HARK the SHARK: Realized Volatility Modeling with Measurement Errors and Nonlinear Dependencies

Giuseppe Buccheri[a,*], Fulvio Corsi[b,†]

[a] *Scuola Normale Superiore*

[b] *University of Pisa and City University of London*

Online Appendix

*Corresponding author: giuseppe.buccheri@sns.it, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126, Pisa (Italy), tel. 050-509091

†fulvio.corsi@gmail.com.

A Simulation study

The linear DGP used in Section (2.3) does note provide a good description of empirical data, since realized volatility exhibits several evidences of nonlinear dependencies. In order to perform a more realistic analysis, we do an extensive simulation study using the two-factor stochastic volatility model of Huang and Tauchen (2005). The same DGP was also employed by Bollerslev et al. (2016) to test the HARQ model (see their Appendix A). We simulate one-second data on a regularly spaced grid of 23400 timestamps per day and then compute daily realized variances by summing up $M = 39, 78, 390$ intraday squared returns, corresponding to 10, 5, 1-minute sampling frequencies. The simulated RV series have $T = 3000$ observations. The in-sample analysis is performed on the last 1000 observations. The out-of-sample analysis is performed on the same sample of 1000 observations, but the models are estimated on the first 2000 observations. We generate $N = 1000$ Monte-Carlo realizations and compare the following models: HAR, HARQ, HAR_{log}, HARK, SHAR, SHARK, ARMA(1,1), HARST. The ARMA(1,1) and HARST are estimated on $\log(RV_t)$. The first part of Table (1) compares the average in-sample mean square error (MSE), mean absolute error (MAE) and QLIKE.

The true IV_t used in the computation of the loss measures is evaluated by summing up 1-sec squared returns[1]. We also show the average signal-to-noise ratio estimated by the HARK and SHARK models.

We first discuss the in-sample results. The HAR_{log} outperforms the HAR. Indeed, in contrast to the linear DGP in Section (2.3), the simulated series has nonlinear dynamics. In this case, as it has been shown on S&P500 data, the HAR_{log} provides a better dynamic specification, which translates into superior in-sample estimates. As expected, the HARQ provides better estimates than the HAR in terms of all the three loss measures. Note that the relative MSE of the HARQ approaches one as M becomes large. This is due to the effect of measurement errors. However, the relative QLIKE is significantly lower than one even at $M = 390$, a scenario in which measurement errors are very small. This result shows that the HARQ also captures nonlinear dependencies.

Except for the fact that the HARK outperforms all other models in terms of MAE, in this in-sample analysis there are no significant advantages in using the HARK, SHAR and SHARK in place of the HAR_{log}. The reason is twofold. On the one hand, estimation errors are small, as indicated by the large values of the signal-to-noise ratio δ. The amount of noise on real data is larger, as indicated by the lower values of δ that we have found in our empirical application in Section (3). Second, the filtered time-varying parameters resulting from estimating the SHAR and the SHARK are less erratic than what found on empirical data, where the HAR parameters tend to change significantly over time (see figures 7, and 10-12 in the paper).

It is more interesting to look at the out-of-sample results, summarized in the second part of Table (1). Both the HAR_{log} and the HARQ perform better than the HAR, but the HAR_{log} outperforms the HARQ in terms of all the three loss measures. This implies that, if one is interested in forecasting, modeling logarithmic time series through the HAR_{log} is more effective than correcting for measurement errors through the HARQ.

The HARK outperforms the HAR_{log} in terms of MAE for $M = 39$ and $M = 78$, where measurement errors are more relevant. In contrast to the in-sample analysis, introducing time-varying parameters allows to improve over

[1]The simulated prices are not contaminated by microstructure noise. The realized variance is therefore a consistent and unbiased estimator of the true IV. In presence of microstructure noise, one only needs to replace realized variance with a robust estimator (such as the two-scale estimator of Zhang et al. 2005) and using a consistent estimator for the variance of the error as a proxy of h_t.

	HAR	HARQ	HAR_{log}	HARK	SHAR	SHARK	ARMA	HARST
				In-sample				
				$M = 39$				
MSE	1.0000	**0.97620**	0.98331	0.99210	0.98376	0.99994	0.98866	0.98321
MAE	1.0000	0.98676	0.97516	**0.96248**	0.97643	0.96478	0.97781	0.97445
QLIKE	1.0000	0.97062	**0.9499**	0.96401	0.95050	0.96581	0.97911	0.95131
Signal-to-Noise	-	-	-	7.54	-	8.00	-	-
				$M = 78$				
MSE	1.0000	**0.97981**	0.98710	0.99148	0.98916	0.99474	0.99101	0.98858
MAE	1.0000	0.98754	0.97748	**0.97048**	0.97937	0.97310	0.98172	0.97863
QLIKE	1.0000	0.96751	**0.9487**	0.95535	0.94931	0.95572	0.95910	0.95201
Signal-to-Noise	-	-	-	13.62	-	14.30	-	-
				$M = 390$				
MSE	1.0000	**0.98222**	0.98884	0.99017	0.99069	0.99024	0.99587	0.99124
MAE	1.0000	0.98870	0.97817	**0.97639**	0.97982	0.97905	0.98165	0.97862
QLIKE	1.0000	0.96532	**0.9446**	0.94634	0.94502	0.9457	0.95673	0.95663
Signal-to-Noise	-	-	-	61.17	-	62.70	-	-
				Out-of-sample				
				$M = 39$				
MSE	1.0000	0.98932	0.96745	0.97617	**0.96002**	0.98566	0.97910	0.96123
MAE	1.0000	0.98654	0.96869	0.95723	0.97210	**0.95588**	0.97443	0.97131
QLIKE	1.0000	0.97024	0.92573	0.94130	**0.92118**	0.94763	0.95523	0.93001
				$M = 78$				
MSE	1.0000	0.99365	0.97322	0.97758	0.96914	0.98080	0.98011	**0.96547**
MAE	1.0000	0.98796	0.97130	0.96511	0.97498	**0.96405**	0.97899	0.97501
QLIKE	1.0000	0.96036	0.92468	0.93155	**0.92112**	0.93358	0.93251	0.93100
				$M = 390$				
MSE	1.0000	0.99821	0.97634	0.99161	**0.97368**	0.97642	0.97793	0.97694
MAE	1.0000	0.98822	**0.97275**	0.97605	0.97701	0.97313	0.98662	0.98100
QLIKE	1.0000	0.95587	0.92332	0.94955	**0.91962**	0.92288	0.93122	0.93283

Table 1: Average relative in-sample and out-of-sample losses of HAR, HARQ, HAR_{log}, HARK, SHAR, SHARK, ARMA(1,1), HARST models on 1000 simulated daily RV data. Signal-to-Noise denotes the average signal-to-noise ratio estimated by the HARK and the SHARK.

the HAR_{log} and the HARK. The SHAR has lowest MSE and QLIKE, while the SHARK provides the lowest MAE. Therefore, even in presence of weak nonlinear dependencies, modeling the residual misspecification of the HAR_{log} and the HARK leads to significantly better out-of-sample forecasts. The ARMA(1,1) outperforms the HAR. This is due to the logarithmic transformation. Indeed, it is in turn outperformed by the HAR_{log}. Finally. we note that the performance of the HARST is close to that of the SHAR, as both models capture nonlinear dependencies.

B Summary statistics of the dataset used in the empirical application

Stock	Symbol	Min	Mean	Median	Max
SP500		0.0177	0.9243	0.4808	38.2914
Citigroup	C	0.1473	9.6360	2.5460	972.4663
Morgan Stanley	MS	0.1847	11.1067	2.8120	1.64e+03
Goldman Sachs	GS	0.2056	4.9457	1.6992	394.4546
JPMorgan Chase	JPM	0.1035	5.2330	1.6949	254.1726
Bank of America	BAC	0.1088	7.8898	2.1562	377.5072
ConocoPhillips	COP	0.1303	2.8366	1.4524	191.3988
Exxon Mobil	XOM	0.1130	1.9655	0.9875	135.4296
Chevron	CVX	0.1052	2.2589	1.1585	142.7586
Schlumberger	SLB	0.3042	4.6566	2.7218	165.4489
General Electric	GE	0.1088	3.2878	1.1553	172.7190
CBS Corporation	CBS	0.2013	5.5533	2.2906	165.6725
Walt Disney	DIS	0.1380	2.2760	1.1164	112.5801
Halliburton Company	HAL	0.1964	5.5976	3.1641	205.0786
Johnson & Johnson	JNJ	0.0668	0.9579	0.5167	49.6791
McDonald's	MCD	0.0867	1.3879	0.7031	124.4459
Pfizer	PFE	0.1620	1.9173	1.1373	64.0720
Verizon Communications	VZ	0.1229	1.8900	0.9336	108.7225
Wal-Mart	WMT	0.1134	1.4193	0.7576	73.2718

Table 2: Summary statistics of realized variances computed with 5-minute returns of S&P500 and NYSE stocks data used in the empirical analysis in Section (3).

C Leverage

Corsi and Renò (2012) pointed out the importance of considering leverage effects when modeling realized volatility. In particular, they found that the forecasting performance of the HAR can be significantly improved by introducing a persistent leverage effect with an approximate long-memory behavior similar to that of volatility itself.

Let r_t denote the log-return at time t. The LHAR model extends the standard HAR to also describe the asymmetric dependence between volatility and returns, with future volatility being more responsive to negative returns:

$$\begin{aligned} RV^l_{t+1} =& \beta_0 + \beta_1 RV^l_t + \beta_2 RV^l_{t-1|t-4} + \beta_3 RV^l_{t-5|t-21} \\ &+ \gamma_1 r^-_t + \gamma_2 r^-_{t-1|t-4} + \gamma_3 r^-_{t-5|t-21} + \eta_{t+1} \end{aligned} \tag{C.1}$$

where $r_{t_1|t_2}$ is computed as $RV_{t_1|t_2}$ and $r^-_{t_1|t_2} = \min(r_{t_1|t_2}, 0)$.

Since returns are observable, the leverage terms in eq. (C.1) can be easily accounted for in our framework. In particular, the SHAR can be written as:

$$\begin{aligned} RV^l_{t+1} =& \beta_{0,t+1} + \beta_{1,t+1} RV^l_t + \beta_{2,t+1} RV^l_{t-1|t-4} + \beta_{3,t+1} RV^l_{t-5|t-21} \\ &+ \gamma_1 r^-_t + \gamma_2 r^-_{t-1|t-4} + \gamma_3 r^-_{t-5|t-21} + \eta_{t+1} \end{aligned} \tag{C.2}$$

$\eta_{t+1} \sim \text{NID}(0, q_{t+1})$. Model (C.2) can be handled as the SHAR, provided that one includes the leverage terms in the conditional mean $\mu_{t|t-1}$ in eq. (24). The additional parameters λ_1, λ_2, λ_3 are part of Θ, the vector of static parameters that is estimated by maximum likelihood. In the HARK and SHARK models, the new transition equation becomes

$$\alpha_{t+1} = c^{(l)}_t + T_t \alpha_t + \eta_t, \quad \eta_t \sim \text{NID}(0, Q_t) \tag{C.3}$$

where:

$$c^{(l)}_t = c_t + \begin{pmatrix} \gamma_1 r^-_t + \gamma_2 r^-_{t-1|t-4} + \gamma_3 r^-_{t-5|t-21} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{C.4}$$

Estimation proceeds in the same way as described in Section (2.4) and (2.6), provided that one replaces c_t with $c^{(l)}_t$ in the Kalman filter recursions. As in the SHAR, the leverage parameters are estimated by maximum likelihood, together with the remaining parameters. The impact of leverage on volatility forecasting with the $\text{HAR}_{\log}$ has been investigated in the econometric literature (see e.g. Corsi and Renò 2012) and produces similar forecasting improvements when implemented in our framework.

D Kalman filter recursions for the HARK model

We first introduce some notation. The n-dimensional identity matrix is denoted as $\mathbb{I}_n$. We use $\otimes$ to denote the Kronecker product between two matrices. The operator vec$(\cdot)$, applied to an $m \times n$ matrix A, stacks the columns of A into an mn column vector while the operator vech$(\cdot)$, applied to a symmetric $n \times n$ matrix B, stacks all the $n(n-1)/2$ upper (or lower) diagonal elements into a column vector. We also introduce the commutation matrix C_{mn}, i.e. the $mn \times mn$ matrix such that $C_{mn}\text{vec}A = \text{vec}A'$ for every $m \times n$ matrix A. The derivative of an $m \times n$ matrix function $F(X)$ with respect to the $p \times q$ matrix X is defined as in Abadir and Magnus (2005), i.e. as the $mn \times pq$ matrix computed as $\partial\text{vec}(F(X))/\partial\text{vec}(X)'$.

Let $\mathcal{F}_t$, a_t, P_t defined as in Section (2.4). The Kalman filter recursions for model (14), (15) are given by:

$$\begin{aligned} v_t &= RV_t - Z a_t & F_t &= Z P_t Z' + h_t \\ a_{t+1} &= c + T a_t + K_t v_t & P_{t+1} &= T P_t (T - K_t Z)' + Q \end{aligned} \tag{D.1}$$

where $K_t = TP_tZ'F_t^{-1}$.

E Computation of ∇_t and $\mathcal{I}_{t|t-1}$ in the SHAR model

It is convenient to introduce the auxiliary vector of time-varying parameters:

$$\tilde{f}_t = \Big(\beta_{0,t}, \beta_{1,t}, \beta_{2,t}, \beta_{3,t}, q_t\Big)' \tag{E.1}$$

The latter is related to f_t by the following link-function:

$$\tilde{f}_t = \mathcal{L}(f_t) = \left[f_t^1, f_t^2, f_t^3, f_t^4, \exp(f_t^5)\right]' \tag{E.2}$$

The Jacobian of the transformation is:

$$J_{\mathcal{L}} = \frac{\partial \tilde{f}_t}{\partial {f_t}'} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \exp(f_t^5) \end{pmatrix} \tag{E.3}$$

Note that, using the chain rule, ∇_t and $\mathcal{I}_{t|t-1}$ can be expressed as:

$$\nabla_t = J_{\mathcal{L}}\tilde{\nabla}_t, \quad \mathcal{I}_{t|t-1} = J_{\mathcal{L}}\tilde{\mathcal{I}}_{t|t-1}J_{\mathcal{L}} \tag{E.4}$$

where:

$$\tilde{\nabla}_t = \left[\frac{\partial \log p(RV_t|\tilde{f}_t, \mathcal{B}_{t-1}, \Theta)}{\partial \tilde{f}_t'}\right]', \quad \tilde{\mathcal{I}}_{t|t-1} = \mathrm{E}[\tilde{\nabla}_t\tilde{\nabla}_t'] \tag{E.5}$$

Thus, it is simpler to compute $\tilde{\nabla}_t$ and $\tilde{\mathcal{I}}_{t|t-1}$ and then using eq. (E.5) to recover ∇_t and $\mathcal{I}_{t|t-1}$. By direct differentiation of the conditional log-likelihood (23), we have:

$$\tilde{\nabla}_t = \begin{pmatrix} \frac{\eta_t}{q_t} \\ \frac{\eta_t}{q_t}RV_{t-1} \\ \frac{\eta_t}{q_t}RV_{t-2|t-5} \\ \frac{\eta_t}{q_t}RV_{t-6|t-22} \\ -\frac{1}{2q_t} + \frac{1}{2}\frac{\eta_t^2}{q_t^2} \end{pmatrix} \tag{E.6}$$

where $\eta_t = RV_t - \mu_{t|t-1}$. The information matrix $\tilde{\mathcal{I}}_{t|t-1}$ is easily computed by noting that $\mathrm{E}_{t|t-1}[\eta_t^2] = q_t$ and $\mathrm{E}_{t|t-1}[\eta_t^4] = 3q_t^2$. We have:

$$\tilde{\mathcal{I}}_{t|t-1} =$$

$$\frac{1}{q_t}\begin{pmatrix} 1 & RV_{t-1} & RV_{t-2|t-5} & RV_{t-6|t-22} & 0 \\ RV_{t-1} & RV_{t-1}^2 & RV_{t-1}RV_{t-2|t-5} & RV_{t-1}RV_{t-6|t-22} & 0 \\ RV_{t-2|t-5} & RV_{t-1}RV_{t-2|t-5} & RV_{t-2|t-5}^2 & RV_{t-2|t-5}RV_{t-6|t-22} & 0 \\ RV_{t-6|t-22} & RV_{t-1}RV_{t-6|t-22} & RV_{t-2|t-5}RV_{t-6|t-22} & RV_{t-6|t-22}^2 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2q_t} \end{pmatrix} \tag{E.7}$$

Note that $\mathcal{I}_{t|t-1}$ has rank 2. In order to compute the scaling matrix S_t we thus take the Moore–Penrose inverse.

F Kalman filter recursions for the SHARK model

Let $\mathcal{F}_t$, a_{t+1} and P_{t+1} as in Section (D). Then, the Kalman filter recursions for model (27), (28) are given by:

$$\begin{aligned} v_t &= RV_t - Za_t & F_t &= ZP_tZ' + h_t \\ a_{t+1} &= c_{t+1} + T_{t+1}a_t + K_tv_t & P_{t+1} &= T_{t+1}P_t(T_{t+1} - K_tZ)' + Q_{t+1} \end{aligned} \tag{F.1}$$

where $K_t = T_{t+1}P_tZ'F_t^{-1}$. The difference with the Kalman filter recursions in Section (D) is that now c_t, T_t and Q_t are time-varying and are updated at each step using the eq. (30) in the paper.

G Computation of $\dot{v}_t$ and $\dot{F}_t$ in the SHARK model

As done with the SHAR, in order to obtain positive variances we introduce the auxiliary vector of time-varying parameters:

$$\tilde{f}_t = \mathcal{L}(f_t) = \left[f_t^1, f_t^2, f_t^3, f_t^4, \exp(f_t^5)\right]' \tag{G.1}$$

The Jacobian $J_{\mathcal{L}}$ is the same as in eq. (E.3) and eq. (E.4) still holds. Consequently, we first compute $\tilde{\nabla}_t$ and $\tilde{\mathcal{I}}_{t|t-1}$ as defined in eq. (E.5). They have the same form as in eq. (32), but now $\dot{v}_t = \partial v_t/\partial \tilde{f}_t'$ and $\dot{F}_t = \partial \text{vec}(F_t)/\partial \tilde{f}_t'$ denote derivatives with respect to $\tilde{f}_t$. Following Delle Monache et al. (2016), they can be computed through the following set of equations:

$$\dot{v}_t = -Z\dot{a}_t \tag{G.2}$$

$$\dot{F}_t = (Z \otimes Z)\dot{P}_t \tag{G.3}$$

$$\dot{a}_t = \dot{c}_t + (a_{t-1}' \otimes \mathbb{I}_n)\dot{T}_t + v_{t-1}\dot{K}_{t-1} \tag{G.4}$$

$$\begin{aligned} \dot{P}_t = &(\mathbb{I}_n \otimes T_tP_{t-1})C_{n,n}\dot{T}_t + (T_tP_{t-1} \otimes \mathbb{I}_n)\dot{T}_t - (K_{t-1}ZP_{t-1} \otimes \mathbb{I}_n)\dot{T}_t \\ &- (\mathbb{I}_n \otimes T_tP_{t-1}Z')\dot{K}_{t-1} + \dot{Q}_t \end{aligned} \tag{G.5}$$

$$\dot{K}_{t-1} = (F_{t-1}^{-1}ZP_{t-1} \otimes \mathbb{I}_n)\dot{T}_t \tag{G.6}$$

where $\dot{c}_t = \frac{\partial a_t}{\partial \tilde{f}_t'}$ is a $n \times 5$ matrix and $\dot{T}_t = \frac{\partial \text{vec}T_t}{\partial \tilde{f}_t'}$, $\dot{Q}_t = \frac{\partial \text{vec}Q_t}{\partial \tilde{f}_t'}$ are $n^2 \times 5$. They are selection matrix of the form:

$$\{\dot{A}_t\}_{i,j} = \begin{cases} 1, & \text{if } \{\text{vec}A_t\}_i = \tilde{f}_t^j \\ 0, & \text{else} \end{cases} \tag{G.7}$$

where A_t generically denotes c_t, T_t and Q_t.

H Proof of propositions 1 and 2

H.1 Proposition 1

The proof is made by induction on j. First, let us prove the formula for $j = 2$. We have:

$$\begin{aligned}
\mathrm{E}_t[\exp(Z\overline{RV}_{t+2})] &= \mathrm{E}_t[\exp(Z(c_l + T_l\overline{RV}_{t+1} + \overline{\eta}_{t+1}))] \\
&= \mathrm{E}_t[\exp(Z(c_l + T_l(c_l + T_l\overline{RV}_t + \overline{\eta}_t) + \overline{\eta}_{t+1}))] \\
&= \exp\left[Z((\mathbb{I}_n + T_l)c_l + T^2\overline{RV}_t) + \frac{1}{2}Z(Q_l + T_lQ_lT_l')Z'\right]
\end{aligned}$$

where we have used the moment generating function of the multivariate normal distribution in the last line. So the formula is true for $j = 2$. Now let us assume that it is true for $j = k$ and prove that it is also true for $j = k + 1$:

$$\begin{aligned}
&\mathrm{E}_t[\exp(Z\overline{RV}_{t+k+1})] = \\
&= \mathrm{E}_t[\exp(Z(c_l + T_l\overline{RV}_{t+k} + \overline{\eta}_{t+k}))] \\
&= \exp(Zc_l)\mathrm{E}_t[\exp(ZT_l\overline{RV}_{t+k})]\mathrm{E}_t[\exp(Z\overline{\eta}_{t+l})] \\
&= \exp(Zc_l)\exp\left\{Z\left[(T_l + \cdots + T_l^k)c_l + T_l^{k+1}\overline{RV}_t^l\right] + \frac{1}{2}Z\left[T_lQ_lT_l' + \cdots + T^kQ_l(T^k)'\right]Z'\right\}\exp(ZQ_lZ')
\end{aligned}$$

where we have used the induction hypothesis to compute $\exp(ZT_l\overline{RV}_{t+k})$. Upon multiplication of the three terms we get:

$$\begin{aligned}
&\mathrm{E}_t[\exp(Z\overline{RV}_{t+k+1})] = \\
&= \exp\left\{Z\left[(\mathbb{I}_n + \cdots + T_l^k)c_l + T_l^j\overline{RV}_t^{k+1}\right] + \frac{1}{2}Z\left[Q_l + \cdots + T^kQ_l(T^k)'\right]Z'\right\}
\end{aligned}$$

which is the formula in Proposition 1 for $j = k + 1$ ■

H.2 Proposition 2

The proof follows exactly the same steps as those in Proposition 1. Let us prove the formula for $j = 2$. We have:

$$\begin{aligned}
\mathrm{E}_t[\exp(Z\alpha_{t+2})] &= \mathrm{E}_t[\exp(Z(c + T\alpha_{t+1} + \eta_{t+1}))] \\
&= \exp(Zc)\mathrm{E}_t[\exp(ZT\alpha_{t+1})]\mathrm{E}_t[\exp(Z\eta_{t+1})]
\end{aligned}$$

where the last step follows from independence between α_{t+1} and η_{t+1}. Using the fact that $\alpha_{t+1}|\mathcal{F}_t \sim \mathrm{N}(a_{t+1}, P_{t+1})$, we have:

$$\begin{aligned}
&\mathrm{E}_t[\exp(Z\alpha_{t+2})] = \\
&= \exp(Zc)\exp\left(ZTa_{t+1} + \frac{1}{2}ZTP_{t+1}T'Z'\right)\exp(ZQZ') \\
&= \exp\left[Z(c + Ta_{t+1}) + \frac{1}{2}Z(TP_{t+1}T' + Q)Z'\right]
\end{aligned}$$

which is the formula in Proposition 2 for $j = 2$. We now assume that the formula is true for $j = k$ and prove that it is also true for $j = k + 1$:

$$\begin{aligned}
&\mathrm{E}_t[\exp(Z\alpha_{t+k+1})] = \\
&= \mathrm{E}_t[\exp(Z(c + T\alpha_{t+k} + \eta_{t+k}))] \\
&= \exp(Zc)\mathrm{E}_t[\exp(ZT\alpha_{t+k})]\mathrm{E}_t[\exp(Z\eta_{t+k})]
\end{aligned}$$

The last term is equal to $\exp(\frac{1}{2}ZQZ')$. Based on the induction hypothesis, the second term is given by:

$$\begin{aligned}
\mathrm{E}_t[\exp(ZT\alpha_{t+k})] = \exp\big\{&Z\left[(T + \cdots + T^{k-1})c + T^k a_{t+1}\right] + \\
&\frac{1}{2}Z\left[T^k P_{t+1}(T^k)' + TQT' + \cdots + T^{k-1}Q(T^{k-1})'\right]Z'\Big\}
\end{aligned}$$

By summing the three exponents we end up with:

$$\begin{aligned}
\mathrm{E}_t[\exp(ZT\alpha_{t+k})] = \exp\big\{&Z\left[(\mathbb{I}_n + \cdots + T^{k-1})c + T^k a_{t+1}\right] + \\
&\frac{1}{2}Z\left[T^k P_{t+1}(T^k)' + Q + \cdots + T^{k-1}Q(T^{k-1})'\right]Z'\Big\}
\end{aligned}$$

which is the formula in Proposition 2 for $j = k + 1$ ■

References

Abadir, K., Magnus, J., 2005. Matrix Algebra. Econometric Exercises. Cambridge University Press.

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

Corsi, F., Renò, R., 2012. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. Journal of Business & Economic Statistics 30 (3), 368–380.

Delle Monache, D., Petrella, I., Venditti, F., 2016. Adaptive state space models with applications to the business cycle and financial stress. CEPR Discussion Paper (DP11599).

Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variance. Journal of Financial Econometrics 3 (4), 456–499.

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