Forecasting the US Dollar/Euro Exchange Rate:

The Economic Value of Combining Fundamentals, Technical Analysis, and Order Flow

Master Thesis
Quantitative Finance
Thijs Klein Tank
299399

Supervisor:
Prof. dr. Dick van Dijk (EUR)

Co-reader:
Assistant prof. Michel van der Wel (EUR)

Department of Econometrics & Management Science
Erasmus University, Rotterdam
August 15, 2012

Abstract

I analyze the economic value of combining three types of information to forecast the US dollar/euro exchange rate: macroeconomic fundamentals as used in structural exchange rate models, information from historical daily prices as used in technical trading rules, and intraday transactional data as used in order flow models. The out-of-sample period starts in January 2005 and ends March 2010. I find that fundamental Taylor rules yield a significant 5.38% annual return out-of-sample. Technical channel rules result in an annual out-of-sample return of 3.52%, but the return is insignificant at the 5% level. Order flow results, based on intraday CME futures data and the popular tick test to sign trades, disappoint with an annual return of minus 2.82% with a t-value of -0.62. As a consequence, combining Taylor rules with channel rules, order flow, or both does not improve the economic performance.

Keywords: forecasting, foreign exchange rates, Taylor rules, technical trading, order flow

Contents

1 Introduction

For decades, both academics and market participants diligently explore methods to forecast exchange rates. Previous research extensively analyzes the predictive contact of two types of information: news on macroeconomic fundamentals as used in structural exchange rate models, and information from historical prices as used in technical trading rules.

In their seminal paper, Meese and Rogoff (1983a,b) show that exchange rate forecasts from fundamental models do not outperform a random walk. Cheung et al. (2005) systematically evaluate five structural models for the period 1973 to 2000. They also find that none consistently outperform the random walk at any horizon. However, Molodtsova and Papell (2008) and Molodtsova et al. (2008) find positive results using Taylor rules, which Chen and Chou (2010) confirm.

Park and Irwin (2007) conduct a comprehensive literature study on technical forecasting models. They find that modern studies, with improved testing procedures, show positive evidence of technical trading rule profitability. But profitability of relatively simple technical rules tends to deteriorate or disappear in forex markets. More complex rules still show positive results, but profits also seem to diminish over time.

More recently, order flow gains interest as a potential third predictor of short term exchange rate movements. Order flow is a measure of net buying pressure. There is overwhelming academic evidence that order flow is the mechanism by which private information becomes embedded in exchange rates, see Evans and Lyons (2002) and Evans and Lyons (2009), among others. The forecasting abilities of order flow are less studied and results are mixed. Where for example Evans and Lyons (2009) and Rime et al. (2010) present positive results, Reitz and Schmidt (2007) and Sager and Taylor (2008) find little evidence that order flow helps to forecast exchange rates.

Academic research often considers the predictive ability of exchange rate forecasting models in isolation. But Gehrig and Menkhoff (2006) show that most professional traders combine macroeconomic fundamentals and technical analysis to arrive at investment decisions. de Zwart et al. (2006) find supportive evidence that coupling these two types of information results in more profitable trading strategies. The survey of Gehrig and Menkhoff (2004) documents that flow analysis is an independent extra type of information for currency professionals, in addition to fundamentals and technicals.

In this thesis report, I examine whether combining fundamental and technical models with order flow information benefits the economic performance of trading rule strategies in the most heavily traded global currency pair: the US dollar/euro. More traditional performance measures include risk-adjusted returns such as measured by the Sharpe ratio. Berge et al. (2010) and Jordà and Taylor (2011) recently introduce other economic performance measures. I evaluate both traditional and newer types of economic measures. Also, I incorporate transaction costs by defining turnover for the trading signals.

I study literature to find which specific methods perform best for a range of currency pairs and forecasting horizons. To generate fundamental trading signals, I select Taylor rules. Molodtsova and Papell (2008) evaluate the performance of Taylor rule exchange rate predictions based on statistical measures. I contribute to their research by investigating economic measures. Furthermore, they do not combine model variants, whereas I develop a composite Taylor rule trading signal. Thirdly, they do not combine forecasts with technical trading rules. For technical signals, I select channel rules. Fourthly, I investigate whether adding order flow information improves forecasts. Data availability is the limiting factor for order flow information. I avoid flow data restrictions by using futures data from the Chicago Mercantile Exchange (CME).

I use an in-sample period from January 1999 to January 2005 to determine the optimal parameters for the fundamental Taylor rules and technical channel rules. The out-of-sample period starts in January 2005 and ends in March 2010. Based on the in-sample results, I set up one optimal composite buy/sell signal for both the Taylor rule and channel rule models. I also set up an order flow signal based on the tick test.

The composite Taylor model generates a significant out-of-sample annual return of 5.38% with a Sharpe ratio of 0.74 and t-value of 1.68. The composite channel model yields a 3.52% return with a Sharpe of 0.43. But with a t-value of 0.98, the channel return is not significant at the 5% level. The flow model returns an insignificant, negative 2.82% per year. Combining the signals provides diversification benefits. But the results of the channel and order flow signals are too weak compared to Taylor results to improve the economic Taylor rule out-of-sample performance measures.

2 Literature

2.1 Fundamental models

Fundamental models try to explain or forecast exchange rates using economic fundamentals based on open-economy macro theory. Seminal work by Meese and Rogoff (1983a,b) examines the out-of-sample performance of three 1970's vintage empirical sticky- and flexible-price monetary models. They find that none outperforms a random walk model at one to twelve month horizons. Since Mark (1995), a number of studies find evidence of greater predictability at longer horizons. However, Kilian (1999) questions these findings.

Cheung et al. (2005) systematically evaluate five structural models using quarterly exchange rate data for the period 1973 to 2000: purchasing power parity, uncovered interest rate parity, a sticky-price model, a productivity differential model, and a composite model. They conclude that none of these models consistently outperform the random walk at any horizon. Rogoff (2009) cautiously states, in an update to his and Meese's seminal paper, that the most promising forecasting models are those based on purchasing power parity or the current account. Though he notes that these mainly predict the real exchange rate, rather than the quoted nominal exchange rate.

Molodtsova and Papell (2008) do find evidence of short-term exchange rate predictability using models that incorporate Taylor rule fundamentals. The Taylor rule, introduced by Taylor (1993), models interest rate setting policies. At the same time, they also do not find much evidence of predictability using interest rate, monetary, and purchasing power parity fundamentals. Recent findings by Chen and Chou (2010) confirm the outperformance of the Taylor-rule model over conventional exchange rate models.

For the US dollar/euro specifically, Molodtsova et al. (2008) confirm that Taylor rule fundamentals provide evidence of exchange rate predictability from 1999 to 2007. Taylor rules seem to provide a reasonable approximation of interest rate setting in the US and euro area. Molodtsova et al. (2008) also find that predictability increases with real-time data compared to revised data, is about the same with inflation forecasts as with actual inflation rates, and weakens if output gap growth is included in the forecasting regression.

I conclude that economics literature mostly fails to find convincing evidence that forecasts from fundamental structural exchange rate models can outperform the random walk. But the recent development of Taylor rule models provides a promising fundamental approach. I will therefore use Taylor rules as the fundamental model component.

2.2 Technical models

Technical models try to forecast price movements using only past prices, volume, and open interest. Predictions emerge from strict 'if-then' trading rules, the most well-known being moving average and filter rules.

Academic economists are skeptic. Technical analysis relies solely on past exchange rate movements and therefore violates the weak-form of the efficient market hypothesis. But technical analysts point to the semi-strong efficient market hypothesis in favor of their métier. They argue that it is pointless to investigate news and fundamentals, as this information is already incorporated in prices. A comparison of surveys from Gehrig and Menkhoff (2006) with Taylor and Allen (1992) shows that technical analysis has gained popularity over time, even among currency dealers. Gehrig and Menkhoff (2006) state that charting may be called the 'workhorse' of forex dealers. In fund management, it ranks second only after fundamental analysis.

Park and Irwin (2007) review historical research on technical analysis. They report that early studies, from 1960 to 1987, frequently show sizable net profits for futures and forex markets. But most of the early empirical researchers use unreliable testing procedures including data-snooping, ex post selection of trading rules, and risk ignorance. Modern studies, from 1988 to 2004, improve these procedures. Park and Irwin (2007) examine a total of 38 modern studies on forex markets. The find that 24 report positive, 6 mixed, and 8 negative results. The annual returns range from 5% to 10% net of transaction costs. A wide variety of strategies such as moving averages, channel, filters, and genetically formulated rules consistently generate profits until the early 1990s. Several recent studies confirm the result, but also report that technical trading profits decline or disappear over time. Park and Irwin (2007) conclude that moving average and channel rules are the most consistent profitable strategies for futures markets.

Menkhoff and Taylor (2007) find firstly that transaction and interest rate costs do not necessarily eliminate the profitability of technical analysis, secondly that technical strategies tend to be more profitable with volatile currencies, and thirdly that the performance of technical trading rules is highly unstable. Menkhoff and Taylor (2007) report the same phenomenon as Park and Irwin (2007): there is evidence that profits from technical strategies decline over time. Qi and Wu (2006) add evidence to this statement. They analyze the profitability and statistical significance of 2,127 technical trading rules. Traditional moving average rules, profitable in the 1970s, become much less profitable in the 1990s, even after allowing for a reduction in transaction costs over time.

Neely et al. (2009) replicate exchange rate forecasting models from early papers with data up to 2005. They use filter rules, moving averages, channels, ARIMA-models, genetic programming (GP), and Markov switching models and correct for declining trading costs. Neely et al. (2009) find that the returns for filter and moving average rules decline dramatically over time, in some cases to the point that rules earn significant negative returns. But they do not find statistically significant declining trends in the net returns of less-studied or more complex rules, such as channel, ARIMA, genetic programming, and Markov rules. Neely et al. (2009) note that technical traders commonly use channel rules.

Decades of research mostly shows that forecasts from fundamental models do not outperform a random walk. It is therefore not surprising that nowadays, technical analysis is a popular forecasting tool among currency market professionals. Park and Irwin (2007) show that modern academic studies on technical foreign exchange forecasting mostly report positive results, but performance of rules seems to decline over time. Channel rules stand out as a technical approach. These rules display a longer record of relatively positive and stable performance. Hence I use channel rules as the technical model.

2.3 Order flow

Discouraging results of fundamental models and declining technical profits move researchers into new directions. Evans and Lyons (2002) augment traditional macro analysis with some price determination microeconomics. This leads to a new class of models that highlight new variables that macro models omit. Order flow is the most important such variable. Flow is the net of buyer-initiated and seller-initiated orders. Gehrig and Menkhoff (2004) show that flow analysis is also a major and independent third tool for forex professionals, in addition to fundamental and technical analysis.

2.3.1 A source of information?

Over the last decade, a series of papers report that order flow is an important component in the exchange rate pricing mechanism. Lyons (2001) and Lyons (2002) finds that there is considerable evidence that order flow accounts for the lion's share of floating exchange-rate movements, even when based on data sets that include only a fraction of market-wide flow. Marsh and O'Rourke (2005) confirm these findings. Lyons (2001) concludes that the price impact of forex orders from mutual funds and hedge funds appears significantly higher than from non-financial corporations. Evans and Lyons (2002) find that their model is quite successful in clarifying order flow's role in transmitting information to price.

Researchers also find indications that order flow information could be useful to forecast exchange rates. Using data from Citibank from 1993 to 1999, Evans and Lyons (2005) show that order flow forecasts consistently outperform both a standard macro model and and a random walk. Evans and Lyons (2009) show that flows carry information on the future of macro variables that drive the risk premium, such as GDP growth, money growth, and inflation. They find that the incremental macro information helps to forecast exchange rates. Again, they use the Citibank flow data. Evans (2010) establishes a link between the high-frequency dynamics of spot exchange rates and developments in the macro economy. Once more, he uses the same order flow data from Citibank as Evans and Lyons (2005) and Evans and Lyons (2009). This makes the positive evidence less convincing. It also reveals the inaccessible nature of order flow data. Rime et al. (2010) investigate Reuters interdealer data from 2004 to 2005. They find that order flow is a useful predictor of daily movements in exchange rates. Rime et al. (2010) conclude that the information in order flow cannot be captured by simple momentum or forward bias strategies.

Besides these positive findings, other researchers express doubt. Reitz and Schmidt (2007) examine different customers using tick data from a small German bank from October 2002 to September 2003. Although they find evidence in favor of the order flow information aggregation process, they do not share the widespread optimism that customer order flow is easily exploitable for speculative purposes. Sager and Taylor (2008) investigate the relationship between order flow and subsequent exchange rate returns. They use both inter-dealer and commercially available customer order flow data, also separating customer order flow data by customer group. They find little evidence that order flow can predict exchange rate movements out of sample. Sager and Taylor (2008) state that only dealers who observe order flow on a real-time, continuous and unfiltered tick-by-tick basis might be able to profit from order flow information.

2.3.2 Data accessibility

Obtaining historical order flow data poses a challenge as the data is either proprietary or very expensive. Let alone to obtain real-time data, which investors require for practical application of forecasting methods. I study a fast-growing subsection of the forex market for which I can readily obtain historical and observe real-time transaction data: exchange traded currency derivatives. I use Chicago Mercantile Exchange (CME) futures data on the euro/US dollar from 2003 to 2010.

According to the BIS (2010) Triennial survey, daily volume in exchange traded derivatives amounts to $ 168 billion in April 2010 on a total daily forex turnover of nearly $ 4 trillion. The survey also reports that in 2010, the US dollar/euro comprises 28% of total global foreign exchange market turnover, followed by 14% in the US dollar/yen and 9% in the US dollar/sterling. King and Rime (2010) note that algorithmic trading through electronic execution venues is the main growth driver of currency derivatives. The CME provides algorithmic traders with an electronic interface in 2002. The platform leads to a sharp increase in turnover from 2003 onwards. CME's average daily turnover in forex products rises from $ 40 billion in 2005 and $ 80 billion in 2008 to $ 110 billion in 2010. EBS is a popular electronic interdealer platform that can serve as a comparison. Daily EBS volume is $ 140 billion in 2005 and $ 210 billion in 2008. But the global financial crisis causes EBS volume to plummet to $ 135 billion in 2009, while CME volume manages to stabilize. If the current growth trends continue, CME volume overtakes EBS volume within the next few years.

Rosenberg and Traub (2006) find evidence that currency futures transactions contain more information than they expect based on market size. Futures order flow also appears highly correlated with interdealer spot order flow. Tse et al. (2006) find that CME's electronic platform plays a dominant role in price discovery for the euro. Price discovery is the ability of a market to provide information about prices. But Cabrera et al. (2009) find that EBS spot volume is more informative than CME traded futures. They conclude that the results are justifiable by the volume size differences. However, the BIS (2010) survey shows that exchange-traded futures are becoming an integral part of the global currency market.

2.3.3 Data signing

Nowadays, nearly all CME futures trading takes place on the electronic GLOBEX platform. Real-time CME exchange data is publicly available for a relatively low monthly fee[1]. Traders observe best bid and best ask quotes in real-time, together with transactional data. They can immediately deduce the sign of the trade: buyer or seller initiated.

Rosenberg and Traub (2006) use a quote rule signing algorithm by Hasbrouck and Ho (1987) and Hasbrouck (1988), which states that buyer initiated trades occur at the ask and seller initiated trades at the bid. Hasbrouck (1991) generalizes this approach for trades inside or outside the spread. He uses the bid-ask midpoint to distinguish buyer from seller initiated trades.

[1] $ 78 monthly `http://www.esignal.com/exchanges/default.aspx?name=North_South_AmericanRegions&tc=`

Bid and ask data is not available for my historical transactional data series. As the total data comprises over 100 million intraday observations, computational burden is an important consideration. Lee and Ready (1991) develop a fast and nowadays popular algorithm to infer the sign of a trade in the absence of quote data: the tick test. I use the tick test to sign trades.

2.4 Combining forecasts

Gehrig and Menkhoff (2006) and Menkhoff and Taylor (2007) show that forex dealers and fund managers base trading decisions on a combination of fundamental, technical, and flow information. Forex professionals predominantly use fundamentals at longer forecasting horizons (months to years), charts for shorter-term forecasting horizons (days to weeks), while order flows dominate at the shortest-term (minutes to hours).

Timmermann (2006) lists a number of compelling reasons to combine forecasts. On average, pooled forecasts outperform predictions from the single best model. Secondly, structural breaks may affect individual forecasts very differently. Thirdly, it is implausible that the same model dominates all others at all points in time. Notice the similarity to the classical portfolio diversification argument for risk reduction.

Timmermann (2006) notes that in many cases, forecasters can make dramatic performance improvements by plainly averaging predictions. Simple combination schemes are difficult to beat, as more complex methods introduce parameter estimation errors. de Zwart et al. (2006) show that applying equal weights to chartist and fundamentalist strategies in emerging currency markets generates more consistent and stable results than the individual strategies. Heterogeneous agents models assign weights according to their past performance. de Zwart et al. (2006) find that these models increase profitability only modestly. Based on these findings, I use equal weights to combine the forecasting signals.

3 Models

I set up daily buy/sell signals $z_t \in (-1,+1)$ using fundamental Taylor rules, technical channel rules, and order flow. Here, $z_t = +1$ indicates a full long position in the US dollar and $z_t = -1$ indicates a full short. $z_t = 0$ indicates a neutral or 'flat' position.

For Taylor rules, I define eight model specifications and four output gap specifications for a total of thirty-two model specifications. For channel rules, L is the single model parameter. L defines the length of the trading range for which a breakout generates a signal. I use the tick test to set up signals for the order flow model.

I analyze the ex post signal performance based on three groups of economic performance measures:

- Return
  - Mean annualized return
  - Sharpe ratio
  - Skewness
- Transaction costs
  - Turnover
  - Break-even transaction costs
- Direction
  - Accuracy
  - Gain/loss ratio

3.1 Fundamental model: Taylor rules

3.1.1 Specification

Taylor (1993) introduces the Taylor rule, which models interest rate setting policies. The rule states that a central bank adjusts the short-run nominal interest rate in response to changes in both inflation and the output gap. The output gap is the difference between GDP and potential GDP. Potential GDP is the output which could be realized without giving rise to inflationary pressures. Instead of GDP, which is only available on a quarterly basis, economists sometimes use proxies such as monthly productivity or unemployment figures.

Molodtsova and Papell (2008) list different model specifications for Taylor rules based on:

Symmetry Asymmetric models differ from symmetric models in that they include the difference between the real exchange rate and the purchasing power parity target exchange rate.

Smoothing Models with smoothing include a lagged interest rate differential to assume that the interest rate only partially adjusts to its target within the period.

Homogeneity Homogeneous models have equal response coefficients to reflect that two central banks respond identically to changes in in inflation, output gap, and interest rate smoothing. In heterogeneous models, the coefficients appear separately. Both types of models can either be symmetric or asymmetric, and with or without smoothing.

Constant If the two central banks have identical inflation targets and equilibrium real interest rates, there is no constant in the equation. Otherwise, there is a constant.

Fundamentals for exchange rate predictions arise when central banks set short term interest rates according to the Taylor rule. The expression for the Taylor rule is:

$$i_t^* = \pi_t - \phi(\pi_t - \pi_t^*) + \gamma y_t + e^* \tag{1}$$

with π_t the inflation rate, y_t the output gap, and e^* the equilibrium level of the real interest rate. The central bank aims for interest target rate i_t^* and inflation target π_t^*.

According to the rule, the central bank raises the short-term interest rate if inflation rises above its desired level. And raises the rate if economic output is above potential output. The target level for y_t is 0, because output cannot permanently exceed potential output. Taylor (1993) assumes equal importance of inflation and output with weights 0.5. And he sets both the equilibrium real interest rate and inflation target to 2 percent, as economists generally believe that moderate inflation is most beneficial to an economy.

Following Molodtsova and Papell (2008), I rewrite equation (1) by setting $\mu = r^* - \phi\pi^*$:

$$i_t^* = \mu + \lambda\pi_t + \gamma y_t \tag{2}$$

with $\lambda = 1 + \phi$.

Molodtsova et al. (2008) compare the policy of the ECB with the Bundesbank using Taylor rules. They find that the ECB, like the Fed but unlike the Bundesbank, does not put much weight on the exchange rate when setting interest rates. Hence I will exclude asymmetric model specifications.

I introduce smoothing by assuming that the observable interest rate i_t adjusts to the target i_t^* as follows:

$$i_t = (1-\rho)i_t^* + \rho i_{t-1} + \nu_t \tag{3}$$

with ρ the smoothing parameter. For no smoothing, $\rho = 0$.

Substitute equation (2) in (3):

$$i_t = (1-\rho)(\mu + \lambda\pi_t + \gamma y_t) + \rho i_{t-1} + \nu_t \tag{4}$$

I construct the interest rate differential to provide the Taylor-rule-based forecasting equation:

$$i_t^{US} - i_t^{EUR} = \alpha + \alpha_{US,\pi}\pi_t^{US} - \alpha_{EUR,\pi}\pi_t^{EUR} + \alpha_{US,y}y_t^{US} - \alpha_{EUR,y}y_t^{EUR} + \rho_{US}i_{t-1}^{US} - \rho_{EUR}i_{t-1}^{EUR} + \eta_t \quad (5)$$

where α is a constant. Here, $\alpha_\pi = \lambda(1-\rho)$ and $\alpha_y = \gamma(1-\rho)$.

Define s_t as the log spot price of one US dollar in euros at time t. Molodtsova and Papell (2008) study literature and postulate that any event that causes a central bank to raise interest rates produces both immediate and forecasted local currency appreciation. They implement this presumption in equation (5) to produce an exchange rate forecasting equation:

$$s_{t+1} - s_t = \omega + \omega_{US,\pi}\pi_t^{US} - \omega_{EUR,\pi}\pi_t^{EUR} + \omega_{US,y}y_t^{US} - \omega_{EUR,y}y_t^{EUR} + \omega_{US,i}i_{t-1}^{US} - \omega_{EUR,i}i_{t-1}^{EUR} + \eta_t \quad (6)$$

Note that the signs of coefficients in equations (5) and (6) are identical. This reflects the presumption that anything that causes the US interest rate to rise relative to the eurozone will cause immediate and forecasted US dollar appreciation.

The Taylor exchange rate forecasting specification of equation (6) is symmetric, with smoothing, heterogeneous, and includes a constant. For no smoothing, $\omega_{EUR,i} = \omega_{US,i} = 0$. For the homogeneous specification, $\omega_{US,\pi} = \omega_{EUR,\pi}$, $\omega_{US,y} = \omega_{EUR,y}$, and $\omega_{EUR,i} = \omega_{US,i}$. Finally, for the specification without constant, $\omega = 0$.

I evaluate the total of eight resulting symmetric Taylor model specifications. The specifications are either homogeneous or heterogeneous, either have smoothing or not, and either have a constant or not.

3.1.2 Output gap

We do not know which definition of potential output central banks use in their interest rate reaction functions. I therefore also consider four output gap measures y_t for every Taylor model specification. I follow Molodtsova and Papell (2008)'s approach to evaluate:

- deviations of the log industrial production from a linear trend;
- deviations of the log industrial production from a quadratic trend;
- deviations of the log industrial production from a Hodrick and Prescott (1997) (HP) filter trend with a smoothing parameter $\lambda = 14,400$.

Furthermore, I add an extra HP specification: deviations of the log industrial production from a Hodrick and Prescott (1997) (HP) filter trend with a smoothing parameter $\lambda = 129,600$ as suggested by Ravn and Uhlig (2002). I outline considerations for the HP filter in the next subsection.

To determine the log industrial production trend, I use all available production data (see chapter 4) with an expanding time window. I determine the output gap y_t by comparing the log industrial production trend 'potential' value with the actual observed value.

3.1.3 HP filter

The HP filter separates the cyclical component of a time series from raw data. The resulting smoothed-curve representation of a time series is more sensitive to long-term than to short-term fluctuations.

For estimates at point t, the HP filter uses data from before and after time t. For a real-time estimate of the output gap, the HP filter becomes one-sided. Gerlach (2001) shows that this gives rise to end-point problems, where the HP filter tends to treat the latest data point as the 'new normal'. This is especially severe at turning points of the business cycle.

I mitigate the end-of-sample problem by using a methodology by Clausen and Meier (2005), that Molodtsova et al. (2008) adopt. I forecast and backcast the industrial production series by twelve months in both directions, assuming that the data follows an AR(4) autoregressive process

$$X_t = c + \sum_{i=1}^{4} \varphi_i X_{t-i} + \varepsilon_t \tag{7}$$

where $\varphi_1, \ldots, \varphi_4$ are the model parameters, c is a constant, and ε_t is white noise. Then, I replace the last twelve data points with the backcasts and I add twelve point forecasts. On the resulting series, I apply the HP filter.

For the HP filter, let q_t denote the log values of a time series variable. The series q_t is made up of a trend component τ_t and a cyclical component c_t, such that $q_t = \tau_t + c_t$. Given an adequately chosen positive λ, there is a trend component that solves

$$min \left\{ \sum_{t=1}^{T} (q_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} \left[(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}) \right]^2 \right\}. \tag{8}$$

The first term in equation (8) penalizes the cyclical component. The second term penalizes variations in the second derivative (growth rate) of the trend component. A higher λ increases the penalty.

Hodrick and Prescott (1997) advise a value of $\lambda = 1,600$ for quarterly data and $\lambda = 14,400$ for monthly data. But Ravn and Uhlig (2002) show that while most while researchers use $\lambda = 1,600$ for quarterly data, there is less agreement when moving to other frequencies. Ravn and Uhlig (2002) derive a value for monthly data of $\lambda = 129,600$. I evaluate both settings for λ. I denote the HP filter with Hodrick and Prescott (1997) value for λ as 'HP' and Ravn and Uhlig (2002) value for λ as 'HP-R'.

3.1.4 Signal

I use OLS regressions to estimate the coefficients of the Taylor-rule-based forecasting equation (6) with a window expanding from three to six year in-sample and a six year rolling window out-of-sample, see section 4.2. Then, I construct one-month ahead return forecasts $\hat{r}_{t+1}^{Taylor}$.

In total, I have eight different Taylor rule specifications with four measures of output gap per specification. This yields thirty-two different model specifications M.

I develop the Taylor signal T_t for model M as:

$$T_t(M) = \begin{cases} 1 & if\ \hat{r}_{t+1}^{Taylor} \geq 0 \\ -1 & if\ \hat{r}_{t+1}^{Taylor} < 0 \end{cases} \tag{9}$$

I align monthly Taylor signals with daily technical and order flow signals by converting to daily signals. The monthly signal continues to hold for all trading days after the signal day until the next monthly signal day. Section 4.1 outlines considerations for the signal day. I assume that on every 15th day of the month, inflation and industrial production figures are available for the preceding month.

3.2 Technical model: channel rules

Channel rules take a long position when the price exceeds the maximum price observed over the previous L days. And go short when the price drops below the minimum price over the last L days. L is the single model parameter.

The channel rule signal for the US dollar/euro $C_t(L)$ using a range of L days is:

$$C_t(L) = \begin{cases} 1 & if\ s_t > max(s_{t-1}, \ldots, s_{t-L}), \\ -1 & if\ s_t < min(s_{t-1}, \ldots, s_{t-L}), \\ C_{t-1}(L) & otherwise. \end{cases} \tag{10}$$

3.3 Order flow

Order flow is the net of buyer-initiated and seller-initiated orders. It is measure of net buying pressure. Hasbrouck (1991) uses the bid-ask midpoint to distinguish active purchases and sales, see section 2.3.3. He labels a trade above the midpoint as buyer-initiated.

3.3.1 Signing data

As my available historical futures data lacks bid-offer data, I cannot readily determine the mid price. Lee and Ready (1991) develop a fast and nowadays popular algorithm to infer the sign of a trade in the absence of quote data: the tick test. Let $q_{t^*} \in \{-1,1\}$ denote the sign of the transaction at intraday time t^*, with $q_{t^*} = +1$ a buyer-initiated trade and $q_{t^*} = -1$ a seller initiated trade. And denote p_{t^*} the intraday euro/US dollar CME future transaction price at time t^*.

The tick test defines the sign as:

$$q_{t^*} = \begin{cases} 1 & if\ p_{t^*} > p_{t^*-1}, \\ -1 & if\ p_{t^*} < p_{t^*-1}, \\ q_{t^*-1} & if\ p_{t^*} = p_{t^*-1}. \end{cases} \tag{11}$$

3.3.2 Signal

For day t, denote the last intraday transaction as $t^* = \lambda$. The daily order flow signal O_t is:

$$O_t = \begin{cases} 1 & if\ \sum_{t^*=1}^{\lambda} q_{t^*} < 0, \\ -1 & if\ \sum_{t^*=1}^{\lambda} q_{t^*} \geq 0. \end{cases} \tag{12}$$

Note that I reverse the signs to ensure that a trading day with positive US dollar buying pressure also triggers an order flow model buy signal. The CME transactional data is the price of one euro in US dollars p_{t^*} that generates trade sign q_{t^*}. But I observe the inverse price, the log spot price of one US dollar in euros s_t, for all models.

Suppose that large, professional market participants are more informed than smaller retail traders, see Lyons (2001). And that transactions with a higher number of future contracts are therefore more informative for constructing exchange rate forecasts than small transactions. I test this hypothesis by multiplying the trade signs q_{t^*} with the transactional volume at t^* for the signal O_t. However, applying this approach to the total order flow dataset results in a different sign for one day out of 2094 days only. Hence I disregard volume weighing.

3.4 Performance measures

The total return r_t on a long position in the US dollar for the period $t-1$ to t is

$$r_t = s_t - s_{t-1} + i^{US}_{t-1} - i^{EUR}_{t-1} \quad (13)$$

with i_t the cash interest deposit rate. The term $i^{US}_{t-1} - i^{EUR}_{t-1}$ represents the interest rate differential: the difference between interest received from the long position in the dollar and interest paid on the associated short position in the euro over the holding period.

The one-period return, based on the signal z_t generated at the start of the period, is

$$r_{z,t} = z_{t-1} r_t \quad (14)$$

The return $r_{z,t}$ is an excess return. To determine the economic significance of the signals, I observe the first, second, and third moment of returns. The first moment is simply the **mean annualized return**. I conduct a t-test to evaluate whether the reported mean return is significantly different from zero.

The second moment of returns is commonly known as the **Sharpe ratio** (SR). The Sharpe ratio is the mean excess return divided by the standard deviation of the excess return. The Sharpe penalizes returns based on volatility: the higher the volatility accompanying a return, the lower the Sharpe.

The third moment of returns, or **skewness**, is well-known in statistics but less common as an economic performance measure. But for example Berge et al. (2010) use skewness to observe whether returns are evenly distributed around the mean. Negative skewness implies a long left tail of the return distribution. Thus a model with a high negative skew is subject to a risk of pronounced period crashes that could ruin leveraged trading accounts.

Transaction costs dent gross returns, sometimes to the point that positive gross returns become negative. Transaction costs include broker fees, clearing fees, regulatory fees, and slippage. Slippage is the cost of not being able to buy a full position at mid price. Usually, brokers state commissions in basis points over the total trade value, with a moderate minimum fee per order[2]. For professional traders, the total sum of traded order sizes is therefore more relevant than the number of transactions.

I define **turnover** as the number of times a full position is traded. For example: take the following range r_1 of consecutive signals: [0.75, 0.5, -0.8, 0.6]. Range r_1 requires four transactions with a turnover of (0.75 + 0.25 + 1.3 + 1.4 =) 3.7. Now suppose r_2 is [0.3, 0.5, 0.2, -0.1]. Again there are four transactions, but the turnover is only 1.1. In practice, although the number of transactions is equal, the signals of r_1 would be over three times as expensive to trade as r_2.

I use turnover to determine the **break-even transaction costs**. I define break-even costs as the strategy's average annual return divided by the average annual turnover. High break-even costs are a favorable indication that a model generates returns with low turnover.

[2] See for example: `http://www.interactivebrokers.com/en/p.php?f=commission&p=fx`

Furthermore, Jordà and Taylor (2011) introduce direction-based performance measures based on Receiver Operating Statistics (ROC) as used in signal detection theory. Denote $d_t \in \{-1,+1\}$ the correct ex post direction of a trade, with $d_t = 1$ for long and $d_t = -1$ for short. Let c be a threshold parameter for an investor applied to a trade signal $x_t \in (-\infty, \infty)$. The value of c depends on an investor's preference and attitude towards risk, as well as the distribution of returns. Fundamental to the ROC method are the four different outcomes for the signal x_t:

1. True Positive (TP) = $P(x_t \geq c | d_t = +1)$
2. True Negative (TP) = $P(x_t < c | d_t = -1)$
3. False Positive (FP) = $P(x_t \geq c | d_t = -1)$
4. False Negative (FN) = $P(x_t < c | d_t = +1)$

The ROC curve visually depicts false positives versus true positives for the total spectrum of $-\infty < c < \infty$. Jordà and Taylor (2011) introduce an economic variant of this curve: the Correct Classification (CC) frontier. Following ROC theory, they develop the area under the Correct Classification frontier, or AUC (area under curve), as a measure of the directional ability of the trading strategy.

I modify Jordà and Taylor (2011)'s general approach by assuming that currency investors generally do not distinguish between true positives and true negatives on the one hand, and false positives and false negatives on the other. Suppose a trader is long in the US dollar, with an accompanying short in the euro, and the US dollar/euro rises (true positive). This is equally valuable as when the trader is long in the euro, with an accompanying short in the US dollar, when the US dollar/euro falls (true negative).

From ROC theory, I derive the rate of correct directional trade forecasts or **accuracy** ACC as:

$$ACC = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{15}$$

I am interested in the correct ex post trade directions with $r_{z,t} = c = 0$. Then, accuracy is the number of ex post correct trades divided by the total number of trades. An accuracy of 0.50 indicates random or 'coin flip' directional performance, while 1.00 indicates perfect directional performance. Using a t-test, I evaluate whether the accuracy is significantly different from 0.50. The accuracy is useful for example to determine whether a model is suitable to trade using binary options strategies.

Suppose a trader does not use binary options, but plainly goes long or short in the US dollar/euro. Even if a model has a high accuracy, it could turn out as a poor investment strategy in case the many winning trades are worth pennies and the few losing ones worth dollars. Similar to Jordà and Taylor (2012), I use the **gain-loss ratio** that Bernardo and Ledoit (2000) introduce. The ex post gain-loss ratio is the model's mean of positive daily excess returns divided by the mean of negative daily excess returns. Naturally, models with a high gain-loss ratio are desirable to a trader or investor.

4 Data

4.1 Overview

Table 1 provides an overview of the data and available time periods. The daily interest rate is the quoted annualized rate divided by 360 days. One calendar year contains on average 251 trading days.

Data	Database	Ticker	Source	Frequency	Start	End
EURUSD	FRED	DEXUSEU	St. Louis Fed	daily	1/4/1999	5/11/2012
EURUSD		EC	CME	tick	6/30/2003	3/30/2010
Eonia	DS	EUEONIA	ECB / EBF	daily	1/4/1999	6/22/2012
Fed funds	DS	FRFEDFD	NY Fed	daily	7/1/1954	6/22/2012
CPI EU	DS	EMCONPRCF	Eurostat	monthly	2/15/1960	5/15/2012
CPI US	DS	USCONPRCE	US BLS	monthly	2/15/1950	5/15/2012
Industr. prod. EU	OECD			monthly	2/15/1985	4/15/2012
Industr. prod. US	OECD			monthly	2/15/1960	4/15/2012

Table 1: *Data overview: FRED is the Federal Reserve Economic Database. DS is Datastream. OECD is the OECD Real-Time Data and Revisions Database.*

EURUSD The price of one euro in dollar terms. I use daily data (noon buying rate New York time) and tick data for order flow (including Chicago transaction times). I take the log of the inverse EURUSD price to obtain s_t. I ensure that the daily order flow signal O_t contains all transactions from 11am Chicago at $t-1$ up to 11 am Chicago at t to correspond with the time of the daily spot price.

Eonia Euro OverNight Index Average. Eonia is the effective short term reference interest rate for the euro. It is the weighted average of all overnight unsecured lending transactions undertaken in the euro interbank market.

Fed Funds Federal funds rate. The interest rate at which depository institutions lend balances at the Federal Reserve to other depository institutions overnight. The daily effective federal funds rate is a weighted average of rates on trades through New York brokers.

CPI I measure annual inflation as the 12-month growth rate of the consumer price index (CPI). For the EU, I use the Harmonized Index of Consumer Prices. 'Harmonized' indicates that every member state calculates the inflation figure as outlined under Article 121 of the Treaty of Amsterdam.

Industrial production Molodtsova and Papell (2008) use monthly industrial production to determine the output gap in the Taylor model, as GDP data is only available quarterly. I take the same approach. The industrial production data is seasonally adjusted for both eurozone and US.

Both CPI[3] and industrial production[4] data are available around mid-month for the preceding month. I assume that on every 15th (or next business date in case exchange is closed) of the month, inflation and industrial production figures for the previous month are available.

CPI series are final when issued[5]. But industrial production figures, especially when seasonally adjusted, are generally prone to a series of revisions after initial release, see Swanson and van Dijk (2006). Therefore, I use real-time data from the OECD that, at a certain date, only includes revisions that were known at that point in time. So every for every new month, I essentially use a full new vector of industrial production figures instead of merely adding a data point.

4.2 Sample period

I require overlap for the data of all forecasting models for the out-of-sample period. The order flow tick data restricts the end of the out-of-sample period to March 30, 2010. The last full month of signals from the Taylor model within the order flow tick data range ends March 16, 2010, which defines the end of the out-of-sample period.

Taylor and channel rules need a sufficient in-sample period to estimate and evaluate coefficients. For Taylor rules, Molodtsova and Papell (2008) use rolling regressions and ten years of monthly data per regression, which implies 120 observations. In their paper on the euro/US dollar, Molodtsova et al. (2008) use 34 quarterly observations per regression. They use an in-sample period from 1991 to 1999 and an out-of-sample period from 2000 to 2007.

I use monthly instead of quarterly economic data. Furthermore, I have more than two years of extra data available (restricted by order flow) compared to Molodtsova et al. (2008). Therefore, I start the in-sample period at the time the euro comes into existence: January 4, 1999. I start the out-of-sample period January 18, 2005. This is the first business day after the 15th, so that the out-of-sample period starts with a new Taylor rule signal. Though not strictly necessary, I start and end mid-month to mirror Molodtsova et al. (2008)'s monthly Taylor rule predictions as closely as possible. This setting provides an out-of-sample length of over five years and six years in-sample.

I require a rolling window with sufficient observations for the out-of-sample Taylor signal. At the same time, I need to obtain in-sample results to evaluate the thirty-two Taylor model specifications. Hence I split the six year in-sample period in half. I use the first three years of data for the first regression. I prefer to use a rolling window longer than three years out-of-sample. Therefore, I use an expanding time window to estimate the monthly regression parameters and signals for the remaining three in-sample years. At the start of the out-of-sample period, I roll over to a fixed six year rolling window. This provides a minimum of 36 monthly observations in-sample and 72 monthly observations per regression out-of-sample.

[3] US CPI release dates: `http://www.bls.gov/schedule/archives/cpi_nr.htm`
[4] US industrial production release dates: `http://www.federalreserve.gov/releases/g17/release_dates.htm`
[5] Corrections to the CPI: p.5 `http://www.bls.gov/opub/hom/pdf/homch17.pdf`

Channel rules need a period of at least L trading days after January 4, 1999 to generate the first signal. Thus the in-sample period decreases with one day for every one-step increase in L. Still, in-sample length is larger than five years for $L = 200$.

I do not need to estimate coefficients for the order flow in-sample period. Also, I prefer an as large as possible out-of-sample period to evaluate all three sources of information. I therefore limit the in-sample period for order flow to just over one and a half years. It spans a bit over a fifth of the total period for which flow data is available. I nevertheless evaluate in-sample flow results.

4.3 Summary statistics

Table 2 reports the summary statistics for the returns of a long position in the US dollar and a short position in the euro.

US dollar return	Start	End	Mean	Spot	Rate diff.	St.dev.	Skewness	Kurtosis
Full Sample	1/4/1999	3/15/2010	-1.15	-1.29	0.14	10.30	-0.18	5.46
In-sample	1/4/1999	1/14/2005	-1.67	-1.72	0.05	10.20	-0.01	3.62
Out-of-sample	1/18/2005	3/15/2010	-0.54	-0.79	0.25	10.42	-0.37	7.43

Table 2: *Summary statistics of US dollar/euro returns. The annualized mean return is the sum of contributions from the annualized spot exchange rate return and the annualized interest rate differential.*

Annualized US dollar spot returns are negative for both the in-sample and out-of-sample period. The positive US dollar interest rate differential mitigates these negative returns only slightly. The kurtosis is much higher than three over the full sample. This indicates a high peak and fat tails in the return distributions. Unreported Jarque-Bera test results confirm that none of the returns are from a normal distribution.

The 2008 global financial crisis and its aftermath define part of the out-of-sample period. This event explains the higher kurtosis and skewness of the out-of-sample period versus the in-sample period.

5 Results

In this chapter, I first evaluate in-sample results. I select a range of favorable model specifications to use out-of-sample for Taylor and channel rules. I arrive at one composite trading signal for both information sources. For Taylor rules, I compose the signal from six of the thirty-two Taylor specifications. For channel rules, L is the number of days in the trading range. I select the mean of $L = 20$ up to and including $L = 79$ as composite channel signal.

Next, I evaluate out-of-sample results. The composite Taylor model generates a significant annual return of 5.38% with a Sharpe ratio of 0.74 and t-value of 1.68. Break-even costs are 1.90% and imply that net returns will not differ much from gross returns. Accuracy is 0.5758 and highly significant with a t-value of 5.53.

The composite channel model yields a 3.52% return with a Sharpe of 0.43. But with a t-value of 0.98, the channel return is not significant at the 5% level. The flow model returns an insignificant, negative 2.82% per year.

Combining the signals provides diversification benefits. But the results of the channel and order flow signals are too weak compared to Taylor results to improve the economic Taylor rule measures.

5.1 In-sample

5.1.1 Fundamental model: Taylor rules

Table 3 displays in-sample results for Taylor models. We require a t-value 't-ret' ≥ 1.65 to be able to state that the chance is larger than 95% that the return and Sharpe ratio are significantly different from zero. Similarly, we require a t-value 't-ACC' ≥ 1.65 to be able to state that the accuracy ACC is significantly different from 0.50 at the 5% level.

For eighteen out of thirty-two Taylor model specifications, mean returns are significant at the 5% level and range from 10.07% to 14.70% per year. Furthermore, four specifications yield a 9.29% return that borders on significance with a t-value of 1.60. Accompanying Sharpe ratios for the eighteen significant models range from 1.74 to 2.54. Generally, the higher the return, the more negative skewness we observe. This is not favorable, but skewness for the significant returns varies modestly between -0.18 to -0.27.

Turnover for the eighteen best models ranges from 0.34 to 3.08 per year and is on average lower than for model specifications with insignificant returns. This results in high accompanying break-even transaction costs that range from 3.27% to close to 43% per year. The t-value for daily accuracy is above one for seven model specifications, but none are significant at the 5% level. Accompanying gain/loss ratios for these seven models range from 1.12 to 1.14.

	Taylor model	y	return (ann %)	t-ret	SR	skew	turnov (ann)	BETC (%)	ACC (day)	t-ACC	g/l (day)
1	he-s-c	lin	6.41	1.11	0.65	-0.16	3.77	1.70	0.5007	0.04	1.11
	he-s-c	qua	6.41	1.11	0.65	-0.16	3.77	1.70	0.5007	0.04	1.11
	he-s-c	HP	11.77	2.03	1.19	-0.22	2.40	4.91	0.5171	0.92	1.13
	he-s-c	HP-R	11.77	2.03	1.19	-0.22	2.40	4.91	0.5171	0.92	1.13
2	ho-s-c	lin	9.29	1.60	0.94	-0.18	3.77	2.47	0.5034	0.18	1.15
	ho-s-c	qua	9.29	1.60	0.94	-0.18	3.77	2.47	0.5034	0.18	1.15
	ho-s-c	HP	11.77	2.03	1.19	-0.22	2.40	4.91	0.5171	0.92	1.13
	ho-s-c	HP-R	13.92	2.41	1.41	-0.26	1.03	13.55	0.5252	1.37	1.14
3	he-ns-c	lin	10.07	1.74	1.02	-0.18	3.08	3.27	0.5075	0.41	1.14
	he-ns-c	qua	10.07	1.74	1.02	-0.18	3.08	3.27	0.5075	0.41	1.14
	he-ns-c	HP	11.77	2.03	1.19	-0.22	2.40	4.91	0.5171	0.92	1.13
	he-ns-c	HP-R	11.77	2.03	1.19	-0.22	2.40	4.91	0.5171	0.92	1.13
4	ho-ns-c	lin	1.58	0.27	0.16	0.01	5.14	0.31	0.4966	-0.18	1.04
	ho-ns-c	qua	1.58	0.27	0.16	0.01	5.14	0.31	0.4966	-0.18	1.04
	ho-ns-c	HP	4.68	0.81	0.47	-0.09	6.50	0.72	0.5034	0.18	1.07
	ho-ns-c	HP-R	6.03	1.04	0.61	-0.11	5.14	1.17	0.5020	0.11	1.09
5	he-ns-nc	lin	10.07	1.74	1.02	-0.18	3.08	3.27	0.5075	0.41	1.14
	he-ns-nc	qua	10.07	1.74	1.02	-0.18	3.08	3.27	0.5075	0.41	1.14
	he-ns-nc	HP	12.18	2.11	1.23	-0.23	1.71	7.11	0.5225	1.22	1.12
	he-ns-nc	HP-R	14.70	2.54	1.49	-0.27	0.34	42.93	0.5293	1.59	1.13
6	ho-ns-nc	lin	6.42	1.11	0.65	-0.13	5.13	1.25	0.5102	0.55	1.07
	ho-ns-nc	qua	6.42	1.11	0.65	-0.13	5.13	1.25	0.5102	0.55	1.07
	ho-ns-nc	HP	-0.19	-0.03	-0.02	0.16	2.40	-0.08	0.4980	-0.11	1.01
	ho-ns-nc	HP-R	-0.19	-0.03	-0.02	0.16	2.40	-0.08	0.4980	-0.11	1.01
7	he-s-nc	lin	9.29	1.60	0.94	-0.18	3.77	2.47	0.5034	0.18	1.15
	he-s-nc	qua	9.29	1.60	0.94	-0.18	3.77	2.47	0.5034	0.18	1.15
	he-s-nc	HP	14.70	2.54	1.49	-0.27	0.34	42.93	0.5293	1.59	1.13
	he-s-nc	HP-R	14.70	2.54	1.49	-0.27	0.34	42.93	0.5293	1.59	1.13
8	ho-s-nc	lin	11.87	2.05	1.20	-0.23	1.03	11.56	0.5157	0.85	1.14
	ho-s-nc	qua	11.87	2.05	1.20	-0.23	1.03	11.56	0.5157	0.85	1.14
	ho-s-nc	HP	14.70	2.54	1.49	-0.27	0.34	42.93	0.5293	1.59	1.13
	ho-s-nc	HP-R	13.92	2.41	1.41	-0.26	1.02	13.55	0.5252	1.37	1.14
A	HP(-R)		13.14	2.32	1.36	-0.25	1.14	11.51	0.5252	1.37	1.13
B	Best (HP-R)		14.44	2.51	1.47	-0.27	0.34	42.17	0.5293	1.59	1.13

Table 3: *In-sample Taylor rules results. The symmetric model specification is either heterogeneous ('he') or homogeneous ('ho'), and may include smoothing ('s') and a constant ('c'). The output gap y is the deviation of industrial production from a linear, a quadratic, HP-filtered trend with either* $\lambda = 14,400$ *('HP') or* $\lambda = 129,600$ *('HP-R'). Annualized mean returns include t-value, Sharpe ratio, and skewness. 'BETC' are break-even transaction costs. 'ACC' is accuracy of direction, while 'g/l' is the gain/loss ratio. I refer to the text for details on model combinations 'A' and 'B'.*

The homogeneous model specifications without smoothing, models ‘4’ (with a constant) and ‘6’ (without a constant) in the table, generate insignificant results for all output gap specifications. These Taylor rule specifications might not be accurate enough to model both central bank’s interest setting policies. Therefore, I disregard model specifications 4 and 6 for the composite Taylor signal.

I evaluate the output gap specifications for the remaining six models. Both HP filter settings outperform linear and quadratic output gap specifications for all six models significantly, with all t-values above two. The Ravn and Uhlig (2002) value for smoothing parameter λ provides better results than Hodrick and Prescott (1997)’s value for two out of six models, worse for one model, and identical for three models.

Cross-correlations (not displayed in a table) for both HP specification returns across the six models range from 0.88 to 1.00, with an average of 0.96. This indicates that trading strategies hardly benefit from diversification if I combine all HP and HP-R signals for models 1 to 3, 5, 7, and 8. The results for this ‘HP(-R) model A’ combination at bottom of table 3 show that there is indeed next to no improvement in performance measures compared to the best models.

Cross-correlations for the four top model specifications with an annual 14.70% return are one. That means that combining the signals yields exactly the same results. The same holds for the two 13.92% annual returns, which have a 0.98 correlation with the 14.70% returns. However, for a longer time frame signals from the 14.70% and 13.92% returns might generate different returns at some point. We don’t know the underlying process and ‘real’ model specification. It therefore makes sense to combine these six best signals in combination ‘B’ instead of just picking one 14.70% return model, even as the results in the table for ‘B’ show no improvement over the 14.70% return models.

So for Taylor rules, I select the composite out of sample Taylor signal ‘B’ as the mean of the:

1. symmetric, homogeneous model with smoothing and a constant, with y_t = HP-R;
2. symmetric, heterogeneous model without smoothing and without a constant, with y_t = HP-R;
3. symmetric, heterogeneous model with smoothing and without a constant, with y_t = HP;
4. symmetric, heterogeneous model with smoothing and without a constant, with y_t = HP-R;
5. symmetric, homogeneous model with smoothing and without a constant, with y_t = HP;
6. symmetric, homogeneous model with smoothing and without a constant, with y_t = HP-R.

5.1.2 Technical model: channel rules

For the in-sample period of channel rules, I vary the number of trading days in the channel range L from 1 day to 200 days in steps of one day to evaluate performance .

Figure 1 displays results for the full range of L. For $L = 1$ to $L = 20$, break-even costs are practically zero as turnover is very high. Performance is instable. From $L = 20$ onwards, performance stabilizes. Turnover gradually declines, which improves break-even costs. As L passes 80 days, performance diminishes to recover near $L = 135$. As L increases, we observe jumps in the performance measures. The rules generate less and less of the same trades. All measures flatten out after $L = 150$.

The returns and break-even costs from $L = 135$ onwards are significant at the 5% level, with t-values above 1.70. But turnover is just 0.6 per year for $L \geq 143$. The performance is based on a few trades. Also, rules with high L capture very long term moves, while users of technical analysis focus mainly on short to mid-term forecasts. Therefore, I concentrate on the $L = 20$ to $L = 80$ range.

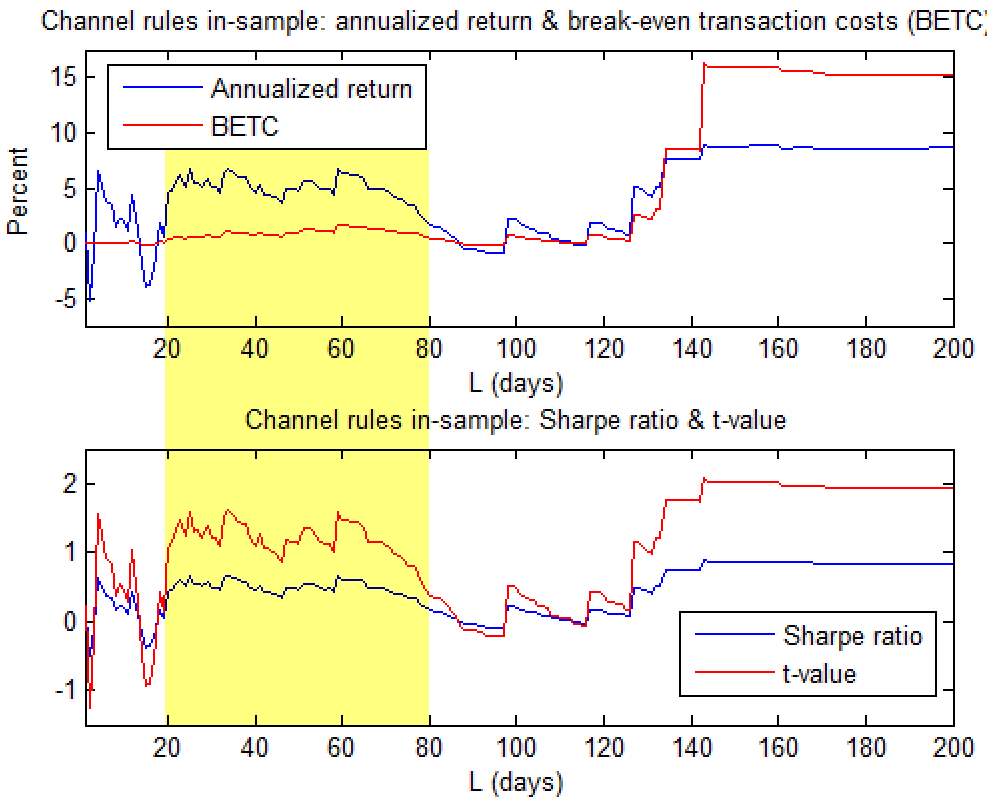

Figure 1: *In-sample channel rules results. L defines the number of trading days in the channel range. I highlight L = 20 to L = 80 in yellow.*

Table 4 displays the results for $L = 15$ up to and including $L = 84$, averaged per basket of five L values. For $L = 15$ up to $L = 19$, the annualized mean return is -1.52%. For $L = 20$ to $L = 69$, annual returns are fairly stable around 5.3%. But even for the highest t-value of 1.61 for $L = 34$, the returns are just short of being signicantly different from zero at the 5% level. Nevertheless, all average t-values are above 1. Turnover falls from 11.7 to 4.0 as L increases. Break-even costs rise steadily to 1.28%, before declining after $L = 70$.

For $L = 75$ up to and including $L = 79$, the return of 2.98% with break-even costs of 0.74% is the last favorable basket. Based on these results, I select the mean of all channel signals for $L = 20$ up to and including $L = 79$, in steps one day, for the composite channel signal.

L (days)	**return (ann %)**	***t*-ret**	**SR**	**skew**	**turnover (ann)**	**BETC (%)**	**ACC (day)**	***t*-ACC**	**g/l (day)**
15 - 19	-1.52	-0.36	-0.15	-0.16	19.01	-0.06	0.4931	-0.53	1.00
20 - 24	5.20	1.24	0.51	-0.19	11.70	0.45	0.5102	0.79	1.05
25 - 29	5.72	1.36	0.56	-0.22	9.72	0.59	0.5150	1.16	1.03
30 - 34	5.55	1.32	0.54	-0.21	8.14	0.71	0.5164	1.27	1.02
35 - 39	5.88	1.40	0.58	-0.21	6.61	0.89	0.5164	1.26	1.03
40 - 44	4.58	1.09	0.45	-0.23	6.08	0.76	0.5179	1.37	1.01
45 - 49	4.49	1.06	0.44	-0.22	5.56	0.82	0.5201	1.54	1.01
50 - 54	5.39	1.27	0.53	-0.23	4.76	1.14	0.5222	1.70	1.00
55 - 59	5.14	1.21	0.50	-0.23	4.50	1.16	0.5217	1.66	1.00
60 - 64	6.16	1.45	0.60	-0.27	3.96	1.55	0.5260	1.98	1.00
65 - 69	5.10	1.20	0.50	-0.27	3.98	1.28	0.5232	1.77	1.00
70 - 74	4.23	0.99	0.41	-0.27	3.99	1.06	0.5184	1.40	1.00
75 - 79	2.98	0.69	0.29	-0.26	4.00	0.74	0.5108	0.82	1.01
80 - 84	1.39	0.32	0.14	-0.27	4.02	0.35	0.5060	0.45	1.01

Table 4: *In-sample channel rules results. L defines the number of trading days in the channel range. Annualized mean returns include t-value, Sharpe ratio, and skewness. 'BETC' are break-even transaction costs. 'ACC' is accuracy of direction, while 'g/l' is the gain/loss ratio.*

5.1.3 Order flow

The in-sample period for order flow is just one and a half years as I do not need to estimate coefficients. Table 5 reports results. The yearly return is 3.26%, but insignificant with a t-value of 0.39. Turnover is very high at nearly 250, resulting in break-even costs close to zero. Accuracy is nearly 0.51, but also insignificant. Only the skew measure of -0.09 compares favorable versus the Taylor and channel results.

return (ann %)	***t*-ret**	**SR**	**skew**	**turnover (ann)**	**BETC (%)**	**ACC (day)**	***t*-ACC**	**g/l (day)**
3.26	0.39	0.31	-0.09	249.70	0.01	0.5090	0.41	1.01

Table 5: *In-sample order flow in-sample results. Annualized mean returns include t-value, Sharpe ratio, and skewness. 'BETC' are break-even transaction costs. 'ACC' is accuracy of direction, while 'g/l' is the gain/loss ratio.*

5.2 Out of sample

5.2.1 Taylor rules, channel rules, and order flow

Table 6 shows that the **composite Taylor model** generates a significant positive annual return of 5.38% with a Sharpe ratio of 0.74 out-of-sample. With an annual turnover of 2.83, break-even transaction costs are 1.90% or 190 basis points. This compares to a real-world brokerage fee of 0.1 to 0.2 basis points at a large online broker at the time of writing[6]. The spread in the US dollar/euro is generally one 'pip', or one basis point with the US dollar/euro at parity. Thus Taylor out of sample net returns differ only marginally from gross returns.

Accuracy is high and significant at 0.5758 with a t-value 5.53. However, the gain/loss ratio of 1.01 indicates that the return of the average winning trade is nearly equal the return of the average losing trade.

signal	return (ann %)	t-ret	SR	skew	turnover (ann)	BETC (%)	ACC (day)	t-ACC	g/l (day)
Taylor (T)	5.38	1.68	0.74	-0.34	2.83	1.90	0.5758	5.53	1.01
channel (C)	3.52	0.98	0.43	-0.22	6.45	0.54	0.4935	-0.47	1.13
flow (F)	-2.82	-0.62	-0.27	-0.09	249.65	-0.01	0.4842	-1.14	1.02
T&C	4.45	1.57	0.69	-0.29	4.61	0.96	0.5466	3.37	1.03
T&C&F	2.02	0.81	0.36	-0.68	85.94	0.02	0.5381	2.75	1.02

Table 6: *Out-of-sample performance of the channel, Taylor, and order flow composite signals. Annualized mean returns include t-value, Sharpe ratio, and skewness. 'BETC' are break-even transaction costs. 'ACC' is accuracy of direction, while 'g/l' is the gain/loss ratio.*

The **composite channel model** yields a 3.52% return per year with a Sharpe of 0.43, but the return is insignificant at the 5% level. Skewness is -0.22, a bit more beneficial than the -0.34 for the composite Taylor model. The channel model turnover is 6.45, more than double the Taylor turnover. This results in break-even costs of 0.54%. Direction accuracy is not statistically different from a coin flip, but the average winning trade return is 1.13 the size of the average losing trade.

The **order flow model** also yields an insignificant return of minus 2.82% per year with a Sharpe of -0.27. Turnover per year is 249.65, practically equal to in-sample order flow turnover. This again results in break-even costs close to zero. Accuracy is 0.4842 with a t-value below minus one.

I evaluate the stability of the composite signal returns in figure 2. I plot the out-of-sample cumulative total returns of the US dollar/euro as well as composite Taylor, channel, and order flow signals. I highlight a fast rise of US dollar price in the second half of 2008, due to a deepening of the global financial crisis, in yellow.

[6] Interactive Brokers Commissions - Forex: `http://www.interactivebrokers.com/en/p.php?f=commission&p=fx`

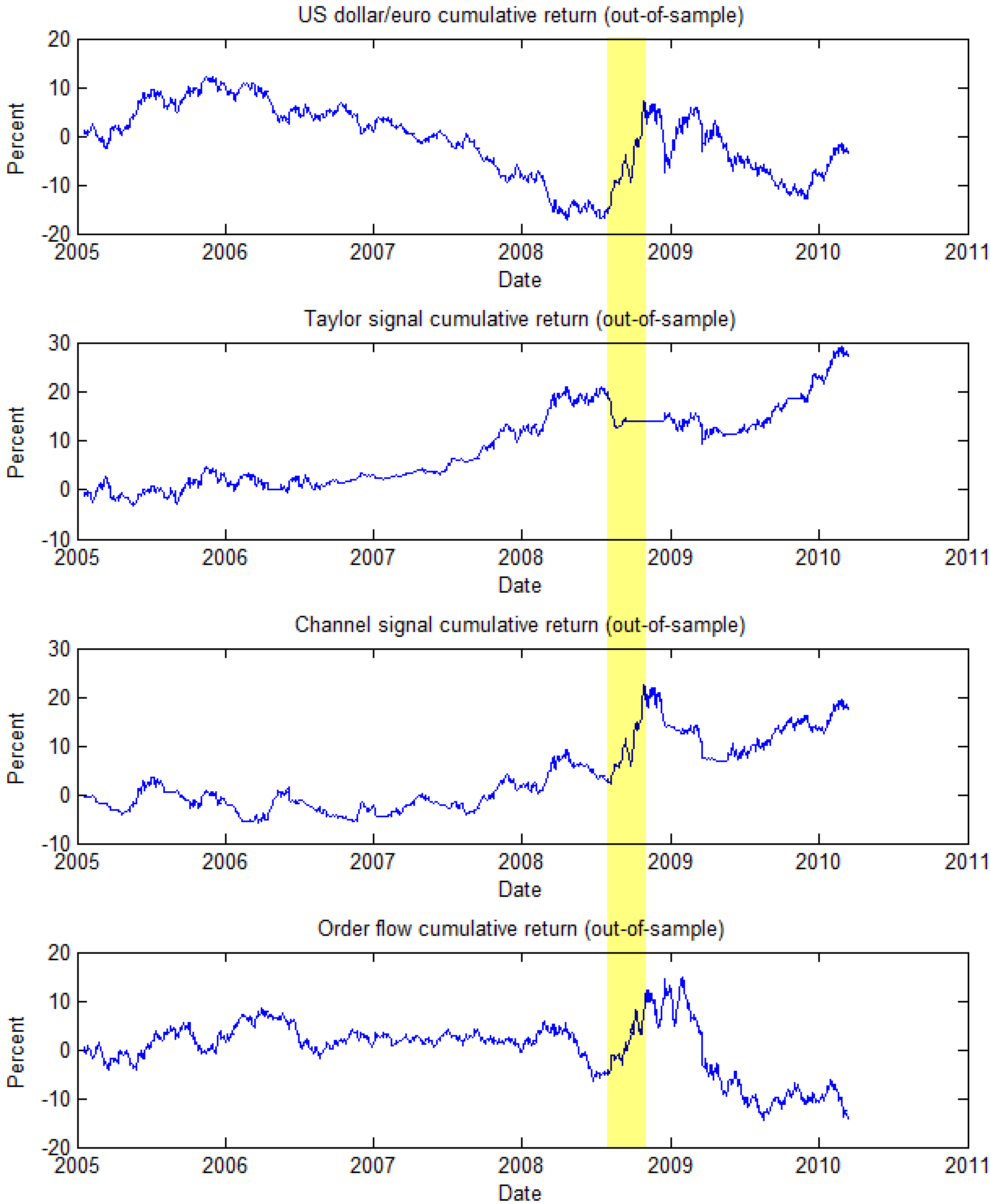

Figure 2: *Out of sample cumulative returns. I highlight the deepening of the global financial crisis in yellow.*

The composite Taylor signal does not profit from this deepening of the crisis, but performance is relatively stable over full out-of-sample period. Declines from peak cumulative returns, or ‘drawdowns’, are limited. Mixed signals for the underlying Taylor signals result in low to zero exposure during the volatile period in the crisis.

From 2005 to 2007, the cumulative composite channel signal return hovers around the zero mark. But returns improve in 2008 and accelerate during the deepening of the crisis, due to high US dollar/euro price momentum. The following sideways movement gives back a large part of the profits, but returns improve into the end of the sample. Overall, the channel signal is not as stable as the Taylor signal. Its performance depends for a large part on the strong US dollar run in the second half of 2008.

The latter is also true for order flow. But overall, the cumulative composite flow return never manages to distance itself from the zero mark. CME US dollar/euro futures volume rises over the out of sample period. We might expect that 2009, the full year in the sample that contains most flow data, provides most information. On the contrary, this year delivers the largest negative yearly return. Returns do not improve into 2010.

5.2.2 Combining fundamental, technical, and order flow information

Table 7 displays cross-correlations for the three composite out of sample signals. Correlation of the Taylor (T) and channel (C) signals is 0.39. Combining these signals should therefore lower volatility. Indeed, the Sharpe ratio of the combined ‘T&C’ signal in table 6 is 0.69, close to the 0.74 Taylor Sharpe, but a significant improvement over the 0.43 Channel Sharpe. Still, the single composite Taylor signal outperforms the combined T&C signal. Only skewness and the gain/loss ratio benefit somewhat from combining fundamental Taylor with technical channel signals.

	Taylor	channel	flow
Taylor	1.00		
channel	0.39	1.00	
flow	0.16	-0.05	1.00

Table 7: *Out-of-sample cross-correlations of the channel, Taylor, and order flow composite signal returns.*

The composite order flow signal is practically uncorrelated with both Taylor and channel signals with correlations of 0.16 and -0.05. But naturally, given the poor performance of the flow signal, adding order flow in the the combined ‘T&C&F’ signal in table 6 does not improve the economic performance of neither the single Taylor signal, nor the single channel signal, nor the combined T&C signal.

Still, the Sharpe of the T&C&F return is 0.36 with a t-value of 0.81. And accuracy is significantly better than random with a t-value of 2.75, despite that two-thirds of the T&C&F signal consists of signals that display insignificant directional performance.

6 Conclusion

Academic research extensively analyzes the forecasting abilities of macroeconomic fundamentals and technical trading rules, but often in isolation and using statistical measures. However, surveys indicate that currency market participants combine fundamentals and technicals to arrive at investment decisions. More recently, order flow gains interest as a potential third predictor of short term exchange rate movements.

In this report, I analyze the economic value of combining fundamental, technical, and order flow information to forecast US dollar/euro returns. Molodtsova and Papell (2008) and Molodtsova et al. (2008) find positive results for exchange rate forecasting using Taylor rules. I confirm these results for the US dollar/euro from January 2002 to March 2010. I also find that the forecasts are profitable and relatively stable from an economic point of view. Furthermore, I show that choosing a different smoothing parameter for the HP filter in the output gap specification following Ravn and Uhlig (2002) can be beneficial.

However, combining the Taylor results with channel rules and order flow does not improve economic results. Taylor rules yield a significant 5.38% out-of-sample, from January 2005 to March 2010. While channel rules still result in an annual out-of-sample return of 3.51% with a t-value of 0.98, order flow results disappoint with an annual return of minus 2.82% with a t-value of -0.62. Also, it appears that order flow result do not improve over time, even though CME futures volume rises from \$ 40 billion in 2005 to \$ 110 billion in 2010, see BIS (2010).

Taylor rules show that, contrary to Meese and Rogoff (1983a,b), fundamentals can be useful for exchange rate forecasting. Still, other promising fundamental directions are available. The naive carry trade exploits interest rates differentials. These trades blew in up when the global financial crisis hit in 2008. Recent research from Jordà and Taylor (2012) shows that slightly more fundamentals-refined carry trade strategies would have generated strong and sustained profits during the crisis.

Other methods to sign trades in the absence of bid-ask quotes might improve order flow results. Wel et al. (2009) develop a new likelihood-based approach. This state space approach with regime switching is more than ten times faster than the MCMC simulation approach by Hasbrouck (2004).

Also, other methods to combine forecasts might lead to improvements. Kim and Swanson (2011) assess the predictive accuracy of a large group of forecast combination techniques. They confirm that shrinkage models, such as Bayesian Model Averaging (BMA), outperform. Furthermore, diversification of combined promising fundamental models, perhaps with technical and order flow information, across a portfolio of multiple currencies, might reduce the variance and skewness of currency trading strategies.

References

Berge, T., Jordà, O., and Taylor, A. (2010). Currency carry trades. *NBER Working Paper 16491*.

Bernardo, A. and Ledoit, O. (2000). Gain, loss, and asset pricing. *Journal of Political Economy*, 108(1):144–172.

BIS (2010). Triennial central bank survey: Report on global foreign exchange market activity in 2010. `http://www.bis.org/publ/rpfxf10t.pdf`.

Cabrera, J., Wang, T., and Yang, J. (2009). Do futures lead price discovery in electronic foreign exchange markets? *Journal of Futures Markets*, 29:137–156.

Chen, S.-S. and Chou, Y.-H. (2010). Exchange rates and fundamentals: Evidence from long-horizon regression tests. *Oxford Bulletin of Economics and Statistics*, 72:63–88.

Cheung, Y.-W., Chinn, M., and Pascual, A. (2005). Empirical exchange rate models of the nineties: Are any fit to survive? *Journal of International Money and Finance*, 24:1150–1175.

Clausen, J. and Meier, C.-P. (2005). Did the bundesbank follow a taylor rule? an analysis based on real-time data. *Swiss Journal of Economics and Statistics*, 127:213–246.

de Zwart, G., Markwat, T., Swinkels, L., and van Dijk, D. (2006). The economic value of fundamental and technical information in emerging currency markets. *Journal of International Money and Finance*, 28:581–604.

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

Evans, M. and Lyons, R. (2002). Order flow and exchange rate dynamics. *Journal of Political Economy*, 110:170–180.

Evans, M. and Lyons, R. (2005). Meese-rogoff redux: Micro-based exchange-rate forecasting. *American Economic Review*, 95:405–414.

Evans, M. and Lyons, R. (2009). Forecasting exchange rate fundamentals with order flow. *Working Paper*.

Gehrig, T. and Menkhoff, L. (2004). The use of order flow analysis in foreign exchange: exploratory evidence. *Journal of International Money and Finance*, 23:573–594.

Gehrig, T. and Menkhoff, L. (2006). Extended evidence on the use of technical analysis in foreign exchange. *International Journal of Finance and Economics*, 11:327–338.

Gerlach, P. (2001). The global output gap: measurement issues and regional disparities. *BIS Quarterly Review*, June:29–37.

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

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

Hasbrouck, J. (2004). Liquidity in the futures pits: Inferring market dynamics from incomplete data. *Journal of financial and Quantitative Analysis*, 39:305–326.

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

Hodrick, R. and Prescott, E. (1997). Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit, and Banking 29*, pages 1–16.

Jordà, O. and Taylor, A. (2011). Performance evaluation of zero net-investment strategies. *NBER Working Paper 17150*.

Jordà, O. and Taylor, A. (2012). The carry trade and fundamentals: Nothing to fear but feer itself. *Journal of International Economics*.

Kilian, L. (1999). Exchange rates and monetary fundamentals: What do we learn from long-horizon regressions? *Journal of Applied Econometrics*, 14:491–510.

Kim, H. and Swanson, N. (2011). Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence. *Working paper, Rutgers University, Department of Economics*.

King, M. and Rime, D. (2010). The $ 4 trillion question: what explains fx growth since the 2007 survey? *BIS Quarterly Review*, pages 27–42.

Lee, C. and Ready, M. (1991). Inferring trade direction from intraday data. *Journal of Finance*, 46(2):733–746.

Lyons, R. (2001). The microstructure approach to exchange rates. *MIT Press*.

Lyons, R. (2002). New perspective on fx markets: Order-flow analysis. *Journal of International Finance*, 4:303–320.

Mark, N. (1995). Exchange rate and fundamentals: Evidence on long-horizon predictability. *American Economic Review*, 85:201–218.

Marsh, I. and O'Rourke, C. (2005). Customer order flow and exchange rate movements: is there really information content? *Working paper, Cass Business School*.

Meese, R. and Rogoff, K. (1983a). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, 14:3–24.

Meese, R. and Rogoff, K. (1983b). *In: Frankel, J., Galli, G., Giovanni, A. The Microstructure of Foreign Exchange Markets*, chapter The out-of-sample failure of empirical exchange rate models: Sampling error or misspecification?, pages 6–105. University of Chicago for NBER.

Menkhoff, L. and Taylor, M. (2007). The obstinate passion of foreign exchange professionals: Technical analysis. *Journal of Economic Literature*, 45:936–972.

Molodtsova, T., Nikolsko-Rzhevskyy, A., and Papell, D. (2008). Taylor rules and the euro. *MPRA Paper prepared for the European Area Business Cycle Network Conference*, 11348.

Molodtsova, T. and Papell, D. (2008). Out-of-sample exchange rate predictability with taylor rule fundamentals. *Journal of International Economics*, 64:1067–1084.

Neely, C., Weller, P., and Ulrich, J. (2009). The adaptive market hypothesis: Evidence from the foreign exchange market. *Journal of Financial and Quantitative Analysis*, 44:467–488.

Park, C.-H. and Irwin, S. (2007). What do we know about the profitability of technical analysis? *Journal of Economic Surveys*, 21:786–826.

Qi, M. and Wu, Y. (2006). Technical trading-rule profitability, data snooping, and reality check: Evidence from the foreign exchange market. *Journal of Money, Credit, and Banking*, 38:2135–2158.

Ravn, M. and Uhlig, H. (2002). On adjusting the hodrick-prescott filter for the frequency of observations. *The Review of Economics and Statistics*, 84 (2):371–380.

Reitz, S. and Schmidt, M. (2007). End-user order flow and exchange rate dynamics. *Deutsche Bundesbank Discussion Paper*, 5.

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

Rogoff, K. (2009). Exchange rates in the modern floating era: what do we really know? *Review of World Economics*, 145:1–12.

Rosenberg, J. and Traub, L. (2006). Price discovery in the foreign currency futures and spot market. *FRB of New York Staff Report*, 262.

Sager, M. and Taylor, M. (2008). Commercially available order flow data and exchange rate movements: caveat emptor. *Journal of Money, Credit and Banking*, 40:583–625.

Swanson, N. and van Dijk, D. (2006). Are statistical reporting agencies getting it right? data rationality and business cycle asymmetry. *Journal of Business & Economic Statistics*, 24(1).

Taylor, J. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy 39*, pages 195–214.

Taylor, M. and Allen, H. (1992). The use of technical analysis in the foreign exchange market. *Journal of International Money and Finance*, 11:304–314.

Timmermann, A. (2006). Forecast combinations. *In: 'Handbook of Economic Forecasting' by Elliot, G., Granger, C.W.J., Timmermann, A.*, pages 135–196.

Tse, Y., Xiang, J., and Fund, J. (2006). Price discovery in the foreign exchange futures markets. *Journal of Futures Markets*, 26:1131–1143.

Wel, M. v. d., Menkveld, A., and Sarkar, A. (2009). Are market makers uninformed and passive? signing trades in the absence of quotes. *FRB of New York Staff Report*, 395.