Bank i Kredyt 50(1), 2019, 83-106

BANK I KREDYT
BANK & CREDIT

Sentiment-induced regime switching in density forecasts of emerging markets' exchange rates. Calibrated simulation trumps estimated autoregression

Krystian Jaworski*

Submitted: 9 April 2018. Accepted: 15 October 2018.

Abstract

Our contribution to existing research is that we propose a novel method to generate density forecasts of foreign exchange rates using Monte Carlo simulation with regime-switching depending on global financial markets' sentiment. The proposed approach has been examined in a one-month ahead forecasting exercise for 22 emerging market currency rates vs. the US dollar. The key findings of our paper are as follows. We show that: (1) our forecasting method is properly calibrated based on a variety of tests and is also suitable for Value-at-Risk analysis; (2) according to the log predictive density score density forecasts produced with our method are superior to random walk forecasts in the case of all 22 analysed currency pairs, and for 7 exchange rates this advantage is statistically significant; (3) in the case of 19 analysed currency pairs our method performs better than the threshold autoregressive model (TAR) with market sentiment as the threshold variable, and for 11 exchange rates this forecasting edge is statistically significant; (4) in the case of 15 analysed currency pairs the proposed approach yields better results than the AR(1)-GARCH(1,1) benchmark, but in none of the cases this difference is statistically significant. The conducted evaluation of the proposed approach suggests that such tool can be suitable for economists, risk managers, econometricians, or policy makers focused on producing accurate density forecasts of foreign exchange rates.

Keywords: evaluating forecasts, regime switching, density forecast, model selection, Value at Risk

JEL: G17, C14, C53

* Warsaw School of Economics, Department of World Economy, Chair of Economics II; e-mail: kjawor@sgh.waw.pl.

1 Introduction

Amisano and Giacomini (2007) proposed the following definition of a density forecast: “it is an estimate of the future probability distribution of a random variable, conditional on the information available at the time the forecast is made. It thus represents a complete characterization of the uncertainty associated with the forecast, as opposed to a point forecast, which provides no information about the uncertainty of the prediction”.

The use of density forecasts has recently become common in various areas of economics. Density forecasts, which have been used in meteorology for a long time, are increasingly used, for instance, in the fields of energy economics (Huurman, Ravazzolo, Zhou 2012), demand management (Taylor 2012), finance (Ghosh, Bera 2015; Hallam, Olmo 2014; Kitsul, Wright 2013; Shackleton, Taylor, Yu 2010), and macroeconomics (Aastveit et al. 2014; Clark 2011; Herbst, Schorfheide 2012; Kolasa, Rubaszek, Skrzypczyński 2012; Wolters 2015). Financial markets are an ideal candidate for producing density forecasts due to a wide array of high-frequency time series. Risk management is most frequently involved, as density forecasts are used with the aim of estimating portfolio risk (Amisano, Giacomini 2007).

The foreign exchange (FX) market is one of the most essential financial markets in the world, with volume of trading surpassing any other (Hong, Li, Zhao 2007). Groen and Matsumoto (2004) make the point that forecasting exchange rates is therefore important for both market participants and policy makers.

Since the original work by Meese and Rogoff (1983), many studies have been dedicated to the production and evaluation of exchange rate point forecasts, and the well-established view is that usually a simple random walk is the best forecasting model. In addition, though point forecasts garner most of attention, density and interval forecasts of FX rates are also of importance to market participants.

A portion of the literature highlights the influence of sentiment in global financial markets on exchange rates. Fama (1984), Dumas and Solnik (1995), and Hodrick (1989) emphasise the importance of investor risk appetite in the analysis of FX rates. Expanding upon this finding Brunnermeier, Nagel, and Pedersen (2009) point out that carry trades have a tendency to generate losses when global risk increases. Liu, Margaritis and Tourani-Rad (2012) established that FX rates behave asymmetrically in reaction to shifts in global risk aversion. Carry trade currencies have a tendency to strengthen only moderately when market risk decreases but depreciate sharply when conditions deteriorate, i.e., the quotations of high-yield currencies are in the habit of “going up by the stairs and coming down by the elevators”. Hopper (1997) saw that exchange rates seem to be influenced by market sentiment rather than by economic fundamentals. Cairns, Ho and McCauley (2007) show that most of the currencies exhibit significant sensitivity towards volatility indicators. Moreover, Kohler (2010) states that “financial crises are often associated with significant movements in exchange rates, which reflect both increasing risk aversion and changes in the perceived risk of investing in certain currencies”.

In this study we follow on these two strands of literature (density forecasting and the influence of sentiment in global markets on FX rates). The objective of this study is to provide a simple, although effective and universal, framework for constructing density forecasts of exchange rates in the emerging market. For this purpose we use a Monte Carlo simulation based on historical, daily exchange rate returns capturing changes of sentiment in the financial markets (allowing for regime switching). Our aim is not to show how the forecast accuracy improves after we account for changing market sentiment, but to propose a new forecasting approach.

We test our procedure on the exchange rates of 22 emerging market currencies against the US dollar. We chose one month ahead (end of next month) density forecasts, mainly because forecasts with such a time horizon are usually most requested by market participants, and they are also gathered for constructing market consensus by news agencies (e.g. Reuters, Bloomberg); moreover portfolio optimization by investors usually occurs with similar frequency. Still, our method can be easily extended for longer time horizons. Our forecasts are evaluated using popular tests available in the literature and are compared against benchmarks (a random walk, threshold autoregressive models and AR-GARCH models).

The key findings of our paper are as follows. We show that: (1) our forecasting method is properly calibrated based on a variety of tests and is also suitable for Value-at-Risk analysis; (2) the density forecasts produced with our method are superior to the random walk forecasts in the case of all 22 analysed currency pairs, and for 7 exchange rates this advantage is statistically significant; (3) in the case of 19 analysed currency pairs our method performs better than the threshold autoregressive model (TAR) with market sentiment as the threshold variable, and for 11 exchange rates this forecasting edge is statistically significant. Given that our approach and the TAR model are built upon a similar underlying premise such results make a case for the calibration of a model rather than its estimation; (4) in the case of 15 analysed currency pairs the proposed approach yields better results than the AR(1)-GARCH(1,1) benchmark, but in none of the cases this difference is statistically significant. The conducted evaluation of the proposed approach suggests that such a tool can be suitable for economists, risk managers, econometricians, or policy makers focused on producing accurate density forecasts of foreign exchange rates.

The remainder of this paper is organized as follows. Section 2 summarizes the present state of the art and introduces the algorithm used in our forecasting procedure. In Section 3 we evaluate our density forecasts using popular tests from the literature. Section 4 concludes.

2 State of the art and the proposed forecasting algorithm

Tay and Wallis (2000) provide a review of the density forecasting literature. The literature on the density forecasting of FX rates is quite limited. General studies (Boero, Marrocu 2004; Christoffersen, Mazzotta 2005; Clews, Panigirtzoglou, Proudman 2000, Diebold, Hahn, Tay 1999; Sarno, Valente 2005) mainly emphasise the FX rate density forecasts that are based on parametric densities. Usually forecasting exercises use high-frequency data, and the multi-step-ahead density forecasts are rarely examined. Recent studies show that – contrary to point forecasts – the simple random walk can be beaten by nonlinear models with regard to the accuracy of out-of-sample density forecasts (Balke, Ma, Wohar 2013; Hong, Li, Zhao 2007).

This paper contributes to the relevant literature in that we propose an approach which takes into consideration the influence of sentiment in global financial markets while constructing density forecasts of exchange rates. Our forecasting framework is outlined below.

We assume that the FX market on each day is in one of three states (regimes) – neutral/normal, "risk-on" or "risk-off". "Risk-on", "risk-off" correspond to investors' sentiment connected with the level of global market risk (risk aversion). When risk is perceived as low, market participants have a tendency to participate in higher-risk investments ("risk-on"). When risk is regarded as high, market participants usually tend to escape towards so-called save heavens, i.e. lower-risk investments ("risk-off"). Otherwise, we consider that markets are in a "neutral" stance.

The method to determine the regime on the particular day is arbitrary. To do so we consider the value of the VIX index (the "fear gauge"), a widespread indicator of the implied volatility of S&P500 index options. The VIX quantifies investors' expectations of equity market volatility over the next 30-day period. High VIX readings indicate that market participants anticipate large changes of option prices in any direction. VIX quotations will hover around low levels when market participants expect neither a serious downside risk nor a considerable upside potential for prices of options. Historically, the value of VIX was positively correlated with risk aversion (Whaley 2000).

2.1 Data description and transformation

Our empirical analysis is conducted with a cross-section comprising 22 emerging markets' currencies rates vs. the US dollar (i.e. USDPLN, USDRUB, etc.; the list of all currency pairs is provided in Table 1) over the period from January 1999 to December 2015. More specifically, we use over 4200 daily observation for each of the foreign exchange rates. We also use daily close values of the VIX indicator regarding the same period. All data were obtained via the Thomson Reuters database. Table 1 provides summary statistics on the data used in the analysis.

We consider that if, on a given day, the VIX stands above the 3rd quartile of its historical daily values it is a "risk-off" day/stance. If the VIX stands below the 1st quartile, it is a "risk-on" day. Anything between these two values is considered a neutral state (regime). Once again, this is just one of the possible approaches to define the global markets' sentiment regime. Other methods might as well lead to superior forecasting performance. This is an avenue for further research. Nevertheless, it is worth noting that Orlowski (2017) performed a Bai-Perron threshold test (allowing a maximum of one threshold) for the daily series of the VIX market. The test has generated a VIX threshold of 23.89 (i.e. the threshold between tranquil and turbulent days), which is similar to the 3rd quartile of the VIX (24.24), i.e. the threshold between "normal" and "risk-off" days). Such findings support our approach.

To calculate the VIX quartiles and the resulting regimes we use the full sample (i.e. every available daily observation up to the point when the forecast is made). It means that the regimes' threshold values (VIX quartiles) are different depending on the FX forecasting period in question. In Figure 1 you can see the VIX index divided into three regimes, assuming that the FX rate at the end of December 2015 was being forecasted. It means that VIX data up to the end of November 2015 were being used to calculate the regimes. Once again, if we had constructed a forecast for a different time period, a different VIX sample would have been used. This is a pseudo real-time approach.

Using historical data we can calculate a transition matrix [P] between these three states. A 3×3 matrix used to describe the frequencies of transitions between two given states (p_{ij} day after day). For clarity, let's denote "risk-on" = 1, "neutral" = 2 and "risk-off" = 3. In formulas (1-3) i and j indicate the numbers assigned to the regimes persisting on the previous (s_{t-1}) and current (s_t) day, respectively.

$$P=\begin{bmatrix} p_{11} & \cdots & p_{13} \\ \vdots & \ddots & \vdots \\ p_{31} & \cdots & p_{33} \end{bmatrix} \quad p_{ij}=\Pr\left(s_t=j|s_{t-1}=i\right) \tag{1}$$

In Table 2 we present the empirical transition values calculated on the full sample up to the end of November 2015. Please note that a situation when the regime changed directly between “risk-on” and “risk-off” (or the other way around) on two consecutive days without reaching a neutral state in-between has not occurred historically. Effectively, for each state the probability that on the next day the regime will not change compared to today is equal to ca. 91.5%. For “risk-on” and “risk-off” the probability of regime switch to “neutral” is equal in both cases to ca. 8.5%. For “neutral” the probability of change to “risk-off” or “risk-on” on the next day is almost the same and amounts to 4.2%.

Also we calculate a matrix of cumulative probabilities of transitions [C]. It will be used later:

$$C = \begin{bmatrix} c_{11} & \cdots & c_{13} \\ \vdots & \ddots & \vdots \\ c_{31} & \cdots & c_{33} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{11}+p_{12} & p_{11}+p_{12}+p_{13} \\ p_{21} & p_{21}+p_{22} & p_{21}+p_{22}+p_{23} \\ p_{31} & p_{31}+p_{32} & p_{31}+p_{32}+p_{33} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{11}+p_{12} & 1 \\ p_{21} & p_{21}+p_{22} & 1 \\ p_{31} & p_{31}+p_{32} & 1 \end{bmatrix} \tag{2}$$

$$c_{ij} = \Pr\left(s_t \leq j | s_{t-1} = i\right) \tag{3}$$

In Table 3 we present the empirical cumulative probabilities of transition calculated on the full sample (up to the end of November 2015).

For a given FX rate [FX_t] (e.g. USDPLN) we calculate its daily logarithmic returns for the same sample as in the case of the VIX – every available daily observation up to the point when the forecast is made.

$$r_t = \log\left(FX_t\right) - \log\left(FX_{t-1}\right) \tag{4}$$

We divide the daily logarithmic returns into three separate groups (empirical distributions) according to the state in which they occurred: “risk-on returns” $f(r^1)$, “normal returns” $f(r^2)$ and “risk--off returns” $f(r^3)$.

We present the histogram (density function) of the returns for the USDPLN calculated as above in Figure 2 (daily observations up to the end of November). The X-axis (daily logarithmic returns) is limited to the range between -4% and +4% for better presentation. But, of course, it must be noted that the main differences between the three distributions occur in the tails (which are not visible on the chart).

Once the data are transformed, they will be used later to construct the FX density forecast. The proposed methodology does not impose any parametric distribution for the returns of the foreign exchange rate and (instead) let the density forecast to be data-driven and based on the risk-regime under consideration. It must be noted that constructing a FX forecast for different time periods requires multiple calculations of regimes’ threshold values, P and C matrices, as well as the division of FX returns into three regimes – each time on a different sample (daily observations up to the point when the forecast is made). The detailed role of these elements in the forecasting algorithm is presented in Section 2.2.

2.2 Forecasting algorithm

1. At the end of month m we check how many trading days [h] are in the month $m + 1$, i.e. the month, for the end of which we would like to generate the FX rate forecast (e.g. $h = 20$ days).

2. As a starting point of the forecast we take the close FX rate of the last trading day of the month m [FX_0].

3. We also note the regime that persisted on this day [s_0], using the approach outlined in Section 2.1.

4. In order to construct a forecast of the FX rate on the first day of the $m + 1$ month [FX_1] we first need to know (simulate) in what state the markets are on this (1st) day. To do so, we first calculate P and C matrices as well as divide the daily returns into three groups, as it is outlined in Section 2.1, using every available daily observation up to the point when the forecast is constructed.

5. Using the transition matrix C we compute the regime on the first day [s_1] of the $m + 1$ month. Depending on the state s_0, we choose one row of the C matrix. If s_0 is a risk-on state, we choose the 1st row, if neutral, the 2nd row and if risk-off, the 3rd row.

6. We randomly draw a number [x] from a uniformly distributed range [0; 1]. Then we select the smallest element of the row indicated in point 5 that is larger than or equal to x.

7. The number of elements that we chose (1st, 2nd or 3rd) determines the regime on the first day of the $m + 1$ month (s_1). If we chose the 1st element, s_1 is a risk-on state; if we chose the 2nd element is neutral; if we chose the 3rd element, s_1 is a risk-off state.

8. Depending on what state ("risk-on", "neutral" or "risk-off") occurs on the first day of the current month (s_1) according to our simulation, we randomly choose a daily percentage return [r^*] from either $f(r^1)$ or $f(r^2)$ or $f(r^3)$, respectively.

9. Then we use it to obtain FX_1 as follows

$$FX_1 = FX_0 * \left(1 + r^*\right) \tag{5}$$

10. In the same way (first randomly obtaining the regime on the next day using the transition matrix and then a return from this particular state) we can recursively calculate the FX rate values for all the remaining ($h - 1$) days of the $m + 1$ month.

$$FX_\tau = FX_{\tau-1} * \left(1 + r^*\right), \quad \tau = 2, 3, \ldots, h \tag{6}$$

Please note that r^* (dependent on the state occurring one day earlier and the transition matrix) is randomly chosen in each iteration, and is (usually) different in each iteration.

Then using the Monte Carlo approach, we repeat the whole process of forecasting (steps 1-10) N times. The only restraint on N is the time required for calculation. We use $N = 15000$. By doing so we get a simulated distribution of the one-month ahead ($m + 1$) forecast of FX rate (N instances of FX_h). The algorithm may be used for longer horizons, but as we outlined earlier, this is not the aim of this paper.

3 Calculation and evaluation of density forecasts

We have tested the out-of-sample forecasting accuracy of this algorithm by constructing 72 one month ahead density forecasts for the end of each month in 2010–2015, for each of 22 emerging markets' FX rates. The first out-of-sample forecast (for the end of January 2010) was constructed with a model using all the available data regarding the VIX and a given FX rate from the 1990–2009 period. Further forecasts are constructed on the rolling sample (the window moving by one-month steps). The use of a rolling sample is vindicated due to the fact that the data-generating process in the financial markets is unstable and often changes as time passes. At the same time the sample should be possibly long to properly capture the wide range of VIX values used in calculating the regimes. The above means that we follow a pseudo real-time forecasting design.

The aim of this paper is to evaluate density forecasts. Therefore extensive investigation of point forecast accuracy (using the mean of the density forecast) is not required in this paper. Still it is always informative to know to what extent superior performance in terms of density forecasts is driven by better point forecasts and to what extent by better calibration in terms of forecast dispersion. Therefore, we present the brief results regarding root-mean-square errors of point forecasts obtained using the proposed approach and compare them to random walk forecasts.

In the case of 12 out of 22 currency pairs the proposed approach yielded a lower RMSE compared to the random walk (see Table 4) at a one-month horizon. However, based on the original Diebold-Mariano-West (1995, 1996) test, as well as its modified version by Harvey, Leybourne and Newbold (1997), which is more suitable for small samples, the difference was statistically significant only in the case of two exchange rates (i.e. USDEGP and USDZAR) at the 5% significance level. Out of the 10 remaining currency pairs for which the point forecasts constructed using the proposed model were inferior to random walk, only in the case of USDPHP the difference was statistically significant at 5%. Such results are consistent with the general view in the literature (Meese, Rogoff 1983) that random walk cannot usually be beaten in out-of-sample FX point forecasting.

To evaluate the quality of the density forecast we follow the novel approach outlined in Gaglianone and Marins (2017). Their approach to density forecast evaluation concentrates on two dimensions. The first one is the full-density analysis, which is a shape evaluation based on the entire estimated density. In this part of evaluation they investigate – among other things – coverage rates, Berkowitz (2001) density test, and the model ranking from log predictive density scores. The second dimension is local analysis, which evaluates the tails of the densities, i.e. the so-called VaR measures. VaR backtesting is performed using Kupiec (1995), Christoffersen (1998) and VQR (2011) tests.

3.1 Coverage rates

Clark (2011) points out that coverage rates constitute a good first step in the evaluation of density forecasts and, more specifically, in the evaluation of the accuracy of interval forecasts. Other studies such as Giordani and Villani (2010) also observe that interval forecasts are a valid test of density forecast calibration.

Table 5 summarizes the frequency with which the realized FX rates were in the 70% highest density interval (highest density region) calculated using the proposed approach. Thus a correct 70% coverage rate interval should correspond to a frequency of ca. 70% in the table. A frequency of less (more) than 70% indicates that, in the case of the analysed sample, the estimated density is too narrow (wide). The table provides p-values for the null hypothesis of appropriate coverage, which means empirical coverage equal to 70%, based on t-statistics (standard errors calculated with the Newey-West estimator). These p-values are supplied as an approximate indicator of the significance of deviations from correct coverage.

Table 5 shows that the proposed forecasting approach yields correct interval forecasts (i.e. coverage rates equal approximately 70%) for 14 out of 22 exchange rates. For other 8 FX rates the null hypothesis is rejected (at the 5% confidence level). For USDEGP, USDIDR, USDKRW, USDTHB and USDTRY intervals turned out to be too wide, with actual observations residing within the intervals more often than the nominal 70% rate. On the other hand, in the case of USDCNY, USDMYR and USDRUB the intervals are too narrow. These results are superior to those calculated using the random walk forecast.[1] For random walk density forecast, the null hypothesis is rejected 11 out of 22 times.

3.2 Berkowitz (2001) test

Berkowitz (2001) proposed a density test, which utilises a probability integral transformation (PIT). The test aims at verifying whether true and hypothesized probability distributions correspond to each other, i.e., $H_0 : F_{L_0}(l) = F_L(l)$. Here, F_{L_0} symbolises the hypothesized cumulative distribution function and F_L stands for the true cumulative distribution function. The normalized forecast error is defined as $\tilde{z}_{t+1} = \Phi^{-1}(z_{t+1})$, where $\tilde{z}_{t+1}$ denotes the PIT of a 1-step ahead forecast error and Φ^{-1} is the inverse of the standard normal distribution. Under the null, it is assumed that the PITs are independently and normally distributed. If the null is not rejected, the Berkowitz (2001) test suggests correct calibration of the density model.

Regarding the Berkowitz test, Table 6 reveals that the forecasts constructed using the proposed approach are not rejected for 16 out of 22 exchange rates (at 5% confidence level), which suggests correct calibration of the density model. In the case of random walk forecasts, models for 12 exchange rates are not rejected.

3.3 Value at Risk backtesting

We also evaluate the prognostic accuracy of our model using the local analysis technique. This concept entails the examination of the performance of the model in the tails of the density forecast distribution. Employing a given forecasting model to produce the full density of the FX rate it might generate, for example, a "satisfactory" risk measure for the right tail of the distribution (i.e. at high percentiles) but simultaneously it can produce "inadequate" risk measure at the left tail of the distribution. Accordingly, we subsequently evaluate the proposed forecasting approach based on its performance for a range of chosen percentiles of the conditional distribution. Each of them may be viewed as a VaR measure (see Christofersen, Hahn, Inoue 2001).

Although VaR gained importance as a risk management tool and is currently extensively utilised for that purpose, this approach has repeatedly been criticized for being unable to generate reliable risk estimates. Implementation of VaR models usually entails making assumptions about the functioning of the financial markets. Furthermore, the VaR model attempts to predict future financial markets' quotations using historical observations, which naturally do not always correctly represent future market conditions.

[1] A random walk model without drift paired up with a normal distribution to generate a density forecast. The random walk point forecast indicates the expected value of the distribution, and the variance of the distribution is implied by the variance of past point forecast errors in the sample.

Therefore, VaR models are effective only when they forecast future risks correctly. In the interest of confirming that the estimates obtained using the VaR approach are reliable and consistent, models ought to be backtested with the use of suitable statistical procedures. Backtesting is a practice in which realized profits and losses are confronted with VaR evaluations. Jorion (2001) fittingly describes backtesting as "reality checks". If the VaR evaluations are not valid, the proposed approach must be reviewed for a flawed specification, improper assumptions or an incorrect modelling method, among others. An assortment of various evaluation procedures has been developed for backtesting purposes. In this study, we utilise Kupiec (1995), Christoffersen (1998), and VQR (2011) tests in order to evaluate the proposed forecasting approach.

Kupiec test

The Kupiec (1995) test is one of the most widespread VaR backtests. This approach concentrates just on the frequency of VaR violations, in other words, the percentage of observations for which the VaR threshold is overstepped in the analysed sample. It is assumed that the probability of a given share of violations occurring follows a Binomial distribution. If the actual number of observed violations fits within acceptable statistical limits the model is approved, and discarded otherwise (Campbell 2006). The test utilises the likelihood ratio approach and the chi-squared distribution to test the hypothesis (Dowd 2006). The null hypothesis suggests a proper calibration of the model.

Christoffersen test

A different, essential element of VaR model evaluation is the verification that data points running above VaR estimates are serially independent, i.e. dispersed over the entire sample. A correctly specified VaR model is able to prevent the interdependence of outliers by promptly responding to shifts in asset price volatilities and correlations. Vast literature is devoted to VaR evaluation procedures that take into consideration the clustering of VaR limit violations. One of such methods is the Christoffersen (1998) test. The null hypothesis suggests a proper calibration of the model.

VQR test

The majority of the VaR backtests outlined in the literature is usually based on binary variables, for example whether or not there was a VaR limit violation. By using such techniques one loses a lot of details about the available data. Therefore we follow the Gaglianone et al. (2011) procedure that does not depend exclusively on binary variables. Utilising a quantile regression model, the suggested approach enables the researcher to determine time intervals when there was an elevated risk exposure. The null hypothesis of the proposed VQR test suggests that the model accurately estimates the actual τ^*th percentile of the distribution. Please note that the null hypothesis is not suggesting that the model accurately estimates the whole distribution. Thus, it is feasible that the model might perform correctly at a selected percentile, but inadequately in other cases. Due to the complexity of the test and its assumptions, and given the space restriction, a full overview of the methodology will not be provided. A more detailed description of this testing procedure can be found in Gaglianone (2011).

We aggregate the results (see Table 7) of the three aforementioned tests with respect to the low quantile $\tau = 0.1$ (the ability to correctly represent the appreciation risk of an emerging market currency against the US dollar) or higher quantile $\tau = 0.9$ (the devaluation risk of an emerging market currency).

For 7 out of 22 exchange rates the null is not rejected (simultaneously) in any of the three aforementioned statistical tests (neither for low nor for high quantiles). Additionally, it should be noted that for additional 2 exchange rates (9 in total) the proposed approach is appropriate to estimate percentiles belonging to the lower portion of the FX rate distribution (concentrating on the appreciation of the local currency versus the USD, $\tau = 0.1$). Also for 6 additional FX rates (13 in total), the proposed approach makes it possible to represent the right tail of the distribution (related to the depreciation of the FX rate,) correctly according to all three tests. Only in the case of 3 exchange rates (USDCNY, USDKRW, USDTHB) none of the three tests indicate proper calibration for either the lower or the higher quantile.

3.4 Log predictive density scores

Once we established that our approach produces properly calibrated density forecasts, we can compare its performance against different benchmarks. To do so we employ the log predictive density score (LPDS). This measure provides a way to classify the analysed models (different benchmarks) regarding their accuracy (correct calibration). The LPDS of the model/benchmark m for the forecast of FX rate in the horizon h is given as:

$$LPDS_{m,h} = T^{-1}\sum_{t=1}^{T}\ln\left(\hat{f}^{m}_{t+h,t}\left(Y_{t+h}\right)\right)$$

where $\hat{f}^{m}_{t+h,t}$ is the density forecast for the exchange rate in period $t+h$ calculated with model m and utilising information set available at period t.

The mentioned density is assessed at the observed FX rate Y_{t+h} and log averaged using the out-of--sample observations. Adolfson, Linde and Villani (2005) note that higher LPDS points to a superior model/benchmark.

Amisano and Giacomini (2007) developed a likelihood ratio test for confronting out-of-sample performance of two rival density forecasts. They recommended calculating scoring rules, which are loss functions established on the basis of the probability forecast and the actual outcome of the FX rate. The proposed test sets side by side the LPDS between two rival benchmarks. The null hypothesis assumes equal LPDS for both models (i.e. density forecasts are equally good). The alternative suggests that the performance of the model with higher LPDS is statistically superior to its counterpart (the model with lower LPDS).

Benchmark models

We compare our model (Baseline) using the LPDS criterion against 9 benchmarks – a naïve forecast and 8 econometric models. The 8 estimated models adopted in this study belong either to the class of threshold autoregressive (TAR) models or generalized auto-regressive conditional heteroskedasticity (GARCH) models.

The TAR models were chosen as benchmark because they are the most natural competitor to the approach proposed by us. In both cases the exchange rate volatility is analysed in the framework of changing regimes based on market sentiment. Comparing these two approaches will indicate whether the calibration of parameters proposed by us performs better in comparison to the situation in which the parameters are estimated within the TAR setup. The AR(1)-GARCH(1,1) benchmark was used due to its popularity in the literature and universality across different FX rates. At first glance it seems inappropriate to compare an approach that utilises an exogenous variable (i.e. VIX), with a model that depends solely on the past values of the endogenous variable. However, despite this handicap, AR-GARCH specification is usually adequate to capture FX rates volatility. There are numerous examples in the literature, when a simple AR-GARCH model performs as well as, or even better (in an out-of-sample forecasting exercise) than, more advanced threshold models (e.g. Pippenger, Goering 1998; Boero, Marrocu 2004). Therefore, we decided to use AR(1)-GARCH(1,1) as a final benchmark regarding the forecasting performance of our approach.

The naïve forecast consists in a random walk model without drift (RW). We joined it up with a normal distribution to be able to generate a density forecast. The random walk point forecast indicates the expected value of the distribution, and the variance of the distribution is implied by the variance of past point forecast errors in the sample.

The first estimated benchmark (B1) is a three-regime TAR model, with VIX Volatility Index as a threshold variable. The number of regimes was chosen to correspond to the number of regimes in the Baseline model. The model was estimated on monthly data. The TAR model can be represented as follows:

$$\Delta y_t = \begin{cases} \phi_0^{(1)} + \phi_1^{(1)} \Delta y_{t-1} + \check{n}_t^{(1)} & \text{if} \quad x_t \leq r_1 \\ \phi_0^{(2)} + \phi_1^{(2)} \Delta y_{t-1} + \check{n}_t^{(2)} & \text{if} \quad r_1 < x_t \leq r_2 \\ \phi_0^{(3)} + \phi_1^{(3)} \Delta y_{t-1} + \check{n}_t^{(3)} & \text{if} \quad x_t > r_2 \end{cases} \tag{8}$$

where y_t is the logarithm of the FX rate (e.g. USDPLN), x_t is a threshold variable, in this case the VIX Volatility Index, r_j represents the threshold values, $\phi_0^{(j)}$ and $\phi_1^{(j)}$ are the estimated coefficients of the linear models specific for each of the three regimes, and $\check{n}_t^{(j)}$ are assumed as IID(0, $\sigma^{2(j)}$).

In order to generate density forecasts a bootstrap method is used. Residuals are resampled from the empirical residuals of the estimated model.

The second estimated benchmark (B2) is very similar to the B1 benchmark. The only difference is that in order to generate density forecasts the Monte-Carlo method is used, where residuals are taken from a normal distribution with standard deviation equal to the standard deviation of the residuals.

Benchmarks B3 and B4 are analogous to benchmarks B1 and B2, respectively, but are estimated on daily, instead of monthly, data.

The fifth estimated benchmark (B5) is an AR(1)-GARCH(1,1), normal distribution model estimated on the daily data (workdays).

$$\Delta y_t = \alpha + \beta \Delta y_{t-1} + \varepsilon_t \tag{9}$$

$$\sigma_t^2 = \omega + \gamma \sigma_{t-1}^2 + \delta \varepsilon_{t-1}^2 \tag{10}$$

$$\varepsilon_t \sim N\left(0, \sigma_t^2\right) \tag{11}$$

where σ_t^2 is the conditional variance and ε_t is the residual following a normal distribution.

The sixth benchmark (B6) is also an AR(1)-GARCH(1,1) model, but with residuals (ε_t) that are Student's t distributed. Benchmarks seven and eight (B7, B8) are the same models as B5 and B6, respectively but estimated on monthly data instead of daily observations. In each case the benchmark models were estimated on the same (rolling) sample as the baseline model – a pseudo real-time forecasting design.

Forecasting competition results

The LPDS ranking (left side of Table 8) point, in general, to the proposed approach as the best model for forecasting majority (14 out of 22) of exchange rates. For the remaining 8 out of 22 exchange rates benchmark B1 or B8 are performing best. It is also noted that benchmarks B2–B7 and the RW forecast are usually overwhelmed in the majority of cases.

The Amissano and Giacomini (2007) test suggests that in the case of 11 currency pairs (USDCLP, USDEGP, USDKRW, USDMYR, USDPEN, USDRUB, USDTWD, USDVND, USDZAR, USDRON, USDCZK, grey cells in the left column on the right side of Table 8) the proposed approach performs statistically better that the best TAR model (i.e. among the B1–B4 benchmarks). Such results make a case for the calibration of model parameters, instead of estimating them to achieve better forecasting accuracy.[2] Such properties were also observed in the case of a half-life purchasing power parity model that was calibrated, rather than estimated (Ca'Zorzi, Rubaszek 2018). Only in the case of USDTHB the TAR benchmark was statistically better than the baseline model. For the remaining 10 currency pairs our approach yielded higher LPDS than the best TAR model, but the difference was not statistically significant.

We also compared the proposed approach against the random walk (RW) forecast. In 7 cases (USDCOP, USDIDR, USDKRW, USDTHB, USDRUB, USDPHP, USDVND; grey cells in the middle column on the right side of Table 8) the proposed model is statistically better than the random walk forecast. In the remaining 15 cases, the LPDS of our method is also higher than in the case of RW, but the Amisano and Giacomini test signals no statistically significant difference between the models.

The final check is the comparison with AR(1)-GARCH(1,1) benchmarks. In the case of 15 analysed currency pairs the proposed approach yields better results than the AR(1)-GARCH(1,1) benchmark, but in none of these cases the difference is statistically significant. Also in only three cases (USDIDR, USDINR, USDTHB) the baseline model performs significantly worse than the best benchmark. This result confirms the mainstream view presented in the literature that the GARCH model is usually an adequate approach to capture FX rates volatility and to prepare a density forecast.

One must also remember that the difficulty to single out a statistically superior model is not surprising taking into consideration the possibly low power of the utilised evaluation approach, on account of a somewhat short sample length (only 72 out-of-sample data points) to perform density forecast comparisons. On the other hand, in typical evaluations of density forecasts of financial markets' indicators hundreds or thousands of observations are being used – e.g. daily returns (Gaglianone, Marins 2017).

[2] The average estimated lower threshold value for the VIX (dividing risk-off and normal regimes) across all currency pairs equaled 16.52 (ranging from 12.42 to 25.05), whereas the calibrated value was 14.52. The average estimated higher threshold value for the VIX (dividing the normal and risk-on regimes) across all currency pairs equaled 21.82 (ranging from 14.02 to 26.20), whereas the calibrated value was 24.34.

4 Conclusions

The key contribution of the paper is a novel approach to produce density forecasts for emerging markets accounting for changing market sentiment. The proposed method applies a Monte Carlo simulation with regime-switching depending on global financial markets' sentiment. Using multiple density forecast evaluation tools, the approach has been examined in a one-month ahead forecasting exercise for 22 emerging market currency rates vs. the US dollar.

Based on evaluation criteria regarding full-density (coverage ratios and the Berkowitz test) the proposed model is properly calibrated for most exchange rates (14 and 18, respectively; at the 5% confidence level). Also backtesting with Kupiec, Christoffersen and VQR tests indicates the usefulness of the proposed approach for a VaR analysis of the majority of the emerging markets' exchange rates.

According to the LPDS scores, the forecasting performance of the proposed approach is superior to the random walk forecast for all the 22 analysed exchange rates and more accurate than the best TAR benchmark in the case of 19 analysed currency pairs. Based on the Amisano and Giacomini (2007) test, this advantage is statistically significant in the case of 7 and 11 exchange rates, respectively. Given that our approach and the TAR model are built upon a very similar underlying premise, such results make a case for the calibration of the model rather than its estimation (mainly the threshold values of the regimes). In the case of 15 analysed currency pairs the proposed approach yields better results than the AR(1)-GARCH(1,1) benchmark, but in none of the cases this difference is statistically significant.

The conducted evaluation of the proposed model indicates that such a tool can be suitable for economists, risk managers, econometricians or policy makers focusing on producing density forecasts of foreign exchange rates. This paper contributes to the relevant literature in that we propose a new approach to constructing density forecasts of exchange rates. The model makes it possible to capture asymmetric changes and fat tails in foreign exchange rates. It should be highlighted that the proposed methodology does not impose any parametric distribution on the returns of the foreign exchange rate and (instead) lets the density forecast to be data-driven and based on the risk-regime under consideration.

Although the results show that the proposed method may not constitute a universally valid tool to produce density forecasts for all exchange rates, the evaluation process still indicates that it yields promising results for the majority of currency pairs. Moreover, the proposed approach allows great flexibility. The possible modification of the procedure may include different definitions of financial markets' stances or the introduction of more regimes. This is an avenue for further research, which is likely to enhance the forecasting performance of the presented approach.

References

Aastveit K.A., Gerdrup K.R., Jore A.S., Thorsrud L.A. (2014), Nowcasting GDP in real time: a density combination approach, *Journal of Business & Economic Statistics*, 32(1), 48–68.

Adolfson M., Linde J., Villani M. (2005), *Forecasting performance of an open economy Dynamic Stochastic General Equilibrium Model*, Sveriges Riksbank Working Paper, 190.

Amisano G., Giacomini R. (2007), Comparing density forecasts via weighted Likelihood Ratio Tests, *Journal of Business and Economic Statistics*, 25, 177–190.

Balke N.S., Ma J., Wohar M.E. (2013), The contribution of economic fundamentals to movements in exchange rates, *Journal of International Economics,* 90(1), 1–16.

Berkowitz J. (2001), *Testing Density Forecasts with Applications to Risk Management,* University of California.

Boero G., Marrocu E. (2004), The performance of SETAR models: a regime conditional evaluation of point, interval and density forecasts, *International Journal of Forecasting*, 20, 305–320.

Brunnrmeier M.K., Nagel S., Pedersen L.H. (2009), Carry trades and currency crashes, *Macroeconomic Annual*, 3(1), 313–348.

Cairns J., Ho C., McCauley R. (2007), Exchange rates and global volatility: implications for Asia--Pacific currencies, *BIS Quarterly Review*, March.

Campbell S.D. (2006), A review of backtesting and backtesting procedures, *The Journal of Risk,* 9(2), 1–17.

Ca'Zorzi M., Rubaszek M. (2018), *Exchange rate forecasting on a Napkin*, ECB Working Paper, 2151.

Christoffersen P., Hahn J., Inoue A. (2001), Testing and comparing value-a-risk measures, *Journal of Empirical Finance*, 8, 325–342.

Christoffersen P. (1998), Evaluating interval forecasts, *International Economic Review*, 39, 841–862.

Christoffersen P.F., Mazzotta S. (2005), The accuracy of density forecasts from foreign exchange options, *Journal of Financial Econometrics*, 3, 578–605.

Clark T.E. (2011), Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility, *Journal of Business & Economic Statistics*, 29(3), 327–341.

Clews R., Panigirtzoglou N., Proudman J. (2000), Recent developments in extracting information from options markets, *Bank of England Quarterly Bulletin*, February.

Diebold F.X., Hahn J., Tay A.S. (1999), Multivariate density forecast evaluation and calibration in financial risk management: high-frequency returns of foreign exchange, *Review of Economics and Statistics*, 81, 661–673.

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

Dowd K. (2006), Retrospective assessment of Value at Risk, in: M.K. Ong (ed.), *Risk Management*: *A Modern Perspective*, Academic Press.

Dumas B., Solnik B. (1995), The world price of foreign exchange risk, *Journal of Finance*, 2, 445–479.

Fama E.F. (1984), Forward and spot exchange rates, *Journal of Monetary Economics*, 14, 319–338.

Gaglianone W., Lima L., Linton O., Smith D. (2011), Evaluating value-at-risk models via quantile regression, *Journal of Business & Economic Statistics*, *Taylor & Francis Journals*, 29(1), 150–160.

Gaglianone W., Marins J. (2017), Evaluation of exchange rate point and density forecasts: an application to Brazil, *International Journal of Forecasting*, 33(3), 707–728.

Ghosh A., Bera A.K. (2015), *Density forecast evaluation for dependent financial data: Theory and applications*, China International Conference in Finance 2015, July 9–12, 1–57, Research Collection Lee Kong Chian School Of Business.

Giordani P., Mattias V. (2010), Forecasting macroeconomic time series with locally adaptive signal extraction, *International Journal of Forecasting*, 26(2), 312–325.

Groen J., Matsumoto A. (2004), Real exchange rate persistence and systematic monetary policy behaviour, *Bank of England Working Papers*, 231.

Hallam M., Olmo J. (2014), Semiparametric density forecasts of daily financial returns from intraday data, *Journal of Financial Econometrics*, 12(2), 408–432.

Harvey D., Leybourne S., Newbold P. (1997), Testing the equality of prediction mean squared errors, *International Journal of Forecasting*, 13(2), 281–291.

Herbst E., Schorfheide F. (2012), Evaluating DSGE model forecasts of comovements, *Journal of Econometrics*, 171(2), 152–166.

Hodrick R.J. (1989), Risk, uncertainty, and exchange rate, *Journal of Monetary Economics*, 23, 433–459.

Hong Y., Li H., Zhao F. (2007), Can the random walk be beaten in out-of-sample density forecasts: evidence from intraday foreign exchange rates, *Journal of Econometrics*, 141, 736–776.

Hopper G. (1997), What determines the exchange rate: economic factors or market sentiment?, *Business Review*, 5, 17–29.

Huurman C., Ravazzolo F., Zhou C. (2012), The power of weather, *Computational Statistics & Data Analysis*, 56(11), 3793–3807.

Jorion P. (2001), *Value at Risk, the New Benchmark for Managing Financial Risk*, McGraw-Hill.

Kitsul Y., Wright J.H. (2013), The economics of options-implied inflation probability density functions, *Journal of Financial Economics*, 110(3), 696–711.

Kohler M. (2010), Exchange rates during financial crises, *BIS Quarterly Review*, March.

Kolasa M., Rubaszek M., Skrzypczyński P. (2012), Putting the New Keynesian DSGE model to the real-time forecasting test, *Journal of Money, Credit and Banking*, 44(7), 1301–1324.

Kupiec P. (1995), Techniques for verifying the accuracy of risk management models, *Journal of Derivatives*, 3, 73–84.

Liu M.H., Margaritis D., Tourani-Rad A. (2012), Risk appetite, carry trade and exchange rates, *Global Finance Journal*, 23, 48–63.

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

Orlowski L.T. (2017), Volatility of commodity futures prices and market-implied inflation expectations, *Journal of International Financial Markets*, *Institutions and Money*, 51, 133–141.

Pippenger M.K., Goering G.E. (1998), Exchange rate forecasting: results from a threshold autoregressive model, *Open Economies Review*, 9(2), 157–170.

Sarno L., Valente G. (2005), Empirical exchange rate models and currency risk: some evidence from density forecasts, *Journal of International Money and Finance*, 24(2), 363–385.

Shackleton M.B., Taylor S.J., Yu P. (2010), A multi-horizon comparison of density forecasts for the S&P 500 using index returns and option prices, *Journal of Banking & Finance*, 34(11), 2678–2693.

Tay A., Wallis K. (2000), Density forecasting: a survey, *Journal of Forecasting*, 19, 235–254.

Taylor J.W. (2012), Density forecasting of intraday call center arrivals using models based on exponential smoothing, *Management Science*, 58(3), 534–549.

Whaley R.E. (2000), The investor fear gauge, *Journal of Portfolio Management*, Spring, 12–17.

Wolters M.H. (2015), Evaluating point and density forecasts of DSGE models, *Journal of Applied Econometrics*, 30(1), 74–96.

Appendix

Table 1

Descriptive statistics

Currency pair	Full name	Mean	Std	Min	Max	Skew	Kurt
		in %					
USDBRL	Brazilian real	0.028	1.148	-10.407	8.81	0.218	12.806
USDCLP	Chilean peso	0.010	0.627	-3.599	4.638	0.336	6.910
USDCNY	Chinese yuan	-0.006	0.092	-2.032	1.816	-0.437	105.165
USDCOP	Colombian peso	0.017	0.694	-6.375	5.129	0.198	11.527
USDEGP	Egyptian pound	0.020	0.405	-3.021	15.540	15.258	548.272
USDIDR	Indonesian rupiah	0.013	0.779	-10.004	8.690	-0.449	28.281
USDINR	Indian rupee	0.011	0.421	-3.551	3.701	0.175	12.323
USDKRW	South Korean won	0.000	0.681	-11.481	10.135	-0.103	41.713
USDMXN	Mexican peso	0.012	0.691	-6.535	7.880	0.670	16.765
USDMYR	Malaysian ringgit	0.003	0.343	-3.657	2.026	-0.435	12.618
USDPEN	Peruvian sol	0.001	0.296	-3.518	3.768	0.220	24.009
USDPHP	Philippine piso	0.005	0.475	-15.125	3.975	-7.759	256.658
USDRUB	Russian ruble	0.027	0.782	-12.877	12.397	0.785	51.432
USDTHB	Thai baht	0.000	1.012	-12.961	12.477	-0.090	75.692
USDTRY	Turkish lira	0.052	1.276	-25.131	35.667	5.648	233.897
USDTWD	New Taiwan dollar	0.000	0.282	-2.753	1.983	-0.078	10.509
USDVND	Vietnamese dong	0.011	0.215	-4.671	6.075	7.167	286.245
USDZAR	South African rand	0.022	1.134	-11.033	16.213	0.792	18.419
USDPLN	Polish zloty	0.003	0.915	-6.071	7.898	0.408	8.113
USDRON	Romanian leu	0.031	0.764	-8.586	7.709	0.326	14.879
USDHUF	Hungarian forint	0.007	0.950	-4.965	7.911	0.405	6.914
USDCZK	Czech koruna	-0.005	0.816	-5.191	5.445	0.146	6.378
VIX	VIX volatility index	20.886	8.703	9.890	80.860	2.013	9.724

Note: the table presents the mean, standard deviation (Std), minimum value (Min), maximum value (Max), skewness (Skew), kurtosis (Kurt) of currency returns and the VIX volatility index for the full sample (4226 daily observations for the period from January 1999 to December 2015).

Table 2
Transition matrix calculated on the full sample (in %)

	Risk-on	**Neutral**	**Risk-off**
Risk-on	91.6	8.4	0.0
Neutral	4.2	91.6	4.2
Risk-off	0.0	8.6	91.4

Table 3
Cumulative probabilities of transition calculated on the full sample (in %)

	Risk-on	**Neutral**	**Risk-off**
Risk-on	91.9	100.0	100.0
Neutral	4.0	95.8	100.0
Risk-off	0.0	8.4	100.0

Table 4
Tests of equal point forecast accuracy (as measured by the RMSE). Results of the Diebold-Mariano-West (1995, 1996) test and its modification by Harvey, Leybourne and Newbold (1997)

	RMSE (Baseline vs. RW) in %	**DMW (p-value)**	**DMW – Harvey (p-value)**
USDBRL	99.1	0.6	0.6
USDCLP	99.4	0.27	0.27
USDCNY	100.5	0.8	0.8
USDCOP	100.9	0.05	0.06
USDEGP	95.2	0.04	0.04
USDIDR	97.7	0.57	0.57
USDINR	99.3	0.51	0.51
USDKRW	100.7	0.24	0.24
USDMXN	99.5	0.8	0.8
USDMYR	99.2	0.31	0.32
USDPEN	103.2	0.24	0.24
USDPHP	102.2	0.03	0.03
USDRUB	96.9	0.37	0.38
USDTHB	99.5	0.48	0.49
USDTRY	100.7	0.93	0.93
USDTWD	100.1	0.87	0.87
USDVND	97.6	0.41	0.41
USDZAR	96.4	0.02	0.02
USDPLN	100.3	0.35	0.36
USDRON	101.4	0.68	0.68
USDHUF	100.0	0.99	0.99
USDCZK	100.6	0.19	0.19

Note: the first column shows the RMSE of the baseline model in relation to the RMSE of the random walk forecast. The shaded cells indicate exchange rates for which the null hypothesis was not rejected at the 5% significance level.

Table 5

Full density coverage rate (70%) – percentage of actual outcomes inside the 70% interval band

	70% coverage rate (p-value)	
USDBRL	0.68	(0.73)
USDCLP	0.63	(0.2)
USDCNY	0.56	(0.02)
USDCOP	0.69	(0.92)
USDEGP	0.85	(0)
USDIDR	0.93	(0)
USDINR	0.63	(0.2)
USDKRW	0.81	(0.03)
USDMXN	0.75	(0.33)
USDMYR	0.57	(0.03)
USDPEN	0.76	(0.21)
USDPHP	0.72	(0.68)
USDRUB	0.57	(0.03)
USDTHB	0.99	(0)
USDTRY	0.88	(0)
USDTWD	0.67	(0.55)
USDVND	0.74	(0.49)
USDZAR	0.79	(0.06)
USDPLN	0.74	(0.49)
USDRON	0.69	(0.92)
USDHUF	0.71	(0.88)
USDCZK	0.75	(0.33)

Note: the table includes in parentheses the p-values for the null of correct coverage (empirical rate = theoretical rate of 70%), based on t-statistics using standard errors computed with the Newey-West estimator. The shaded cells indicate exchange rates for which the null hypothesis was not rejected at the 5% significance level.

Table 6
Berkowitz (2001) density test

	Berkowitz test p-value
USDBRL	0.46
USDCLP	0.48
USDCNY	0.06
USDCOP	0.41
USDEGP	0.00
USDIDR	0.00
USDINR	0.32
USDKRW	0.04
USDMXN	0.47
USDMYR	0.89
USDPEN	0.08
USDPHP	0.46
USDRUB	0.09
USDTHB	0.00
USDTRY	0.00
USDTWD	0.85
USDVND	0.23
USDZAR	0.01
USDPLN	0.19
USDRON	0.49
USDHUF	0.34
USDCZK	0.46

Note: the shaded cells indicate exchange rates for which the null hypothesis was not rejected at the 5% significance level.

Table 7
Kupiec (1995), Christofferesn (1998) and VQR (2011) backtests for selected percentiles

	tau = 0.1			**tau = 0.9**		
	Kupiec	**Christoffersen**	**VQR**	**Kupiec**	**Christoffersen**	**VQR**
USDBRL	0.36	0.09	0.02	0.63	0.17	0.41
USDCLP	0.30	0.23	0.55	0.16	0.05	0.24
USDCNY	0.02	0.05	0.05	0.04	0.00	0.00
USDCOP	0.17	0.04	0.03	0.08	0.05	0.04
USDEGP	0.02	0.00	0.00	0.76	0.13	0.06
USDIDR	0.00	0.00	0.00	0.00	0.00	0.00
USDINR	0.76	0.40	0.48	0.04	0.05	0.04
USDKRW	0.02	0.00	0.00	0.02	0.00	0.00
USDMXN	0.36	0.09	0.46	0.63	0.10	0.81
USDMYR	0.94	0.12	0.83	0.16	0.33	0.33
USDPEN	0.17	0.04	0.00	0.94	0.29	0.58
USDPHP	0.94	0.12	0.04	0.49	0.51	0.09
USDRUB	0.01	0.01	0.01	0.30	0.04	0.01
USDTHB	0.00	0.00	0.00	0.00	0.00	0.00
USDTRY	0.36	0.09	0.26	0.00	0.00	0.00
USDTWD	0.49	0.12	0.25	0.49	0.16	0.49
USDVND	0.17	0.04	0.02	0.63	0.10	0.01
USDZAR	0.06	0.01	0.00	0.36	0.46	0.45
USDPLN	0.94	0.28	0.13	0.36	0.09	0.49
USDRON	0.76	0.40	0.97	0.94	0.28	0.82
USDHUF	0.36	0.09	0.37	0.94	0.28	0.08
USDCZK	0.36	0.09	0.01	0.49	0.12	0.23

Notes:
The table presents p-values for the Kupiec (1995), Christoffersen (1998) and VQR (2011) test. Shaded cells indicate results for which the null is not rejected at the 5% significance level.

Table 8
On the left side: the rank of models based on LPDS. On the right side: the results (p-values) of the Amisano-Giacomini (2007) test

	Base-line	RW	TAR benchmarks				GARCH benchmarks				Amisano-Giacomini test		
			B1	B2	B3	B4	B5	B6	B7	B8	Baseline vs. best TAR	Baseline vs. RW	Baseline vs. best GARCH
USDBRL	2	4	1	5	8	7	9	10	6	3	Baseline vs. B1 (0.794)	Baseline vs. RW (0.477)	Baseline vs. B8 (0.623)
USDCLP	1	2	6	5	8	7	10	9	4	3	**Baseline vs. B2 (0.004)**	Baseline vs. RW (0.501)	Baseline vs. B8 (0.113)
USDCNY	6	7	10	9	8	5	3	2	4	1	Baseline vs. B4 (0.615)	Baseline vs. RW (0.765)	Baseline vs. B8 (0.103)
USDCOP	1	6	5	4	8	7	9	10	3	2	Baseline vs. B2 (0.319)	**Baseline vs. RW (0.023)**	Baseline vs. B8 (0.781)
USDEGP	1	4	6	7	8	9	5	10	3	2	**Baseline vs. B1 (0.035)**	Baseline vs. RW (0.056)	Baseline vs. B8 (0.325)
USDIDR	5	7	3	4	6	8	9	10	2	1	Baseline vs. B1 (0.303)	**Baseline vs. RW (0)**	Baseline vs. B8 (0)
USDINR	3	5	7	8	10	6	9	4	2	1	Baseline vs. B4 (0.062)	Baseline vs. RW (0.175)	Baseline vs. B8 (0.039)
USDKRW	1	4	7	6	5	8	10	9	3	2	**Baseline vs. B3 (0)**	**Baseline vs. RW (0)**	Baseline vs. B8 (0.442)
USDMXN	1	3	5	4	7	8	9	10	6	2	Baseline vs. B2 (0.125)	Baseline vs. RW (0.235)	Baseline vs. B8 (0.48)
USDMYR	1	2	9	6	10	8	7	4	3	5	**Baseline vs. B2 (0.032)**	Baseline vs. RW (0.553)	Baseline vs. B7 (0.201)
USDPEN	3	4	8	7	6	5	9	10	2	1	**Baseline vs. B4 (0)**	Baseline vs. RW (0.092)	Baseline vs. B8 (0.122)
USDPHP	1	4	6	5	7	8	10	9	2	3	Baseline vs. B2 (0.212)	**Baseline vs. RW (0.002)**	Baseline vs. B7 (0.933)
USDRUB	1	4	9	7	10	8	5	6	3	2	**Baseline vs. B2 (0.001)**	**Baseline vs. RW (0.011)**	Baseline vs. B8 (0.783)
USDTHB	5	6	1	2	7	10	9	8	4	3	Baseline vs. B1 (0)	**Baseline vs. RW (0)**	Baseline vs. B8 (0.001)
USDTRY	4	2	5	6	7	8	10	9	3	1	Baseline vs. B1 (0.251)	Baseline vs. RW (0.598)	Baseline vs. B8 (0.245)
USDTWD	1	3	5	6	8	7	9	10	4	2	**Baseline vs. B1 (0.018)**	Baseline vs. RW (0.815)	Baseline vs. B8 (0.954)
USDVND	1	3	8	7	5	10	9	6	4	2	**Baseline vs. B3 (0.003)**	**Baseline vs. RW (0.007)**	Baseline vs. B8 (0.259)
USDZAR	1	6	4	3	7	8	10	9	5	2	**Baseline vs. B2 (0.005)**	Baseline vs. RW (0.175)	Baseline vs. B8 (0.598)
USDPLN	2	6	3	4	7	8	10	9	5	1	Baseline vs. B1 (0.673)	Baseline vs. RW (0.21)	Baseline vs. B8 (0.427)
USDRON	1	4	6	5	8	7	10	9	3	2	**Baseline vs. B2 (0.01)**	Baseline vs. RW (0.069)	Baseline vs. B8 (0.168)
USDHUF	1	4	5	6	7	8	10	9	3	2	Baseline vs. B1 (0.102)	Baseline vs. RW (0.243)	Baseline vs. B8 (0.703)
USDCZK	1	2	6	5	7	8	10	9	4	3	**Baseline vs. B2 (0.017)**	Baseline vs. RW (0.656)	Baseline vs. B8 (0.395)

Notes:
On the left side, the best model according to the LPDS rank ordering (i.e. from higher to lower LPDS figures) is highlighted in grey for each exchange rate. On the right side, the baseline model is compared with three different benchmarks using the Amisano-Giacomini test: with the best threshold autoregressive (TAR) model for a given exchange rate, the random walk forecast and the best AR-GARCH benchmark. Grey cells indicate that the baseline model is statistically (at 5%) better than the corresponding benchmark; black cells indicate that the baseline model performs statistically worse than the corresponding benchmark. White cells indicate no statistically significant difference between the density forecasts of the baseline model and the corresponding benchmark.

Figure 1
Historical values of VIX with regime segmentation

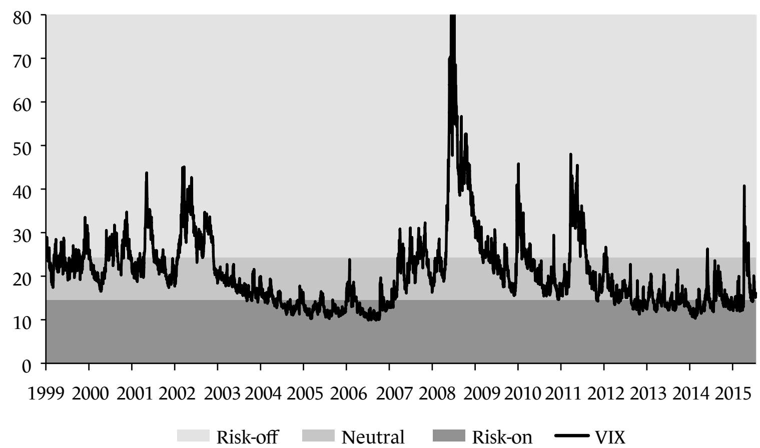

Figure 2
Histogram of USDPLN daily logarithmic returns