td>-1187</td>
<td>1420</td>
</tr>
<tr>
<td>1 hour</td>
<td>price change</td>
<td>0.0529</td>
<td>17.5085</td>
<td>-0.05</td>
<td>11.31</td>
<td>-149</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>4.7985</td>
<td>51.0708</td>
<td>-0.16</td>
<td>19.65</td>
<td>-580</td>
<td>596</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>27.2751</td>
<td>134.3025</td>
<td>1.92</td>
<td>23.83</td>
<td>-816</td>
<td>1683</td>
</tr>
<tr>
<td>2 hours</td>
<td>price change</td>
<td>0.1066</td>
<td>24.9975</td>
<td>0.12</td>
<td>12.24</td>
<td>-175</td>
<td>216</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>9.6410</td>
<td>80.1774</td>
<td>-0.26</td>
<td>14.25</td>
<td>-569</td>
<td>684</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>54.5446</td>
<td>221.4514</td>
<td>2.05</td>
<td>20.63</td>
<td>-1154</td>
<td>2166</td>
</tr>
<tr>
<td>4 hours</td>
<td>price change</td>
<td>0.2169</td>
<td>36.1813</td>
<td>-0.08</td>
<td>8.47</td>
<td>-208</td>
<td>209</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>18.9308</td>
<td>132.8098</td>
<td>-0.27</td>
<td>12.57</td>
<td>-809</td>
<td>996</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>108.7047</td>
<td>353.7327</td>
<td>1.18</td>
<td>10.76</td>
<td>-1762</td>
<td>2338</td>
</tr>
<tr>
<td>8 hours</td>
<td>price change</td>
<td>0.4077</td>
<td>49.5924</td>
<td>0.10</td>
<td>5.46</td>
<td>-190</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>37.5213</td>
<td>218.0006</td>
<td>-0.69</td>
<td>8.62</td>
<td>-1120</td>
<td>961</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>216.5477</td>
<td>529.6934</td>
<td>0.75</td>
<td>6.34</td>
<td>-1730</td>
<td>2458</td>
</tr>
<tr>
<td>1 day</td>
<td>price change</td>
<td>0.7605</td>
<td>88.2221</td>
<td>-0.24</td>
<td>3.03</td>
<td>-257</td>
<td>251</td>
</tr>
<tr>
<td></td>
<td>std. order flow</td>
<td>110.4371</td>
<td>493.7761</td>
<td>-1.13</td>
<td>7.95</td>
<td>-2267</td>
<td>1469</td>
</tr>
<tr>
<td></td>
<td>ref. order flow</td>
<td>634.9880</td>
<td>1051.5044</td>
<td>-0.05</td>
<td>3.76</td>
<td>-2702</td>
<td>3238</td>
</tr>
</tbody>
</table>

Table 3: Descriptive Statistics for EUR/USD.

First we see that there is a diurnally seasonality pattern. It corresponds to a standard market activity: a positive peak at 3 o’clock, when the European traders enter the market, a slight decrease between 6-7 o’clock, when there is lunch time in Europe, a large positive peak 9-10 o’clock, when the American traders enter the market and when there are most macroeconomic news announcements, a decrease until 16 o’clock, when the American traders leave the market and again a slight recovery between 20 and 22 o’clock, when Asian traders start to trade. We observe this pattern for positive and negative (inverse pattern) order flows for both order flow...
measures (lower four panels). The pattern is more pronounced for the refined order flow measure than for the standard order flow measure. However, for the standard net order flow measure, which is only very slightly negatively skewed, the seasonality functions for positive and negative order flow cancel each other. For the positively skewed refined order flow measure this cancellation works only partially, so that a diurnally seasonality pattern remains. Please, keep in mind that we postulate the same seasonality pattern for every weekday, because on a daily frequency we have only 163 observations.

![Diurnal Seasonality](image)

**Figure 1:** Diurnally seasonality in the standard (1st column) and the refined (2nd column) net order flow measure, computed on a 10 min aggregation level.
Figure 2 and Figure 3 depict the empirical bivariate autocorrelation functions up to 30 lags between price changes and standard net order flow as well as refined net order flow for the considered frequencies. For every frequency there are 4 panels, the upper left panel depicts the autocorrelation function of the particular order flow measure, the lower right panel depicts the autocorrelation function for price changes. For these two, we plotted lag 1 up to lag 30. The lower left panel depicts the cross-correlation function of lagged order flow with price changes and the upper right panel depicts the cross-correlation function of lagged price changes with order flow. For these two, we plotted lag 0 up to lag 29. The value at lag 0 is in both cross-panels the same and represents the contemporaneous correlation between order flow and price changes. The analysis of the bivariate autocorrelation functions allows us to get some ideas of the dynamic interaction and behavior of order flow and price changes. The following observations are worth to point out: i) There is only significant positive contemporaneous correlation between the standard order flow measure and price changes for aggregation levels up to 10 min. Thereafter up to 1 day there is no significant contemporaneous correlation anymore (Figure 2). In contrast, there are significant positive contemporaneous correlations between the refined order flow measure and price changes for all aggregation levels. Furthermore for frequencies (1 min to 10 min) these correlations are significantly larger than those stemming from the standard order flow measure. This observation is in line with Table 2 and indicates that the refined order flow measure has a higher contemporaneous correlation with price changes than the standard one. ii) For the refined order flow (Figure 3), we observe in the upper right panels (1 min to 2 hours) further significant cross-autocorrelations, which indicate that future order flow is driven by current price changes. This effect, however, seems to be a short term effect since autocorrelation coefficients are significant between order flow and future price changes, which are up to 1.5 hours in the future. This is in line with the structure of our customer dataset that consists of a certain part of private investors, which we expect to trade based on historical developments (chart analysis). Furthermore, this observation partially supports the hypothesis that in the short run the Granger causality goes from order flow to price changes and not vice versa. This effect and therefore this Granger causality direction is also observable for the standard order flow measure (Figure 2), however it is not that pronounced and only observable up to 10 min. iii) The lower left panels, allow us to analyze the opposite effect, cross-autocorrelations between current order flow and future price changes. For this direction we observe only hardly significant autocorrelations. There is only negative
Figure 2: Empirical multivariate autocorrelation function of price changes and standard order flow for different aggregation levels. For every frequency there are 4 panels, the upper left panel depicts the autocorrelation function (lag: 1–30) of the particular order flow measure, the lower right panel depicts the autocorrelation function (lag: 1–30) for price changes. The lower left panel depicts the cross-correlation function (lag: 0–29) of lagged order flow with price changes and the upper right panel depicts the cross-correlation function (lag: 0–29) of lagged price changes with order flow. The dotted lines mark the approximate 99% confidence bounds, computed as $\pm \frac{2.58}{\sqrt{T}}$, where $T$ denotes the particular number of observations.
Figure 3: Empirical multivariate autocorrelation function of price changes and refined order flow for different aggregation levels. For every frequency there are 4 panels, the upper left panel depicts the autocorrelation function (lag: 1–30) of the particular order flow measure, the lower right panel depicts the autocorrelation function (lag: 1–30) for price changes. The lower left panel depicts the cross-correlation function (lag: 0–29) of lagged order flow with price changes and the upper right panel depicts the cross-correlation function (lag: 0–29) of lagged price changes with order flow. The dotted lines mark the approximate 99% confidence bounds, computed as $\pm 2.58 \sqrt{\frac{T}{T}}$, where $T$ denotes the particular number of observations.
significant first order correlation for aggregation levels up to 5 min for both order flow measures, which may be induced by the bid-ask bounce of the price change series itself. Compare the negative first order autocorrelations in the lower right panel (up to 10 min). Since this is a known phenomenon, we pay no further attention to the observed effect. iv) In the upper right panel, we can observe the autocorrelation function of the order flow measures themselves. On a 1 min frequency, we observe for the refined order flow measure a very slow declining autocorrelation function. For the standard order flow measure the autocorrelation function declines also very slowly but on a smaller level. An interesting feature is that for the refined order flow, the lower the frequencies the more sharply declining becomes the autocorrelation function, whereas for the standard order flow the autocorrelation function becomes more significant. This is an indication that for the refined order flow measure more information on the current value is contained in the nearer lagged values (in the recent history), whereas for the standard order flow measure, information on the current value is more dispersed in the past.

Based on the results of our explorative analysis we propose a two-dimensional Structural Vector Autoregressive (SVAR) model to explain the dynamics of order flow and price changes. Although price changes and both order flow measures are discrete variables, the histograms in Figures 5 to 7 (Appendix) depict that the range is fairly large (except for 1 min and 2 min frequencies). Moreover, in the estimation both order flow measures are used in their de-seasonalized form, where we assumed an additive seasonality function. This seasonality adjustment can be seen as a transformation from the discrete space to the real space. Due to these arguments, we consider a SVAR model, which lives on the real space as appropriate and we do not account for the discreteness of the price change time series. The SVAR model finally takes the following form:

\[
\Delta p_t = c_1 + \phi_{11}^{(1)} \Delta p_{t-1} + \phi_{12}^{(1)} x_{t-1}^k + \ldots + \phi_{11}^{(p)} \Delta p_{t-p} + \phi_{12}^{(p)} x_{t-p}^k + \varepsilon_{1t},
\]

\[
x_{t}^k = \phi_{21}^{(0)} \Delta p_t + c_2 + \phi_{21}^{(1)} \Delta p_{t-1} + \phi_{22}^{(1)} x_{t-1}^k + \ldots + \phi_{21}^{(p)} \Delta p_{t-p} + \phi_{22}^{(p)} x_{t-p}^k + \varepsilon_{2t},
\]

with the corresponding vector notation:

\[
\begin{pmatrix}
\Delta p_t \\
x_t^k
\end{pmatrix} = c + \Phi_p(L) \begin{pmatrix}
\Delta p_t \\
x_t^k
\end{pmatrix} + \varepsilon_t, \quad \text{where } \Phi^{(0)} = \begin{pmatrix}
0 & 0 \\
0 & \phi_{21}^{(0)}
\end{pmatrix}.
\]

Thereby \(\Delta p_t\) denotes the price change from \(t - 1\) to \(t\) and \(x_t^k\) denotes the net order flow between \(t - 1\) and \(t\). The superscript \(k \in \{S, R\}\) refers to the standard (S) order flow or to the refined (R) order flow measure. \(\varepsilon_t\) is assumed to be bivariate normal.
distributed, with zero mean, zero covariance and no restrictions on the variances. The identification of the SVAR model is achieved, assuming \( \Phi^{(0)} = \begin{pmatrix} 0 & 0 \\ \phi_{21}^{(0)} & 0 \end{pmatrix} \).

This identification is justified by the construction of the refined order flow measure, which emphasizes the instantaneous causality direction from price changes to order flow.\(^7\)

| Parameter \( \phi_{21}^{(0)} \) | Frequency |
|---|---|---|
|  | 10 min | 1 hour | 1 day |
| \( c' \) | (0.0088 -0.0356) | (0.0491 -0.2685) | (4.9817 379.5833***) |
| \( \Phi^{(1)} \) | \begin{pmatrix} -0.0769 *** & 0.0060 *** \\ 0.1076 *** & 0.1237 *** \end{pmatrix} | \begin{pmatrix} 0.0170 & -0.0006 \\ 0.2633 ** & 0.2507 *** \end{pmatrix} | \begin{pmatrix} -0.0200 & -0.0059 \\ -2.5269 *** & 0.4138 *** \end{pmatrix} |
| \( \Phi^{(2)} \) | \begin{pmatrix} -0.0234 *** & -0.0023 \\ 0.2395 *** & 0.0855 *** \end{pmatrix} | \begin{pmatrix} 0.0571 ** & -0.0094 *** \\ -0.6548 *** & 0.0937 *** \end{pmatrix} | \begin{pmatrix} \end{pmatrix} |
| \( \Phi^{(3)} \) | \begin{pmatrix} 0.0214 ** & -0.0019 \\ 0.0918 *** & 0.0518 *** \end{pmatrix} | \begin{pmatrix} 0.0120 & -0.0010 \\ -0.4808 *** & 0.0801 *** \end{pmatrix} | \begin{pmatrix} \end{pmatrix} |
| \( \Phi^{(4)} \) | \begin{pmatrix} 0.0337 *** & -0.0073 *** \\ -0.0244 & 0.0342 *** \end{pmatrix} | \begin{pmatrix} \end{pmatrix} | \begin{pmatrix} \end{pmatrix} |
| \( \Phi^{(5)} \) | \begin{pmatrix} 0.0233 *** & 0.0006 \\ -0.1084 *** & 0.0262 *** \end{pmatrix} | \begin{pmatrix} \end{pmatrix} | \begin{pmatrix} \end{pmatrix} |
| \( \Phi^{(6)} \) | \begin{pmatrix} 0.0080 & 0.0004 \\ -0.1079 *** & 0.0594 *** \end{pmatrix} | \begin{pmatrix} \end{pmatrix} | \begin{pmatrix} \end{pmatrix} |
| \( \Phi^{(7)} \) | \begin{pmatrix} 0.0066 & -0.0015 \\ -0.1509 *** & 0.0442 *** \end{pmatrix} | \begin{pmatrix} \end{pmatrix} | \begin{pmatrix} \end{pmatrix} |
| \( \Phi^{(8)} \) | \begin{pmatrix} -0.0068 & 0.0005 \\ -0.1289 *** & 0.0268 *** \end{pmatrix} | \begin{pmatrix} \end{pmatrix} | \begin{pmatrix} \end{pmatrix} |
| SIC | 10.6809 | 14.7226 | 22.5035 |
| Ljung-Box (Residuals) | LB(50): 292.7 *** | LB(20): 104.2 *** | LB(20): 59.2 |
| | LB(100): 527.8 *** | LB(50): 292.4 *** | LB(30): 98.9 |
| | LB(200): 1049.0 *** | LB(100): 532.5 *** | LB(50): 197.2 * |

**Table 4:** Estimation results of the SVAR models for refined net order flow and price changes, the Ljung-Box statistics refer to the multivariate Ljung-Box statistic, which is \( \chi^2 \) distributed with degrees of freedom equal to 4 times the lag length minus the number of estimated parameters. *, ** and *** denote significance levels of 10%, 5% and 1%.

\(^7\)Compare the construction of the order flow measures discussion.
### Table 5: Estimation results of the SVAR models for standard net order flow and price changes, the Ljung-Box statistics refer to the multivariate Ljung-Box statistic, which is $\chi^2$ distributed with degrees of freedom equal to 4 times the lag length minus the number of estimated parameters. *, ** and *** denote significance levels of 10%, 5% and 1%.

In Table 4 and Table 5 we present the estimation results of the SVAR model with the refined order flow measure and with the standard order flow measure, respectively. The results are given for three aggregation levels (10 min, 1 hour and 1 day). The selection of the number of lags in the SVAR model is based on the Schwarz Information Criterion (SIC). i) For the refined order flow measure $\phi^{(0)}_{21}$ is for all 3 frequencies significant on a 1% level. For the standard order flow measure $\phi^{(0)}_{21}$ is significant on a 1% level for 10 min data and on a 5% level for hourly and daily aggregation levels. This coefficient captures the contemporaneous correlation between price changes and order flow, and due to the chosen identification, quantifies how much of the current order flow is explained by the current price change. Similar to the explorative analysis of the bivariate empirical autocorrelation functions, we see (referring to the significance levels) that for lower frequencies (1 hour, 1 day)
the refined order flow measure explains more. ii) Considering on a 10 min level the 
\( \phi_{11}^{(i)} \), \( i \geq 1 \) coefficients, we see that we have found the bid-ask bounce (coefficients 
for \( i = 1,2 \) negative), which is somehow compensated over time (coefficients for 
\( i = 3,4,5 \) positive). On an hourly and daily level this effect disappears in both 
regressions. Considering the \( \phi_{22}^{(i)} \) coefficients that capture the autocorrelation structure 
in the order flow measures, we see that (on all three frequencies) all of them 
are significant, which also corresponds to the results of the explorative analysis. iii) 
Considering the cross effects: \( \phi_{12}^{(i)} \) captures the effect of lagged order flows on current 
price changes, with a few exceptions all of these coefficients are insignificant. This 
means that one is not able to explain a future price change with current order flow. 
This observation is quite natural and is completely in line with market efficiency 
arguments. On the other hand, considering the \( \phi_{21}^{(i)} \) coefficients, we observe for both 
order flow measures mainly significant effects. This corresponds to our explorative 
analysis, too and shows that future order flow is influenced by current and lagged 
price changes. Even on a daily level this effect is significant.

We evaluate the goodness-of-fit of our regressions by considering the empirical bi-
variate autocorrelation functions of the residuals (Figure 4), and by verifying the 
multivariate Ljung-Box statistics (Tables 4 and 5). We observe that for both order 
flow measures a large part of the auto- and the cross-correlations are explained by 
the simple SVAR models. However, on the higher frequencies (10 min and 1 hour), 
there are still some correlation coefficients which are significantly different from zero. 
This means that there is still some dynamics left that cannot be captured by the 
selected specifications. This observation (for both order flows) is confirmed by the 
multivariate Ljung-Box statistics for the residual series. On a 10 min and 1 hour 
level, the null hypothesis that there is autocorrelation up to the specific lag cannot 
be reject on standard significance levels, for lags 50,100,200 (10 min) and lags 20, 
50, 100 (1 hour). We can reject on a daily level for the refined order flow the null 
on a 10 % level (20, 30 lags) and on a 5% level (50 lags). For the standard order 
flow measure (daily level) the null can be rejected on a 1 % level (20, 30 lags) and 
on a 5% level (50 lags). Although, these observations confirm the patterns in the 
empirical bivariate autocorrelation functions on a daily level, please keep in mind 
that on a daily level our results are based on only 163 observations, whereas on 
hourly data we have 3911 observations and on a 10 min aggregation level there are 
23471 observations. We also tried to fit an appropriate SVAR model for 1 min data. 
Unfortunately we have not been able to find a parsimoniously parameterized model 
in the class of SVAR models, based on the SIC and by considering the residual prop-
erties. Altogether, we can derive the following conclusion: The SVAR models are well suited to explain a considerable amount of the dynamics pattern (first moments) on low and moderate frequencies. But the higher the frequency the more inflexible appear the SVAR models considered here. However, this observation seems to be quite natural, since we have considered only plain SVAR models, without further regressors to capture for example news effects, inventory control problems, or additional short term phenomena which can be expected to be compensated over time. Therefore, the analysis above is meant to receive a first impression on the dynamics of order flow and price changes on different aggregation levels for a typical internet trading platform customer dataset.
Residuals: std. order flow and price change

Residuals: ref. order flow and price change

10 min

1 hour

1 day

Figure 4: Empirical multivariate autocorrelation function for residuals of refined order flow and price changes (1st column) as well as of standard order flow and price changes (2nd column) for different aggregation levels. The dotted lines mark the approximate 99% confidence bounds, computed as $\pm 2.58 \sqrt{\frac{T}{T}}$, where $T$ denotes the number of observations (163 daily, 3911 hourly and 23471 10-minutely).
5 Conclusion

In this paper we analyze the dynamics of order flow and price changes with the help of a customer dataset from OANDA FXTrade, a prominent internet currency trading platform. The features of the dataset are that a certain part of the customers consists of “small” private investors. Moreover, the trading mechanism on OANDA FXTrade allows us to analyze order flow for distinct types of orders. In particular, we can distinguish between orders (market, limit, stop-loss, take-profit) to close a previous established position or to open a new position in a currency pair. This information enables us to construct a refined net order flow measure, which we compare against the standard net order flow measure (difference between buyer initiated trades and seller initiated trades). Both net order flow measures are computed for the EUR/USD exchange rate on 11 different frequencies from 1 min to 1 day and the interaction with the corresponding price change series is analyzed on these frequencies.

The explorative analysis shows that the refined net order flow measure is right-skewed in upward trending price periods and left skewed in downward trending price periods. This is quite natural because in upward trending price periods there should be more positive order flows, whereas in negative trending price periods there should be more negative order flows. For the standard net order flow measure we have not found this property. The skewness, however, induces a seasonality pattern in upward and downward trend price periods. This can be explained using the following argument: Absolute net order flow is subject to diurnally seasonality, which induces a seasonality pattern for negative and positive order flows. Since these seasonality patterns do not cancel each other in downward and upward trend price periods, the diurnally seasonality is carried forward to the net order flow measure.

Furthermore, we find that the refined net order flow measure is better explained by contemporaneous price changes than the standard one. We favor the contemporaneous causality direction from price changes to order flow instead of the opposite direction, which conforms to Osler (2002). Several arguments are presented why this contemporaneous causality direction is favored, especially on high aggregation levels. For all different aggregation levels, there exists in fact a significant positive contemporaneous correlation between the refined net order flow measure and price changes, whereas there is only a significant positive contemporaneous correlation between the standard net order flow measure and price changes for aggregation lev-
els till 10 minutes. Moreover, we find apart from this contemporaneous correlation, there is Granger causality between price changes and order flow, whereas a Granger causality relationship from order flow to price changes can only hardly be found. In addition, we try to explain the dynamics of order flow and price changes with SVAR models. These models are able to explain a large part of the dynamics of the net order flow and the price changes for low frequencies (10 min up to 1 day). However, for higher frequencies these specifications seem to be not flexible enough to capture all components of the dynamics.
References


Figure 5: Histograms of the price changes computed on different frequencies.
Figure 6: Histograms of the standard net order flow measure computed on different frequencies.
Figure 7: Histograms of the refined net order flow measure computed on different frequencies.