Microstructure Order Flow: Statistical and Economic Evaluation of Nonlinear Forecasts

Mario Cerrato*, Hyunsok Kim*[1] and Ronald MacDonald*

University of Glasgow, Department of Economics, Adam Smith building.

December 6, 2010

[1]Corresponding author University of Glasgow, Department of Economics, Adam Smith building. h.kim.3@research.gla.ac.uk

Abstract

   In this paper we propose a novel empirical extension of the standard market microstructure order flow model. The main idea is that heterogeneity of beliefs in the foreign exchange market can cause model instability and such instability has not been fully accounted for in the existing empirical literature. We investigate this issue using two different data sets and focusing on out- of-sample forecasts. Forecasting power is measured using standard statistical tests and, additionally, using an alternative approach based on measuring the economic value of forecasts after building a portfolio of assets. We find there is a substantial economic value on conditioning on the proposed models.

   JEL Classification:F31; F41; G10
   Keywords: microstructure, order flow, forecasting

1   Introduction

There is something of a consensus in the exchange rate literature that macro based models of the exchange rate fail to outperform a simple random walk model in an out-of- sample forecasting context (see, for example, Meese and Rogoff, 2002). Given this, many researchers have turned to a market microstructure approach to provide alternative insights into the forecasting behaviour of exchange rates. For example, Evans and Lyons (2002b), Evans and Lyons (2005b) and Sager and Taylor (2008) use such an approach and provide mixed evidence that microstructure models (i.e. order flow models) can do better than a simple random walk in out of sample forecasts. The main conclusion of Evans and Lyons (2002b) is that order flow is a significant determinant of exchange rates and can be also used to forecast exchange rates out of sample. However, Sager and Taylor (2008) finds little empirical evidence supporting these conclusions after employing interdealer and commercially available order flow data.

A related but slightly different strand of the market microstructure literature investigates the issue of whether the strength of the relationship between order flow and exchange rates is dependent upon prevailing market conditions or the announcement of macroeconomic news. For example, Love and Payne (2003) examines the role of order flow in the transmission of news regarding published macro fundamentals and finds that information that is contemporaneously released to all market participants is partially impounded into prices via the microstructure order flow. However, this is clearly at odds with rational expectations. Bacchetta and Wincoop (2006) and Rime et al. (2010) argue that macroeconomic information impacts on exchange rates both directly, as in a standard macro model, but also indirectly via order flow. Thus, order flow can be viewed as a random variable which maps disperse information in the market in to price discovery. In particular, since the order flow of the FX market consists of different participants, displaying significant heterogeneity in terms of risk-return expectations and informational asymmetries, the customer order flow represents the primary source of private information that is assumed to represent future innovations in fundamental exchange rate determinants.

The above microstructure models provide some useful insights into the foreign exchange market, but there are still several unanswered questions. For example, the success of microstructure models in out of sample forecasts has primarily been achieved when the information is publicly and simultaneously released to all market participants contemporaneously. However, since the information of the state of the economy available at a given point may takes some time before it affects the exchange rate, it is probably preferable to consider a lagged order flow model as in Sager and Taylor (2008). Additionally, since different market participants trade using private as well as public information, expectations about the new equilibrium exchange rate are formed based on a combination of macroeconomic fundamentals and market microstructure variables.

In this paper we try to shed some light on some of the issues raised above. Firstly, we propose various extensions to the existing order flow model which should accommodate model instability. That is, if order flow does reflect heterogeneous beliefs about the current and future state of the economy, and if currency markets do not

2

discover order flow in real time but only through a gradual learning process, the heterogeneity in the market can cause model instability. This important point has been largely neglected in the literature. For example, Rime et al. (2010) employs a Probit model and shows that order flow may be linked to macroeconomic fundamentals both via a direct link, as in classical exchange rate theory, and via order flow, as in the microstructure approach to the foreign exchange market. We attempt to capture this effect using time-varying parameter (TVP), structural change (STR) and smooth transition (STAR) models. This is also in line with Gradojevic and Yang (2006).

Also, and as pointed out by Sarno and Valente (2009), parameter instability caused by instabilities in macro fundamentals, and agents' heterogeneity, or swings in expectations about future values of the exchange rate, make it difficult to select a predictive model. We show that our model specifications can address this issue. In particular, our study suggests the inclusion of microstructure variables and nonlinear models produces out-of-sample forecasts which are superior to those from a random walk model.

Finally, we evaluate our out of sample forecasts using statistical tests, such as the root mean squared forecast error (hereafter RMSFE), and the Diebold-Mariano (hereafter DM) tests, as well as mean-variance analysis as a standard measure of portfolio performance, as in Fleming et al. (2001), Han (2006), Della Corte et al. (2009) and Rime et al. (2010).

The remainder of this paper is organized as follows. In the next section, we provide a brief literature review. Section 3 describes the link between order flow and exchange rates and statistical evaluation methods. The forecasting setup and the investor's asset allocation problem are described in Section 4, and the results on the statistical and economic evaluation of the forecasting models that condition on order flow are reported in Section 5. The final section concludes the paper and recommends further research.

2   A brief review of exchange rate predictability issues from a microstructure perspective.

Microstructure models view order flow as a random variable which maps heterogeneous disperse information into price discovery. Thus, relative to macro based exchange rate models, order flow in the microstructural approach represents the missing link between exchange rates changes and changes in economic conditions. Consider the following (contemporaneous) order flow model,

$$\Delta s_t = \beta_1 \Delta (i_t - i_t^*) + \beta_2 X_t + \varepsilon_t \tag{1}$$

Using the above model, Evans and Lyons (2002b) report significant explanatory power when the mark–dollar and the yen–dollar exchange rates are considered. The empirical analysis of Evans and Lyons (2002a) is extended to an additional seven exchange rates and they report explanatory power ranging from 0.00% to 68%. They also report a high out of sample power of the order flow model when compared to a

simple random walk model. Killeen et al. (2006) also reports significant explanatory power of the order flow model which is consistent with the results of Evans and Lyons (2002b).

Payne and Vitale (2003) points out that the model above is not very relevant in practice as it assumes perfect foresight. Indeed, using central bank order flow for the Swiss franc–dollar rate over the sample period 1986–95, they show that although inter-dealer order flow has a significant contemporaneous correlation with exchange rate returns, its predictive power is minimal. Recently Sager and Taylor (2008) investigates this issue further in a large empirical study. They argue that the announcement of public information is impounded in prices with a delay. Thus, they suggest the following modification of the above model which they call the "publication lag" model:

$$\Delta s_t = \beta_1 \Delta(i_{t-1} - i^*_{t-1}) + \beta_2 X_{t-1} + \varepsilon_t, \tag{2}$$

After undertaking a large empirical analysis, they show that the (lagged) order flow model has very little (in sample) explanatory power and cannot outperform a simple random walk model in forecasting exchange rates at different horizons. Additionally, they show widespread evidence of a Granger-causal relationship that runs from exchange rate returns to customer order flow. This result is consistent with the empirical evidence of Engel and West (2005), which have found some support for the link between fundamentals and exchange rate in the other direction: exchange rates can help forecast the fundamentals.

More recently, Cerrato et al. (2009) uses weekly customer order flow for nine of the most liquid currencies and investigates the in-sample and out-of-sample forecasting power of the order flow models. While empirical results using aggregate data are in line with Sager and Taylor (2008)[1], using disaggregate data seems to increase the in-sample and out-of-sample forecasting power of the order flow model.

A number of researchers (Evans and Lyons (2005a), Evans and Lyons (2008), and Love and Payne (2008)) have provided empirical evidence that macro news triggers trading that reveals dispersed information, which in turn affects currency prices, and a number of papers have sought to clarify the relationship between the release of economic news and order flow. For example, the hypothesis suggested by Rime et al. (2010) is that the heterogenous interpretation of macroeconomic news may lead market makers to make inferences differently and that the order flow incorporates this information gradually. Based on this observation, Rime et al. (2010) proposes the following direct (3) and indirect (4) specifications:

$$\Delta s_t = \alpha_1 + \sum_{n=1}^{N} \beta_n \mathbf{NEWS}_{n,t} + u_t, \tag{3}$$

and

$$\Delta s_t = \alpha_1 + \sum_{n=1}^{N} \beta_n \mathbf{NEWS}_{n,t} + \gamma_1 \Delta X_t + u_t, \tag{4}$$

---

[1] However, the in-sample results, using the contemporaneous order flow model, strongly support such a model. In effect, with weekly data, the lagged model might be too restrictive.

Both of the above models show evidence that exchange rate fluctuations are linked to macroeconomic fundamentals via the direct link, as in traditional exchange rate theory, and via order flow, as in the microstructure approach to the foreign exchange rate. The equation (3) implies that the heterogeneous interpretation of market information directly affects the asset price if the order flow fully contains macroeconomic news as implied by typical microstructure models. However, as shown in Love and Payne (2008), order flows partly reflect a heterogeneous interpretation of macroeconomic news and (4) specifies the effects between news and order flows. This modeling approach can provide some explanation for the link between macroeconomic fundamentals and exchange rates examined in Bacchetta and Wincoop (2006) and Evans and Lyons (2008). Note that the finding of significant explanatory power for macroeconomic news on the exchange rate does not automatically imply that order flow information is redundant. (e.g. Rime et al. (2010)). The addition of order flow significantly increases the explanatory power of the model. Rime et al. (2010) demonstrates that macroeconomic news can explain exchange rates changes to the same extent that they explain order flow.

Overall, the empirical literature in this area seems to have produced conflicting results and we believe a key reason for this could lie in the way the models are estimated. For example, for the news models mentioned above, news is constructed using monthly macroeconomic data. However, with high frequency data that approach is not feasible and so an alternative specification is required, which can properly capture shifts in expectations. The aim of the present study is to shed some light on these issues and address some problems that in our view have been neglected when modelling exchange rates dynamics. In particular, most of the studies cited above have mainly focused on linear models and a direct relationship between the exchange rate and order flow. We believe these models are very restrictive with high frequency data sets. In this paper we propose a novel structural break model which we believe clarifies the role of heterogeneous information and relaxes the linearity assumption.

2.1 Model instability

3 Empirical models and evaluation

The models introduced in the previous section suggest that shifts in expectations can cause model instability. Very few papers have considered this an important issue (see for example Rime et al. (2010)). However, as we shall discuss in this section, the models proposed in this paper have the additional advantage of being able to accommodate a shift in expectations caused by non-observable fundamentals. We propose three different models which address this important issue and test them in out of sample exercises.

3.1 Time-varying parameter model

The first model we consider is a variation of the standard model considered in the literature. The idea is that if an economic announcement affects order flow, this will

cause a change in the parameters governing the exchange rate forecasts. Thus, we suggest the following time varying parameter model.

In a time-varying parameter model the dynamic for exchange rate returns is driven by the following regression

$$\Delta s_{t+k} = \alpha + \beta_t X_{t-1} + \varepsilon_{t+k}.$$

The parameters of the model are estimated in the usual way, using the first n observations. The estimates are then updated in each subsequent observation, $s_{n+1}, s_{n+2}...s_T$. The main difference with the approach used in the literature is that, this model uses a different recursive filter.[2] That is, once the tth observation becomes available, β_t may be obtained from β_{t-1} without the matrix inversion implied by OLS (ordinary least squares).

3.2 Smooth transition model

The second specification we use is completely new in the literature. We propose a non-linear model where the band of inaction caused by low relative risk aversion, generates slow adjustment to the equilibrium. As shown in Rime et al. (2010)'s investigation, order flows aggregate changes in market expectations with regard to macro fundamentals, and the relation between order flows and expectations is estimated to be significant. We attempt to filter the change in expectations caused by macro news using our transition function.

We employ the smooth transition function, $CMK - STAR$ recently suggested by Cerrato et al. (ming)

$$\Delta s_{t+k} = \alpha + \beta S(\theta) X_{t-1} + \varepsilon_{t+k},$$

where
$$S(\theta) = [1 + \exp\{\gamma_1(X_{t-1} - c_1)\mathbf{I}_t - \gamma_2(X_{t-1} - c_2)(1 - \mathbf{I}_t)\}]^{-1},$$

and θ represents parameter set to be estimated. The function $S(\theta)$ allows for both threshold effects and smooth transition movements of X_{t-1}. In the central regime, when $-c < X_{t-d} < c$, $S(X_{t-d}, \theta) = 0$. In the limiting outer regimes, when $X_{t-d} < -c$ and $c < X_{t-d}$, $S(X_{t-d}, \theta) = 1$. The specification given by $S(\theta)$ allows the transition depending on X_{t-1}. Thus, if the news directly affects order flow and expectations are heterogeneous, the transition depending on the order flow, X_{t-1} should be able to capture this effect. We use the above model in our forecasting exercises.

---

[2] Given the basic setup
$$y_t = X_t \beta_t + \varepsilon_t$$

The relevant formulae are driven by

$$\beta_t = \beta_{t-1} + \left(X'_{t-1}X_{t-1}\right)^{-1} x_t \left(y_t - X_t\beta_{t-1}\right)/f_t$$

where $f_t = 1 + x'_t \left(X'_{t-1}X_{t-1}\right)^{-1} x_t$ and $X_t = (x_1, x_2, ..., x_t)$

3.3 A structural change model

The above models suggest a direct link between macroeconomic news and exchange rates. In contrast to the direct specification, when order flow is partly reflecting macroeconomic news and expectations, we suggest an alternative model. This modeling approach is very convenient when using high frequency data or unobservable fundamentals. The model we consider incorporates structural breaks due to shifts in expectations by allowing a shift in the mean process

$$\Delta s_{t+k} = \alpha_1 + \alpha_2 S(\theta) + \beta X_{t-1} + \varepsilon_{t+k},$$

where

$$S(\theta) = \left[1 + \exp\left\{-\gamma_1^2 \left(t - c_1 T\right)^2\right\}\right] \left[1 - \exp\left\{-\gamma_2^2 \left(t - c_2 T\right)^2\right\}\right] - 1.$$

The transition function $S(\theta)$ traverses the interval $(-1, 1)$ and the timing of the transition is determined by c_i. The speed at which the function moves between -1 and 1 changes with γ. As discussed in Cerrato et al. (2010), this model is able to capture structural changes taking place in different regimes. If $c_1 < c_2$, $0 < S_t(\theta) < 1$ when $t = c_1 T$, and $-1 < S_t(\theta) < 0$ when $t = c_2 T$. In the limiting or no structural change state $S_t(\theta) = 0$, the model collapses to $\Delta s_{t+k} = a_1 + \beta X_{t-1} + \varepsilon_{t+k}$, and is consistent with the linear model proposed in previous studies. On the contrary, when the structural changes take place because of omitted economic fundamentals such as macro news, or a different interpretation of them, the model becomes $\Delta s_{t+k} = a_1 + \alpha_2 S(\cdot) + \beta X_{t-1} + \varepsilon_{t+k}$. The mean process is determined by the value $S_t(\theta)$. Thus, this structural change model might be viewed as a reasonable approximation of model instability caused by omitted variables, when fundamentals have an indirect link to order flows.

3.4 Forecast evaluation

We assess the out of sample forecasts produced by the three models above in different ways. Firstly, we use the root mean squared forecast error (RMSFE):

$$RMSFE = \sqrt{\frac{\varepsilon'_{t+k}\varepsilon_{t+k}}{T}}.$$

Additionally, we also construct a test statistic for comparing the forecasting performance of the models relative to a simple random walk (RW). Given two forecasts, the RW forecast and the forecast provided by the alternative models (hereafter AM), the ratio of RMSFE against RW can be used to evaluate the out of sample forecasts. We also support this test using the Diebold and Mariano (1995) test. This test allows us to compare the forecasting accuracy of two competing models. Defining $d_t = g(\varepsilon_{1,t}) - g(\varepsilon_{2,t})$ where $t = 1, ..., n$, the Diebold-Mariano test statistic is

$$DM = \frac{\bar{d}}{\left[var\left(\bar{d}\right)\right]^{\frac{1}{2}}}$$

where $\bar{d} = n^{-1} \sum_{t=1}^{n} d_t$ and $var\left(\bar{d}\right)$ represents the asymptotic (long-run) variance of $\sqrt{T}\bar{d}$.

Diebold and Mariano (1995) shows that under the null of equal predictive accuracy, $DM \sim N(0,1)$, and we can reject the null of equal predictive accuracy at the 5% level if

$$|DM| > 1.96.$$

We use the Diebold-Mariano test to assess the out of sample forecasts of our models with respect to a simple Random Walk model RW.

4 Economic value of exchange rate predictability

Most of the previous studies have focused on evaluating the statistical performance rather than the economic significance of a nonlinear approach. Here we also examine the latter and specifically examine the economic value of nonlinear models to risk-averse investors. To measure the economic value of the out of sample forecasts, we address the issue of whether our three models can be used practically by assessing the forecasts where a portfolio of assets is rebalanced according to a trading rule at each time t.

4.1 Portfolio weights of a mean-variance framework

In order to measure the economic performance of a portfolio it is standard to use Sharpe ratios. However, as Marquering and Verbeek (2004) and Han (2006) note, Sharpe ratios can underestimate the performance of dynamically managed portfolios. This happens because Sharpe ratios are calculated using the average standard deviation of the realized returns, which overestimates the conditional risk (standard deviation) faced by an investor at each point in time. Consequently, Sharpe ratios cannot properly quantify the economic gains of a dynamic strategy.

As an alternative measure of forecasting performance, we use a mean-variance framework and calculate the performance fee to quantify the economic gain from using the exchange rate models introduced above with respect to a simple random walk model. The framework for our analysis is straightforward. We consider an investor who uses a mean-variance optimization rule to allocate funds across assets. The investor's objective is to maximize the expected return matching a target expected volatility.

Allowing for weekly rebalancing, the solution to the investor's portfolio problem is a dynamic trading strategy that specifies the optimal asset weights. Implementing this strategy requires estimates of both the conditional expected returns and the conditional covariance matrix. If the conditional expected return and covariance are constant, the optimal portfolio weights w will be constant over time. However, when the conditional expected return and covariance are defined as recursive estimates, investors will rebalance their portfolio weights and change strategies. Thus, in terms of one-step ahead forecasts, we treat the expected returns as the conditional mean, $\mu_{t+1|t} = E_t\left[r_{t+1} \mid \mathcal{F}_t\right]$ and let the variation in the

portfolio weights be driven purely by changes in the conditional covariance matrix, $\sum_{t+1|t} = E_t\left[\left(r_{t+1} - \mu_{t+1|t}\right)\left(r_{t+1} - \mu_{t+1|t}\right)' \mid \mathcal{F}_t\right]$ where $\mathcal{F}_t$ represents the current information set.

To maximize the conditional expected return, $\mu_{t+1|t}$ subject to a given level of conditional volatility σ_p^*, investors solve the following problem at time t,

$$\max_{w_t}\left\{\mu_{p,t+1} = w_t'\mu_{t+1|t} + \left(1 - w_t'\mathbf{1}\right)r_f\right\}$$
$$\text{s.t. } \left(\sigma_p^*\right)^2 = w_t'\sum\nolimits_{t+1|t} w_t$$

where $\mu_{p,t+1}$ and σ_p^* denote the conditional mean and variance of the portfolio return, $r_{p,t+1}$ of risky assets. In the present setting, w_t is the portfolio weights on the risky assets, and r_f is the return on the riskless asset. Among the trading strategies such as the minimum variance and maximum return, the above mean-variance analysis solves for the weight that maximizes conditional return where the portfolio variance equal to a fixed target.

After constructing the covariance matrix of the portfolio, we determine the weights by maximizing the conditional mean of the portfolio return. The solution to this problem yields the following risky asset weights,

$$w_t = \frac{\sigma_p^*}{\sqrt{C_t}}\sum\nolimits_{t+1|t}^{-1}\left(\mu_{t+1|t} - \mathbf{1}r_f\right)$$

where $C_t = \left(\mu_{t+1|t} - \mathbf{1}r_f\right)'\sum_{t+1|t}^{-1}\left(\mu_{t+1|t} - \mathbf{1}r_f\right)$. The optimal weights will vary across the models depending on the conditional mean and volatility. That is, the trading strategy identifies the rebalanced portfolio that optimizes maximum conditional expected return subject to the conditional variance-covariance.

In our analysis, the benchmark against which we compare the model specifications is a simple RW. In other words, our objective is to evaluate whether there is any economic value in conditioning on microstructure order flow and non-linear models and, if so, which of the four specifications including RW has superior forecasting power.

4.2 Performance measures under quadratic utility

To measure the performance of a trading strategy, using a generalization of West et al. (1993)'s method, Fleming et al. (2001) suggest comparing the performance of the dynamic strategies to that of the unconditional mean-variance efficient static strategy. The latter is based on the relation between mean-variance analysis and quadratic utility. Using a second-order approximation to the investor's true utility function, the investor's realized utility is defined as

$$U(W_{t+1}) = W_{t+1} - \frac{\lambda}{2}W_{t+1}^2 = W_t R_{p,t+1} - \frac{\lambda}{2}W_t^2 R_{p,t+1}^2,$$

where W_{t+1} is the investor's wealth at $t+1$, $R_{p,t+1}$ is the gross portfolio return, equal to $1 + r_{p,t+1}$ and λ represents absolute risk preference.

In our empirical exercise we fix the value of relative risk aversion (RRA) as

9

$\delta = \frac{\lambda W_t}{1 - \lambda W_t}$. Given the level of initial wealth W_0, the average realized utility is then defined as

$$\bar{U}(\cdot) = W_0 \sum_{t=0}^{T-1} \left\{ R_{p,t+1} - \frac{\delta}{2(1+\delta)} R_{p,t+1}^2 \right\},$$

where δ is constant. The average realized utility ($\bar{U}$) can be used to consistently estimate the expected utility generated at the given level of initial wealth, W_0, and value of relative risk aversion (RRA), δ. If the value of RRA is assumed to be $\delta = \{2,6\}$ and the initial wealth is fixed at $W_0 = 1$, we can standardize the investor problem of maximum conditional expected return and assess the economic value of our FX strategies in the context of asset allocation.

To measure the economic value of our FX strategies, we use the average utility and compute the performance fee as suggested in Fleming et al. (2001). The selected pairs of portfolios, RW against alternatives are evaluated by equating the average utilities. That is, if an investor is indifferent between holding a portfolio where the optimal weights have been computed using a simple RW and an alternative portfolio using a more "sophisticated" approach, then the value of Φ can be interpreted as the performance fee that the investor would be willing to pay to switch from the RW to the alternative model, such as TVP, STAR and STR. The performance fee, Φ, is defined as:

$$\sum_{t=0}^{T-1} \left\{ \left(R_{p,t+1}^{AM} - \Phi\right) - \frac{\delta}{2(1+\delta)} \left(R_{p,t+1}^{AM} - \Phi\right)^2 \right\} = \sum_{t=0}^{T-1} \left\{ R_{p,t+1}^{RW} - \frac{\delta}{2(1+\delta)} \left(R_{p,t+1}^{RW}\right)^2 \right\},$$

where $R_{p,t+1}^{RW}$ is the gross portfolio return obtained using forecasts from the benchmark RW model, and $R_{p,t+1}^{AM}$ is the gross portfolio return constructed using the forecasts from the alternative models. Thus, the utility-based criterion measures how much the investor is willing to pay for conditioning on order flow, as in the AM strategy, for the purpose of forecasting exchange rate returns. In the context of this maximum return dynamic strategy, we can compute both the in-sample and the out-of-sample performance fee, Φ.

4.3   Transaction costs

In the literature, transaction costs are generally assumed given and not estimated. For example, Marquering and Verbeek (2004) consider three levels of transaction costs, 0.1%, 0.5%, and 1%, representing low, medium, and high costs, respectively. Our empirical models use dynamic strategies and in this context transaction costs can play a significant role in determining returns and comparative utility gains where individuals rebalance their portfolios. Thus, instead of assuming a given cost, we follow the method introduced by Han (2006), Della Corte et al. (2009) and Rime et al. (2010), and calculate the break-even transaction costs,

$$\tau \sum_{j=0}^{9} \left| w_t^j - w_{t-1}^j \frac{1 + r_{t+1}^j}{R_{p,t+1}} \right|,$$

which make the investors indifferent between the dynamic and buy-and-hold strategies in terms of utility. In the present setting, the break-even transaction cost, τ, is the minimum proportional cost that cancels out the utility advantage of a given strategy.

Using the above mean-variance quadratic-utility framework, we design a global strategy consisting of an US investor holding a portfolio of 10 currencies: one domestic (United States), and nine foreign currencies. The investor is exposed to currency risk. We employ each of the 4 models to forecast the one step ahead period of the exchange rate returns. Thereafter, we dynamically rebalance our portfolio by computing the new optimal weights for the maximum return strategy conditioned on the forecasts of each model. In the analysis, the yields of the riskless bonds are proxied by the LIBOR rates.

We report the performance fees for the combinations corresponding to the following cases: (1) three sets of target annualized portfolio volatilities $\sigma_p^* = \{8\%, 10\%, 12\%\}$; (2) all pairs of 3 models against RW; and (3) degrees of RRA $\delta = \{2, 6\}$. We report our estimates of Φ and break-even transaction cost τ as annualized fees expressed in basis points.

5 Estimation and empirical results

5.1 Data and preliminary test

In this study we use two different datasets. The first data set is the customer order flow data set used in Cerrato et al. (2009). The data set consists of customer (weekly frequency) order flows from UBS and covers the period November, 02 2001 - November, 23 2007 for nine of the most liquid currencies. This is the largest data set ever used in the exchange rate microstructure literature. The data set is aggregated across currency pairs with customers split into 4 classifications: asset managers, hedge funds, corporate and private clients. The currencies considered are the Canadian Dollar (CAD), the Swiss Franc (CHF), the Euro (EUR), the Australian Dollar (AUD), the New Zealand Dollar (NZD), the UK Pound (GBP), the Japanese Yen (JPY), the Norwegian Krone (NOK) and the Swedish Krone (SEK). We use the three month LIBOR rate collected from Bloomberg to approximate the risk-free rate.

Since all rates are foreign currency per US dollar, a positive coefficient indicates dollar buying (foreign currency selling), the rate will increase as the foreign currency weakens. Conversly, a decline in this rate represents a strengthening of the foreign currency relative to the US dollar. Descriptive statistics for this data set are reported in Cerrato et al. (2009). Since exchange rates are found to be $I(1)$, we employ log differenced rates. We have used this data set to assess the in sample predictive power of the three models introduced above. Results were not different to what is already reported in Cerrato et al. (2009) and therefore are not reported in this study to save space[3].

Linearity tests against STAR nonlinearity for the order flow are reported in Table

---

[3] These results are available upon request.

	Linearity test for the STAR model				
	aggregate	disaggregate			
		AM	CO	HF	PC
EUR/dollar	$10.198^\dagger$	4.022	1.713	$4.794^\dagger$	0.161
JPY/dollar	4.393	2.022	1.002	$10.517^\dagger$	$11.476^\dagger$
GBP/dollar	$13.046^\dagger$	$32.893^\dagger$	$6.698^\dagger$	1.518	3.789
CHF/dollar	$10.885^\dagger$	$5.943^\dagger$	$17.234^\dagger$	$5.669^\dagger$	0.073
AUD/dollar	3.725	$9.074^\dagger$	$64.932^\dagger$	2.875	$23.236^\dagger$
CAD/dollar	3.939	$13.249^\dagger$	1.689	$4.705^\dagger$	$5.471^\dagger$
NOK/dollar	$22.766^\dagger$	1.818	2.147	0.645	$17.980^\dagger$
SEK/dollar	$15.545^\dagger$	$8.687^\dagger$	$13.278^\dagger$	0.083	3.802
NZD/dollar	$36.289^\dagger$	$7.843^\dagger$	$32.099^\dagger$	$18.601^\dagger$	3.631

Table 1: Linearity test to the aggregate and disaggregate order flows

(1). We use the approach as suggested in Harvey and Leybourne (2007). To implement this test, we select the AR order in the regression using a general-to-specific methodology and a 10%-significance level, (4.605), with a maximum permitted AR order of four and a minimum order of two. We find evidence of nonlinearity for six aggregate order flows and more than half the disaggregate order flows. Thus, more than half of the series analyzed exhibit evidence of nonlinearity and this suggests that nonlinear models may be appropriate.

5.1.1 Evans and Lyons' data set: out of sample forecasts

The second data set considered in this study is the one used in Evans and Lyons (2002b). It contains 80 daily observations on inter-dealer order flow for the mark–dollar and yen–dollar during the period May 1–August 29, 1996. These data were originally collected from the Reuters D2000-1 inter-dealer service and are defined as the difference between the number of buyer-initiated and seller-initiated trades. Thus, in contrast to the data set discussed in the prvious section, this data set consists of interdealer order flow

We start with the out-of-sample forecasts and compare the forecasts using the order flow model as in Evans and Lyons (2002b) and thereafter using the methodologies for model instability as discussed in the previous sections.

Table (2) shows the empirical results. We use a recursive approach to computing forecasts and root mean square errors. At the 1 and 2-week horizons, the Evans–Lyons model, which addresses the publication lag issue, does not outperform the random walk. Our models show a significant predictive power for weekly exchange rate returns at any horizon.

5.2 Customer order flow data: out of sample forecasts

5.2.1 Aggregate order flow

We now turn to the UBS customer order flow data and repeat the forecasting exercise as in the previous section. The out-of-sample predictions are reported in Table

			Out of sample prediction results: forecasts of lagged models using Evans and Lyons Data set									
		RW	E&L	E&L/RW	TVP	TVP/RW	DM	STAR	STAR/RW	DM	STR	STR/RW	DM
DEM/dollar	k =1	0.4332	0.5213	1.2075	0.3966	0.9155	0.3123	0.4184	0.9657	1.3353	0.4078	0.9412	1.0868
	5	0.9700	1.8073	1.8631	0.9212	0.9497		0.9321	0.9610		0.8537	0.8801	
	10	1.5162	2.7752	1.8303	1.3593	0.8965		1.3323	0.8787		1.2175	0.8030	
JPN/dollar	k =1	0.4019	0.4191	1.0430	0.3815	0.9494	0.2777	0.3655	0.9094	1.7169	0.3834	0.9540	0.7069
	5	0.9513	1.8123	1.9050	0.9342	0.9820		0.8287	0.8711		0.9288	0.9763	
	10	1.383	3.2625	2.3499	1.1595	0.8835		0.7786	0.5932		1.0154	0.7737	

[a]

Table 2: Estimated Results for TVP, STAR and STR with Evans and Lyons Dataset

[a]Note: The table reports the ratio of RMSFE for exchange rate forecasts at various horizons against RW with TVP: $\Delta y_{t+k} = \alpha + \beta_{1t}\Delta(i_t - i_t^*) + \beta_{2t}X_t$, STAR: $\Delta y_{t+k} = \alpha + \beta_1\Delta(i_t - i_t^*) + \beta_2 S(.)X_t$ and STR: $\Delta y_{t+k} = \alpha_1 + \alpha_2 S(.) + \beta_1\Delta(i_t - i_t^*) + \beta_2 X_t$ respectively.

(3). As in the previous section, the out-of-sample exercise involves two steps: (1) the initial parameter estimation for the first 267 observations, and (2) sequential weekly updating of the parameter estimates for the rest of the out-of-sample period. In other words, the forecasts at any given week are constructed according to a recursive procedure that is conditional only upon information up to the date of the forecast. The model is then successively re-estimated as the date on which forecasts are conditioned moves through the data set. Hence the design of the out-of-sample exercise is computationally intensive.

At all the horizons, except for GBP, the RMSFE statistics computed using the TVP, STAR and STR are slightly lower than those associated with the random walk forecasts. The Diebold-Mariano test statistic shows that only NZD is significant at the 5% level.

Thus, the empirical results in this section show very little evidence of forecasting power for the order flow model.

5.2.2 Disaggregate order flow

Evans and Lyons (2005b) argue that the lack of success in generating results generally supportive of the core hypotheses of the market microstructure literature may be due to using aggregate customer order flow data. For example, the heterogeneities in the customer segment of the foreign exchange market imply that different customers may react to news in different ways. Sager and Taylor (2008) points out that knowledge of the types of customers prevalent in the market at any given time, and of the ways in which they trade and interact with the wider market, should help understanding of the behavior of an exchange rate at that time.

In this section, following Evans and Lyons (2005b), Sager and Taylor (2008) and Cerrato et al. (2009), we test whether the predictive performance of the order flow model can be improved using disaggregate customer data.

The results of asset managers, with the TVP, STAR and STR models are reported in Table (4). All the series that demonstrate nonlinearity produce an $RMSFE$ ratio which is less than 1. The most striking contrast between the results reported in Tables (3) and (4) is the additional rejection of AUD and CAD in Diebold-Mariano test. This is slightly better than the results of estimated aggregate order flows and can at least partly be explained by multiple structural changes that have been manipulated to ensure customer heterogeneity.

Table (5) and (6) reports the forecasts from the TVP, STAR and STR models for corporate clients and hedge funds, respectively. Except for the CHF with the STR model (see hedge fund), in all cases the RMSFE ratios are less than 1. However, only for CAD (see hedge funds) can the hypothesis that the RMSFE ratios is less than one be rejected at the 10% level with the Diebold-Mariano statistic.

Summing up, the empirical evidence from the previous sections shows a weak empirical evidence that the order flow model can overcome a simple random walk model in out of sample forecasts.

14

		$y_{t+k} - y_t = \alpha + \beta_t X_t$				$y_{t+k} - y_t = \alpha + \beta S(.)X_t$			$y_{t+k} - y_t = \alpha_1 + \alpha_2 S(.) + \beta X_t$		
		RW	TVP	TVP/RW	DM	STAR	STAR/RW	DM	STR	STR/RW	DM
EUR/dollar	k =1	0.8165	0.8041	0.9848	0.13	0.7767	0.9512	0.97	0.7652	0.9372	1.40
	2	1.1761	1.0704	0.9101		1.0805	0.9187		1.0750	0.9141	
	3	1.5422	1.3282	0.8612		1.3218	0.8571		1.3422	0.8703	
	4	1.8963	1.6106	0.8493		1.6318	0.8606		1.6585	0.8746	
GBP/dollar	k =1	0.8980	0.9055	1.0084	-0.06	0.8899	0.9910	0.30	1.0812	1.2040	0.15
	2	1.1837	1.2522	1.0579		1.2263	1.0360		1.1933	1.0081	
	3	1.2362	1.3032	1.0542		1.2955	1.0479		1.3208	1.0684	
	4	1.4269	1.4714	1.0312		1.4209	0.9958		1.4644	1.0263	
CHF/dollar	k =1	0.9651	0.9258	0.9592	0.26	0.9434	0.9775	0.67	0.9539	0.9884	0.60
	2	1.3064	1.2197	0.9336		1.2553	0.9609		1.2917	0.9888	
	3	1.7263	1.5029	0.8706		1.6169	0.9366		1.6856	0.9764	
	4	1.9992	1.7050	0.8529		1.8457	0.9232		1.9965	0.9987	
NOK/dollar	k =1	1.2806	1.2360	0.9651	0.27	1.2522	0.9778	1.15	1.2666	0.9891	1.18
	2	1.8348	1.7370	0.9467		1.7796	0.9699		1.8198	0.9918	
	3	2.3171	2.2241	0.9599		2.2263	0.9608		2.3030	0.9939	
	4	2.8380	2.6468	0.9326		2.5912	0.9130		2.7728	0.9770	
SEK/dollar	k =1	1.0853	1.0806	0.9956	0.03	1.0784	0.9936	0.26	1.0836	0.9984	0.11
	2	1.6874	1.6256	0.9633		1.6406	0.9722		1.6537	0.9800	
	3	2.2432	2.1364	0.9524		2.1391	0.9536		2.1845	0.9738	
	4	2.6823	2.5523	0.9516		2.5498	0.9506		2.6035	0.9706	
NZD/dollar	k =1	2.0665	1.9246	0.9313	0.42	2.0291	0.9819	1.34	2.0411	0.9877	2.43**
	2	2.9120	2.8220	0.9691		2.8729	0.9866		2.8926	0.9933	
	3	3.5876	3.4556	0.9632		3.4894	0.9726		3.5264	0.9830	
	4	4.3568	4.1361	0.9493		4.1954	0.9630		4.2819	0.9828	

Out of sample prediction results: forecasts of lagged models using aggregate order flow

Table 3: Estimated Results for STAR and STR with Aggregate Order Flow

			$y_{t+k} - y_t = \alpha + \beta_t X_t$				$y_{t+k} - y_t = \alpha + \beta S(.)X_t$			$y_{t+k} - y_t = \alpha_1 + \alpha_2 S(.) + \beta X_t$		
		RW	TVP	TVP/RW	DM	STAR	STAR/RW	DM	STR	STR/RW	DM
GBP/dollar	k =1	0.8980	0.8687	0.9674	0.26	0.8840	0.9844	0.49	0.8825	0.9827	0.60
	2	1.1837	1.2255	1.0353		1.2279	1.0374		1.0421	0.8803	
	3	1.2362	1.2881	1.0420		1.2915	1.0447		1.3013	1.0526	
	4	1.4269	1.4602	1.0234		1.4782	1.0359		1.4663	1.0276	
CHF/dollar	k =1	0.9651	0.9086	0.9414	0.39	0.9496	0.9839	0.93	0.9622	0.9970	0.78
	2	1.3064	1.2310	0.9423		1.2556	0.9611		1.2994	0.9946	
	3	1.7263	1.5312	0.8869		1.6370	0.9482		1.7070	0.9888	
	4	1.9992	1.7171	0.8589		1.8469	0.9238		1.9815	0.9911	
AUD/dollar	k =1	1.6960	1.6276	0.9597	0.25	1.6725	0.9861	0.96	1.6699	0.9846	1.97**
	2	2.4159	2.3195	0.9601		2.3450	0.9707		2.3496	0.9726	
	3	2.9776	2.7620	0.9276		2.8687	0.9634		2.8806	0.9674	
	4	3.5769	3.3396	0.9337		3.4226	0.9569		3.4723	0.9708	
CAD/dollar	k =1	1.1611	1.0798	0.9300	0.45	1.1096	0.9557	2.18	1.1197	0.9643	1.64*
	2	1.9189	1.6888	0.8801		1.7990	0.9375		1.8172	0.9470	
	3	2.4335	2.0534	0.8438		2.1779	0.8949		2.2029	0.9052	
	4	2.8512	2.4694	0.8661		2.5104	0.8805		2.5438	0.8922	
SEK/dollar	k =1	1.0853	1.0467	0.9645	0.30	1.0685	0.9846	0.70	1.0753	0.9907	0.97
	2	1.6874	1.5712	0.9311		1.6314	0.9668		1.6467	0.9759	
	3	2.2432	2.1195	0.9448		2.1350	0.9518		2.1849	0.9740	
	4	2.6823	2.5499	0.9507		2.5389	0.9466		2.6106	0.9733	
NZD/dollar	k =1	2.0665	1.9985	0.9671	0.21	2.0513	0.9926	0.63	2.0469	0.9905	1.09
	2	2.9120	2.8519	0.9794		2.8737	0.9868		2.8757	0.9875	
	3	3.5876	3.5244	0.9824		3.5221	0.9818		3.5038	0.9766	
	4	4.3568	4.2676	0.9795		4.2477	0.9750		4.2368	0.9725	

Out of sample prediction results: forecasts of lagged models using disaggregate data (AM asset manager)

Table 4: Estimated Results for STAR and STR with Disaggregate Order Flow - AM asset managers

16

Out of sample prediction results: forecasts of lagged models using disaggregate order flow (CO corporate clients)											
			$y_{t+k} - y_t = \alpha + \beta_t X_t$			$y_{t+k} - y_t = \alpha + \beta S(.) X_t$			$y_{t+k} - y_t = \alpha_1 + \alpha_2 S(.) + \beta X_t$		
		RW	TVP	TVP/RW	DM	STAR	STAR/RW	DM	STR	STR/RW	DM
GBP/dollar	k =1	0.8980	0.8907	0.9918	0.07	0.8697	0.9685	1.28	0.8729	0.9721	1.49
	2	1.1837	1.1463	0.9684		1.1804	0.9972		1.1696	0.9881	
	3	1.2362	1.1552	0.9344		1.1429	0.9245		1.1471	0.9279	
	4	1.4269	1.3278	0.9305		1.3480	0.9447		1.3184	0.9239	
CHF/dollar	k =1	0.9651	0.9466	0.9808	0.13	0.9305	0.9641	1.21	0.9429	0.9770	0.96
	2	1.3064	1.2386	0.9481		1.1932	0.9134		1.2554	0.9609	
	3	1.7263	1.5057	0.8722		1.5890	0.9204		1.6619	0.9627	
	4	1.9992	1.7356	0.8681		1.8184	0.9095		1.9371	0.9689	
AUD/dollar	k =1	1.6960	1.6580	0.9776	0.14	1.6845	0.9932	0.43	1.6859	0.9940	0.80
	2	2.4159	2.2956	0.9502		2.3681	0.9802		2.3951	0.9914	
	3	2.9776	2.7520	0.9242		2.8693	0.9636		2.9285	0.9835	
	4	3.5769	3.3033	0.9235		3.3779	0.9443		3.4927	0.9765	
SEK/dollar	k =1	1.0853	1.0179	0.9379	0.51	1.0609	0.9775	0.44	1.0593	0.9761	0.58
	2	1.6874	1.5810	0.9369		1.7233	1.0213		1.7086	1.0125	
	3	2.2432	2.1529	0.9597		2.3326	1.0398		2.3292	1.0383	
	4	2.6823	2.6715	0.9960		2.7673	1.0317		2.7947	1.0419	
NZD/dollar	k =1	2.0665	2.0193	0.9772	0.14	2.0627	0.9981	0.10	2.0534	0.9936	0.63
	2	2.9120	2.8764	0.9878		2.8859	0.9910		2.8733	0.9867	
	3	3.5876	3.4798	0.9700		3.5010	0.9759		3.5495	0.9894	
	4	4.3568	4.2384	0.9728		4.2293	0.9707		4.3062	0.9884	

Table 5: Estimated Results for STAR and STR with Disaggregate Order Flow - CO corporate clients

Out of sample prediction results: forecasts of lagged models using disaggregate order flow (HF hedge funds)											
			$y_{t+k} - y_t = \alpha + \beta_t X_t$			$y_{t+k} - y_t = \alpha + \beta S(.) X_t$			$y_{t+k} - y_t = \alpha_1 + \alpha_2 S(.) + \beta X_t$		
		RW	TVP	TVP/RW	DM	STAR	STAR/RW	DM	STR	STR/RW	DM
EUR/dollar	k =1	0.8165	0.7879	0.9650	0.29	0.7768	0.9514	1.36	0.7967	0.9757	0.95
	2	1.1761	1.0823	0.9203		1.0862	0.9236		1.1338	0.9641	
	3	1.5422	1.3557	0.8790		1.3751	0.8916		1.4531	0.9422	
	4	1.8963	1.6279	0.8585		1.6495	0.8699		1.7927	0.9454	
JPY/dollar	k =1	1.4010	1.3240	0.9450	0.31	1.3851	0.9886	0.94	1.3843	0.9881	1.40
	2	1.6165	1.5253	0.9436		1.5939	0.9860		1.5789	0.9767	
	3	2.0704	1.8661	0.9013		2.0355	0.9832		2.0203	0.9758	
	4	2.5014	2.4427	0.9765		2.5473	1.0183		2.5178	1.0065	
CHF/dollar	k =1	0.9651	0.9273	0.9608	0.25	0.9605	0.9952	0.25	0.9652	1.0001	-0.01
	2	1.3064	1.2534	0.9594		1.2842	0.9830		1.3062	0.9998	
	3	1.7263	1.4976	0.8675		1.6429	0.9517		1.7084	0.9896	
	4	1.9992	1.7300	0.8653		1.8762	0.9385		2.0033	1.0021	
CAD/dollar	k =1	1.1611	1.0909	0.9396	0.38	1.1272	0.9708	1.67*	1.1323	0.9752	1.25
	2	1.9189	1.7343	0.9038		1.8458	0.9619		1.8614	0.9700	
	3	2.4335	2.0981	0.8622		2.2649	0.9307		2.2984	0.9445	
	4	2.8512	2.5219	0.8845		2.5657	0.8999		2.6051	0.9137	
NZD/dollar	k =1	2.0665	1.9743	0.9554	0.28	2.0482	0.9911	0.64	2.0497	0.9918	1.52
	2	2.9120	2.8301	0.9719		2.8852	0.9908		2.8999	0.9959	
	3	3.5876	3.4154	0.9520		3.4868	0.9719		3.5118	0.9789	
	4	4.3568	4.0425	0.9279		4.2441	0.9741		4.3012	0.9873	

Table 6: Estimated Results for STAR and STR with Disaggregate Order Flow - HF hedge funds

Out of sample prediction results: forecasts of lagged models using disaggregate order flow (PC private client)											
		$y_{t+k} - y_t = \alpha + \beta_t X_t$				$y_{t+k} - y_t = \alpha + \beta S(.) X_t$			$y_{t+k} - y_t = \alpha_1 + \alpha_2 S(.) + \beta X_t$		
	RW	TVP	TVP/RW	DM	STAR	STAR/RW	DM	STR	STR/RW	DM
JPY/dollar k = 1	1.4010	1.2933	0.9231	0.43	1.3714	0.9789	1.54	1.3929	0.9942	0.47
2	1.6165	1.5917	0.9847		1.6050	0.9929		1.6501	1.0208	
3	2.0704	1.8900	0.9128		2.0600	0.9950		2.1359	1.0316	
4	2.5014	2.4764	0.9900		2.5534	1.0208		2.6537	1.0609	
AUD/dollar k = 1	1.6960	1.5726	0.9272	0.45	1.6500	0.9729	1.83*	1.6658	0.9822	2.51**
2	2.4159	2.2930	0.9491		2.3531	0.9740		2.3754	0.9833	
3	2.9776	2.7927	0.9379		2.8409	0.9541		2.9092	0.9770	
4	3.5769	3.3276	0.9303		3.4057	0.9521		3.4924	0.9764	
CAD/dollar k = 1	1.1611	1.0204	0.8788	0.76	1.0808	0.9309	2.89***	1.0949	0.9430	2.88***
2	1.9189	1.5745	0.8205		1.7286	0.9008		1.7760	0.9255	
3	2.4335	1.9758	0.8119		2.2099	0.9081		2.2706	0.9330	
4	2.8512	2.4414	0.8563		2.5118	0.8810		2.5974	0.9110	
NOK/dollar k = 1	1.2806	1.2387	0.9673	0.25	1.2576	0.9820	0.95	1.2576	0.9820	1.38
2	1.8348	1.7319	0.9439		1.7573	0.9578		1.7746	0.9672	
3	2.3171	2.1881	0.9444		2.1477	0.9269		2.1990	0.9491	
4	2.8380	2.6614	0.9378		2.6303	0.9268		2.7076	0.9541	

Table 7: Estimated Results for STAR and STR with Disaggregate Order Flow - HF hedge funds

5.3 Economic evaluation

5.3.1 Evans and Lyons' dataset

In this and the following sections we build a portfolio of currencies and measure the out of sample forecasting performance using the mean variance approach introduced in the previous sections. We start with the Evans and Lyons data set. Results are presented in Table (8). Panel A of table (8) contains the out-of-sample annualized Sharpe Ratios for the nonlinear models. We build an efficient portfolio by investing in the daily return of two currencies, the German DM and Japanese Yen, and using the two exchange rates to convert the portfolio return into US dollars. The maximum return strategies are evaluated at three target portfolio return volatilities, 8%, 10%, and 12%. For instance, at $\sigma_p^* = 10\%$, the out-of-sample Sharpe Ratios are 0.41 for TVP, 1.86 for STAR, and 2.43 for STR. Thus, we can conclude that in terms of economic value the models perform better than a RW.

Panel B of Table (8) contains the out-of-sample performance fees, Φ, and the break-even transaction costs τ^{BE}. The fees denote the amount an investor with quadratic utility and a degree of relative risk aversion equal to 2 and 6 would be willing to pay for switching from the RW model to an alternative model. The target portfolio volatilities are set at 8%, 10%, and 12%. τ^{BE} is defined as the minimum proportional cost that cancels out the utility advantage of a strategy. The fees are expressed in annual basis points. As an example, setting $\sigma_p^* = 10\%$ and $\delta = 2$ the results indicate the out-of-sample fees for switching from the RW model to the nonlinear models are 116 bps for TVP, 100 bps for STAR and 91 bps for STR. Both economic evaluations using the Sharpe Ratio and performance fees confirm that our TVP, STAR, and STR models consistently outperform a RW in out-of-sample forecasts.

5.3.2 Aggregate and disaggregate customer order flows

The empirical results for the UBS dataset are reported in Table (9). We calculate the performance fee and this is reported in the Table (9). We estimate the fees assuming different degrees of relative risk aversion, specifically $\delta = 2$ and $\delta = 6$.

The out-of-sample performance fees are displayed in Table (9) and suggest that there is still high economic value in nonlinear specifications. This is a new and important result, which is in contrast to the seminal contribution of Meese and Rogoff (1983). Specifically, at $\sigma_p^* = 10\%$ and $\delta = 2$, the performance fees for switching from RW to an alternative model are 1793 bps for TVP, 1951 bps for $STAR$ and 1149 bps for STR , when aggregate order flow is used. We can therefore conclude that there is a substantial economic value out-of-sample against the naive random walk model and in favor of conditioning on the order flows with nonlinearity. Thus, there is clear out-of-sample economic value relative to the naive random walk benchmark.

If transaction costs are sufficiently high, the period-by-period fluctuations in the dynamic weights of an optimal strategy will render the strategy too costly to implement relative to the static random walk model. We address this concern by computing the break-even transaction cost, τ, as the minimum proportional cost

Panel A: Sharpe Ratios for Out of Sample Forecasts

σ_p^*	TVP	STAR	STR
8%	0.5146	2.3324	3.0456
10%	0.4117	1.8659	2.4365
12%	0.3431	1.5549	2.0304

Panel B: Performance fee for in-sample forecasts against RW

	TVP				STAR				STR			
	Φ_2	τ_2^{BE}	Φ_6	τ_6^{BE}	Φ_2	τ_2^{BE}	Φ_6	τ_6^{BE}	Φ_2	τ_2^{BE}	Φ_6	τ_6^{BE}
8%	110.3	20.1	110.9	20.2	97.7	8.5	98.4	8.6	90.2	65.0	90.9	65.5
10%	116.3	16.9	116.9	17.0	100.8	7.0	101.4	7.1	91.5	52.7	92.1	53.1
12%	122.2	14.8	122.8	14.9	103.9	6.0	104.5	6.0	92.7	44.5	93.3	44.8

Table 8: Economic Evaluations for the TVP, STAR and STR forecasts with Order Flows

Performance fee for out-of-sample forecasts against RW

	TVP				Aggregate STAR				STR			
σ_p^*	Φ_2	τ_2^{BE}	Φ_6	τ_6^{BE}	Φ_2	τ_2^{BE}	Φ_6	τ_6^{BE}	Φ_2	τ_2^{BE}	Φ_6	τ_6^{BE}
8%	1556.7	104.6	1557.2	104.6	1686.1	172.9	1686.6	172.9	1040.4	48.4	1041.0	48.4
10%	1793.6	96.4	1794.1	96.4	1951.8	159.9	1952.2	160.0	1149.3	42.8	1149.9	42.8
12%	2030.3	90.9	2030.8	90.9	2216.8	151.2	2217.3	151.2	1258.3	39.0	1258.9	39.1
					AM(Asset Manager)							
8%	2031.2	9.1	2031.7	9.1	1619.1	18.5	1619.7	18.5	1348.8	30.6	1349.3	30.6
10%	2387.9	8.5	2388.4	8.5	1872.8	17.1	1873.3	17.1	1534.9	27.9	1535.4	27.9
12%	2744.6	8.2	2745.1	8.2	2126.5	16.2	2127.0	16.2	1721.0	26.0	1721.5	26.1
					CO(Corporate Client)							
8%	1525.7	50.7	1526.2	50.8	1345.8	84.1	1346.3	84.2	1536.6	43.0	1537.1	43.1
10%	1756.0	46.7	1756.5	46.7	1531.1	76.6	1531.6	76.6	1769.7	39.7	1770.2	39.7
12%	1986.4	44.0	1986.9	44.0	1716.5	71.5	1717.0	71.6	2002.7	37.4	2003.2	37.4
					HF(Hedge Fund)							
8%	1409.9	34.1	1410.4	34.1	1250.7	22.2	1251.2	22.3	1557.5	76.5	1558.0	76.6
10%	1611.3	31.1	1611.8	31.1	1412.3	20.1	1412.8	20.1	1795.7	70.6	1796.1	70.6
12%	1812.6	29.2	1813.1	29.2	1573.8	18.6	1574.3	18.7	2033.8	66.6	2034.3	66.7
					PC (Private Client)							
8%	1324.8	33.7	1325.3	33.7	1293.3	27.5	1293.8	27.5	1233.1	56.7	1233.6	56.8
10%	1504.9	30.6	1505.4	30.6	1465.4	24.9	1465.9	24.9	1390.2	51.2	1390.8	51.2
12%	1684.9	28.5	1685.4	28.6	1637.5	23.2	1638.0	23.2	1547.4	47.5	1547.9	47.5

Table 9: Economic Value for the TVP, STAR and STR forecasts with Order Flows

that cancels out the utility advantage of a given strategy. In comparing a dynamic strategy with the static random walk strategy, an investor who pays a transaction cost lower than τ will prefer the dynamic strategy.

The out-of-sample break-even transaction costs are reported in Table (9). It is clear from this that for the TVP, $STAR$ and STR models transaction costs are reasonably high. They tend to be higher than 50 bps. Marquering and Verbeek (2004) argue that, at the reasonably high transaction cost of 50 bps, there is still significant out-of-sample economic value in empirical models that condition on the microstructure order flows, especially under nonlinear specification. Therefore, given the values of Φ, we conclude that the out-of-sample economic value we have reported is robust to reasonably high transaction costs.

5.4   Summary of results

Thus, the empirical results presented above can be summarized as follows: (1) the nonlinear models consistently outperform a random walk model when RMSFEs are considered; (2) when a portfolio of currencies is considered, after conditioning on the microstructure order flow models introduced above, there is clear empirical evidence that these models have a higher economic value than a simple random walk model; (3) the economic value of the forecasts increases after conditioning on the nonlinear models.

6   Robustness

In this section we conduct some robustness tests to check that our results are not driven by a specific model specification. Table (10) presents Sharpe Ratios of the out-of-sample performance for the aggregate and disaggregate order flow models. Conditioning on STAR models generally outperforms the benchmark RW under all scenarios. Overall these empirical results are in line with the ones reported in the previous section.

The order flow models we have used above did not contain the interest rates differential. As an additional check, we have also repeated all the empirical applications as above using the same approaches but using the interest rates differential as an additional regressor. The empirical results are in line with what is already reported and therefore not given here to save space[4].

7   Conclusion

This paper makes several contributions to the literature on exchange rates forecasting. We focus on the initiating customer trades and extend the order flow model to account for model instability. In a microstructure context, Gradojevic and Yang (2006) highlights the necessity of embodying information in a nonlinear way. Our empirical results show that order flow, which is related to the economic fundamentals, has some forecasting power to forecast exchange rate returns when forecasts

---

[4]These empirical results are available upon request.

Sharpe Ratios for Out of Sample Forecasts				
		Aggregate		
	σ_p^*	TVP	STAR	STR
Aggregate	8%	0.9230	1.2809	0.6902
	10%	0.7384	1.0248	0.5522
	12%	0.6153	0.8540	0.4601
		Disaggregate		
AM(Asset Manager)	8%	0.7232	0.9621	0.7186
	10%	0.5786	0.7696	0.5749
	12%	0.4821	0.6414	0.4791
CO(Corporate Client)	8%	0.6272	1.0383	0.7000
	10%	0.5018	0.8306	0.5600
	12%	0.4181	0.6922	0.4667
HF(Hedge Fund)	8%	0.8504	0.7337	0.4633
	10%	0.6803	0.5870	0.3707
	12%	0.5669	0.4892	0.3089
PC (Private Client)	8%	0.5201	1.0096	0.6626
	10%	0.4161	0.8077	0.5301
	12%	0.3467	0.6731	0.4417

Table 10: Sharpe Ratios for the TVP, STAR and STR forecasts with Order Flows

are evaluated using standard statistical methods. As we have discussed, this result is consistent with the hypothesis that order flow variation can be explained using macroeconomic news (for example, in our structured break model the effect of macroeconomic news is captured by a shift in the drift parameter). We use two statistical criteria to evaluate model forecasts: the RMSFEs and also the Diebold-Mariano test.

In addition, and more importantly, we assess the economic value of exchange rate forecasts. We find that the predictive ability of the microstructure order flow has substantial economic value in a dynamic portfolio allocation context and that nonlinear models outperform the naive RW model. We believe these are new and important results which have not been previously documented.

References

Bacchetta, P. and E. V. Wincoop (2006). Can information heterogeneity explain the exchange rate determination puzzle? *American Economic Review 96*, 552–576.

Cerrato, M., H. Kim, and R. MacDonald (2010). Equilibrium exchange rate determination and multiple structural changes. *University of Glasgow Dept. of Economics Discussion Papers 2010-14*.

Cerrato, M., H. Kim, and R. MacDonald (forthcoming). 3-regime asymmetric STAR modeling and exchange rate reversion. *Journal of Money Credit and Banking*.

Cerrato, M., N. Sarantis, and A. Saunders (2009). An investigation of customer order flow in the foreign exchange market. *University of Glasgow Dept. of Economics Discussion Papers 2009-25*.

Della Corte, P., L. Sarno, and I. Tsiakas (2009). An economic evaluation of empirical exchange rate models. *Review of Financial Studies 22*, 3491–3530.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 13*, 253–263.

Engel, C. and K. D. West (2005). Exchange rates and fundamentals. *Journal of Political Economy 113*, 486–517.

Evans, M. D. D. and R. K. Lyons (2002a). Informational integration and FX trading. *Journal of International Money and Finance 21*, 807–831.

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

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

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

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

Fleming, J., C. Kirby, and B. Ostdiek (2001). The economic value of volatility timing. *The Journal of Finance LVI*, 329–352.

Gradojevic, N. and J. Yang (2006). Non-linear, non-parametric, nonfundamental exchange rate forecasting. *Journal of Forecasting 25*, 227–245.

Han, Y. (2006). Asset allocation with a high dimensional latent factor stochastic volatility model. *The Review of Financial Studies 19*, 237–271.

Harvey, D. I. and S. J. Leybourne (2007). Testing for time series linearity. *Econometrics Journal 10*, 149–65.

Killeen, W. P., R. K. Lyons, and M. J. Moore (2006). Fixed versus flexible: Lessons from EMS order flow. *Fixed versus flexible: Lessons from EMS order flow 25*, 551–579.

Love, R. and R. Payne (2003). Macroeconomic news, order flows and exchange rates. *Financial Markets Group Working Papers*.

Love, R. and R. Payne (2008). Macroeconomic news, order flows, and exchange rates. *Journal of Financial and Quantitave Analysis 43*, 467–488.

Marquering, W. and M. Verbeek (2004). The economic value of predicting stock index returns and volatility. *Journal of Financial and Quantitative Analysis 39*, 407–429.

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

Payne, R. and P. Vitale (2003). A transaction level study of the effects of central bank intervention on exchange rates. *Journal of International Economics 61*, 331–352.

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

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

Sarno, L. and G. Valente (2009). Exchange rates and fundamentals: Footloose or evolving relationship? *Journal of the European Economic Association 7*, 786–830.

West, K. D., H. J. Edison, and D. Cho (1993). A utility-based comparison of some models of exchange rate volatility. *Journal of International Economics 35*, 23–45.