Appendix for "Forecasting Equity Index Volatility by Measuring the Linkage among Component Stocks"

March 1, 2019

Abstract

In this appendix, seven sections are included, which provide further details regarding (1) the proof of theorem 1, (2) simulation studies, (3) description of random forest, (4) details of the NASDAQ 100 constituents, (5) details on the exogenous variable, (6) more empirical results, and (7) more robustness check results.

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

A Proof of Theorem 1

The appendix provides the mathematical derivation of Theorem 1. For the original model (17), it can be rewritten as in the matrix notation

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{D}\boldsymbol{\alpha}_i + \psi_i \boldsymbol{y}_{i,-1} + \boldsymbol{X}_i \boldsymbol{\beta}_i + \boldsymbol{F}\boldsymbol{\lambda}_i + \boldsymbol{\epsilon}_i \\ &= \boldsymbol{D}\boldsymbol{\alpha}_i + \Xi_i \boldsymbol{\theta}_i + \boldsymbol{F}\boldsymbol{\lambda}_i + \boldsymbol{\epsilon}_i, \end{aligned} \tag{A1}$$

where $\boldsymbol{y}_i$, $\boldsymbol{D}$, $\boldsymbol{y}_{i,-1}$, $\boldsymbol{X}_i$, $\boldsymbol{F}$, Ξ_i, and all the coefficients are properly defined in the main text. Using the augmented regression (22) and following Chudik and Pesaran (2015), the above equation (A1) can also be rewritten as

$$\boldsymbol{y}_i = \Xi_i \boldsymbol{\theta}_i + \bar{\boldsymbol{Q}}\delta_i + \boldsymbol{\epsilon}_i + \boldsymbol{\eta}_i + \boldsymbol{\xi}_i, \tag{A2}$$

where $\bar{\boldsymbol{Q}}$ is defined in (23), $\delta_i = \left(\boldsymbol{\alpha}_i^{*\prime}, \delta_{i0}^{\prime}, \ldots, \delta_{ip_T}^{\prime}\right)^{\prime}$, $\boldsymbol{\eta}_i$ is $(T - p_T) \times 1$ vector with its elements given by $\sum_{l=p_T+1}^{\infty} \delta_{il}^{\prime} \bar{\boldsymbol{z}}_{t-l}$, and

$$\begin{aligned} \boldsymbol{\xi}_i &= \boldsymbol{D}\boldsymbol{\alpha}_i + \boldsymbol{F}\boldsymbol{\lambda}_i - \bar{\boldsymbol{Q}}\delta_i - \boldsymbol{\eta}_i \\ &= \boldsymbol{F}\boldsymbol{\lambda}_i - \bar{\boldsymbol{Z}}\delta_i(L) - \boldsymbol{D}\delta_i^{\prime}(1)\,\overline{\bar{\boldsymbol{B}}}, \end{aligned}$$

where $\bar{\boldsymbol{Z}} = \left(\bar{\boldsymbol{z}}_{p_T+1}, \ldots, \bar{\boldsymbol{z}}_T\right)^{\prime}$ and $\overline{\bar{\boldsymbol{B}}} = \frac{1}{N}\sum_{i=1}^{N}\sum_{l=0}^{\infty} \Phi_{il} \bar{\boldsymbol{B}}_i$ with $\bar{\boldsymbol{B}}_i = \boldsymbol{A}_{0i}^{-1}\boldsymbol{B}_i$. Note that $\boldsymbol{A}_{0i}$ and $\boldsymbol{B}_i$ are defined in equation (16).

Define

$$\boldsymbol{M}_H = \boldsymbol{I}_{I-p_T} - \boldsymbol{H}\left(\boldsymbol{H}^{\prime}\boldsymbol{H}\right)^{+}\boldsymbol{H}^{\prime},$$

with

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{d}_{i,p_T+1}^{\prime} & \bar{\boldsymbol{h}}_{p_T+1}^{\prime} & \bar{\boldsymbol{h}}_{p_T}^{\prime} & \cdots & \bar{\boldsymbol{h}}_1^{\prime} \\ \boldsymbol{d}_{i,p_T+2}^{\prime} & \bar{\boldsymbol{h}}_{p_T+2}^{\prime} & \bar{\boldsymbol{h}}_{p_T+1}^{\prime} & \cdots & \bar{\boldsymbol{h}}_2^{\prime} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{d}_{iT}^{\prime} & \bar{\boldsymbol{h}}_T^{\prime} & \bar{\boldsymbol{h}}_{T-1}^{\prime} & \cdots & \bar{\boldsymbol{h}}_{T-p_T}^{\prime} \end{pmatrix},$$

where $\bar{\boldsymbol{h}}_t = \Lambda(\mathbb{L})\boldsymbol{f}_t + \bar{\boldsymbol{d}}_t$ with $\Lambda(\mathbb{L}) = \frac{1}{N}\sum_{i=1}^{N}\Phi_i(\mathbb{L})\Gamma_i$ and $\bar{\boldsymbol{d}}_t = \frac{1}{N}\sum_{i=1}^{N}\Phi_i(\mathbb{L})\boldsymbol{B}_i\boldsymbol{d}_t$.

Substituting (A2) into (24) and noticing that $\boldsymbol{M}_{\bar{Q}}\bar{\boldsymbol{Q}} = \boldsymbol{0}$ yields

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{i,CCE} &= \left(\Xi_i^{\prime}\boldsymbol{M}_{\bar{Q}}\Xi_i\right)^{-1}\Xi_i^{\prime}\boldsymbol{M}_{\bar{Q}}\left(\Xi_i\boldsymbol{\theta}_i + \boldsymbol{\epsilon}_i + \boldsymbol{\eta}_i + \boldsymbol{\xi}_i\right) \\ &= \boldsymbol{\theta}_i + \left(\frac{1}{T^2}\Xi_i^{\prime}\boldsymbol{M}_{\bar{Q}}\Xi_i\right)^{-1}\frac{1}{T^2}\Xi_i^{\prime}\boldsymbol{M}_{\bar{Q}}\boldsymbol{\epsilon}_i + \left(\frac{1}{T^2}\Xi_i^{\prime}\boldsymbol{M}_{\bar{Q}}\Xi_i\right)^{-1}\frac{1}{T^2}\Xi_i^{\prime}\boldsymbol{M}_{\bar{Q}}\boldsymbol{\eta}_i \\ &\quad + \left(\frac{1}{T^2}\Xi_i^{\prime}\boldsymbol{M}_{\bar{Q}}\Xi_i\right)^{-1}\frac{1}{T^2}\Xi_i^{\prime}\boldsymbol{M}_{\bar{Q}}\boldsymbol{\xi}_i. \end{aligned} \tag{A3}$$

Under the asymptotics that $(N, T) \to \infty$ and $p_T = O\left(T^{1/3}\right)$, by using the standard asymptotic analysis in Chudik and Pesaran (2015) (P411),[1] it can be shown that for $i = 1, 2, \ldots, N$,

$$\begin{aligned} \frac{1}{T^2}\Xi_i' \boldsymbol{M}_{\bar{Q}} \Xi_i &= \frac{1}{T^2}\Xi_i' \boldsymbol{M}_H \Xi_i + O_p\left(\frac{1}{\sqrt{N}}\right) \\ &\to_p \Sigma_{i,H}, \end{aligned} \tag{A4}$$

where $\Sigma_{i,H} = \text{plim}_{T\to\infty} \frac{1}{T^2}\Xi_i' \boldsymbol{M}_H \Xi_i$ is a nonstochastic matrix with full rank. Similarly, it is straightforward to show that

$$\frac{1}{T^2}\Xi_i' \boldsymbol{M}_H \boldsymbol{\epsilon}_i \to_p 0, \tag{A5}$$

$$\frac{1}{T^2}\Xi_i' \boldsymbol{M}_H \boldsymbol{\xi}_i \to_p 0, \tag{A6}$$

$$\frac{1}{T^2}\Xi_i' \boldsymbol{M}_H \boldsymbol{\eta}_i \to_p 0. \tag{A7}$$

Consequently, substituting (A4)-(A7) into (A3), we obtain

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

as $(N, T) \to \infty$ and $p_T = O\left(T^{1/3}\right)$.

B Simulation studies

The main objective is to examine the finite sample performance of the CCE estimator in our framework. Since the proposed model can be treated as a panel retricted AR model, the data are generated as

$$y_{it} = c_{yi} + \phi_i^{(1)} \bar{y}_{it}^{(1)} + \phi_i^{(2)} \bar{y}_{it}^{(2)} + x_{it}\beta_i + \gamma_{i1} f_{1t} + \gamma_{i2} f_{2t} + \epsilon_{it}, \tag{A8}$$

and

$$x_{it} = c_{xi} + c_{1i} y_{it-1} + c_{2i} y_{it-2} + \gamma_{x1i} f_{1t} + \gamma_{x2i} f_{2t} + v_{it}, \tag{A9}$$

where $i = 1, 2, \ldots, N$ denote the cross-section units, $t = -99, \ldots, 0, 1, \ldots, T$ denotes time, and the first 100 observations are discarded as burn-in replications. Equation (A8) is a HAR-type model with the maximum lag order set to 2 ($L = 2$). We note that it can also be rewritten as

$$y_{it} = c_{yi} + \psi_{1i} y_{it-1} + \psi_{2i} y_{it-2} + x_{it}\beta_i + \gamma_{i1} f_{1t} + \gamma_{i2} f_{2t} + \epsilon_{it}, \tag{A10}$$

[1] The order of the limit is of T^2 is simply because $\boldsymbol{f}_t$ is nonstationary since $\sum_{t=1}^{T}\sum_{s=1}^{T} E\left(\boldsymbol{f}_t \boldsymbol{f}_s'\right) = O\left(T^2\right)$, and because $\Xi_i, \boldsymbol{H}, \bar{\boldsymbol{Q}}$ are function of $\boldsymbol{f}_t$.

with the following restrictions on the parameters $\psi_{1i} = \phi_i^{(1)} + \frac{\phi_i^{(2)}}{2}$ and $\psi_{2i} = \frac{\phi_i^{(2)}}{2}$.

In the above models (A8) and (A9), the fixed effects are generated as $c_{yi} \sim IIDN\,(0,1)$ and $c_{xi} \sim IIDN\,(0,1)$. All factor loadings are generated independently as $\gamma_{i1}, \gamma_{i2} \sim IIDU\,(0,0.5)$ and $\gamma_{x1i}, \gamma_{x2i} \sim IIDU\,(0,1)$, where the rank condition in Assumption 4 is met. For simplicity, we consider homogeneous coefficients across the panels and let $\phi^{(1)} = 0.3$, $\phi^{(2)} = 0.6$, and $\beta = 2$.[2] The idiosyncratic errors ϵ_{it} and v_{it} are drawn as $\sim IIDN\left(0, \sigma_{\epsilon i}^2\right)$, with $\sigma_{\epsilon i}^2$ being random draws from $\chi^2\,(1)$ distribution.

Consistent with the previous set-ups about $\boldsymbol{f}_t$, we consider the generation of unobserved common factors follows a simple independent stationary AR(1) processes, where f_{1t} and f_{2t} are given by

$$\begin{aligned} f_{1t} &= \rho_{f1l} f_{1t-1} + \eta_{1t}, \\ f_{2t} &= \rho_{f1l} f_{2t-1} + \eta_{2t},; \;\; t = 2, \ldots, T, \end{aligned}$$

for $\rho_{f1l} = 0.5$ and $\rho_{f2l} = 0.8$. The error terms η_{1t} and η_{2t} are generated as $\sim IIDN\,(0,1)$.

We consider the following combinations of (N,T) as $N = 20, 100$ and T varies from 100 to 1000. We set p_T equal to the integer part of $T^{1/3}$, designated as $p_T = \lfloor T^{1/3} \rfloor$ in the experiment. The focus is on the estimates of the average parameter values $\beta = E\,(\beta_i)$, $\phi^{(1)} = E\left(\phi_i^{(1)}\right)$ and $\phi^{(2)} = E\left(\phi_i^{(2)}\right)$, resulting from the CCE estimation. We report the bias and the root mean square error (RMSE) of the coefficient estimates from 1000 replications. The results are presented in Figure A1. Subplots (a) to (d) correspond to the results for the bias under $N = 20$, the RMSE under $N = 20$, the bias under $N = 100$, and the RMSE under $N = 100$, respectively. For each subplot, the results are plotted against the varying sample size T. The solid line, the dash-dotted line, and the dash line, correspond to β, $\phi^{(1)}$, and $\phi^{(2)}$, respectively.

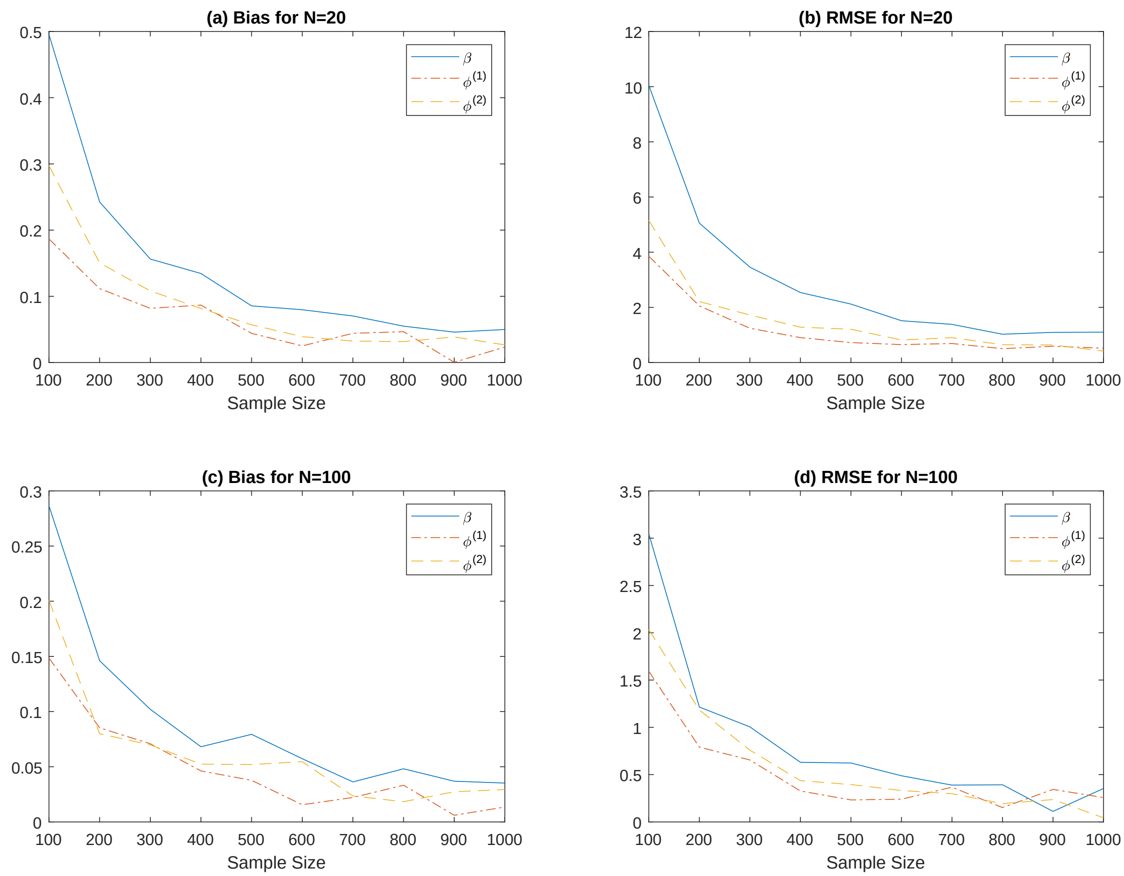
Looking across the rows with the same values of N, we find that both the bias and the RMSE decrease towards zero as T increases. When we compare the subplots with different values of N, both the bias and the RMSE decline as N increases. The simulation results imply that the CCE estimator is consistent under the HARP framework.

C Random forest

The building block of random forest is called the regression tree (RT) proposed by Breiman, Friedman, and Stone (1984). Note that the full name of the method is Classification and Regression Trees (CART), in which Classification mostly deals with the categorical response of non-numeric symbols and texts and Regression Trees concentrate purely on quantitative

[2]Note that the eigenvalues of ψ_{1i} and ψ_{2i} must be less than one, so as to satisfy the condition in Assumption 3. Under this specification, the DGP (A10) is stationary.

Figure A1: Simulation Results



responses variables. Given the numerical nature of our data set, we only consider the second part of CART.

The trick in applying the RT is to find the best split. Consider a sample of $\{y_t, x_t\}_{t=1}^{n}$. A simple regression will yield a sum of squared residuals, SSR_0. Suppose we can split the original sample into two sub-samples such that $n = n_1 + n_2$. The RT method finds the best split of a sample to minimize the sum of squared residuals (SSR) from the two sub-samples. That is, the SSR values computed from each sub-sample should follow: $\text{SSR}_1 + \text{SSR}_2 \leq \text{SSR}_0$. We can continue splitting until we reach a pre-determined boundary.

Random forest by Breiman (2001) involves a training process based on the RT where the level of training is predetermined. The random forest algorithm is summarized as below:

(i) Take a random sample with replacement from the data.

(ii) Construct a regression tree. As each tree is constructed, we take a random sample (without replacement) of q predictors out of the total K $(q < K)$ predictors before each node is split.

(iii) Use the regression tree to make forecasts, $\hat{f}$.

(iv) Repeat steps (i) to (iii), $b = 1, ..., B$ times and obtain $\hat{f}^b$ for each b.

(v) Take a simple average of the B forecasts $\hat{f}_{\text{BAG}} = \frac{1}{B}\sum_{b=1}^{B}\hat{f}^b$ and consider the averaged value $\hat{f}_{\text{BAG}}$ as the final forecast.

For most of the part, the more bootstrap samples in the training process, the better the forecast accuracy. However, more bootstrap samples means longer computational time. A balance needs to be found between accuracy, and time costs and constraints.

The above algorithm is usually executed for the cross-sectional data. When the data is time series having dependent observations, we need to replace step (i) with specific bootstrap methods for time series based on different assumptions. A straightforward way is to bootstrap the residuals instead of observations. For observations following a stationary Markov chain with finite state space, Kulperger and Rao (1989) initiated the Markov bootstrap method. If we are not willing to assume a specific structural form for time series (e.g., stationary and weakly dependent), we can use the moving block bootstrap (MBB) method formulated by Künsch (1989). Instead of performing single-data resampling, Künsch (1989) advocated the idea of resampling blocks of observations at a time. By retaining the neighboring observations together within each block, the dependence structure of the random variable at short lag distances is preserved. See Kreiss and Lahiri (2012) for a detailed literature review.

In this paper, we use block bootstrap to construct $B = 100$ samples. The input variables are the 22 lagged periods of $\hat{\boldsymbol{f}}_t$. We follow the conventional set-up and randomly use (roughly) one third of the total inputs variables for each split, that is $q = \lfloor K/3 \rfloor$.

D Details of the NASDAQ 100 constituents

In the empirical exercise, we include the individual stocks that composed of the NASDAQ100 index to obtain unobserved common factors $\boldsymbol{f}_t$. The NASDAQ100 Index includes 104 of the largest domestic and international non-financial companies listed on the NASDAQ Stock Market based on market capitalization. The Index reflects companies across major industry groups including computer hardware and software, telecommunications, retail/wholesale trade and biotechnology. Their NASDAQ tickers, associated company names, and industries are listed in the second, third, and fourth columns of Table A1, respectively.

Table A1: Descriptions of 104 component stocks

No.	Ticker Symbol	Company Name	GICS Sector
1	ATVI*	Activision Blizzard	Information Technology
2	ADBE*	Adobe Systems Incorporated	Information Technology
3	ALXN*	Alexion Pharmaceuticals	Health Care
4	ALGN*	Align Technology, Inc	Health Care
5	GOOGL*	Alphabet Inc. Class A	Information Technology
6	GOOG	Alphabet Inc. Class C	Information Technology
7	AMZN*	Amazon.com, Inc	Consumer Discretionary
8	AAL*	American Airlines Group	Industrials
9	AMGN	Amgen Inc	Health Care
10	ADI*	Analog Devices	Information Technology
11	AAPL*	Apple Inc	Information Technology
12	AMAT*	Applied Materials, Inc	Information Technology
13	ASML	ASML Holding	Information Technology
14	ADSK*	Autodesk, Inc	Information Technology
15	ADP*	Automatic Data Processing, Inc	Information Technology
16	BIDU*	Baidu.com, Inc	Information Technology
17	BIIB*	Biogen, Inc	Health Care
18	BMRN*	BioMarin Pharmaceutical, Inc	Health Care
19	AVGO	Broadcom Limited	Information Technology
20	CA*	CA Technologies	Information Technology
21	CDNS	Cadence Design Systems, Inc.	Information Technology
22	CELG*	Celgene Corporation	Health Care
23	CERN*	Cerner Corporation	Health Care
24	CHTR*	Charter Communications, Inc	Consumer Discretionary
25	CHKP*	Check Point Software Technologies Ltd.	Information Technology
26	CTAS*	Cintas Corporation	Industrials
27	CSCO*	Cisco Systems, Inc	Information Technology
28	CTXS*	Citrix Systems, Inc	Information Technology
29	CTSH*	Cognizant Technology Solutions Corporation	Information Technology
30	CMCSA*	Comcast Corporation	Consumer Discretionary
31	COST*	Costco Wholesale Corporation	Consumer Staples
32	CSX*	CSX Corporation	Industrials
33	CTRP*	CTrip International	Consumer Discretionary
34	XRAY	Dentsply Sirona	Health Care
35	DISCA*	Discovery Communications, Inc.	Consumer Discretionary
36	DLTR*	Dollar Tree, Inc.	Consumer Discretionary
37	EBAY*	eBay Inc.	Information Technology
38	EA*	Electronic Arts	Information Technology
39	EXPE*	Expedia, Inc.	Consumer Discretionary
40	ESRX*	Express Scripts, Inc.	Health Care
41	FB*	Facebook, Inc.	Information Technology
42	FAST*	Fastenal Company	Industrials
43	FISV*	Fiserv, Inc.	Information Technology
44	GILD*	Gilead Sciences, Inc.	Health Care
45	HAS*	Hasbro, Inc.	Consumer Discretionary
46	HSIC*	Henry Schein, Inc.	Health Care
47	HOLX*	Hologic, Inc.	Health Care
48	IDXX*	IDEXX Laboratories, Inc.	Health Care
49	ILMN*	Illumina, Inc.	Health Care
50	INCY*	Incyte Corporation	Health Care
51	INTC*	Intel Corporation	Information Technology
52	INTU*	Intuit, Inc.	Information Technology

Ticker with an asterisk indicates that the component is included in our estimation.

Table A1: Descriptions of 104 component stocks (continued)

No.	Ticker Symbol	Company Name	GICS Sector
53	ISRG*	Intuitive Surgical Inc.	Health Care
54	JBHT*	J.B. Hunt Transport Services, Inc.	Industrials
55	JD	JD.com	Consumer Discretionary
56	KLAC*	KLA-Tencor Corporation	Information Technology
57	LRCX*	Lam Research, Inc.	Information Technology
58	LBTYA*	Liberty Global plc Ordinary A	Consumer Discretionary
59	LBTYK*	Liberty Global plc Ordinary C	Consumer Discretionary
60	LVNTA	Liberty Interactive	Consumer Discretionary
61	QVCA*	Liberty Interactive	Consumer Discretionary
62	MAR*	Marriott International, Inc.	Consumer Discretionary
63	MXIM	Maxim Integrated Products	Information Technology
64	MELI	MercadoLibre	Information Technology
65	MCHP*	Microchip Technology	Information Technology
66	MU*	Micron Technology, Inc.	Information Technology
67	MSFT*	Microsoft Corporation	Information Technology
68	MDLZ*	Mondelēz International	Consumer Staples
69	MNST*	Monster Beverage	Consumer Staples
70	MYL*	Mylan N.V.	Health Care
71	NTES*	NetEase, Inc.	Information Technology
72	NFLX*	Netflix	Information Technology
73	NVDA*	NVIDIA Corporation	Information Technology
74	ORLY*	O'Reilly Automotive, Inc.	Consumer Discretionary
75	PCAR*	PACCAR Inc.	Industrials
76	PAYX*	Paychex, Inc.	Information Technology
77	PYPL*	PayPal Holdings, Inc.	Information Technology
78	QCOM*	QUALCOMM Incorporated	Information Technology
79	REGN*	Regeneron Pharmaceuticals	Health Care
80	ROST*	Ross Stores Inc.	Consumer Discretionary
81	STX*	Seagate Technology Holdings	Information Technology
82	SHPG*	Shire plc	Health Care
83	SIRI*	Sirius XM Radio, Inc.	Consumer Discretionary
84	SWKS*	Skyworks Solutions, Inc.	Information Technology
85	SBUX*	Starbucks Corporation	Consumer Discretionary
86	SYMC*	Symantec Corporation	Information Technology
87	SNPS*	Synopsys, Inc.	Information Technology
88	TMUS*	T-Mobile US	Telecommunication Services
89	TTWO*	Take-Two Interactive, Inc.	Information Technology
90	TSLA	Tesla, Inc.	Consumer Discretionary
91	TXN*	Texas Instruments, Inc.	Information Technology
92	KHC	The Kraft Heinz Company	Consumer Staples
93	PCLN*	The Priceline Group	Consumer Discretionary
94	FOXA*	Twenty-First Century Fox Class A	Consumer Discretionary
95	FOX*	Twenty-First Century Fox Class B	Consumer Discretionary
96	ULTA	Ulta Beauty	Consumer Discretionary
97	VRSK	Verisk Analytics	Industrials
98	VRTX*	Vertex Pharmaceuticals	Health Care
99	VOD*	Vodafone Group, plc.	Telecommunication Services
100	WBA*	Walgreens Boots Alliance	Consumer Staples
101	WDC*	Western Digital	Information Technology
102	WDAY	Workday, Inc.	Information Technology
103	WYNN*	Wynn Resorts	Consumer Discretionary
104	XLNX*	Xilinx, Inc.	Information Technology

Ticker with an asterisk indicates that the component is included in our estimation.

E Discussions on the exogenous variable z_t

As defined in equations (10), the main prediction equation of the RV include observed common effects $\boldsymbol{d}_t$. In our exercises, the exogenous variables $\boldsymbol{d}_t$ consist of the following five variables:

(i) OIL: The logarithm of one-month crude oil future contract index;

(ii) USDI: The logarithm of trade-weighted average of the foreign exchange value of the US dollar index;

(iii) CS: Credit Spread, the excess yield of the Moody's seasoned Baa corporate bond over the Moody's seasoned Aaa corporate bond;

(iv) TS: Term Spread, the difference between the 10-Year and 3-month treasury constant maturity rates;

(v) FFD: Federal Fund Deviation, the difference between the effective and target Federal Fund rates.

Descriptive statistics of $\boldsymbol{d}_t$ are presented in Table A2. We report the logarithm of one-month crude oil future contract index, and trade-weighted average of the foreign exchange value of the US dollar index, instead of their observed values.

Table A2: Descriptive statistics for observed common factors

Statistics	Oil	USDI	CS	TS	FFD
Mean	4.2878	4.3721	1.1540	2.1856	0.0211
Median	4.3886	4.3364	1.0100	2.1400	-0.0050
Maximum	4.9821	4.5737	3.5000	3.8300	0.9700
Minimum	3.2869	4.2198	0.5300	-0.1800	-1.3300
Std. Dev.	0.3400	0.0962	0.5392	0.7973	0.1260
Skewness	-0.5168	0.5994	2.3585	-0.2107	-3.8771
Kurtosis	2.1868	1.9600	8.9998	2.9317	31.7740
Jarque-Bera	0.0000	0.0000	0.0000	0.0000	0.0000
ADF Test	0.4864	0.4407	0.7668	0.0348	0.0000
*ADF Test on Δ^1	0.0000	0.0000	0.0000	0.0000	0.0000

$*$ We also perform ADF test on the first difference (Δ^1) of each series.

As indicated by the p-values of the ADF test in Table A2, all the series except the term spread and the federal fund deviation are nonstationary. For these nonstationary data series, we use either first differenced or log differenced series in the exercises. As reported in the last row of Table A2, all series after proper transformation are stationary. Moreover, the results of the Jarque-Bera test imply that none of the above series is normally distributed.

F More Empirical Results

In this section, we present supplemental results for the main text. Figure A2 presents sample autocorrelation analysis of unobserved common factors $\boldsymbol{f}_t$.

G More Robustness Check Results

G.1 Other industry indices

In this section, we consider applying the HARP method to two other industry indices, the Dow Jones Transportation Average (DJTA) and the Dow Jones Utility Average (DJUA). The DJTA index is an sector-specific index and much older than the DJIA. It tracks the stock prices of twenty transportation corporations. The DJUA index is a stock index derived from Dow Jones Indexes, that keeps track of the performance of 15 prominent U.S. utility companies. Data at 1-minute sampling interval is provided by Pitrading Incorporation and covers the same period as the main results. The two indexes are computed from the underlying group 15 and 20 stocks, respectively. All forecasts are generated using rolling window regressions based on 1,000 observations, and parameter estimates are updated daily.

The results for the two indices are reported in Tables A3 to A6. For the DJUA, the HARP forecast is always the winner and significantly improves the out-of-sample forecast performance. For the DJTA in Table A3, the HARP model outperforms for all but the longest horizon, where it is surpassed by the HAR-CJ model. Nevertheless, the GW test in Table A4 implies that the HARP forecast is not significantly out-performed by the HAR-CJ forecast. Overall, the above conclusions are in agreement with the results for the NASDAQ 100 ETF.

G.2 Different rolling window length

We consider a different rolling window length $U = 500$. The outcomes for the out-of sample comparison are reported in Table A7 and the GW test results are contained in Table A8. It can be seen that the performance of the HARP model still exceeds that of peer models and the improvement is significant in most cases. Even in the exception of $h = 22$, where the HARP model has higher QLIKE and MAFE losses than the HAR-RS-II model, the dominance is not statistically significant.

Figure A2: Sample autocorrelation analysis of unobserved common factors

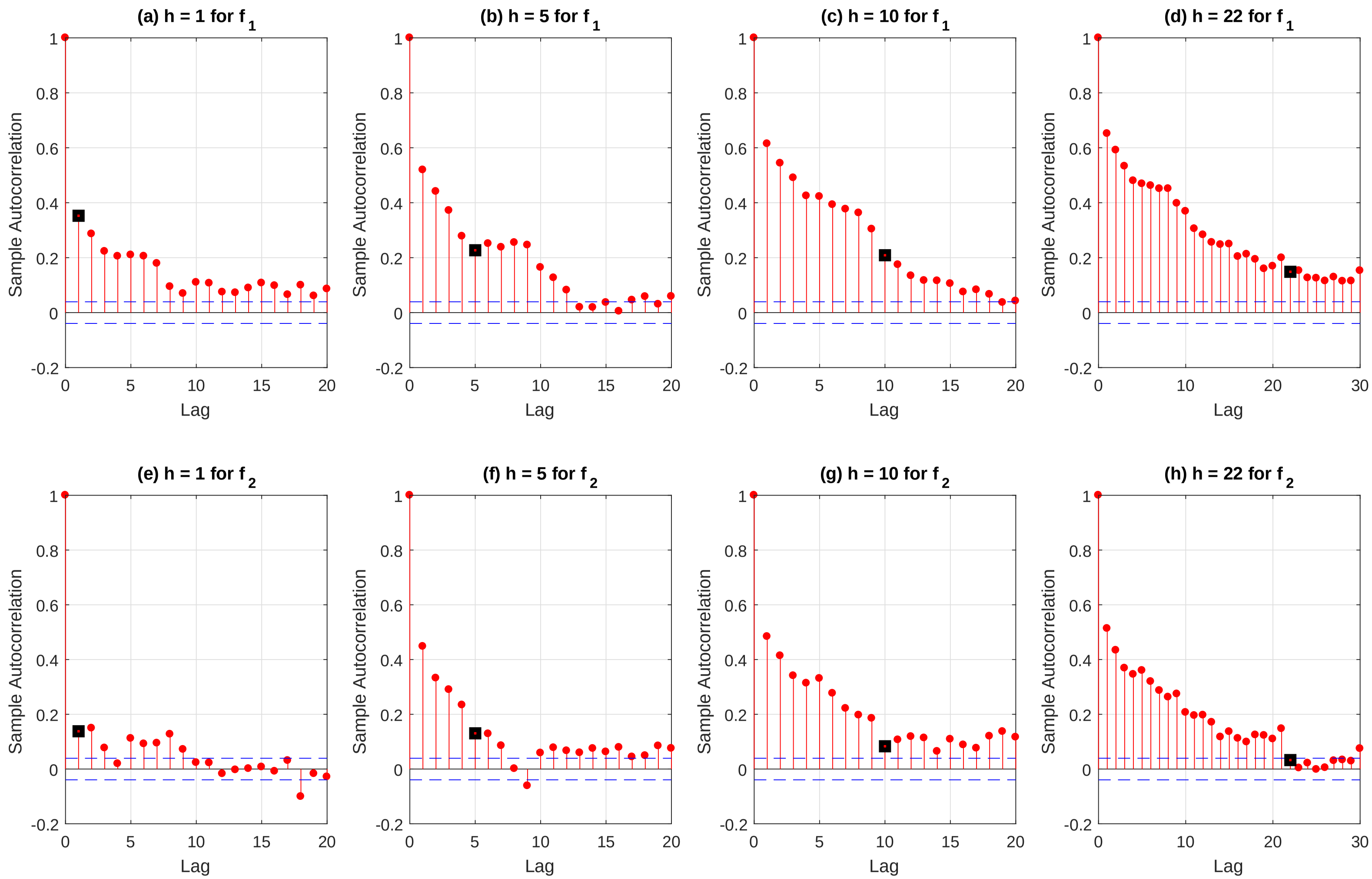


Table A3: Out-of-sample forecast comparison of models on the DJTA

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
AR(22)	2.3917	0.0867	1.2233	1.5465	0.8527
HAR	2.3083	0.0856	1.2097	1.5193	0.8578
HAR-J	2.3291	0.0857	1.2174	1.5261	0.8566
HAR-CJ	2.2123	0.0831	1.1783	1.4874	0.8637
HAR-RS-I	2.2975	0.0853	1.2067	1.5158	0.8585
HAR-RS-II	2.2995	0.0849	1.2036	1.5164	0.8584
HAR-SJ-I	2.3179	0.0857	1.2152	1.5225	0.8572
HAR-SJ-II	2.3291	0.0859	1.2173	1.5261	0.8566
HARP	**2.1314**	**0.0800**	**1.1524**	**1.4599**	**0.8687**
Panel B: $h = 5$					
AR(22)	3.4155	0.1122	1.4674	1.8481	0.7895
HAR	3.2081	0.1090	1.4363	1.7911	0.8022
HAR-J	3.1848	0.1082	1.4265	1.7846	0.8037
HAR-CJ	2.9741	0.1048	1.4314	1.7245	0.8167
HAR-RS-I	3.2147	0.1092	1.4383	1.7929	0.8018
HAR-RS-II	3.2034	0.1089	1.4353	1.7898	0.8025
HAR-SJ-I	3.2262	0.1093	1.4401	1.7962	0.8011
HAR-SJ-II	3.2250	0.1091	1.4370	1.7958	0.8012
HARP	**2.8719**	**0.1020**	**1.3422**	**1.6947**	**0.8230**
Panel C: $h = 10$					
AR(22)	3.6647	0.1181	1.5235	1.9143	0.7710
HAR	3.4681	0.1157	1.4928	1.8623	0.7833
HAR-J	3.4369	0.1150	1.4787	1.8539	0.7853
HAR-CJ	3.2592	0.1117	1.4340	1.8053	0.7964
HAR-RS-I	3.4902	0.1170	1.4952	1.8682	0.7819
HAR-RS-II	3.4673	0.1159	1.4888	1.8621	0.7834
HAR-SJ-I	3.4896	0.1169	1.4947	1.8680	0.7820
HAR-SJ-II	3.4995	0.1194	1.4949	1.8707	0.7813
HARP	**3.2110**	**0.1092**	**1.3504**	**1.7696**	**0.8031**
Panel D: $h = 22$					
AR(22)	3.7752	0.1213	1.5326	1.9430	0.7591
HAR	3.5811	0.1196	1.5006	1.8924	0.7715
HAR-J	3.5072	0.1180	1.4756	1.8728	0.7762
HAR-CJ	**3.3557**	**0.1161**	**1.4457**	**1.8318**	**0.7859**
HAR-RS-I	3.5838	0.1197	1.5011	1.8931	0.7713
HAR-RS-II	3.5867	0.1199	1.5001	1.8939	0.7711
HAR-SJ-I	3.5842	0.1197	1.5004	1.8932	0.7713
HAR-SJ-II	3.5894	0.1199	1.5010	1.8946	0.7710
HARP	3.4060	0.1235	1.4803	1.8455	0.7827

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 A3 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 A4: The Giacomini-White test for the mean absolute forecast errors-the DJTA

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	0.6922	-	-	-	-	-	-	-
HAR-J	0.7749	0.2798	-	-	-	-	-	-
HAR-CJ	0.0288	0.0012	0.0049	-	-	-	-	-
HAR-RS-I	0.5051	0.1475	0.1726	0.0007	-	-	-	-
HAR-RS-II	0.3444	0.3687	0.1392	0.0008	0.5204	-	-	-
HAR-SJ-I	0.5547	0.6407	0.3935	0.0023	0.9150	0.6455	-	-
HAR-SJ-II	0.4975	0.5508	0.3597	0.0018	0.7901	0.7082	0.6124	-
HARP	**0.0015**	**0.0023**	**0.0021**	**0.0004**	**0.0031**	**0.0031**	**0.0036**	**0.0039**
Panel B: $h = 5$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.0891	-	-	-	-	-	-	-
HAR-J	0.0002	0.0012	-	-	-	-	-	-
HAR-CJ	0.0000	0.0000	0.0103	-	-	-	-	-
HAR-RS-I	0.0634	0.2489	0.0028	0.0000	-	-	-	-
HAR-RS-II	0.1330	0.7678	0.0086	0.0000	0.9558	-	-	-
HAR-SJ-I	0.1461	0.6206	0.0129	0.0000	0.8266	0.8362	-	-
HAR-SJ-II	0.1420	0.5758	0.0405	0.0004	0.6966	0.7446	0.7141	-
HARP	**0.0009**	**0.0003**	**0.0000**	**0.0000**	**0.0002**	**0.0006**	**0.0002**	**0.0000**
Panel C: $h = 10$	-							
AR(22)		-	-	-	-	-	-	-
HAR	0.2469	-	-	-	-	-	-	-
HAR-J	0.0051	0.0020	-	-	-	-	-	-
HAR-CJ	0.0002	0.0001	0.0042	-	-	-	-	-
HAR-RS-I	0.5180	0.1475	0.0009	0.0000	-	-	-	-
HAR-RS-II	0.5251	0.7368	0.0011	0.0000	0.7423	-	-	-
HAR-SJ-I	0.3566	0.9111	0.0055	0.0001	0.4638	0.7862	-	-
HAR-SJ-II	0.3573	0.9359	0.0101	0.0001	0.5765	0.8345	0.9822	-
HARP	**0.0373**	**0.0047**	**0.0005**	**0.0000**	**0.0062**	**0.0064**	**0.0035**	**0.0025**
Panel D: $h = 22$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.4313	-	-	-	-	-	-	-
HAR-J	0.0005	0.0029	-	-	-	-	-	-
HAR-CJ	0.0000	0.0000	0.0134	-	-	-	-	-
HAR-RS-I	0.5089	0.7391	0.0011	0.0000	-	-	-	-
HAR-RS-II	0.0172	0.0612	0.0076	0.0000	0.0045	-	-	-
HAR-SJ-I	0.3860	0.7154	0.0007	0.0000	0.5091	0.1118	-	-
HAR-SJ-II	0.5685	0.8848	0.0034	0.0000	0.7817	0.2349	0.8003	-
HARP	0.3042	0.6614	0.9135	0.3486	0.6516	0.6692	0.6642	0.6545

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 A4, respectively. Bold numbers indicate the null hypothesis can be rejected at 5% level of significance.

Table A5: Out-of-sample forecast comparison of models on the DJUA

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
AR(22)	1.8050	0.1600	0.7263	1.3435	0.5558
HAR	1.3840	0.1537	0.6963	1.1764	0.6594
HAR-J	1.5376	0.1522	0.6919	1.2400	0.6216
HAR-CJ	1.5280	0.1401	0.6470	1.2361	0.6239
HAR-RS-I	1.4039	0.1528	0.6948	1.1849	0.6545
HAR-RS-II	1.3633	0.1520	0.6911	1.1676	0.6645
HAR-SJ-I	1.4879	0.1530	0.6976	1.2198	0.6338
HAR-SJ-II	1.6251	0.1546	0.6951	1.2748	0.6000
HARP	**0.9025**	**0.1161**	**0.5739**	**0.9500**	**0.7779**
Panel B: $h = 5$					
AR(22)	1.8842	0.1972	0.7861	1.3727	0.5371
HAR	1.4272	0.1792	0.7593	1.1947	0.6493
HAR-J	1.6328	0.1748	0.7550	1.2778	0.5988
HAR-CJ	1.6274	0.1629	0.7103	1.2757	0.6002
HAR-RS-I	1.5180	0.1785	0.7631	1.2321	0.6271
HAR-RS-II	1.4539	0.1771	0.7568	1.2058	0.6428
HAR-SJ-I	1.6587	0.1779	0.7670	1.2879	0.5925
HAR-SJ-II	1.8553	0.1759	0.7651	1.3621	0.5442
HARP	**1.0741**	**0.1337**	**0.6265**	**1.0364**	**0.7361**
Panel C: $h = 10$					
AR(22)	2.3475	0.1259	0.6668	1.5322	0.4217
HAR	1.6403	0.1232	0.6277	1.2808	0.5959
HAR-J	2.3147	0.1200	0.6256	1.5214	0.4298
HAR-CJ	2.3098	0.1093	0.5835	1.5198	0.4310
HAR-RS-I	2.0190	0.1206	0.6271	1.4209	0.5026
HAR-RS-II	1.7546	0.1171	0.6074	1.3246	0.5677
HAR-SJ-I	2.4402	0.1203	0.6311	1.5621	0.3988
HAR-SJ-II	2.9313	0.1185	0.6270	1.7121	0.2779
HARP	**1.0734**	**0.1067**	**0.5211**	**1.0360**	**0.7356**
Panel D: $h = 22$					
AR(22)	1.7005	0.5322	0.8153	1.3040	0.5802
HAR	1.4691	0.1984	0.8026	1.2121	0.6373
HAR-J	1.4821	0.1963	0.7978	1.2174	0.6341
HAR-CJ	1.4516	0.1890	0.7718	1.2048	0.6416
HAR-RS-I	1.4818	0.1984	0.8035	1.2173	0.6342
HAR-RS-II	1.4847	0.1984	0.8031	1.2185	0.6335
HAR-SJ-I	1.4972	0.1983	0.8050	1.2236	0.6304
HAR-SJ-II	1.5252	0.1984	0.8050	1.2350	0.6235
HARP	**1.3842**	**0.1927**	**0.7172**	**1.1765**	**0.6583**

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 Utility 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 A5 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 A6: The Giacomini-White test for the mean absolute forecast errors-the DJUA

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	0.0045	-	-	-	-	-	-	-
HAR-J	0.0092	0.7681	-	-	-	-	-	-
HAR-CJ	0.0000	0.0000	0.0000	-	-	-	-	-
HAR-RS-I	0.0066	0.8907	0.6761	0.0000	-	-	-	-
HAR-RS-II	0.0001	0.0003	0.0146	0.0040	0.0003	-	-	-
HAR-SJ-I	0.0287	0.6886	0.0043	0.0000	0.3451	0.0037	-	-
HAR-SJ-II	0.0371	0.9491	0.8007	0.0000	0.9836	0.0981	0.4052	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel B: $h = 5$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.0481	-	-	-	-	-	-	-
HAR-J	0.0122	0.1801	-	-	-	-	-	-
HAR-CJ	0.0000	0.0000	0.0000	-	-	-	-	-
HAR-RS-I	0.0347	0.3125	0.3477	0.0000	-	-	-	-
HAR-RS-II	0.0227	0.0935	0.8711	0.0000	0.2375	-	-	-
HAR-SJ-I	0.0430	0.6063	0.0062	0.0000	0.0867	0.1005	-	-
HAR-SJ-II	0.0205	0.7891	0.1507	0.0000	0.9589	0.4855	0.3478	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel C: $h = 10$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.0976	-	-	-	-	-	-	-
HAR-J	0.0245	0.4716	-	-	-	-	-	-
HAR-CJ	0.0000	0.0000	0.0000	-	-	-	-	-
HAR-RS-I	0.1207	0.2080	0.0285	0.0000	-	-	-	-
HAR-RS-II	0.0614	0.5284	0.7671	0.0000	0.0974	-	-	-
HAR-SJ-I	0.1696	0.1772	0.0000	0.0000	0.1756	0.0681	-	-
HAR-SJ-II	0.1203	0.5240	0.0044	0.0000	0.7492	0.3190	0.5925	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel D: $h = 22$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.3673	-	-	-	-	-	-	-
HAR-J	0.1685	0.0253	-	-	-	-	-	-
HAR-CJ	0.0020	0.0000	0.0001	-	-	-	-	-
HAR-RS-I	0.3771	0.4783	0.0005	0.0000	-	-	-	-
HAR-RS-II	0.3584	0.8184	0.0111	0.0000	0.7998	-	-	-
HAR-SJ-I	0.4238	0.1763	0.0000	0.0000	0.0577	0.2312	-	-
HAR-SJ-II	0.3877	0.4555	0.0000	0.0000	0.4882	0.4232	0.9749	-
HARP	**0.0001**	**0.0005**	**0.0002**	**0.0011**	**0.0009**	**0.0007**	**0.0000**	**0.0009**

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 A6, respectively. Bold numbers indicate the null hypothesis can be rejected at 5% level of significance.

Table A7: Out-of-sample forecast comparison of models for RV of the NASDAQ 100 ETF with $U = 500$

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
AR(22)	0.5452	0.1707	0.3426	0.7384	0.4890
HAR	0.4205	0.1371	0.3342	0.6485	0.6059
HAR-J	0.3367	0.1299	0.3171	0.5802	0.6845
HAR-CJ	0.3321	0.1332	0.3172	0.5763	0.6887
HAR-RS-I	0.2539	0.1247	0.3047	0.5039	0.7620
HAR-RS-II	0.3263	0.1251	0.3049	0.5712	0.6942
HAR-SJ-I	0.2526	0.1734	0.3063	0.5026	0.7632
HAR-SJ-II	0.2562	0.1266	0.3062	0.5062	0.7599
HARP	**0.2374**	**0.1217**	**0.2863**	**0.4872**	**0.7775**
Panel B: $h = 5$					
AR(22)	0.5555	0.2562	0.4279	0.7453	0.4778
HAR	0.4679	0.2656	0.4138	0.6840	0.5602
HAR-J	0.4573	0.2391	0.4099	0.6763	0.5701
HAR-CJ	0.4354	0.4851	0.3985	0.6598	0.5907
HAR-RS-I	0.4516	0.2412	0.4016	0.6720	0.5755
HAR-RS-II	0.4302	0.2433	0.3983	0.6559	0.5956
HAR-SJ-I	0.4550	0.4001	0.4021	0.6745	0.5723
HAR-SJ-II	0.4878	0.2563	0.3999	0.6984	0.5414
HARP	**0.3419**	**0.1852**	**0.3536**	**0.5847**	**0.6786**
Panel C: $h = 10$					
AR(22)	0.6564	0.3331	0.4822	0.8102	0.3815
HAR	0.5423	0.2892	0.4562	0.7364	0.4890
HAR-J	0.5398	0.2929	0.4557	0.7347	0.4913
HAR-CJ	0.5324	0.3349	0.4475	0.7296	0.4983
HAR-RS-I	0.5438	0.2838	0.4508	0.7374	0.4876
HAR-RS-II	0.5251	0.2850	0.4471	0.7246	0.5052
HAR-SJ-I	0.5465	0.2840	0.4515	0.7392	0.4850
HAR-SJ-II	0.5953	0.2934	0.4523	0.7715	0.4390
HARP	**0.4705**	**0.2831**	**0.4175**	**0.6859**	**0.5566**
Panel D: $h = 22$					
AR(22)	0.8013	0.6595	0.5366	0.8952	0.2417
HAR	0.6857	1.0469	0.5098	0.8281	0.3511
HAR-J	0.6841	0.5663	0.5087	0.8271	0.3527
HAR-CJ	0.7025	0.6423	0.5150	0.8382	0.3352
HAR-RS-I	0.7022	0.4706	0.5110	0.8380	0.3355
HAR-RS-II	0.6932	**0.4653**	**0.5084**	0.8326	0.3440
HAR-SJ-I	0.7046	0.4669	0.5114	0.8394	0.3332
HAR-SJ-II	0.7101	0.5084	0.5099	0.8427	0.3280
HARP	**0.6554**	0.5367	0.5134	**0.8096**	**0.3798**

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 500 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 A7 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 A8: The Giacomini-White test for the mean absolute forecast errors-the NASDAQ 100 ETF with $U = 500$

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	0.2693	-	-	-	-	-	-	-
HAR-J	0.0022	0.0000	-	-	-	-	-	-
HAR-CJ	0.0026	0.0002	0.9731	-	-	-	-	-
HAR-RS-I	0.0003	0.0001	0.0117	0.0246	-	-	-	-
HAR-RS-II	0.0000	0.0000	0.0001	0.0036	0.9533	-	-	-
HAR-SJ-I	0.0009	0.0006	0.0560	0.0823	0.1016	0.7993	-	-
HAR-SJ-II	0.0006	0.0002	0.0364	0.0625	0.0739	0.7992	0.8868	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel B: $h = 5$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.0488	-	-	-	-	-	-	-
HAR-J	0.0169	0.0535	-	-	-	-	-	-
HAR-CJ	0.0004	0.0046	0.0241	-	-	-	-	-
HAR-RS-I	0.0008	0.0000	0.0002	0.5474	-	-	-	-
HAR-RS-II	0.0004	0.0004	0.0007	0.9737	0.2484	-	-	-
HAR-SJ-I	0.0009	0.0000	0.0004	0.4862	0.1197	0.2167	-	-
HAR-SJ-II	0.0005	0.0002	0.0091	0.8094	0.5700	0.7651	0.4291	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel C: $h = 10$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.0643	-	-	-	-	-	-	-
HAR-J	0.0656	0.7500	-	-	-	-	-	-
HAR-CJ	0.0133	0.0766	0.0882	-	-	-	-	-
HAR-RS-I	0.0278	0.0055	0.0183	0.5145	-	-	-	-
HAR-RS-II	0.0142	0.0005	0.0004	0.9429	0.0338	-	-	-
HAR-SJ-I	0.0315	0.0149	0.0409	0.4410	0.0175	0.0217	-	-
HAR-SJ-II	0.0309	0.3286	0.4754	0.4803	0.6489	0.2494	0.7973	-
HARP	**0.0003**	**0.0010**	**0.0014**	**0.0136**	**0.0041**	**0.0071**	**0.0036**	**0.0043**
Panel D: $h = 22$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.1448	-	-	-	-	-	-	-
HAR-J	0.1313	0.1820	-	-	-	-	-	-
HAR-CJ	0.2221	0.1462	0.0871	-	-	-	-	-
HAR-RS-I	0.1314	0.6326	0.4167	0.2970	-	-	-	-
HAR-RS-II	0.0951	0.6770	0.9355	0.1262	0.1048	-	-	-
HAR-SJ-I	0.1355	0.5515	0.3680	0.3492	0.1249	0.0697	-	-
HAR-SJ-II	0.1074	0.9628	0.7428	0.2524	0.3988	0.3994	0.2140	-
HARP	0.4160	0.8733	0.8361	0.9461	0.9159	0.8206	0.9293	0.8771

This table report the results of the modified Giacomini-White test (Giacomini and White, 2006) based on a window length of 500 observations. The GW test 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 A8, respectively. Bold numbers indicate the null hypothesis can be rejected at 5% level of significance.

G.3 Subsample evidence

In this section we partition the entire sample period in half: the first subsample period runs from May 22, 2007 to May 20, 2012, the second subsample period runs from May 21, 2012 to October 20, 2017. We carry out the similar out-of-sample analysis for the two subsamples in Tables A9-A12, respectively. The previous conclusions remain basically unchanged under the subsamples. The only exception happens to the second subsample period, when the HARP model does not significantly improves the out-of-sample forecast performance for $h = 10$ and 22, relative to some of semivariance-based models.

Table A9: Out-of-sample results for the first subsample period of the NASDAQ 100 ETF

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
AR(22)	0.8466	0.1659	0.6031	0.9201	0.7433
HAR	0.9111	0.1799	0.6391	0.9545	0.7237
HAR-J	0.9193	0.1763	0.6263	0.9588	0.7212
HAR-CJ	0.8873	0.1744	0.6211	0.9420	0.7310
HAR-RS-I	0.8408	0.1681	0.6086	0.9169	0.7451
HAR-RS-II	0.7501	0.1554	0.5682	0.8661	0.7725
HAR-SJ-I	0.8401	0.1692	0.6115	0.9166	0.7453
HAR-SJ-II	0.8424	0.1680	0.6081	0.9178	0.7446
HARP	**0.7410**	**0.1389**	**0.5501**	**0.8603**	**0.7783**
Panel B: $h = 5$					
AR(22)	1.5386	0.2896	0.8297	1.2404	0.5414
HAR	1.5358	0.2934	0.8403	1.2393	0.5422
HAR-J	1.5331	0.2923	0.8380	1.2382	0.5430
HAR-CJ	1.4840	0.2954	0.7944	1.2182	0.5577
HAR-RS-I	1.5201	0.2885	0.8228	1.2329	0.5469
HAR-RS-II	1.5289	0.2880	0.8188	1.2365	0.5443
HAR-SJ-I	1.5198	0.2884	0.8225	1.2328	0.5470
HAR-SJ-II	1.4759	0.2753	0.8027	1.2149	0.5601
HARP	**1.3180**	**0.2359**	**0.7400**	**1.1480**	**0.6072**
Panel C: $h = 10$					
AR(22)	1.7476	0.3348	0.9065	1.3220	0.4893
HAR	1.7278	0.3410	0.9224	1.3145	0.4951
HAR-J	1.7280	0.3409	0.9224	1.3145	0.4950
HAR-CJ	1.7023	0.3518	0.8832	1.3047	0.5025
HAR-RS-I	1.7286	0.3417	0.9130	1.3148	0.4948
HAR-RS-II	1.7494	0.3416	0.9124	1.3226	0.4887
HAR-SJ-I	1.7291	0.3420	0.9141	1.3149	0.4947
HAR-SJ-II	1.7040	0.3363	0.9010	1.3054	0.5020
HARP	**1.4072**	**0.2442**	**0.7335**	**1.1862**	**0.5887**
Panel D: $h = 22$					
AR(22)	2.1870	0.4318	1.1385	1.4788	0.3879
HAR	2.1938	0.4378	1.1546	1.4811	0.3860
HAR-J	2.1663	0.4343	1.1419	1.4718	0.3937
HAR-CJ	2.2136	0.4579	1.1426	1.4878	0.3804
HAR-RS-I	2.2031	0.4408	1.1570	1.4843	0.3833
HAR-RS-II	2.1858	0.4358	1.1467	1.4785	0.3882
HAR-SJ-I	2.2049	0.4409	1.1576	1.4849	0.3829
HAR-SJ-II	2.1831	0.4379	1.1442	1.4775	0.3889
HARP	**1.7227**	**0.3537**	**0.8601**	**1.3125**	**0.5178**

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 first half sample period of the NASDAQ 100 ETF spanning from May 22, 2007 to May 20, 2012 (a total of 1,313 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 A9 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 A10: The Giacomini-White test for the mean absolute forecast errors-the first subsample period of 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	0.0070	-	-	-	-	-	-	-
HAR-J	0.0947	0.0006	-	-	-	-	-	-
HAR-CJ	0.1960	0.0035	0.4243	-	-	-	-	-
HAR-RS-I	0.7147	0.0005	0.0616	0.2116	-	-	-	-
HAR-RS-II	0.0824	0.0000	0.0009	0.0030	0.0025	-	-	-
HAR-SJ-I	0.5743	0.0012	0.1193	0.3275	0.0089	0.0013	-	-
HAR-SJ-II	0.7348	0.0003	0.0527	0.1832	0.7530	0.0032	0.0034	-
HARP	**0.0016**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.2771	**0.0000**	**0.0000**
Panel B: $h = 5$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.2557	-	-	-	-	-	-	-
HAR-J	0.3647	0.0007	-	-	-	-	-	-
HAR-CJ	0.0157	0.0002	0.0004	-	-	-	-	-
HAR-RS-I	0.5679	0.0417	0.0738	0.0535	-	-	-	-
HAR-RS-II	0.3894	0.0332	0.0540	0.1116	0.3312	-	-	-
HAR-SJ-I	0.5515	0.0385	0.0684	0.0560	0.0002	0.3719	-	-
HAR-SJ-II	0.0336	0.0003	0.0006	0.6053	0.0020	0.0385	0.0022	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0026**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel C: $h = 10$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.1936	-	-	-	-	-	-	-
HAR-J	0.1891	0.9537	-	-	-	-	-	-
HAR-CJ	0.2029	0.0053	0.0060	-	-	-	-	-
HAR-RS-I	0.5973	0.1460	0.1533	0.0378	-	-	-	-
HAR-RS-II	0.5042	0.2711	0.2715	0.0774	0.9289	-	-	-
HAR-SJ-I	0.5366	0.1858	0.1910	0.0332	0.1524	0.8093	-	-
HAR-SJ-II	0.6048	0.0043	0.0049	0.2032	0.0043	0.0405	0.0041	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0001**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Panel D: $h = 22$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.2513	-	-	-	-	-	-	-
HAR-J	0.7785	0.0041	-	-	-	-	-	-
HAR-CJ	0.8408	0.4407	0.9584	-	-	-	-	-
HAR-RS-I	0.1291	0.6897	0.0046	0.3066	-	-	-	-
HAR-RS-II	0.1948	0.4841	0.6211	0.8301	0.2879	-	-	-
HAR-SJ-I	0.1181	0.6245	0.0040	0.2939	0.0648	0.2631	-	-
HAR-SJ-II	0.5836	0.2135	0.6582	0.9006	0.0115	0.7570	0.0095	-
HARP	**0.0001**	**0.0003**	**0.0003**	**0.0002**	**0.0001**	**0.0001**	**0.0001**	**0.0001**

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 A10, respectively. Bold numbers indicate the null hypothesis can be rejected at 5% level of significance.

Table A11: Out-of-sample results for the second subsample period of the NASDAQ 100 ETF

Method	MSFE	QLIKE	MAFE	SDFE	Pseudo R^2
Panel A: $h = 1$					
AR(22)	0.0962	0.2218	0.2042	0.3101	0.5248
HAR	0.0945	0.2189	0.2071	0.3074	0.5330
HAR-J	0.0847	0.1967	0.1882	0.2910	0.5814
HAR-CJ	0.0801	0.1958	0.1829	0.2831	0.6040
HAR-RS-I	0.0829	0.1686	0.1675	0.2879	0.5904
HAR-RS-II	0.0875	0.1736	0.1698	0.2958	0.5675
HAR-SJ-I	0.0824	0.1697	0.1687	0.2871	0.5927
HAR-SJ-II	0.0850	0.1711	0.1696	0.2916	0.5799
HARP	**0.0707**	**0.1622**	**0.1597**	**0.2659**	**0.6506**
Panel B: $h = 5$					
AR(22)	0.1113	0.2636	0.2266	0.3336	0.4417
HAR	0.1108	0.2617	0.2252	0.3329	0.4441
HAR-J	0.1134	0.2676	0.2212	0.3368	0.4309
HAR-CJ	0.1135	0.2723	0.2157	0.3369	0.4305
HAR-RS-I	0.1143	0.2572	0.2128	0.3381	0.4264
HAR-RS-II	0.1148	0.2569	0.2137	0.3388	0.4240
HAR-SJ-I	0.1143	0.2578	0.2139	0.3381	0.4266
HAR-SJ-II	0.1163	0.2593	0.2156	0.3411	0.4164
HARP	**0.0794**	**0.2138**	**0.1933**	**0.2817**	**0.6018**
Panel C: $h = 10$					
AR(22)	0.1162	0.2811	0.2289	0.3409	0.4183
HAR	0.1147	0.2794	0.2260	0.3387	0.4257
HAR-J	0.1180	0.2854	0.2251	0.3435	0.4093
HAR-CJ	0.1173	0.2815	0.2248	0.3424	0.4131
HAR-RS-I	0.1156	0.2738	0.2182	0.3400	0.4212
HAR-RS-II	0.1156	0.2737	0.2179	0.3399	0.4216
HAR-SJ-I	0.1161	0.2752	0.2192	0.3408	0.4187
HAR-SJ-II	0.1161	0.2754	0.2191	0.3407	0.4191
HARP	**0.0903**	**0.2519**	**0.2097**	**0.3005**	**0.5480**
Panel D: $h = 22$					
AR(22)	0.1228	0.5661	0.2343	0.3504	0.3980
HAR	0.1216	0.6142	0.2331	0.3487	0.4037
HAR-J	0.1218	0.5569	0.2323	0.3489	0.4029
HAR-CJ	0.1222	0.6243	0.2342	0.3495	0.4008
HAR-RS-I	0.1216	0.8873	**0.2304**	0.3486	0.4039
HAR-RS-II	0.1220	0.5910	0.2311	0.3492	0.4019
HAR-SJ-I	0.1216	0.9297	0.2307	0.3487	0.4038
HAR-SJ-II	0.1217	1.1937	0.2308	0.3489	0.4031
HARP	**0.1107**	**0.3821**	0.2364	**0.3327**	**0.4572**

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 second half sample period of the NASDAQ 100 ETF spanning from May 21, 2012 to October 20, 2017 (a total of 1,312 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 A11 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 A12: The Giacomini-White test for the mean absolute forecast errors-the second subsample period of 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	0.1996	-	-	-	-	-	-	-
HAR-J	0.0060	0.0012	-	-	-	-	-	-
HAR-CJ	0.0009	0.0001	0.0531	-	-	-	-	-
HAR-RS-I	0.0000	0.0000	0.0000	0.0007	-	-	-	-
HAR-RS-II	0.0001	0.0000	0.0005	0.0087	0.0974	-	-	-
HAR-SJ-I	0.0000	0.0000	0.0000	0.0005	0.1520	0.4763	-	-
HAR-SJ-II	0.0000	0.0000	0.0002	0.0046	0.0545	0.8461	0.3893	-
HARP	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0466**	**0.0344**	**0.0309**	**0.0322**
Panel B: $h = 5$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.2052	-	-	-	-	-	-	-
HAR-J	0.1455	0.3050	-	-	-	-	-	-
HAR-CJ	0.0275	0.0700	0.0136	-	-	-	-	-
HAR-RS-I	0.0534	0.0974	0.0666	0.4205	-	-	-	-
HAR-RS-II	0.0783	0.1318	0.1101	0.5981	0.1765	-	-	-
HAR-SJ-I	0.0618	0.1140	0.0767	0.5816	0.0420	0.8007	-	-
HAR-SJ-II	0.1325	0.2097	0.2356	0.9795	0.0011	0.0496	0.0891	-
HARP	**0.0000**	**0.0001**	**0.0004**	**0.0079**	**0.0357**	**0.0280**	**0.0227**	**0.0182**
Panel C: $h = 10$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.1036	-	-	-	-	-	-	-
HAR-J	0.2565	0.7530	-	-	-	-	-	-
HAR-CJ	0.3163	0.7499	0.8695	-	-	-	-	-
HAR-RS-I	0.0276	0.0863	0.0072	0.0163	-	-	-	-
HAR-RS-II	0.0256	0.0779	0.0067	0.0161	0.4802	-	-	-
HAR-SJ-I	0.0398	0.1225	0.0103	0.0308	0.0022	0.0188	-	-
HAR-SJ-II	0.0365	0.1137	0.0075	0.0253	0.0381	0.0370	0.4687	-
HARP	**0.0142**	**0.0252**	**0.0476**	**0.0466**	0.2955	0.3139	0.2412	0.2467
Panel D: $h = 22$								
AR(22)	-	-	-	-	-	-	-	-
HAR	0.1260	-	-	-	-	-	-	-
HAR-J	0.0078	0.3675	-	-	-	-	-	-
HAR-CJ	0.9752	0.4150	0.2106	-	-	-	-	-
HAR-RS-I	0.0713	0.2784	0.3021	0.1433	-	-	-	-
HAR-RS-II	0.1866	0.4706	0.5731	0.2674	0.3071	-	-	-
HAR-SJ-I	0.0561	0.2795	0.3015	0.1392	0.3120	0.6365	-	-
HAR-SJ-II	0.0556	0.2851	0.3059	0.1401	0.3140	0.7115	0.5225	-
HARP	0.8157	0.7046	0.6545	0.8036	0.5606	0.6153	0.5743	0.5790

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 A12, respectively. Bold numbers indicate the null hypothesis can be rejected at 5% level of significance.

References

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., 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.

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

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.