3-2019

Forecasting equity index volatility by measuring the linkage among component stocks

Yue QIU

Tian XIE

Jun YU
Singapore Management University, yujun@smu.edu.sg

Qiankun ZHOU

SMU ECONOMICS & STATISTICS

Forecasting Equity Index Volatility by Measuring the Linkage among Component Stocks

Yue Qiu, Tian Xie, Jun Yu, Qiankun Zhou

March 2019

Paper No. 07-2019

Forecasting Equity Index Volatility by Measuring the Linkage among Component Stocks

Yue Qiu,* Tian Xie,† Jun Yu,† Qiankun Zhou‡

March 2, 2019

Abstract

The linkage among the realized volatilities across component stocks are important when modeling and forecasting the relevant index volatility. In this paper, the linkage is measured via an extended Common Correlated Effects (CCE) approach under a panel heterogeneous autoregression model where unobserved common factors in errors are assumed. Consistency of the CCE estimator is obtained. The common factors are extracted using the principal component analysis. Empirical studies show that realized volatility models exploiting the linkage effects lead to significantly better out-of-sample forecast performance, for example, an up to 32% increase in the pseudo R^2. We also conduct various forecasting exercises on the the linkage variables that compare conventional regression methods with popular machine learning techniques.

JEL classification: C31, C32, G12, G17
Keywords: Volatility Forecasting; Heterogeneous autoregression; Common correlated effect; Factor analysis; Random forest

*WISE and School of Economics, Xiamen University, Xiamen, Fujian 361005, China.
†School of Economics and Lee Kong Chian School of Business, Singapore Management University.
‡Department of Economics, Louisiana State University.

1 Introduction

Volatility forecasting is central to financial institutions and market regulators. Portfolio managers tends to maximize returns when facing risk limits. With the development of realized variation based on high-frequency data, we are able to better measure financial market volatility. From then on, various volatility forecasting models have been put forward in the literature, such as the renowned fractionally integrated autoregressive moving average models (ARFIMA) used in Andersen, Bollerslev, Diebold, and Labys (2001) and the heterogeneous autoregressive (HAR) model proposed by Corsi (2009). No matter how complicate forms the above models can take, most of them rely on the asset-specific realized volatility histories. On the other hand, the comovement and spillover effect of risk across assets is well documented in the existing volatility literature and has been modeled by multivariate GARCH and stochastic volatility models.[1]

This paper exploits the linkages among the realized volatilities across component stocks to improve the corresponding stock index volatility forecasting. We propose a heterogeneous panel HAR (HARP) model assuming unobserved common factors to grasp the linkages. Our framework is based on the CCE estimator of Pesaran (2006) and allows us to extract unobserved common factors from the residuals of our econometric model. We then prove consistency of the CCE estimator within our framework. Regarding the specification of cross-sectional unit regressors, we follow the realized semivariance models of Patton and Sheppard (2015). Another important step for our forecasting implementation is to model the dynamics of unobserved common factors. We conduct various forecasting exercises and compare regression methods with popular machine learning techniques.

We consider empirical applications to many equity indices, including the NASDAQ 100

[1] See Bauwens, Laurent, and Rombouts (2006) and Asai, McAleer, and Yu (2006) for more detailed reviews and discussions.

exchange traded fund, the Dow Jones Industrial Average, the Dow Jones Transportation Average, and the Dow Jones Utility Average. Several novel findings are summarized as follows. First, the cross-sectional correlation of realized volatilities does exist for the considered component stocks. This finding holds irrespective of the underlying models.[2] Second, the in-sample results suggest that the role of unobserved common factors at explaining future volatility is nontrivial. They even carry partially the information contained in realized semivariances, especially the negative realized semivariances. Third and perhaps more importantly, we show that incorporating unobserved common factors into the HAR-type regressions leads to large and significant improvements in forecast accuracy.

It should be noted that there are other empirical applications of the CCE estimator in the literature. For example, the CCE approach allows Kapetanios and Pesaran (2004) to estimate asset return equations with both observed and unobserved common factors. Bernoth and Pick (2011) utilize the CCE framework to model the linkages between bank and insurance companies so as to improve forecasting the systemic risk. Chudik, Mohaddes, Pesaran, and Raissi (2017) develop tests for debt threshold effects in the context of dynamic heterogeneous panel data models, where the CCE estimator produces unbiased and consistent coefficient estimates for threshold variables. However, we are not aware of any application of the CCE estimator to the problem of volatility forecasting.

This paper makes several contributions in several strands of literature. The first contribution is to extend the HAR model to a heterogeneous panel data model to exploit the linkages among the realized volatilities across component stocks. Secondly, we introduce the CCE approach to estimation of the panel data model and establish consistency of the CCE estimator. Thirdly, we contribute to the literature pioneered by Bollerslev, Hood, Huss,

[2]After some normalization of daily realized volatilites for a wide range of asset classes, Bollerslev, Hood, Huss, and Pedersen (2018) reach a conclusion closely related to ours. Their “normalized risk measuresâ€ exhibit almost identical unconditional distributions and similar highly persistent autocorrelation functions when comparing across assets and asset classes.

and Pedersen (2018) that seek to exploit strong commonality in volatility in a "global risk factor" to improve forecast accuracy. Bollerslev, Hood, Huss, and Pedersen (2018) defined the global risk factor as the average normalized realized volatilities across all assets, whereas our linkage factors are extracted using the principal component analysis from the residuals of the CCE framework. Both approaches try to utilize the panel data information to better volatility forecasts, but from different angles. We show that the model that exploits the linkage effects can lead to up to 32% increase in the pseudo R^2 relative to models without accommodating the linkage effects for volatility forecasting.

In the next section, we review a list of reference HAR-type models. Section 3 discusses the econometric approach and the forecasting procedure. Section 4 describes the data, which are analyzed in Section 5. Section 6 conducts some robustness checks. Finally, Section 7 concludes. An online supplement contains extra empirical and theoretical results.

2 A review of reference models

Before moving into the panel-based HAR model, it is useful to review some HAR-type reference models, which act either as our basic specification for the CCE estimator or as comparison models in the subsequent exercise. Following Andersen and Bollerslev (1998), the M-sample daily realized variance (RV) at day t can be calculated by summing the corresponding M equally spaced intra-daily squared returns $r_{t,j}$. Here, the subscript t indexes the day, and j indicates the time interval within day t,

$$\mathrm{RV}_t \equiv \sum_{j=1}^{M} r_{t,j}^2, \quad \text{for } t = 1, 2, ..., T,\ j = 1, 2, ..., M, \tag{1}$$

where $r_{t,j} = p_{t,j} - p_{t,j-1}$ with $p_{t,j}$ being the log-price at time (t, j).

To model realized variation, a series of HAR-type models were invented in the literature, see Corsi, Audrino, and Renò (2012) for a survey. The basic HAR model was introduced by Corsi (2009) and has gained great popularity because of its estimation simplicity and outstanding out-of-sample performance. The basic HAR model in Corsi (2009) postulates that the h-step-ahead daily RV_{t+h} can be modeled by

$$\mathrm{RV}_{t+h} = \beta_0 + \beta_d \mathrm{RV}_t^{(1)} + \beta_w \mathrm{RV}_t^{(5)} + \beta_m \mathrm{RV}_t^{(22)} + \varsigma_{t+h}, \tag{2}$$

where the explanatory variables can take the general form of $\mathrm{RV}_t^{(l)}$. It is defined by

$$\mathrm{RV}_t^{(l)} \equiv l^{-1} \sum_{s=1}^{l} \mathrm{RV}_{t-s} \tag{3}$$

as the l period averages of daily RV, the βs are the coefficients, and $\{\varsigma_t\}_t$ is the error term. Since each $\mathrm{RV}_t^{(l)}$ can be regarded as a volatility cascade, generated by the actions of distinct types of market participants trading at daily, weekly or monthly frequencies (Müller et al, 1993), the lag structure in the HAR model is fixed at some lag index vector $\boldsymbol{l} = [1, 5, 22]$.

Andersen, Bollerslev, and Diebold (2007) extend the standard HAR model from two perspectives. First, they added the daily jump component J_t to Equation (2) to explicitly capture its impacts. The extended model is denoted as the HAR-J model,

$$\mathrm{RV}_{t+h} = \beta_0 + \beta_d \mathrm{RV}_t^{(1)} + \beta_w \mathrm{RV}_t^{(5)} + \beta_m \mathrm{RV}_t^{(22)} + \beta^j \mathrm{J}_t + \varsigma_{t+h}, \tag{4}$$

where the empirical measurement of the squared jumps is $\mathrm{J}_t = \max(\mathrm{RV}_t - \mathrm{BPV}_t, 0)$, and the realized bipower variation (BPV) is defined as $\mathrm{BPV}_t \equiv (2/\pi)^{-1} \sum_{j=2}^{M} |r_{t,j-1}||r_{t,j}|$. Second, through a decomposition of RV by the $Z_{1,t}$ statistic in Huang and Tauchen (2005) into the continuous sample path and the jump components, they extend the HAR-J model by explicitly incorporating the above two types of volatility components. The $Z_{1,t}$ statistic distinguishes the "significant" jumps CJ_t from continuous sample path components CSP_t:

$$\begin{aligned}\mathrm{CSP}_t &\equiv \mathbb{I}(Z_t \leq \Phi_\alpha)\cdot \mathrm{RV}_t + \mathbb{I}(Z_t \leq \Phi_\alpha)\cdot \mathrm{BPV}_t,\\ \mathrm{CJ}_t &\equiv \mathbb{I}(Z_t > \Phi_\alpha)\cdot \max(\mathrm{RV}_t - \mathrm{BPV}_t, 0),\end{aligned}$$

where Z_t is the ratio-statistic defined in Huang and Tauchen (2005), and Φ_α is the cumulative distribution function (CDF) of a standard Gaussian distribution with α level of significance. The daily, weekly, and monthly average components of CSP_t and CJ_t are then constructed in the same manner as $\mathrm{RV}^{(l)}$ in Equation (3). The model specification for the continuous HAR-J, namely, HAR-CJ, is given by

$$\mathrm{RV}_{t+h} = \beta_0 + \beta_d^c \mathrm{CSP}_t^{(1)} + \beta_w^c \mathrm{CSP}_t^{(5)} + \beta_m^c \mathrm{CSP}_t^{(22)} + \beta_d^j \mathrm{CJ}_t^{(1)} + \beta_w^j \mathrm{CJ}_t^{(5)} + \beta_m^j \mathrm{CJ}_t^{(22)} + \varsigma_{t+h}. \tag{5}$$

Note that compared with the HAR-J model, the HAR-CJ model explicitly controls for the weekly and monthly components of continuous jumps. Thus, the HAR-J model can be treated as a special and restrictive case of the HAR-CJ model for $\beta_d = \beta_d^c + \beta_d^j$, $\beta^j = \beta_d^j$, $\beta_w = \beta_w^c + \beta_w^j$, and $\beta_m = \beta_m^c + \beta_m^j$.

To capture the information from signed high-frequency variation, Patton and Sheppard (2015) developed a series of realized semivariance HAR (HAR-RS) models. The first one, HAR-RS-I model, completely decomposes the $\mathrm{RV}^{(1)}$ in Equation (2) into two asymmetric semi-variances, RS_t^+ and RS_t^-,

$$\mathrm{RV}_{t+h} = \beta_0 + \beta_d^+ \mathrm{RS}_t^+ + \beta_d^- \mathrm{RS}_t^- + \beta_w \mathrm{RV}_t^{(5)} + \beta_m \mathrm{RV}_t^{(22)} + \varsigma_{t+h}, \tag{6}$$

where $\mathrm{RS}_t^- = \sum_{j=1}^M r_{t,j}^2 \cdot \mathbb{I}(r_{t,j} < 0)$ and $\mathrm{RS}_t^+ = \sum_{j=1}^M r_{t,j}^2 \cdot \mathbb{I}(r_{t,j} > 0)$. To verify the actual effects of signed variations, they include an additional term capturing the leverage effect,

$\mathrm{RV}_t^{(1)} \cdot \mathbb{I}(r_t < 0)$. The second model in Equation (7) is denoted as HAR-RS-II,

$$\mathrm{RV}_{t+h} = \beta_0 + \beta_1 \mathrm{RV}_t^{(1)} \cdot \mathbb{I}(r_t < 0) + \beta_d^+ \mathrm{RS}_t^+ + \beta_d^- \mathrm{RS}_t^- + \beta_w \mathrm{RV}_t^{(5)} + \beta_m \mathrm{RV}_t^{(22)} + \varsigma_{t+h}. \quad (7)$$

The third and fourth models in Patton and Sheppard (2015), denoted as HAR-SJ-I (Equation (8)) and HAR-SJ-II (Equation (9)), respectively, examine the role that decomposing realized variances into signed jump variations and bipower variation (BPV) can play in forecasting volatility:

$$\mathrm{RV}_{t+h} = \beta_0 + \beta_d^j \mathrm{SJ}_t + \beta_d^{bpv} \mathrm{BPV}_t + \beta_w \mathrm{RV}_t^{(5)} + \beta_m \mathrm{RV}_t^{(22)} + \varsigma_{t+h}, \quad (8)$$

$$\mathrm{RV}_{t+h} = \beta_0 + \beta_d^{j-} \mathrm{SJ}_t^- + \beta_d^{j+} \mathrm{SJ}_t^+ + \beta_d^{bpv} \mathrm{BPV}_t + \beta_w \mathrm{RV}_t^{(5)} + \beta_m \mathrm{RV}_t^{(22)} + \varsigma_{t+h}, \quad (9)$$

where $\mathrm{SJ}_t = \mathrm{RS}_t^+ - \mathrm{RS}_t^-$, $\mathrm{SJ}_t^+ = \mathrm{SJ}_t \cdot \mathbb{I}(\mathrm{SJ}_t > 0)$, and $\mathrm{SJ}_t^- = \mathrm{SJ}_t \cdot \mathbb{I}(\mathrm{SJ}_t < 0)$. The HAR-SJ-II model further extends the HAR-SJ-I model by distinguishing the effect of a positive jump variation from that of a negative jump variation.

In practice, RV_{t+h} is unobservable at time t. Hence both the dependent and explanatory variables of the models in Section 2 must take h-period of lags.

3 The panel HAR model

In this section we construct a heterogeneous panel HAR model with error cross-sectional dependence, which is an extension of the framework of Chudik and Pesaran (2015). Let y_{it} be the realized variance of the i^{th} individual asset at time t for $i = 1, ..., N$; $t = 1, ..., T$. Suppose that y_{it} is described by the following heterogeneous dynamic panel data model,

$$y_{it} = \boldsymbol{\alpha}_i' \boldsymbol{d}_t + \sum_{l \in \mathcal{L}} \phi_i^{(l)} \bar{y}_{it}^{(l)} + \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + u_{it}, \quad (10)$$

$$\bar{y}_{it}^{(l)} \;=\; l^{-1}\sum_{s=1}^{l} y_{i,t-s}, \tag{11}$$

for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$, where $\boldsymbol{d}_t$ is a $r \times 1$ vector of observed common effects, including deterministics such as intercepts or seasonal dummies, $\boldsymbol{x}_{it}$ is a $k \times 1$ vector of regressors specific to cross-sectional unit i at time t, and $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ are parameter vectors. $y_{i,t-s}$ represents the s^{th} lag of y_{it}, and $\bar{y}_{it}^{(l)}$ is the HAR component, which is the average of previous l periods of y_{it}. $\phi_i^{(l)}$ is the coefficient for $\bar{y}_{it}^{(l)}$, $\mathcal{L}$ is the lag index vector of l and we let $L = \max(\mathcal{L})$.[3] Furthermore, we assume that the realized variance of individual stocks is correlated beyond what can be explained by the observed determinants because the error term, u_{it}, comprises m unobserved common factors,[4]

$$u_{it} = \gamma_i' \boldsymbol{f}_t + \epsilon_{it}, \tag{12}$$

where γ_i is the $m \times 1$ vector of factor loadings, $\boldsymbol{f}_t$ is the $m \times 1$ vector of unobserved common factors that could themselves be serially correlated, and ϵ_{it} are the idiosyncratic errors assumed to be independently distributed of $(\boldsymbol{d}_t, \boldsymbol{x}_{it})$ and uncorrelated with the factors. Assume that $\boldsymbol{f}_t$ has vector autoregressive form,

$$\boldsymbol{f}_t = \Phi_f \boldsymbol{f}_{t-1} + \zeta_t. \tag{13}$$

Following Pesaran (2006) and Chudik and Pesaran (2015), the unobserved factors, $\boldsymbol{f}_t$, can be also correlated with $(\bar{y}_{it}^{(l)}, \boldsymbol{d}_t, \boldsymbol{x}_{it})$. To permit such a possibility, we assume a fairly

[3]Following the convention of the HAR-RV literature, we set $L = 22$ in our empirical exercise. In this literature, it is common to set $\mathcal{L} = [1, 5, 22]$ to claim that tomorrow's realized variance can be a sum of daily, weekly, and monthly averages of past realized variances.

[4]The recent paper by Bollerslev, Hood, Huss, and Pedersen (2018) interprets the common factors as combined economic forces from the investor sentiment, the variance risk premium and the news surprise variable.

general model for individual specific regressors, $\boldsymbol{x}_{it}$,

$$\boldsymbol{x}_{it} = \boldsymbol{\Pi}_i \boldsymbol{y}_{it,-L} + \boldsymbol{\Lambda}_i' \boldsymbol{d}_t + \boldsymbol{\Gamma}_i' \boldsymbol{f}_t + \boldsymbol{v}_{it}, \tag{14}$$

where $\boldsymbol{y}_{it,-L} = (y_{i,t-1}, \ldots, y_{i,t-L})'$, $\boldsymbol{\Lambda}_i$ and $\boldsymbol{\Gamma}_i$ are $r \times k$ and $m \times k$ matrix of factor loadings for observed and unobserved factors, respectively, $\boldsymbol{\Pi}_i$ is a $k \times L$ matrix of unknown coefficients, and $\boldsymbol{v}_{it}$ is assumed to follow a general linear covariance stationary process distributed independently of the idiosyncratic errors, ϵ_{it}. The equations (10) to (14) hitherto set out our panel HAR volatility forecasting model, denoted as the HARP model.[5]

3.1 The CCE estimator and its consistency

Clearly, forecasts of y_{it} need estimates of the parameters and the unobserved common factors. Unfortnately, conventional panel estimators of equation (10) yield inconsistent estimates of coefficients due to the correlation of regressors $(\boldsymbol{x}_{it}, \bar{y}_{it}^{(l)})$ and error terms u_{it}. To see this, since $\boldsymbol{f}_t$ is assumed to be serially correlated, the error term in (10), u_{it}, is thus serially correlated through (12), which further entails the correlation of u_{it} and $y_{i,t-s}$. Since the HAR components $\bar{y}_{it}^{(l)}$ are linear functions of $y_{i,t-s}$, they are correlated with the error terms u_{it}. It is more apparent to witness the correlation between $\boldsymbol{x}_{it}$ and u_{it}, regarding the fact that they share common factors $\boldsymbol{f}_t$.

In this section we address the issue of inconsistency and demonstrate how to estimate the slope coefficients $(\phi_i^{(1)}, ..., \phi_i^{(L)}, \boldsymbol{\beta}_i)$ from (10), employing the CCE estimator proposed by Pesaran (2006). Moreover, we prove that the CCE estimator holds its consistency in the HAR specification. Drawing from Chudik and Pesaran (2015), we posit the following assumptions for equations (10), (12) and (14).

[5] Note that before we pin down the underlying specification of $\boldsymbol{x}_{it}$, the HARP model can be quite general to accommodate other HAR-type models in the literature.

Assumption 1 *Individual-Specific Errors: The individual-specific errors* ϵ_{it} *and* $\boldsymbol{v}_{jt'}$ *are distributed independently for all* i, j, t*, and* t'*. We further assume that for each* i*,* ϵ_{it} *has uniformly bounded positive variance,* $\sup_i \sigma_i^2 \leq K < \infty$*, for some constant* K*, and* $\boldsymbol{v}_{it}$ *has covariance matrices,* Σ_{v_i}*, which are non-singular and satisfy* $\sup_i \|\Sigma_{v_i}\| \leq K < \infty$*, where* $\|\cdot\|$ *denotes the spectral norm. Both* ϵ_{it} *and* $\boldsymbol{v}_{it}$ *have finite fourth-order cumulants.*

Assumption 2 *Factor Loadings: The unobserved factor loadings* γ_i *and* $\boldsymbol{\Gamma}_i$ *are independent and identically distributed across* i*, and of the individual specific errors* ϵ_{jt} *and* $\boldsymbol{v}_{jt}$*, the common factors* $(\boldsymbol{d}_t', \boldsymbol{f}_t')$ *for all* i, j*, and* t*. Also,* γ_i *and* $\boldsymbol{\Gamma}_i$ *have fixed means* γ *and* $\boldsymbol{\Gamma}$*, respectively, and finite variances.*

Remark Assumption 1 and 2 are standard in the literature of panel data model with common factor error structure; see, Pesaran (2006) and Chudik and Pesaran (2015).

We note that equations (10) and (12) can be combined, and rewritten as

$$
\begin{aligned}
y_{it} &= \boldsymbol{\alpha}_i'\boldsymbol{d}_t + \sum_{l=1}^{L} \psi_{i,l} y_{i,t-s} + \boldsymbol{\beta}_i'\boldsymbol{x}_{it} + \boldsymbol{\gamma}_i'\boldsymbol{f}_t + \epsilon_{it} \\
&= \boldsymbol{\alpha}_i'\boldsymbol{d}_t + \boldsymbol{\psi}_i'\boldsymbol{y}_{it,-L} + \boldsymbol{\beta}_i'\boldsymbol{x}_{it} + \boldsymbol{\gamma}_i'\boldsymbol{f}_t + \epsilon_{it},
\end{aligned} \tag{15}
$$

where $\boldsymbol{y}_{it,-L} = (y_{i,t-1}, \ldots, y_{i,t-L})'$ and $\boldsymbol{\psi}_i = (\psi_{1,i}, \ldots, \psi_{L,i})'$ are given by

$$
\begin{aligned}
\psi_{i,1} &= \phi_i^{(1)} + \frac{\phi_i^{(2)}}{2} + \cdots + \frac{\phi_i^{(L)}}{L}, \quad \psi_{i,2} = \frac{\phi_i^{(2)}}{2} + \cdots + \frac{\phi_i^{(L)}}{L}, \\
&\vdots \\
\psi_{i,L-1} &= \frac{\phi_i^{(L-1)}}{L-1} + \frac{\phi_i^{(L)}}{L}, \quad \psi_{i,L} = \frac{\phi_i^{(L)}}{L}.
\end{aligned}
$$

As a result, we note that equation (15) can be viewed as a panel restricted AR(L) model with error cross sectional dependence.

The parameters of interest in (15) are $\boldsymbol{\psi}_i$ and $\boldsymbol{\beta}_i$ while $\boldsymbol{f}_t$ is unobserved. To estimate $\boldsymbol{\psi}_i$ and $\boldsymbol{\beta}_i$ in model (15), we can adopt the the common correlated effects (CCE) estimator proposed by Pesaran (2006), which is further extended by Chudik and Pesaran (2015) to dynamic panel data models. To this end, let $\boldsymbol{z}_{it} = \left(y_{it}, \boldsymbol{x}'_{it}\right)'$, and rewrite (14) and (15) in the below system of equations

$$\boldsymbol{A}_{0i}\boldsymbol{z}_{it} = \boldsymbol{B}_i\boldsymbol{d}_t + \sum_{l=1}^{L} \boldsymbol{A}_{i,l}\boldsymbol{z}_{it-l} + \boldsymbol{C}_i\boldsymbol{f}_t + \boldsymbol{e}_{it}, \tag{16}$$

where

$$\boldsymbol{A}_{0i} = \begin{pmatrix} 1 & -\boldsymbol{\beta}'_i \\ \underset{k\times 1}{\boldsymbol{0}} & I_k \end{pmatrix}, \boldsymbol{B}_i = \begin{pmatrix} \boldsymbol{\alpha}'_i \\ \boldsymbol{\Lambda}'_i \end{pmatrix}, \boldsymbol{C}_i = \begin{pmatrix} \gamma'_i \\ \boldsymbol{\Gamma}'_i \end{pmatrix}, \boldsymbol{e}_{it} = \begin{pmatrix} \epsilon_{it} \\ \boldsymbol{v}_{it} \end{pmatrix},$$

$$\boldsymbol{A}_{i,l} = \begin{pmatrix} \psi_{i,l} & \underset{1\times k}{\boldsymbol{0}} \\ \boldsymbol{\Pi}_i^{(l)} & \underset{k\times k}{\boldsymbol{0}} \end{pmatrix}, \text{ for } l = 1, \ldots, L, \text{ and } \boldsymbol{\Pi}_i^{(l)} \text{ denotes the } l\text{th column of } \boldsymbol{\Pi}_i.$$

For equation (16), we note that $\boldsymbol{A}_{0i}$ is invertible for any i, therefore multiplying both sides of (16) by $\boldsymbol{A}_{0i}^{-1}$ yields

$$\boldsymbol{z}_{it} = \bar{\boldsymbol{B}}_i\boldsymbol{d}_t + \sum_{l=1}^{L} \bar{\boldsymbol{A}}_{i,l}\boldsymbol{z}_{it-l} + \boldsymbol{A}_{0i}^{-1}\boldsymbol{C}_i\boldsymbol{f}_t + \boldsymbol{e}_{zit}, \tag{17}$$

where $\bar{\boldsymbol{B}}_i = \boldsymbol{A}_{0i}^{-1}\boldsymbol{B}_i$, $\bar{\boldsymbol{A}}_{i,l} = \boldsymbol{A}_{0i}^{-1}\boldsymbol{A}_{i,l}$ and $\boldsymbol{e}_{zit} = \boldsymbol{A}_{0i}^{-1}\boldsymbol{e}_{it}$. Equation (17) is a reduced form VAR(L) model of $\boldsymbol{z}_{it}$. We make the following assumption regarding the invertibility of $\bar{A}_{i,l}$.

Assumption 3 *All of the roots of* $\boldsymbol{I}_{k+1} - \bar{\boldsymbol{A}}_{i,1}z - \bar{\boldsymbol{A}}_{i,2}z^2 - \cdots - \bar{\boldsymbol{A}}_{i,L}z^L = 0$ *lie outside the unit circle* $i = 1, \ldots, N$.

Remark The above assumption is called stability condition and under which $\boldsymbol{z}_{it}$ is stable (Lukepohl, 2005).

Define $\bar{\boldsymbol{A}}_i(\mathbb{L})$ as $\bar{\boldsymbol{A}}_i(\mathbb{L}) = \boldsymbol{I}_{k+1} - \bar{\boldsymbol{A}}_{i,1}\mathbb{L} - \bar{\boldsymbol{A}}_{i,2}\mathbb{L}^2 - \cdots - \bar{\boldsymbol{A}}_{i,L}\mathbb{L}^L$, where $\mathbb{L}$ being the lag operator with $\mathbb{L}y_t = y_{t-1}$. Under Assumption 3, there exists $\Phi_i(\mathbb{L}) = \sum_{l=0}^{\infty}\Phi_{il}\mathbb{L}^l$ such that $\Phi_i(\mathbb{L})\bar{\boldsymbol{A}}_i(\mathbb{L}) = \boldsymbol{I}_{k+1}$ and the coefficient matrices of $\Phi_i(\mathbb{L})$ are absolutely summable (Lukepohl, 2005). Consequently, we have

$$\boldsymbol{z}_{it} = \sum_{l=0}^{\infty}\Phi_{il}\mathbb{L}^l\left(\bar{\boldsymbol{B}}_i\boldsymbol{d}_t + \boldsymbol{A}_{0i}^{-1}\boldsymbol{C}_i\boldsymbol{f}_t + \boldsymbol{e}_{zit}\right), \tag{18}$$

for $i = 1, 2, \ldots, N$. Taking cross-section averages of the above and making use of the fact that the elements of $\boldsymbol{e}_{zit}$ are weakly cross-sectionally dependent, [6] we have

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{l=0}^{\infty}\Phi_{il}\mathbb{L}^l\boldsymbol{e}_{zit} = O_p\left(N^{-1/2}\right).$$

Furthermore, we can obtain

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N}\sum_{l=0}^{\infty}\Phi_{il}\mathbb{L}^l\boldsymbol{A}_{0i}^{-1}\boldsymbol{C}_i\boldsymbol{f}_t &= \sum_{l=0}^{\infty}\frac{1}{N}\sum_{i=1}^{N}\Phi_{il}\mathbb{L}^l\boldsymbol{A}_{0i}^{-1}\boldsymbol{C}_i\boldsymbol{f}_t \\
&= \sum_{l=0}^{\infty}\mathbb{L}^l E\left(\Phi_{il}\boldsymbol{A}_{0i}^{-1}\boldsymbol{C}_i\right)\boldsymbol{f}_t + O_p\left(N^{-1/2}\right) \\
&= \Lambda(\mathbb{L})\boldsymbol{C}\boldsymbol{f}_t + O_p\left(N^{-1/2}\right),
\end{aligned}$$

where the inverse of polynomial $\Lambda(\mathbb{L}) = \sum_{l=0}^{\infty}\Lambda_l\mathbb{L}^l$ with $\Lambda_l = E\left(\Phi_{il}\boldsymbol{A}_{0i}^{-1}\right)$ exists, and $\boldsymbol{C} = E(\boldsymbol{C}_i)$. Consequently, for cross-section averages of (18), we obtain

$$\bar{\boldsymbol{z}}_t = \bar{\bar{\boldsymbol{B}}}\boldsymbol{d}_t + \Lambda(\mathbb{L})\boldsymbol{C}\boldsymbol{f}_t + O_p\left(N^{-1/2}\right), \tag{19}$$

where $\bar{\boldsymbol{z}}_t = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{z}_{it}$ and $\bar{\bar{\boldsymbol{B}}} = \frac{1}{N}\sum_{i=1}^{N}\sum_{l=0}^{\infty}\Phi_{il}\bar{\boldsymbol{B}}_i$. As established by Assumption 6 of

[6] Pesaran (2006) discusses the more general case of cross-section weighted averages, which could be readily adapted to our case. However, to simplify the exposition we restrict our attention to simple averages throughout. Furthermore, the asymptotic variance matrix for slope coefficients $\left(\boldsymbol{\psi}_i', \boldsymbol{\beta}_i'\right)'$ does not depend on the weights.

Chudik and Pesaran (2015), we make the following assumption to ensure the invertibility of $\boldsymbol{C}'\boldsymbol{C}$ and consequently the viable estimation of unit-specific coefficients.

Assumption 4 *$(k+1) \times m$ dimensional matrix $\boldsymbol{C} = (\boldsymbol{\gamma}, \boldsymbol{\Gamma})'$ has full column rank.*

Rearranging (19) for the expression of $\boldsymbol{f}_t$, we have

$$\boldsymbol{f}_t = G(\mathbb{L})\left(\bar{\boldsymbol{z}}_t - \bar{\overline{\boldsymbol{B}}}\boldsymbol{d}_t\right) + O_p\left(N^{-1/2}\right), \tag{20}$$

where $G(\mathbb{L}) = \left(\boldsymbol{C}'\boldsymbol{C}\right)^{-1}\boldsymbol{C}'\Lambda(\mathbb{L})^{-1}$. Substituting (20) into (15) yields

$$y_{it} = \boldsymbol{\alpha}_i^{*\prime}\boldsymbol{d}_t + \boldsymbol{\psi}_i'\boldsymbol{y}_{it,-L} + \boldsymbol{\beta}_i'\boldsymbol{x}_{it} + \boldsymbol{\delta}_i'(\mathbb{L})\,\bar{\boldsymbol{z}}_t + \epsilon_{it} + O_p\left(N^{-1/2}\right), \tag{21}$$

where $\boldsymbol{\alpha}_i^* = \boldsymbol{\alpha}_i - \bar{\overline{\boldsymbol{B}}}'\boldsymbol{\delta}_i(\mathbb{L})$ and $\boldsymbol{\delta}_i(\mathbb{L}) = \sum_{l=0}^{\infty}\boldsymbol{\delta}_{il}\mathbb{L}^l = G(\mathbb{L})'\boldsymbol{\gamma}_i$.

Since $\boldsymbol{\delta}_i'(\mathbb{L})\,\bar{\boldsymbol{z}}_t$ contains an infinite order of lags of $\bar{\boldsymbol{z}}_t$, one can practically use p_T lags in the estimation (assuming $p_T > L$). The following augmented regression is thus obtained

$$y_{it} = \boldsymbol{\alpha}_i^{*\prime}\boldsymbol{d}_t + \boldsymbol{\psi}_i'\boldsymbol{y}_{it,-L} + \boldsymbol{\beta}_i'\boldsymbol{x}_{it} + \sum_{l=0}^{p_T}\boldsymbol{\delta}_{il}'\bar{\boldsymbol{z}}_{t-l} + e_{y,it}, \tag{22}$$

where $\boldsymbol{\alpha}_i^*$ and $\boldsymbol{\delta}_{il}$ $(l = 0, 1, \ldots, p_T)$ are nuisance parameters. The composite errors $e_{y,it}$ consist of three components: an idiosyncratic term, ϵ_{it}, an error component due to the truncation of infinite polynomial distributed lag function at p_T, $\boldsymbol{\delta}_i'(\mathbb{L})$, and an error component due to the approximation of unobserved common factors, specifically

$$e_{y,it} = \epsilon_{it} + \sum_{l=p_T+1}^{\infty}\boldsymbol{\delta}_{il}'\bar{\boldsymbol{z}}_{t-l} + O_p\left(N^{-1/2}\right).$$

The parameter of interests are $\boldsymbol{\theta}_i = \left(\boldsymbol{\psi}_i', \boldsymbol{\beta}_i'\right)'$, which can be estimated by the least squares estimation based on the cross-sectionally augmented regression (22). Define

$$\Xi_i = \begin{pmatrix} \boldsymbol{y}'_{ip_T+1,-L} & \boldsymbol{x}'_{i,p_T+1} \\ \boldsymbol{y}'_{ip_T+2,-L} & \boldsymbol{x}'_{i,p_T+2} \\ \vdots & \vdots \\ \boldsymbol{y}'_{iT,-L} & \boldsymbol{x}'_{i,T} \end{pmatrix}, \bar{\boldsymbol{Q}} = \begin{pmatrix} \boldsymbol{d}'_{p_T+1} & \bar{\boldsymbol{z}}'_{p_T+1} & \bar{\boldsymbol{z}}'_{p_T} & \cdots & \bar{\boldsymbol{z}}'_1 \\ \boldsymbol{d}'_{p_T+2} & \bar{\boldsymbol{z}}'_{p_T+2} & \bar{\boldsymbol{z}}'_{p_T+1} & \cdots & \bar{\boldsymbol{z}}'_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{d}'_T & \bar{\boldsymbol{z}}'_T & \bar{\boldsymbol{z}}'_{T-1} & \cdots & \bar{\boldsymbol{z}}'_{T-p_T} \end{pmatrix}, \quad (23)$$

and the projection matrix $\boldsymbol{M}_{\bar{Q}} = \boldsymbol{I}_{T-p_T} - \bar{\boldsymbol{Q}}\left(\bar{\boldsymbol{Q}}'\bar{\boldsymbol{Q}}\right)^{+}\bar{\boldsymbol{Q}}'$, where $\boldsymbol{I}_{T-p_T}$ is a $(T-p_T)\times(T-p_T)$ dimensional identity matrix, and $\boldsymbol{A}^{+}$ represents the Moore-Penrose generalized inverse of $\boldsymbol{A}$. Based on (22), the CCE estimator of $\boldsymbol{\theta}_i$ based on (22) is given by

$$\hat{\boldsymbol{\theta}}_{i,CCE} = \left(\Xi_i'\boldsymbol{M}_{\bar{Q}}\Xi_i\right)^{-1}\Xi_i'\boldsymbol{M}_{\bar{Q}}\boldsymbol{y}_i, \quad (24)$$

where $\boldsymbol{y}_i = \left(y_{i,p_T+1}, y_{i,p_T+2}, \ldots, y_{i,T}\right)'$. We manage to establish the consistency of CCE estimates for the HARP model in the ensuing theorem.

Theorem 1 *Suppose y_{it} for $i = 1,2,\ldots,N$ and $t = 1,2,\ldots,T$, are generated following the HARP models (3), (12)-(14). Under the Assumptions 1-4, as $(N,T)\to\infty$ and $p_T = O\left(T^{1/3}\right)$, we have*

$$\hat{\boldsymbol{\theta}}_{i,CCE} \to_p \boldsymbol{\theta}_i,$$

where $\hat{\boldsymbol{\theta}}_{i,CCE} = \left(\hat{\boldsymbol{\psi}}_i', \hat{\boldsymbol{\beta}}_i'\right)'$ is described by (24).

Remark The proof of Theorem 1 is provided in Appendix A. Since the parameters of interest $\phi_i^{(l)}$ in model (10) are linear transformations of the first coefficient vector in $\boldsymbol{\theta}_{i,CCE}$, based on the above theorem and the Slutsky's theorem, we can claim that the CCE estimator $\phi_i^{(l)}$ is consistent as $T\to\infty$ for $i = 1,\ldots,N$ and $l = 1,\ldots,L$. In Appendix B, we conduct Monte Carlo simulation to investigate the general performance of the CCE estimator under the HARP framework.

3.2 Forecasting realized volatility

We are interested in the forecast of $y_{i,T+h}$ conditional the information up to time T. Without loss of generality, y_{it} can also be described by the following direct forecasting model[7]

$$y_{i,t+h} = \boldsymbol{\alpha}_i' \boldsymbol{d}_t + \boldsymbol{\psi}_i' \boldsymbol{y}_{it,-L} + \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + \gamma_i' \boldsymbol{f}_{t+h} + \epsilon_{it}. \tag{25}$$

The forecast of $y_{i,T+h}$ contingent on the information up to time T is therefore

$$\hat{y}_{i,T+h|T} = \hat{\boldsymbol{\alpha}}_i' \boldsymbol{d}_T + \hat{\boldsymbol{\psi}}_i' \boldsymbol{y}_{iT,-L} + \hat{\boldsymbol{\beta}}_i' \boldsymbol{x}_{iT} + \hat{\gamma}_i' \hat{\boldsymbol{f}}_{T+h|T}, \tag{26}$$

where $\hat{\boldsymbol{\alpha}}_i$, $\hat{\psi}_i$, $\hat{\boldsymbol{\beta}}_i$, $\hat{\gamma}_i$ are the estimates from (10) and (12), and $\hat{\boldsymbol{f}}_{T+h|T}$ is a forecast of $\boldsymbol{f}_{T+h}$.

The forecast requires consistent estimation of the unobserved factors and hence the consistent estimation of parameters. Since the consistency of slope coefficients $\hat{\boldsymbol{\theta}}_{i,CCE}$ has been established in Section 3.1 without a prior knowledge of $\boldsymbol{f}_t$, we solve the forecasting problem by employing a two stage process: first, we estimate the parameters $\left(\boldsymbol{\psi}_i', \boldsymbol{\beta}_i'\right)'$ and unobserved common factors $\boldsymbol{f_t}$, utilizing the CCE estimator and principal components. The factor estimates from this procedure are consistent (Pesaran, 2006). In the second stage, the factor estimates can be used directly to estimate the parameters $\boldsymbol{\alpha}_i$, ψ_i, $\boldsymbol{\beta}_i$ and $\boldsymbol{\gamma_i}$ in (25) by OLS. We then obtain forecasts of $\boldsymbol{f_{t+h}}$ under hypothetical processes of $\boldsymbol{f_t}$.[8]

Given the consistent estimation of $\boldsymbol{\psi_i}$ and $\boldsymbol{\beta}_i$, we can obtain $g_{it} = \boldsymbol{\alpha}_i' \boldsymbol{d}_{t-h} + u_{it}$, as

$$\hat{g}_{it} = y_{it} - \hat{\boldsymbol{\psi}}_i' \boldsymbol{y}_{it-h,-L} - \hat{\boldsymbol{\beta}}_i' \boldsymbol{x}_{it-h}. \tag{27}$$

After acquiring the residuals, $\hat{g}_{it}$, an estimate of u_{it} is produced by integrating out the

[7]Assuming certain vector autoregressive processes for $\boldsymbol{d}_t$ and $\boldsymbol{f}_t$, we can iterate on (10) and (12), and then conduct recursive substitutions to yield the direct forecasting model. Bernoth and Pick (2009) provide a detailed mathematical induction on a similar question.

[8]Section 5.1 provides a detailed discussion on various forecasting schemes of $\boldsymbol{f_t}$ and their implications.

common observed factors, $\boldsymbol{d}_{t-h}$,

$$\hat{\boldsymbol{u}}_{\boldsymbol{i}} = \boldsymbol{M}_D \hat{\boldsymbol{g}}_{\boldsymbol{i}}, \tag{28}$$

where $\hat{\boldsymbol{u}}_{\boldsymbol{i}} = (\hat{u}_{i,h+1}, \hat{u}_{i,h+2}, \ldots, \hat{u}_{i,T})$, $\boldsymbol{M}_D = \boldsymbol{I} - \boldsymbol{D}\left(\boldsymbol{D}'\boldsymbol{D}\right)^{-1}\boldsymbol{D}'$ and $\boldsymbol{D} = \left(\boldsymbol{d_1}', \boldsymbol{d_2}', \ldots, \boldsymbol{d_{T-h}}'\right)$. Bernoth and Pick (2011) pointed out that the orthogonality assumption of $\boldsymbol{d}_t$ to $\boldsymbol{f}_t$ is necessary to guarantee unbiasedness of the parameter estimates of the common factors, $\hat{\boldsymbol{\alpha}}_i$, although an absence of it will not bias the forecasts resulting from the above procedure.

With estimates of the residuals $\hat{u}_{it}$, we are able to extract the unobserved common factors, $\boldsymbol{f}_t$, using the principal component analysis. In our application, consistent estimation of ψ_i and $\boldsymbol{\beta}_i$ is guaranteed under any fixed number of unobserved factors, m.[9] However, since $\boldsymbol{f}_t$ and its factor loadings still matter for the forecasts by equation (25), an estimate of m seems essential. To resolve this, we apply the Bai and Ng's (2002) method to the residuals, $\hat{\boldsymbol{u}}_{\boldsymbol{i}}$ given in (28), which yields a choice of $m = 2$.[10]

For forecasting $y_{i,t+h}$, a prediction of $\boldsymbol{f}_{t+h}$ is required, since the principal component analysis only generates $\hat{\boldsymbol{f}}_t$ up to time T. Based on Pesaran, Pick, and Allan (2011) and Stock and Watson (2002), we assumes a VAR model of $\hat{\boldsymbol{f}}_t$ to predict $\hat{\boldsymbol{f}}_{t+h}$, where the number of lags is selected by the Bayes Information Criterion (BIC). We also use the standard HAR model and the random forest method to forecast $\hat{\boldsymbol{f}}_t$.[11]

3.3 Assessment of the forecasts

Forecast performance of y_{it} is evaluated using the following criteria:

$$\text{MAFE}(h) = \frac{1}{M}\sum_{j=1}^{M} |e_{iT_j,h}|, \tag{29}$$

[9]As suggested by Pesaran (2006), the number of unobserved factors, m, only becomes a practical issue if the focus of the analysis is on the factor loadings, for instance, the parameters of asset pricing factors.

[10]We also test the forecasts for different values of m and the results remain qualitatively similar.

[11]A detailed description of the random forest technique is provided in Appendix C.

$$\text{MSFE}(h) = \frac{1}{M}\sum_{j=1}^{M} e_{iT_j,h}^2, \tag{30}$$

$$\text{SDFE}(h) = \sqrt{\frac{1}{M-1}\left(e_{iT_j,h} - \frac{1}{M}\sum_{j=1}^{M} e_{iT_j,h}\right)^2}, \tag{31}$$

where $e_{iT_j,h} = y_{iT_j,h} - \hat{y}_{iT_j,h}$ is the forecast error, $j = 1, 2, \ldots, M$, and $\hat{y}_{iT_j,h}$ is the h-day ahead forecast with information up to T_j, where T_j stands for the last observation in each of the M rolling windows. Another widely-adopted method for evaluation is by means of the R^2-criterion of the Mincer-Zarnowitz regression,[12] given by

$$y_{iT_j,h} = a + b\hat{y}_{iT_j,h} + u_{T_j}, \text{for } j = 1, 2, \ldots, M, \tag{32}$$

Note that we choose the level-regression (32) over the log-regression, because Hansen and Lunde (2006) have argued that the R^2 from a regression of $\log(y_{iT_j,h})$ on a constant and $\log(\hat{y}_{iT_j,h})$ is unlikely to induce the same ranking of volatility models as the R^2 from the infeasible regressions (with the true volatility), unless a proportionate relationship exists between the estimated and true values of volatility.[13]

Based on the findings of Hansen and Lunde (2006) and Patton (2011), we also compute the expected values of gaussian quasi-likelihood (QLIKE),

$$\text{QLIKE}(h) = \log \hat{y}_{iT_j,h} + \frac{y_{iT_j,h}}{\hat{y}_{iT_j,h}}, \text{for } j = 1, 2, \ldots, M, \tag{33}$$

The QLIKE function, along with the MSFE loss, has been demonstrated to be robust to noise in the proxy for volatility in Patton (2011). Moreover, Patton and Sheppard (2009) find that relative to the MSFE loss, the QLIKE loss has better power properties under the

[12] Interested readers may refer to Mincer and Zarnowitz (1969) for more details.

[13] Hansen and Lunde (2006) prove that if the proxy, $\tilde{\sigma}_t^2$, and the true volatility, σ_t^2 , satisfy the equation $\tilde{\sigma}_t^2 = (1 - v_t)\sigma_t^2$ for some random variable, v_t, the ranking of volatility models remains unaffected by the measurement errors of $\tilde{\sigma}_t^2$.

Diebold-Mariano test. In the last place, we complement the above results by running the unconditional Giacomini and White (2006) test for the mean absolute forecast error, in order to test the equal predictive ability of a pair of models.

4 Data descriptions

In the empirical exercise, we consider an application to the realized variance of the NASDAQ 100 Trust exchange-traded fund (ETF) tracking the NASDAQ 100 Index, with ticker symbol QQQ. To avoid the concerns of data mining, the information on unobserved common factors is extracted solely from a panel data of the NASDAQ 100 constituents.

The data on the NASDAQ 100 ETF consist of high frequency transaction prices from May 22, 2007 to October 20, 2017, which totals 2,625 observations. The 104 constituents cover industry groups ranging from computer hardware and software, telecommunications, retail/wholesale trade, and biotechnology. Since the components of the NASDAQ 100 index are varying over time, we only include the stocks that always belong to the index during our sample period, in order to keep the panel balanced. In total, 89 stocks are selected. A more detailed description of these stocks is given in Appendix D, including their ticker symbols, names, and Global Industry Classification (GIC) sectors. The whole data set and its relevant information are provided by Pitrading Inc., which base their source information from New York Stock Exchange's TAQ database. All the above data are obtained at one-minute increments. The intraday prices are then used to calculate daily realized variance measures by equation (1).

To have an abundance of time series, we use the original data at one-minute intervals in our primary analysis. However, we are aware that a too high sampling frequency might cause microstructure bias to distort volatility estimates from its true daily variance. The previous work by Liu, Patton, and Sheppard (2015) offers some evidence on why 5-minute RV is typically considered as the benchmark.[14] For the justification of using the 1-minute data, we construct volatility signature plots for the NASDAQ 100 ETF (QQQ) and the eight representative stocks in Figure 1, to check if 1-minute RV is an appropriate alternative to 5-minute RV.[15] A description of the full company names and their weights is included in Table 1. The main pattern is that volatility signature plots for all the considered assets are flat, which is especially the case for QQQ. This indicates no apparent variations from 1-minute sampling, relative to 5-minute sampling, at least for our data sets.[16]

Table 1: Descriptions of liquid and illiquid stocks

Ticker	Company Name	Weights in the NASDAQ 100 (%)
Liquid Stocks		
AAPL	Apple Inc.	12.39
FB	Facebook Inc.	4.85
INTC	Intel Corporation	2.60
NVDA	NVIDIA Corporation	1.80
Iliquid Stocks		
CTAS	Cintas Corporation	0.27
IDXX	IDEXX Laboratories Inc.	0.25
ISRG	Intuitive Surgical Inc.	0.70
JBHT	J.B. Hunt Transport Services Inc.	0.16

We describe summary statistics of the RV series for the NASDAQ100 ETF in column 2 of Table 2. Due to the vast number of the NASDAQ 100 Index components, we only present the statistics of six representative stocks in columns 3 to 8 of Table 2. Table 2 documents the results of the sample mean, median, minimum, maximum, standard deviation, skewness, and kurtosis for the RV series over the full sample periods. Table 2 also reports the

[14] Liu, Patton, and Sheppard (2015) conduct a comprehensive study for over 400 different realized measures, with a wide range of sampling frequencies, and they apply these to 11 years of daily data on 31 individual financial assets. Overall, they find it difficult to significantly outperform 5-minute RV.

[15] The volatility signature plot was first introduced by Andersen, Bollerslev, Diebold, and Labys (2000) to provide some guidance on the optimal sampling frequency.

[16] Liu, Patton, and Sheppard (2015) conclude that as long as the assets are liquid enough, one-minute sampling are nevertheless sparse to avoid the problem of microstructure bias. For the robustness check, we also conduct a similar empirical analysis on five-minute sampling data in Section 6.2 and the HARP model still outperforms the risk models solely based on its own realized volatility components.

p-values[17] of the Jarque-Bera test for normality and those of the augmented Dickey-Fuller (ADF) test for unit root. The null hypotheses of a normal distribution and a unit root are strongly rejected in all cases, whereas the other statistics disperse over a wide range.

Table 2: Summary statistics for the RV of the NASDAQ100 ETF and 6 representative stocks

Statistic	QQQ	Tickers of representative stocks					
		AAL	ALXN	DISCA	ISRG	QCOM	XLNX
Mean	1.0536	17.0766	4.7820	3.5654	4.1525	2.3330	2.7815
Median	0.6119	7.2913	3.4835	2.0676	2.6538	1.4461	1.7512
Maximum	9.8061	669.3273	135.1100	583.1942	65.4187	34.2651	33.9595
Minimum	0.0687	0.5071	0.4013	0.4374	0.3030	0.1598	0.3030
Standard Deviation	1.2467	30.9578	5.0603	13.5319	4.6312	2.7424	2.9660
Skewness	3.0783	7.8892	9.2852	35.4074	4.0143	4.2497	3.5105
Kurtosis	15.1628	117.0782	189.5792	1416.5222	29.7366	31.5784	21.4986
Jarque-Bera	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010
ADF	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010

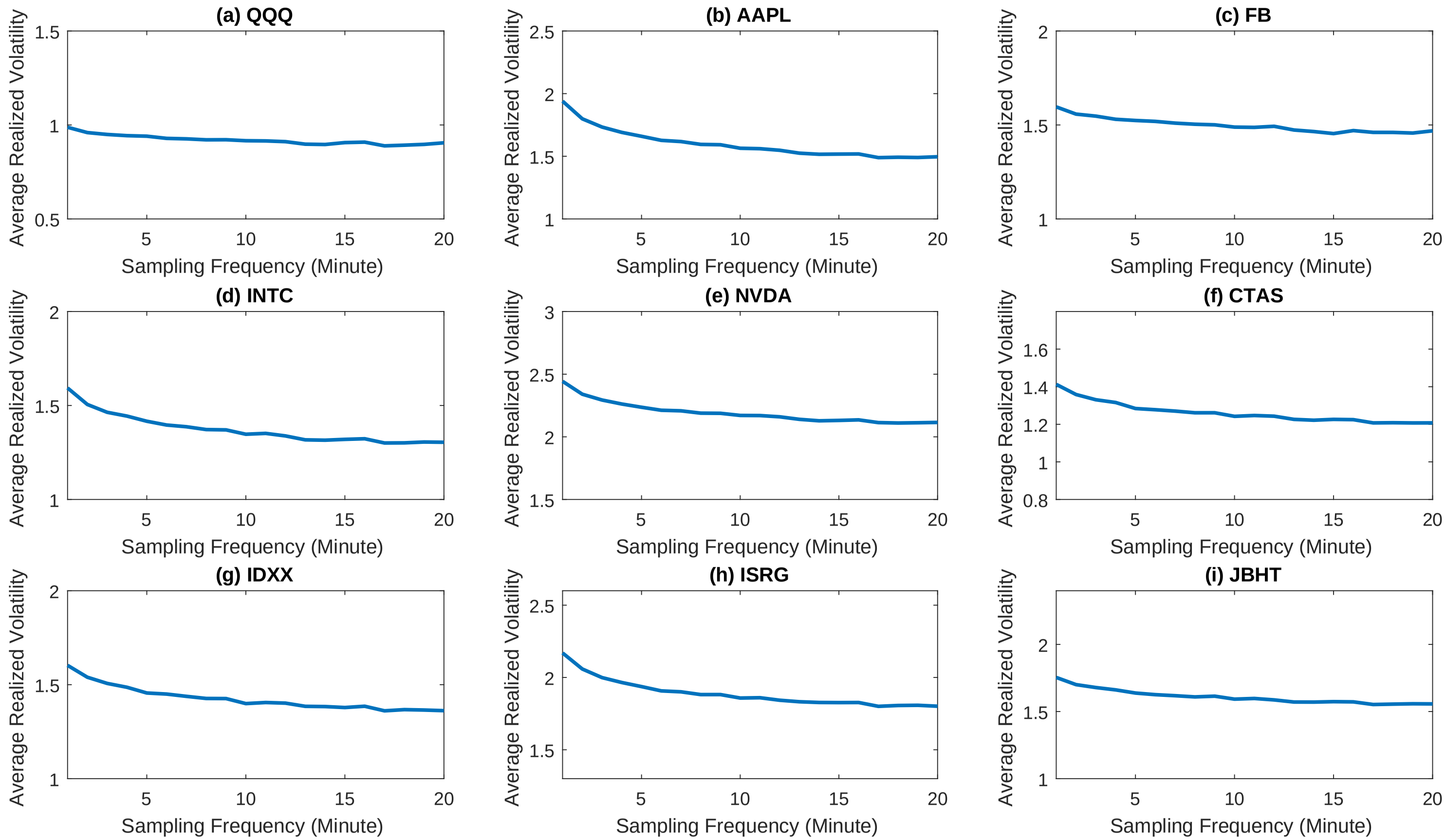
The second column of Table 2 contains statistics for the realized variance of the NASDAQ100 ETF from May 22, 2007 to October 20, 2017, for a total of 2,625 observations. The statistics of six component stocks of the NASDAQ 100 index are presented in columns 3 to 8. For JB and ADF test statistics that are outside tabulated critical values, we report maximum (0.999) or minimum (0.001) p-values. All the statistics here are computed based on the data at one-minute increments.

4.1 Observed common factors

Note that the main forecast equation (10) contains a set of observed macroeconomic factors $\boldsymbol{d}_t$. Inspired by recent research by Fernandes, Medeiros, and Scharth (2014), we include the following predictors in $\boldsymbol{d}_t$ contemporaneously: (i) the j-day continuously compounded return on the one-month crude oil futures contract for $j = 1, 5, 10, 22, 66$ (oil j-day return); (ii) the first difference of the logarithm of the trade-weighted average of the foreign exchange value of the US dollar index against the Australian dollar, Canadian dollar, Swiss franc, euro, British pound sterling, Japanese yen, and Swedish krona (USD change); (iii) the excess yield of the Moody's seasoned Baa corporate bond over the Moody's seasoned Aaa corporate bond (credit spread); (iv) the difference between the 10-Year and 3-month Treasury constant maturity rates (term spread); (v) and the difference between

[17] In our exercises, we set the lower bound of the p-values of the Jarque-Bera and the ADF tests at 0.001. Values less than 0.001 are truncated at 0.001.

Figure 1: Volatility signature plot for representative stocks and the NASDAQ 100 ETF



Panels (a) to (i) plot the averaged daily realized volatility (on the y-axis) for the NASDAQ 100 ETF, four highly liquid components (AAPL, FB, INTC, and NVDA) and four relatively iliquid components (CTAS, IDXX, ISRG, and JBHT) of the NASDAQ 100 Index over their corresponding sampling intervals (on the x-axis). The sampling intervals run from 1 minute to 20 minutes.

the effective and target federal fund rates (FF deviation). While both oil prices (Oil) and term spread (TS) are concerned with different dimensions of the overall market conditions, USD change (USDI) and FF deviation (FED) are both linked to US macroeconomic conditions. Descriptive statistics for $\boldsymbol{d}_t$ are available in Appendix E.

5 Empirical analysis

In this section, we conduct an empirical exercise to thoroughly examine both the in-sample and the out-of-sample performance of the HARP model. This is in comparison with the performance of the autoregressive model (AR) and a battery of HAR-type models reviewed in Section 2. The rivalry models are listed as follows:

(i) AR model: the simple autoregression model AR(22);

(ii) HAR model: defined in (2);

(iii) HAR-J model: defined in (4);

(iv) HAR-CJ model: defined in (5);

(v) HAR-RS-I model: defined in (6);

(vi) HAR-RS-II model: defined in (7);

(vii) HAR-SJ-I model: defined in (8);

(viii) HAR-SJ-II model: defined in (9).

In the first instance, we need to verify the existence of the possible error cross-sectional dependence. In Table 3 we report the cross-section dependence (CD) test of Pesaran (2004, 2015) and their p-values, which are based on the average of pair-wise correlations of the residuals from regressions of individual stock volatility. For all regressions and forecast horizons, these error terms show a considerable degree of cross-sectional dependence.[18] This implies that even after controlling for the predictors of volatility in our data set, sizable cross-section dependance remains across component stocks, which supports the utilization of this information to improve the accuracy of forecasts.

Table 3: Results for cross-section dependence test for RV of the NASDAQ 100 constituents

Method	$h = 1$		$h = 5$		$h = 10$		$h = 22$	
	CD	p-value	CD	p-value	CD	p-value	CD	p-value
RW	5386.1615	0.0000	6661.1541	0.0000	6844.1628	0.0000	7426.2886	0.0000
AR(22)	4178.6508	0.0000	5563.3119	0.0000	6031.7421	0.0000	6814.7767	0.0000
HAR	4491.6446	0.0000	5719.2537	0.0000	6167.5320	0.0000	6957.3473	0.0000
HAR-J	4384.6859	0.0000	5655.8968	0.0000	6107.9819	0.0000	6895.6076	0.0000
HAR-CJ	4251.0679	0.0000	5474.3930	0.0000	5880.9295	0.0000	6596.3800	0.0000
HAR-RS-I	4386.1947	0.0000	5641.6365	0.0000	6118.8177	0.0000	6919.7391	0.0000
HAR-RS-II	4179.7576	0.0000	5530.1978	0.0000	6063.0118	0.0000	6884.2000	0.0000
HAR-SJ-I	4403.8828	0.0000	5646.2349	0.0000	6132.6288	0.0000	6926.7192	0.0000
HAR-SJ-II	4351.2087	0.0000	5611.6945	0.0000	6103.8054	0.0000	6900.0884	0.0000

CD is short for the CD test statistic of cross-section independence applied to the residuals of the asset specific regression. RW indicates a random walk model. A large CD indicates that the residuals of certain model estimation are correlated across individual component stocks.

With an explicit control of error cross-sectional dependence, we implement the HARP model built upon the specification of HAR-RS-II, due to its sound out-of-sample performance documented in Patton and Sheppard (2015). This particularly implies that unit-specific regressors $\boldsymbol{x}_{it}$ are $(\mathrm{RV}_{it}^{(1)} \cdot \mathbb{I}(r_t < 0), \mathrm{RS}_{it}^{+}, \mathrm{RS}_{it}^{-})'$, while other regressors in (10) are well defined above.[19] This is our benchmark forecasting specification in the sequel. In the following analysis, we mostly concentrate on forecasting the RV of the NASDAQ 100 ETF, where the information of unobserved common effects are extracted from the volatility of the NASDAQ 100 components. We want to see if the comovements of component stocks genuinely assist in predicting the index fund volatility.

For all of the exercises, we conduct an in-sample exercise with the full sample data and a rolling window out-of-sample exercise. The window length is set at 1000. We also test

[18] Under the null of weak error cross-sectional dependence, the CD statistics are asymptotically distributed as $N(0, 1)$.

[19] The lagged dependent variables $\bar{y}_{it}^{(l)}$ include $RV_{it}^{(5)}$ and $RV_{it}^{(22)}$, while $\boldsymbol{d}_t$ are observed macroeconomic factors, of which the details are explained in Section 4.1.

other values of the window length and the results remain similar.[20] Each of the above candidate models is applied to the data set, and a series of h days ahead forecasts are obtained.[21] We consider both short-horizon and long-horizon forecasts with $h = 1, 5, 10$ and 22. For comparing the forecast performance, we specifically employ the five statistics in Section 3.3: (i) the mean squared forecast error (MSFE); (ii) the standard deviation of forecast error (SDFE); (iii) the mean absolute forecast error (MAFE); (iv) the Mincer-Zarnowitz pseudo-R^2 and (v) the QLIKE function for each model at each forecast horizon.

5.1 The benchmark scheme of predicting $\hat{f}_{t+h}$

Next, we need to choose a benchmark scheme for forecasting $\hat{f}_{t+h}$. In order to accomplish this, an investigation on the statistical properties of f_t is necessary. We start by estimating f_t based on the HARP specification and the CCE estimator, across various forecast horizons ($h = 1, 5, 10, 22$). We report summary statistics of f_t in Table 4.[22] As can be seen clearly, the null hypotheses of a normal distribution and a unit root process are strongly rejected for the two principal components (f_1 and f_2). Except for f_2 at $h = 1$, the factors have highly persistent autocorrelation structures in other cases.

In Section 3.2, we consider four methods of forecasting $\hat{f}_{t+h}$: the AR model, the HAR model, the VAR model,[23] and the random forest method. With the HARP model and its benchmark specification (i.e., HAR-RS-II), the relative out-of-sample performance from the above three ways of forecasting $\hat{f}_{t+h}$ is compared in Table 5. Looking across the columns, we see that the random forest method performs the very best overall, followed by the

[20] We test the results with the window length of 500 observations. Tables are provided in Appendix G.2.

[21] Note that, to compute h-day ahead forecasts, we employ a direct forecast approach in which we use RV_{t+h} in the above models. This approach permits us to produce multi-step ahead forecasts without imposing any assumption about future realizations on the explanatory variables.

[22] The sample autocorrelation functions of f_t is reported in Figure A2 of Appendix F.

[23] We also tried to apply VARMA model to forecast f_{t+h}. However, the VARMA model requires the data to be highly stationary. Since we are using a rolling window exercise, this implies that f_t needs to be highly stationary in every roll, which is not guaranteed in practice. Hence, we use the VAR model instead.

Table 4: Summary statistics for the unobserved common factors

Statistic	$h=1$		$h=5$		$h=10$		$h=22$	
	f_1	f_2	f_1	f_2	f_1	f_2	f_1	f_2
Mean	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Median	0.2038	0.1382	0.2290	0.1189	0.2524	-0.0393	-0.2678	0.0697
Maximum	14.1083	7.2167	5.4090	20.6654	5.0550	12.5781	12.3710	15.2257
Minimum	-14.1560	-26.5783	-18.9432	-21.5940	-15.4713	-15.9108	-6.7179	-13.2345
Standard Deviation	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Skewness	-2.5078	-11.2995	-6.4684	-0.8306	-5.0666	-2.0137	4.0513	0.7770
Kurtosis	56.8944	252.7725	88.6692	189.6395	54.0926	64.9030	35.8895	49.9837
Jarque-Bera	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010
ADF Test	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010

Table 4 contains statistics for the unobserved common factors computed based on the CCE estimator and applied to RVs of the NASDAQ 100 constituents. For each forecast horizon ($h = 1, 5, 10, 22.$), we extract two principal components. For JB and ADF test statistics that are outside tabulated critical values, we report maximum (0.999) or minimum (0.001) p-values.

HAR model. The Giacomini-White (GW) tests based on the mean absolute forecast errors, reported in Table 6, again corroborate these conclusions. Now taking the random forest method as the benchmark, the pairwise tests show that the forecasts from this method result in significantly lower errors than the forecasts from all the other methods, except for the forecast horizon of $h = 1$. Based on the above set of analyses, we decide to employ the random forest method as the benchmark scheme, to predict $\hat{f}_{t+h}$ in the following empirical exercises.

5.2 The in-sample and out-of-sample results for the NASDAQ 100 ETF

We begin our discussions by considering the in-sample results in Table 7, where the HARP model is compared with its baseline specification without unobserved common factors, the HAR-RS-II model. The race is evaluated by the in-sample predictive R^2's. We find that, in the HARP model with f_t, the coefficients on negative realized semivariance decrease substantially in magnitude but remain significant for all horizons, and the effect is much milder for the coefficients on positive realized semivariance. Akin to a substitution effect, the coefficients on all f_t are large and significant. Moreover, the HARP model explains 32% (at $h = 1$) to 174% (at $h = 22$) more of the variation in future volatility than the

Table 5: The out-of-sample forecast comparison for different ways of forecasting $\hat{f}_{t+h}$

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
$HARP_{AR}$	0.2190	0.1330	0.2753	0.4680	0.7767
$HARP_{HAR}$	0.2160	**0.1307**	0.2722	0.4647	0.7799
$HARP_{VAR}$	0.2159	0.1375	0.2775	0.4646	0.7800
$HARP_{RF}$	**0.2114**	0.1339	**0.2698**	**0.4598**	**0.7845**
Panel B: $h = 5$					
$HARP_{AR}$	0.3666	0.2085	0.3758	0.6055	0.6272
$HARP_{HAR}$	0.3622	0.2072	0.3724	0.6018	0.6316
$HARP_{VAR}$	0.3564	0.1956	0.3604	0.5970	0.6375
$HARP_{RF}$	**0.3433**	**0.1931**	**0.3409**	**0.5859**	**0.6509**
Panel C: $h = 10$					
$HARP_{AR}$	0.4075	**0.2398**	0.4024	0.6384	0.5862
$HARP_{HAR}$	0.4075	0.2449	0.4059	0.6383	0.5862
$HARP_{VAR}$	0.4150	0.2453	0.3790	0.6442	0.5785
$HARP_{RF}$	**0.4058**	0.2654	**0.3683**	**0.6370**	**0.5879**
Panel D: $h = 22$					
$HARP_{AR}$	0.4670	**0.2863**	0.4595	0.6834	0.5261
$HARP_{HAR}$	**0.4629**	0.3061	0.4614	**0.6804**	**0.5303**
$HARP_{VAR}$	0.5193	0.3587	0.4410	0.7206	0.4731
$HARP_{RF}$	0.4788	0.3748	**0.4266**	0.6919	0.5142

This table reports the out-of-sample results for predicting h-day future realized variation using the different models of forecasting $\hat{f}_{t+h}$. The candidate models are the HAR model ($HARP_{HAR}$), the AR(h) model ($HARP_{AR}$), and the random forecast method ($HARP_{RF}$). The results are based on the transaction data of the NASDAQ 100 ETF spanning from May 22, 2007 to October 20, 2017 (a total of 2,625 observations). We use a rolling window of 1000 observations to estimate the coefficients of the HARP model, and evaluate the out-of-sample forecast performance at four horizons ($h = 1$, $h = 5$, $h = 10$ and $h = 22$). Each panel in Table 5 corresponds to a specific forecast horizon, which varies from 1 day to 22 days. Bold numbers indicate the best performing model by each criterion at each forecast horizon.

Table 6: The Giacomini-White test for the mean absolute forecast errors-different ways of forecasting $\hat{f}_{t+h}$

Method	$HARP_{AR}$	$HARP_{HAR}$	$HARP_{VAR}$	$HARP_{AR}$	$HARP_{HAR}$	$HARP_{VAR}$
		$h = 1$			$h = 5$	
$HARP_{AR}$	-	-	-	-	-	-
$HARP_{HAR}$	0.0244	-	-	0.1274	-	-
$HARP_{VAR}$	0.4173	0.0185	-	0.0097	0.0266	
$HARP_{RF}$	0.0573	0.3722	0.4001	**0.0000**	**0.0000**	**0.0000**
		$h = 10$			$h = 22$	
$HARP_{AR}$	-	-	-	-	-	-
$HARP_{HAR}$	0.3843	-	-	0.6001	-	-
$HARP_{VAR}$	0.0088	0.0021	-	0.3418	0.2872	-
$HARP_{RF}$	**0.0003**	**0.0001**	**0.0141**	**0.0476**	**0.0325**	0.1026

The modified Giacomini-White test (Giacomini and White, 2006) is implemented to test the null hypothesis that the *row method* (in vertical headings) performs equally well as the *column method* (in horizontal headings) in terms of the absolute forecast error. Corresponding p values for a number of forecasting horizons ($h = 1, 5, 10, 22$) are reported in Panels A to D of Table 6, respectively. Bold numbers indicate the null hypothesis can be rejected at 5% level of significance.

model not containing $\boldsymbol{f}_t$. The above finding indicates that aside from some additional useful information for prediction, the unobserved common factors partially capture the information contained in positive and negative realized semivariances.

Table 7: In-sample results for the benchmark specification with and without unobserved common factors

	$h=1$		$h=5$		$h=10$		$h=22$	
	M	M'	M	M'	M	M'	M	M'
Panel A: Feasible common factors								
Constant	0.3700	0.4505	0.4317	0.5096	0.4845	0.5260	0.6013	0.6428
	(0.0413)	(0.0237)	(0.0508)	(0.0310)	(0.0565)	(0.0352)	(0.0629)	(0.0379)
OIL	0.4373	-0.2574	-0.9595	-0.9611	0.4554	-0.1422	-0.7933	-1.0313
	(0.6334)	(0.3622)	(0.7781)	(0.4736)	(0.8639)	(0.5377)	(0.9577)	(0.5770)
USD	-1.7878	1.2087	-1.3487	0.7342	-3.2909	-1.2032	0.9706	2.3014
	(3.0697)	(1.7551)	(3.7716)	(2.2960)	(4.1871)	(2.6061)	(4.6416)	(2.7970)
CS	-1.3083	-0.8762	-0.0156	0.7781	0.4170	1.1349	0.9932	1.6086
	(0.6697)	(0.3829)	(0.8226)	(0.5009)	(0.9131)	(0.5683)	(1.0124)	(0.6104)
TS	-0.0590	-0.0254	-0.0419	0.0049	-0.0298	-0.0233	-0.0303	-0.0196
	(0.0180)	(0.0103)	(0.0222)	(0.0135)	(0.0246)	(0.0154)	(0.0273)	(0.0165)
FFD	1.7295	1.2395	1.0488	0.3062	1.5633	1.5272	1.7559	1.6254
	(0.1547)	(0.0889)	(0.1901)	(0.1165)	(0.2112)	(0.1348)	(0.2344)	(0.1429)
Panel B: Cross-section specific regressors								
$\mathrm{RV}_t^{(1)} \cdot \mathbb{I}(r_t < 0)$	0.0778	0.0798	-0.1214	-0.0585	-0.0954	-0.0634	-0.0147	0.0068
	(0.0333)	(0.0190)	(0.0409)	(0.0249)	(0.0453)	(0.0282)	(0.0502)	(0.0303)
RS_t^+	0.1129	0.1315	0.0926	0.0995	0.2715	0.1749	0.0196	0.0176
	(0.0501)	(0.0287)	(0.0616)	(0.0375)	(0.0683)	(0.0426)	(0.0756)	(0.0457)
RS_t^-	0.2733	0.0957	0.1447	0.0673	0.1229	0.0459	0.1368	0.0731
	(0.0163)	(0.0097)	(0.0201)	(0.0123)	(0.0223)	(0.0139)	(0.0247)	(0.0149)
$\mathrm{RV}_t^{(5)}$	0.2379	0.1930	0.3287	0.1714	0.1177	0.0674	0.1744	0.0746
	(0.0220)	(0.0126)	(0.0270)	(0.0166)	(0.0300)	(0.0187)	(0.0332)	(0.0201)
$\mathrm{RV}_t^{(22)}$	0.2710	0.2490	0.2548	0.2466	0.3098	0.3770	0.2320	0.2923
	(0.0165)	(0.0094)	(0.0203)	(0.0123)	(0.0225)	(0.0149)	(0.0249)	(0.0151)
Panel C: Estimated unobserved common factors								
Common effect f_1		-0.4718		-0.6600		-0.7052		0.8185
		(0.0086)		(0.0108)		(0.0124)		(0.0132)
Common effect f_2		-0.4150		-0.3149		0.3062		-0.2731
		(0.0081)		(0.0109)		(0.0133)		(0.0136)
Panel D: Goodness of fit								
R^2	0.6765	0.8944	0.5127	0.8196	0.4008	0.7681	0.2686	0.7347
Adj. R^2	0.6753	0.8939	0.5108	0.8188	0.3985	0.7670	0.2658	0.7335

This table reports the in-sample results for predicting the h-day future realized volatility using the HAR-RS-II model and the HARP model. The estimation is based on the full sample data of the NASDAQ 100 ETF and considers a range of forecast horizons ($h = 1, 5, 10, 22$). Panel A reports the coefficient estimates of observed common factors and their standard errors (in parentheses). Panel B reports the coefficient estimates of unit-specific regressors and their standard errors (in brackets), while panel C reports the relevant results for unobserved common factors. The bottom panel provides the in-sample predictive R^2 and adjusted R^2.

Table 8 presents some descriptive results of the out-of-sample evaluation for forecasts 1, 5, 10 and 22 days ahead. In particular, we report the MSFE, SDFE, MAFE, QLIKE as well as the pseudo R^2 from the rolling-window regressions for the HARP model and for the set

Table 8: Out-of-sample forecast comparison of models for RV of the NASDAQ 100 ETF

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
AR(22)	0.4428	0.1587	0.3282	0.6654	0.5487
HAR	0.4190	0.1569	0.3250	0.6473	0.5729
HAR-J	0.3783	0.1487	0.3114	0.6150	0.6145
HAR-CJ	0.3575	0.1491	0.3096	0.5979	0.6356
HAR-RS-I	0.2341	0.1371	0.2884	0.4839	0.7614
HAR-RS-II	0.2302	0.1349	0.2819	0.4798	0.7654
HAR-SJ-I	0.2498	0.1465	0.2927	0.4998	0.7454
HAR-SJ-II	0.2333	0.1382	0.2891	0.4830	0.7622
HARP	**0.2114**	**0.1339**	**0.2698**	**0.4598**	**0.7845**
Panel B: $h = 5$					
AR(22)	0.5623	0.2618	0.4247	0.7499	0.4280
HAR	0.4761	0.2547	0.4101	0.6900	0.5158
HAR-J	0.4768	0.2728	0.4052	0.6905	0.5151
HAR-CJ	0.4578	0.3255	0.3942	0.6766	0.5343
HAR-RS-I	0.4780	0.2548	0.3984	0.6914	0.5138
HAR-RS-II	0.4632	0.2612	0.3965	0.6806	0.5289
HAR-SJ-I	0.4874	0.2577	0.3999	0.6981	0.5043
HAR-SJ-II	0.4462	0.2443	0.3925	0.6680	0.5461
HARP	**0.3433**	**0.1931**	**0.3409**	**0.5859**	**0.6509**
Panel C: $h = 10$					
AR(22)	0.5782	0.3126	0.4460	0.7604	0.4128
HAR	0.5426	0.3322	0.4398	0.7366	0.4490
HAR-J	0.5223	0.9742	0.4368	0.7227	0.4696
HAR-CJ	0.5222	0.3965	0.4276	0.7226	0.4697
HAR-RS-I	0.5107	0.3472	0.4310	0.7146	0.4814
HAR-RS-II	0.5027	0.3552	0.4269	0.7090	0.4895
HAR-SJ-I	0.5056	0.3622	0.4308	0.7110	0.4866
HAR-SJ-II	0.9043	0.3170	0.4415	0.9509	0.0817
HARP	**0.4058**	**0.2654**	**0.3683**	**0.6370**	**0.5879**
Panel D: $h = 22$					
AR(22)	0.6179	0.3958	0.4941	0.7861	0.3730
HAR	0.6012	0.3952	0.4899	0.7754	0.3900
HAR-J	0.6112	0.3815	0.4885	0.7818	0.3799
HAR-CJ	0.6250	0.3982	0.4883	0.7906	0.3658
HAR-RS-I	0.6035	0.4550	0.4899	0.7768	0.3877
HAR-RS-II	0.6045	0.3888	0.4872	0.7775	0.3866
HAR-SJ-I	0.6017	0.4642	0.4898	0.7757	0.3894
HAR-SJ-II	0.6055	0.5216	0.4887	0.7781	0.3856
HARP	**0.4788**	**0.3748**	**0.4266**	**0.6919**	**0.5142**

This table reports the out-of-sample results for predicting h-day future realized variation using the different predictor variables and risk models. The results are based on the transaction data of the NASDAQ 100 ETF spanning from May 22, 2007 to October 20, 2017 (a total of 2,625 observations). We use a rolling window of 1000 observations to estimate the coefficients of the models, and evaluate the out-of-sample forecast performance at four horizons ($h = 1$, $h = 5$, $h = 10$ and $h = 22$). Each panel in Table 8 corresponds to a specific forecast horizon, which varies from 1 day to 22 days. Bold numbers indicate the best performing model by each criterion at each forecast horizon.

of candidate models discussed in Section 2. We find consistent ranking of models across all forecast horizons: the AR(22) performs the worst, followed by models with high-frequency intraday data (i.e., the HAR, the HAR-J and the HAR-CJ), while the more sophisticated RV models perform better. The HARP model has the best performance in all cases of h. To further examine whether the out-performance is statistically significant, we perform the modified GW test in Table 9. The outcome indicates that the out-performance of the HARP model is statistically significant at all four forecasting horizons.

6 Robustness checks

This section presents three out-of-sample checks on the conclusions from the previous section, with the identical set of volatility models. The first is an application to the Dow Jones Industrial Average (DJIA). The second is a test of the main results on the NASDAQ 100 ETF with a different sampling frequency, a 5-minute sampling interval. The third is an exercise using the principal components (PCs) approach, an alternative way of estimating unobserved common factors. Apart from the above trials, we also test our findings by other industrial indices in Appendix G.1, by a different window length in Appendix G.2, and by varying the sample period in Appendix G.3.

6.1 Evidence on DJIA

The previous sections presented results for the NASDAQ 100 ETF. In this section we report out-of-sample results for the DJIA index. Data at 1-minute sampling interval is provided by Pitrading Incorporation and covers the same period as the main results. The DJIA is a weighted average index that indicates the value of 30 large, publicly owned companies

Table 9: The Giacomini-White test for the mean absolute forecast errors-the NASDAQ 100 ETF

Method	AR(22)	HAR	HAR-J	HAR-CJ	HAR-RS-I	HAR-RS-II	HAR-SJ-I	HAR-SJ-II
Panel A: $h = 1$								
AR(22)	-	-	-	-	-	-	-	-
HAR-	0.5936	-	-	-	-	-	-	-
HAR-J	0.0068	0.0000	-	-	-	-	-	-
HAR-CJ	0.0008	0.0000	0.4332	-	-	-	-	-
HAR-RS-I	0.0001	0.0001	0.0050	0.0058	-	-	-	-
HAR-RS-II	0.0000	0.0000	0.0000	0.0000	0.0304	-	-	-
HAR-SJ-I	0.0000	0.0000	0.0020	0.0028	0.0656	0.0000	-	-
HAR-SJ-II	0.0002	0.0002	0.0068	0.0086	0.3482	0.0190	0.1289	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0060**	**0.0000**	**0.0000**
Panel B: $h = 5$								
AR(22)	-	-	-	-	-	-	-	-
HAR-	0.1284	-	-	-	-	-	-	-
HAR-J	0.0475	0.0062	-	-	-	-	-	-
HAR-CJ	0.0028	0.0002	0.0077	-	-	-	-	-
HAR-RS-I	0.0139	0.0003	0.0027	0.3819	-	-	-	-
HAR-RS-II	0.0093	0.0000	0.0003	0.5971	0.1716	-	-	-
HAR-SJ-I	0.0178	0.0030	0.0260	0.2597	0.0175	0.0714	-	-
HAR-SJ-II	0.0034	0.0000	0.0000	0.7023	0.0072	0.0176	0.0052	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel C: $h = 10$								
AR(22)	-	-	-	-	-	-	-	-
HAR-	0.4620	-	-	-	-	-	-	-
HAR-J	0.3554	0.2268	-	-	-	-	-	-
HAR-CJ	0.0477	0.0042	0.0154	-	-	-	-	-
HAR-RS-I	0.1665	0.0236	0.0022	0.4104	-	-	-	-
HAR-RS-II	0.1279	0.0359	0.0165	0.9043	0.1295	-	-	-
HAR-SJ-I	0.1858	0.0536	0.0185	0.4746	0.8771	0.0495	-	-
HAR-SJ-II	0.6800	0.8733	0.7063	0.2821	0.4451	0.3663	0.4666	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0001**
Panel D: $h = 22$								
AR(22)	-	-	-	-	-	-	-	-
HAR-	0.6077	-	-	-	-	-	-	-
HAR-J	0.4487	0.5539	-	-	-	-	-	-
HAR-CJ	0.5025	0.6427	0.9489	-	-	-	-	-
HAR-RS-I	0.5918	0.9796	0.4787	0.6039	-	-	-	-
HAR-RS-II	0.3305	0.3026	0.4892	0.7376	0.1795	-	-	-
HAR-SJ-I	0.5900	0.8755	0.6093	0.6654	0.8190	0.2828	-	-
HAR-SJ-II	0.4660	0.5700	0.8699	0.8888	0.3776	0.2854	0.5575	-
HARP	**0.0075**	**0.0127**	**0.0136**	**0.0140**	**0.0124**	**0.0131**	**0.0126**	**0.0120**

The modified Giacomini-White test (Giacomini and White, 2006) is implemented to test the null hypothesis that the *row method* (in vertical headings) performs equally well as the *column method* (in horizontal headings) in terms of the absolute forecast error. Corresponding p values for a number of forecasting horizons ($h = 1, 5, 10, 22$) are reported in Panels A to D of Table 9, respectively. Bold numbers indicate the null hypothesis can be rejected at 5% level of significance.

based in the United States. All forecasts are generated using rolling window regressions based on 1,000 observations, and parameter estimates are updated daily.

The results are reported in Tables 10 and 11. We note that the HARP forecast is always the winner and significantly improves the out-of-sample forecast performance. Relative

Table 10: Out-of-sample forecast comparison of models on the DJIA

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
AR(22)	1.1207	0.2401	0.3852	1.0586	0.2990
HAR	0.9000	0.2177	0.3707	0.9487	0.4370
HAR-J	0.8720	0.2116	0.3648	0.9338	0.4545
HAR-CJ	0.9153	0.2154	0.3722	0.9567	0.4275
HAR-RS-I	0.9776	0.2063	0.3636	0.9887	0.3885
HAR-RS-II	0.8223	0.2049	0.3467	0.9068	0.4857
HAR-SJ-I	0.9325	0.2089	0.3662	0.9656	0.4168
HAR-SJ-II	1.1121	0.2036	0.3653	1.0546	0.3044
HARP	**0.5487**	**0.1852**	**0.3005**	**0.7407**	**0.6568**
Panel B: $h = 5$					
AR(22)	1.3034	0.3465	0.4744	1.1416	0.1859
HAR	0.8782	0.3252	0.4391	0.9371	0.4515
HAR-J	0.8822	0.3200	0.4388	0.9392	0.4490
HAR-CJ	0.9012	0.3502	0.4391	0.9493	0.4371
HAR-RS-I	0.8617	0.3206	0.4337	0.9283	0.4618
HAR-RS-II	0.8222	0.3157	0.4306	0.9068	0.4864
HAR-SJ-I	0.8471	0.3194	0.4333	0.9204	0.4709
HAR-SJ-II	0.7602	0.3105	0.4254	0.8719	0.5252
HARP	**0.5644**	**0.2209**	**0.3435**	**0.7513**	**0.6475**
Panel C: $h = 10$					
AR(22)	1.0535	0.4052	0.5045	1.0264	0.3432
HAR	0.8767	0.3756	0.4849	0.9363	0.4534
HAR-J	0.8740	0.3732	0.4822	0.9349	0.4551
HAR-CJ	0.9050	0.3761	0.4830	0.9513	0.4358
HAR-RS-I	0.8780	0.3724	0.4835	0.9370	0.4526
HAR-RS-II	0.8853	0.3763	0.4832	0.9409	0.4481
HAR-SJ-I	0.8708	0.3732	0.4837	0.9332	0.4571
HAR-SJ-II	0.8077	0.3692	0.4778	0.8987	0.4964
HARP	**0.6043**	**0.3014**	**0.3780**	**0.7773**	**0.6233**
Panel D: $h = 22$					
AR(22)	2.0514	0.5088	0.6419	1.4323	-0.2733
HAR	1.1346	0.4721	0.5842	1.0652	0.2958
HAR-J	1.1211	0.4865	0.5820	1.0588	0.3042
HAR-CJ	1.1576	0.4725	0.5947	1.0759	0.2815
HAR-RS-I	1.1450	0.5333	0.5847	1.0700	0.2894
HAR-RS-II	1.0788	0.4880	0.5840	1.0386	0.3304
HAR-SJ-I	1.1261	0.6599	0.5844	1.0612	0.3011
HAR-SJ-II	1.0841	0.5133	0.5801	1.0412	0.3272
HARP	**0.7172**	**0.3415**	**0.4420**	**0.8469**	**0.5548**

This table reports the out-of-sample results for predicting h-day future realized variation using the different predictor variables and risk models. The results are based on data of the Dow Jones Industrial Average spanning from May 22, 2007 to October 20, 2017 (a total of 2,625 observations). We use a rolling window of 1000 observations to estimate the coefficients of the models, and evaluate the out-of-sample forecast performance at four horizons ($h = 1$, $h = 5$, $h = 10$ and $h = 22$). Each panel in Table 10 corresponds to a specific forecast horizon, which varies from 1 day to 22 days. Bold numbers indicate the best performing model by each criterion at each forecast horizon.

to the best semivariance-based specification, the HARP alternative generates gains in the out-of-sample R^2 of between 35.2% ($h = 1$) and 69.6% ($h = 22$). The GW test in Table 11 implies that the improvement by HARP is significant. Overall, the above conclusions are in agreement with the results for the NASDAQ 100 ETF.

Table 11: The Giacomini-White test for the mean absolute forecast errors-the DJIA

Method	AR(22)	HAR	HAR-J	HAR-CJ	HAR-RS-I	HAR-RS-II	HAR-SJ-I	HAR-SJ-II
Panel A: $h = 1$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.0959	-	-	-	-	-	-	-
HAR-J	0.0239	0.0266	-	-	-	-	-	-
HAR-CJ	0.1254	0.3795	0.0016	-	-	-	-	-
HAR-RS-I	0.0129	0.0034	0.7456	0.0033	-	-	-	-
HAR-RS-II	0.0000	0.0000	0.0009	0.0000	0.0001	-	-	-
HAR-SJ-I	0.0316	0.1331	0.7198	0.0833	0.1478	0.0000	-	-
HAR-SJ-II	0.0324	0.1526	0.9165	0.0919	0.4069	0.0004	0.7530	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel B: $h = 5$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.0413	-	-	-	-	-	-	-
HAR-J	0.0357	0.6419	-	-	-	-	-	-
HAR-CJ	0.0299	0.9917	0.9169	-	-	-	-	-
HAR-RS-I	0.0182	0.0014	0.0039	0.1231	-	-	-	-
HAR-RS-II	0.0133	0.0395	0.0419	0.0979	0.4131	-	-	-
HAR-SJ-I	0.0184	0.0013	0.0045	0.1095	0.4816	0.4684	-	-
HAR-SJ-II	0.0099	0.0001	0.0003	0.0122	0.0169	0.2224	0.0162	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel C: $h = 10$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.0370	-	-	-	-	-	-	-
HAR-J	0.0168	0.0010	-	-	-	-	-	-
HAR-CJ	0.0172	0.6185	0.8431	-	-	-	-	-
HAR-RS-I	0.0234	0.3035	0.3525	0.8879	-	-	-	-
HAR-RS-II	0.0214	0.4243	0.5734	0.9466	0.8606	-	-	-
HAR-SJ-I	0.0262	0.4239	0.3272	0.8486	0.5191	0.7606	-	-
HAR-SJ-II	0.0050	0.0064	0.0889	0.3101	0.0472	0.1004	0.0365	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel D: $h = 22$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.1786	-	-	-	-	-	-	-
HAR-J	0.1690	0.3074	-	-	-	-	-	-
HAR-CJ	0.2649	0.0354	0.0164	-	-	-	-	-
HAR-RS-I	0.1720	0.7118	0.3231	0.0451	-	-	-	-
HAR-RS-II	0.1663	0.9334	0.4615	0.0432	0.7895	-	-	-
HAR-SJ-I	0.1728	0.8626	0.3431	0.0412	0.4813	0.8687	-	-
HAR-SJ-II	0.1527	0.0357	0.3910	0.0078	0.0311	0.1347	0.0220	-
HARP	**0.0003**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**

The modified Giacomini-White test (Giacomini and White, 2006) is implemented to test the null hypothesis that the *row method* (in vertical headings) performs equally well as the *column method* (in horizontal headings) in terms of the absolute forecast error. Corresponding p values for a number of forecasting horizons ($h = 1, 5, 10, 22$) are reported in Panels A to D of Table 11, respectively. Bold numbers indicate the null hypothesis can be rejected at 5% level of significance.

6.2 Variation in the sampling frequency

In this section, we examine the robustness of the HARP model to the intraday RVs constructed from a 5-minute sampling frequency. This choice directly reflects the sampling frequency used in much of the existing realized volatility literature.[24] Summary statistics of the 5-minute RVs of the NASDAQ 100 ETF are reported in Table 12.

Table 12: Summary statistics for the 5-min RV of the NASDAQ100 ETF and selected stocks

Statistic	QQQ	Tickers of representative stocks					
		AAL	ALXN	DISCA	ISRG	QCOM	XLNX
Mean	0.9913	14.8403	4.2783	3.3142	3.5345	2.1977	2.5813
Median	0.5586	6.2083	2.9420	1.8320	2.0853	1.3135	1.5721
Maximum	9.8942	1007.9313	228.1957	896.5039	56.6374	54.5670	35.3205
Minimum	0.0461	0.4526	0.3081	0.2676	0.2545	0.1351	0.2156
Standard Deviation	1.2470	32.3428	6.4480	17.9007	4.4494	2.8914	2.9424
Skewness	3.2858	14.2782	18.8513	47.9847	4.2328	5.6266	3.6790
Kurtosis	16.9410	370.0102	587.4505	2393.3160	31.4911	63.2058	23.8816
Jarque-Bera	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010
ADF Test	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010

The second column contains statistics for the RV of the NASDAQ100 ETF from May 22, 2007 to October 20, 2017. The statistics of 6 component stocks are presented in columns 3 to 8. For JB and ADF test statistics that are outside tabulated critical values, we report maximum (0.999) or minimum (0.001) p-values.

Comparing to the statistics based on 1-minute intraday data in Table 2, we notice very minor changes. We then duplicate the rolling window regressions relying on the the 5-minute data. Tables 13 and 14 contain the results from the forecasting analysis. The losses for all the forecast criteria are systematically lower than those in the 1-minute case. We also notice that the results are qualitatively the same as those based on the 1-minute data. The experiment confirms that the robustness of HARP under different sampling frequency.

6.3 An alternative way of estimating f_t

In this section, we consider a direct estimation of unobserved common factors from RVs of the NASDAQ 100 components, using the PC analysis. Following Kapetanios and Pesaran

[24] See also the theoretical comparisons of various volatility estimators in Andersen, Bollerslev, and Meddahi (2011) and Ghysels and Sinko (2011) from a forecasting perspective. Liu, Patton, and Sheppard (2015) provide a comprehensive empirical study of 400 volatility estimators across multiple assets.

Table 13: Out-of-sample forecast comparison of models on the NASDAQ 100 ETF sampled at a 5-minute frequency

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
AR(22)	1.1111	0.2327	0.4401	1.3125	-0.8863
HAR	1.0463	0.2257	0.3824	1.0229	-0.1457
HAR-J	0.7040	0.1963	0.3428	0.8390	0.2291
HAR-CJ	0.7178	0.1881	0.3293	0.8472	0.2139
HAR-RS-I	0.3090	0.1906	0.3269	0.5559	0.6616
HAR-RS-II	0.6451	0.1847	0.3255	0.8032	0.2936
HAR-SJ-I	0.2812	0.1909	0.3249	0.5303	0.6921
HAR-SJ-II	0.2641	0.1853	0.3125	0.5139	0.7108
HARP	**0.2400**	**0.1757**	**0.2948**	**0.4899**	**0.7372**
Panel B: $h = 5$					
AR(22)	1.1099	0.4246	0.4834	1.0535	-0.2161
HAR	0.5493	0.3099	0.4292	0.7411	0.3982
HAR-J	0.4898	0.3009	0.4148	0.6999	0.4633
HAR-CJ	0.5215	0.3006	0.4083	0.7221	0.4286
HAR-RS-I	0.4766	0.2976	0.4125	0.6903	0.4778
HAR-RS-II	0.4808	0.2976	0.4102	0.6934	0.4732
HAR-SJ-I	0.4804	0.3016	0.4130	0.6931	0.4736
HAR-SJ-II	0.4776	0.2927	0.4051	0.6911	0.4767
HARP	**0.4051**	**0.2545**	**0.3702**	**0.6365**	**0.5561**
Panel C: $h = 10$					
AR(22)	0.8593	0.3482	0.4736	0.9270	0.0609
HAR	0.6494	0.3418	0.4403	0.8058	0.2903
HAR-J	0.7303	0.3389	0.4366	0.8546	0.2019
HAR-CJ	0.7456	0.3509	0.4363	0.8635	0.1852
HAR-RS-I	0.5661	0.3377	0.4309	0.7524	0.3814
HAR-RS-II	0.5166	0.3393	0.4262	0.7188	0.4354
HAR-SJ-I	0.5110	0.3390	0.4289	0.7149	0.4415
HAR-SJ-II	1.0454	0.3296	0.4359	1.0224	-0.1424
HARP	**0.4236**	**0.2895**	**0.3825**	**0.6508**	**0.5371**
Panel D: $h = 22$					
AR(22)	0.6860	0.4409	0.5087	0.8283	0.2520
HAR	0.5850	0.4175	0.4847	0.7649	0.3621
HAR-J	0.5858	0.4740	0.4824	0.7654	0.3613
HAR-CJ	0.6163	0.4414	0.4887	0.7850	0.3280
HAR-RS-I	0.5884	0.5270	0.4839	0.7671	0.3584
HAR-RS-II	0.6486	0.4211	0.4870	0.8054	0.2928
HAR-SJ-I	0.5862	0.4890	0.4840	0.7656	0.3609
HAR-SJ-II	0.5940	0.4826	0.4808	0.7707	0.3524
HARP	**0.4520**	**0.3329**	**0.4339**	**0.6723**	**0.5072**

This table reports the out-of-sample results for predicting h-day future realized variation using the different predictor variables and risk models. The results are based on data of the Dow Jones Transportation Average spanning from May 22, 2007 to October 20, 2017 (a total of 2,625 observations). We use a rolling window of 1000 observations to estimate the coefficients of the models, and evaluate the out-of-sample forecast performance at four horizons ($h = 1$, $h = 5$, $h = 10$ and $h = 22$). Each panel in Table 13 corresponds to a specific forecast horizon, which varies from 1 day to 22 days. Bold numbers indicate the best performing model by each criterion at each forecast horizon.

Table 14: The Giacomini-White test for the mean absolute forecast error-the NASDAQ 100 ETF at a 5-minute sampling frequency

Method	AR(22)	HAR	HAR-J	HAR-CJ	HAR-RS-I	HAR-RS-II	HAR-SJ-I	HAR-SJ-II
Panel A: $h = 1$								
AR(22)								
HAR	0.0013							
HAR-J	0.0000	0.0000						
HAR-CJ	0.0000	0.0000	0.0000					
HAR-RS-I	0.0000	0.0012	0.1849	0.8437				
HAR-RS-II	0.0000	0.0000	0.0000	0.3558	0.9096			
HAR-SJ-I	0.0000	0.0048	0.2370	0.7770	0.5794	0.9649		
HAR-SJ-II	0.0000	0.0008	0.0542	0.3025	0.0013	0.3993	0.0000	
HARP	**0.0000**	**0.0000**	**0.0029**	**0.0379**	**0.0000**	**0.0033**	**0.0000**	**0.0000**
Panel B: $h = 5$								
AR(22)								
HAR	0.0593							
HAR-J	0.0261	0.0009						
HAR-CJ	0.0151	0.0003	0.1482					
HAR-RS-I	0.0250	0.0016	0.3114	0.4493				
HAR-RS-II	0.0201	0.0005	0.1141	0.7354	0.2580			
HAR-SJ-I	0.0252	0.0013	0.3025	0.3534	0.5359	0.2264		
HAR-SJ-II	0.0118	0.0000	0.0001	0.5128	0.0065	0.0582	0.0015	
HARP	**0.0008**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel C: $h = 10$								
AR(22)								
HAR	0.1192							
HAR-J	0.0863	0.2069						
HAR-CJ	0.0969	0.6042	0.9643					
HAR-RS-I	0.0848	0.0651	0.3394	0.6036				
HAR-RS-II	0.0776	0.0802	0.2500	0.4410	0.1607			
HAR-SJ-I	0.0956	0.1605	0.4023	0.5783	0.5424	0.0345		
HAR-SJ-II	0.0594	0.6190	0.9218	0.9444	0.6916	0.5334	0.6575	
HARP	**0.0024**	**0.0000**	**0.0000**	**0.0008**	**0.0000**	**0.0000**	**0.0000**	**0.0036**
Panel D: $h = 22$								
AR(22)								
HAR	0.1951							
HAR-J	0.1514	0.1019						
HAR-CJ	0.2441	0.4948	0.2226					
HAR-RS-I	0.1770	0.4383	0.1499	0.3582				
HAR-RS-II	0.1928	0.6754	0.3407	0.6512	0.5147			
HAR-SJ-I	0.1854	0.5111	0.1387	0.4013	0.7541	0.5737		
HAR-SJ-II	0.1180	0.2433	0.5217	0.0897	0.2890	0.0853	0.3150	
HARP	**0.0046**	**0.0045**	**0.0058**	**0.0034**	**0.0055**	**0.0034**	**0.0057**	**0.0053**

The modified Giacomini-White test (Giacomini and White, 2006) is implemented to test the null hypothesis that the *row method* (in vertical headings) performs equally well as the *column method* (in horizontal headings) in terms of the absolute forecast error. Corresponding p values for a number of forecasting horizons ($h = 1, 5, 10, 22$) are reported in Panels A to D of Table 14, respectively. Bold numbers indicate the null hypothesis can be rejected at 5% level of significance.

(2004), this is implemented in a two-stage procedure, where in the first stage principal components for RVs of the NASDAQ 100 components are obtained as in Stock and Watson (2002), and in the second stage the model for the NASDAQ 100 ETF is estimated augmenting the observed regressors with the estimated PCs. It can be seen that this approach immediately generates the unobserved factors instead of approximating them by cross-section averages of the dependent variable and the observed regressors. Also unlike the CCE method, we need to determine the number of factors initially to implement the PC augmented procedure. This can be solved using the criteria in Bai and Ng (2002).

A theoretical comparison is provided in Kapetanios and Pesaran (2004) on the small sample performance of the CCE method and the PC approach. After a series of Monte Carlo experiments, Kapetanios and Pesaran (2004) conclude that the PC augmented method does not perform as well as the CCE estimator, and can be subject to substantial bias in small samples. They attribute this deficiency to the small sample errors in the number of factors selected by the Bai and Ng procedure. Because of the above discussion, it becomes more appealing to conduct an empirical comparison of the two approaches.

To compare things on a equal basis, we adopt the same model specification (i.e., the HAR-RS-II model) for the two procedures. The out-of-sample forecast results are presented in Tables 15, which shows that the HARP forecasts based on the CCE method still dominate those relying on the PC procedure. We conduct GW test and obtain the p-values $[0.0002, 0.0005, 0.0261, 0.5612]$ respectively for $h = 1, 5, 10$, and 22. The above conclusion is compatible with the theoretical finding in Kapetanios and Pesaran (2004).

7 Conclusion

In this paper, we argue that the linkages among component stock volatilities are important for forecasting the relevant index or index fund volatility. We develop a panel-based HAR model assuming unobserved common factors across cross-sectional units to capture the comovements in realized volatility. The framework configuration draws from the CCE-type estimators of Pesaran (2006) and Chudik and Pesaran (2015). It is shown that the

Table 15: Out-of-sample comparison of the HARP model and the PC method on the NASDAQ 100 ETF

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
HARPCA	0.2382	0.8017	0.2887	0.4881	0.7577
HARP	**0.2129**	**0.1307**	**0.2709**	**0.4614**	**0.7831**
Panel B: $h = 5$					
HARPCA	0.4310	1.3620	0.3681	0.6565	0.5616
HARP	**0.3458**	**0.2072**	**0.3409**	**0.5880**	**0.6483**
Panel C: $h = 10$					
HARPCA	0.4903	0.9658	0.3941	0.7002	0.5021
HARP	**0.4071**	**0.2449**	**0.3662**	**0.6380**	**0.5866**
Panel D: $h = 22$					
HARPCA	0.5650	1.1543	0.4390	0.7516	0.4267
HARP	**0.4835**	**0.3061**	**0.4258**	**0.6954**	**0.5093**

The results are based on data of the NASDAQ 100 ETF spanning from May 22, 2007 to October 20, 2017 (a total of 2,625 observations). HARPCA denotes estimating the unobserved factors with the principal component analysis. Bold numbers indicate the best performing model by each criterion at each forecast horizon.

CCE estimator is consistent. Monte Carlo studies show that the CCE estimator has good finite sample performance.

We illustrate the relevance of the proposed HARP model by an empirical application to forecasting the realized volatility of the NASDAQ 100 ETF. The in-sample analysis disclose that the unobserved factors from the panel regression play an important role. They capture information that are not already contained in the asset-specific realized volatility histories, such as the sentiment of the financial market, the news effect, or the varying risk premium.[25] Taking the unobserved factors into account leads up to 174% increase in the in-sample R^2. In the out-of-sample exercise, the HARP model that includes this latent common factors significantly outperforms models that do not. Moreover, the HARP model produces up to 32% increase in the pseudo R^2 of the forecasts of the index fund volatility. Our findings are robust to different stock indices and an alternative way of estimating the unobserved common factors.

[25] See, for example, Baker and Wurgler (2006) and Baker, Wurgler, and Yuan (2012) for the importance of investor sentiment in explaining the cross-section of stock returns.

References

ANDERSEN, T. G., AND T. BOLLERSLEV (1998): "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39(4), 885–905.

ANDERSEN, T. G., T. BOLLERSLEV, AND F. X. DIEBOLD (2007): "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility," *The Review of Economics and Statistics*, 89(4), 701–720.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2000): "Great Realizations," *RISK*, 13, 105–108.

——— (2001): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96(453), 42–55.

ANDERSEN, T. G., T. BOLLERSLEV, AND N. MEDDAHI (2011): "Realized Volatility Forecasting and Market Microstructure Noise," *Journal of Econometrics*, 160.

ASAI, M., M. MCALEER, AND J. YU (2006): "Multivariate Stochastic Volatility: A survey," *Econometric Reviews*, 25, 145–175.

BAI, J., AND S. NG (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1), 191–221.

BAKER, M., AND J. WURGLER (2006): "Investor Sentiment and the Cross-Section of Stock Returns," *The Journal of Finance*, 61(4), 1645–1680.

BAKER, M., J. WURGLER, AND Y. YUAN (2012): "Global, local, and contagious investor sentiment," *Journal of Financial Economics*, 104(2), 272 – 287, Special Issue on Investor Sentiment.

BAUWENS, L., S. LAURENT, AND J. V. K. ROMBOUTS (2006): "Multivariate GARCH models: A survey," *Journal of Applied Econometrics*, 21(1), 79–109.

BERNOTH, K., AND A. PICK (2009): "Forecasting the fragility of the banking and insurance sectors," De Nederlandsche Bank 202, DNB Working Paper.

——— (2011): "Forecasting the fragility of the banking and insurance sectors," *Journal of Banking and Finance*, 35(4), 807 – 818.

BOLLERSLEV, T., B. HOOD, J. HUSS, AND L. H. PEDERSEN (2018): "Risk Everywhere: Modeling and Managing Volatility," *The Review of Financial Studies*, 31(7), 2729–2773.

BREIMAN, L. (2001): "Random Forests," *Machine Learning*, 45, 5–32.

BREIMAN, L., J. FRIEDMAN, AND C. J. STONE (1984): *Classification and Regression Trees*. Chapman and Hall/CRC.

CHUDIK, A., K. MOHADDES, M. H. PESARAN, AND M. RAISSI (2017): “Is There a Debt-threshold Effect on Output Growth?,” *Review of Economics and Statistics*, 99(1), 135–150.

CHUDIK, A., AND M. H. PESARAN (2015): “Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors,” *Journal of Econometrics*, 188(2), 393 – 420, Heterogeneity in Panel Data and in Nonparametric Analysis in honor of Professor Cheng Hsiao.

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

CORSI, F., F. AUDRINO, AND R. RENÒ (2012): “HAR modeling for realized volatility forecasting,” in *Handbook of Volatility Models and Their Applications*, pp. 363–382. John Wiley & Sons, Inc.

ENGLE, R. F., T. ITO, AND W.-L. LIN (1990): “Meteor Showers or Heat Waves? Heteroskedastic Intra-daily Volatility in the Foreign Exchange Market,” *Econometrica*, 58(3), 525–542.

FERNANDES, M., M. C. MEDEIROS, AND M. SCHARTH (2014): “Modeling and Predicting the CBOE Market Volatility Index,” *Journal of Banking & Finance*, 40, 1–10.

GHYSELS, E., AND A. SINKO (2011): “Volatility Forecasting and Microstructure Noise,” *Journal of Econometrics*, 160.

GIACOMINI, R., AND H. WHITE (2006): “Tests of Conditional Predictive Ability,” *Econometrica*, 74(6), 1545–1578.

HANSEN, P. R., AND A. LUNDE (2006): “Consistent ranking of volatility models,” *Journal of Econometrics*, 131(1-2), 97–121.

HUANG, X., AND G. TAUCHEN (2005): “The Relative Contribution of Jumps to Total Price Variance,” *Journal of Financial Econometrics*, 3(4), 456–499.

KAPETANIOS, G., AND M. H. PESARAN (2004): “Alternative Approaches to Estimation and Inference in Large Multifactor Panels : Small Sample Results with an Application to Modelling of Asset Returns,” CESifo Discussion Paper 1416.

KAROLYI, A. G. (1995): “A Multivariate GARCH Model of International Transmissions of Stock Returns and Volatility: The Case of the United States and Canada,” *Journal of Business & Economic Statistics*, 13(1), 11–25.

KREISS, J.-P., AND S. N. LAHIRI (2012): “Bootstrap Methods for Time Series,” in *Time Series Analysis: Methods and Applications, Volume 30*, ed. by T. S. Rao, S. S. Rao, and C. Rao, chap. 1, pp. 3–26. North Holland.

KULPERGER, R. J., AND B. L. S. P. RAO (1989): “Bootstrapping a Finite State Markov Chain,” *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 51(2), 178–191.

KÜNSCH, H. R. (1989): “The Jackknife and the Bootstrap for General Stationary Observations,” *The Annals of Statistics*, 17(3), 1217–1241.

LIU, Y. L., J. A. PATTON, AND K. SHEPPARD (2015): “Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes,” *Journal of Econometrics*, 187.

LUKEPOHL, H. (2005): *New Introduction to Multiple Time Series Analysis*. Springer-Verlag Berlin Heidelberg.

MINCER, J., AND V. ZARNOWITZ (1969): “The Evaluation of Economic Forecasts,” in *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, pp. 3–46. National Bureau of Economic Research, Inc.

MÜLLER, U. A., M. M. DACOROGNA, R. D. DAVÉ, O. V. PICTET, R. B. OLSEN, AND J. WARD (1993): “Fractals and intrinsic time - a challenge to econometricians,” Discussion paper.

PATTON, A. J. (2011): “Volatility forecast comparison using imperfect volatility proxies,” *Journal of Econometrics*, 160(1), 246 – 256.

PATTON, A. J., AND K. SHEPPARD (2009): “Optimal combinations of realised volatility estimators,” *International Journal of Forecasting*, 25(2), 218 – 238.

——— (2015): “Good Volatility, Bad Volatility: Signed Jumps and The Persistence of Volatility,” *The Review of Economics and Statistics*, 97(3), 683–697.

PESARAN, M. H. (2004): “General Diagnostic Tests for Cross Section Dependence in Panels,” IZA Discussion Paper 1240.

——— (2006): “Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure,” *Econometrica*, 74(4), 967–1012.

——— (2015): “Testing Weak Cross-Sectional Dependence in Large Panels,” *Econometric Reviews*, 34(6-10), 1089–1117.

PESARAN, M. H., A. PICK, AND T. ALLAN (2011): “Variable selection, estimation and inference for multi-period forecasting problems,” *Journal of Econometrics*, 164.

STOCK, J. H., AND M. W. WATSON (2002): “Macroeconomic Forecasting Using Diffusion Indexes,” *Journal of Business & Economic Statistics*, 20(2), 147–162.