Barunik, Jozef; Krehlik, Tomas; Vacha, Lukas

Working Paper

Modeling and forecasting exchange rate volatility in time-frequency domain

FinMaP-Working Paper, No. 55

Provided in Cooperation with:
Collaborative EU Project FinMaP - Financial Distortions and Macroeconomic Performance, Kiel University et al.

Suggested Citation: Barunik, Jozef; Krehlik, Tomas; Vacha, Lukas (2016) : Modeling and forecasting exchange rate volatility in time-frequency domain, FinMaP-Working Paper, No. 55, Kiel University, FinMaP - Financial Distortions and Macroeconomic Performance, Kiel

This Version is available at:
https://hdl.handle.net/10419/125827

FinMaP-Working Paper No.55

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 612955

FINMAP –

FINANCIAL DISTORTIONS AND MACROECONOMIC PERFORMANCE: EXPECTATIONS, CONSTRAINTS AND INTERACTION OF AGENTS

DATE: 01/13/2016

TITLE

Modeling and forecasting exchange rate volatility in time-frequency domain

by: Jozef Barunik, Tomas Krehlik, Lukas Vacha

ABSTRACT

This paper proposes an enhanced approach to modeling and forecasting volatility using high frequency data. Using a forecasting model based on Realized GARCH with multiple time-frequency decomposed realized volatility measures, we study the influence of different timescales on volatility forecasts. The decomposition of volatility into several timescales approximates the behaviour of traders at corresponding investment horizons. The proposed methodology is moreover able to account for impact of jumps due to a recently proposed jump wavelet two scale realized volatility estimator. We propose a realized Jump-GARCH models estimated in two versions using maximum likelihood as well as observation-driven estimation framework of generalized autoregressive score. We compare forecasts using several popular realized volatility measures on foreign exchange rate futures data covering the recent financial crisis. Our results indicate that disentangling jump variation from the integrated variation is important for forecasting performance.
An interesting insight into the volatility process is also provided by its multiscale decomposition. We find that most of the information for future volatility comes from high frequency part of the spectra representing very short investment horizons. Our newly proposed models outperform statistically the popular as well conventional models in both one-day and multi-period-ahead forecasting.

Keywords: Realized GARCH, wavelet decomposition, jumps, multi-period-ahead volatility forecasting

AUTHORS

1. **Jozef Barunik**
   Institute of Economic Studies, Charles University,
   Opletalova 21, 110 00,
   Prague, Czech Republic

   Department of Econometrics, IITA,
   The Czech Academy of Sciences,
   Pod Vodarenskou Vezi 4, 182 00,
   Prague, Czech Republic

   Email: barunik@utia.cas.cz

2. **Tomas Krehlik**
   Institute of Economic Studies, Charles University,
   Opletalova 21, 110 00,
   Prague, Czech Republic

   Department of Econometrics, IITA,
   The Czech Academy of Sciences,
   Pod Vodarenskou Vezi 4, 182 00,
   Prague, Czech Republic

3. **Lukas Vacha**
   Institute of Economic Studies, Charles University,
   Opletalova 21, 110 00,
   Prague, Czech Republic

Modeling and forecasting exchange rate volatility in time-frequency domain[☆]

Jozef Barunik[a,b,*], Tomas Krehlik[a,b], Lukas Vacha[a,b]

[a]*Institute of Economic Studies, Charles University, Opletalova 21, 110 00, Prague, Czech Republic*
[b]*Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodarenskou Vezi 4, 182 00, Prague, Czech Republic*

Abstract

This paper proposes an enhanced approach to modeling and forecasting volatility using high frequency data. Using a forecasting model based on Realized GARCH with multiple time-frequency decomposed realized volatility measures, we study the influence of different timescales on volatility forecasts. The decomposition of volatility into several timescales approximates the behaviour of traders at corresponding investment horizons. The proposed methodology is moreover able to account for impact of jumps due to a recently proposed jump wavelet two scale realized volatility estimator. We propose a realized Jump-GARCH models estimated in two versions using maximum likelihood as well as observation-driven estimation framework of generalized autoregressive score. We compare forecasts using several popular realized volatility measures on foreign exchange rate futures data covering the recent financial crisis. Our results indicate that disentangling jump variation from the integrated variation is important for forecasting performance. An interesting insight into the volatility process is also provided by its multiscale decomposition. We find that most of the information for future volatility comes from high frequency part of the spectra representing very short investment horizons. Our newly proposed models outperform statistically the popular as well conventional models in both one-day and multi-period-ahead forecasting.

Keywords: Realized GARCH, wavelet decomposition, jumps, multi-period-ahead volatility forecasting

July 28, 2015

JEL: C14, C53, G17

1. Introduction

In contrast to the conventional framework of a generalized autoregressive conditional heteroscedasticity (GARCH) model, volatility is directly observed and can be used for forecasting when high frequency data are applied.[1] While Hansen and Lunde (2005) argue that GARCH(1,1) can hardly be beaten by any other model, recent active research shows that with help of high frequency measures, we can improve the volatility forecasts significantly. Mcmillan and Speight (2012) for example utilize intra-day data and show that we can obtain forecasts superior to forecasts from GARCH(1,1). Louzis et al. (2013) assesses the informational content of alternative realized volatility estimators using Realized GARCH in Value-at-Risk prediction. We aim to extend this line of research by investigating the importance of disentangling jump variation and integrated variance in recently developed framework, which combines appeal of a widely used GARCH(1,1) and high frequency data. Moreover, we employ recently developed multiscale estimators which decompose volatility into several investment horizons[2] and allow us to study the influence of intraday investment horizons on the volatility forecasts.

Traders on financial markets make their decisions over different time horizons, for example, minutes, hours, days, or even longer such as months and years (LeBaron, 2001; Ramsey, 2002; Gençay et al., 2005; Corsi, 2009). Nevertheless, majority of the empirical literature studies the relationships in the time domain only aggregating the behavior across all investment

☆We are grateful to the editor Lorenzo Peccati and two anonymous referees for many useful comments and suggestions, which greatly improved the paper. We are also grateful to David Veredas and Karel Najzar, and seminar participants at the Modeling High Frequency Data in Finance 3 in New York (July 2011) and Computational and Financial Econometrics in Oviedo (December 2012) for many useful discussions. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. FP7-SSH- 612955 (FinMaP). Support from the Czech Science Foundation under the 13-32263S project is gratefully acknowledged.

*Corresponding author, Tel. +420(776)259273, Email address: barunik@utia.cas.cz

[1] A vast quantity of literature on several aspects of estimating volatility using high frequency data is nicely surveyed by McAleer and Medeiros (2008).

[2] An investment horizon refers to the length of time that an investor expects to hold a security or portfolio.

horizons. A notable exception is the Heterogenous Autoregressive approach (HAR) for realized volatility proposed by Corsi (2009). Although staying in the time domain solely, Corsi (2009) builds his model on the idea of the investors' heterogeneity.

In our work, we ask if wavelet decomposition can provide better insight into the foreign exchange volatility modeling and forecasting.[3] Wavelets are often successfully used as a de-noising tool (Haven et al., 2012; Sun and Meinl, 2012). One particularly appealing feature of wavelets is that they can be embedded into stochastic processes, as shown by Antoniou and Gustafson (1999). Thus we can conveniently use them to extend the theory of realized measures to obtain decomposed volatility as shown by Fan and Wang (2007), or Barunik and Vacha (2014). One of the common issues with the interpretation of wavelets in economic applications is that they are filter, thus they can hardly be used for forecasting in econometrics. Models based on wavelets are often outperformed by simple benchmark models, as shown by Fernandez (2008). Rather, they can provide a useful "lens" into the spectral properties of the time series. Our wavelet-based estimator of realized volatility uses wavelets only to decompose the daily variation of the returns using intraday information, hence the problem with forecasts is no longer an issue. As wavelets are used to measure realized volatility at different investment horizons, we can construct a forecasting model based on the wavelet decomposed volatility conveniently.

Several attempts to use wavelets in the estimation of realized variation have emerged in the past few years. Høg and Lunde (2003) were the first to suggest a wavelet estimator of realized variance. Capobianco (2004), for example, proposes to use a wavelet transform as a comparable estimator of quadratic variation. Subbotin (2008) uses wavelets to decompose volatility into a multi-horizon scale. One exception is the work of Fan and Wang (2007), who were the first to use the wavelet-based realized variance estimator together with the methodology for estimation of jumps. In Barunik and Vacha (2014), we revisit and extend this work and using large Monte Carlo study we show that this estimator improves forecasting of the volatility substantially when compared to other estimators. Moreover, in Barunik and Vacha (2014) we attempt to use the estimators to decompose stock market volatility into several investment horizons in a non-parametric way.

Motivated by previous results, this paper focuses on proposing a model

[3]Our interest is in return variation, although models attempting to capture the prices directly may be of interest.

which will improve the modeling and forecasting of foreign exchange volatility. Similarly to Lanne (2007), Andersen et al. (2011), and Sévi (2014) we use the decomposition of the quadratic variation with the intention of building a more accurate forecasting model. Our approach is very different though, as we use wavelets to decompose the integrated volatility into several investment horizons and jumps. Moreover, we employ recently proposed realized GARCH framework of Hansen et al. (2012). In contrast to popular HAR framework of Corsi (2009), Realized GARCH allows to model jointly returns and realized measures of volatility, while key feature is a measurement equation that relates the realized measure to the conditional variance of returns. In addition, we benchmark our approach to several measures of realized volatility and jumps, namely realized volatility estimator proposed by Andersen et al. (2003), the bipower variation estimator of Barndorff-Nielsen and Shephard (2004), the median realized volatility of Andersen et al. (2012), and finally jump wavelet two-scale realized variance (JWTSRV) estimator of Barunik and Vacha (2014) in the framework of Realized GARCH, and we find significant differences in volatility forecasts, while our JWTSRV estimator brings the largest improvement. We use Realized GARCH models of Hansen et al. (2012) as well as realized GAS of Huang et al. (2014) based on the observation-driven estimation framework of generalized autoregressive score models to build a realized Jump-GARCH modeling strategy. In addition, we also utilize Realized GARCH with multiple realized measures (Hansen and Huang, 2012) to build a time-frequency model for forecasting volatility.

The main contribution of the paper is thus threefold.[4] First, we propose several model extensions to utilize jumps in the popular Realized GARCH frameworks, as well as build time-frequency model for forecasting volatility. Second, we use several popular measures as a benchmark to our time-frequency model. Third, we bring interesting empirical comparison of all frameworks in multiple-period-ahead forecasting exercise. We show that the most important information influencing the future volatility in foreign exchange is carried by the high frequency part of the spectra representing very short investment horizons. This decomposition gives us an interesting

[4]Note that our research adds to recent operation research contributions using wavelets in denoising of financial data, specifically high frequency data research (Haven et al., 2012; Sun and Meinl, 2012; Marroquı et al., 2013), literature developing volatility models (Pun et al., 2015; Date and Islyaev, 2015), literature contributing to forecasting volatility (Sévi, 2014; Charles, 2010; Christodoulakis, 2007) and studying stock market returns (Wang et al., 2015; Doyle and Chen, 2013; Yang and Bessler, 2008; Buckley and Long, 2015).

insight into the volatility process. Our newly proposed time-frequency models and Jump-GARCH models outperforms the existing modeling strategies significantly.

2. Theoretical framework for time-frequency decomposition of realized volatility

While most time series models are naturally set in the time domain, wavelet transform help us to enrich the analysis of quadratic variation by the frequency domain. Traders of the foreign exchange markets are operating with heterogeneous expectations, ranging from minutes to days, or even weeks and months. Hence volatility dynamics should be understood not only in time but at different investment horizons as well. In this section, we introduce a multiscale estimator that will allow these features and is moreover able to separate the continuous part of the price process containing noise from the jump variation. We will briefly introduce general ideas of constructing the estimator, while for the details necessary to understand the derivation of the estimator using wavelet theory, we refer to Barunik and Vacha (2014). In addition, we introduce several other estimators commonly used in the literature, which will serve as a benchmarks to us in the empirical application.

In the analysis, we assume that the latent logarithmic asset price follows a standard jump-diffusion process contamined with microstructure noise. Let y_t be the observed logarithmic prices evolving over $0 \leq t \leq T$, which will have two components; the latent, so-called "true log-price process", $dp_t = \mu_t dt + \sigma_t dW_t + \xi_t dq_t$, and zero mean *i.i.d.* microstructure noise, ϵ_t, with variance η^2. In a latent process, q_t is a Poisson process uncorrelated with W_t, and the magnitude of the jump, denoted as J_l, is controlled by factor $\xi_t \sim N(\bar{\xi}, \sigma_\xi^2)$. Thus, the observed price process is $y_t = p_t + \epsilon_t$.

The quadratic return variation over the interval $[t-h, t]$, for $0 \leq h \leq t \leq T$ associated with the price process y_t can be naturally decomposed into two parts: integrated variance of the latent price process, $IV_{t,h}$ and jump variation $JV_{t,h}$

$$QV_{t,h} = \underbrace{\int_{t-h}^{t} \sigma_s^2 ds}_{IV_{t,h}} + \underbrace{\sum_{t-h \leq l \leq t} J_l^2}_{JV_{t,h}} \tag{1}$$

As detailed by Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002a), quadratic variation is a natural measure of variability in the loga-

rithmic price process. A simple consistent estimator of the overall quadratic variation under the assumption of zero noise contamination in the price process is provided by the well-known realized variance, introduced by Andersen and Bollerslev (1998). The realized variance over $[t-h,t]$ can be estimated as

$$\widehat{QV}_{t,h}^{(RV)} = \sum_{k=1}^{N} \left(\Delta_k y_t\right)^2 \ , \tag{2}$$

where $\Delta_k y_t = y_{t-h+\left(\frac{k}{N}\right)h} - y_{t-h+\left(\frac{k-1}{N}\right)h}$ is the k-th intraday return in the $[t-h,t]$ and N is the number of intraday observations. The estimator in Eq.(2) converges in probability to $IV_{t,h} + JV_{t,h}$ as $N \to \infty$ (Andersen and Bollerslev, 1998; Andersen et al., 2001, 2003; Barndorff-Nielsen and Shephard, 2001, 2002a,b).

While the observed price process y_t is contamined with noise and jumps in real data, we need to account for this, as the main object of interest is the IV_t part of quadratic variation. Zhang et al. (2005) propose solution to the noise contamination by introducing the so-called two-scale realized volatility (TSRV henceforth) estimator. They adopt a methodology for estimation of the quadratic variation utilizing all of the available data using an idea of precise bias estimation. The two-scale realized variation over $[t-h,t]$ is measured by

$$\widehat{QV}_{t,h}^{(TSRV)} = \widehat{QV}_{t,h}^{(average)} - \frac{\bar{N}}{N}\widehat{QV}_{t,h}^{(all)} \ , \tag{3}$$

where $\widehat{QV}_{t,h}^{(all)}$ is computed as in Eq. (2) on all available data and $\widehat{QV}_{t,h}^{(average)}$ is constructed by averaging the estimators $\widehat{QV}_{t,h}^{(g)}$ obtained on G grids of average size $\bar{N} = N/G$ as $\widehat{QV}_{t,h}^{(average)} = \frac{1}{G}\sum_{g=1}^{G}\widehat{QV}_{t,h}^{(g)}$, where the original grid of observation times, $M = \{t_1, \ldots, t_N\}$ is subsampled to $M^{(g)}$, $g = 1, \ldots, G$, where $N/G \to \infty$ as $N \to \infty$. For example, $M^{(1)}$ will start at the first observation and take an observation every 5 minutes, $M^{(2)}$ will start at the second observation and take an observation every 5 minutes, ets. Finally, we average these estimators through the subsamples, so the variation of the estimator is averaged.

The estimator in Eq. (3) provides the first consistent and asymptotic estimator of the quadratic variation of p_t with rate of convergence $N^{-1/6}$. Zhang et al. (2005) also provide the theory for optimal choice of G grids, $G^* = cN^{2/3}$, where the constant c can be set to minimize the total asymptotic variance.

Since we are interested in decomposing quadratic variation into the integrated variance and jump variation component, we introduce a methodology for jump detection. Recent evidence from the volatility forecasting literature indicates that two sources of variation in the price process substantially differ and impact future volatility in different ways. Before introducing our estimator, we introduce two commonly used estimators of volatility and integrated variation, which will be used as benchmark in the empirical exercise.

Barndorff-Nielsen and Shephard (2004, 2006) develop bipower variation estimator (BV), which can detect the presence of jumps in high-frequency data. The main idea of the BV estimator is to compare two measures of the integrated variance, one containing the jump variation and the other being robust to jumps and hence containing only the integrated variation part. In our work, we use the Andersen et al. (2011) adjustment of the original Barndorff-Nielsen and Shephard (2004) estimator, which helps render it robust to certain types of microstructure noise. The bipower variation over $[t-h, t]$ is defined by

$$\widehat{IV}_{t,h}^{(BV)} = \mu_1^{-2} \frac{N}{N-2} \sum_{k=3}^{N} |\Delta_{k-2} y_t| . |\Delta_k y_t|, \tag{4}$$

where $\mu_a = \pi/2 = E(|Z|^a)$, and $Z \sim N(0,1)$, $a \geq 0$ and $\widehat{IV}_{t,h}^{(BV)} \rightarrow \int_{t-h}^{t} \sigma_s^2 ds$. Therefore, $\widehat{IV}_{t,h}^{(BV)}$ provides a consistent estimator of the integrated variance. Although $\widehat{QV}_{t,h}^{(RV)}$ provides a consistent estimator of the integrated variance plus the jump variation, the jump variation may be estimated consistently as the difference between the realized variance and the realized bipower variation

$$\text{plim}_{N \to \infty} \left(\widehat{QV}_{t,h}^{(RV)} - \widehat{IV}_{t,h}^{(BV)} \right) = JV_{t,h}. \tag{5}$$

Under the assumption of no jump and some other regularity conditions, Barndorff-Nielsen and Shephard (2006) provide the joint asymptotic distri-

bution of the jump variation.[5] Using this theory, the contribution of the jump variation to the quadratic variation of the price process is measured by

$$\widehat{JV}_{t,h}^{(BV)} = \mathbb{1}_{\{\mathcal{Z}_{t,h}^{(BV)} > \Phi_\alpha\}} \left(\widehat{QV}_{t,h}^{(RV)} - \widehat{IV}_{t,h}^{(BV)} \right), \tag{6}$$

where $\mathbb{1}_{\{\mathcal{Z}_{t,h}^{(BV)} > \Phi_\alpha\}}$ denotes the indicator function and Φ_α refers to the chosen critical value from the standard normal distribution. The measure of integrated variance is defined as

$$\widehat{IV}_{t,h}^{(CBV)} = \mathbb{1}_{\{\mathcal{Z}_{t,h}^{(BV)} \leq \Phi_\alpha\}} \widehat{QV}_{t,h}^{(RV)} + \mathbb{1}_{\{\mathcal{Z}_{t,h}^{(BV)} > \Phi_\alpha\}} \widehat{IV}_{t,h}^{(BV)}, \tag{7}$$

ensuring that the jump measure and the continuous part add up to the estimated variance without jumps. Another popular estimator, which estimates the integrated volatility in the presence of jumps is the median realized volatility (MedRV), introduced by Andersen et al. (2012):

$$\widehat{IV}_{t,h}^{(MedRV)} = \frac{\pi}{6 - 4\sqrt{3} + \pi} \left(\frac{N}{N-2} \right) \sum_{k=3}^{N} \text{med} \left(|\Delta_{k-2} y_t|, |\Delta_{k-1} y_t|, |\Delta_k y_t| \right)^2. \tag{8}$$

Under the assumption of no jump and some other regularity conditions, Andersen et al. (2012) provide the joint asymptotic distribution of the jump variation[6] analogously to the BV estimator. The integrated variance and

[5] Under the null hypothesis of no within-day jumps,

$$\mathcal{Z}_{t,h}^{(BV)} = \frac{\frac{\widehat{QV}_{t,h}^{(RV)} - \widehat{IV}_{t,h}^{(BV)}}{\widehat{QV}_{t,h}^{(RV)}}}{\sqrt{\left(\left(\frac{\pi}{2}\right)^2 + \pi - 5 \right) \frac{1}{N} \max \left(1, \frac{\widehat{TQ}_{t,h}^{(BV)}}{\left(\widehat{IV}_{t,h}^{(BV)} \right)^2} \right)}},$$

where $\widehat{TQ}_{t,h}^{(BV)} = N \mu_{4/3}^{-3} \left(\frac{N}{N-4} \right) \sum_{k=5}^{N} |\Delta_{k-4} y_t|^{4/3} |\Delta_{k-3} y_t|^{4/3} |\Delta_{k-2} y_t|^{4/3}$ is asymptotically standard normally distributed.

[6] Under the null hypothesis of no within-day jumps,

$$\mathcal{Z}_{t,h}^{(MedRV)} = \frac{\frac{\widehat{QV}_{t,h}^{(RV)} - \widehat{IV}_{t,h}^{(MedRV)}}{\widehat{QV}_{t,h}^{(RV)}}}{\sqrt{0.96 \frac{1}{N} \max \left(1, \frac{\widehat{TQ}_{t,h}^{(MedRV)}}{\left(\widehat{IV}_{t,h}^{(MedRV)} \right)^2} \right)}},$$

jump variation can be consistently estimated as

$$\widehat{IV}_{t,h}^{(CMedRV)} = \mathbb{1}_{\{\mathcal{Z}_{t,h}^{(MedRV)} \leq \Phi_\alpha\}} \widehat{QV}_{t,h}^{(RV)} + \mathbb{1}_{\{\mathcal{Z}_{t,h}^{(MedRV)} > \Phi_\alpha\}} \widehat{IV}_{t,h}^{(MedRV)}, \quad (9)$$

$$\widehat{JV}_{t,h}^{(MedRV)} = \mathbb{1}_{\{\mathcal{Z}_{t,h}^{(MedRV)} > \Phi_\alpha\}} \left(\widehat{QV}_{t,h}^{(RV)} - \widehat{IV}_{t,h}^{(MedRV)} \right). \quad (10)$$

2.1. Estimation of jumps and time-frequency realized variance

Fan and Wang (2007) use a different approach to realized volatility measurement. They use wavelets in order to separate jump variation from the price process, as well as for estimation of the integrated variance on the jump–adjusted data. In addition, wavelet methodology offers decomposition of the estimated volatility into scales representing investment horizons. Therefore, we can observe how particular investment horizon contributes to the total variance. For a more detailed description of the wavelet transform used as a building block of the estimation see Appendix A. In the empirical section, we aim to study information content of investment horizons for volatility forecasting, thus we describe the wavelet jump detection and then introduce the wavelet estimator of integrated variance of Barunik and Vacha (2014), which allows to decompose the volatility into several investment horizons.

As in the previous Section, we assume the sample path of the price process $y_t = p_t + \epsilon_t$ will have a finite number of jumps. Following results of Wang (1995) on the wavelet jump detection and further extension of Fan and Wang (2007) to stochastic processes, we apply wavelet transform to detect jumps. Using effective localization properties of the wavelets, Fan and Wang (2007) show a way how to distinguish between continuous and jump part of the stochastic price process with *i.i.d.* additive noise. They use the first scale of the discrete wavelet transform (the highest frequency) where the price process p_t dominates the noise ϵ_t only close to a jump location, otherwise it is very small. In order to detect dominating part of the process y_t, Fan and Wang (2007) use the universal threshold of Donoho and Johnstone (1994), defined as $D_t = d_t\sqrt{2\log N}$, where $d_t = median\{|\mathcal{W}_{1,k}|\}/0.6745$ for $k \in [1, N]$ is a robust estimate of standard deviation.[7] When the absolute value of the wavelet coefficient at the first scale is greater than a threshold

where $\widehat{TQ}_{t,h}^{(MedRV)} = \frac{3\pi N}{9\pi+72-52\sqrt{3}} \left(\frac{N}{N-2}\right) \sum_{k=3}^{N} \text{med}\left(|\Delta_{k-2}y_t|, |\Delta_{k-1}y_t|, |\Delta_k y_t|\right)^4$ is asymptotically standard normally distributed.

[7] See Appendix A.1 for the definition of wavelet tranform.

D_t then the noise ϵ_t is relatively small and the dominance of p_t is caused by the jump part, therefore a jump is detected.

In the empirical part we adapt Fan and Wang (2007) procedure to the "maximal overlap discrete wavelet transform" (MODWT). As a result a robust estimate of standard deviation has to be modified as: $d_t = \sqrt{2}\ median\{|\mathcal{W}_{1,k}|\}/0.6745$ for $k \in [1, N]$.[8] Since we use the MODWT, we have k wavelet coefficients at the first scale, which corresponds to number of intraday observations, i.e., $k = 1, \ldots, N$. In case the absolute value of the wavelet coefficient $|\mathcal{W}_{1,k}|$ is greater[9] then the threshold D_t than a jump with size $\Delta_k J_t$ is detected as

$$\Delta_k J_t = \left(y_{t-h+\left(\frac{k}{N}\right)h} - y_{t-h+\left(\frac{k-1}{N}\right)h} \right) \mathbb{1}_{\left\{|\mathcal{W}_{1,k}| > D_t\right\}} \quad k \in [1, N]. \tag{11}$$

Following Fan and Wang (2007), the jump variation over $[t-h, t]$ in the discrete time is estimated as the sum of squares of all the estimated jump sizes,

$$\widehat{JV}_{t,h} = \sum_{k=1}^{N} \left(\Delta_k J_t\right)^2. \tag{12}$$

Fan and Wang (2007) prove that using (12), we are able to estimate the jump variation from the process consistently with the convergence rate of $N^{-1/4}$.

Having precisely detected jumps, we proceed to jump adjustment of the observed price process y_t over $[t-h, t]$. We adjust the data for jumps by subtracting the intraday jumps from the price process as:

$$\Delta_k y_t^{(J)} = \Delta_k y_t - \Delta_k J_t, \quad k = 1, \ldots N, \tag{13}$$

where N is the number of intraday observations.

Finally, the volatility can be computed using the jump-adjusted wavelet two-scale realized variance (JWTSRV) estimator on the jump adjusted data $\Delta_k y_t^{(J)}$. The estimator utilizes the TSRV approach of Zhang et al. (2005) as well as the wavelet jump detection method. Another advantage of the estimator is, that it decomposes the integrated variance into $J^m + 1$ components, therefore we are able to study the dynamics of volatility at various

[8] For more information about universal thresholds applied on the MODWT see (Percival and Walden, 2000) and Gençay et al. (2002).

[9] Using the MODWT filters, we need to slightly correct the position of the wavelet coefficients to get the precise jump position, see Percival and Mofjeld (1997).

investment horizons. Following Barunik and Vacha (2014), we define the JWTSRV estimator over $[t-h,t]$, on the jump-adjusted data as:

$$\widehat{IV}_{t,h}^{(JWTSRV)} = \sum_{j=1}^{J^m+1} \widehat{IV}_{j,t,h}^{(JWTSRV)} = \sum_{j=1}^{J^m+1} \left(\widehat{IV}_{j,t,h}^{(average)} - \frac{\bar{N}}{N} \widehat{IV}_{j,t,h}^{(all)} \right), \quad (14)$$

where $\widehat{IV}_{j,t,h}^{(average)} = \frac{1}{G} \sum_{g=1}^{G} \sum_{k=1}^{N} \left(\mathcal{W}_{j,k}^{(g)} \right)^2$ is obtained from wavelet coefficient estimates on a grid of size $\bar{N} = N/G$, and $\widehat{IV}_{j,t,h}^{(all)} = \sum_{k=1}^{N} \left(\mathcal{W}_{j,k} \right)^2$ is the wavelet realized variance estimator at a scale j on all the jump-adjusted observed data, $\Delta_k y_t^{(J)}$. $\mathcal{W}_{j,k}$ denotes the MODWT wavelet coefficient at scale j with position k obtained over $[t-h,t]$.

Barunik and Vacha (2014) show that the JWTSRV is consistent estimator of the integrated variance as it converges in probability to the integrated variance of the process p_t, and they test the small sample performance of the estimator in a large Monte Carlo study. The JWTSRV is found to be able to recover true integrated variance from the noisy process with jumps very precisely. Moreover, the JWTSRV estimator is also tested in forecasting exercise, which confirms to improve forecasting of the integrated variance substantially.

2.1.1. Bootstrapping the jump test using JWTSRV

Although Fan and Wang (2007) showed the effectiveness of the wavelet jump detection, distribution properties of the estimated jump variation, and hence any test statistic stay unknown. In order to test for the presence of jumps using JWTSRV estimator, we propose to use the bootstrap test. Main reason for bootstrapping the jump test is that consistent estimator for the integrated quarticity is not analytically available for JWTSRV estimator. More importantly, finite sample properties of the jump tests based on functions of realized volatility estimators can be considerably improved using bootstrap, as noted by Dovonon et al. (2014).

In order to obtain the bootstrapped distribution of test statistic under the assumption of no jumps, we generate k intraday returns using estimated integrated part of the quadratic variation as $\Delta_k y_t^* = \sqrt{(1/k)\widehat{IV}_t^{(JWTSRV)}}\eta_{i,t}$, with $\eta_{i,t} \sim N(0,1)$ generated independently. $\widehat{QV}_t^{(RV^*)}$ and $\widehat{QV}_t^{(JWTSRV^*)}$ are then estimated on the given day t. Generating $b = 1, \ldots, B$ realizations, we

obtain

$$\mathcal{Z}^*_{t,h}(b) = \frac{\widehat{QV}^{(RV^*)}_{t,h} - \widehat{IV}^{(JWTSRV^*)}_{t,h}}{\widehat{QV}^{(RV^*)}_{t,h}}, \tag{15}$$

which can be used to construct a bootstrap statistic to test the null hypothesis of no jumps as:

$$\mathcal{Z}^{(JWTSRV)}_{t,h} = \frac{\frac{\widehat{QV}^{(RV^*)}_{t,h} - \widehat{IV}^{(JWTSRV^*)}_{t,h}}{\widehat{QV}^{(RV^*)}_{t,h}} - E(\mathcal{Z}^*_{t,h}(b))}{\sqrt{Var(\mathcal{Z}^*_{t,h}(b))}}. \tag{16}$$

The bootstrap expectation and variance both depend on the data. We will rely on the assumptions of Dovonon et al. (2014), who show that under general conditions, this statistics will be normally distributed with limiting variance one, although they provide this result for the BV estimator. While we leave the rigorous treatment of this approach with JWTSRV estimator for the future work, we have studied the properties of the bootstrap test using simulations, which are available upon request from authors.

The integrated variance and jump variation can then be consistently estimated as

$$\widehat{IV}^{(CJWTSRV)}_{t,h} = \mathbb{1}_{\{\mathcal{Z}^{(JWTSRV)}_{t,h} \leq \Phi_\alpha\}} \widehat{QV}^{(RV)}_{t,h} + \mathbb{1}_{\{\mathcal{Z}^{(JWTSRV)}_{t,h} > \Phi_\alpha\}} \widehat{IV}^{(JWTSRV)}_{t,h} \tag{17}$$

$$\widehat{JV}^{(JWTSRV)}_{t,h} = \mathbb{1}_{\{\mathcal{Z}^{(JWTSRV)}_{t,h} > \Phi_\alpha\}} \left(\widehat{QV}^{(RV)}_{t,h} - \widehat{IV}^{(JWTSRV)}_{t,h} \right). \tag{18}$$

3. A forecasting model based on decomposed integrated volatilities and jumps

Similarly to Lanne (2007), Andersen et al. (2011), and Sévi (2014), we use the decomposition of the quadratic variation with the intention of building a more accurate forecasting model. Our approach is very different though, as we use wavelets to decompose the integrated volatility into several investment horizons and jumps first. Then, we employ recently proposed Realized GARCH framework of Hansen et al. (2012) and its variants. Realized GARCH allows to model jointly returns and realized measures of volatility, while key feature is a measurement equation that relates the realized measure to the conditional variance of returns. We use the decomposed realized measures in the Realized GARCH, and expect that our modification will result in better in-sample fits of the data as well as out-of-sample

forecasts. For comparison, we also use other estimators and study how they improve the forecasting ability of Realized GARCH.

3.1. Realized Jump GARCH framework for forecasting

The key object of interest in GARCH family is the conditional variance, $h_t = var(r_t|\mathcal{F}_{t-1})$, where r_t is a time series of returns. While in a standard GARCH(1,1) model the conditional variance, h_t is dependent on its past h_{t-1} and r^2_{t-1}, Hansen et al. (2012) propose to utilize realized measures of volatility and make h_t dependent on them as well. The authors introduce so-called measurement equation which ties the realized measure to latent volatility. The general framework of Realized GARCH(p, q) models is well connected to existing literature in Hansen et al. (2012). Here, we restrict ourselves to the simple log-linear specification of Realized GARCH$(1, 1)$ with Gaussian innovations which we will use to build our model.

Realized GARCH makes use of realized measures of volatility to help forecast the latent volatility process. In the previous sections, we have motivated several estimators, which allow us to disentangle continuous part and jump part of the quadratic return variation. While both parts may carry important information about future volatility, we propose a modified framework, which includes both.

There are essentially two possible treatments of jumps in the Realized GARCH framework, depending on the belief about its endogenous or exogenous nature. In a large study, Chatrath et al. (2014); Lahaye et al. (2011) show that currency jumps can be explained by U.S. macro announcements using the realized measures. This provides a good empirical evidence about the exogenous nature of jump arrivals. By addition of estimated jumps into the variance equation, we propose a Realized Jump-GARCH(1,1) model (Realized J-GARCH) given by

$$
\begin{aligned}
r_t &= \sqrt{h_t} z_t, && (19)\\
\log(h_t) &= \omega + \beta \log(h_{t-1}) + \gamma \log(x_{t-1}) + \gamma_J \log(1 + JV_{t-1}), && (20)\\
\log(x_t) &= \xi + \phi \log(h_t) + \tau_1 z_t + \tau_2 z_t^2 + u_t, && (21)
\end{aligned}
$$

where r_t is the return, x_t and JV_t are estimated continuous and jump components of quadratic variation using BV, MedRV, or JWTSRV realized measures, and z_t and u_t come from Gaussian normal distribution and are mutually independent. $\tau_1 z_t + \tau_2 z_t^2$ is leverage function. If jumps have a significant impact on volatility forecasts, γ_J coefficient should be significantly

different from zero. For $\gamma_J = 0$, the model reduces to the original Realized GARCH.

Hansen et al. (2012) motivate possibility of obtaining feasible multi-period-ahead forecasts as one of the main advantages of this framework. Multi-period-ahead predictions with the Realized GARCH model are straightforward with the use of vector autoregression structure for $log(h_t)$ and $log(x_t)$. In this paper, we follow this simple approach. In order to obtain multiple-period-ahead forecasts, we need to include jump component to the forecasting structure. Once we are treating jumps as an exogenous process, we simply use the ARMA structure for the $\log(1 + JV_{t-1})$, which allows to obtain the multiple-period-ahead forecasts analogously to the Realized GARCH model.

3.2. Realized GARCH model based on decomposed integrated volatiles

In addition to jumps, we also utilize decomposition of JWTSRV to see which investment horizon has impact on the future volatility as well. We also expect each volatility component at different investment horizon to carry different information, which should again help to enhance the final forecasts. To be able to fully explore the decompositions, we use the extension of Realized Exponential GARCH model that can utilize multiple realized volatility measures introduced by Hansen and Huang (2012). The realized EGARCH model with $j = 1, \ldots, J^m + 1$ volatility components at different investment horizons estimated using JWTSRV in $x_{j,t}$ is

$$
\begin{aligned}
r_t &= \sqrt{h_t} z_t, && (22)\\
\log(h_t) &= \omega + \beta \log(h_{t-1}) + \tau(z_{t-1}) + \gamma' u_{t-1}, && (23)\\
\log(x_{j,t}) &= \xi_j + \phi_j \log(h_t) + \delta_{(j)}(z_t) + u_{j,t}, && (24)
\end{aligned}
$$

where $z_t \sim N(0,1)$, and $u_t \sim N(0, \Sigma)$ are mutually and serially independent, and $u_t = (u_{1,t}, \ldots, u_{j,t})'$, and $\tau(z_t) = \tau_1 z_t + \tau_2(z_t^2 - 1)$, and $\delta_{(j)}(z_t) = \delta_{j,1} z_t + \delta_{j,2}(z_t^2 - 1)$.

Note that the model is different as the $\log(h_t)$ equation has the u_{t-1} instead of realized measure, and includes leverage function. For the case when $j = 1$, model is equivalent to the previous one, and by simple substitution, we can obtain the relation of parameters directly (Hansen and Huang, 2012). Hence the model with multiple equations is just a generalization of the previous work, which allows us to fully utilize the decomposed volatility into several investment horizons, and so parameters in vector γ'

will provide a good guide for significance of various investment horizons on volatility forecasts.

All the models are estimated by quasi-maximum likelihood farmework (QMLE) and can be easily generalized by assuming different distributions of z_t and u_t. Hansen et al. (2012) provide the asymptotic properties of the QMLE, while (Hansen and Huang, 2012) extend it to the framework with multiple realized measures, although the work is currently unfinished. The quasi log-likelihood function is given by

$$\ell(r,x;\theta,\Sigma) = -\frac{1}{2}\sum_{t=1}^{T}\left(\underbrace{\log(2\pi)+\log(h_t)+z_t^2}_{=\ell(r)} + \underbrace{K\log(2\pi)+\log(|\Sigma|)+u_t'\Sigma^{-1}u_t}_{=\ell(x|r)}\right), \tag{25}$$

where θ holds set of parameters to be estimated by maximizing the quasi log-likelihood with respect to θ and Σ. The log-likelihood can be divided in two according to the contribution of realized measures to the log-likelihood value, $\ell(x|r)$ and contribution of returns, $\ell(r)$. In the empirical analysis, we report the two values as we use conventional GARCH model as a benchmark, so we are able to compare the fits. It is again straightforward to obtain multiple-period-ahead point forecasts using estimated parameters. For the details, see for example Lunde and Olesen (2013)

3.3. Generalized Autoregressive Score Model with Realized Measures of Volatility and Jumps

Recently introduced observation-driven estimation framework of Generalized Autoregressive Score (GAS) models due to Creal et al. (2013) has recently gained considerable popularity. Huang et al. (2014) propose a new observation-driven time-varying parameter Realized GARCH, in which the dynamic latent factor is updated by the scaled local density score as a function of past daily returns and realized variance. The new framework is robust to extreme outliers in observations, hence it may serve as a robustness check to our modeling strategy. We again add jumps to the original model, obtaining Realized Jump GAS Model as

$$\begin{aligned} r_t &= \sqrt{h_t}z_t, && (26)\\ \log(x_t) &= \xi + \phi\log(h_t) + d_1(z_t^2-1) + d_2 z_t + \sigma u_t, && (27)\\ \log(h_{t+1}) &= \omega + \alpha\mathcal{S}_t\Delta_t + \beta\log(h_t) + \gamma_J\log(1+JV_{t-1}), && (28) \end{aligned}$$

where x_t and JV_t are estimated continuous and jump components of quadratic variation using BV, MedRV, or JWTSRV estimators, and z_t and u_t come from Gaussian normal distribution and are mutually independent. $d_1(z_t^2 - 1) + d_2 z_t$ is leverage function that introduces dependence between the return shock and volatility shock. The main change in comparison to previous models is in the dynamics of the latent volatility, driven by the dynamic score, where $\Delta_t = \partial \ln p\left(r_t, \log(x_t) | \mathcal{F}_{t-1}; \log(h_t), \theta\right) / \partial \log(h_t)$ is the conditional score at time t and $\mathcal{S}_t = -E_{t-1}[\partial^2 \ln p\left(r_t, \log(x_t) | \mathcal{F}_{t-1}; \log(h_t), \theta\right) / \partial^2 \log(h_t)^2]^{-1}$ is the scaling matrix. Analogously to the QMLE framework, likelihood can be separated to two parts, which we report in order to be able to compare the fits. Assuming both z_t, and u_t follow independent standardized normal distributions, dynamic score reduces to $\Delta_t = \frac{1}{2}(z_t^2 - 1) + \frac{1}{\sigma} u_t \left(\phi + d_1 z_t^2 - \frac{d_2}{2} z_t\right)$, $\mathcal{S}_t^{-1} = \frac{1}{\sigma^2}\left(\phi^2 + 3d_1^2 + \frac{d_2^2}{4} - 2d_1\phi\right) + \frac{1}{2}$. Assuming exogenous ARMA structure for jumps, multiple-period-ahead forecasts are again obtained readily.

3.4. Forecast evaluation using different realized variance measures

To test significant differences of competing models, we use the Model Confidence Set (MCS) methodology of Hansen et al. (2011). Given a set of forecasting models, $\mathcal{M}_0$, we identify the model confidence set $\widehat{\mathcal{M}}^*_{1-\alpha} \subset \mathcal{M}_0$, which is the set of models that contain the best forecasting model given a level of confidence α. For a given model $i \in \mathcal{M}_0$, the p-value is the threshold confidence level. Model i belongs to the MCS only if $\widehat{p}_i \geq \alpha$. MCS methodology repeatedly tests the null hypothesis of equal forecasting accuracy

$$H_{0,\mathcal{M}} : E[L_{i,t} - L_{j,t}] = 0, \qquad \text{for all } i, j \in \mathcal{M}$$

with $L_{i,t}$ being an appropriate loss function of the i-th model. Starting with the full set of models, $\mathcal{M} = \mathcal{M}_0$, this procedure sequentially eliminates the worst-performing model from $\mathcal{M}$ when the null is rejected. The surviving set of models then belong to the model confidence set $\widehat{\mathcal{M}}^*_{1-\alpha}$. Following Hansen et al. (2011), we implement the MCS using a stationary bootstrap with an average block length of 20 days.[10] Two robust loss functions, mean square error (MSE) and QLIKEare used in the MCS (Patton, 2011),while root mean square error (RMSE) is reported in the Tables.

[10] We have used different block lengths, including the ones depending on the forecasting horizons, to assess the robustness of the results, without any change in the final results. These results are available from the authors upon request.

4. Empirical application: Does decomposition bring any improvement in volatility forecasting?

4.1. Data description

Foreign exchange future contracts are traded on the Chicago Mercantile Exchange (CME) on a 24-hour basis. As these markets are among the most liquid, they are suitable for analysis of high-frequency data. We will estimate the realized volatility of British pound (GBP), Swiss franc (CHF) and euro (EUR) futures. All contracts are quoted in the unit value of the foreign currency in US dollars. It is advantageous to use currency futures data for the analysis instead of spot currency prices, as they embed interest rate differentials and do not suffer from additional microstructure noise coming from over-the-counter trading. The cleaned data are available from Tick Data, Inc.[11]

It is important to look first at the changes in the trading system before we proceed with the estimation on the data. In August 2003, for example, the CME launched the Globex trading platform, and for the first time ever in a single month, the trading volume on the electronic trading platform exceeded 1 million contracts every day. On Monday, December 18, 2006, the CME Globex(R) electronic trading platform started offering nearly continuous trading. More precisely, the trading cycle became 23 hours a day (from 5:00 pm on the previous day until 4:00 pm on current day, with a one-hour break in continuous trading), from 5:00 pm on Sunday until 4:00 pm on Friday. These changes certainly had a dramatic impact on trading activity and the amount of information available, resulting in difficulties in comparing the estimators on the pre-2003 data, the 2003–2006 data and the post–2006 data. For this reason, we restrict our analysis to a sample period extending from January 2, 2007 through August 20, 2014, which contains the most recent financial crisis. The futures contracts we use are automatically rolled over to provide continuous price records, so we do not have to deal with different maturities.

The tick-by-tick transactions are recorded in Chicago Time, referred to as Central Standard Time (CST). Therefore, in a given day, trading activity starts at 5:00 pm CST in Asia, continues in Europe followed by North America, and finally closes at 4:00 pm in Australia. To exclude potential jumps due to the one-hour gap in trading, we redefine the day in accordance with the electronic trading system. Moreover, we eliminate transactions

[11] http://www.tickdata.com/

executed on Saturdays and Sundays, US federal holidays, December 24 to 26, and December 31 to January 2, because of the low activity on these days, which could lead to estimation bias. Finally, we are left with 1902 days in the sample. Looking more deeply at higher frequencies, we find a large amount of multiple transactions happening exactly at the same time stamp. We use the arithmetic average for all observations with the same time stamp.

Having prepared the data, we can estimate the integrated volatility using different estimators and use them within proposed forecasting framework. For each futures contract, the daily quadratic variation is estimated using the realized variance estimator. Integrated variance and jump variation are estimated with the bipower variation, median estimator, and finally our jump wavelet two-scale realized variance estimator. All the estimators are adjusted for small sample bias. For convenience, we refer to the estimators in the description of the results as RV, BV, MedRV and JWTSRV, respectively, while the BV, MedRV, and JWTSRV estimators are used for decomposition of continuous and jump part of quadratic variation, and JWTSRV for decomposition to various investment horizons. We use sampling frequency of 5-minutes.

The decomposition of volatility into the continuous and jump part is depicted by Figure 1, which provide the returns, estimated jump and finally integrated variance components using JWTSRV estimator for all three futures pairs. Figure 2 shows the further decomposition into several investment horizons. For better illustration, we annualize the square root of the integrated variance in order to get the annualized volatility and we compute the components of the volatility on several investment horizons. Figure 2 (a) to (e) show the investment horizons of up to 10 minutes, 10 to 20 minutes, 20 to 40 minutes, 40 to 80 minutes and up to 1 day, respectively. It is very interesting that most of the volatility (around 50%) comes from the 5-minute to 10-minute investment horizons band which is a new empirical insight. Moreover, the longer the investment horizon, the lower the contribution of the variance to the total variation.

4.2. In-sample fits

The main results of estimation and forecasting are presented in this section. The estimation strategy is as follows. For each of three forex futures considered, namely GBP, CHF and EUR, we first estimate benchmark GARCH(1,1) model. Then, we estimate the Realized GARCH (1,1) with RV, which will serve as a benchmark model to our Realized Jump

GARCH(1,1) with BV, MedRV, and JWTSRV. All these models are estimated using QMLE and GAS model frameworks. Finally, we add Realized GARCH model with multiple JWTSRV components to see the impact of investment horizons on forecasts.[12]

Tables 1, 2 and 3 contain in-sample fits for GBP futures, CHF futures and EUR futures on the full sample respectively. By observing partial log-likelihood $\ell(r)$, we can see immediately that all the Realized GARCH models bring significant improvement to the conventional GARCH(1,1) without high frequency realized measures, reported by the first column (in testing significance of the difference, we restrict ourselves to use simple log-likelihood ratio test).

When we focus on comparison of Realized GARCH models, we can observe further significant differences. Our Jump-GARCH brings small improvements to the $l(r)$ consistent with the literature, but large improvements in terms of $l(x|r)$ when compared to the benchmark Realized GARCH with RV. As to the comparison of QMLE and GAS specifications, original QMLE model outperforms GAS in terms of likelihood slightly. These observations hold for all three futures used in the study.

Further comparison of the Realized Jump-GARCH models with three different realized measures reveals that JWTSRV and MedRV largely outperform BV, with JWTSRV bringing largest gains for CHF futures, and MedRV winning the race for the rest. While log-likelihoods $\ell(r, x)$ uncover rather large differences between the models, parameter estimates for the different realized measures are very similar to each other, and are consistent with the estimates found in the literature.

The most important parameter γ_J is significantly different from zero for BV and MedRV estimators, but not for JWTSRV estimator. We explain this by more strict statistics for testing the null hypothesis of no jumps in comparison to MedRV and BV, while we use bootstrap, which corrects the statistics for small sample distortions. As pointed out by Dovonon et al. (2014), the differences maybe quite severe. Even with this result, we can

[12] All models are estimated with NLopt-BOBYQA optimization algorithm using *nloptr* package in R version 3.2.1. Comparison of the models' elapsed times (running on a MacBook computer with Intel Core i5 2.6 GHz CPU) computed as 95% trimmed mean from 30 runs is 0.25s for GAS models and 0.22s for MLE models. We note that we report the elapsed times from our rather inefficient algorithm, which become proportionally slower with increasing number of parameters. Elapsed times can be significantly improved with use of efficient optimization algorithms.

conclude that jumps bring significant improvement in the modeling and Realized Jump-GARCH(1,1) outperforms benchmark Realized GARCH.

Finally, we focus on the Realized GARCH model with multiple measures, where we use volatility decompositions to several investment horizons due to our JWTSRV measure, and also include RV representing full quadratic variation. We find γ_j coefficients statistically different from zero for all three futures. This means that volatility further decomposed to several investment horizons carry significant contribution to the future latent volatility. Coefficient is largest at the first scale, following the second, and the rest. This points us to the result that mainly volatility from highest frequency impacts the future volatility.

Turning our attention to ϕ_j, we can see that it is close to one (within standard errors) for all investment horizons. Note however how ξ_j decreases with decreasing scale. This mirrors the different contributions of the energy (variance of each volatility at different investment horizon j to total variance) to the latent volatility. From Eq. (14) we know, that volatility components at different horizons j always sum up to the total volatility. But Realized GARCH model use logarithmic transforms, which do not hold this property. Hence, the expected value of the parameter ξ_j will logically be a total constant minus $\log(1/2^j)$, as JWTSRV is simply sum of squared wavelet coefficients on intraday return, which is driven mainly by Brownian motion. This points us to the conclusion that the most of the information can be found in the high frequency part of the spetral density of returns.

4.3. Multi-period-ahead forecasting results

Motivated by a good in-sample performance of the models, we study if inclusion of jumps in the model improves the volatility forecasts in our newly proposed Realized Jump-GARCH models. We also wait to see if the model with multiple investment horizons improves volatility forecasts, and finally, it will be interesting to find out if the log-likelihood gains also translate to good forecasting performance of the models.

We use all the Realized GARCH models to produce $h = \{1, 5, 10\}$-day-ahead forecasts based on rolling basis. Table 4 compares RMSE of all the models. To see if the forecasts are statistically different, we use the Model Confidence Set (MCS) with two robust loss functions, MSE and QLIKE. Models, which are included in the MCS with the use of both loss functions are highlighted in bold. In addition, we provide ranking of the models according to the both loss functions within MCS in the superscript. First

number is ranking due to MSE, second one is ranking of the models due to QLIKE.

Turning to the results in the Table 4, we can see that Realized GARCH model with multiple investment horizons is never rejected[13] from the Model Confidence set by neither of the loss functions. Moreover, for GBP and EUR futures, it ranks as the best forecasting model with exception of forecasting horizon of 10 days, when it ranks as second according to QMLE. The model also delivers lowest RMSE of the forecasts, and ranks second to fifth with CHF futures outperformed mainly by GAS estimates.

Another model, which is never rejected by neither of the loss functions from MCS is the Realized Jump GAS model with our JWTSRV. For all three futures and all forecasting horizons, the model ranks as second best, to eighth best depending on the loss function. Similar results are delivered also with the use of MedRV, when the model often ranks third to fifth best, with one exception of forecasting EUR futures at horizons of five days. Realized Jump GAS model with BV is the third best model, as it is rejected from the MCS only for EUR futures with QLIKE loss function.

Most of the Realized (Jump) GARCH models estimated using QMLE are rejected from MCS by one of the loss functions. The only exception is CHF forecasts at 10-day-ahead horizon, when the Realized Jump GARCH model with BV measure ranks best using both loss functions.

Overall, the log-likelihood gains from QMLE estimates do not translate to better out-of-sample forecasts, as GAS outperforms the MLE models. Realized Jump-GARCH largely outperforms benchmark Realized GARCH with RV, and finally our multiple horizon model outperforms all the models delivering lowest loss functions most of the times. Thus jump variation as well as further decomposition of volatility to different scales bring significant improvement to the volatility forecasts in all tested forex futures.

5. Conclusion

In this paper, we investigate how the decomposed integrated volatilities and jumps influence the future volatility using realized GARCH framework.

[13]Note that we use h_t as predictor of volatility in the models. Some researchers report that restricting parameters $\phi = 1$ may improve the predictive performance in this case. We, however do not document significant improvement on our dataset, and make the results with restricted models available upon request.

Utilizing a jump wavelet two scale realized volatility estimator, which measures foreign exchange volatility in the time-frequency domain, we study the influence of intra-day investment horizons on daily volatility forecasts.

After the introduction of wavelet-based estimation of quadratic variation together with forecasting model, we compare our estimators to several most popular estimators, namely, realized variance, bipower variation, and median realized volatility in the forecasting exercise. Using several Realized GARCH specifications estimated by QMLE, GAS, and multiple realized measures, the wavelet-based estimator proves to bring significant improvement in the volatility forecasts. Models incorporating jumps improve forecasting ability significantly. Next, we find that while realized Jump GAS models do not outperform other models in terms of in-sample fits, they largely outperform the MLE-based estimates in the forecasts at all forecasting horizons.

Concluding the empirical findings, we show that our wavelet-based estimators bring a significant improvement to the volatility estimation and forecasting. It also offers a new method of time-frequency modeling of realized volatility which helps us to better understand the dynamics of stock market behavior. Specifically, it uncovers that most of the volatility is created on higher frequencies.

References

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

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association 96*(453), 42–55.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Andersen, T., T. Bollerslev, and X. Huang (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *Journal of Econometrics 160*(1), 176–189.

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

Antoniou, I. and K. Gustafson (1999). Wavelets and stochastic processes. *Mathematics and Computers in Simulation 49*(1), 81–104.

Barndorff-Nielsen, O. and N. Shephard (2001). Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B 63*(2), 167–241.

Barndorff-Nielsen, O. and N. Shephard (2002a). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*(2), 253–280.

Barndorff-Nielsen, O. and N. Shephard (2002b). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics 17*(5), 457–477.

Barndorff-Nielsen, O. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics 2*(1), 1–37.

Barndorff-Nielsen, O. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics 4*(1), 1–30.

Barunik, J. and L. Vacha (2014). Realized wavelet-based estimation of integrated variance and jumps in the presence of noise. *To Appear in Quantittive Finance*.

Buckley, W. S. and H. Long (2015). A discontinuous mispricing model under asymmetric information. *European Journal of Operational Research 243*(3), 944 – 955.

Capobianco, E. (2004). Multiscale stochastic dynamics in finance. *Physica A* (344), 122–127.

Charles, A. (2010). The day-of-the-week effects on the volatility: The role of the asymmetry. *European Journal of Operational Research 202*(1), 143 – 152.

Chatrath, A., H. Miao, S. Ramchander, and S. Villupuram (2014). Currency jumps, cojumps and the role of macro news. *Journal of International Money and Finance 40*, 42–62.

Christodoulakis, G. A. (2007). Common volatility and correlation clustering in asset returns. *European Journal of Operational Research 182*(3), 1263 – 1284.

Chui, C. (1992). *An Inroduction to Wavelets*. Academic Press, New York.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics 7*(2), 174–196.

Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*(5), 777–795.

Date, P. and S. Islyaev (2015). A fast calibrating volatility model for option pricing. *European Journal of Operational Research 243*(2), 599 – 606.

Daubechies, I. (1992). *Ten lectures on wavelets*. SIAM.

Donoho, D. L. and I. M. Johnstone (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrica 81*(3), 425–455.

Dovonon, P., S. Gonçalves, U. Hounyo, and N. Meddahi (2014). Bootstrapping high-frequency jump tests.

Doyle, J. R. and C. H. Chen (2013). Patterns in stock market movements tested as random number generators. *European Journal of Operational Research 227*(1), 122 – 132.

Fan, J. and Y. Wang (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association 102*(480), 1349–1362.

Fernandez, V. (2008). Traditional versus novel forecasting techniques: How much do we gain? *Journal of Forecasting 27*(7), 637–648.

Gençay, R., F. Selçuk, and B. Whitcher (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics.* Academic Press.

Gençay, R., F. Selcuk, and B. Whitcher (2005). Multiscale systematic risk. *Journal of International Money and Finance 24*(1), 55–70.

Hansen, P. R. and Z. Huang (2012). Exponential garch modeling with realized measures of volatility.

Hansen, P. R., Z. Huang, and H. H. Shek (2012). Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics 27*(6), 877–906.

Hansen, P. R. and A. Lunde (2005). A forecast comparison of volatility models: does anything beat a GARCH (1, 1)? *Journal of Applied Econometrics 20*(7), 873–889.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Haven, E., X. Liu, and L. Shen (2012). De-noising option prices with the wavelet method. *European Journal of Operational Research 222*(1), 104–112.

Høg, E. and A. Lunde (2003). Wavelet estimation of integrated volatility. *Working Paper. Aarhus School of Business*.

Huang, Z., T. Wang, and X. Zhang (2014). Generalized autoregressive score model with realized measures of volatility. *Available at SSRN 2461831*.

Lahaye, J., S. Laurent, and C. J. Neely (2011). Jumps, cojumps and macro announcements. *Journal of Applied Econometrics 26*(6), 893–921.

Lanne, M. (2007). Forecasting realized exchange rate volatility by decomposition. *International Journal of Forecasting 23*(2), 307–320.

LeBaron, B. (2001). Stochastic volatility as a simple generator of apparent financial power laws and long memory. *Quantitative Finance 1*(6), 621–631.

Louzis, D. P., S. Xanthopoulos-Sisinis, and A. P. Refenes (2013). The role of high-frequency intra-daily data, daily range and implied volatility in multi-period value-at-risk forecasting. *To Appear in Journal of Forecasting*.

Lunde, A. and K. V. Olesen (2013). Modeling and forecasting the volatility of energy forward returns-evidence from the nordic power exchange. Technical report, School of Economics and Management, University of Aarhus.

Mallat, S. (1998). *A wavelet tour of signal processing.* Academic Press.

Marroqu, N., M. Moreno, et al. (2013). Optimizing bounds on security prices in incomplete markets. does stochastic volatility specification matter? *European Journal of Operational Research 225*(3), 429–442.

McAleer, M. and M. Medeiros (2008). Realized volatility: A review. *Econometric Reviews* (27), 10–45.

Mcmillan, D. G. and A. E. Speight (2012). Daily FX Volatility Forecasts: Can the GARCH (1, 1) Model be Beaten using High-Frequency Data? *Journal of Forecasting 31*(4), 330–343.

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

Percival, D. B. and H. Mofjeld (1997). Analysis of subtidal coastal sea level fluctuations using wavelets. *Journal of the American Statistical Association 92*(439), 886–880.

Percival, D. B. and A. T. Walden (2000). *Wavelet Methods for Time series Analysis*. Cambridge University Press.

Pun, C. S., S. F. Chung, and H. Y. Wong (2015). Variance swap with mean reversion, multifactor stochastic volatility and jumps. *European Journal of Operational Research 245*(2), 571 – 580.

Ramsey, J. B. (2002). Wavelets in economics and finance: Past and future. *Studies in Nonlinear Dynamics & Econometrics 6*(3).

Serroukh, A., A. T. Walden, and D. B. Percival (2000). Statistical properties and uses of the wavelet variance estimator for the scale analysis of time series. *Journal of the American Statistical Association 95*, 184–196.

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

Subbotin, A. (2008). A multi-horizon scale for volatility. Technical report, Documents de travail du Centre d'Economie de la Sorbonne, Université Panthéon-Sorbonne (Paris 1).

Sun, E. W. and T. Meinl (2012). A new wavelet-based denoising algorithm for high-frequency financial data mining. *European Journal of Operational Research 217*(3), 589–599.

Wang, P., B. Zhang, and Y. Zhou (2015). Asymmetries in stock markets. *European Journal of Operational Research 241*(3), 749 – 762.

Wang, Y. (1995). Jump and sharp cusp detection via wavelets. *Biometrika 82*(2), 385–397.

Yang, J. and D. A. Bessler (2008). Contagion around the october 1987 stock market crash. *European Journal of Operational Research 184*(1), 291 – 310.

Zhang, L., P. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.

Appendix:Tables

Table 1: Results for the GBP futures: in-sample fits of GARCH(1,1), Realized GARCH(1,1) with RV, Realized Jump-GARCH with BV, MedRV, and JWTSRV estimated using MLE (Realized (Jump) GARCH) and GAS (Realized (Jump) GAS), and finally Realized GARCH with multiple $\widehat{IV}_{j,t}^{(JWTSRV)}$ volatility decompositions on different investment horizons. Robust standard errors are reported in parentheses.

	GARCH	Realized (Jump) GARCH				Realized (Jump) GAS					Realized GARCH with multiple JWTSRV$_j$					
		RV	BV	MedRV	JWTSRV	RV	BV	MedRV	JWTSRV		RV	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
ω	0.092	0.040	0.095	0.122	0.161	0.013	0.008	0.010	0.014	ω				0.019		
	(0.059)	(0.030)	(0.034)	(0.036)	(0.036)	(0.008)	(0.009)	(0.009)	(0.010)					(0.009)		
β	0.951	0.757	0.731	0.716	0.707	0.996	0.996	0.996	0.995	β				0.994		
	(0.007)	(0.018)	(0.020)	(0.021)	(0.021)	(0.002)	(0.003)	(0.003)	(0.003)					(0.003)		
γ	0.046	0.224	0.235	0.244	0.242	0.237	0.261	0.275	0.283	γ_j	-0.011	0.128	0.050	0.008	0.023	0.013
	(0.007)	(0.018)	(0.019)	(0.020)	(0.020)	(0.017)	(0.019)	(0.020)	(0.020)		(0.021)	(0.029)	(0.024)	(0.016)	(0.011)	(0.008)
γ_J			0.015	0.008	0.002		0.017	0.009	0.002							
			(0.005)	(0.004)	(0.004)		(0.005)	(0.005)	(0.005)							
τ_1													-0.027			
													(0.005)			
τ_2													0.032			
													(0.004)			
ξ		-0.128	-0.368	-0.455	-0.602	-0.009	-0.229	-0.320	-0.450	ξ_j	-0.273	-1.252	-2.052	-2.875	-3.659	-3.876
		(0.131)	(0.145)	(0.153)	(0.155)	(0.064)	(0.136)	(0.143)	(0.145)		(0.125)	(0.134)	(0.136)	(0.140)	(0.144)	(0.139)
ϕ		1.070	1.127	1.147	1.189	1.040	1.091	1.113	1.150	ϕ_j	1.112	1.168	1.180	1.199	1.198	1.231
		(0.039)	(0.043)	(0.045)	(0.046)	(0.021)	(0.041)	(0.043)	(0.043)		(0.036)	(0.039)	(0.040)	(0.041)	(0.042)	(0.041)
τ_1/d_1		-0.017	-0.024	-0.025	-0.028	0.079	0.064	0.056	0.058	$\delta_{j,1}$	-0.014	-0.028	-0.026	-0.021	-0.020	-0.032
		(0.008)	(0.007)	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)		(0.008)	(0.007)	(0.008)	(0.010)	(0.013)	(0.015)
τ_2/d_2		0.087	0.072	0.066	0.068	-0.003	-0.008	-0.009	-0.010	$\delta_{j,2}$	0.087	0.060	0.065	0.074	0.087	0.209
		(0.006)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)		(0.006)	(0.005)	(0.006)	(0.007)	(0.009)	(0.012)
$\ell(x\|r)$		-613	-533	-486	-506	-625	-546	-501	-523					-2543		
$\ell(r)$	-5849	-5825	-5824	-5823	-5825	-5825	-5825	-5823	-5825					-5825		

Table 2: Results for the CHF futures: in-sample fits of GARCH(1,1), Realized GARCH(1,1) with RV, Realized Jump-GARCH with BV, MedRV, and JWTSRV estimated using MLE (Realized (Jump) GARCH) and GAS (Realized (Jump) GAS), and finally Realized GARCH with multiple $\widehat{IV}_{j,t}^{(JWTSRV)}$ volatility decompositions on different investment horizons. Robust standard errors are reported in parentheses.

	GARCH	Realized (Jump) GARCH				Realized (Jump) GAS					Realized GARCH with multiple JWTSRV$_j$					
		RV	BV	MedRV	JWTSRV	RV	BV	MedRV	JWTSRV		RV	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
ω	0.092	-0.140	-0.095	-0.108	-0.098	0.018	0.016	0.019	0.023	ω				0.033		
	(0.087)	(0.031)	(0.034)	(0.038)	(0.036)	(0.012)	(0.013)	(0.013)	(0.013)					(0.014)		
β	0.936	0.760	0.729	0.707	0.719	0.995	0.994	0.993	0.994	β				0.991		
	(0.008)	(0.018)	(0.019)	(0.020)	(0.020)	(0.003)	(0.003)	(0.003)	(0.003)					(0.004)		
γ	0.066	0.276	0.297	0.324	0.311	0.235	0.267	0.275	0.268	γ_j	0.030	0.262	-0.020	0.068	-0.010	0.031
	(0.008)	(0.021)	(0.022)	(0.023)	(0.022)	(0.018)	(0.019)	(0.019)	(0.019)		(0.033)	(0.043)	(0.037)	(0.026)	(0.018)	(0.012)
γ_J			0.016	0.012	0.004		0.017	0.013	0.002							
			(0.007)	(0.006)	(0.005)		(0.007)	(0.006)	(0.006)							
										τ_1				0.021		
														(0.007)		
										τ_2				0.030		
														(0.005)		
ξ		0.602	0.389	0.412	0.399	0.786	0.572	0.571	0.536	ξ_j	0.677	-0.200	-0.964	-1.711	-2.534	-2.802
		(0.092)	(0.099)	(0.102)	(0.101)	(0.075)	(0.084)	(0.094)	(0.093)		(0.079)	(0.083)	(0.085)	(0.088)	(0.096)	(0.101)
ϕ		0.844	0.885	0.875	0.878	0.791	0.833	0.828	0.837 ϕ_j		0.823	0.846	0.846	0.842	0.850	0.886
		(0.024)	(0.026)	(0.027)	(0.027)	(0.020)	(0.022)	(0.025)	(0.025)		(0.021)	(0.022)	(0.022)	(0.023)	(0.026)	(0.027)
τ_1/d_1		0.030	0.020	0.024	0.019	0.083	0.077	0.059	0.059	$\delta_{j,1}$	0.028	0.022	0.016	0.011	0.025	0.001
		(0.009)	(0.008)	(0.008)	(0.008)	(0.005)	(0.005)	(0.004)	(0.004)		(0.008)	(0.007)	(0.008)	(0.010)	(0.012)	(0.007)
τ_2/d_2		0.098	0.089	0.073	0.074	0.030	0.020	0.025	0.017	$\delta_{j,2}$	0.100	0.071	0.075	0.084	0.087	0.185
		(0.006)	(0.006)	(0.005)	(0.005)	(0.008)	(0.008)	(0.008)	(0.008)		(0.006)	(0.005)	(0.006)	(0.007)	(0.008)	(0.012)
$\ell(x\|r)$		-825	-753	-718	-688	-832	-760	-731	-705					-3286		
$\ell(r)$	-6199	-6164	-6164	-6163	-6164	- 6173	-6170	-6170	-6169					-6167		

Table 3: Results for the EUR futures: in-sample fits of GARCH(1,1), Realized GARCH(1,1) with RV, Realized Jump-GARCH with BV, MedRV, and JWTSRV estimated using MLE (Realized (Jump) GARCH) and GAS (Realized (Jump) GAS), and finally Realized GARCH with multiple $\widehat{IV}_{j,t}^{(JWTSRV)}$ volatility decompositions on different investment horizons. Robust standard errors are reported in parentheses.

	GARCH	Realized (Jump) GARCH				Realized (Jump) GAS					Realized GARCH with multiple JWTSRV$_j$					
		RV	BV	MedRV	JWTSRV	RV	BV	MedRV	JWTSRV		RV	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
ω	0.039	-0.063	-0.022	0.011	0.047	0.026	0.018	0.020	0.028	ω				0.039		
	(0.048)	(0.038)	(0.043)	(0.079)	(0.041)	(0.012)	(0.013)	(0.013)	(0.015)					(0.013)		
β	0.955	0.734	0.698	0.683	0.684	0.992	0.993	0.992	0.992	β				0.989		
	(0.006)	(0.020)	(0.022)	(0.022)	(0.022)	(0.003)	(0.004)	(0.003)	(0.004)					(0.004)		
γ	0.045	0.286	0.313	0.321	0.311	0.260	0.296	0.308	0.308	γ_j	-0.016	0.182	0.036	0.049	-0.008	0.041
	(0.006)	(0.024)	(0.025)	(0.029)	(0.024)	(0.020)	(0.021)	(0.021)	(0.021)		(0.028)	(0.037)	(0.032)	(0.022)	(0.016)	(0.011)
γ_J			0.018	0.012	0.001		0.019	0.014	0.001							
			(0.008)	(0.007)	(0.003)		(0.008)	(0.007)	(0.009)							
										τ_1				-0.027		
														(0.006)		
										τ_2				0.051		
														(0.005)		
ξ		0.302	0.122	0.030	-0.063	0.452	0.297	0.230	0.101	ξ_j	0.155	-0.866	-1.683	-2.434	-3.307	-3.456
		(0.114)	(0.124)	(0.224)	(0.127)	(0.099)	(0.107)	(0.112)	(0.118)		(0.118)	(0.128)	(0.131)	(0.135)	(0.143)	(0.135)
ϕ		0.907	0.945	0.965	0.990	0.869	0.901	0.916	0.950	ϕ_j	0.948	1.015	1.032	1.034	1.064	1.077
		(0.033)	(0.035)	(0.062)	(0.036)	(0.029)	(0.031)	(0.032)	(0.034)		(0.033)	(0.036)	(0.037)	(0.038)	(0.040)	(0.038)
τ_1/d_1		0.001	0.001	-0.004	-0.009	0.088	0.076	0.069	0.067	$\delta_{j,1}$	0.001	-0.007	-0.007	-0.002	0.014	-0.012
		(0.005)	(0.005)	(0.008)	(0.008)	(0.006)	(0.006)	(0.005)	(0.005)		(0.008)	(0.008)	(0.009)	(0.011)	(0.014)	(0.016)
τ_2/d_2		0.093	0.083	0.077	0.075	0.010	0.011	0.008	0.003	$\delta_{j,2}$	0.093	0.072	0.076	0.073	0.075	0.207
		(0.006)	(0.006)	(0.005)	(0.005)	(0.008)	(0.007)	(0.007)	(0.007)		(0.006)	(0.005)	(0.006)	(0.007)	(0.009)	(0.012)
$\ell(x\|r)$		-715	-655	-582	-600	-710	-651	-582	-601					-2555		
$\ell(r)$	-6016	-5994	-5993	-5992	-5990	-5996	-5996	-5994	-5992					-5995		

Table 4: RMSE ($\times 10^{-4}$) from all forecasts for the GBP, CHF, and EUR at different forecasting horizons $h = \{1, 5, 10\}$. Forecasts which fall into the 10% Model Confidence Set (MCS) using both robust MSE and QLIKE loss functions are in bold. In addition, ranking of the models included in the MCS is provided in the superscript, first is ranking using MSE, second using QLIKE.

	Realized (Jump) GARCH				Realized (Jump) GAS				Multiple
	RV	BV	MedRV	JWTSRV	RV	BV	MedRV	JWTSRV	JWTSRV_j
GBP									
$h = 1$	$1.007^{(9)}$	$1.003^{(7)}$	$1.003^{(6)}$	$1.004^{(8)}$	$\mathbf{0.998}^{(5,5)}$	$\mathbf{0.988}^{(4,4)}$	$\mathbf{0.986}^{(3,3)}$	$\mathbf{0.982}^{(2,2)}$	$\mathbf{0.972}^{(1,1)}$
$h = 5$	$\mathbf{0.644}^{(6,5)}$	0.658	0.660	0.661	$\mathbf{0.643}^{(5,3)}$	$\mathbf{0.639}^{(2,2)}$	$\mathbf{0.641}^{(3,4)}$	$\mathbf{0.642}^{(4,6)}$	$\mathbf{0.633}^{(1,1)}$
$h = 10$	$\mathbf{0.561}^{(2,1)}$	$\mathbf{0.566}^{(3,3)}$	$\mathbf{0.569}^{(5,4)}$	$\mathbf{0.574}^{(8,7)}$	$\mathbf{0.572}^{(7,6)}$	$\mathbf{0.568}^{(4,5)}$	$\mathbf{0.571}^{(6,8)}$	$\mathbf{0.575}^{(9,9)}$	$\mathbf{0.558}^{(1,2)}$
CHF									
$h = 1$	$1.497^{(3)}$	$1.459^{(1)}$	$1.517^{(6)}$	$1.496^{(2)}$	$1.591^{(3)}$	$\mathbf{1.497}^{(4,1)}$	$\mathbf{1.530}^{(7,5)}$	$\mathbf{1.538}^{(8,2)}$	$\mathbf{1.509}^{(5,4)}$
$h = 5$	$\mathbf{1.087}^{(5,5)}$	$\mathbf{1.027}^{(1,4)}$	$1.095^{(6)}$	$1.073^{(2)}$	$1.211^{(6)}$	$\mathbf{1.077}^{(3,1)}$	$\mathbf{1.120}^{(7,7)}$	$\mathbf{1.140}^{(8,3)}$	$\mathbf{1.087}^{(4,2)}$
$h = 10$	$\mathbf{1.155}^{(7,2)}$	$\mathbf{1.079}^{(1,1)}$	$\mathbf{1.119}^{(3,7)}$	$\mathbf{1.125}^{(4,5)}$	$1.224^{(9)}$	$\mathbf{1.107}^{(2,4)}$	$\mathbf{1.136}^{(6,8)}$	$\mathbf{1.163}^{(8,6)}$	$\mathbf{1.126}^{(5,3)}$
EUR									
$h = 1$	$1.280^{(9)}$	$1.273^{(6)}$	$1.253^{(4)}$	$\mathbf{1.230}^{(2,3)}$	$\mathbf{1.279}^{(8,5)}$	$1.278^{(7)}$	$\mathbf{1.264}^{(5,4)}$	$\mathbf{1.244}^{(3,2)}$	$\mathbf{1.221}^{(1,1)}$
$h = 5$	0.981	$0.970^{(8)}$	$0.943^{(4)}$	$0.927^{(2)}$	$0.960^{(6)}$	$0.961^{(7)}$	$0.946^{(5)}$	$\mathbf{0.930}^{(3,2)}$	$\mathbf{0.911}^{(1,1)}$
$h = 10$	0.974	0.936	$0.894^{(6)}$	$0.905^{(3)}$	$0.876^{(5)}$	$0.870^{(7)}$	$\mathbf{0.848}^{(3,4)}$	$\mathbf{0.848}^{(2,1)}$	$\mathbf{0.848}^{(1,2)}$

Appendix: Figures

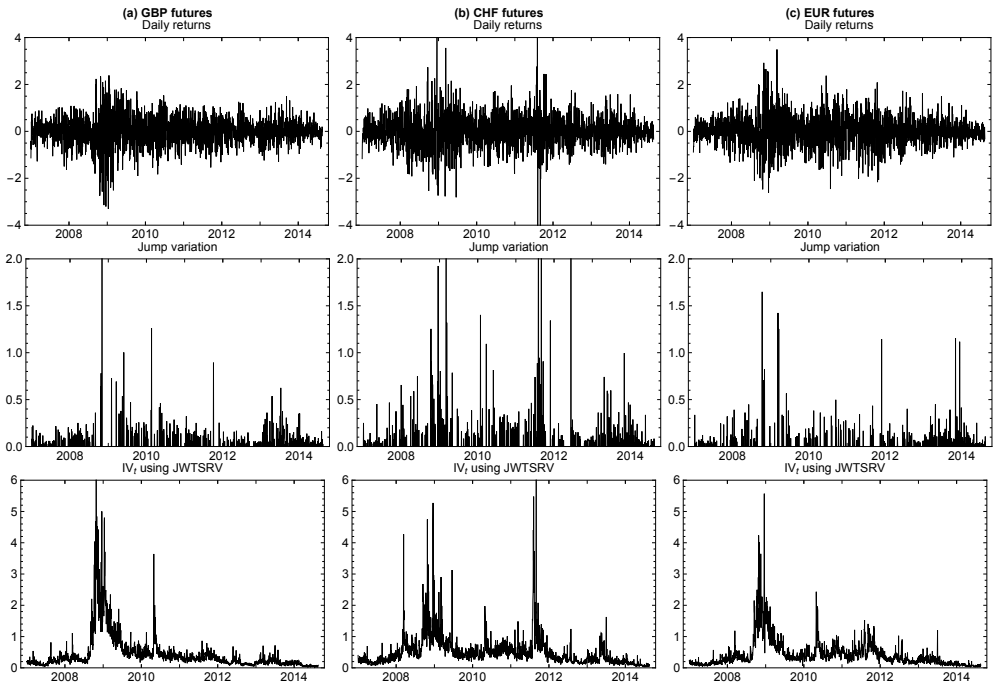

Figure 1: Daily returns, estimated jump variation and IV_t estimated by JWTSRV for (a) GBP, (b) CHF and (c) EUR futures.

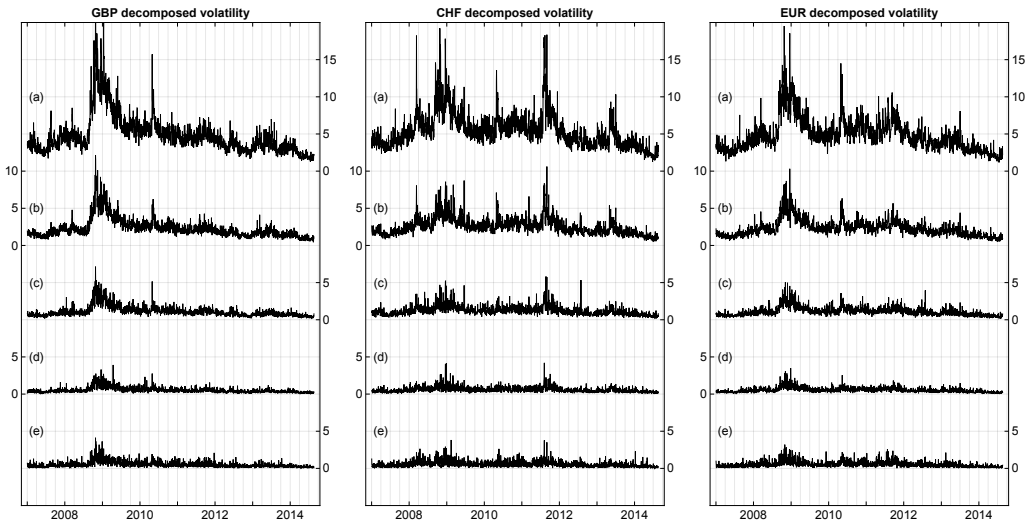

Figure 2: Decomposed annualized volatility (by 252 days) of GBP, CHF and EUR futures using JWTSRV, (a) volatility on investment horizons up to 10 minutes, (b) volatility on investment horizons of 10 to 20 minutes, (c) volatility on investment horizons of 20 to 40 minutes, (d) volatility on investment horizons of 40 to 80 minutes, (e) volatility on investment horizons up to 1 day. Note that sum of components (a), (b), (c), (d) and (e) give total volatility.

Appendix A. Wavelet transform

In this Appendix we briefly introduce basic ideas of wavelet transform. Let us begin with the continuous wavelet transform which is a cornerstone of the wavelet analysis. Further we introduce a special form of discrete wavelet transform called the "maximal overlap discrete wavelet transform" (MODWT) that we use in empirical part. Following Daubechies (1992) and Chui (1992) we define doubly-indexed wavelet function – a wavelet[14] as:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{j}}\psi\left(\frac{t-k}{j}\right) \in L^2(\mathbb{R}), \tag{A.1}$$

where index k determines the exact position of the wavelet in time, whereas the scaling index j controls how the wavelet is stretched or dilated, i.e., frequency resolution of the wavelet. The continuous wavelet transform, $W_{j,k}$, is a projection of a wavelet function $\psi_{j,k}$ onto the time series $y(t) \in L^2(\mathbb{R})$:

$$W_{j,k} = \int_{\infty}^{\infty} y(t)\overline{\psi}_{j,k}(t)dt. \tag{A.2}$$

Hence, Eq.(A.2) tranforms $y(t)$, time-domain process, into $W_{j,k}$ which is time-frequency (or time-scale) space, where k is the position in time and j corresponds to a specific frequency. Using the wavelet coefficients $W_{j,k}$ we can subsequently recover the time series $y(t)$ as follows:

$$y(t) = \frac{1}{C_\psi}\int_0^{\infty}\left[\int_{-\infty}^{\infty} W_{j,k}\psi_{j,k}(t)dk\right]\frac{dj}{j^2}, \quad k{>}0. \tag{A.3}$$

The continuous wavelet transform preserves variance of the analyzed time series. It is an important property that allows us to work with the decomposed wavelet variance. Thus we can write:

$$x^2 = \frac{1}{C_\psi}\int_0^{\infty}\left[\int_{-\infty}^{\infty} |W_{j,k}|^2 dk\right]\frac{dj}{j^2}. \tag{A.4}$$

For a more detailed introduction to continuous wavelet transform and wavelets, see Daubechies (1992), Chui (1992), and Percival and Walden (2000).

[14] An important conditions a wavelet function must fulfill is the admissibility condition: $C_\psi = \int_0^{\infty}\frac{|\Psi(f)|^2}{f}df < \infty$, where $\Psi(f)$ is the Fourier transform of a wavelet $\psi(.)$. For more details about wavelet filer conditions see Daubechies (1992)

In empirical applications we work with discrete time series, thus some form of discretization is needed. The discrete wavelet transform (DWT)[15], which is a parsimonious form of the continuous wavelet transform allows for an analysis of discrete time series where only a bounded number of scales is required. The discrete version of wavelet transform has, however, some serious limitation that make its application to real time series rather difficult. These are mainly the sample size restriction to the power of two and the starting point sensitivity of the wavelet transform.

Appendix A.1. Maximal overlap discrete wavelet transform

The MODWT is in some cases superior to the DWT for empirical data analysis. For example, the problem of sample length restriction is connected with downsampling procedure of the DWT. However, the construction of MODWT does not use downsampling, thus vectors of the wavelet coefficients at all scales have equal length, corresponding to the length of transformed time series. As a consequence, the MODWT is not restricted to any sample size. In addition, the MODWT is a translation-invariant; therefore, it is not sensitive to the choice of the starting point of the examined time series. Similarly as the CWT, the MODWT wavelet and scaling coefficients can be used for analysis of variance of a time series in the time-frequency domain. Statistical properties of the MODWT variance estimators for non-stationary and non-Gaussian processes are discussed in detail in Serroukh et al. (2000). For additional details on the MODWT, see Mallat (1998) and Percival and Walden (2000).

For computation of the MODWT coefficients we apply the pyramid algorithm of Mallat (1998). The procedure is based on filtering time series with MODWT wavelet filters; the filtered time series is then filtered again in a subsequent stages to obtain other wavelet scales. These scales contain information localized at corresponding frequency bands of analyzed time series.

Let us briefly introduce the pyramid algorithm. In the first stage, the wavelet coefficients are obtained via circular filtering of time series y_t using the MODWT wavelet and scaling filters $h_{1,l}$ and $g_{1,l}$ (Percival and Walden, 2000):

$$\mathcal{W}_{1,k} \equiv \sum_{l=0}^{L-1} h_{1,l}\ y_{k-l\ mod N}, \qquad \mathcal{V}_{1,k} \equiv \sum_{l=0}^{L-1} g_{1,l}\ y_{k-l\ mod N}, \tag{A.5}$$

[15]For a definition and detailed discussion of the discrete wavelet transform, see Mallat (1998), Percival and Walden (2000), and Gençay et al. (2002).

where $L_j = 2^{j-1}(L-1)+1$ defines a width of the wavelet and scaling filters.[16] After the first stage we obtain the wavelet and scaling coefficients at the first scale ($j = 1$). The algorithm continues with the second stage where instead of y_t we filter the sequence of scaling coefficients from the first stage $\mathcal{V}_{1,k}$, using the MODWT wavelet and scaling filters $h_{2,l}$ and $g_{2,l}$ for the second scale, i.e.,

$$\mathcal{W}_{2,k} \equiv \sum_{l=0}^{L-1} h_{2,l}\mathcal{V}_{1,k-l \ modN}, \qquad \mathcal{V}_{2,k} \equiv \sum_{l=0}^{L-1} g_{2,l}\mathcal{V}_{1,k-l \ modN}. \tag{A.6}$$

We may continue with more stages until the level of decomposition is $j \leq log_2(N)$. For example, in case we need two levels of decomposition, i.e, we apply two stages, we obtain two vectors of wavelet coefficients; $\mathcal{W}_{1,k}$, $\mathcal{W}_{2,k}$ and a vector of the scaling coefficients at scale two $\mathcal{V}_{2,k}$, where $k = 0, 1, \ldots, N-1$. Vectors of wavelet and scaling coefficients reflect variations at specific frequency bands. Generally, $\mathcal{W}_{j,\cdot}$ represents a frequency band $f[1/2^{j+1}, 1/2^j]$, whereas $\mathcal{V}_{j,\cdot}$ represents a frequency band $f[0, 1/2^{j+1}]$.

[16] For more information about wavelet filters see for example Percival and Walden (2000).