Variance Risk Premium and Expected Return: Empirical Analysis on China 50ETF

Nov 2025

1 Experiment 1: Baseline Regression

The first experiment directly follows the basic setup of *Bollerslev, Tauchen, and Zhou (2009)*. We define the 30-day variance risk premium (VRP) as the difference between the implied variance and the realized variance:

$$\mathrm{VRP}_t = IV_{t,30d}^2 - RV_{t,30d}.$$

We then regress the future 30-day return on this contemporaneous VRP using non-overlapping monthly observations:

$$r_{t\to t+30} = \alpha + \beta\,\mathrm{VRP}_t + \varepsilon_t.$$

The scatter plot below shows the weak correlation between the VRP and subsequent returns.

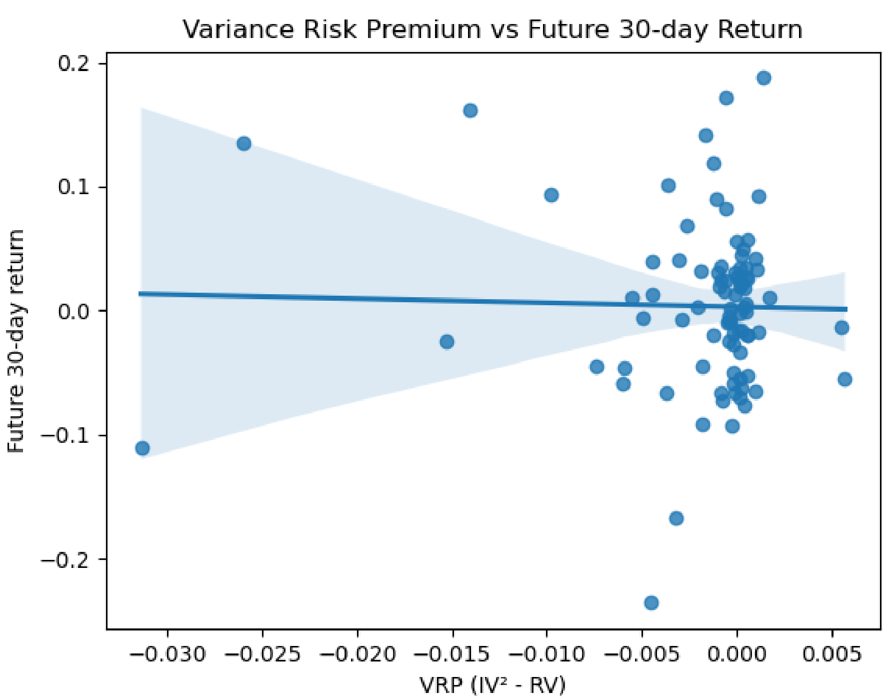

Figure 1: Baseline Regression: VRP vs Future 30-day Return (Non-overlapping)

Table 1: Baseline Regression Results (Non-overlapping)

Variable	Coefficient	Std. Error	t-stat	p-value
Constant	0.0045	0.006	0.761	0.447
VRP_30d	1.1168	2.171	0.514	0.607
R^2		0.006		
Adj. R^2		0.006		
No. of Obs.		86 (non-overlapping)		

The coefficient on VRP is positive but statistically insignificant, suggesting that the variance risk premium has limited predictive power for future returns in this sample.

2 Experiment 2: Overlapping Sample and Newey–West Correction

Since the sample is short and using non-overlapping observations reduces data points, we follow *Bollerslev et al. (2009)* and employ overlapping daily observations. However, overlapping windows introduce serial correlation in residuals because each 30-day return shares 29 days with the next. To address this, we apply the **Newey–West heteroskedasticity and autocorrelation consistent (HAC)** standard error estimator, which adjusts the covariance matrix up to a chosen lag length (30 in our case):

$$\widehat{\mathrm{Var}}(\hat{\beta}) = (X'X)^{-1}X'\widehat{\Omega}X(X'X)^{-1}, \quad \widehat{\Omega} = \sum_{k=-q}^{q} w_k \Gamma_k,$$

where $q = 30$ and w_k are Bartlett weights.

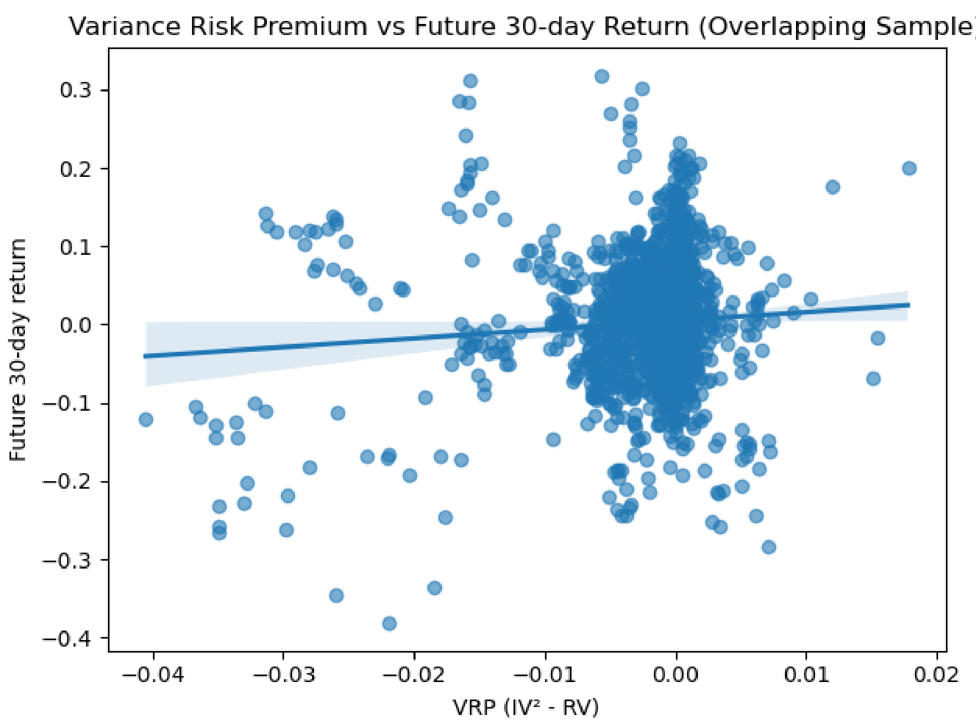

Figure 2: Overlapping Regression with Newey–West (30 lags)

Table 2: Regression Results: Overlapping Sample with Newey–West (30 lags)

Variable	Coefficient	Std. Error	t-stat	p-value
Constant	0.0045	0.006	0.761	0.447
VRP_30d	1.1168	2.171	0.514	0.607
R^2		0.006		
Adj. R^2		0.006		
No. of Obs.		2,584 (overlapping)		

While the coefficient remains insignificant, the correction ensures valid inference despite serial correlation. This step establishes a statistically sound baseline for models using overlapping horizons.

3 Experiment 3: HAR Model and Expected VRP

In this section, we refine the measure of expected variance by modeling the realized variance dynamics directly via the **Heterogeneous Autoregressive (HAR)** model of *Corsi (2009)*:

$$RV_t = \beta_0 + \beta_d RV_{t-1} + \beta_w RV_{t-1}^{(w)} + \beta_m RV_{t-1}^{(m)} + \varepsilon_t,$$

where $RV_{t-1}^{(w)}$ and $RV_{t-1}^{(m)}$ denote 5-day and 22-day moving averages of realized variance, respectively.

Using this model, we forecast the next 30-day realized variance $E_t[\widehat{RV_{t+30}}]$, and define an *expected variance risk premium*:

$$\text{VRP}_t^{exp} = IV_{t,30d}^2 - E_t[\widehat{RV_{t+30}}].$$

We then repeat the predictive regression:

$$r_{t \to t+30} = \alpha + \gamma\, \text{VRP}_t^{exp} + \varepsilon_t.$$

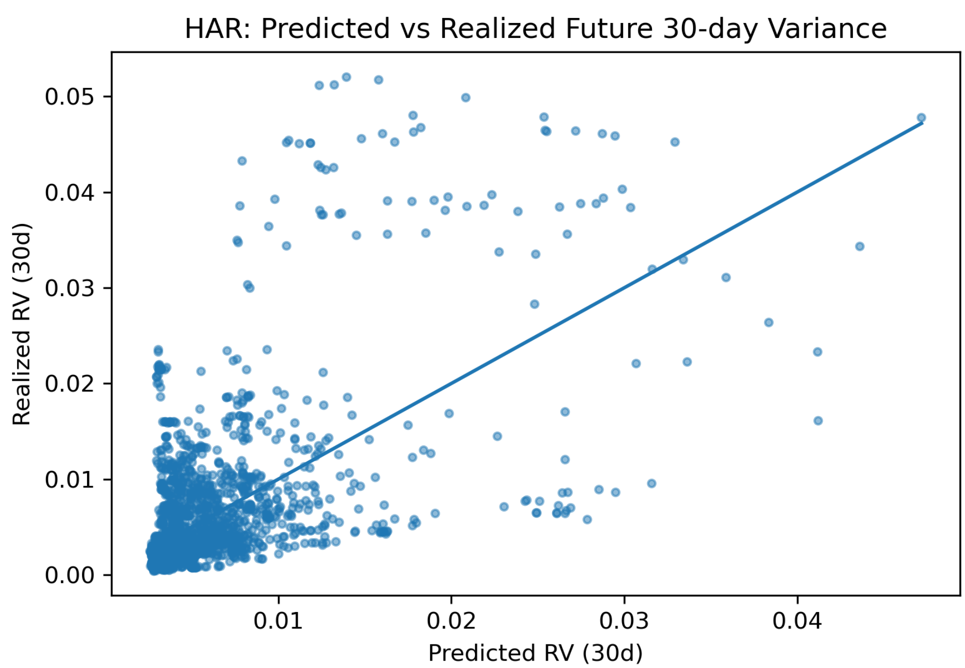

Figure 3: HAR Fit: Predicted vs Realized Future 30-day Variance

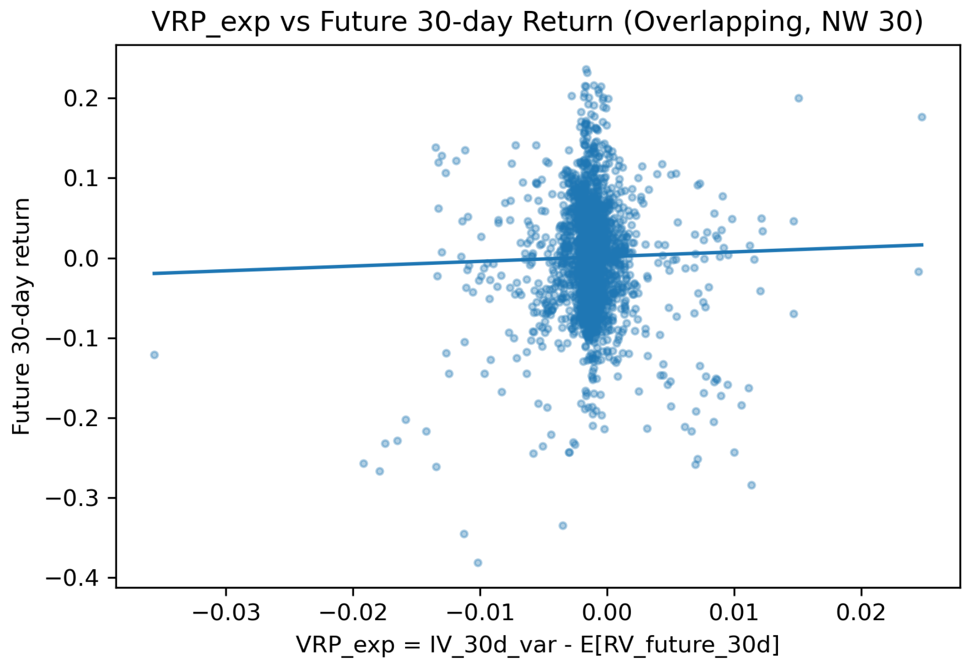

Figure 4: VRP^{exp} vs Future 30-day Return (Overlapping, NW 30)

Table 3: HAR Model for Future 30-day Realized Variance

Variable	Coefficient	Std. Error	t-stat	p-value
const	0.0024	0.0001	17.192	0.000
rv_d	1.6684	0.244	6.829	0.000
rv_w	1.0306	0.482	2.136	0.033
rv_m	13.9674	0.597	23.397	0.000
R^2		0.363		
Adj. R^2		0.363		
No. of Obs.		2,551		

Table 4: Return Regression using HAR-based VRP^{exp}

	Overlapping (NW 30)		Non-overlapping	
Variable	Coef.	p-value	Coef.	p-value
Const	0.0017	0.806	0.0056	0.508
VRP_exp_30d	0.5915	0.844	2.0006	0.444
R^2	0.001		0.007	
No. Obs.	2,551		86	

Discussion

The HAR model captures significant persistence in realized variance ($R^2 = 0.36$). However, the expected VRP constructed from this model fails to predict future returns—both overlapping and non-overlapping regressions produce insignificant coefficients. This suggests that, while volatility is predictable, the variance risk premium itself does not systematically forecast excess returns in the Chinese 50ETF market.

4 Future Work

The current study establishes a three-step baseline framework for analyzing the Variance Risk Premium (VRP) in the Chinese market. Our empirical results show that the raw ex-post VRP ($IV^2 - RV$) exhibits limited predictive power for future 30-day returns. However, after introducing the Heterogeneous Auto-Regressive (HAR) model to forecast future realized variance, we observe that the fitted HAR model performs well in explaining realized volatility dynamics ($R^2 = 0.36$), and provides a reasonable ex-ante measure of expected variance. This transformation from ex-post to ex-ante VRP lays the foundation for deeper economic interpretation and model-based applications.

Going forward, we will use HAR-predicted variance as the standard method to construct VRP:

$$VRP_t^{exp} = IV_t^2 - E_t[\widehat{RV_{t+30}}],$$

and further explore how VRP can be utilized beyond its contemporaneous relationship with returns.

Specifically, the next phase of our research will focus on two directions:

1. **Time-Series Properties.** We will study the dynamic behavior of VRP over time, including its persistence, cyclical components, and correlation with macroeconomic uncertainty. Prior works such as *Bollerslev, Tauchen, and Zhou (2009)* and *Pyun (2018)* highlight that VRP serves as a short-term predictor of equity returns and captures time-varying risk aversion. Building on these findings, we will analyze whether HAR-based ex-ante VRP in the Chinese market similarly exhibits predictive power that depends on the leverage effect and volatility shocks.

2. **Cross-Sectional Properties.** We will investigate whether VRP can explain return spreads across assets or sectors, extending the empirical design from Fama–French-style factor models. Following the framework of *Feunou et al. (2017)* and *Han and Zhou (2011)*, we aim to decompose VRP into its *bad* and *good* components, and examine whether downside-related VRP (bad VRP) drives risk compensation in the cross-section. Additionally, we plan to test whether VRP differentials across firms capture heterogeneity in exposure to aggregate volatility risk, as suggested by *Bollerslev et al. (2014)* and *Londono and Zhou (2017).*

Overall, while the baseline regressions suggest weak contemporaneous predictability, the integration of HAR-based expected variance establishes a robust foundation for using VRP as an informational variable in both the time-series and cross-sectional domains. These directions will help uncover the deeper structure behind variance compensation and its economic meaning in the Chinese equity market.

References

[1] Bollerslev, T., Tauchen, G., and Zhou, H. (2009). *Expected Stock Returns and Variance Risk Premia.* The Review of Financial Studies, 22(11), 4463–4492.

[2] Feunou, B., Lopez Aliouchkin, R., Tédongap, R., and Xi, L. (2017). *Variance Premium, Downside Risk, and Expected Stock Returns.* Bank of Canada Staff Working Paper 2017-58.

[3] Han, B., and Zhou, Y. (2011). *Variance Risk Premium and Cross-Section of Stock Returns.* SSRN Working Paper.

[4] Pyun, S. (2018). *Variance Risk in Aggregate Stock Returns and Time-Varying Return Predictability.* National University of Singapore Working Paper.

[5] Londono, J., and Zhou, H. (2017). *Variance Risk Premia and International Equity Returns.* Journal of Financial and Quantitative Analysis, 52(5), 2129–2151.