

Supplemental Appendix for

Bespoke Realized Volatility:
Tailored Measures of Risk for Volatility Prediction

by Andrew J. Patton and Haozhe Zhang

This version: December 30, 2022

This appendix contains six subsections, presenting additional details and analyses relevant to the main paper.

S.1. Data cleaning details

We use trade data from the NYSE Trade and Quote (TAQ) database. We follow Barndorff-Nielsen, Hansen, Lunde and Shephard (2009) to clean this data, using the following rules:

1. Keep only entries with time step between 9:30 am to 4:00 pm (when the exchange is open).
2. Delete entries with zero transaction prices.
3. Retain entries originating from NYSE and NASDAQ only.
4. Delete entries with corrected trades ($\text{CORR} \neq 0$).
5. Delete entries with unusual sale condition (COND with letter code except for letters E and F).
6. We use the median price if multiple entries have the same time stamp.

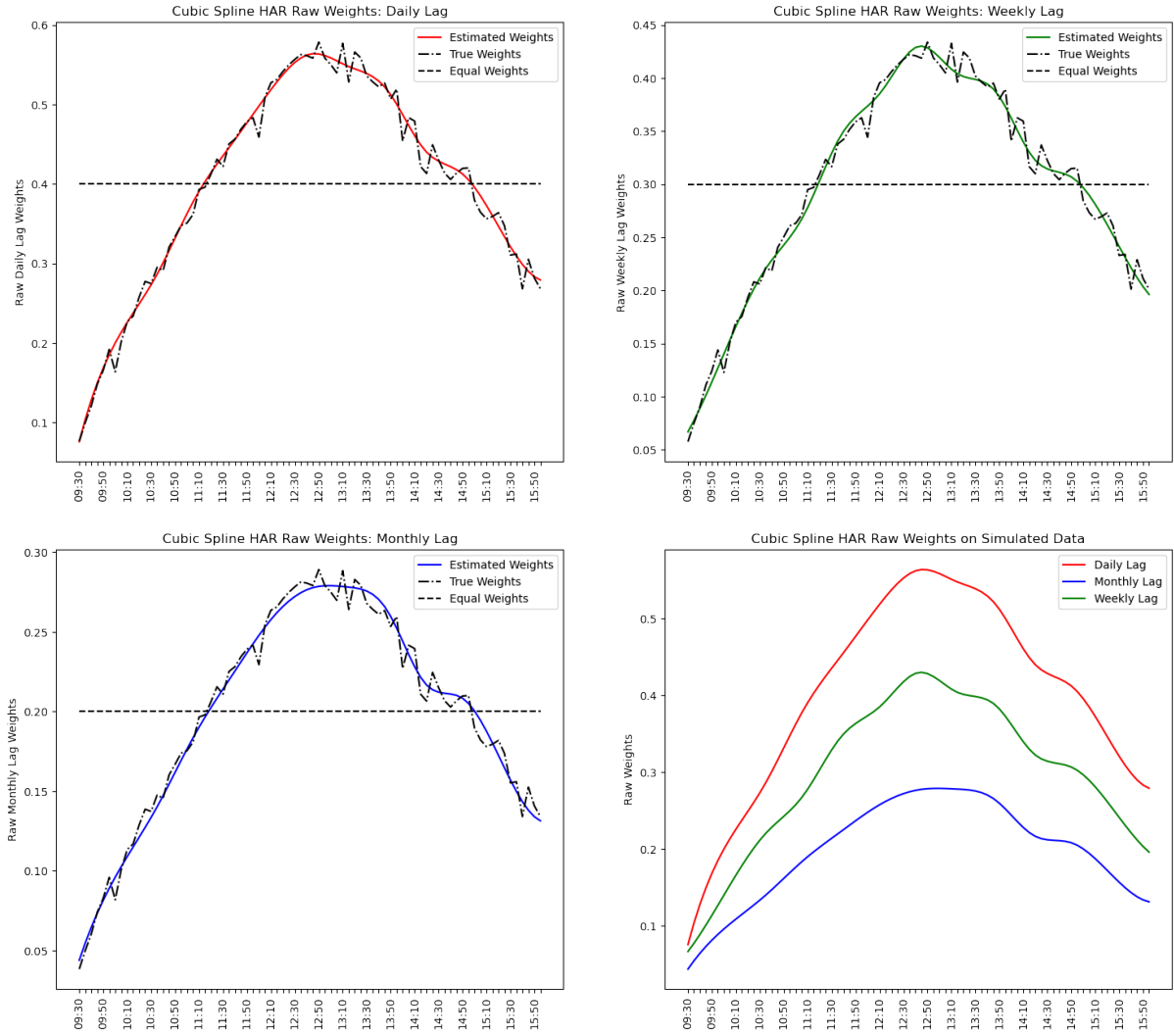
S.2. Simulation study

In this section we simulate data to illustrate the ability of the proposed Cubic HAR model to recover the true coefficients. We assume that high-frequency returns, $r_{i,t}$, are i.i.d. $N(0,1/M)$, where M is the number of high-frequency returns each trade day. We combine these high frequency returns to obtain realized variances using time-of-day (TOD) weights, ω_i , as in equation 7 of the main paper, and we use a HAR model to simulate the time series dynamics of RV.

$$RV_t = \beta_0 + \beta_d \sum_{i=1}^M \omega_i r_{i,t-1}^2 + \beta_w \frac{1}{4} \sum_{j=2}^5 \sum_{i=1}^M \omega_i r_{i,t-j}^2 + \beta_m \frac{1}{16} \sum_{j=6}^{21} \sum_{i=1}^M \omega_i r_{i,t-j}^2 + \epsilon_t \quad (\text{S.1})$$

where $\epsilon_t \sim N(0,1)$. We set $\beta_0 = 0$, $\beta_d = 0.4$, $\beta_w = 0.3$, $\beta_m = 0.3$, and $M = 78$. We impose that ω_i follows the empirical TOD weights presented in Figure 6. Then we use the cubic HAR model with the exact same set-up (hyperparameter grid and estimation algorithm) to estimate the coefficients on the simulated data. Figure S.1 shows that when the true data generating process indeed is a linear combination with the TOD weights, our cubic HAR model will recover the true weights almost perfectly. (As the true weights in this simulation are not a smooth function of time, the cubic spline model clearly cannot obtain a perfect fit.) This simulation results give us confidence that the proposed Cubic HAR has enough estimation power to recover the optimal weighting scheme for realized variance forecasting.

Figure S.1: Cubic HAR vs True Weights: Simulated Data



Note: This figure compares the estimated optimal weights (solid line) based on simulated data based with the true weights (dash-dotted line).

S.3. Alternative target for shrinkage

In the analysis in Section 3.2 of the main paper, we consider shrinkage methods (Ridge, LASSO and elastic net) that shrink the estimated parameters towards zero. This is a standard shrinkage target in high-dimensional estimation, but in our application an interesting alternative target is to shrink the parameters towards the benchmark HAR parameters. To do this, we re-write the regularized regression models as:

$$RV_t = \beta_0 + \widetilde{RV}_{t-1}(\beta_{HAR}^d + \gamma^d) + \frac{1}{4} \sum_{j=2}^5 \widetilde{RV}_{t-j}(\beta_{HAR}^w + \gamma^w) + \frac{1}{16} \sum_{j=6}^{21} \widetilde{RV}_{t-j}(\beta_{HAR}^m + \gamma^m) + e_t \quad (\text{S.1})$$

That is, we split the coefficients on high frequency returns into the HAR coefficients and a perturbation term to be estimated. Then, instead of regularizing the entire coefficient, we only penalize the perturbation terms, and shrink these towards zero. That is, the penalty term becomes:

$$\alpha(\lambda \|\gamma\|_1 + (1 - \lambda) \|\gamma\|_2) \quad (\text{S.2})$$

where $\gamma = [\gamma^d, \gamma^w, \gamma^m]$. The estimation procedure and hyperparameters grid are identical to those in the original regularized regression analysis. We compare the cubic HAR model with these alternative shrinkage estimators in Table S.1. (The OLS and HAR results are identical to those in Table 2 of the main paper, and are included here for ease of comparison.) With this alternative target the optimal degree of shrinkage is found to be large, and the Ridge, LASSO and elastic net models are all shrunk almost all the way to the benchmark HAR parameters. Given that, it is unsurprising that the cubic HAR continues to significantly out-perform these alternative regularized models, as shown in Table S.1, which have performance comparable to the benchmark HAR model.

Table S.1: Alternative regularized regression models: Shrink towards HAR

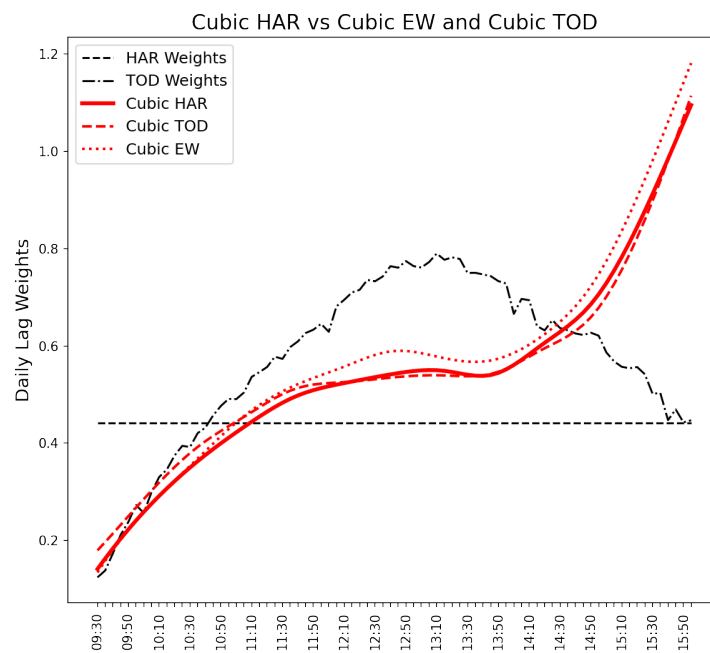
CubicHAR vs:	DM Losses		DM Wins		DM t-stat Panel
	Total	Signif	Total	Signif	
<i>Ridge-Alt</i>	166	44	720	423	-17.3
<i>LASSO-Alt</i>	165	46	721	437	-4.2
<i>Elastic net-Alt</i>	170	46	716	412	-16.4
<i>OLS</i>	54	17	832	564	-6.4
<i>HAR</i>	155	41	731	470	-21.7

Note: This table reports individual and panel Diebold-Mariano tests comparing the cubic HAR model against competing models across 886 S&P 500 stocks. A positive panel DM t-statistic indicates that the competing model out-performs the cubic HAR model, while a negative t-statistic indicates the opposite. The models labeled “-Alt” shrink the estimated coefficients towards the HAR model coefficients, rather than towards zero as in Table 2 of the main paper.

S.4. Visualizing the optimal Cubic TOD and Cubic EW weights

This section presents the optimal daily weights for the cubic HAR, cubic TOD and cubic EW models, which differ in what shapes are estimated or imposed on the weekly and monthly weights. We observe the cubic TOD and cubic EW daily weights are almost identical to the cubic HAR weights for daily lags, which shows that restricting the shapes of weekly and monthly weights has almost zero affect on the estimated daily weights.

Figure S.2: Cubic TOD and Cubic EW weights



Note: This figure compares estimated cubic TOD and cubic EW weights against the cubic HAR model weights, averaged across the 886 S&P 500 stocks in our sample.

S.5. Forecast comparisons using mean-squared error loss

Table S.2: HAR vs other models performance on S&P500. MSE loss.

HAR vs:	DM Losses		DM Wins		DM t-stat Panel
	Total	Signif	Total	Signif	
<i>Fully Flexible</i>	421	169	465	106	0.8
<i>Flexible HAR</i>	726	408	160	11	3.3
<i>Cubic HAR</i>	714	406	172	7	1.9
<i>TOD HAR</i>	653	348	233	62	2.2
<i>Ridge</i>	462	186	424	120	-0.1
<i>LASSO</i>	203	46	683	338	-4.1
<i>Elastic net</i>	434	166	452	136	-2.0
<i>OLS</i>	64	7	822	507	-3.3

Note: This table reports individual and panel Diebold-Mariano tests comparing the baseline HAR model against competing models across 886 S&P 500 stocks, using MSE loss. A positive panel DM t-statistic indicates that the competing model out-performs the HAR model, while a negative t-statistic indicates the opposite. This table is related to Table 1 of the main paper, which uses QLIKE loss.

Table S.3: Cubic HAR vs other models performance on S&P 500. MSE loss.

Cubic HAR vs:	DM Losses		DM Wins		DM t-stat Panel
	Total	Signif	Total	Signif	
<i>Fully Flexible</i>	151	16	735	383	-0.7
<i>Flexible HAR</i>	411	55	475	90	0.0
<i>TOD HAR</i>	405	69	481	196	-1.0
<i>HAR</i>	172	7	714	406	-1.9
<i>Ridge</i>	171	12	715	349	-1.7
<i>LASSO</i>	66	2	820	542	-5.2
<i>Elastic net</i>	159	8	727	372	-3.4
<i>OLS</i>	23	3	863	664	-3.3

Note: This table reports individual and panel Diebold-Mariano tests comparing the cubic HAR model against competing models across 886 S&P 500 stocks, using MSE loss. A positive panel DM t-statistic indicates that the competing model out-performs the cubic HAR model, while a negative t-statistic indicates the opposite. This table is related to Table 2 of the main paper, which uses QLIKE loss.

S.6. Multi-days ahead forecasting comparisons: TOD HAR vs HAR

Table S.4: Multidays Ahead Volatility Forecasting: TOD HAR vs HAR

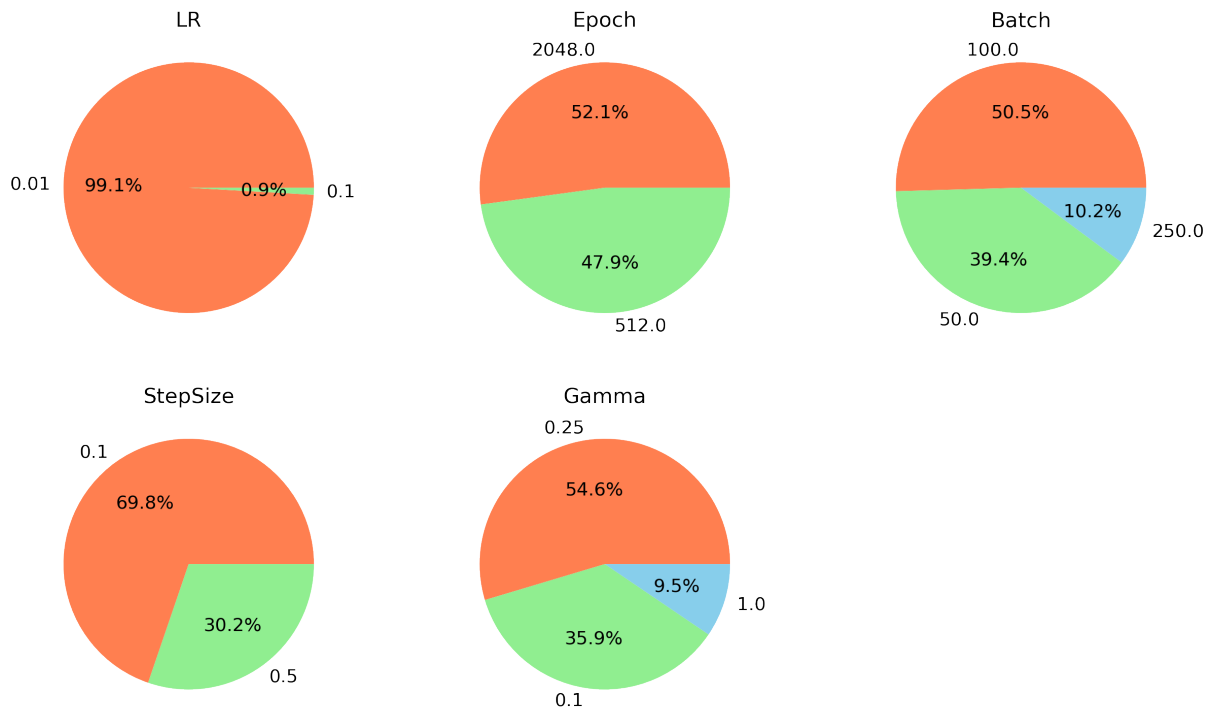
TOD vs HAR Horizon	DM Losses		DM Wins		DM t-stat Panel
	Total	Signif	Total	Signif	
<i>1-Day</i>	206	63	680	458	-18.6
<i>2-Day</i>	253	72	633	349	-12.5
<i>3-Day</i>	257	74	629	314	-10.1
<i>4-Day</i>	264	74	622	293	-13.6
<i>5-Day</i>	270	71	616	298	-11.5
<i>20-Day</i>	253	68	633	345	-6.8
<i>60-Day</i>	345	126	541	255	-0.2

Note: In this table, we report the individual and panel Diebold Mariano tests results of TOD HAR model against the HAR model in the S&P500 cross section for longer horizon forecasting. Note that negative test statistics favors the TOD HAR model.

S.7. Visualization of the hyperparameters for deep learning based models

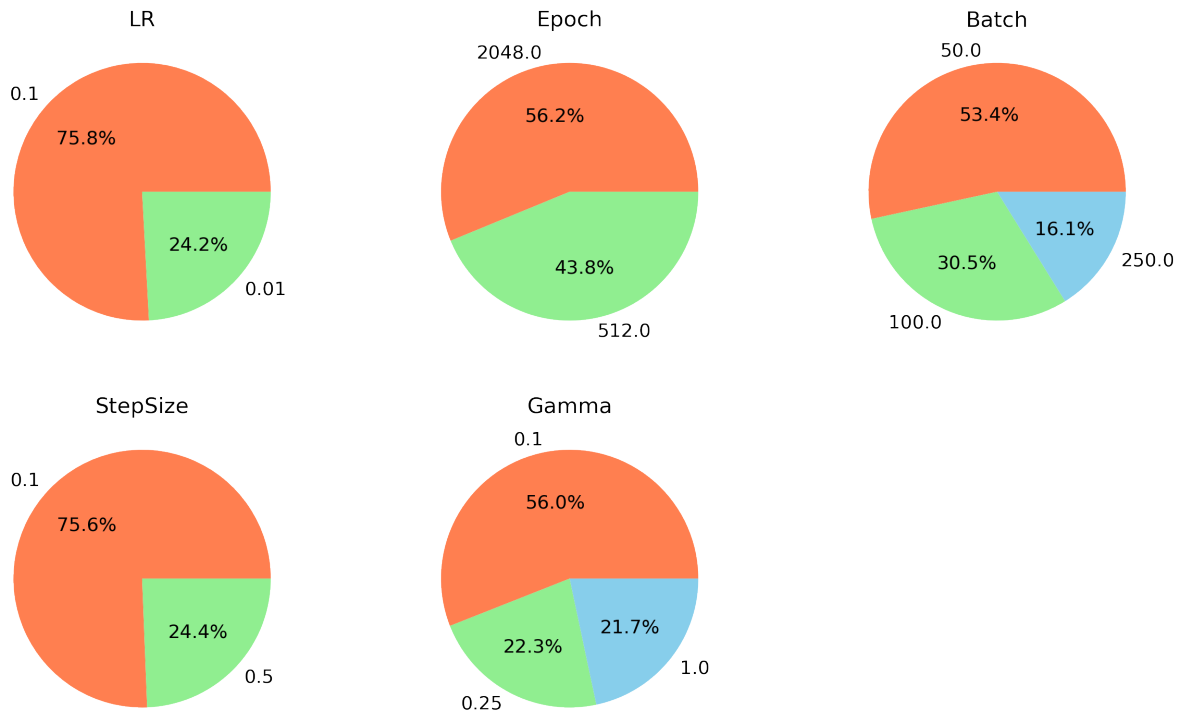
In this section, we also report visualization of the cross sectional variation in the deep learning based models.

Figure S.3: Fully flexible model hyperparameters



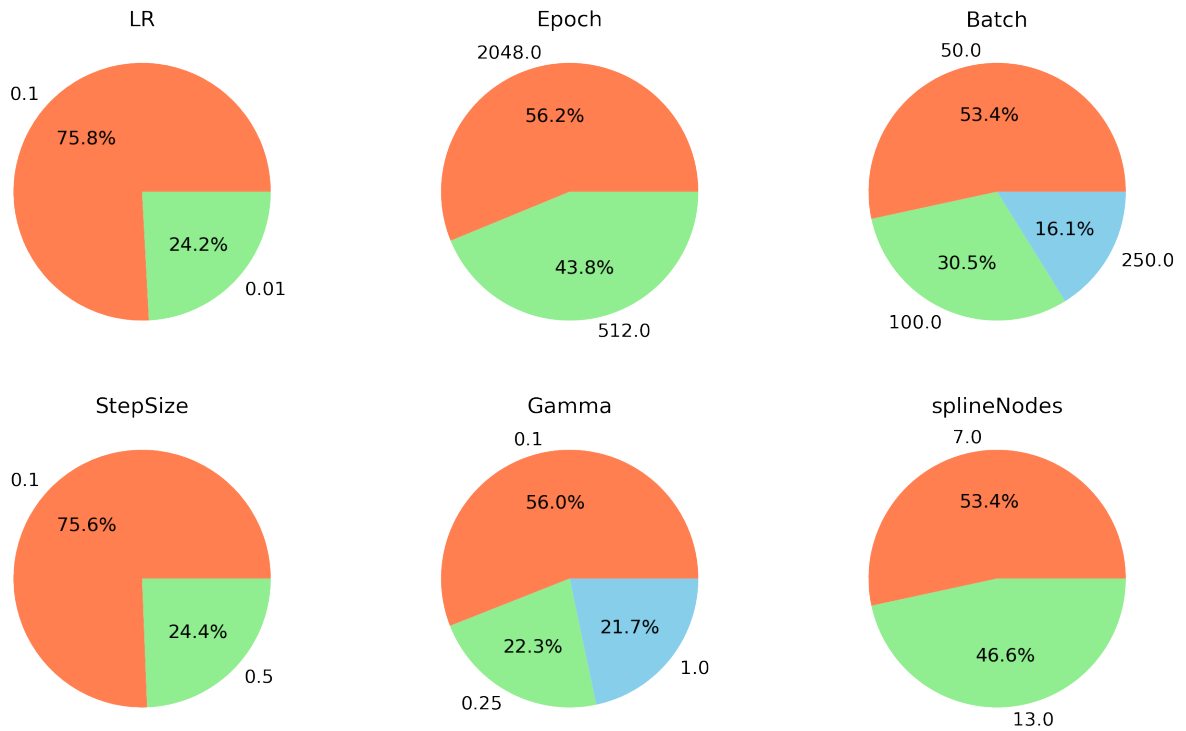
Note: This figure displays the cross-sectional variation of the optimal hyperparameters for the fully flexible model.

Figure S.4: Flexible HAR model hyperparameters



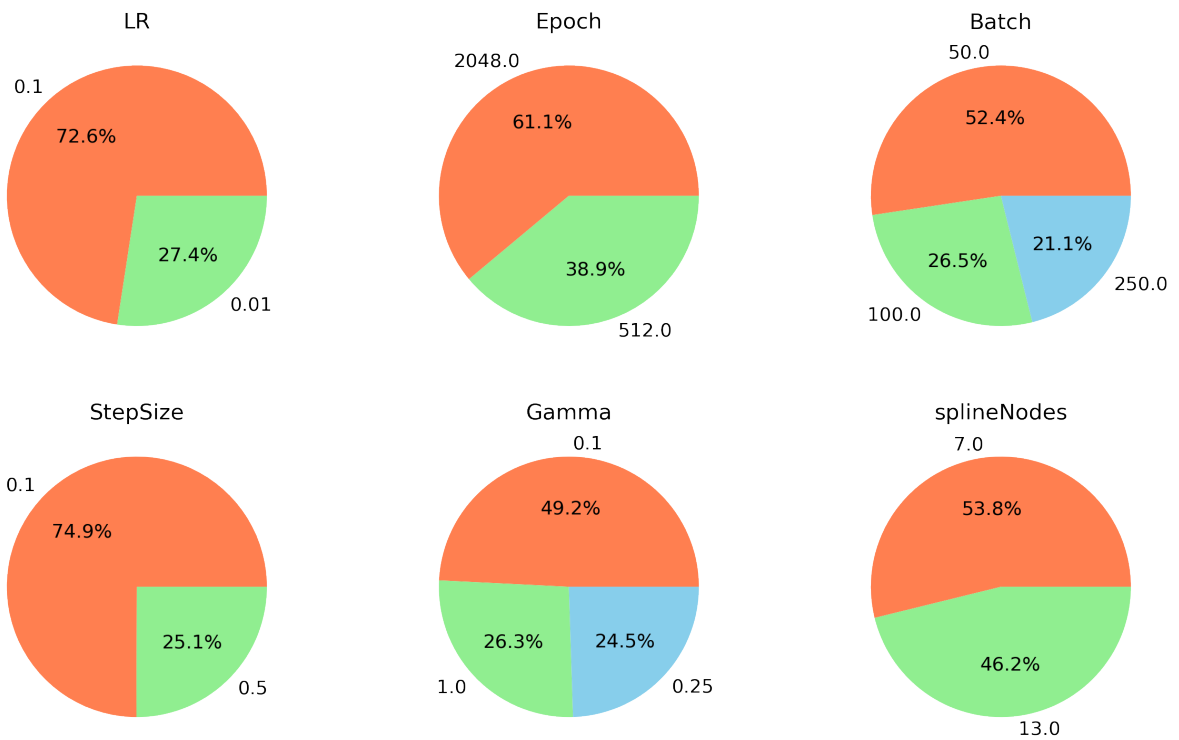
Note: This figure displays the cross-sectional variation of the optimal hyperparameters for the flexible HAR model.

Figure S.5: Cubic HAR model hyperparameters



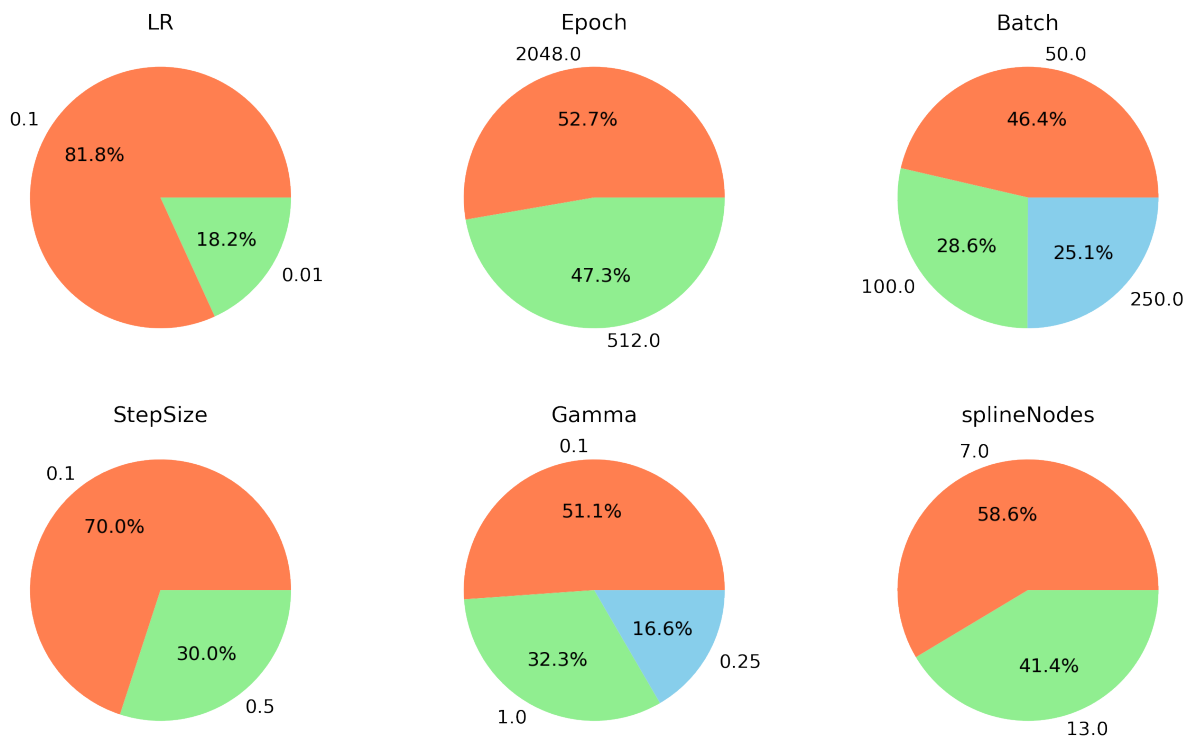
Note: This figure displays the cross-sectional variation of the optimal hyperparameters for the cubic HAR model.

Figure S.6: Cubic-TOD model hyperparameters



Note: This figure displays the cross-sectional variation of the optimal hyperparameters for the cubic-TOD model.

Figure S.7: Cubic-EW model hyperparameters



Note: This figure displays the cross-sectional variation of the optimal hyperparameters for the cubic-EW model.