Journal of International Money and Finance 79 (2017) 1–19

Contents lists available at ScienceDirect

Journal of International Money and Finance

journal homepage: www.elsevier.com/locate/jimf

The role of jumps and leverage in forecasting volatility in international equity markets ☆

Daniel Buncic[a,*], Katja I.M. Gisler[b]

[a] Sveriges Riksbank, Sweden
[b] Swiss Life, Switzerland

ARTICLE INFO

Article history:
Available online 5 September 2017

JEL classification:
C13
C15
C50
F30
F44

Keywords:
Realized volatility
Jumps
The leverage effect
HAR modelling and forecasting
International equity markets

ABSTRACT

We analyze the importance of jumps and the leverage effect on forecasts of realized volatility in a large cross-section of 18 international equity markets, using daily realized measures data from the Oxford-Man Realized Library, and two widely employed empirical models for realized volatility that allow for jumps and leverage. Our out-of-sample forecast evaluation results show that the separation of realized volatility into a continuous and a discontinuous (jump) component is important for the S&P 500, but of rather limited value for the remaining 17 international equity markets that we analyze. Only for 6 equity markets are significant and sizable forecast improvements realized at the one-step-ahead horizon, which, nevertheless, deteriorate quickly and abruptly as the prediction horizon increases. The inclusion of the leverage effect, on the other hand, has a much larger impact on all 18 international equity markets. Forecast gains are not only highly significant, but also sizeable, with gains remaining significant for forecast horizons of up to one month ahead.

1. Introduction

Measuring and forecasting financial market volatility is central to asset pricing, asset allocation, and risk management. With the availability of high-frequency data, new non-parametric estimators known as *realized volatility* have been developed to provide better measurements of volatility. Subsequently, a large body of literature has emerged and proposed a variety of suitable models of realized volatility that are designed to capture prominent stylized facts such as volatility persistence, fat tails induced by jumps, and the leverage effect.

One of the most widely used empirical models for realized volatility is the Heterogeneous Autoregressive (HAR) model of Corsi (2009). Because the HAR model captures the persistence of volatility in an extremely parsimonious and intuitive way, it has become the benchmark model to beat in out-of-sample forecast evaluations. Since its introduction, the HAR model has undergone numerous refinements in order to better capture additional stylized facts of volatility. For instance, Andersen

☆ We are grateful to Roberto Renó, Frank Diebold, Paolo Giordani, Kevin Sheppard, Xin Zhang and seminar participants at the University of St. Gallen and the University of Pennsylvania for comments that helped to improve the paper. The opinions expressed in this article are the sole responsibility of the authors and should not be interpreted as reflecting the views of Sveriges Riksbank.

* Corresponding author at: Research and Modelling Division, Financial Stability Department, Sveriges Riksbank, SE-103 37 Stockholm, Sweden.
E-mail addresses: daniel.buncic@riksbank.se (D. Buncic), katjagisler@gmx.ch (K.I.M. Gisler).
URL: http://www.danielbuncic.com (D. Buncic).

http://dx.doi.org/10.1016/j.jimonfin.2017.09.001
0261-5606/

et al. (2007) propose a model for realized volatility that is based on the HAR framework of Corsi (2009), but which separates the continuous and the discontinuous (jump) parts of realized volatility. Corsi et al. (2012) extend the model of Andersen et al. (2007) to also allow for the leverage effect. More recently, Patton and Sheppard (2015) consider signed jumps as well as good and bad volatility measures in their specification. The overall conclusion from these studies is that jumps, the leverage effect and the separation of volatility into good and bad measures are all important for the modelling and forecasting of realized volatility.

The majority of empirical studies that assess the importance of adding jump and leverage components to the HAR model have focused exclusively on U.S. based asset classes, that is, mainly the S&P 500 index, its constituents, or S&P 500 futures.[1] To the best of our knowledge, there exists no empirical study so far that analyses the role of jumps and the leverage effect for realized volatility forecasts at an international cross-sectional equity market level. Prokopczuk et al. (2016) have recently found jumps to be overall statistically significant in-sample, but of no predictive value when constructing out-of-sample forecasts of realized volatility for four energy futures. Given this overall negative finding, it is of interest to quantify how much of a predictive improvement in volatility one can expect from the inclusion of jumps and the leverage effect when considering volatility forecasts in international equity markets.

The goal of this study is to fill this gap in the literature. We provide a first comprehensive analysis of the importance of jumps and the leverage effect on forecasts of realized volatility in a large cross-section of 18 international equity markets. We use daily realized measures data from the Oxford-Man Realized Library and two widely employed empirical models for realized volatility that allow for jumps and the leverage effect. Considering a sample period from January 3, 2000, to July 13, 2016 and taking the simple HAR model of Corsi (2009) as the benchmark model for realized volatility, we proceed by first analysing the in-sample value of separating realized volatility into a continuous and discontinuous (jump) component as well as adding the leverage effect to the model, and then assess their predictive contributions in an out-of-sample forecast evaluation. Our in-sample findings show that, with the exception of the S&P 500, the separation of realized volatility into a continuous and a jump component results in rather mild (in-sample) R^2 improvements of at most 1.37 percentage points (Brazilian Bovespa index), with most of the improvements being less than 25 basis points. In contrast, the benefit of adding the leverage effect to the HAR model is considerably larger, resulting in (in-sample) R^2 increases of at least 1.23 percentage points (South Korean KOSPI index), and up to 3.90 percentage points (Brazilian Bovespa index), with more than half of the improvements being greater than 1.92 percentage points.

Our in-sample findings also hold out-of-sample. The contribution of jumps on forecasts of realized volatility is strong for the S&P 500, but rather weak and mostly unimportant for the remaining 17 international equity markets. For the S&P 500 index, the out-of-sample R^2 is over 8% at the one-step-ahead horizon. However, only for 6 of the 17 non-U.S. equity markets do we find statistically significant improvements, with out-of-sample R^2 values, nevertheless, being considerably lower, ranging between 1.65% and 4.19%. Moreover, these improvements are also only short-lived. The inclusion of the leverage effect in the HAR model substantially improves out-of-sample forecasts for all 18 international equity markets that we study. At the one-step-ahead horizon, we find uniformly larger out-of-sample forecast improvements across all 18 equity markets, producing out-of-sample R^2 values between 2.71% and 12.18%. These improvements are statistically significant and have a lasting impact that can affect forecasts as far as 1 month ahead. More specifically, we find significant improvements for 16, 8, and 2 of the 18 equity markets at the 5, 10, and 22 days-ahead forecast horizons, respectively. In summary of our analysis, we find the separation of realized volatility into a continuous and a discontinuous (jump) component to be beneficial for the S&P 500 equity index, but only of limited value for the remaining non-U.S. equity markets. The inclusion of the leverage effect, nevertheless, leads uniformly to substantially better forecast of realized volatility in all 18 international equity markets that we study.

The remainder of the paper is structured as follows. In Section 2 we describe the volatility modelling strategy that we implement. Here we briefly outline the theoretical background that links realized volatility to the theory of quadratic variation and also describe in detail the empirical econometric framework that we employ to model jumps and the leverage effect. The data that is used in the study is described in Section 3. The importance of jumps and the leverage effect for the determination of realized volatility in international equity markets is assessed in Section 4, where we implement an in-sample fitting, as well as an out-of-sample forecast evaluation analysis. Section 6 concludes the study.

2. Volatility modelling framework

Before we describe the empirical model that we use to assess the importance of jump and leverage effects on realized volatility forecasts in international equity markets, we briefly outline the background that links empirical realized volatility and jumps to its theoretical counterpart, quadratic variation.

2.1. Realized volatility, bipower variation and jumps

Denote by p_t the logarithm (log) of an asset price at time t. The log-price p_t is assumed to follow a continuous-time diffusion process driven by Brownian motion, with its dynamics determined by the following stochastic differential equation:

[1] Note that Andersen et al. (2007) assess the model also on 30 year treasure futures and the DM/$ (Deutsch Mark/U.S. Dollar) spot exchange rate.

$$dp_t = \mu_t dt + \sigma_t dW_t + \kappa_t dq_t, \tag{1}$$

where μ_t is a predictable and locally bounded drift term, σ_t is a càdlàg volatility process which is bounded away from zero, W_t is standard Brownian motion, and q_t is a counting process with (possibly) time varying intensity. The quadratic variation (QV) of the log-price process p_t is given by:

$$\mathrm{QV}_t = \underbrace{\int_0^t \sigma_s ds}_{\text{continuous component}} + \underbrace{\sum\nolimits_{0<s\leqslant t} \kappa_s^2}_{\text{discontinuous component}}, \tag{2}$$

where the first term $\int_0^t \sigma_s ds$ on the right-hand side of (2) is the integrated variance (IV) of the process and captures the continuous component of quadratic variation, while the second term is the squared jump component between 0 and t, and corresponds to the discontinuous component of quadratic variation in (2).

An estimate of QV_t in (2) is referred to as realized variance. By definition, the simplest consistent estimator of realized variance is the sum of discretely observed squared intraday log returns. More formally, let $r_{t,i} = p_{t,i} - p_{t,i-1}$ denote the time t log-return observed in the ith interval of an equidistant grid, with a total of M intervals. Then, the realized variance (RV) estimator is defined as:

$$\mathrm{RV}_t = \sum_{i=1}^{M} r_{t,i}^2. \tag{3}$$

As the intraday sampling frequency increases, realized variance converges uniformly in probability to the quadratic variation process, or, more formally:

$$\operatorname*{plim}_{M\to\infty} \mathrm{RV}_t = \mathrm{QV}_t, \tag{4}$$

(see Andersen et al. (2001, 2003), among others). RV_t is thus a consistent estimator of quadratic variation. One drawback of QV_t, and thus also its estimator RV_t, is that it captures both the integrated variance as well as the jump variation of the process p_t. This can be seen from the definition of QV_t in (2). Due to this, several jump robust estimators designed to capture only the continuous component of quadratic variation have been proposed. The most widely used one is bipower variation (BPV_t) of Barndorff-Nielsen and Shephard (2004, 2005, 2006). Formally, bipower variation is defined as:

$$\mathrm{BPV}_t = \frac{\pi}{2}\left[\frac{M}{M-2}\right]\sum_{i=2}^{M} |r_{t,i}||r_{t,i-1}|, \tag{5}$$

where $\left[\frac{M}{M-2}\right]$ is a finite sample bias correction term.

Since RV_t estimates both, the continuous as well as the discontinuous (jump) component of quadratic variation, while BPV_t captures only the continuous component, the jump component can be identified simply by the difference of the two estimators as:

$$\operatorname*{plim}_{M\to\infty}(\mathrm{RV}_t - \mathrm{BPV}_t) = \sum_{0<s\leqslant t} \kappa^2(s). \tag{6}$$

In finite samples, $(\mathrm{RV}_t - \mathrm{BPV}_t)$ can be negative. To ensure positivity of the daily jump estimate, Barndorff-Nielsen and Shephard (2004) and Andersen et al. (2007) suggest to use only the positive part of $(\mathrm{RV}_t - \mathrm{BPV}_t)$, that is, to employ the following truncated jump definition:

$$J_t = \max\{\mathrm{RV}_t - \mathrm{BPV}_t, 0\}. \tag{7}$$

The continuous component then follows as:

$$\begin{aligned} C_t &= \mathrm{RV}_t - J_t \\ &= \mathrm{RV}_t - \max\{\mathrm{RV}_t - \mathrm{BPV}_t, 0\}. \end{aligned} \tag{8}$$

We will work with the continuous and discontinuous components as defined in (7) and (8) in our analysis, which are in the spirit of the original definitions formulated in Barndorff-Nielsen and Shephard (2004).[2]

2.2. Empirical volatility models

We use the Heterogeneous Autoregressive (HAR) class of models introduced by Corsi (2009) into the realized variance modelling literature as our empirical framework to asses the importance of jump and leverage effects on realized volatility

[2] Note here that one can also identify the jump component J_t as only those that are deemed '*significant*' from a jump detection test. There exist several such tests in the literature (see Huang and Tauchen (2005), Barndorff-Nielsen and Shephard (2006) or Corsi and Renó (2012) and others). However, the construction of these tests require high-frequency intraday data. Since we are working with a publicly available database that contains aggregate daily realized measures and we do not have access to the high-frequency intraday data from which these were constructed, it is not feasible for us to implement such jump tests.

forecasts in international equity markets. The HAR model has a cascade type lag structure, where volatility at any point in time is constructed as a linear combination of a daily, weekly and monthly volatility component. This temporal cascade structure is motivated by the so-called Heterogeneous Market Hypothesis (HMH) of Müller et al. (1993), where agents are assumed to have different time horizons for trading (see Corsi (2009) for a more detailed discussion).

Formally, the HAR model that we employ takes the following form[3]:

$$\log(\mathrm{RV}_{t+1}) = \beta_{0,\mathrm{RV}} + \beta^{(d)}_{\mathrm{RV}} \log(\mathrm{RV}^{(d)}_t) + \beta^{(w)}_{\mathrm{RV}} \log(\mathrm{RV}^{(w)}_t) + \beta^{(m)}_{\mathrm{RV}} \log(\mathrm{RV}^{(m)}_t) + \epsilon^{\mathrm{RV}}_{t+1}, \tag{9}$$

where $\epsilon^{\mathrm{RV}}_{t+1}$ is an innovation term, and the individual daily, weekly, and monthly HAR components are, respectively, defined as:

$$\begin{aligned}
\log(\mathrm{RV}^{(d)}_t) &= \log(\mathrm{RV}_t),\\
\log(\mathrm{RV}^{(w)}_t) &= \frac{1}{5}\sum_{j=1}^{5} \log(\mathrm{RV}_{t+1-j}), \quad \text{and}\\
\log(\mathrm{RV}^{(m)}_t) &= \frac{1}{22}\sum_{j=1}^{22} \log(\mathrm{RV}_{t+1-j}).
\end{aligned}$$

One of the key attractions of the HAR model in (9) is its computational simplicity. It can be estimated by ordinary least squares (OLS) regression, once the daily, weekly, and monthly volatility components are constructed. Moreover, it is an extremely difficult to beat benchmark model for realized volatility in out-of-sample forecast evaluations (see Corsi et al. (2012) for a recent survey).

We consider two different HAR model specifications to assess the importance of jumps and the leverage effect on realized volatility. The two specifications that we consider are:

(i) the HAR-CJ model of Andersen et al. (2007), and
(ii) the HAR-CJL model of Corsi and Renó (2012).[4]

The HAR-CJ model of Andersen et al. (2007) separates each of the realized volatility components of the HAR model into its continuous and discontinuous part, having the form:

$$\begin{aligned}
\log(\mathrm{RV}_{t+1}) &= \beta_{0,\mathrm{CJ}}\\
&\quad + \overbrace{\beta^{(d)}_{\mathrm{CJ}} \log(C^{(d)}_t) + \beta^{(w)}_{\mathrm{CJ}} \log(C^{(w)}_t) + \beta^{(m)}_{\mathrm{CJ}} \log(C^{(m)}_t)}^{\text{HAR structure to capture the continuous part}} + \underbrace{\gamma^{(d)}_{\mathrm{CJ}} \log(1+J^{(d)}_t) + \gamma^{(w)}_{\mathrm{CJ}} \log(1+J^{(w)}_t) + \gamma^{(m)}_{\mathrm{CJ}} \log(1+J^{(m)}_t)}_{\text{HAR structure to capture the discontinuous/jump part}} + \epsilon^{\mathrm{CJ}}_{t+1},
\end{aligned} \tag{10}$$

where the $\left\{\log(C^{(n)}_t)\right\}_{n\in\{d,w,m\}}$ HAR components in (10) of the continuous part are constructed in the same way as for $\left\{\log(\mathrm{RV}^{(n)}_t)\right\}_{n\in\{d,w,m\}}$. In the construction of the corresponding daily, weekly and monthly HAR jump components we follow Corsi and Renó (2012) and compute these as aggregates, that is, as $J^{(d)}_t = J_t$, $J^{(w)}_t = \sum_{j=1}^{5} J_{t-j+1}$ and $J^{(m)}_t = \sum_{j=1}^{22} J_{t-j+1}$, rather than averages.

The HAR-CJL model of Corsi and Renó (2012) adds a leverage effect to the HAR-CJ model in (10), taking the form:

$$\begin{aligned}
\log(\mathrm{RV}_{t+1}) &= \beta_{0,\mathrm{CJL}} + \beta^{(d)}_{\mathrm{CJL}} \log(C^{(d)}_t) + \beta^{(w)}_{\mathrm{CJL}} \log(C^{(w)}_t) + \beta^{(m)}_{\mathrm{CJL}} \log(C^{(m)}_t) + \gamma^{(d)}_{\mathrm{CJL}} \log(1+J^{(d)}_t) + \gamma^{(w)}_{\mathrm{CJL}} \log(1+J^{(w)}_t) + \gamma^{(m)}_{\mathrm{CJL}}\\
&\quad \times \log(1+J^{(m)}_t) + \underbrace{\alpha^{(d)}_{\mathrm{CJL}} r^{(d)-}_t + \alpha^{(w)}_{\mathrm{CJL}} r^{(w)-}_t + \alpha^{(m)}_{\mathrm{CJL}} r^{(m)-}_t}_{\text{HAR structure to capture the leverage effect}} + \epsilon^{\mathrm{CJL}}_{t+1},
\end{aligned} \tag{11}$$

where $r^{(n)-}_t = \min\{r^{(n)}_t, 0\}$ for all $n \in \{d,w,m\}$ is a HAR structured leverage effect, with $r^{(d)}_t = r_t, r^{(w)}_t = \frac{1}{5}\sum_{j=1}^{5} r_{t-j+1}$ and $r^{(m)}_t = \frac{1}{22}\sum_{j=1}^{22} r_{t-j+1}$ being the daily, weekly and monthly returns. The two innovation terms in these models are denoted by $\epsilon^{\mathrm{CJ}}_{t+1}$ and $\epsilon^{\mathrm{CJL}}_{t+1}$, respectively.

[3] We follow the suggestions in Andersen et al. (2007) and work with log-transformed RV data. It is well known that log-transforming data has a number of appealing properties such as, for instance, stabilizing the data, which makes the use of linear time series models more appropriate in our context. The disadvantage of not log-transforming the data can be most easily appreciated by looking at a time-series plot of the recursively estimated HAR model parameters needed to construct the out-of-sample forecasts. These move around erratically when not working with log-transformed RV data, and evolve smoothly over time with log-transformed RV data.

[4] This model is referred to as the Leverage HAR with Continuous volatility and Jumps and is abbreviated as LHAR-CJ in Corsi and Renó (2012). For consistency with the notation and labelling that we are using, we will abbreviate this model as HAR-CJL instead, which can be read as HAR with Continuous volatility, Jumps and Leverage.

For notational simplicity, we will work with the following more compactly re-written form of the three HAR models defined in (9)–(11) in the analysis that follows:

$$\text{HAR-RV}: y_{t+1} = \boldsymbol{x}_t^{\text{RV}} \boldsymbol{\beta}_{\text{RV}} + \epsilon_{t+1}^{\text{RV}} \tag{12a}$$

$$\text{HAR-CJ}: y_{t+1} = \boldsymbol{x}_t^{C} \boldsymbol{\beta}_{\text{CJ}} + \boldsymbol{x}_t^{J} \boldsymbol{\gamma}_{\text{CJ}} + \epsilon_{t+1}^{\text{CJ}} \tag{12b}$$

$$\text{HAR-CJL}: y_{t+1} = \boldsymbol{x}_t^{C} \boldsymbol{\beta}_{\text{CJL}} + \boldsymbol{x}_t^{J} \boldsymbol{\gamma}_{\text{CJL}} + \boldsymbol{x}_t^{L} \boldsymbol{\alpha}_{\text{CJL}} + \epsilon_{t+1}^{\text{CJL}} \tag{12c}$$

where $y_{t+1} = \log(\text{RV}_{t+1})$ denotes the logarithm of realized volatility (RV), and the two (1×4) vectors of (log transformed) realized volatility and continuous HAR components (including an intercept term) are defined, respectively, as:

$$\boldsymbol{x}_t^{\text{RV}} = \left[1 \log(\text{RV}_t^{(d)}) \log(\text{RV}_t^{(w)}) \log(\text{RV}_t^{(m)})\right], \quad \text{and}$$
$$\boldsymbol{x}_t^{C} = \left[1 \log(C_t^{(d)}) \log(C_t^{(w)}) \log(C_t^{(m)})\right].$$

The two (1×3) vectors, which contain the HAR structured jump and leverage components, are denoted by:

$$\boldsymbol{x}_t^{J} = \left[\log(1+J^{(d)}) \log(1+J^{(w)}) \log(1+J^{(m)})\right], \quad \text{and}$$
$$\boldsymbol{x}_t^{L} = \left[r_t^{(d-)} r_t^{(w-)} r_t^{(m-)}\right].$$

The $\boldsymbol{\beta}_{(\cdot)}, \boldsymbol{\gamma}_{(\cdot)}$, and $\boldsymbol{\alpha}_{(\cdot)}$ terms in (12) are corresponding conformable parameter vectors.

3. Data

We use daily realized measures data from the publicly available Oxford-Man Institute's Quantitative Finance Realized Library of Heber et al. (2009) in our analysis. The library contains realized volatility data for 4 U.S. and 17 foreign (non-U.S.) equity price indices from January 3, 2000 to the present. Our last update of the entire data set is on July 13, 2016.[5] The RV and BPV measures that we employ are based on 5 min intraday return intervals (the estimators under the heading '∗.rv' and '∗.bv5' in the Oxford-Man Realized Library under each equity market data block). We transform all variance measures to annualised volatilities.[6] The leverage effect is based on log-returns computed from the provided close-prices (under the heading '∗.closeprice' in the Oxford-Man Realized Library) for each equity market, and is multiplied by 100 to be expressed in percentage terms.

The 17 international (non-U.S.) equity markets that are included in the Oxford-Man Realized Library are: the FTSE 100 (United Kingdom), the Nikkei 225 (Japan), the DAX 30 (Germany), the All Ordinaries (Australia), the CAC 40 (France), the Hang Seng (Hong Kong), the KOSPI (South Korea), the AEX (The Netherlands), the Swiss Market Index (Switzerland), the IBEX 35 (Spain), the S&P CNX Nifty (India), the IPC Mexico (Mexico), the Bovespa (Brazil), the S&P TSX (Canada), the Euro STOXX 50 (Euro area), the FT Straits Times (Singapore), and the FTSE MIB (Italy). For the U.S., we use the S&P 500 as the representative equity price index, which is the most widely analyzed equity index in the world. Existing results on the S&P 500 therefore serve as a benchmark that our cross-sectional findings can be compared to. We provide standard summary statistics as well as time series plots of all data that we use in our study as background information in an online appendix to conserve space.[7]

4. Assessing the value of the jump and leverage effect

We follow the extensive forecasting literature and assess the importance of jumps and leverage effects on realized volatility in international equity markets by initially conducting an in-sample analysis, and then extend our evaluation to an out-of-sample forecast environment.

4.1. In-sample evaluation

We fit all three HAR models in (12) over the full available sample period for each of the 18 considered equity markets to initially gauge the magnitude and significance of the estimated parameters of the relations in (12). These estimation results

[5] The data are available from http://realized.oxford-man.ox.ac.uk/data. Note that the indices have in general marginally different starting and ending dates due to differences in public holidays and trading day closings on the last date the data were updated.

[6] This is done by taking the volatility measures from the Oxford-Man Library and re-scaling them by $100^2 \times 252$, and then taking the square root to be interpreted as annualised volatility (in percentage terms). That is, annualized RV and BPV are computed as '$(*.\text{rv}' \times 100^2 \times 252)^{1/2}$ and '$(*.\text{bv5}' \times 100^2 \times 252)^{1/2}$. Note here also that Liu et al. (2015) have emphasised in a recent study that a simple five minute interval based RV estimator is the most consistent estimator of realized volatility and that there is very little evidence to suggest that it can be outperformed by any of the other realized measures that are considered. We therefore use the five minute based RV as well as BPV estimators provided in the Oxford-Man realized measures library.

[7] This online appendix is available from http://www.danielbuncic.com/pdf/jumps_online_appendix.pdf. The Oxford-Man realized measures database is also described in more detail in Section 3 of Buncic and Gisler (2016), which also describes the time periods where some data are missing from the Oxford-Man RV library. The online appendix reports the proportion of missing data for each equity market index that is used in the fifth column.

are reported in Tables 1–3, respectively. In each of these tables, the first three columns show the foreign equity index of interest, the time period over which the models were fitted, as well as the available sample size (denoted by T). In the fourth to penultimate columns, point estimates of the parameters of interest $\boldsymbol{\beta}, \boldsymbol{\gamma}$, and $\boldsymbol{\alpha}$ are reported, with (2-sided) p-values based on a Heteroskedasticity and Autocorrelation Consistent (HAC) variance/covariance matrix estimator in square brackets below the estimates. The last column shows the sample R^2 values, with round brackets below the results showing also the change in R^2 relative to the benchmark HAR-RV model, that is, $R^2_{\mathcal{M}} - R^2_{\mathrm{RV}}, \forall \mathcal{M} \in \{\mathrm{CJ}, \mathrm{CJL}\}$ for the HAR-CJ and HAR-CJL models.[8]

Although the estimation results reported in Tables 1–3, are rather self-contained, and as such, do not require much further discussion, let us point out some key results that are interesting to highlight here.[9] From the HAR-RV model estimation results in Table 1, it is evident that the baseline model captures a substantial part of the overall variation in realized volatility in all 18 equity markets that we analyze. The full sample R^2 values are between 53% for the Brazilian Bovespa and as high as 79% for the South Korean KOSPI, with an average cross-sectional R^2 value of around 70%. Examining the size of the estimates of the daily, weekly and monthly components, a fairly homogeneous picture emerges across all 18 equity markets. On average, the weekly component receives the largest weight of approximately 0.42, with the daily one receiving the second largest of around 0.33, while the monthly component has the smallest weight of about 0.20. By and large, our cross-sectional HAR-RV estimates are in line with the range of estimates typically reported for the S&P 500 of around 0.3, 0.4 and 0.2 for the daily, weekly and monthly components, respectively (see for instance Andersen et al. (2007), Corsi et al. (2012) and Patton and Sheppard (2015)). Moreover, from the p-values below the estimates, it is clear that all components are highly significant for all 18 international equity markets. Thus, the baseline HAR-RV model constitutes a rather good fit to RV data.

How much does the in-sample fit improve when we separate the *'empirical'* QV process into a continuous and a jump component? The estimation results for the HAR-CJ model shown in Table 2 seem to suggest only little improvement in general. Examining initially the R^2 values to get a broad sense of how much of an improvement in model fit is realized from this separation, one can see that the largest improvement, relative to the baseline HAR-RV, is obtained for the S&P 500, with a gain in R^2 of about 2.5%, followed by the Bovespa, the IPC Mexico and the Indian S&P CNX Nifty, all three yielding R^2 gains of about 1.3%. For the 14 remaining equity markets that we consider, the improvement is less than 1%, with 12 of them showing improvements of less than 0.5%, of which 9 improve by less than 0.25%. Comparing the magnitude of the estimated $\boldsymbol{\beta}_{\mathrm{CJ}}$ and $\boldsymbol{\beta}_{\mathrm{RV}}$ parameters, it can be seen that there are only small differences, with the S&P 500 showing the biggest change in all three HAR components. Overall, the components on the continuous part remain highly significant. Estimation results for the HAR jump components ($\boldsymbol{\gamma}_{\mathrm{CJ}}$) are rather mixed. Only 12, 7 and 12 of the 18 cross-sectional daily, weekly, and monthly components are significantly greater than zero at the 10% level. Nevertheless, due to the large sample size that is available for estimation, despite their small magnitude for most of the equity markets, we can view the improvements from this separation in fit as population values, and thus also statistically significant when tested jointly via an F-test (see the discussion in footNote 10 below for a more detailed explanation).

A finding that is particularly interesting to see here is the negative (and significant) daily jump component for the S&P 500 and the All Ordinaries. Andersen et al. (2007) (see Table 4B on page 716 of their paper) also report a negative daily (and weekly) estimate of the jump component. Such negative parameter estimates have the rather counterintuitive interpretation that jumps observed today have a decreasing effect on future realized volatility.

Estimation results for the HAR-CJL model reported in Table 3 show that leverage effects have a considerably more important impact on RV than jumps. This can be seen not only from the larger R^2 values relative to the HAR-RV benchmark model, but also from the overall *'significance'* of the coefficients on the HAR leverage components. All R^2_{CJL} values improve by at least one additional percentage point relative to the HAR-CJ model, with results for 4 of the 18 equity markets improving by more than two percentage points. The largest gains are realized for the Bovespa, the Euro STOXX 50 and the All Ordinaries, with gains of around 2.5%.[10]

All daily and weekly leverage components have a highly significant impact on future realized volatility (p-values ≤ 0.001), with coefficients showing the expected negative sign that is consistent with the traditional leverage effect narrative of a current decline in equity prices (return < 0) leading to future (expected) increases in (realized) volatility. It is inter-

[8] We use a standard Bartlett Kernel and a Newey and West (1994) rule of thumb bandwidth set equal to $4(T/100)^{2/9}$ in the computation of the HAC standard errors. Due to the large samples that are available to us, these results are extremely robust to different bandwidth choices.

[9] We are not aware of any other study that provides a large international assessment of how well the baseline HAR-RV model fits to realized volatility data, and what the cross-sectional distribution of the daily, weekly and monthly HAR components for RV looks like. The estimates reported in Table 2 in Buncic and Gisler (2016) condition on U.S. RV as well as VIX information, thus are somewhat different in magnitude.

[10] Note here, that, for consistency of the reporting of the tables, we show the gain relative to the HAR-RV benchmark model in the last column of Table 3. Nevertheless, it is easy to compute the gain relative to the HAR-CJ model as $R^2_{\mathrm{CJL}} - R^2_{\mathrm{CJ}}$. Also, we only report R^2 values here, rather than an adjusted R^2 and/or an F-statistic to gauge the *'significance'*. The reason for that is simply that we have rather large sample sizes (over 4000 observations for most of the equity markets) so that we can view the small differences in R^2 nearly as population increases. For instance, the relation between the R^2 and the adjusted R^2 ($\overline{R}^2$ henceforth) is: $\overline{R}^2 = 1 - \frac{(T-1)}{(T-K)}[1 - R^2]$. When $T = 4000$ and $K = 10$ (number of regressors plus constant), the ratio $\frac{(T-1)}{(T-K)} = \frac{4000-1}{4000-10} \approx 1.002255$ so that for a baseline R^2 of 0.70 we obtain an $\overline{R}^2$ of 0.699323, a difference to the baseline R^2 of less than 6.77×10^{-4}. Similarly, considering an F-test for nested models, the F-statistic can be computed from the R^2 values of the two models being compared as $F\text{-statistic} = \frac{(R^2_{\mathrm{CJL}} - R^2_{\mathrm{CJ}})}{(1 - R^2_{\mathrm{CJL}})} \frac{(T-K)}{\#\text{ of restrictions}}$. With an R^2_{CJL} value of 0.70, this yields $(R^2_{\mathrm{CJL}} - R^2_{\mathrm{CJ}}) \times \frac{1}{(1-0.7)} \frac{(4000-10)}{3} \approx (R^2_{\mathrm{CJL}} - R^2_{\mathrm{CJ}}) \times 4433$. Because of the 4433 term, even the smallest increase in model fit of $(R^2_{\mathrm{CJL}} - R^2_{\mathrm{CJ}}) = 0.001$ (0.1 percentage point increase in R^2) yields an F-statistic of 4.4333, which would be deemed significant at the 1% value (a 1% critical value from $F_{(3,3990)} \approx 3.7865$). We can therefore view the increases in R^2 as if they were population increases due to the large sample sizes that are available to us.

Table 1
HAR – RV model parameter estimates over the full sample period.

Equity index	No. stocks	Full sample period	T	$\hat{\beta}_{RV}$				R^2_{RV}
S&P 500	500	03.02.2000–13.07.2016	4107	0.1383	0.3248	0.4323	0.1878	0.6888
United States				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
FTSE 100	100	03.02.2000–13.07.2016	4128	0.0956	0.3204	0.4617	0.1780	0.7609
United Kingdom				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
Nikkei 225	225	04.02.2000–13.07.2016	3982	0.1730	0.3900	0.3408	0.2030	0.6514
Japan				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
DAX 30	30	02.02.2000–13.07.2016	4160	0.1209	0.3508	0.4185	0.1869	0.7441
Germany				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
All Ordinaries	500	04.02.2000–13.07.2016	4107	0.1306	0.1132	0.5262	0.2994	0.5912
Australia				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
CAC 40	40	02.02.2000–13.07.2016	4186	0.1266	0.3503	0.4482	0.1547	0.7384
France				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
Hang Seng	50	02.02.2000–13.07.2016	3783	0.1427	0.1821	0.4599	0.3013	0.6243
Hong Kong				[0.0001]	[0.0000]	[0.0000]	[0.0000]	
KOSPI	760	03.02.2000–13.07.2016	4046	0.0811	0.3627	0.3717	0.2341	0.7925
South Korea				[0.0001]	[0.0000]	[0.0000]	[0.0000]	
AEX	25	02.02.2000–13.07.2016	4185	0.1186	0.3629	0.4518	0.1396	0.7530
The Netherlands				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
Swiss Market Index	20	03.02.2000–13.07.2016	4113	0.1055	0.3639	0.4662	0.1267	0.7721
Switzerland				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
IBEX 35	35	03.02.2000–13.07.2016	4151	0.1131	0.3612	0.4136	0.1842	0.7531
Spain				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
S&P CNX Nifty	50	07.08.2002–13.07.2016	3424	0.1605	0.3284	0.3755	0.2365	0.6451
India				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
IPC Mexico	35	02.02.2000–13.07.2016	4111	0.1645	0.2101	0.3982	0.3234	0.5570
Mexico				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
Bovespa	50	03.02.2000–12.07.2016	4019	0.2964	0.3128	0.4033	0.1847	0.5281
Brazil				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
S&P TSX	250	04.06.2002–13.07.2016	3522	0.1014	0.2909	0.4166	0.2458	0.6968
Canada				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
Euro STOXX 50	50	02.02.2000–13.07.2016	4163	0.1670	0.2967	0.4430	0.1998	0.6539
Euro Area				[0.0000]	[0.0000]	[0.0000]	[0.0000]	
FT Straits Times	30	02.02.2000–18.09.2015	3857	0.0970	0.3029	0.3914	0.2638	0.7247
Singapore				[0.0001]	[0.0000]	[0.0000]	[0.0000]	
FTSE MIB	40	02.02.2000–12.07.2016	4144	0.1204	0.3610	0.4199	0.1738	0.7376
Italy				[0.0000]	[0.0000]	[0.0000]	[0.0000]	

Notes: This table reports OLS regression estimates of the HAR – RV model parameters in (12a) for each foreign equity index. Columns one to four show the equity indices, the number of stocks in each index, the corresponding full sample fitting periods, as well as the sample size T. Columns 5–8 show OLS estimates of $\boldsymbol{\beta}_{RV}$ in (12a), together with (2-sided) p-values computed with Heteroskedasticity and Autocorrelation (HAC) robust standard errors in square brackets below the estimates. The last column shows the in-sample R^2 values, providing information about the fit of the model.

esting to observe here that the estimates of the weekly leverage component are about twice the size of the daily ones. The importance of the monthly leverage component is somewhat lower, with *'only'* 13 of the 18 estimates being significant at the 10% level.[11] Overall, our results for the leverage effect are consistent with previous findings in the literature (see Figlewski and Wang (2000), Bollerslev et al. (2006) and Corsi et al. (2012) among others).

The inclusion of the leverage component to the HAR model also affects the relative importance of the jump component in the model. To more formally asses how much value added remains after adding the jump components to the model once leverage is included, we perform a formal F-test of the null hypothesis $\mathcal{H}_0 : \gamma_{CJL} = \mathbf{0}_3$, and add those results to the last column of Table 3 (asymptotic HAC robust p-values are provided in the parenthesis below the F-test statistic). As we can see from these results, only for 6 equity markets is the null hypothesis of the jump components being irrelevant in-sample strongly rejected. Despite these *'strong'* statistical rejections for some of the equity markets, it is evident from a comparison of the R^2 values, that the improvement can be minimal. For instance, for the FTSE100 and the FT Straits Times (Singapore), the restriction is strongly rejected, but the improvement in R^2 when separating the continuous and jump parts is less than 0.1%. Examining later on the cumulative sum of squared forecast errors in the out-of-sample evaluation, highlights even further the irrelevance of the jump component for volatility in most of the equity markets that we study.

4.2. Out-of-sample forecast evaluation

We now present out-of-sample forecast evaluation results to assess the value of jump and leverage effects on realized volatility in a predictive context. We initially outline the general prediction environment and the evaluation criteria that we utilize, and then proceed to present the out-of-sample forecast evaluation results.

[11] Also, the estimate for the S&P CNX Nifty is positive and significant, thereby counter to economic intuition.

Table 2
HAR – CJ model parameter estimates over the full sample period.

Equity index	Full sample period	T	$\hat{\beta}_{\mathrm{CJ}}$				$\hat{\gamma}_{\mathrm{CJ}}$			R^2_{CJ}
S&P 500	03.02.2000–13.07.2016	4107	0.1916	0.4440	0.3516	0.1106	−0.0267	0.0126	0.0436	0.7141
United States			[0.0000]	[0.0000]	[0.0000]	[0.0001]	[0.0008]	[0.2374]	[0.0051]	(0.0253)
FTSE 100	03.02.2000–13.07.2016	4128	0.1295	0.3247	0.4465	0.1428	0.0275	0.0138	0.0385	0.7618
United Kingdom			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0072]	[0.1598]	[0.0001]	(0.0008)
Nikkei 225	04.02.2000–13.07.2016	3982	0.2522	0.3911	0.3202	0.1703	0.0251	0.0145	0.0284	0.6539
Japan			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0027]	[0.2494]	[0.1113]	(0.0026)
DAX 30	02.02.2000–13.07.2016	4160	0.1767	0.3779	0.3805	0.1507	0.0139	0.0189	0.0307	0.7486
Germany			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0613]	[0.0424]	[0.0114]	(0.0044)
All Ordinaries	04.02.2000–13.07.2016	4107	0.2061	0.1635	0.4868	0.2667	−0.0654	0.0021	0.0200	0.5998
Australia			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.8754]	[0.2295]	(0.0086)
CAC 40	02.02.2000–13.07.2016	4186	0.2038	0.3494	0.4157	0.1558	0.0173	0.0333	−0.0000	0.7387
France			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0655]	[0.0007]	[0.9991]	(0.0004)
Hang Seng	02.02.2000–13.07.2016	3783	0.2017	0.1901	0.4332	0.2758	0.0059	0.0251	0.0257	0.6265
Hong Kong			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.5084]	[0.0222]	[0.0673]	(0.0023)
KOSPI	03.02.2000–13.07.2016	4046	0.2120	0.3566	0.3499	0.2263	0.0307	0.0146	−0.0036	0.7933
South Korea			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0003]	[0.1652]	[0.8082]	(0.0008)
AEX	02.02.2000–13.07.2016	4185	0.1811	0.3635	0.4258	0.1354	0.0282	0.0173	0.0086	0.7533
The Netherlands			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0072]	[0.0831]	[0.4891]	(0.0002)
Swiss Market Index	03.02.2000–13.07.2016	4113	0.1585	0.3690	0.4423	0.1000	0.0319	0.0104	0.0304	0.7739
Switzerland			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0019]	[0.3379]	[0.0379]	(0.0018)
IBEX 35	03.02.2000–13.07.2016	4151	0.1875	0.3762	0.3798	0.1737	0.0095	0.0238	0.0052	0.7550
Spain			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.2636]	[0.0226]	[0.7177]	(0.0020)
S&P CNX Nifty	07.08.2002–13.07.2016	3424	0.1925	0.3929	0.3371	0.1657	0.0109	0.0064	0.0468	0.6574
India			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.1934]	[0.6231]	[0.0032]	(0.0123)
IPC Mexico	02.02.2000–13.07.2016	4111	0.1719	0.2787	0.3983	0.2181	−0.0140	0.0166	0.0479	0.5703
Mexico			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.1309]	[0.1078]	[0.0000]	(0.0133)
Bovespa	03.02.2000–12.07.2016	4019	0.3634	0.3521	0.3557	0.1395	−0.0044	0.0282	0.0315	0.5418
Brazil			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.5137]	[0.0010]	[0.0031]	(0.0137)
S&P TSX	04.06.2002–13.07.2016	3522	0.1193	0.3377	0.3987	0.1688	0.0091	−0.0021	0.0588	0.7063
Canada			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.5116]	[0.8778]	[0.0000]	(0.0095)
Euro STOXX 50	02.02.2000–13.07.2016	4163	0.2634	0.3027	0.4220	0.1593	0.0213	0.0176	0.0259	0.6564
Euro Area			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0099]	[0.1028]	[0.0428]	(0.0026)
FT Straits Times	02.02.2000–18.09.2015	3857	0.1543	0.3018	0.3654	0.2171	0.0435	0.0145	0.0508	0.7248
Singapore			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0001]	[0.1644]	[0.0000]	(0.0001)
FTSE MIB	02.02.2000–12.07.2016	4144	0.1858	0.3699	0.3755	0.1557	0.0220	0.0359	0.0216	0.7400
Italy			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0135]	[0.0002]	[0.0823]	(0.0025)

Notes: This table reports OLS regression estimates of the HAR – CJ model parameters in (12b) for each foreign equity index. Columns one to four show the equity indices, the number of stocks in contains, the corresponding full sample fitting periods, as well as the sample size T. Columns 4–10 show OLS estimates of $\boldsymbol{\beta}_{\mathrm{CJ}}$ and γ_{CJ} in (12b), together with (2-sided) p-values computed with Heteroskedasticity and Autocorrelation (HAC) robust standard errors in square brackets below the estimates. The last column shows the in-sample R^2 values of the fit of the model, with the improvement in R^2 relative to the benchmark HAR – RV model in (12a) ($R^2_{\mathrm{CJ}} - R^2_{\mathrm{RV}}$) below in brackets.

4.2.1. Prediction setting

We implement a standard direct forecasting approach in our out-of-sample evaluation (see Andersen et al. (2007), Corsi and Renó (2012) and many others). To outline how this is implemented, let us formalise the notation and define the (normalised) h-period $\log(\mathrm{RV}_t)$ series as:

$$y_t^{(h)} = \frac{1}{h}\sum_{j=1}^{h} \log(\mathrm{RV}_{t-j+1}), \tag{13}$$

Table 3
HAR – CJL model parameter estimates over the full sample period.

Equity index	Full sample period	T	$\hat{\beta}_{CJL}$				$\hat{\gamma}_{CJL}$			$\hat{\alpha}_{CJL}$			$\underline{R}^2_{CJL}$	F-test
S&P 500	03.02.2000–13.07.2016	4107	0.3389	0.3358	0.3251	0.1681	−0.0298	0.0173	0.0375	−0.0552	−0.1226	−0.0946	0.7303	7.2484
United States			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0001]	[0.0948]	[0.0121]	[0.0000]	[0.0000]	[0.0055]	(0.0415)	(0.0001)
FTSE 100	03.02.2000–13.07.2016	4128	0.2177	0.2118	0.4102	0.2312	0.0198	0.0152	0.0418	−0.0539	−0.1052	−0.0712	0.7750	8.4944
United Kingdom			[0.0000]	[0JIMF.0000]	[0.0000]	[0.0000]	[0.0554]	[0.1034]	[0.0000]	[0.0000]	[0.0000]	[0.0345]	(0.0141)	(0.0000)
Nikkei 225	04.02.2000–13.07.2016	3982	0.3708	0.3011	0.2880	0.2331	0.0160	0.0140	0.0254	−0.0363	−0.0678	−0.0828	0.6706	2.1123
Japan			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0533]	[0.2495]	[0.1424]	[0.0000]	[0.0000]	[0.0006]	(0.0193)	(0.0965)
DAX 30	02.02.2000–13.07.2016	4160	0.3145	0.2676	0.3359	0.2394	0.0075	0.0154	0.0310	−0.0410	−0.0732	−0.1004	0.7626	3.0193
Germany			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.3039]	[0.0936]	[0.0098]	[0.0000]	[0.0000]	[0.0000]	(0.0184)	(0.0286)
All Ordinaries	04.02.2000–13.07.2016	4107	0.3611	0.0703	0.3975	0.3590	−0.0705	−0.0059	0.0138	−0.0739	−0.1460	−0.1950	0.6243	3.4019
Australia			[0.0000]	[0.0007]	[0.0000]	[0.0000]	[0.0000]	[0.6512]	[0.4054]	[0.0000]	[0.0000]	[0.0002]	(0.0331)	(0.0170)
CAC 40	02.02.2000–13.07.2016	4186	0.3461	0.2284	0.3845	0.2428	0.0112	0.0301	−0.0056	−0.0466	−0.0874	−0.0844	0.7559	2.9816
France			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.2233]	[0.0008]	[0.6720]	[0.0000]	[0.0000]	[0.0008]	(0.0175)	(0.0301)
Hang Seng	02.02.2000–13.07.2016	3783	0.2906	0.1473	0.3755	0.3322	0.0007	0.0163	0.0283	−0.0174	−0.0595	−0.0488	0.6367	1.3735
Hong Kong			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.9328]	[0.1424]	[0.0532]	[0.0010]	[0.0000]	[0.1371]	(0.0124)	(0.2489)
KOSPI	03.02.2000–13.07.2016	4046	0.2784	0.2745	0.3430	0.2731	0.0202	0.0128	0.0008	−0.0430	−0.0555	−0.0042	0.8049	1.2409
South Korea			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0138]	[0.2177]	[0.9579]	[0.0000]	[0.0000]	[0.8531]	(0.0123)	(0.2932)
AEX	02.02.2000–13.07.2016	4185	0.3028	0.2429	0.4117	0.2025	0.0131	0.0119	0.0139	−0.0542	−0.0819	−0.0401	0.7704	0.9164
The Netherlands			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.1858]	[0.2184]	[0.2488]	[0.0000]	[0.0000]	[0.0559]	(0.0173)	(0.4320)
Swiss Market Index	03.02.2000–13.07.2016	4113	0.2934	0.2306	0.4071	0.1987	0.0174	0.0069	0.0326	−0.0505	−0.0928	−0.1266	0.7913	1.5773
Switzerland			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0556]	[0.4980]	[0.0182]	[0.0000]	[0.0000]	[0.0001]	(0.0192)	(0.1927)
IBEX 35	03.02.2000–13.07.2016	4151	0.2873	0.2450	0.3845	0.2475	0.0075	0.0189	0.0051	−0.0512	−0.0806	−0.0328	0.7723	1.1521
Spain			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.3546]	[0.0642]	[0.7150]	[0.0000]	[0.0000]	[0.2357]	(0.0192)	(0.3266)
S&P CNX Nifty	07.08.2002–13.07.2016	3424	0.3089	0.2543	0.3342	0.2495	0.0037	0.0129	0.0407	−0.0681	−0.0897	0.0580	0.6813	1.5746
India			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.6536]	[0.2556]	[0.0122]	[0.0000]	[0.0000]	[0.0520]	(0.0362)	(0.1934)
IPC Mexico	02.02.2000–13.07.2016	4111	0.2820	0.2037	0.3451	0.2608	−0.0195	0.0153	0.0613	−0.0455	−0.1046	−0.0873	0.5894	6.4667
Mexico			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0268]	[0.1221]	[0.0000]	[0.0000]	[0.0000]	[0.0207]	(0.0323)	(0.0002)
Bovespa	03.02.2000–12.07.2016	4019	0.4990	0.2619	0.3352	0.1817	−0.0105	0.0299	0.0362	−0.0360	−0.0644	−0.0390	0.5671	7.7047
Brazil			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.1025]	[0.0004]	[0.0009]	[0.0000]	[0.0000]	[0.0446]	(0.0390)	(0.0000)
S&P TSX	04.06.2002–13.07.2016	3522	0.2135	0.2307	0.3662	0.2321	−0.0005	−0.0024	0.0665	−0.0678	−0.1263	−0.0526	0.7238	6.3384
Canada			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.9677]	[0.8461]	[0.0000]	[0.0000]	[0.0000]	[0.2170]	(0.0270)	(0.0003)
Euro STOXX 50	02.02.2000–13.07.2016	4163	0.4707	0.1852	0.3439	0.2660	0.0091	0.0159	0.0174	−0.0486	−0.1060	−0.1775	0.6813	1.6697

(*continued on next page*)

Table 3 (*continued*)

Equity index	Full sample period	T		$\hat{\boldsymbol{\beta}}_{\text{CJL}}$				$\hat{\boldsymbol{\gamma}}_{\text{CJL}}$			$\hat{\boldsymbol{\alpha}}_{\text{CJL}}$		R^2_{CJL}	F-test
Euro Area			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.2593]	[0.1051]	[0.1570]	[0.0000]	[0.0000]	[0.0000]	(0.0274)	(0.1713)
FT Straits Times	02.02.2000–18.09.2015	3857	0.2403	0.2437	0.2954	0.2878	0.0340	0.0045	0.0618	−0.0312	−0.0537	−0.1073	0.7380	12.7760
Singapore			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0014]	[0.6649]	[0.0000]	[0.0000]	[0.0009]	[0.0000]	(0.0133)	(0.0000)
FTSE MIB	02.02.2000–12.07.2016	4144	0.3075	0.2372	0.3569	0.2527	0.0137	0.0307	0.0130	−0.0507	−0.0919	−0.0344	0.7599	2.9897
Italy			[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.1120]	[0.0012]	[0.3011]	[0.0000]	[0.0000]	[0.1412]	(0.0224)	(0.0298)

Notes: This table reports OLS regression estimates of the HAR − CJL model parameters in (12c) for each foreign equity index. Columns one to four show the equity indices, the number of stocks in contains, the corresponding full sample fitting periods, as well as the sample size T. Columns 4–13 show OLS estimates of $\boldsymbol{\beta}_{\text{CJL}}, \boldsymbol{\gamma}_{\text{CJL}}$ and $\boldsymbol{\alpha}_{\text{CJL}}$ in (12c), together with (2-sided) p-values computed with Heteroskedasticity and Autocorrelation (HAC) robust standard errors in square brackets below the estimates. The second last column shows the in-sample R^2 values of the fit of the model, with the improvement in R^2 relative to the benchmark HAR − RV model in (12a) ($R^2_{\text{CJL}} - R^2_{\text{RV}}$) below in brackets. The last column shows F-test statistics (with asymptotic p-values based on HAC robust standard errors in parenthesis below) of the null hypothesis that the HAR jump components are not significantly different from zero.

and re-formulate the predictive relations in (12) for the general h-steps ahead long-horizon regression setting as:

$$y_t^{(h)} = \boldsymbol{x}_{t-h}^{\mathrm{RV}}\boldsymbol{\beta}_{\mathrm{RV}}^{(h)} + \epsilon_{t,\mathrm{RV}}^{(h)} \tag{14a}$$
$$y_t^{(h)} = \boldsymbol{x}_{t-h}^{C}\boldsymbol{\beta}_{\mathrm{CJ}}^{(h)} + \boldsymbol{x}_{t-h}^{J}\boldsymbol{\gamma}_{\mathrm{CJ}}^{(h)} + \epsilon_{t,\mathrm{CJ}}^{(h)} \tag{14b}$$
$$y_t^{(h)} = \boldsymbol{x}_{t-h}^{C}\boldsymbol{\beta}_{\mathrm{CJL}}^{(h)} + \boldsymbol{x}_{t-h}^{J}\boldsymbol{\gamma}_{\mathrm{CJL}}^{(h)} + \boldsymbol{x}_{t-h}^{L}\boldsymbol{\alpha}_{\mathrm{CJL}}^{(h)} + \epsilon_{t,\mathrm{CJL}}^{(h)}, \tag{14c}$$

where direct h-step ahead forecasts are computed as:

$$\hat{y}_{t+h|t}^{\mathrm{RV}} = \boldsymbol{x}_t^{\mathrm{RV}}\hat{\boldsymbol{\beta}}_{\mathrm{RV}}^{(h)} \tag{15a}$$
$$\hat{y}_{t+h|t}^{\mathrm{CJ}} = \boldsymbol{x}_t^{C}\hat{\boldsymbol{\beta}}_{\mathrm{CJ}}^{(h)} + \boldsymbol{x}_t^{J}\hat{\boldsymbol{\gamma}}_{\mathrm{CJ}}^{(h)} \tag{15b}$$
$$\hat{y}_{t+h|t}^{\mathrm{CJL}} = \boldsymbol{x}_t^{C}\hat{\boldsymbol{\beta}}_{\mathrm{CJL}}^{(h)} + \boldsymbol{x}_t^{J}\hat{\boldsymbol{\gamma}}_{\mathrm{CJL}}^{(h)} + \boldsymbol{x}_t^{L}\hat{\boldsymbol{\alpha}}_{\mathrm{CJL}}^{(h)}. \tag{15c}$$

The h superscripts on the $\boldsymbol{\beta}, \boldsymbol{\gamma}$, and $\boldsymbol{\alpha}$ parameters (as well as their estimates) indicate that these are from the h-periods off-set (or long-horizon predictive) regressions in (14). The forecast errors corresponding to the predictions in (15) and the resulting Mean Squared Forecast Errors (MSFEs) are given by:

$$\hat{e}_{t+h|t}^{\mathcal{M}^*} = y_{t+h}^{(h)} - \hat{y}_{t+h|t}^{\mathcal{M}^*}, \quad \text{and} \quad \mathrm{MSFE}_{\mathcal{M}^*} = \frac{1}{T_{os}}\sum_{t=T_{is}}^{T-h}\left(\hat{e}_{t+h|t}^{\mathcal{M}^*}\right)^2, \tag{16}$$

respectively, where $\mathcal{M}^* \in \{\mathrm{RV}, \mathrm{CJ}, \mathrm{CJL}\}$ denotes the set of HAR models that we consider. The terms T_{os} and T_{is} are the number of out-of-sample and in-sample observations, where $T_{os} = T - T_{is} - h + 1$, and T is the full sample size.

We implement a rolling window forecasting scheme, where 750 observations (approximately three years of data) are used as the in-sample fitting period, i.e., $T_{is} = 750$. We view 750 observations as large enough to obtain reasonably precise parameter estimates needed to compute reliable forecasts for the out-of-sample evaluation, while at the same time allowing for enough flexibility to accommodate possible structural changes that may have occurred. Also, using 750 in-sample observations leaves well over 3000 data points for out-of-sample analysis for all but two of the 18 equity markets that we analyze.[12] The out-of-sample forecast evaluation results that we present are thus not sensitive to the choice of the estimation window.

4.2.2. Evaluation criteria

We use three widely employed evaluation criteria to assess the out-of-sample forecast performance of the HAR-CJ and HAR-CJL model relative to the benchmark HAR-RV model. These are:

(i) the Diebold and Mariano (1995) test,
(ii) the out-of-sample R^2 of Campbell and Thompson (2008), and
(iii) plots of the cumulative sum of squared forecast error differences between the models.[13]

The Diebold and Mariano (1995) test (DM-test henceforth) performs a test of unconditional superior predictive ability.[14] Due to the rolling window forecast implementation that we employ, the DM-test is equivalent to the Giacomini and White (2006) testing framework for conditional (and unconditional) predictive ability, which is valid for nested as well as non-nested model comparisons, and also accounts for parameter estimation error. Our implementation of the DM-test is an unconditional one, nevertheless is valid under the asymptotic theory of Giacomini and White (2006) (see the discussion on pages 1549–1551 in Giacomini and White (2006) for additional details).

The simplest way to compute the DM-test is to form the sequence:

$$\mathrm{DM}_{t+h}^{\mathcal{M}} = \left(\hat{e}_{t+h|t}^{\mathrm{RV}}\right)^2 - \left(\hat{e}_{t+h|t}^{\mathcal{M}}\right)^2,$$

where, as before, $\mathcal{M} \in \{\mathrm{CJ}, \mathrm{CJL}\}$ denotes the model that is compared to the benchmark HAR-RV model, and $\hat{e}_{t+h|t}$ are h-step ahead forecast errors. The DM-statistic is then computed as:

$$\mathrm{DM}^{\mathcal{M}}\text{-statistic} = \frac{\overline{\mathrm{DM}}^{\mathcal{M}}}{\sqrt{\mathbb{V}\mathrm{ar}(\overline{\mathrm{DM}}^{\mathcal{M}})}}, \tag{17}$$

[12] For the Canadian S&P TSX and the Indian S&P CNX Nifty, we have at least 2600 out-of-sample data points. Reliable RV data is available only from mid 2002 for these two markets.

[13] These three evaluation criteria are widely used in the RV forecasting literature, see for instance, Corsi and Renó (2012), Patton and Sheppard (2015), and Prokopczuk et al. (2016).

[14] Note here that we are performing a simple pairwise forecast comparison between the competing HAR-CJ or HAR-CJL model and the benchmark HAR-RV model for each equity market and are not comparing forecasts from many different models for a single series. A Diebold and Mariano (1995) type test of unconditional predictive ability is thus sufficient for our purpose of assessing the contribution of the jump and leverage effect to each equity market's volatility forecasts.

where $\overline{\mathrm{DM}}^{\mathcal{M}} = T_{os}^{-1}\sum_{t=T_{is}}^{T-h}\mathrm{DM}_{t+h}^{\mathcal{M}}$ and $\mathbb{V}\mathrm{ar}(\overline{\mathrm{DM}}^{\mathcal{M}})$ is the variance of the sample mean, which can simply be obtained as the HAC robust t-statistic on the intercept term from a regression of DM_{t+h} on a constant. The DM-test implements an unconditional superior predictive ability test by comparing the MSFE of the benchmark HAR-RV model to those from the HAR-CJ or the HAR-CJL models under the null-hypothesis. A rejection of the null hypothesis thus means that forecast errors from the HAR-CJ or the HAR-CJL models are on average significantly smaller than from the benchmark HAR-RV model, suggesting forecast superiority relative to the HAR-RV.

The Campbell and Thompson (2008) out-of-sample R^2 (denoted by R^2_{os} henceforth) is computed as 1 minus the ratio of MSFEs from the model under consideration and the benchmark. That is, the R^2_{os} for model $\mathcal{M}$ is defined as:

$$ {}_{\mathcal{M}}R^2_{os} = 1 - \frac{\mathrm{MSFE}_{\mathcal{M}}}{\mathrm{MSFE}_{\mathrm{RV}}}. \tag{18}$$

Intuitively, ${}_{\mathcal{M}}R^2_{os}$ provides an indication of the reduction in MSFE of the proposed model relative to the benchmark HAR-RV model. When ${}_{\mathcal{M}}R^2_{os} > 0$ (respectively, ${}_{\mathcal{M}}R^2_{os} < 0$), then the proposed model $\mathcal{M}$ performs better (worse) than the benchmark model. The R^2_{os} is frequently used in the equity premium forecasting literature to gauge the economic magnitude (or significance) of the statistical forecast gains that one obtains. In this context, Campbell and Thompson (2008) provide a mapping from the R^2_{os} to the gain in utility that a risk averse investor would obtain (see pages 1524–1526 in Campbell and Thompson (2008) for the exact relations). Since no such relation exists in the volatility forecasting literature, we will judge the size of the R^2_{os} relative to other values found in the volatility forecasting literature, in particularly, the recent study by Patton and Sheppard (2015). They report increases in R^2_{os} in the range of 1.1% and 3% for various semi-variance based model specifications that separate the effect of good and bad volatility (as well as signed jumps) on forecasts of volatility (see their Table 6 and the discussion on page 695).[15] In the discussion that follows, we will consider increases in R^2_{os} relative to the benchmark HAR-RV model of above 2–3% as large, and increases closer to 1% and below as small.

In addition to the DM-statistic and R^2_{os}, we also compute and plot the cumulative difference of the squared forecast errors from the proposed model $\mathcal{M}$ relative to the HAR-RV benchmark over the out-of-sample period. This cumulative difference (denoted by cumSFE) is informative, because it provides a visual overview of the predictive performance of the proposed model over time. We define the cumSFE as:

$$ \mathrm{cumSFE}_{t+h}^{\mathcal{M}} = \sum_{\tau=T_{is}}^{t}\left[\left(\hat{e}^{\mathrm{RV}}_{\tau+h|\tau}\right)^2 - \left(\hat{e}^{\mathcal{M}}_{\tau+h|\tau}\right)^2\right], \quad \forall t = T_{is},\ldots,T-h. \tag{19}$$

Whenever the cumSFE series is above zero, the cumulative sum of the squared forecast errors of the benchmark HAR-RV model is larger than from model $\mathcal{M}$, which indicates that the benchmark's forecasts are less accurate. Moreover, an upward sloping cumSFE sequence means that the proposed model produces consistently better out-of-sample predictions than the benchmark HAR-RV model.

4.2.3. Forecast evaluation results

4.2.3.1. One-step-ahead. One-step-ahead out-of-sample forecast evaluation results for the HAR-CJ and HAR-CJL models for all 18 international equity markets that we consider are presented in the top and bottom panels of Table 4. The first four columns in Table 4 show the equity index of interest, the corresponding country, the actual out-of-sample evaluation period and the effective number of out-of-sample observations T_{os} that are available. In columns five to seven, the $\mathrm{MSFE}_{\mathcal{M}}$, the ratio of MSFEs denoted by $\mathrm{MSFE}_{\mathcal{M}}/\mathrm{MSFE}_{\mathrm{RV}}$, and the ${}_{\mathcal{M}}R^2_{os}$ are shown. In the last two columns, we report the $\mathrm{DM}^{\mathcal{M}}$-statistic and its corresponding one-sided p-value.

The one-step-ahead evaluation results reported in the top panel of Table 4 show a number of interesting results. First, as also found in Andersen et al. (2007), Corsi and Renó (2012), Patton and Sheppard (2015) and others, the decomposition of RV into a continuous and discontinuous jump component is important for the S&P 500. This is evident not only form the highly significant DM-statistic with a value greater than 7, but also from the out-of-sample R^2 of over 8%. Despite this strong result for the S&P 500, for the 17 remaining international equity markets that we study, the results are, nonetheless, rather mixed. In fact, only for 6 of the 17 equity indices are the improvements in out-of-sample forecast performance highly significant, with DM-statistics in excess of 3, and as high as 4.6. These equity indices are the DAX 30, the All Ordinaries, the S&P CNX Nifty, the IPC Mexico, the Bovespa, and the S&P TSX. For the DAX 30 and the All Ordinaries, the overall magnitude of the improvement, as measured by R^2_{os}, is somewhat smaller, with gains of only 1.65% and 1.92%. For the S&P CNX Nifty, the IPC Mexico, the Bovespa, and the S&P TSX, the improvement in R^2_{os} are more substantial between 2.72% and 4.19%. Although the forecast gains for the Euro STOXX 50 and the FTSE MIB are marginally significant at the 10% level, they yield R^2_{os} values of only 0.4% and 0.87%, respectively. For the remaining 9 equity markets, forecast gains are not only statistically insignificant, but R^2_{os} values are rather small and can even be negative (see the FTSE 100,CAC 40, KOSPI and AEX), suggesting a worsening in

[15] Note that their R^2_{os} values in Table 6 are computed relative to a model with (only) a constant term included. To get a comparable measure, one needs to use the R^2_{os} from the HAR model given in column 1 of Table 6 in their paper.

out-of-sample forecast performance relative to the benchmark HAR-RV model. Cross-sectional evidence of out-of-sample forecasts gains from separating the continuous and discontinuous parts of RV are thus at best mixed, providing significant and sizeable improvements for only 6 of the 17 international equity markets that we analyze.

Looking over the forecast evaluation results for the HAR-CJL model reported in the bottom panel of Table 4, it is clear that the addition of a HAR structured leverage component to the predictor set leads uniformly to much greater out-of-sample forecast improvements across all 18 international equity markets. The *'weakest'* improvements are realized for the FTSE 100 and the Hang Seng, with R^2_{os} values of *'only'* 2.71% and 2.87%, with corresponding DM-statistics of 1.64 and 1.76. Note that the HAR-CJ model's R^2_{os} values for these two equity markets were -0.71% and 0.06%. Relative to the HAR-CJ model, the improvements in R^2_{os} for these two equity markets are 3.41% and 2.81%, respectively, which are rather sizeable. Moreover, for 14 of the 18 equity markets, the DM-statistics are in excess of 2.85 (p-value < 0.0025), with R^2_{os} values of at least 4.39%. For 10 of the equity markets, the R^2_{os} values are above 6%, suggesting very strong overall forecast improvements relative to the HAR-RV benchmark.

To provide some visual evidence of the gains that can be obtained from the separation of RV into a continuous and a discontinuous component, as well as the addition of the leverage effect, we show the time series evolution of the cumulative sum of squared forecast errors at the one-step-ahead forecast horizon over the out-of-sample period in Fig. 1.

The cumSFE series for the HAR-CJ model is shown by the thick (orange) line, while for the HAR-CJL model it is drawn by the thin (blue) line in Fig. 1. Both cumSFE series are constructed relative to the benchmark HAR-RV forecast errors. As an initial observation, we can notice from the plots that the HAR-CJL model consistently produces substantially more accurate forecasts over the entire out-of-sample period than the HAR-CJ model, which can be seen from the thin (blue) line being uniformly above the thick (orange) line in Fig. 1. Moreover, the HAR-CJL cumSFE series appears to be largely monotonically increasing over the entire out-of-sample period, while for the HAR-CJ model there are instances where the cumSFE series is hump shaped (for the Euro STOXX 50 for instance) or even downward sloping (for the FTSE 100, CAC 40, KOSPI and AEX).

A second observation that is striking from the cumSFE plots in Fig. 1 is the consistently weak (or non-existing) forecast gain from the HAR-CJ model for the FTSE 100, Nikkei 225, CAC 40, Hang Seng, KOSPI, AEX, Swiss Market Index, IBEX 35, FT Straits Times and FTSE MIB equity markets, over the *entire* out-of-sample period. This can be seen from the rather flat and close to zero or negative cumSFE sequence (thick orange line), which is consistent with the statistical results found from the upper panel of Table 4. For these ten equity markets, the smallest R^2_{os} values are obtained, with results from the corresponding DM-test suggesting that forecasts from the HAR-CJ model are not statistically different from the benchmark HAR-RV model. For the IPC Mexico, Bovespa and S&P TSX, the cumSFE series remain flat around 0 for approximately half of the out-of-sample period, showing a decisively positive slope only from around the end of 2009 onwards. Interestingly, this weaker performance over the first half of the out-of-sample period is not evident from the R^2_{os} values and DM-statistics reported in Table 4, which highlights the benefits of computing and plotting the cumSFE sequence over the out-of-sample period.

4.2.3.2. Multi-step-ahead. Multiple-steps ahead out-of-sample forecast evaluation results for the HAR-CJ and the HAR-CJL models are presented in Tables 5 and 6, respectively.

The evaluation results reported in Tables 5 and 6 are split into three parts, with each of the top, middle and lower parts corresponding to one of the three forecast horizons that we consider, i.e., $5, 10$, and 22 steps-ahead. The column entries contain the same information as the one-step-ahead evaluation results reported in Table 4. Since h-step-ahead forecast errors will be $\mathrm{MA}(h-1)$ processes in general, the DM_{t+h} sequence itself will be autocorrelated for $h > 1$. To construct the HAC standard errors needed to conduct inference on the average of the DM_{t+h} sequence, we employ a *'pre-whitening'* step, using an $\mathrm{ARMA}(1,1)$ as the approximating model for the DM_{t+h} sequence to reduce the initial autocorrelation in the series, and then apply a Quadratic Spectral (QS) kernel based non-parametric HAC estimator on the residuals from the $\mathrm{ARMA}(1,1)$ model. Following Andrews and Monahan (1992), we choose the bandwidth parameter for the QS kernel optimally with an AR(1) as the approximating model, and then *'re-colour'* to obtain the required HAC standard errors for the average of the (out-of-sample) DM_{t+h} sequence.

Looking over the HAR-CJ multiple-steps-ahead forecast evaluation results reported in Table 5, we can see that the forecast improvements relative to the benchmark HAR-RV model deteriorate quickly and substantially as the forecast horizon increases. At the 5 day horizon, forecast gains remain highly significant only for the S&P 500, yielding a DM-statistic of close to 4 and an R^2_{os} value of around 5.7%. For the Euro STOXX 50 and the IPC Mexico, the results are marginally significant at the 10% and 5% levels, respectively, with, nonetheless, rather *'small'* R^2_{os} values of 1.50% and 2.06%. For all other 15 equity markets, the realized forecast gains are statistically insignificant, with 8 of the 15 equity markets in fact yielding negative R^2_{os} values, thereby suggesting a worsening of the forecast performance relative to the benchmark HAR-RV model. At the 10 day horizon, only for the S&P 500 and the Euro STOXX 50 are the forecast gains significant at the 5% and 10% levels, and somewhat sizeable, with R^2_{os} values of around 3.6% and 2.1%. Forecasts for all other equity markets are no better than from the benchmark HAR-RV model. At the 22 day horizon, no statistically significant or sizeable forecast gains from the HAR-CJ model can be realized.

Table 4
One-step-ahead out-of-sample forecast evaluation results (rolling window, $T_{is} = 750$).

Equity index	Country	Out-of-sample period	T_{os}	$\text{MSFE}_{\mathcal{M}}$	$\frac{\text{MSFE}_{\mathcal{M}}}{\text{MSFE}_{\text{RV}}}$	${}_{\mathcal{M}}R^2_{os}$	$DM^{\mathcal{M}}$-stat	p-value
			HAR-CJ model					
S&P 500	United States	26.02.2003–13.07.2016	3358	0.0858	0.9193	0.0807	7.5875	0.0000
FTSE 100	United Kingdom	07.02.2003–13.07.2016	3379	0.0598	1.0071	−0.0071	−1.5459	0.9389
Nikkei 225	Japan	14.03.2003–13.07.2016	3233	0.0703	0.9962	0.0038	0.6637	0.2534
DAX 30	Germany	30.01.2003–13.07.2016	3411	0.0676	0.9835	0.0165	3.2065	0.0007
All Ordinaries	Australia	04.02.2003–13.07.2016	3358	0.0888	0.9808	0.0192	3.0541	0.0011
CAC 40	France	03.02.2003–13.07.2016	3437	0.0632	1.0061	−0.0061	−1.8030	0.9643
Hang Seng	Hong Kong	10.03.2003–13.07.2016	3034	0.0638	0.9994	0.0006	0.1703	0.4324
KOSPI	South Korea	14.03.2003–13.07.2016	3297	0.0540	1.0011	−0.0011	−0.2358	0.5932
AEX	The Netherlands	04.02.2003–13.07.2016	3436	0.0642	1.0067	−0.0067	−2.3241	0.9899
Swiss Market Index	Switzerland	12.02.2003–13.07.2016	3364	0.0469	0.9995	0.0005	0.1694	0.4327
IBEX 35	Spain	21.02.2003–13.07.2016	3402	0.0602	0.9956	0.0044	1.0429	0.1485
S&P CNX Nifty	India	18.08.2005–13.07.2016	2675	0.0797	0.9581	0.0419	4.6028	0.0000
IPC Mexico	Mexico	26.02.2003–13.07.2016	3362	0.0972	0.9687	0.0313	4.2441	0.0000
Bovespa	Brazil	20.03.2003–12.07.2016	3270	0.0638	0.9728	0.0272	3.6086	0.0002
S&P TSX	Canada	06.06.2005–13.07.2016	2773	0.0865	0.9641	0.0359	4.2528	0.0000
Euro STOXX 50	Euro Area	29.01.2003–13.07.2016	3414	0.0907	0.9913	0.0087	1.3781	0.0841
FT Straits Times	Singapore	17.02.2003–18.09.2015	3108	0.0367	0.9991	0.0009	0.1930	0.4235
FTSE MIB	Italy	10.02.2003–12.07.2016	3395	0.0652	0.9960	0.0040	1.3734	0.0848
			HAR-CJL model					
S&P 500	United States	26.02.2003–13.07.2016	3358	0.0820	0.8782	0.1218	8.3413	0.0000
FTSE 100	United Kingdom	07.02.2003–13.07.2016	3379	0.0577	0.9729	0.0271	1.6363	0.0509
Nikkei 225	Japan	14.03.2003–13.07.2016	3233	0.0675	0.9561	0.0439	3.8680	0.0001
DAX 30	Germany	30.01.2003–13.07.2016	3411	0.0650	0.9464	0.0536	3.6009	0.0002
All Ordinaries	Australia	04.02.2003–13.07.2016	3358	0.0837	0.9243	0.0757	5.9432	0.0000
CAC 40	France	03.02.2003–13.07.2016	3437	0.0603	0.9607	0.0393	2.1680	0.0151
Hang Seng	Hong Kong	10.03.2003–13.07.2016	3034	0.0620	0.9713	0.0287	1.7617	0.0391
KOSPI	South Korea	14.03.2003–13.07.2016	3297	0.0507	0.9399	0.0601	5.6921	0.0000
AEX	The Netherlands	04.02.2003–13.07.2016	3436	0.0613	0.9608	0.0392	2.0592	0.0197
Swiss Market Index	Switzerland	12.02.2003–13.07.2016	3364	0.0444	0.9451	0.0549	3.4461	0.0003
IBEX 35	Spain	21.02.2003–13.07.2016	3402	0.0567	0.9380	0.0620	3.9862	0.0000
S&P CNX Nifty	India	18.08.2005–13.07.2016	2675	0.0745	0.8951	0.1049	6.5456	0.0000
IPC Mexico	Mexico	26.02.2003–13.07.2016	3362	0.0926	0.9225	0.0775	6.6448	0.0000
Bovespa	Brazil	20.03.2003–12.07.2016	3270	0.0610	0.9297	0.0703	5.8701	0.0000
S&P TSX	Canada	06.06.2005–13.07.2016	2773	0.0811	0.9045	0.0955	6.0355	0.0000
Euro STOXX 50	Euro Area	29.01.2003–13.07.2016	3414	0.0855	0.9348	0.0652	4.0695	0.0000
FT Straits Times	Singapore	17.02.2003–18.09.2015	3108	0.0351	0.9545	0.0455	2.8570	0.0021
FTSE MIB	Italy	10.02.2003–12.07.2016	3395	0.0611	0.9340	0.0660	5.0756	0.0000

Notes: This table reports one-step-ahead out-of-sample forecast evaluation results for the 18 international equity markets that we consider. The target variable is log(RV). The top and bottom panels show the results for the HAR-CJ and HAR-CJL model, respectively. We use a rolling window forecasting scheme with an in-sample window size of 750 observations. Columns one to four show the equity index of interest, the corresponding country, the actual out-of-sample evaluation period and the effective number of out-of-sample observations T_{os} that are available. In columns five to seven, the $\text{MSFE}_{\mathcal{M}}$, the ratio of MSFEs denoted by $\text{MSFE}_{\mathcal{M}}/\text{MSFE}_{\text{RV}}$, and the ${}_{\mathcal{M}}R^2_{os}$ are shown. In the last two columns, we report the $\text{DM}^{\mathcal{M}}$-statistic and its corresponding one-sided p-value.

For the HAR-CJL model, the forecast evaluation results reported in Table 6 show a rather different picture. We can see that, with the exception of the FTSE 100 and the AEX, forecast improvements remain significant at the 10% level for all other 16 equity markets at the 5 day ahead forecast horizon, with 9 of these being highly significant (p-values $\leqslant 0.01$). Out-of-sample R^2 values remain sizeable, with 12 of the best performing forecasts producing R^2_{os} values of at least 4%. In addition to the S&P 500, 5 day ahead realized volatility forecasts for the Euro STOXX 50 generate also a large R^2_{os} value of close to 10%. Comparing this to the R^2_{os} value of 1.50% from the HAR-CJ model emphasizes the strong gain that the inclusion of the leverage effect has on improving forecasts of RV for the Euro STOXX 50.

At the 10 and 22 day ahead forecast horizons, gains relative to the benchmark HAR-RV deteriorate in general. Nevertheless, for 8 of the 18 equity markets, improvements remain significant at the 10% level and are moreover sizeable, generating R^2_{os} values of at least 2.73% at the 10 day horizon. For the S&P 500 and the Euro STOXX 50, R^2_{os} values are as high as 7.49% and 9.44%, respectively, with corresponding DM-statistics of 3.6 and 3.1. At the 22 day ahead horizon, with the exception of the Euro STOXX 50 and the IPC Mexico, all forecast gains are statistically insignificant, with mostly negative R^2_{os} values. For the Euro STOXX 50 and the IPC Mexico, forecast gains remain not only significant at the 1% and 5% levels, but also sizeable, producing R^2_{os} values of over 7% and 4%. This last finding is particularly interesting and does not appear to have been documented

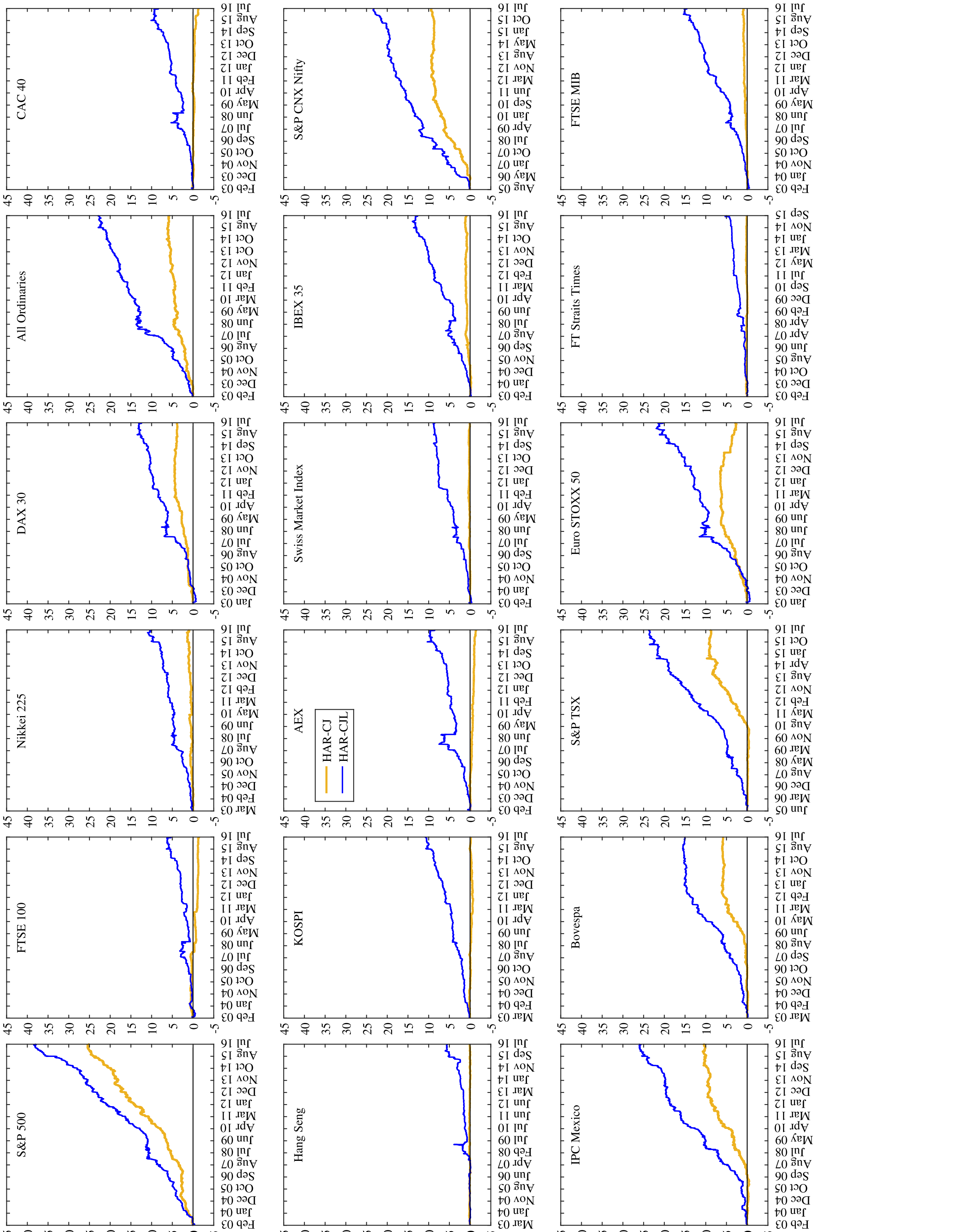

Fig. 1. Time series evolution of the cumulative difference of the squared one-step-ahead forecast errors of the HAR-CJ and HAR-CJL models relative to the benchmark HAR-RV (cumSFE). The thick (orange) line shows the cumSFE sequence for the HAR-CJL model, while the thin (blue) line shows the results for the HAR-CJ model. These are based on a rolling window forecasting scheme with a window size of 750 observations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Multiple-step-ahead out-of-sample forecast evaluation results: HAR-CJ model (rolling window, $T_{is} = 750$).

Equity index	Country	Out-of-sample period	T_{os}	$\mathrm{MSFE}_{\mathcal{M}}$	$\frac{\mathrm{MSFE}_{\mathcal{M}}}{\mathrm{MSFE}_{\mathrm{RV}}}$	${}_{\mathcal{M}}R^2_{os}$	$\mathrm{DM}_{\mathcal{M}}$-stat	p-value
		Forecast horizon h = 5						
S&P 500	United States	04.03.2003–13.07.2016	3354	0.0578	0.9431	0.0569	3.9222	0.0000
FTSE 100	United Kingdom	13.02.2003–13.07.2016	3375	0.0396	1.0174	−0.0174	−1.6342	0.9489
Nikkei 225	Japan	20.03.2003–13.07.2016	3229	0.0489	1.0149	−0.0149	−1.0256	0.8475
DAX 30	Germany	05.02.2003–13.07.2016	3407	0.0445	0.9977	0.0023	0.2726	0.3926
All Ordinaries	Australia	10.02.2003–13.07.2016	3354	0.0399	0.9914	0.0086	0.8862	0.1877
CAC 40	France	07.02.2003–13.07.2016	3433	0.0421	1.0154	−0.0154	−2.3905	0.9916
Hang Seng	Hong Kong	14.03.2003–13.07.2016	3030	0.0329	1.0054	−0.0054	−0.6170	0.7314
KOSPI	South Korea	20.03.2003–13.07.2016	3293	0.0348	0.9999	0.0001	0.0152	0.4939
AEX	The Netherlands	10.02.2003–13.07.2016	3432	0.0448	1.0165	−0.0165	−2.9671	0.9985
Swiss Market Index	Switzerland	18.02.2003–13.07.2016	3360	0.0331	1.0132	−0.0132	−2.3577	0.9908
IBEX 35	Spain	27.02.2003–13.07.2016	3398	0.0404	1.0037	−0.0037	−0.4336	0.6677
S&P CNX Nifty	India	24.08.2005–13.07.2016	2671	0.0481	0.9830	0.0170	1.1528	0.1245
IPC Mexico	Mexico	04.03.2003–13.07.2016	3358	0.0470	0.9794	0.0206	1.6603	0.0484
Bovespa	Brazil	26.03.2003–12.07.2016	3266	0.0387	0.9880	0.0120	1.0207	0.1537
S&P TSX	Canada	10.06.2005–13.07.2016	2769	0.0514	0.9839	0.0161	0.9283	0.1766
Euro STOXX 50	Euro Area	04.02.2003–13.07.2016	3410	0.0535	0.9850	0.0150	1.3845	0.0831
FT Straits Times	Singapore	21.02.2003–18.09.2015	3104	0.0211	0.9925	0.0075	0.7603	0.2235
FTSE MIB	Italy	14.02.2003–12.07.2016	3391	0.0430	1.0021	−0.0021	−0.2869	0.6129
		Forecast horizon h = 10						
S&P 500	United States	11.03.2003–13.07.2016	3349	0.0602	0.9641	0.0359	2.2144	0.0134
FTSE 100	United Kingdom	20.02.2003–13.07.2016	3370	0.0409	1.0132	−0.0132	−1.0432	0.8516
Nikkei 225	Japan	28.03.2003–13.07.2016	3224	0.0491	1.0320	−0.0320	−1.5145	0.9351
DAX 30	Germany	12.02.2003–13.07.2016	3402	0.0457	1.0095	−0.0095	−0.8098	0.7910
All Ordinaries	Australia	17.02.2003–13.07.2016	3349	0.0385	1.0049	−0.0049	−0.3798	0.6479
CAC 40	France	14.02.2003–13.07.2016	3428	0.0439	1.0181	−0.0181	−1.9562	0.9748
Hang Seng	Hong Kong	21.03.2003–13.07.2016	3025	0.0314	1.0156	−0.0156	−1.0143	0.8448
KOSPI	South Korea	27.03.2003–13.07.2016	3288	0.0342	0.9979	0.0021	0.1518	0.4397
AEX	The Netherlands	17.02.2003–13.07.2016	3427	0.0482	1.0224	−0.0224	−3.0345	0.9988
Swiss Market Index	Switzerland	25.02.2003–13.07.2016	3355	0.0361	1.0280	−0.0280	−3.2653	0.9995
IBEX 35	Spain	06.03.2003–13.07.2016	3393	0.0412	1.0082	−0.0082	−0.6931	0.7559
S&P CNX Nifty	India	31.08.2005–13.07.2016	2666	0.0474	1.0114	−0.0114	−0.7589	0.7760
IPC Mexico	Mexico	11.03.2003–13.07.2016	3353	0.0435	0.9968	0.0032	0.2495	0.4015
Bovespa	Brazil	02.04.2003–12.07.2016	3261	0.0380	0.9956	0.0044	0.3083	0.3789
S&P TSX	Canada	17.06.2005–13.07.2016	2764	0.0514	1.0048	−0.0048	−0.2313	0.5914
Euro STOXX 50	Euro Area	11.02.2003–13.07.2016	3405	0.0525	0.9790	0.0210	1.3759	0.0844
FT Straits Times	Singapore	28.02.2003–18.09.2015	3099	0.0208	0.9943	0.0057	0.4225	0.3363
FTSE MIB	Italy	21.02.2003–12.07.2016	3386	0.0447	1.0109	−0.0109	−1.0289	0.8482
		Forecast horizon h = 22						
S&P 500	United States	27.03.2003–13.07.2016	3337	0.0674	0.9895	0.0105	0.5282	0.2987
FTSE 100	United Kingdom	10.03.2003–13.07.2016	3358	0.0470	1.0218	−0.0218	−1.7227	0.9575
Nikkei 225	Japan	17.04.2003–13.07.2016	3212	0.0576	1.0556	−0.0556	−2.2697	0.9884
DAX 30	Germany	28.02.2003–13.07.2016	3390	0.0522	1.0249	−0.0249	−1.6120	0.9465
All Ordinaries	Australia	05.03.2003–13.07.2016	3337	0.0438	1.0194	−0.0194	−1.2082	0.8865
CAC 40	France	04.03.2003–13.07.2016	3416	0.0492	1.0231	−0.0231	−1.6952	0.9550
Hang Seng	Hong Kong	09.04.2003–13.07.2016	3013	0.0320	1.0176	−0.0176	−0.7589	0.7760
KOSPI	South Korea	16.04.2003–13.07.2016	3276	0.0386	0.9988	0.0012	0.0663	0.4736
AEX	The Netherlands	05.03.2003–13.07.2016	3415	0.0563	1.0251	−0.0251	−2.3303	0.9901
Swiss Market Index	Switzerland	13.03.2003–13.07.2016	3343	0.0444	1.0501	−0.0501	−4.3200	1.0000
IBEX 35	Spain	24.03.2003–13.07.2016	3381	0.0448	1.0202	−0.0202	−1.5523	0.9397
S&P CNX Nifty	India	19.09.2005–13.07.2016	2654	0.0517	1.0292	−0.0292	−1.4949	0.9325
IPC Mexico	Mexico	28.03.2003–13.07.2016	3341	0.0451	0.9958	0.0042	0.2817	0.3891
Bovespa	Brazil	22.04.2003–12.07.2016	3249	0.0415	0.9949	0.0051	0.3155	0.3762
S&P TSX	Canada	06.07.2005–13.07.2016	2752	0.0561	1.0015	−0.0015	−0.0597	0.5238
Euro STOXX 50	Euro Area	27.02.2003–13.07.2016	3393	0.0570	0.9927	0.0073	0.3960	0.3461
FT Straits Times	Singapore	18.03.2003–18.09.2015	3087	0.0262	1.0194	−0.0194	−0.9483	0.8285
FTSE MIB	Italy	11.03.2003–12.07.2016	3374	0.0497	1.0261	−0.0261	−2.1462	0.9841

Notes: This table reports the multiple-steps-ahead out-of-sample forecast evaluation results for the 18 international equity markets that we consider. Forecasts for horizons $h = 5, 10$ and 22 are shown in the top, middle and bottom panels, respectively. The target variable is (normalised) multi-period log (RV), as defined in (13). All column entries are the same as described in Table 4. The p-values corresponding to the DM-statistic are computed from HAC robust standard errors, where we use a *pre-whitening* step using an ARMA(1, 1) model for the DM_{t+h} sequence to reduce the initial autocorrelation in the series, and then apply a Quadratic Spectral (QS) kernel based non-parametric HAC estimator on the ARMA(1, 1) residuals. We follow Andrews and Monahan (1992) and choose the bandwidth optimally with an AR(1) as the approximating model, and then *re-colour* to obtain the HAC standard errors of the DM_{t+h} sequence.

Table 6
Multiple-step-ahead out-of-sample forecast evaluation results: HAR-CJL model (rolling window, $T_{is} = 750$).

Equity index	Country	Out-of-sample period	T_{os}	$\text{MSFE}_{\mathcal{M}}$	$\frac{\text{MSFE}_{\mathcal{M}}}{\text{MSFE}_{\text{RV}}}$	${}_{\mathcal{M}}R^2_{os}$	$\text{DM}_{\mathcal{M}}$-stat	p-value
		Forecast horizon h = 5						
S&P 500	United States	04.03.2003–13.07.2016	3354	0.0548	0.8943	0.1057	5.2751	0.0000
FTSE 100	United Kingdom	13.02.2003–13.07.2016	3375	0.0381	0.9811	0.0189	0.8408	0.2002
Nikkei 225	Japan	20.03.2003–13.07.2016	3229	0.0468	0.9713	0.0287	1.3831	0.0833
DAX 30	Germany	05.02.2003–13.07.2016	3407	0.0420	0.9416	0.0584	2.5738	0.0050
All Ordinaries	Australia	10.02.2003–13.07.2016	3354	0.0379	0.9419	0.0581	2.9159	0.0018
CAC 40	France	07.02.2003–13.07.2016	3433	0.0398	0.9594	0.0406	2.0871	0.0184
Hang Seng	Hong Kong	14.03.2003–13.07.2016	3030	0.0317	0.9680	0.0320	1.3781	0.0841
KOSPI	South Korea	20.03.2003–13.07.2016	3293	0.0333	0.9561	0.0439	2.6782	0.0037
AEX	The Netherlands	10.02.2003–13.07.2016	3432	0.0428	0.9710	0.0290	1.0478	0.1474
Swiss Market Index	Switzerland	18.02.2003–13.07.2016	3360	0.0314	0.9607	0.0393	1.9982	0.0228
IBEX 35	Spain	27.02.2003–13.07.2016	3398	0.0388	0.9627	0.0373	1.8718	0.0306
S&P CNX Nifty	India	24.08.2005–13.07.2016	2671	0.0464	0.9476	0.0524	2.3175	0.0102
IPC Mexico	Mexico	04.03.2003–13.07.2016	3358	0.0441	0.9199	0.0801	3.2374	0.0006
Bovespa	Brazil	26.03.2003–12.07.2016	3266	0.0372	0.9506	0.0494	2.5329	0.0057
S&P TSX	Canada	10.06.2005–13.07.2016	2769	0.0480	0.9181	0.0819	3.0119	0.0013
Euro STOXX 50	Euro Area	04.02.2003–13.07.2016	3410	0.0491	0.9035	0.0965	3.2930	0.0005
FT Straits Times	Singapore	21.02.2003–18.09.2015	3104	0.0202	0.9471	0.0529	1.7698	0.0384
FTSE MIB	Italy	14.02.2003–12.07.2016	3391	0.0409	0.9519	0.0481	3.0037	0.0013
		Forecast horizon h = 10						
S&P 500	United States	11.03.2003–13.07.2016	3349	0.0578	0.9251	0.0749	3.5505	0.0002
FTSE 100	United Kingdom	20.02.2003–13.07.2016	3370	0.0398	0.9855	0.0145	0.7103	0.2388
Nikkei 225	Japan	28.03.2003–13.07.2016	3224	0.0476	0.9998	0.0002	0.0075	0.4970
DAX 30	Germany	12.02.2003–13.07.2016	3402	0.0433	0.9563	0.0437	1.9158	0.0277
All Ordinaries	Australia	17.02.2003–13.07.2016	3349	0.0373	0.9744	0.0256	1.2282	0.1097
CAC 40	France	14.02.2003–13.07.2016	3428	0.0418	0.9690	0.0310	1.6908	0.0454
Hang Seng	Hong Kong	21.03.2003–13.07.2016	3025	0.0308	0.9961	0.0039	0.1463	0.4419
KOSPI	South Korea	27.03.2003–13.07.2016	3288	0.0335	0.9774	0.0226	1.2430	0.1069
AEX	The Netherlands	17.02.2003–13.07.2016	3427	0.0464	0.9854	0.0146	0.4881	0.3127
Swiss Market Index	Switzerland	25.02.2003–13.07.2016	3355	0.0347	0.9885	0.0115	0.7004	0.2419
IBEX 35	Spain	06.03.2003–13.07.2016	3393	0.0399	0.9747	0.0253	0.9850	0.1623
S&P CNX Nifty	India	31.08.2005–13.07.2016	2666	0.0471	1.0056	−0.0056	−0.2458	0.5971
IPC Mexico	Mexico	11.03.2003–13.07.2016	3353	0.0416	0.9531	0.0469	1.9800	0.0239
Bovespa	Brazil	02.04.2003–12.07.2016	3261	0.0371	0.9718	0.0282	1.2466	0.1063
S&P TSX	Canada	17.06.2005–13.07.2016	2764	0.0490	0.9562	0.0438	1.5875	0.0562
Euro STOXX 50	Euro Area	11.02.2003–13.07.2016	3405	0.0486	0.9056	0.0944	3.1437	0.0008
FT Straits Times	Singapore	28.02.2003–18.09.2015	3099	0.0199	0.9485	0.0515	1.6213	0.0525
FTSE MIB	Italy	21.02.2003–12.07.2016	3386	0.0430	0.9727	0.0273	1.4167	0.0783
		Forecast horizon h = 22						
S&P 500	United States	27.03.2003–13.07.2016	3337	0.0680	0.9989	0.0011	0.0401	0.4840
FTSE 100	United Kingdom	10.03.2003–13.07.2016	3358	0.0461	1.0029	−0.0029	−0.1405	0.5559
Nikkei 225	Japan	17.04.2003–13.07.2016	3212	0.0575	1.0536	−0.0536	−1.9687	0.9755
DAX 30	Germany	28.02.2003–13.07.2016	3390	0.0501	0.9838	0.0162	0.7232	0.2348
All Ordinaries	Australia	05.03.2003–13.07.2016	3337	0.0439	1.0222	−0.0222	−0.9649	0.8327
CAC 40	France	04.03.2003–13.07.2016	3416	0.0475	0.9882	0.0118	0.6332	0.2633
Hang Seng	Hong Kong	09.04.2003–13.07.2016	3013	0.0318	1.0134	−0.0134	−0.3485	0.6363
KOSPI	South Korea	16.04.2003–13.07.2016	3276	0.0393	1.0166	−0.0166	−0.5201	0.6985
AEX	The Netherlands	05.03.2003–13.07.2016	3415	0.0546	0.9936	0.0064	0.2351	0.4071
Swiss Market Index	Switzerland	13.03.2003–13.07.2016	3343	0.0435	1.0274	−0.0274	−1.7722	0.9618
IBEX 35	Spain	24.03.2003–13.07.2016	3381	0.0446	1.0153	−0.0153	−0.4656	0.6792
S&P CNX Nifty	India	19.09.2005–13.07.2016	2654	0.0528	1.0500	−0.0500	−1.8092	0.9648
IPC Mexico	Mexico	28.03.2003–13.07.2016	3341	0.0434	0.9590	0.0410	1.8165	0.0346
Bovespa	Brazil	22.04.2003–12.07.2016	3249	0.0410	0.9832	0.0168	0.6439	0.2598
S&P TSX	Canada	06.07.2005–13.07.2016	2752	0.0570	1.0178	−0.0178	−0.4127	0.6601
Euro STOXX 50	Euro Area	27.02.2003–13.07.2016	3393	0.0530	0.9245	0.0755	3.1312	0.0009
FT Straits Times	Singapore	18.03.2003–18.09.2015	3087	0.0258	1.0036	−0.0036	−0.1596	0.5634
FTSE MIB	Italy	11.03.2003–12.07.2016	3374	0.0488	1.0075	−0.0075	−0.4146	0.6608

Notes: This table reports the multiple-steps-ahead out-of-sample forecast evaluation results for the 18 international equity markets that we consider. Forecasts for horizons $h = 5, 10$ and 22 are shown in the top, middle and bottom panels, respectively. The target variable is (normalised) multi-period log (RV), as defined in (13). All column entries are the same as described in Table 4. The p-values corresponding to the DM-statistic are computed from HAC robust standard errors, where we use a *pre-whitening* step using an ARMA(1, 1) model for the DM_{t+h} sequence to reduce the initial autocorrelation in the series, and then apply a Quadratic Spectral (QS) kernel based non-parametric HAC estimator on the ARMA(1, 1) residuals. We follow Andrews and Monahan (1992) and choose the bandwidth optimally with an AR(1) as the approximating model, and then *re-colour* to obtain the HAC standard errors of the DM_{t+h} sequence.

in the empirical RV modelling literature so far, as most of the studies have focused on the U.S. market, using either the S&P 500, its constituents or S&P 500 futures data in their analyses.

In summary of the out-of-sample forecast evaluating results that we have presented in this section, it is clear that including the leverage effect leads to substantial improvements in out-of-sample predictions of volatility in all of the 18 international equity markets that we analyze. Moreover, this improvement has a lasting impact and can effect forecasts as far as 1 month ahead. On the contrary, we find the separation of the continuous and jump components to be of limited value in improving out-of-sample forecasts of RV in international equity markets. Only for 6 of the 17 non-U.S. equity markets can statistically significant and sizeable short-horizon forecast improvements over the benchmark HAR-RV model be realized.

5. Discussion

Why is the jump component of such little predictive value for the majority of international equity markets that we consider? One reason could be due to the vastly different number of stocks in each of the indices. For instance, the S&P 500 consists of 500 stocks, while 12 of the other indices are based on 20–50 stocks (see the second column of Table 1 that lists the number of stocks in each equity index).[16] Having an index made up of a much large number of individual stocks could make jumps seemingly more idiosyncratic. Nevertheless, since they are market capitalisation weighted, it is not clear whether a greater number of stocks, with many of them having a rather small weight, would lead to jumps being transmitted to the aggregate index more frequently, or in a more important way. With regards to this, we can notice that the FTSE 100, the Nikkei 225, and the KOSPI are all made up of a large number of stocks, yet show no predictive improvement when jump components are included in the forecasting models for volatility. The only other two equity indices where jumps do show some predictive improvement and which contain a larger number of constituent stocks are the Canadian S&P TSX (250 stocks) and the Australian All Ordinaries (500 stocks). Given that we find evidence of both, that is, a larger number of constituents in an index resulting in significant as well insignificant contributions from jumps to volatility, it is difficult to conclude that the number of stocks is the driving factor behind these results.[17]

What is noticeable for the 6 equity markets for which jumps seem to matter, i.e., the IPC Mexico, Bovespa, the S&P TSX, the DAX 30, S&P CNX Nifty, and the All Ordinaries, is that the first three trade in the same time zone as the S&P 500. The trading hours of the New York Stock Exchange are from 09:30 to 16:00 Eastern Standard Time (EST). The S&P TSX and the IPC Mexico trade over the same time window (09:30–16:00), while the Brazilian Bovespa trades between 10:00 and 17:00 local time, which corresponds to 08:00–15:00 EST. One might thus conjecture that the importance of the jump component for these three equity markets could be driven by jump information originating in the S&P 500, or the U.S. market at least. Given the findings of the recent study by Rapach et al. (2013), this could indeed be the most reasonable explanation.[18] If the arrival of news in the U.S. leads to a jump in U.S. equity prices, markets operating in the same time zone are likely to show a jump response on the same trading day, while for other markets, it would be absorbed in the opening prices the following day.

Given that this seems to be the most likely explanation, one may want to examine how *'correlated'* or alike the jump components in the different equity markets are. This question is, nonetheless, difficult for us to address empirically for the following two reasons. First, we do not have access to high frequency data to be able to construct a co-jump test such as, for instance, in Bollerslev et al. (2008). Second, as an alternative, taking the independently constructed jump components for these markets and for the S&P 500, and then running Tobit (censored) type of regressions of non-U.S. jumps on a constant and S&P 500 jumps to assess this relationship is unlikely to be valid, because of the regressor variable (S&P 500 jumps) also being a censored variable (see Rigobon and Stoker (2009) for a more detailed treatment of the problems that arise from using censored regressors in a Tobit model). Even though one could make an argument to defend the assumption of exogeneity of the S&P 500, due to its size and prevalence, inference is still likely to be non-standard because of the truncated nature of the regressor variables. We have also tried to address this question from an out-of-sample forecast evaluation perspective, by adding the U.S. jump components as additional predictors, but have found rather mixed results that require further analysis, which is beyond the scope of this paper. We therefore delegate the question of how U.S. jumps spillover to other equity markets to future research.

[16] We thank an anonymous referee for pointing this out to us, and for encouraging us to conjecture why these differences in the role of jumps might occur. Note that it is not our intention here to give an in depth account of how the different equity indices are constructed. Nevertheless, for broad informational purposes, we point out that most of these equity indices are market capitalisation weighted indices. Some also incorporate a measure of trading activity or liquidity into either the selection of the stocks included in the index, or the weight construction of the index (for instance, the Swiss Market Index, and the Amsterdam Exchange index (AEX).

[17] Are indices with a smaller number of constituent stocks more likely to be driven by idiosyncratic jump components? The Netherlands AEX index, for instance, is made up of 25 stocks, with 15% of the index weight given to Royal Dutch Shell and Unilever, respectively, with another 11% being determined by ING Group. These three stocks represent 40% of the index. One could thus expect any idiosyncratic shocks to these three individual stocks to have a fairly strong impact, and therefore jumps to be of a frequent occurrence and hence predictive value. Yet, the separation of RV into a continuous and a jump component has no predictive information at all for the AEX. The situation is similar with some of the other indices that are made up of a small number of constituent stocks, and where the index weights are dominated by a hand full of firms. Once again, the only reasonable conclusion from this is that the number of constituent stocks in an index is an unlikely source of our finding of jumps being largely irrelevant.

[18] Rapach et al. (2013) write on page 1635, *"Why does the United States lead much of the world? . . ., since the U.S. equity market is the world's largest, investors likely focus more intently on this market, so that information on macroeconomic fundamentals relevant for equity markets worldwide diffuses gradually from the U.S. market to other countries' markets."*

6. Conclusion

We evaluate the importance of jumps and the leverage effect on forecasts of realized volatility in a large cross-section of 18 international equity markets, using two widely employed empirical models for realized volatility in our analysis. The first one is the HAR-CJ model of Andersen et al. (2007), which includes daily, weekly and monthly HAR components of the continuous and the discontinuous (jump) components of realized volatility as regressors. The second is the HAR-CJL model of Corsi and Renó (2012), which adds a HAR structured leverage effect to the HAR-CJ model of Andersen et al. (2007). Both models are evaluated against the baseline HAR-RV model specification of Corsi (2009).

We find that the separation of realized volatility into a continuous and a discontinuous (jump) component is highly beneficial for the S&P 500, but much less so for the remaining 17 international equity markets that we analyze, where only for 6 of the 17 non-U.S. equity markets statistically significant and sizeable forecast improvements are realized. These forecast gains are further short lived and deteriorate rather quickly as the forecast horizon increases, with forecasts for only 2 of the 17 non-U.S. equity markets remaining statistically significant at the 5-day-ahead horizon, and only one at the 10-day-ahead horizon. These forecast gains are, moreover, rather small in magnitude, yielding out-of-sample R^2 values of only around 2%. For the S&P 500, forecast gains remain significant and sizeable at the 5 and 10-day-ahead horizons, with out-of-sample R^2 values of 5.69% and 3.59%, respectively, but are insignificant at the 22-day-ahead horizon.

Adding a HAR structured leverage effect to the HAR-CJ model leads not only to statistically significant, but also much larger forecast improvements for all 18 international equity markets that we consider. One-step-ahead forecasts are significant and sizeable for all 18 equity markets, yielding out-of-sample R^2 values in the range from 2.71% (FTSE 100) to 12.18% (S&P 500). At the 5-day-ahead horizon, forecast gains remain significant and sizeable for 16 of the 18 equity markets, with out-of-sample R^2 values ranging between 2.87% and 10.57%, while at the 10-day-ahead horizon, 8 remain significant, generating a similar range of out-of-sample R^2 values as at the 5-day-ahead horizon. For the Euro STOXX 50 and the IPC Mexico forecast gains are significant (at the 5% level) and sizeable ($R^2_{os} \geqslant 4.10\%$) even at the 22-day-ahead horizon.

In summary, our findings show that the separation of realized volatility into a continuous and a discontinuous (jump) component is only of marginal value for realized volatility forecasts at an international (cross-sectional) equity market level, while the benefits of including a HAR structured leverage effect are considerably larger and longer lasting.

References

Andersen, Torben G., Bollerslev, Tim, Diebold, Francis X., 2007. Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility. Rev. Econ. Stat. 89 (4), 701–720.

Andersen, Torben G., Bollerslev, Tim, Diebold, Francis X., Ebens, Heiko, 2001. The distribution of realized stock return volatility. J. Financ. Econ. 61 (1), 43–76.

Andersen, Torben G., Bollerslev, Tim, Diebold, Francis X., Labys, Paul, 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.

Andrews, Donald W.K., Monahan, Christopher J., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. Econometrica 60 (4), 953–966.

Barndorff-Nielsen, Ole E., Shephard, Neil, 2004. Power and bipower variation with stochastic volatility and jumps. J. Financ. Economet. 2 (1), 1–37.

Barndorff-Nielsen, Ole E., Shephard, Neil, 2005. Measuring the Impact of Jumps in Multivariate Price Processes using Bipower Covariation. Working Paper. Nuffield College, Oxford University.

Barndorff-Nielsen, Ole E., Shephard, Neil, 2006. Econometrics of testing for jumps in financial economics using bipower variation. J. Financ. Economet. 4 (1), 1–30.

Bollerslev, Tim, Law, Tzuo Hann, Tauchen, George, 2008. Risk, jumps, and diversification. J. Economet. 144 (1), 234–256.

Bollerslev, Tim, Litvinova, Julia, Tauchen, George, 2006. Leverage and volatility feedback effect in high-frequency data. J. Financ. Economet. 4 (3), 353–384.

Buncic, Daniel, Gisler, Katja I.M., 2016. Global equity market volatility spillovers: a broader role for the United States. Int. J. Forecast. 32 (4), 1317–1339.

Campbell, John Y., Thompson, Samuel B., 2008. Predicting excess stock returns out of sample: can anything beat the historical average? Rev. Financ. Stud. 21 (4), 1509–1531.

Corsi, Fulvio, 2009. A simple approximate long-memory model of realized volatility. J. Financ. Economet. 7 (2), 174–196.

Corsi, Fulvio, Audrino, Francesco, Renó, Roberto, 2012. HAR modeling for realized volatility forecasting. In: Bauwens, Luc, Hafner, Christian, Laurent, Sebastien (Eds.), Handbook of Volatility Models and their Applications. John Wiley and Sons, Inc., pp. 363–382.

Corsi, Fulvio, Renó, Roberto, 2012. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. J. Bus. Econ. Stat. 30 (3), 368–380.

Diebold, Francis X., Mariano, Roberto S., 1995. Comparing predictive accuracy. J. Bus. Econ. Stat. 13 (1), 253–263.

Figlewski, Stephen, Wang, Xiaozu, 2000. Is the 'Leverage Effect' a Leverage Effect? SSRN Working Paper No. 256109. Available from: <http://ssrn.com/abstract=256109>.

Giacomini, Raffaella, White, Halbert, 2006. Tests of conditional predictive ability. Econometrica 74 (6), 1545–1578.

Heber, Gerd, Lunde, Asger, Shephard, Neil, Sheppard, Kevin, 2009. Oxford-Man Institute's Realized Library. Oxford-Man Institute, University of Oxford. Available from: <http://realized.oxford-man.ox.ac.uk/>.

Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variance. J. Financ. Economet. 3 (4), 456–499.

Liu, Lily Y., Patton, Andrew J., Sheppard, Kevin, 2015. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. J. Economet. 187 (1), 293–311.

Müller, Ulrich A., Dacorogna, Michel M., Davé, Rakhal D., Olsen, Richard B., Pictet, Olivier V., Ward, John R., 1993. Fractals and intrinsic time – a challenge to econometricians. In: Invited presentation at the 39th International AEA Conference on Real Time Econometrics, 14–15 October 1993, Luxembourg.

Newey, Whitney K., West, Kenneth D., 1994. Automatic lag selection in covariance matrix estimation. Rev. Econ. Stud. 61 (4), 631–653.

Patton, Andrew J., Sheppard, Kevin, 2015. Good volatility, bad volatility: signed jumps and the persistence of volatility. Rev. Econ. Stat. 97 (3), 683–697.

Prokopczuk, Marcel, Symeonidis, Lazaros, Simen, Chardin Wese, 2016. Do jumps matter for volatility forecasting? Evidence from energy markets. J. Fut. Markets 36 (8), 758–792.

Rapach, David E., Strauss, Jack K., Zhou, Guofu, 2013. International stock return predictability: what is the role of the United States? J. Finan. 68 (4), 1633–1662.

Rigobon, Roberto, Stoker, Thomas M., 2009. Bias from censored regressors. J. Bus. Econ. Stat. 27 (3), 340–353.