JEL classification: F31, G15
Keywords: forecasting volatility, foreign exchange, combination forecasts, Central and Eastern Europe, multivariate HAR

Forecasting Exchange Rate Volatility: The Case of the Czech Republic, Hungary and Poland*

Štefan LYÓCSA—University of Economics, Institute of Economics and Management, Bratislava; University of Prešov, Faculty of Management (stefan.lyocsa@gmail.com), Prešov, Slovakia, *corresponding author*

Peter MOLNÁR—University of Stavanger, UiS Business School, Stavanger, Norway (peto.molnar@gmail.com)

Igor FEDORKO—University of Prešov, Faculty of Management, Prešov, Slovakia (igor.fedorko@gmail.com)

Abstract

We study various models for forecasting one-day forward volatility of the exchange rates of the Czech koruna, Hungarian forint and Polish zloty against the euro. We used high-frequency data to calculate realized volatility. We found that our benchmark model, the heterogeneous autoregressive (HAR) model of Corsi (2009) is rarely out-performed, even if we extend the standard HAR model by including signed jumps or substituting continuous and jump components, or if we allow the autoregressive parameter of the HAR model to vary with the estimated degree of the measurement error (Bollerslev et al., 2016). Our results suggest that the preferred forecasting strategy is to average univariate forecasts, as these combination forecasts offer improvements upon the benchmark (CZK/EUR, PLZ/EUR) or do not lead to worse forecasts (HUF/EUR). Extensions of the HAR models with regional and global exchange rate volatilities and multivariate HAR models which also model covariance between exchange rates (Baruník and Čech, 2016) have usually performed worse than the benchmark. Therefore, our study offers little evidence of volatility spillovers, an exception is spillovers from USD/EUR to CZK/EUR and PLZ/EUR and from HUF/EUR to CZK/EUR and from CHF/EUR to PLZ/EUR.

1. Introduction

Understanding and forecasting volatility comprise one of the key topics in economics and finance. The volatility of exchange rates is inevitably related to international trade (e.g. Arize *et al.*, 2000), investments (e.g. Chowdhury, 1993; Caporale *et al.*, 2015), monetary policy (e.g. Gali and Monacelli, 2005) and productivity growth (e.g. Aghion *et al.*, 2009). Apart from the macroeconomic and monetary perspectives, exchange rate volatility is of interest in the financial literature. If investors are unable to defend themselves against adverse movements on the foreign exchange market or if such hedging strategies are costly, an asset pricing perspective dictates that exchange rate volatility is an additional risk factor. Such risks increase the required returns from foreign investments. Altogether, foreign exchange volatility predictions are important for various entities, including investors, traders, banks and policymakers.

* Štefan Lyócsa appreciates the support provided by the Slovak Grant Agency under Grant No. 1/0392/15 and Grant No. 1/0406/17. This work was supported by the Slovak Research and Development Agency under Contract No. APVV-14-0357.

Volatility forecasting is particularly important for the currencies of small developing countries. International trade and foreign investments play particularly important roles in such countries, which usually have more volatile exchange rates. However, even though volatility forecasting has been studied extensively for major exchange rates, much less research has been done for emerging markets, particularly Central and Eastern European (CEE) countries.

We study volatility forecasting for the exchange rates of the Czech koruna (CZK), Hungarian forint (HUF) and Polish zloty (PLZ) against the euro (EUR) using the heterogeneous autoregressive (HAR) class of univariate and multivariate models. More specifically, by using volatility forecasting models we address two questions: 1. What type of one-day forward volatility forecasts (within our set of models) is to be preferred and 2. are exchange rates subject to volatility spillovers in an out-of-sample framework? The answers to these questions might be important for investors and policymakers, as forecasts with lower one-day forward forecasting errors are to be preferred because they provide higher forecasting accuracy, while knowledge about possible volatility spillovers offers insight into whether exchange rate volatility should be studied within a basket of exchange rates or if it suffices to study the exchange rate volatility of individual currencies.

We found that, in general, it is difficult to beat the standard univariate HAR model of Corsi (2009). Based on our results, it also appears that it is a good practice to combine forecasts from different univariate models, as such combination forecasts often lead to better forecasts (CZK/EUR, PLZ/EUR) or not worse forecasts (HUF/EUR) than those from standard HAR models. Forecasters facing model choice uncertainty might therefore benefit from simply averaging forecasts from simple univariate models.

In an out-of-sample framework, our results provided only limited evidence of volatility spillovers within CEE exchange rates or from exchange rates of developed countries to CEE markets. Among 66 models designed to capture volatility spillovers, volatility spillovers were found only from USD/EUR to CZK/EUR and PLZ/EUR, from HUF/EUR to CZK/EUR and from CHF/EUR to PLZ/EUR.

Even though there are several studies that model the volatility of the exchange rates of CEE countries (e.g. Kočenda and Valachy, 2006; Fidrmuc and Horváth, 2008; Schnabl, 2008; Schnabl, 2009; Cuaresma *et al.*, 2010; Arratibel *et al.*, 2011; and Bubák *et al.*, 2011) these papers model volatility for various purposes and, with the exception of Cuaresma *et al.* (2010), none of them studies volatility forecasting. We are the first to utilize realized volatility and other related variables in several state-of-the-art volatility models. The models we use are based on the HAR model of Corsi (2009), which performs well in comparison with more advanced models, yet it is easy to estimate.[1] Cuaresma *et al.* (2010) have attempted to outperform random-walk prediction using models based on foreign exchange options but failed to do so, thus confirming the finding of Meese and Rogoff (1983) that the short-run predictive quality of the random-walk model is superior to relevant economic models. Although it was not our intention to compare the HAR model of Corsi (2009) with random-

[1] Such models have been applied to major stock markets (Christoffersen *et al.*, 2010), major exchange rates (Andresen *et al.*, 2001) and recently also to commodities (Haugom *et al.*, 2014; Birkelund *et al.*, 2015).

walk predictions, we performed such a comparison as a robustness check and, in our study and sample, HAR models were much more successful, as they easily outperform the random-walk model and therefore we set our benchmark higher to the HAR model instead of the random-walk model.

In addition, all of the previous studies except for Bubák *et al.* (2011) are based on daily data. It is known that the most accurate volatility models are based on the concept of realized volatility calculated from high-frequency data, introduced by Andersen and Bollerslev (1998). Bubák *et al.* (2011) focus primarily on volatility transmission, whereas our contribution consists in studying volatility for individual markets and volatility transmission in an out-of-sample framework.

The rest of the paper is organized as follows: In Section 2 we describe the forecasting models used in our study and the forecasting evaluation methods. Data and filtration techniques used to process the high-frequency data are presented in Section 3. Section 4 presents the results and Section 5 concludes the paper.

2. Methodology

Our main quantity of interest is the square root of the realized variance (RV). There are several reasons why we use an estimator of RV, which simply sums squared returns, instead of some more sophisticated alternative. First, there is no consensus in the literature about which estimator is the best, and this simple estimator is the most commonly used. Second, we calculate also positive and negative semi-variance (as explained later), and positive and negative semi-variance adds together to RV only if we calculate RV using a simple estimator. Third, RV is also the preferred choice in out-of-sample comparisons (see Hansen and Lunde, 2006). The realized variance is given by

$$RV_t = \sum_{i=2}^{f} r_{ti}^2 \tag{1}$$

where r is the intraday return and, given a specific sampling frequency, f is the number of intraday prices.

2.1 Univariate Heterogeneous Autoregressive Models

Given the realized variances (RV), our base-line model is the HAR predictive regression model of Corsi (2009):

HAR

$$RV_{t+1}^{0.5} = \beta_1 + \beta_2 RV_t^{0.5} + \beta_3 RV_{t,t-4}^{0.5} + \beta_4 RV_{t,t-21}^{0.5} + e_t \tag{2}$$

where $RV_{t+1}^{0.5}$ is the realized volatility calculated over the trading day t+1, which corresponds to a one-day forward volatility forecast. $RV_t^{0.5}$ is the realized volatility over the previous trading day t, $RV_{t,t-4}^{0.5}$ is the realized volatility averaged across the previous five trading days (t, t–1, t–2, t–3, and t–4) and, similarly, $RV_{t,t-21}^{0.5}$ is the realized volatility averaged across the previous 22 trading days. We are interested in forecasting realized volatility (the square root of realized variance) instead of realized variance for two reasons. First, we expect that linear forecasting models will have better predictive performance with respect to the square root of variance. If there are some extreme observations, the estimation results will be less influenced by them if we study the square root of variance. Second, the GHAR model (to be defined later) is based on modeling elements of a Cholesky factorization from a variance-covariance matrix, which are of similar scale to the realized volatilities.

The HAR model in (2) is our benchmark model instead of the recommended random-walk model of Meese and Rogoff (1983), as the HAR model outperforms the random-walk predictions by a large margin,[2] thus not supporting the stylized finding of Meese and Rogoff (1983) that short-term volatility forecasts are best arrived at with random-walk models.

HAR-SJ

Several recent studies have emphasized the importance of disentanglement of volatility into positive and negative semi-variances (e.g. Barndorff-Nielsen *et al.*, 2010; Patton and Sheppard, 2015; Sévi, 2014), as they might have different predictive content with regard to future variance. It may be argued that negative semi-variances (variances during market declines) tend to be more clustered than positive semi-variances (variances during upward markets), which might lead to improved forecasting when predictive regressions utilize information on positive and negative semi-variances. Let's define the positive and negative semi-variances as

$$\begin{aligned} RV_t^- &= \sum_{i=2}^{f} r_{ti}^2 \times I\left[r_{ti} < 0\right] \\ RV_t^+ &= \sum_{i=2}^{f} r_{ti}^2 \times I\left[r_{ti} \geq 0\right] \end{aligned} \tag{3}$$

where $I[.]$ is the indicator function returning 1 if the condition holds or 0 otherwise. Furthermore, Patton and Sheppard (2015) have shown that given the positive and negative semi-variances, the continuous component of volatility can be removed by subtracting the realized semi-variances. The remaining part is called the signed jump:

$$SJ_t^{0.5} = \begin{cases} +\sqrt{RV_t^+ - RV_t^-}, & RV_t^+ - RV_t^- > 0 \\ -\sqrt{RV_t^+ - RV_t^-}, & RV_t^+ - RV_t^- < 0 \end{cases} \tag{4}$$

Adding signed jumps into Eq. (2) leads to the following specification:

$$RV_{t+1}^{0.5} = \beta_1 + \beta_2 RV_t^{0.5} + \beta_3 SJ_t^{0.5} + \beta_4 RV_{t,t-4}^{0.5} + \beta_5 SJ_{t,t-4}^{0.5} + \beta_6 RV_{t,t-21}^{0.5} + \beta_7 SJ_{t,t-21}^{0.5} + e_t \tag{5}$$

In Eq. (5), the signed jump $SJ_{t,t-4}$ corresponds to the average of signed jumps calculated for the given days t, $t-1$, $t-2$, $t-3$ and $t-4$. The same principle was used for $SJ_{t,t-21}$.

[2] We do not elaborate on these results any further, as the results were very convincing and currently the literature is more focused on HAR models than on random-walk models. For CZK/EUR, the MSFE, QLIKE and MAFE (see Section 2.4 for descriptions of these) were 6.40, 5.72 and 5.79, respectively. For HUF/EUR, the MSFE, QLIKE and MAFE were 20.09, 6.03 and 10.32, respectively. For PLZ/EUR, the MSFE, QLIKE and MAFE were 11.46, 4.94 and 7.35, respectively. All these values are clearly above those produced with the univariate and multivariate HAR models.

HAR-CJ

Since the seminal paper of Barndorff-Nielsen and Shephard (2004), it is acknowledged that the log price process is a combination of a continuous price movement and a jump component, which can be understood as a discontinuity in the price process (Žikeš and Baruník, 2014). Both components can have different predictive power with regard to future variance; more specifically, price discontinuities might lead to overestimation of the persistence of volatility (e.g. Andersen *et al.*, 2007), thus leading to sub-optimal forecasts. Barndorff-Nielsen and Shephard (2004) proposed a consistent estimator of variance due to the continuous price movement, the so-called *bi-power variation*, but in this study we will use the *median realized variance* of Andersen *et al.* (2012), which has more desirable finite-sample properties in the presence of small returns and jumps that are likely to occur for emerging foreign exchange markets:

$$MRV_{t,f} = \frac{\pi}{6-4\sqrt{3}+\pi}\left(\frac{f}{f-2}\right)\sum_{j=2}^{f-1}\left[median\left\{\left|r_{t,j-1}\right|,\left|r_{t,j}\right|,\left|r_{t,j+1}\right|\right\}\right]^2 \tag{6}$$

The estimate of the variance movement attributed to the jump component JC_t can be defined as the difference between the realized variance and the median realized-variance but, as noted by Beine *et al.* (2007), it might be advisable to test for the presence of jumps first, as one might consider small jumps to be a part of the continuous sample path rather than genuine discontinuities. Furthermore, as in finite samples the difference between the realized variance and the median realized variance might be negative, it seems reasonable to place a lower boundary on the jump component (Barndorff-Nielsen and Shephard, 2004). Therefore, our estimate of is

$$JC_t = \max\left\{0,\left(RV_t - MRV_t\right)\times I\left[JT_t > \varphi_{1-\alpha}\right]\right\} \tag{7}$$

where JT_t is the test statistics as defined by Andersen *et al.* (2012):

$$JT_t = \frac{\sqrt{f}\left(RV_t - MRV_t\right)RV_t^{-1}}{\left(0.96\max\left\{1, MRQ_t / RV_t^2\right\}\right)^{1/2}} \tag{8}$$

$\varphi_{1-\alpha}$ is the critical value of the standard normal distribution, and MRQ_t is the median realized-quarticity:

$$MRQ_{t,f} = \frac{3\pi f}{9\pi+72-52\sqrt{3}}\left(\frac{f}{f-2}\right)\sum_{j=2}^{f-1}\left[med\left\{\left|r_{t,j-1}\right|,\left|r_{t,j}\right|,\left|r_{t,j+1}\right|\right\}\right]^4 \tag{9}$$

The continuous component of market volatility is then

$$CC_t = MRV_t \times I\left[JT_t > \varphi_{1-\alpha}\right] + RV \times I\left[JT_t \le \varphi_{1-\alpha}\right] \tag{10}$$

Substituting jump and continuous components into (2) leads to

$$\begin{aligned} RV_{t+1}^{0.5} = {} & \beta_1 + \beta_2 JC_t^{0.5} + \beta_3 CC_t^{0.5} + \beta_4 JC_{t,t-4}^{0.5} + \beta_5 CC_{t,t-4}^{0.5} + \\ & + \beta_6 JC_{t,t-21}^{0.5} + \beta_7 CC_{t,t-21}^{0.5} + e_t \end{aligned} \tag{11}$$

We use specifications (2), (5) and (11) in our study to observe whether taking into account realized semi-variances or disentanglement of the price process into its jump and continuous components increases the predictive power of HAR models when predicting volatility on foreign exchange markets. Specification (2) will be considered to be the benchmark model to which other specifications will be compared.

HAR-Q

In finite samples, the realized variance is subject to a measurement error. Instead of considering the measurement error to be constant through time, Bollerslev *et al.* (2016) have recently proposed taking time variation of the measurement error into account when modeling the realized variance. Their idea was to give higher weight to autoregressive terms if the variance of the measurement error is small and *vice versa*. We use a slightly modified HAR-Q model of Bollerslev *et al.* (2016), which meets our requirement to model the realized volatility:

$$RV_{t+1}^{0.5} = \beta_1 + \beta_2 RV_t^{0.5} + \beta_3 \left(RQ_t^{0.5} RV_t \right)^{0.5} + \beta_4 RV_{t,t-4}^{0.5} + \beta_5 RV_{t,t-21}^{0.5} + e_t \tag{12}$$

where RQ_t is the realized quarticity defined as

$$RQ_t = \frac{f}{3} \sum_{i=2}^{f} r_{ti}^4 \tag{13}$$

The realized volatility $RV_t^{0.5}$ in (12) can be separated as $(\beta_2 + \beta_3 RQ_t^{0.5}) RV_t^{0.5}$, which shows how the realized quarticity has time-varying influence on the weight given to past realized volatilities, which is the central idea of Bollerslev *et al.* (2016).

HAR-X, HAR-SJ-X, HAR-CJ-X, HAR-Q-X

Until now, we have ignored linkages between currency pairs. For example, investors might consider emerging Central and Eastern European markets as a single investment region. This might lead to increased co-movement between exchange rates and volatility transmissions. Therefore, it might be advantageous to incorporate such information into predictive models. Furthermore, the overall uncertainty on the main currency pairs (EUR/USD and EUR/CHF) could transfer into the foreign currency markets of CEE as well. We have therefore expanded models (2), (5), (11) and (12) by including realized volatilities on regional and developed foreign exchange markets. This inclusion allows us to assess whether short-term information from other foreign exchange markets are useful in predicting the volatility of the local currency pair.[3]

For each of the three emerging-market foreign exchange rate volatilities, this leads to specifications in the following form:

$$RV_{t+1}^{0.5} = \beta_1 + \beta_2 RV_t^{0.5} + \beta_3 RVX_t^{0.5} + \beta_4 RV_{t,t-4}^{0.5} + \beta_5 RV_{t,t-21}^{0.5} + e_t \tag{14}$$

$$\begin{aligned} RV_{t+1}^{0.5} = \beta_1 + \beta_2 RV_t^{0.5} + \beta_3 SJ_t^{0.5} + \beta_4 RVX_t^{0.5} + \beta_5 SJX_t^{0.5} + \beta_6 RV_{t,t-4}^{0.5} + \\ + \beta_7 SJ_{t,t-4}^{0.5} + \beta_8 RV_{t,t-21}^{0.5} + \beta_9 SJ_{t,t-21}^{0.5} + e_t \end{aligned} \tag{15}$$

[3] Inclusion of these variables and comparison of the performance with the baseline model essentially comprise a Granger causality test.

$$RV_{t+1}^{0.5} = \beta_1 + \beta_2 JC_t^{0.5} + \beta_3 CC_t^{0.5} + \beta_3 JCX_t^{0.5} + \beta_4 CCX_t^{0.5} + \beta_5 JC_{t,t-4}^{0.5} + \\ + \beta_6 CC_{t,t-4}^{0.5} + \beta_7 JC_{t,t-21}^{0.5} + \beta_8 CC_{t,t-21}^{0.5} + e_t \tag{16}$$

$$RV_{t+1}^{0.5} = \beta_1 + \beta_2 RV_t^{0.5} + \beta_3 \left(RQ_t^{0.5} RV_t \right)^{0.5} + \beta_4 RVX_t^{0.5} + \\ + \beta_5 \left(RQX_t^{0.5} RVX_t \right)^{0.5} + \beta_6 RV_{t,t-4}^{0.5} + \beta_7 RV_{t,t-21}^{0.5} + e_t \tag{17}$$

where RVX_t, SJX_t, JCX_t, CCX_t, RQX_t denote either one of the two remaining emerging foreign exchange markets or one of the two developed markets.[4] For each of the three currency pairs, we have 20 univariate models.

To produce a volatility forecast, a given predictive regression is estimated using a fixed estimation window of a length $w = 180$ *via* OLS. The volatility is predicted by substituting estimated coefficients into the given specification and the last known values of the right-hand side variables. Forecast of a realized volatility of a given model $m = 1, 2, \ldots, 20$ estimated within an estimation window of length 180 days will be denoted as $RV_{t+1}{}^{*}$. Next forecast is produced by rolling the estimation window one observation ahead, estimating the predictive regressions all over again, and substituting the right-hand side coefficients and variables with the last known values.

2.2 Multivariate Heterogeneous Autoregressive Models

A series of recent papers have proposed extensions of the univariate HAR model to multivariate settings; see, for example, Chiriac and Voev (2011), Bauer and Vorking (2011), Fengler and Gisler (2015), and Baruník and Čech (2016). In cases when realized volatilities across different currency pairs are highly correlated, adding lagged realized volatilities from another currency pair into univariate forecasting regressions can lead to noisier forecasts. We decided to follow the work of Baruník and Čech (2016), as their generalized HAR (GHAR) model is based on modeling elements of Cholesky factors of a suitable variance-covariance matrix within a seemingly unrelated regression (SUR) framework. The estimation feature of GHAR makes it particularly interesting, as common factor(s) driving foreign exchange volatility can be exhibited through the dependence in error terms.

For each exchange rate, we have considered four GHAR-2 models and two GHAR-3 models, where the numeric figure denotes the number of exchange rates in the system. For example, the modeling of CZK/EUR included the following four GHAR-2 models: {CZK/EUR, PLZ/EUR}, {CZK/EUR, HUF/EUR}, {CZK/EUR, CHF/EUR} and {CZK/EUR, USD/EUR}, and the following two GHAR-3 models: {CZK/EUR, PLZ/EUR, HUF/EUR}, {CZK/EUR, CHF/EUR, USD/EUR}.

We first define the estimators of the variance-covariance matrix, and then estimation procedure. The standard approach for estimating the variance-covariance matrix is the realized variance-covariance matrix:

[4] For example, if we predict the realized volatility of the Czech koruna using the specification Eq. (14), then one model will be estimated with RVX_t being the HUF/EUR realized volatility, another model where RVX_t is the PLZ/EUR realized volatility, another model with RVX_t being the CHF/EUR realized volatility and finally we estimate one model where RVX_t belong to the USD/EUR realized volatility.

$$\hat{\mathbf{\Sigma}}_t = \sum_{j=1}^{N} \mathbf{r}_{tj} \mathbf{r}_{tj}^T \tag{18}$$

where $\mathbf{r}_{t,j}$ is the column vector of the intraday returns of a given set of exchange rates. In the presence of microstructure noise, this estimator becomes inconsistent and it is not guaranteed that the resulting variance-covariance matrix will be positive and semi-definite (Barndorff-Nielsen and Shephard, 2004). Therefore, we have additionally used the multivariate realized kernel (*MRK*) estimator of Barndorff-Nielsen *et al.* (2011):

$$\hat{\mathbf{\Sigma}}_t^{MRK} = \sum_{h=-(f-1)}^{(f-1)} k\left(\frac{h}{H}\right) \Gamma_{t,h} \tag{19}$$

where

$$\Gamma_{t,h} = \sum_{j=h+1}^{f-1} \mathbf{r}_{t,j} \mathbf{r}_{t,j-h}^{\mathrm{T}} \quad ,h \geq 0 \tag{20}$$

while for $h < 0$, $\Gamma_{t,h} = \Gamma_{t,-h}^{\ T}$, and $k(.)$ is the kernel weighting function, where we use the recommended Parzen scheme. The value of the bandwidth parameter H was selected from a set {2,3,…,8} that led to the lowest mean squared forecast error in the previous 30 out-of-sample observations. Alternatively, if the mean squared error over the previous 30 out-of-sample observations was lower with the realized variance-covariance matrix as defined in (18), we employed that instead of the multivariate realized kernel estimator (19).

If we let q be the number of exchange rates in the system, then $\hat{\mathbf{\Sigma}}_t^{MRK}$ is the q x q matrix that represents the estimated variance-covariances. To ensure that the forecasted volatilities will remain positive, Baruník and Čech (2016) modeled the elements of the Cholesky factor. More specifically, given a Cholesky factorization $\mathbf{P}_t \mathbf{P}_t^T = \hat{\mathbf{\Sigma}}_t^{MRK}$ and the following column vector with $m = q(q-1)/2 + q$ elements:

$$\mathbf{X}_t = vech\left(\mathbf{P}_t\right) \tag{21}$$

where vech(.) corresponds to lower triangular elements of $\mathbf{P}_t$, we are interested in modeling the resulting elements of $\mathbf{X}_t$ within a system of m equations, i.e. we are modeling not only volatilities but also covariances. The following system of equations is estimated within the SUR framework:

$$\begin{pmatrix} \mathbf{X}_{1,t+1} \\ . \\ \mathbf{X}_{m,t+1} \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_{1,t} & 0 & 0 \\ 0 & . & 0 \\ 0 & 0 & \mathbf{Z}_{m,t} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ . \\ \boldsymbol{\beta}_m \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ . \\ \boldsymbol{\varepsilon}_m \end{pmatrix} \tag{22}$$

where $\mathbf{Z}_{i,t} = (1\ \mathrm{X}_{i,t}\ \mathrm{X}_{i,t,t-4}\ \mathrm{X}_{i,t,t-21})$, $\mathrm{X}_{i,t,t-4}$ ($\mathrm{X}_{i,t,t-21}$) correspond to the average values of the i-th element of $\mathbf{X}_t$ over the previous five (22) trading days, $\boldsymbol{\beta}$ is a column vector of coefficients (including the constant), and $\boldsymbol{\varepsilon}_i$ is the vector of disturbances of the i-th equation.

2.3 Forecast Combination

Given a set of unbiased and weakly correlated forecasts, a suitable forecast combination might lead to better forecasting performance in that the resulting volatility of the combination forecast should be lower than that of individual forecasts (e.g. Bates and Granger, 1969; Timmermann, 2006). Combination of forecasts may offer diversification benefits, as forecasts (i) differ with respect to the information sets when producing forecasts and (ii) they are also based on a different assumption about the underlying data-generating process. Besides presenting the results of individual forecasts, we will also explore whether the simple combination of forecasts across different specifications offers forecasting benefits.

We will use two combination forecasts: the arithmetic mean (for combination of five or fewer forecasts) and the trimmed mean (e.g. Stock and Watson, 2004; Genre *et al.*, 2013), where the lowest 20% and the highest 20% of forecasts are removed from the sample before calculating the combined forecast. We denote combination forecasts in *Tables 2–4* as *CF-Mean* and *CF-Trim*.

2.4 Forecast Evaluation

The performance of individual and combination forecasts will be evaluated with respect to the benchmark model, the standard HAR model or the HAR-X model. The individual and combined forecasts will be evaluated along two lines. First, each model will be compared to the benchmark using the mean squared forecast errors (*MSFE*), the measure advocated by Patton (2011) (*QLIKE*) and the mean absolute forecast error (*MAFE*). Second, possible improvement in forecasting will be statistically evaluated using the test of Hansen *et al.* (2011).

For a given model, the loss function defined via the squared forecast error is given by

$$L^{SFE} = \left(\sqrt{RV^*_{t+1}} - \sqrt{RV_{t+1}}\right)^2 \tag{23}$$

where $\sqrt{RV^*_{t+1}}$ is the forecasted squared root of the realized variance and $\sqrt{RV_{t+1}}$ is the observed square root of the realized variance (our proxy for the true variance; see Hansen and Lunde, 2006). The *QLIKE* loss function of Patton (2011) is given by

$$L^{QLIKE} = \sqrt{RV_{t+1}} \,/\, \sqrt{RV^*_{t+1}} - \ln\left(\sqrt{RV_{t+1}} \,/\, \sqrt{RV^*_{t+1}}\right) - 1 \tag{24}$$

Finally, the absolute forecast error is defined as

$$L^{AFE} = \left|\sqrt{RV^*_{t+1}} - \sqrt{RV_{t+1}}\right| \tag{25}$$

The choice of a preferred loss function depends on the application. For example, in financial risk management applications, extreme forecast errors might receive larger weight, which corresponds to the *SFE* loss function. The *QLIKE* takes into account the heteroskedastic nature of forecast errors, which is often observed in the empirical literature, i.e. during periods of extreme volatility, forecast errors tend to be extreme as well. We will therefore report the results for the mean values for all three loss functions.

To improve the comparison between the benchmark and other forecasting models, we also report relative comparisons based on simple indices. If we denote the benchmark HAR model as B and the competing model as A, then

$$R^{MSFE} = \frac{MSFE_A}{MSFE_B}, R^{QLIKE} = \frac{QLIKE_A}{QLIKE_B}, R^{MAFE} = \frac{MAFE_A}{MAFE_B} \quad (26)$$

If we are interested in whether signed jumps lead to superior forecasts in the HAR-X-SJ model, the forecasts should be evaluated with the corresponding HAR-X model instead of the HAR model alone. Therefore, alongside the HAR model, we also considered HAR-X models to act as benchmarks. In *Tables 2–4* we denote the corresponding relative indices as R_X^{MSFE}, R_X^{QLIKE} and R_X^{MAFE}, which differ from (26) only in that the benchmark model in the denominator is the corresponding HAR-X model instead of the HAR model.

Let $d_{A,B,t+1}$ denote the loss differential:

$$d_{A,B,t+1} = L_A - L_B \quad (27)$$

The hypothesis under interest is of the form

$$\begin{aligned} H_0 &: \; E\left[d_{A,B}\right] = 0 \\ H_1 &: \; E\left[d_{A,B}\right] \neq 0 \end{aligned} \quad (28)$$

Now let

$$\bar{d}_{A,B} = \left(T - T_1 + 1\right)^{-1} \sum_{T_1 \le t \le T} d_{A,B,t+1} \quad (29)$$

where T_1 denotes the time index of the first forecast and T the last forecast, i.e. (29) denotes the average loss differential. As within our empirical application we compare always two models (the competing model and the benchmark model), our test statistics are of a similar form as the well-known Diebold and Mariano (1995) test statistics:

$$t_{A,B} = \frac{\bar{d}_{A,B}}{\sqrt{\hat{v}\left(\bar{d}_{A,B}\right)}} \quad (30)$$

However, $\hat{v}\left(\bar{d}_{A,B}\right)$ is the block-bootstrap estimate of the variance of $\bar{d}_{A,B}$. The block length p was set to the maximum number of significant parameters of an $AR(p)$ model fitting the $d_{A,B,t+1}$ series (see Hansen *et al.*, 2011) and the number of resamples was set to 5,000. The distribution of the test statistics under the null hypothesis was again bootstrapped using a similar bootstrap as before.[5]

[5] The test of Hansen *et al.* (2011) was performed using the procedures developed in Bernardi and Catania (2014) for program R.

3. Data

We use high-frequency foreign exchange rates covering a sample period from January 2011 until December 2014 available from GAIN Capital.[6] Our data set consists of three emerging foreign exchange rates from Central and Eastern Europe, CZK/EUR PLZ/EUR and HUF/EUR, and two exchange rates from developed markets, USD/EUR and CHF/EUR. As there are no trading hour limitations on the foreign exchange market, we made an arbitrary choice of setting a rather larger trading window, which starts from 8:00 a.m. central European summer time (CEST) and ends at 10:00 p.m. The corresponding time frame in New York is from 3:00 a.m. and 5:00 p.m. A trading window of fourteen hours should cover the most relevant market activity in both Europe and the US.

Our initial dataset has 3.93 million (CZK/EUR), 9.95 million (PLZ/EUR) and 7.67 million (HUF/EUR) observations of bid and ask tick data. Such large datasets might be prone to human and technical data error. The techniques of Müller *et al.* (1990), Dacorogna *et al.* (1993) and Dacorogna *et al.* (2001) are applied to filter the series with the aim of removing suspicious, potentially outlying observations. Let $P^b_j = \ln(x^b_j)$ denote the logarithm of the tested bid price with index j, and P^b_{j*} the last valid logarithm of a bid price, while $s_j = \ln(x^b_j/x^a_j)$ is the logarithm of the spread (x^a_j is the ask price), s_{j*} is the logarithm of the bid-ask spread of the last valid price and Δt_{jj*} is the time difference between the j-th price and the last valid price j^* measured in days. The first set of filters is as follows:

The bid price with index j is valid if

$$\left|P^b_j - P^b_{j*}\right| < 0.4 \quad \wedge \quad \left|P^b_j - P^b_{j*}\right| < 2.2 s_{j*} + 0.27 \Delta t_{jj*} \tag{31}$$

Similarly, for the ask price we have the condition

$$\left|P^a_j - P^a_{j*}\right| < 0.4 \quad \wedge \quad \left|P^a_j - P^a_{j*}\right| < 2.2 s_{j*} + 0.27 \Delta t_{jj*} \tag{32}$$

But we retained only such observations that also satisfied the following two conditions from the spread filter of Dacorogna *et al.* (2001):

The bid and ask prices with index j are valid if

$$\left|\ln \frac{s_j}{s_{j*}}\right| < 1.5 + 75 \Delta t_{jj*} \quad \wedge \quad \left|\ln \frac{s_j}{s_{j*}}\right| < 5.5 \tag{33}$$

The resulting data were used to create our dataset of prices where for each day we have sampled data with a ten-minute data sampling frequency, which corresponds to $f = 84$ price observations. With a higher sampling frequency, the number of missing data points (particularly at the tails of our trading window) gets larger, as within 14 trading hours there are periods without any activity.[7]

[6] http://ratedata.gaincapital.com/

[7] With a higher sampling frequency we end up with fewer trading day observations if we apply all filters described in this section. For example, we have considered a five-minute sampling frequency, from which we had 74% fewer observations for the PLZ/EUR exchange rate in comparison with the ten-minute sampling frequency, which is extreme, considering that we lost only around 11% for CZK/EUR and 5% for HUF/EUR. Using a fifteen-minute sampling frequency led only to small increases in sample sizes: around 4% for PLZ/EUR, 5% for CZK/EUR and 2% for HUF/EUR.

The intraday returns used in this study are defined as

$$r_{t,i} = \ln\left(z_{t,i} / z_{t,i-1}\right) \tag{34}$$

where $z_{t,i}$ is the average mid-price:

$$z_{t,i} = n\left(t,i\right)^{-1} \sum_{k=1}^{n(t,i)} \frac{x_{tik}^{b} + x_{tik}^{a}}{2} \tag{35}$$

with $n(t,i)$ being the number of valid prices for a given exchange rate on day t, in i-th sampling window. By focusing on returns calculated from the average of the mid-prices, we are targeting the non-synchronicity issues which may arise if only the last known values are taken from each window. Averaging should not only mitigate non-synchronicity issues, but it is also suitable for handling other microstructure noise in the data.

After creating datasets of prices with a ten-minute sampling frequency, we still encountered several occasions with data holes, i.e. a period where no price was recorded. This may be due to holidays, non-trading, technical errors during data recordings, etc. Therefore, each series was checked using two additional filters:

- if more than 16 prices ($z_{t,i}$) on a given day t are missing, remove all observations of that day t;
- if more than eight consecutive prices ($z_{t,i}$) on a given day t are missing, remove all observations of that day t.

The remaining data holes were inputted using linear interpolation. Finally, to allow comparisons, the datasets across all exchange rate series were synchronized. Given the estimation window of 180 trading days (and the additional 30 observations needed to choose the bandwidth; see Section 2.2), the out-of-sample performance of our models is tested on a sample of 585 days. *Table 1* describes our baseline data.

4. Results

We first describe our estimates of realized volatility and its components. Next we turn our attention to the results from the forecasting models. We will refrain from describing in detail all results for all loss functions (they are available in the corresponding *Tables 2–4*), as we will instead focus on the results related to the two empirical questions posed in our study: i. what types of HAR models lead to superior forecasts within our set of models and ii. whether an out-of-sample framework provides evidence for short-term volatility spillovers across emerging foreign exchange markets or from developed to emerging foreign exchange markets.

4.1 Realized Volatilities on Foreign Exchange Markets

Table 1 presents summary statistics for realized intraday returns and realized volatility and its subcomponents (signed jumps and jumps) for various currencies. Realized volatility (*RV*) and signed jumps are plotted in *Figure 1* and continuous and jump components are plotted in *Figure 2*. The figures suggest that volatility was high in 2012 and that we may expect upward volatility spikes rather than downward spikes. Both descriptive statistics and figures indicate that *RV* and its continuous component have rather high autocorrelation, whereas autocorrelation is very low for

Table 1 Descriptive Statistics of Realized Volatilities

*x*1000	Mean	SD	IQ	IIIQ	Kurt.	Skew.	$\rho(1)$	Max.	Date
Czech Koruna									
Realized Volatility	2.66	1.25	1.76	3.46	3.48	0.67	0.76	8.34	1.12.2011
Signed Jump	0.04	1.52	-0.99	1.24	2.88	-0.05	-0.06	4.57	28.11.2011
Continuous component	2.51	1.23	1.63	3.20	3.89	0.85	0.74	8.34	1.12.2011
Jump	0.49	0.78	0.00	1.04	3.58	1.34	0.09	3.26	13.2.2012
Hungarian Forint									
Realized Volatility	4.43	1.83	3.05	5.40	3.92	1.02	0.68	11.74	18.11.2011
Signed Jump	-0.05	2.48	-1.85	1.64	2.94	0.06	0.03	8.02	22.12.2011
Continuous component	4.28	1.82	2.89	5.21	4.04	1.04	0.66	11.74	18.11.2011
Jump	0.51	1.04	0.00	0.00	7.29	2.12	0.00	5.83	31.1.2014
Polish Zloty									
Realized Volatility	3.47	1.52	2.34	4.29	5.25	1.22	0.74	11.34	21.6.2013
Signed Jump	-0.04	1.95	-1.45	1.24	3.52	0.21	-0.02	8.11	21.6.2013
Continuous component	3.36	1.50	2.23	4.14	5.29	1.23	0.73	11.34	21.6.2013
Jump	0.38	0.82	0.00	0.00	9.20	2.39	-0.06	5.89	2.11.2011
US Dollar									
Realized Volatility	3.63	1.32	2.67	4.40	3.47	0.61	0.67	8.88	13.7.2011
Signed Jump	0.03	2.01	-1.55	1.70	2.19	-0.12	-0.01	4.81	5.10.2011
Continuous component	3.50	1.33	2.54	4.30	3.52	0.69	0.69	8.88	13.7.2011
Jump	0.43	0.86	0.00	0.00	6.29	1.97	0.02	4.87	8.7.2013
Swiss Frank									
Realized Volatility	2.10	2.05	0.77	2.81	14.90	2.70	0.90	18.02	18.8.2011
Signed Jump	-0.14	1.49	-0.62	0.48	15.65	-0.52	-0.06	10.61	15.8.2011
Continuous component	2.04	2.01	0.74	2.69	15.40	2.74	0.89	18.02	18.8.2011
Jump	0.19	0.61	0.00	0.00	38.70	5.30	0.10	6.24	16.8.2011

Notes: There are 783 valid observations in our analysis. From these, the initial 180 are used for the first forecast and a forecasting exercise is carried out for the remaining 603 observations. $\rho(1)$ is the value of the autocorrelation coefficient. Prior to calculating the descriptive statistics, the variables were multiplied by 10^3. *IQ* and *IIIQ* denote the first and third quartiles, respectively.

signed jumps and jump components. This implies that including signed jumps and jump components in our models might not improve them, which is confirmed later in the paper.

Among the CEE markets, the highest *RV* was measured for the Hungarian forint, which is 60.5% higher than the RV of the Czech koruna. This might be the consequence of the larger interest rate differentials against the euro rates, which were among the largest in the European Union during the sample period. Based on the data from Eurostat,[8] the short-term, three-month interbank rates in 2013 and 2014 were, on average, at 4.18 and 2.51, in Hungary, and 0.22 and 0.21 in the euro area, while for long-term maturities represented by ten-year bond yields, the rates in 2013 and

[8] http://ec.europa.eu/eurostat/statistics-explained/index.php/Exchange_rates_and_interest_rates

Figure 1 Realized Volatility and Signed Jumps of Foreign Exchange Volatility

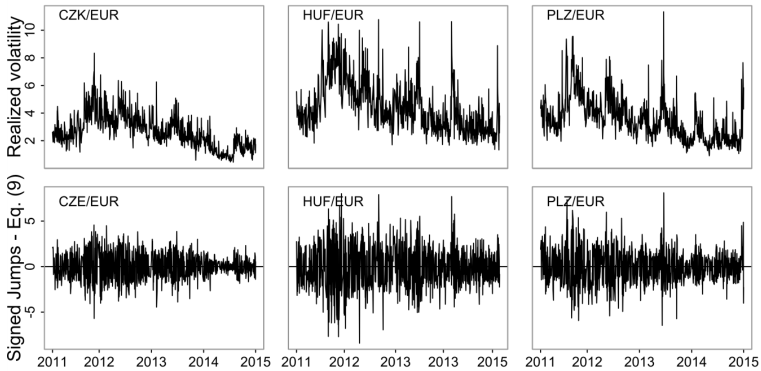
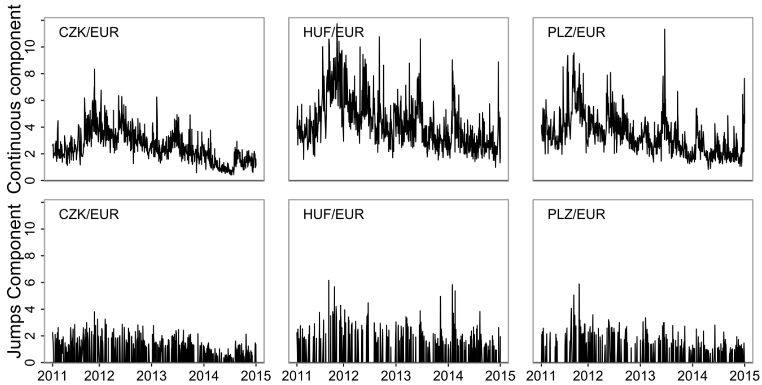


Figure 2 Continuous and Jump Components of Foreign Exchange Volatility



2014 were, on average, 5.92 and 4.81, in Hungary, and 3.00 and 2.05 in the euro area. As was already visible in *Figures 1–2*, *RVs* are skewed to the right.

An interesting observation is that there are not many instances of jumps. We found that the proportion of days with statistically significant jumps is 23% for the Hungarian Forint, 21% for the Polish zloty and 33% for the Czech koruna. Using a decade of high-frequency data of the DEM/USD exchange rate, Andersen *et al.* (2007) found that at the 0.05 significance level, jumps occurred on around 41% of trading days. The persistence of jumps was also small in Andersen *et al.* (2007). Compared to the existing literature, our estimates do not appear to be out of line and we conclude that jumps are not frequent on the CEE foreign exchange market, as they appear on less than one-third of trading days.

4.2 Volatility Spillovers on Foreign Exchange Markets

The results shown in *Tables 2–4* suggest that one-day forward volatility spillovers are rare in our sample of markets. Compared to the benchmark HAR model, statistically significant improvements in forecasting ability were identified for the HAR models of the Czech koruna and Polish zloty if the USD/EUR exchange rate realized

Table 2 Evaluation of Individual and Combined Forecasts: Czech Koruna to Euro

	MSFE	R^{MSFE}	R_X^{MSFE}	QLIKE	R^{QLIKE}	R_X^{QLIKE}	MAFE	R^{MAFE}	R_X^{MAFE}
Standard specification									
HAR	4.203	---	---	3.866	---	---	4.858	---	---
HAR-SJ	4.245	1.010	---	3.875	1.003	---	4.906	1.010[a]	---
HAR-CJ	4.216	1.003	---	4.142	1.072[b]	---	4.888	1.006	---
HAR-Q	4.225	1.005	---	3.890	1.006[a]	---	4.853	0.999	---
CF-Mean	4.178	**0.994**[a]	---	3.871	1.001	---	4.846	0.997	---
Regional exchange rate—HUF/EUR									
HAR-X	4.018	**0.956**[c]	---	3.741	**0.968**[b]	---	4.748	**0.977**[c]	---
HAR-X-SJ	4.069	**0.968**[a]	1.013	3.909	1.011	1.045[a]	4.770	**0.982**[aa]	1.005
HAR-X-CJ	4.126	0.982	1.027[b]	4.252	1.100	1.137[c]	4.812	0.991	1.014[a]
HAR-X-Q	4.054	**0.964**[b]	1.009	3.803	0.984	1.016	4.751	**0.978**[c]	1.001
CF-Mean	4.008	**0.953**[c]	0.997	3.806	0.985	1.017[a]	4.721	**0.972**[c]	**0.994**[a]
Regional exchange rate—PLZ/EUR									
HAR-X	4.216	1.003	---	3.875	1.002	---	4.854	0.999	---
HAR-X-SJ	4.225	1.005	1.002	3.911	1.012	1.009	4.867	1.002	1.003
HAR-X-CJ	4.237	1.008	1.005	4.200	1.086[b]	1.084[b]	4.884	1.005	1.006
HAR-X-Q	4.343	1.033[b]	1.030[b]	3.928	1.016[a]	1.014[b]	4.904	1.009	1.010[a]
CF-Mean	4.173	0.993	0.990[b]	3.876	1.003	1.000	4.821	0.992	**0.993**[b]
Global exchange rate—CHF/EUR									
HAR-X	4.200	0.999	---	3.902	1.009	---	4.848	0.998	---
HAR-X-SJ	4.244	1.010	1.011	3.933	1.017	1.008	4.929	1.015[a]	1.017[b]
HAR-X-CJ	4.242	1.009	1.010	4.366	1.129[a]	1.119[c]	4.885	1.006	1.008
HAR-X-Q	4.360	1.037[b]	1.038[b]	4.130	1.068[b]	1.058[c]	4.893	1.007	1.009[b]
CF-Mean	4.196	0.998	0.995	3.978	1.029	1.026[b]	4.850	0.998	0.999
Global exchange rate—USD/EUR									
HAR-X	4.141	**0.985**[a]	---	3.869	1.001	---	4.821	**0.992**[a]	---
HAR-X-SJ	4.242	1.009	1.024[b]	4.078	1.055[b]	1.054[b]	4.918	1.012[a]	1.020[c]
HAR-X-CJ	4.168	0.992	1.006	4.238	1.096[a]	1.095[b]	4.864	1.001	1.009
HAR-X-Q	4.195	0.998	1.013[a]	3.899	1.009	1.008	4.814	**0.991**[a]	0.999
CF-Mean	4.119	**0.980**[b]	0.977	3.903	1.010	1.007	4.817	**0.991**[a]	0.992
All univariate models									
CF-Trim	4.083	**0.971**[c]	---	3.836	0.992	---	4.785	**0.985**[c]	---
Multivariate models									
GHAR-2									
+ PLZ/EUR	4.362	1.038[c]	---	3.895	1.008	---	5.033	1.036[c]	---
+ HUF/EUR	4.270	1.016	---	3.827	0.990	---	4.955	1.020[c]	---
+ CHF/EUR	4.278	1.018	---	3.875	1.002	---	4.964	1.022[c]	---
+ USD/EUR	4.373	1.040[c]	---	3.898	1.008	---	5.004	1.030[c]	---
CF-Mean	4.290	1.021[a]	---	3.850	0.996	---	4.967	1.022[c]	---

GHAR-3									
+ PLZ/EUR, HUF/EUR	4.341	1.033^{c}	---	3.859	0.998	---	5.041	1.038^{c}	---
+ CHF/EUR, USD/EUR	4.360	1.037^{b}	---	3.890	1.006	---	4.998	1.029^{c}	---
CF-Mean	4.319	1.028^{b}	---	3.850	0.996	---	5.005	1.030^{c}	---
All models									
CF-Trim	4.118	**0.980**^{c}	---	**3.809**	**0.985**^{c}	---	**4.817**	**0.992**^{c}	---

Notes: The table reports the mean squared forecast error multiplied by 107, the average *QLIKE* multiplied by 102, and the mean absolute forecast error multiplied by 104. *R* and R_X are relative indices where the competing forecasting model is divided by the benchmark model, *HAR* in the case of the *R* index and the corresponding *HAR-X* model in the case of the *RX* index. *CF-Mean* corresponds to a combination forecast, i.e. the forecast being the average of corresponding forecasts; *CF-Trim* are values for the combination forecast, i.e. the forecast being the trimmed average of corresponding forecasts. a, b, and c denote statistical significance of the model encompassing test of Hansen *et al.* (2011) at the 10%, 5% and 1% levels, respectively, where we always compare two models: the competing forecasting model and the benchmark model. Bolded values correspond to competing models that outperformed the benchmark.

volatilities were included. Furthermore, the HAR model for the Czech koruna was improved by adding the HUF/EUR exchange rate volatility and the HAR model for the Polish zloty was improved by adding the CHF/EUR exchange rate volatility. Otherwise, we found that inclusion of foreign exchange rate volatilities led to statistically significant deterioration of volatility forecasts instead of improved forecasts.

Among the class of GHAR models, we did not specifications which would have provided consistently improved volatility forecasts. The choice between systems of two (GARH-2) or three (GHAR-3) exchange rates led to very similar outcomes, where results differ with respect to the given loss function and the outperformance of the HAR model is rare and never statistically significant. It appears that, at least within our sample period, the GHAR model was unable to exploit the dependence between foreign exchange market volatilities.

Overall, our results point to the finding that, given our sample of exchange rates and sample period, short-term volatility spillovers are uncommon. Furthermore, it also appears that with regard to one-day forward volatility forecasting, one-by-one exchange rate forecasting is enough.

4.3 Preferred One-Day Forward Forecasting Approach

Compared to the baseline HAR model, the largest improvements were observed for the least volatile market, CZK/EUR (4.7%), followed by the PLZ/EUR (3.0%), while only a small improvement of 0.7% (i.e. no statistically significant improvements at all) was observed for the most volatile market, HUF/EUR. Looking at the *R* and R_X indices allows us to evaluate the performance of the HAR-SJ, HAR-CJ and HAR-Q classes of models. Interestingly, forecasting the performance of HAR-CJ and HAR-Q models, which utilize decomposition of the realized volatility into the continuous and jump component (HAR-CJ) and estimated measurement error (HAR-Q), did not improve upon the benchmark models (HAR and HAR-X). This result was consistent across loss functions, exchange rates and other HAR model specifications. Decomposition of the realized volatility into positive and negative volatility (taken into account through signed jumps; HAR-SJ) appeared to be beneficial only for the PLZ/EUR exchange rate and the MSFE loss function.

Table 3 Evaluation of Individual and Combined Forecasts: Hungarian Forint to Euro

	MSFE	R^{MSFE}	R_X^{MSFE}	QLIKE	R^{QLIKE}	R_X^{QLIKE}	MAFE	R^{MAFE}	R_X^{MAFE}
Standard specification									
HAR	13.779	---	---	3.841	---	---	8.706	---	---
HAR-SJ	13.945	1.012	---	3.898	1.015	---	8.808	1.012[a]	---
HAR-CJ	14.141	1.026[c]	---	3.996	1.040[c]	---	8.841	1.015[b]	---
HAR-Q	13.812	1.002	---	3.847	1.002	---	8.700	0.999	---
CF-Mean	13.753	0.998	---	3.845	1.001	---	8.717	1.001	---
Regional exchange rate— PLZ/EUR									
HAR-X	13.883	1.008	---	3.836	0.999	---	8.682	0.997	---
HAR-X-SJ	14.022	1.018	1.010	3.941	1.026	1.027[a]	8.784	1.009	1.012
HAR-X-CJ	14.404	1.045[c]	1.038[c]	4.053	1.055[c]	1.056[c]	8.866	1.018[b]	1.021[b]
HAR-X-Q	14.039	1.019[b]	1.011[a]	3.924	1.022[b]	1.023[b]	8.721	1.002	1.004
CF-Mean	13.851	1.005	0.998	3.864	1.006	1.007	8.689	0.998	1.001
Regional exchange rate—CZK/EUR									
HAR-X	14.074	1.021[c]	---	3.920	1.020[c]	---	8.813	1.012[c]	---
HAR-X-SJ	14.278	1.036[c]	1.014	4.011	1.044[c]	1.023[a]	8.968	1.030[c]	1.018[b]
HAR-X-CJ	14.267	1.035[c]	1.014	4.029	1.049[c]	1.028[c]	8.938	1.027[c]	1.014[a]
HAR-X-Q	14.198	1.030[c]	1.009	3.982	1.037[c]	1.016	8.824	1.014[b]	1.001
CF-Mean	14.005	1.016[b]	0.995	3.916	1.020[c]	0.999	8.826	1.014[c]	1.001
Global exchange rate—CHF/EUR									
HAR-X	14.025	1.018[b]	---	3.914	1.019[c]	---	8.715	1.001	---
HAR-X-SJ	14.006	1.016	0.999	3.943	1.027[b]	1.008	8.737	1.004	1.002
HAR-X-CJ	14.412	1.046[c]	1.028[c]	4.052	1.055[c]	1.035[c]	8.845	1.016[b]	1.015[b]
HAR-X-Q	14.294	1.037[c]	1.019[c]	3.989	1.039[c]	1.019[b]	8.794	1.010[b]	1.009[b]
CF-Mean	13.952	1.013[a]	0.991	3.902	1.016[b]	0.996	8.709	1.000	0.988
Global exchange rate—USD/EUR									
HAR-X	13.799	1.002	---	3.849	1.002	---	8.646	**0.993[a]**	---
HAR-X-SJ	14.011	1.017	1.015	3.918	1.020[a]	1.018	8.724	1.002	1.009
HAR-X-CJ	14.177	1.029[c]	1.027[b]	4.012	1.045[c]	1.042[c]	8.793	1.010	1.017[b]
HAR-X-Q	14.061	1.021	1.019[a]	3.993	1.040[b]	1.037[b]	8.704	1.000	1.007
CF-Mean	13.758	0.998	0.978	3.855	1.004	0.983	8.655	0.994	0.982
All univariate models									
CF-Trim	13.736	0.997	---	3.838	0.999	---	8.676	0.997	---
Multivariate models									
GHAR-2									
+ PLZ/EUR	14.391	1.044[c]	---	3.955	1.030[b]	---	8.991	1.033[c]	---
+ HUF/EUR	14.428	1.047[c]	---	3.953	1.029[a]	---	8.849	1.016[b]	---
+ CHF/EUR	14.093	1.023	---	3.882	1.011	---	8.864	1.018[a]	---
+ USD/EUR	14.447	1.048[c]	---	3.965	1.032[c]	---	8.883	1.020[c]	---
CF-Mean	14.151	1.027[b]	---	3.896	1.014[a]	---	8.826	1.014[a]	---

GHAR-3									
+ PLZ/EUR, HUF/EUR	14.519	1.054[c]	---	3.987	1.038[b]	---	8.936	1.026[c]	---
+ CHF/EUR, USD/EUR	14.310	1.039[c]	---	3.945	1.027[b]	---	8.854	1.017[b]	---
CF-Mean	14.202	1.031[c]	---	3.919	1.020[a]	---	8.825	1.014[b]	---
All models									
CF-Trim	13.744	0.997	---	3.832	0.998	---	8.689	0.998	---

Notes: The table reports the mean squared forecast error multiplied by 107, the average QLIKE multiplied by 102, and the mean absolute forecast error multiplied by 104. *R* and R_X are relative indices where the competing forecasting model is divided by the benchmark model, *HAR* in the case of the *R* index and the corresponding *HAR-X* model in the case of the *RX* index. *CF-Mean* corresponds to a combination forecast, i.e. the forecast being the average of corresponding forecasts; *CF-Trim* are values for the combination forecast, i.e. the forecast being the trimmed average of corresponding forecasts. a, b and c denote statistical significance of the model encompassing test of Hansen *et al.* (2011) at the 10%, 5% and 1% levels, respectively, where we always compare two models: the competing forecasting model and the benchmark model. Bolded values correspond to competing models that outperformed the benchmark.

Simple combination forecasts of univariate HAR models (CF-Mean, CF-Trim) led, even across otherwise uncompetitive models (in terms of higher mean values of loss functions), to forecasts which often outperformed the benchmark HAR model (across all loss functions). An exception is the HUF/EUR exchange rate, where no statistically significant improvements were observed at all. Therefore, our conclusion is that, given that forecasters face model choice uncertainty, it appears to be preferable to use combination forecasts across (sound) univariate models rather than rely on one model only.

5. Conclusion

We use several models of short-term, one-day forward volatility forecasts of the exchange rates of the Czech koruna (CZK), Hungarian forint (HUF) and Polish zloty (PLN) against the euro. All of these models are based on the concept of realized volatility calculated from high-frequency data. Our main benchmark model is the univariate heterogeneous autoregressive (HAR) model of Corsi (2009). We addressed two questions: i. what type of one-day forward volatility forecasts (within our set of models) are to be preferred, and ii. are exchange rates subject to volatility spillovers?

We use several extensions of the baseline HAR model. We decompose realized volatility into positive and negative semi-volatilities, as well as into continuous and jump components, and we allow the autoregressive parameter of the HAR model to vary with the estimated degree of the measurement error (Bollerslev *et al.*, 2016). As another extension, we also include realized volatilities and components thereof from other local emerging markets and global foreign exchange markets (the Swiss franc and the US dollar), which allows us to study volatility spillovers on foreign exchange markets in an out-of-sample framework. Our analysis is further extended by estimating the multivariate GHAR model of Baruník and Čech (2016), which allows exploiting the cross-sectional dependence between regional and global exchange rates.

Overall, our study reveals that in our sample of markets and time period, it is difficult to outperform the standard HAR model of Corsi (2009), as there is no single univariate or multivariate HAR specification, which would dominate others across all

Table 4 Evaluation of Individual and Combined Forecast: Polish Zloty to Euro

	MSFE	R^{MSFE}	R_X^{MSFE}	QLIKE	R^{QLIKE}	R_X^{QLIKE}	MAFE	R^{MAFE}	R_X^{MAFE}
Standard specification									
HAR	7.975			3.496			6.184		
HAR-SJ	7.752	**0.972**[b]		3.422	0.979		6.095	**0.986**[a]	
HAR-CJ	7.970	0.999		3.488	0.998		6.195	1.002	
HAR-Q	8.109	1.017[a]		3.531	1.010[a]		6.229	1.007[b]	
CF-Mean	7.836	**0.983**		3.425	**0.980**[c]		6.125	**0.990**[c]	
Regional exchange rate—HUF/EUR									
HAR-X	8.071	1.012		3.501	1.002		6.250	1.011	
HAR-X-SJ	7.768	0.974	**0.963**[b]	3.418	0.978	0.976	6.157	0.996	**0.985**[a]
HAR-X-CJ	8.249	1.034	1.022	3.539	1.012	1.011	6.314	1.021[a]	1.010
HAR-X-Q	8.243	1.034	1.021[a]	3.567	1.020	1.019[b]	6.310	1.020[a]	1.010[a]
CF-Mean	7.930	0.994	**0.983**[c]	3.429	**0.981**[a]	**0.979**[c]	6.190	1.001	**0.990**[c]
Regional exchange rate—CZK/EUR									
HAR-X	8.053	1.010		3.534	1.011		6.234	1.008[a]	
HAR-X-SJ	7.844	0.984	**0.974**[b]	3.482	0.996	0.985	6.158	0.996	0.988
HAR-X-CJ	8.101	1.016	1.006	3.567	1.020	1.009	6.269	1.014[a]	1.006
HAR-X-Q	8.237	1.033[b]	1.023[b]	3.593	1.028[b]	1.017[a]	6.268	1.014[a]	1.006
CF-Mean	7.905	0.991	**0.982**[c]	3.465	0.991	**0.981**[c]	6.150	0.994	**0.987**[c]
Global exchange rate—CHF/EUR									
HAR-X	8.038	1.008		3.530	1.010		6.158	0.996	
HAR-X-SJ	7.739	**0.970**[a]	**0.963**[b]	3.429	0.981	0.971	6.096	0.986	0.990
HAR-X-CJ	8.111	1.017	1.009	3.565	1.020	1.010	6.215	1.005	1.009
HAR-X-Q	8.385	1.051[c]	1.043[c]	3.730	1.067[b]	1.057[b]	6.281	1.016[a]	1.020[c]
CF-Mean	7.896	0.990	**0.981**[c]	3.449	0.987	**0.976**[c]	6.109	**0.988**[a]	**0.980**[b]
Global exchange rate—USD/EUR									
HAR-X	7.987	1.002		3.501	1.001		6.168	0.997	
HAR-X-SJ	7.812	**0.980**[a]	**0.978**[a]	3.463	0.991	0.989	6.092	**0.985**[a]	0.988
HAR-X-CJ	8.011	1.005	1.003	3.526	1.009	1.007	6.176	0.999	1.001
HAR-X-Q	8.197	1.028[b]	1.026[b]	3.567	1.020[b]	1.019[b]	6.249	1.010[a]	1.013[c]
CF-Mean	7.851	**0.984**[b]	**0.975**[c]	3.430	**0.981**[a]	**0.971**[b]	6.101	**0.987**[b]	**0.979**[c]
All univariate models									
CF-Trim	7.807	**0.979**[c]		3.409	**0.975**[c]		6.113	**0.988**[c]	
Multivariate models									
GHAR-2									
+ PLZ/EUR	8.719	1.093[c]		3.690	1.056[c]		6.530	1.056[c]	
+ HUF/EUR	8.544	1.071[c]		3.618	1.035[c]		6.453	1.043[c]	
+ CHF/EUR	8.321	1.043[b]		3.523	1.008		6.380	1.032[c]	
+ USD/EUR	8.545	1.072[b]		3.519	1.007		6.456	1.044[c]	
CF-Mean	8.366	1.049[c]		3.537	1.012[b]		6.399	1.035[c]	

GHAR-3						
+ PLZ/EUR, HUF/EUR	8.707	1.092^{c}	3.751	1.073^{c}	6.593	1.066^{c}
+ CHF/EUR, USD/EUR	8.147	1.022	3.454	0.988	6.357	1.028^{b}
CF-Mean	8.273	1.037^{c}	3.545	1.014^{a}	6.421	1.038^{c}
All models						
CF-Trim	7.839	$\mathbf{0.983^{c}}$	3.424	$\mathbf{0.980^{c}}$	6.142	$\mathbf{0.993^{b}}$

Notes: The table reports the mean squared forecast error multiplied by 107, the average QLIKE multiplied by 102, and the mean absolute forecast error multiplied by 104. *R* and R_X are relative indices where the competing forecasting model is divided by the benchmark model, *HAR* in the case of the *R* index and the corresponding *HAR-X* model in the case of the R_X index. *CF-Mean* corresponds to a combination forecast, i.e. the forecast being the average of corresponding forecasts; *CF-Trim* are values for the combination forecast, i.e. the forecast being the trimmed average of corresponding forecasts. a, b and c denote statistical significance of the model encompassing test of Hansen *et al.* (2011) at 10%, 5%, 1% levels, respectively, where we always compare two models: the competing forecasting model and the benchmark model. Bolded values correspond to competing models that out-performed the benchmark.

exchange rates and loss functions (MSFE, QLIKE and MAFE). Furthermore, there is sparse evidence of short-term volatility spillovers.

Our results can be summarized as follows:

- We found that the HUF/EUR exchange rate has the highest realized volatility, followed by PLZ/EUR and CZK/EUR. In accordance with the existing empirical literature, the realized volatilities as well as their continuous components show a large degree of persistence.
- Jump components are not common, as they were recorded for less than one-third of the time (CZK/EUR). Furthermore, contrary to the continuous component, the persistence of the jump component is negligible, suggesting that price discontinuities might be one of the factors increasing forecasting errors.
- Compared to the benchmark models, adding signed jumps (the difference between positive and negative semi-variances) led to statistically improved forecasts of the realized volatility of the PLZ/EUR exchange rate only.
- Decomposition of the realized variance into its continuous and jump components did not improve forecasting performance.
- Taking the estimated measurement error of integrated volatility into account also did not improve upon the performance of the benchmark models.
- In our out-of-sample framework, we found little evidence of volatility spillovers between regional exchange rates and from global to regional exchange rates. An exception is the volatility spillover from USD/EUR to HUF/EUR and PLZ/EUR and from HUF/EUR to CZK/EUR, and from CHF/EUR to PLZ/EUR. Thus the Czech koruna and Polish zloty appear to be more vulnerable to local and global spillovers than the Hungarian forint.
- Multivariate extensions of the HAR model, the GHAR models, have not performed better than univariate models or combination forecasts from univariate models.
- Interestingly, for the PLZ/EUR and CZK/EUR exchange rates, combination forecasts almost always led to improved forecasts. Although the improvements were not of a particularly great magnitude (from 2.5% to 4.7%), they were often statistically significant.

REFERENCES

Aghion P, Bacchetta P, Panciére R, Rogoff K (2009): Exchange rate volatility and productivity growth: The role of financial development. *Journal of Monetary Economics*, 56(4):494–513.

Andersen TG, Bollerslev T (1998): Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4):885–905.

Andersen TG, Bollerslev T, Diebold FX (2007): Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility. *Review of Economics and Statistics*, 89(4):701–720.

Andersen TG, Bollerslev T, Diebold FX, Labys P (2001): Modeling and forecasting realized volatility. *National Bureau of Economic Research*, *Working Paper*, no. 8160.

Andersen TG, Dobrev D, Schaumburg E (2012): Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1):75–93.

Arize AC, Osang T, Slottje DJ (2000): Exchange-Rate Volatility and Foreign Trade: Evidence from Thirteen LDC's. *Journal of Business & Economic Statistics*, 18(1):10–17.

Arrabitel O, Furceri D, Martin R, Zdzienicka A (2011): The effect of nominal exchange rate volatility on real macroeconomic performance in the CEE countries. *Economic Systems*, 35(2): 261–277.

Barndorff-Nielsen OE, Shephard N (2004): Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, 2(1):1–37.

Barndorff-Nielsen O, Kinnebrock S, Shephard N (2010): Measuring downside risk—Realised semivariance. In: Bollerslev T, Russell J, Watson M (Eds): *Volatility and time series econometrics: Essays in honor of Robert F. Engle*. Oxford University Press.

Barndorff-Nielsen, OE, Hansen PH, Lunde A, Shephard N (2011): Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 162(2):149–169.

Baruník J, Čech P (2016): On the modelling and forecasting multivariate realized volatility: Generalized Heterogeneous Autoregressive (GHAR) model. *Journal of Forecasting* (forthcoming).

Bates JM, Granger CWJ (1969): The Combination of Forecasts. *Operational Research Quarterly*, 20(4):451–468.

Bauer GH, Vorkink K (2011): Forecasting multivariate realized stock market volatility. *Journal of Econometrics*, 160(1):93–101.

Beine M, Lahaye J, Laurent S, Neely CJ, Palm FC (2007): Central bank intervention and exchange rate volatility, its continuous and jump components. *International Journal of Finance & Economics*, 12(2):201–223.

Bernardi M, Catania L (2014): The Model Confidence Set package for R. *arXiv*:1410.8504v1.

Birkelund OH, Haugom E, Molnár P, Opdal M, Westgaard S (2015): A comparison of implied and realized volatility in the Nordic power forward market. *Energy Economics*, 48(March):288–294.

Bollerslev T, Patton AJ, Quaedvlieg R (2016): Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Forecasting*, 192(1):1–18.

Bubák V, Kočenda E, Žikeš F (2011): Volatility transmission in emerging European foreign exchange markets. *Journal of Banking & Finance*, 35(11):2829–2841.

Caporale GM, Ali FM, Spagnolo N (2015): Exchange rate uncertainty and international portfolio flows: A multivariate GARCH-in-mean approach. *Journal of International Money and Finance*, 54(June):70–92.

Chiriac R, Voev V (2011): Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics*, 26(6):922–947.

Chowdhury AR (1993): Does Exchange Rate Volatility Depress Trade Flows? Evidence from Error-Correction Models. *Review of Economics and Statistics*, 75(4):700–706.

Christoffersen P, Jacobs K, Mimouni K (2010): Volatility dynamics for the S&P500: Evidence from realized volatility, daily returns, and option prices. *Review of Financial Studies*, 23(8):3141–3189.

Corsi F (2009): A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2):174–196.

Cuaresma JC, Geršl A, Slačík T (2010): Global Financial Crisis and the Puzzling Exchange Rate Path in CEE Countries. *Charles University in Prague, IES Working Paper*, no. 24/2010.

Dacorogna MM, Müller UA, Nagler RJ, Olsen RB, Pictet OV (1993): A geographical model for the daily and weekly seasonal volatility in the FX market. *Journal of International Money and Finance*, 12(4):413–438.

Dacorogna MM, Müller UA, Pictet OV, Vries CG de (2001): Extremal Forex Returns in Extremely Large Data Sets. *Extremes*, 4(2):105–127.

Diebold FX, Mariano RS (1995): Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 20(1):134–144.

Fengler MR, Gisler KIM (2015): A variance spillover analysis without covariances: What do we miss? *Journal of International Money and Finance*, 51(March):174–195.

Fidrmuc J, Horváth R. (2008): Volatility of exchange rates in selected new EU members: Evidence from daily data. *Economic Systems*, 32(1):103–118.

Galí J, Monacelli T (2005): Monetary Policy and Exchange Rate Volatility in a Small Open Economy. *Review of Economic Studies*, 72(3):707–734.

Granger CWJ (1969): Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Hansen PR, Lunde A, Nason JM (2011): The model confidence set. *Econometrica*, 79(2):453–497.

Hansen PR, Lunde A (2006): Consistent ranking of volatility models. *Journal of Econometrics*, 131(1-2):97–121.

Haugom E, Langeland H, Molnár P, Westgaard S (2014): Forecasting volatility of the US oil market. *Journal of Banking & Finance*, 47(October):1–14.

Kočenda E, Valachy J (2006): Exchange rate volatility and regime change: A Visegrad comparison. *Journal of Comparative Economics*, 34(4):727–753.

Meese RA, Rogoff K (1983): Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, 14(1-2):3–24.

Müller UA, Dacorogna MM, Olsen RB, Pictet OV, Schwartz M, Morgenegg C (1990): Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal of Banking & Finance*, 14(6):1189–1208.

Patton AJ (2011): Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256.

Patton AJ, Sheppard K (2015): Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility. *Review of Economics and Statistics*, 97(3):683–697.

Schnabl G (2008): Exchange rate volatility and growth in small open economies at the EMU periphery. *Economic Systems*, 32(1):70–91.

Schnabl G (2009): Exchange rate volatility and growth in emerging Europe and East Asia. *Open Economies Review*, 20(4):565–587.

Sévi B (2014): Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research*, 235(3):643–659.

Stock JH, Watson MW (2004): Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23(6):405–430.

Timmermann A (2006): Forecast combinations. In: Elliott G, Granger C, Timmermann A (Eds): *Handbook of Economic Forecasting*, 1:136–196.

Žikeš F, Baruník J (2014): Semiparametric Conditional Quantile Models for Financial Returns and Realized Volatility. *Journal of Financial Econometrics*, doi: 10.1093/jjfinec/nbu029