Variance Risk Premium and Multi-Factor Predictability of Short-Horizon Returns

Nov 2025

1 Volatility Forecasting Models: HAR and GARCH

Volatility forecasting plays a central role in the construction of the Variance Risk Premium (VRP), defined as the difference between the risk–neutral expectation of future variance (implied variance) and the physical-measure expectation of future realized variance. The latter quantity, $\mathbb{E}_t[\mathrm{RV}_{t,t+30}]$, is unobservable and must be estimated using a statistical model. In this section, we introduce two widely used volatility forecasting models: the Heterogeneous Auto-Regressive (HAR) model and the GARCH(1,1) model.

1.1 HAR Model

The HAR model was introduced by Corsi (2009) to capture the long-memory structure of realized volatility. It models the future realized variance as a linear combination of realized variance measured over heterogeneous time horizons. Specifically, the HAR(1,5,22) specification is given by:

$$\mathrm{RV}_{t+1} = \alpha + \beta_d \mathrm{RV}_t^{(d)} + \beta_w \mathrm{RV}_t^{(w)} + \beta_m \mathrm{RV}_t^{(m)} + \varepsilon_{t+1},$$

where

$$\mathrm{RV}_t^{(d)} = \mathrm{RV}_t, \qquad \mathrm{RV}_t^{(w)} = \frac{1}{5}\sum_{i=0}^{4} \mathrm{RV}_{t-i}, \qquad \mathrm{RV}_t^{(m)} = \frac{1}{22}\sum_{i=0}^{21} \mathrm{RV}_{t-i}.$$

Because financial volatility exhibits persistence at multiple horizons, the HAR structure allows the model to capture long-memory behavior using a simple linear specification. In our VRP computation, the HAR model provides a physical-measure forecast of the next 30-day realized variance:

$$\widehat{\mathrm{RV}}_{t,t+30}^{HAR} = \mathbb{E}_t^{HAR}[\mathrm{RV}_{t,t+30}].$$

1.2 GARCH(1,1) Model

The GARCH(1,1) model of Bollerslev (1986) specifies the conditional variance of daily returns as:

$$r_t = \mu + \varepsilon_t, \qquad \varepsilon_t \mid \mathcal{F}_{t-1} \sim N(0, h_t),$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}.$$

The model captures volatility clustering, a hallmark of financial time series. Under the GARCH structure, the k-step-ahead conditional variance forecast is:

$$\widehat{h}_{t+k} = \omega + (\alpha + \beta)\widehat{h}_{t+k-1}.$$

The 30-day realized variance forecast is obtained by summing 30 daily conditional variance forecasts:

$$\widehat{\mathrm{RV}}_{t,t+30}^{GARCH} = \sum_{k=1}^{30} \widehat{h}_{t+k}.$$

1.3 Purpose for VRP Construction

The Variance Risk Premium is defined as:

$$\mathrm{VRP}_t = \mathrm{IV}_{t,t+30} - \mathbb{E}_t[\mathrm{RV}_{t,t+30}],$$

where $\mathrm{IV}_{t,t+30}$ is implied variance extracted from option prices, and the physical expectation $\mathbb{E}_t[\mathrm{RV}_{t,t+30}]$ is estimated using either HAR or GARCH. Therefore, accurate volatility forecasting is essential for correctly identifying VRP dynamics and evaluating its predictive power for future returns.

2 Results of Rolling Window Forecasting

To evaluate the out-of-sample forecasting performance of HAR and GARCH models, we adopt a rolling-window approach. At each day t, the models are estimated using the most recent 750 trading days (approximately three years of data). Using the fitted model, we compute the forecast of the next 30-day realized variance $\widehat{\mathrm{RV}}_{t,t+30}$.

Table 1 reports the out-of-sample forecasting accuracy for the two models based on the following metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Pearson correlation with realized variance, and the coefficient of determination R^2.

Table 1: Out-of-Sample Forecasting Performance for 30-day Realized Variance

Model	MSE	MAE	RMSE	Corr	R^2
HAR	4.01e-05	4.01e-03	6.33e-03	0.6806	0.4465
GARCH	5.24e-05	4.28e-03	7.24e-03	0.6502	0.2978

The HAR model consistently outperforms GARCH across all metrics. In particular, HAR achieves substantially lower MSE and RMSE, and exhibits a higher correlation with realized volatility. The R^2 indicates that HAR explains approximately 45% of out-of-sample variation, compared with about 30% for GARCH.

Figures 1 and 2 illustrate the realized 30-day volatility and the corresponding out-of-sample forecasts generated by the HAR and GARCH models. Both models successfully capture major volatility cycles, including the spikes during the 2008 Global Financial Crisis, the 2015 market correction, and the 2020 COVID-19 shock.

However, notable differences arise:

- The HAR model tracks realized volatility more closely, especially during moderate volatility regimes. Its multi–horizon structure enables it to adjust more rapidly once volatility enters or exits persistent periods.

- The GARCH forecast is smoother and exhibits stronger inertia, leading to delayed adjustments during sudden volatility spikes. This is expected, as GARCH(1,1) relies solely on past squared returns and past variance, with no heterogeneous horizon components.

- HAR also delivers more accurate volatility levels overall, consistent with the numerical results in Table 1.

Thus, in the further step, we would use HAR model to calculate VRP.

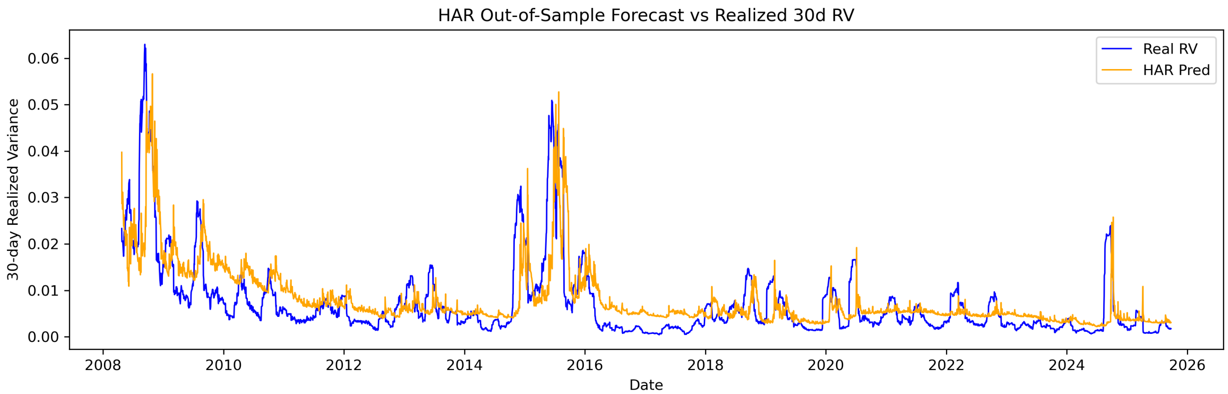

Figure 1: HAR Out-of-Sample Forecast vs Realized 30-day RV

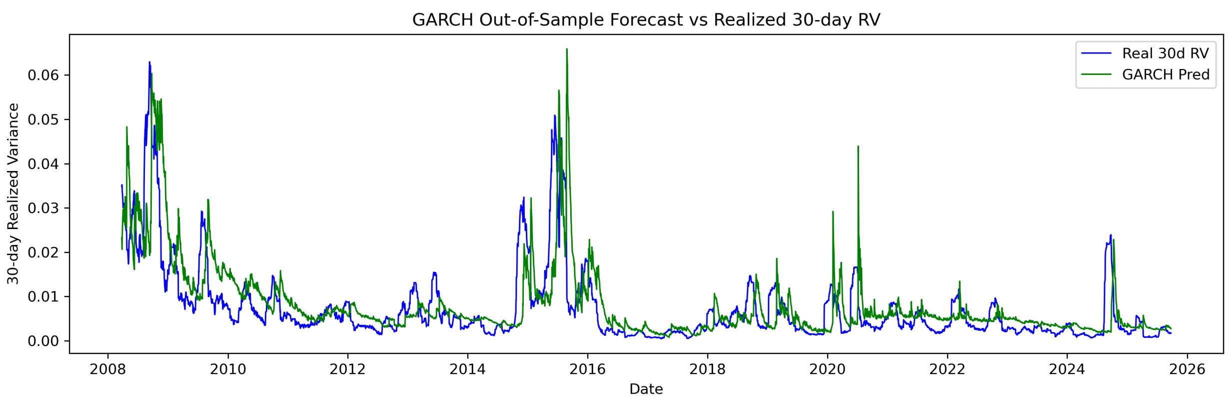

Figure 2: GARCH Out-of-Sample Forecast vs Realized 30-day RV

3 Explaining Next-Day Returns Using VRP

In the previous sections, we constructed forecasts of future 30-day realized volatility using both HAR and GARCH models, and further computed the Variance Risk Premium (VRP). In this section, we examine whether VRP can explain or predict the **next-day (1-day horizon)** index returns.

Across all specifications, including single-factor and multi-factor regressions, the explanatory power remains extremely low. This confirms the well-established finding in the asset pricing literature that daily returns are dominated by noise and are extremely difficult to predict.

3.1 Model 1: Baseline VRP Regression

We begin with the baseline model:

$$r_{t+1} = \alpha + \beta \cdot VRP_t + \varepsilon_t.$$

Regression output is as follows:

Table 2: Model 1: Regression of Next-Day Returns on VRP

Metric	Value
R^2	0.002
Adj. R^2	0.002
N	2584
F-statistic p-value	0.342
Durbin-Watson	1.973
VRP coefficient	0.2137
VRP p-value	0.341
Constant	0.0006
Constant p-value	0.159

Conclusion: VRP shows no statistical significance ($p = 0.341$), and the model explains only 0.2% of next-day returns. This is fully consistent with the literature: VRP is a low-frequency risk factor and does not predict daily returns.

3.2 Model 2: Adding Implied Variance and HAR Forecasts

We next include implied variance and the HAR-predicted realized variance:

$$r_{t+1} = \alpha + \beta_1 VRP_t + \beta_2 IV_t + \beta_3 E[RV_{t+30}] + \varepsilon_t.$$

Table 3: Model 2: Regression with VRP, Implied Variance, and HAR Forecast

Metric	Value
R^2	0.002
Adj. R^2	0.001
N	2584
F-statistic p-value	0.630
Durbin-Watson	1.972
VRP coefficient	0.1435
VRP p-value	0.358
IV_30d_var coef	0.0746
IV_30d_var p-val	0.518
HAR_pred coef	-0.0690
HAR_pred p-val	0.400
Constant	0.0006
Constant p-value	0.298

Conclusion: None of the variables are significant, and strong multicollinearity exists due to the identity $VRP = IV - HAR$, thus adding these variables does not introduce new predictive information.

3.3 Model 3: Adding Technical and Liquidity Factors

We further augment the model using return-based and volatility-based technical indicators, and trading volume:

Table 4: Model 3: Regression with VRP, Volatility, Returns, and Volume Factors

Metric	**Value**
R^2	0.008
Adj. R^2	0.005
N	2584
F-statistic p-value	0.287
Durbin-Watson	1.998
VRP coefficient	0.2095
VRP p-value	0.211
IV_30d_var coef	0.0232
IV_30d_var p-val	0.861
HAR_pred coef	-0.1863
HAR_pred p-val	0.069
ret_d coef	0.0179
ret_d p-val	0.650
ret_5d coef	0.0011
ret_5d p-val	0.947
ret_22d coef	-0.0065
ret_22d p-val	0.290
rv_5d coef	0.7267
rv_5d p-val	0.082
log(volume) coef	-0.0012
log(volume) p-val	0.023
Constant	0.0258
Constant p-value	0.019

While the R^2 increases slightly from 0.002 to 0.008, the explanatory power remains extremely low. The possible reason is slow-moving nature of volatility-structure factors (VRP, IV), or instability of technical indicators in high frequency.

Overall, VRP and related factors do not predict next-day returns.

4 Information Coefficient (IC) Analysis

To further assess the predictive content of all factors, we compute both Pearson and Spearman Information Coefficients between each factor and next-day returns:

Table 5: Information Coefficients (IC) of Factors vs Next-Day Return

Factor	**Pearson IC**	**Spearman IC**
VRP	0.0463	-0.0077
IV_30d_var	0.0082	-0.0156
HAR_pred	-0.0199	0.0024
ret_d	0.0135	-0.0086
ret_5d	-0.0017	-0.0372
ret_22d	-0.0085	-0.0103
rv_5d	0.0273	-0.0251
log(volume)	-0.0249	-0.0471

4.1 Some observations:

- All IC values are within ± 0.05, indicating no meaningful predictive power.
- VRP has a weak Pearson IC (0.046) but its Spearman IC is near zero.
- Technical return factors (ret_5d, ret_22d) display no monotonic relation with next-day returns.
- rv_5d and log(volume) exhibit minor correlation but remain too weak for practical use.

5 Next Steps

During last week's discussion, it was noted that forecasting one-month returns may involve too long a horizon, as VRP decays relatively quickly. However, our results in this section suggest that forecasting next-day returns is also inappropriate, as daily noise overwhelms all predictive signals.

Given the slow-moving nature of volatility risk premium and the strong persistence in volatility dynamics, a more suitable prediction horizon is likely in the range of: 5days, 10days, 22days. These horizons are long enough for risk-premium effects to accumulate, yet short enough to capture volatility-term-structure dynamics without excessive decay.

Also, I would need to try multiple factors that work with VRP.