Statistical and Distributional Evaluation of a VRP Estimator

October 2025

1 Modified Methodology

At day t, what we have is:

- **CNVIX**, denoted by VIX_t, is defined to be $\mathbb{E}_{\mathbb{Q}}[\mathrm{RV}_{t,t+30}]$
- **Realized Volatility**, denoted by $\mathrm{RV}_{t-30,t}$

Now, we want to measure the VRP, defined by:

$$\mathrm{VRP}_t := \mathbb{E}_{\mathbb{Q}}\big[\mathrm{RV}_{t,t+30}\big] - \mathbb{E}_{\mathbb{P}}\big[\mathrm{RV}_{t,t+30}\big] \;\approx\; \mathrm{VIX}_t^2 - \mathbb{E}_{\mathbb{P}}\big[\mathrm{RV}_{t,t+30}\big],$$

The problem is that we do not know the $\mathbb{E}_{\mathbb{P}}\big[\mathrm{RV}_{t,t+30}\big]$, thus we currently provide three kinds of estimators, that is:

- **Trivial Estimator**: $\widehat{\mathbb{E}}_P\left[\mathrm{RV}_{t,t+30}\right] = \mathrm{RV}_{t-30,t}$
- **GARCH Estimator**
- **HAR Estimator**

Now, based on the selection of estimator of realized volatility $\widehat{\mathbb{E}}_P\left[\mathrm{RV}_{t,t+30}\right]$, we may have an estimator of VRP_t, given by:

$$\widehat{\mathrm{VRP}_t} = \mathrm{VIX}_t^2 - \widehat{\mathbb{E}}_P\left[\mathrm{RV}_{t,t+30}\right].$$

Thus, at day $t+30$, we might know the real VRP_t, we can compare the difference between the estimator and the real value.

2 Trivial Estimator

In this part, we directly used the past 30 days of realized volatility to be the estimator.

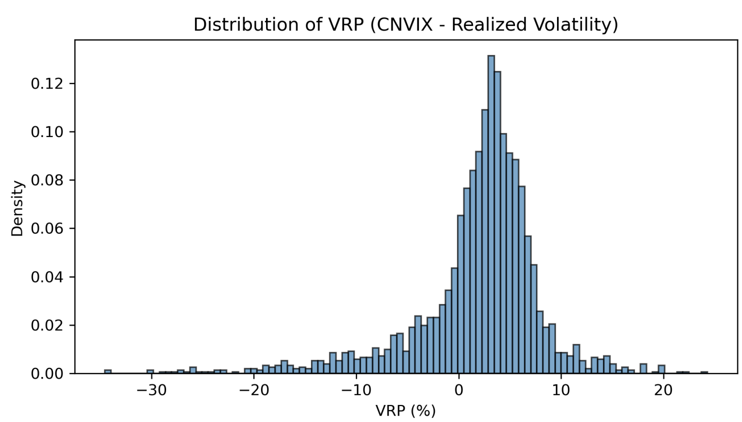

Figure 1: Distribution of $\widehat{VRP_t}$

3 Statistical Analysis of the VRP Estimator

We define the volatility risk premium (VRP) estimator as the difference between the model-implied volatility index (CNVIX) and the realized volatility:

$$\widehat{\mathrm{VRP}}_t = \mathrm{CNVIX}_t - \mathrm{RV}_{t-30,t}, \tag{1}$$

where CNVIX_t denotes the market's expectation of future volatility at time t, and $\mathrm{RV}_{t-30,t}$ is the realized volatility observed over the same 30-day horizon.

Distributional Assumption. We treat $\widehat{\mathrm{VRP}}_t$ as an independent sample drawn from an unknown continuous distribution $f(x;\theta)$ with parameter vector θ. To identify an appropriate parametric family for f, we consider several candidate distributions supported on $\mathbb{R}$ that are widely employed in modeling financial return-type variables, namely the **Normal, Student-t, Laplace, and Skew-Normal distributions**. The immediate motivation for this choice is that these distributions are defined over the entire real line, as opposed to those restricted to positive support. More fundamentally, each of them carries distinct economic interpretability, which will be elaborated upon later.

Maximum Likelihood Estimation. For a given distribution $f(x;\theta)$, the parameters are estimated via maximum likelihood:

$$\hat{\theta} = \arg\max_{\theta} \; L(\theta \mid x_1, \ldots, x_n) = \arg\max_{\theta} \prod_{i=1}^{n} f(x_i;\theta), \tag{2}$$

or equivalently by minimizing the negative log-likelihood:

$$\ell(\theta) = -\sum_{i=1}^{n} \log f(x_i;\theta). \tag{3}$$

The resulting $\hat{\theta}$ provides the maximum likelihood estimates (MLE) for the parameters of each candidate distribution.

Goodness-of-Fit Testing. To evaluate how well each distribution fits the empirical VRP data, we employ the Kolmogorov–Smirnov (KS) statistic, which measures the maximum deviation between the empirical and theoretical cumulative distribution functions:

$$D_n = \sup_x \left| F_n(x) - F(x;\hat{\theta}) \right|, \tag{4}$$

where $F_n(x)$ denotes the empirical CDF and $F(x;\hat{\theta})$ is the fitted CDF based on estimated parameters. A smaller D_n indicates a closer fit between empirical and model distributions.

Distributional Models. For clarity, the probability density functions (PDFs) of the four candidate distributions are summarized below:

- **Normal distribution:**

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right),$$

  where μ and σ represent the mean and standard deviation. It assumes symmetric and light-tailed behavior.

- **Student-t distribution:**

$$f(x;\nu,\mu,\sigma) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi}\sigma} \left[1 + \frac{(x-\mu)^2}{\nu\sigma^2} \right]^{-\frac{\nu+1}{2}},$$

  where ν controls tail thickness. Smaller ν implies heavier tails, capturing extreme variations in volatility risk premia.

- **Laplace distribution:**

$$f(x;\mu,b) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right),$$

which exhibits a sharper peak and thicker tails than the normal distribution, often used to model leptokurtic financial returns.

- **Skew-Normal distribution:**

$$f(x;\alpha,\mu,\sigma) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\alpha\frac{x-\mu}{\sigma}\right),$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal PDF and CDF, respectively. The parameter α allows for skewness, enabling asymmetric modeling of positive and negative VRP shocks.

4 Empirical Hypothesis Testing

Empirical estimation indicates that the Student-t distribution achieves the smallest KS statistic, providing the best overall fit to the observed VRP data. This suggests that the volatility risk premium exhibits heavy-tailed behavior, consistent with the occurrence of extreme deviations in market-implied versus realized volatility.

Table 1: Goodness-of-Fit Results for the VRP Estimator

Distribution	**KS Statistic**	**p-value**	**Estimated Parameters**
Student-t	0.0526	1.48×10^{-6}	$\nu = 1.966,\ \mu = 3.0185,\ \sigma = 3.0564$
Laplace	0.0736	1.96×10^{-12}	—
Skew-Normal	0.1212	4.59×10^{-33}	—
Normal	0.1460	7.07×10^{-48}	—

Notes: The Kolmogorov–Smirnov (KS) statistic measures the maximum deviation between the empirical and fitted cumulative distribution functions. A smaller KS statistic indicates a better fit. The Student-t distribution provides the best fit among candidates, consistent with the heavy-tailed nature of VRP.

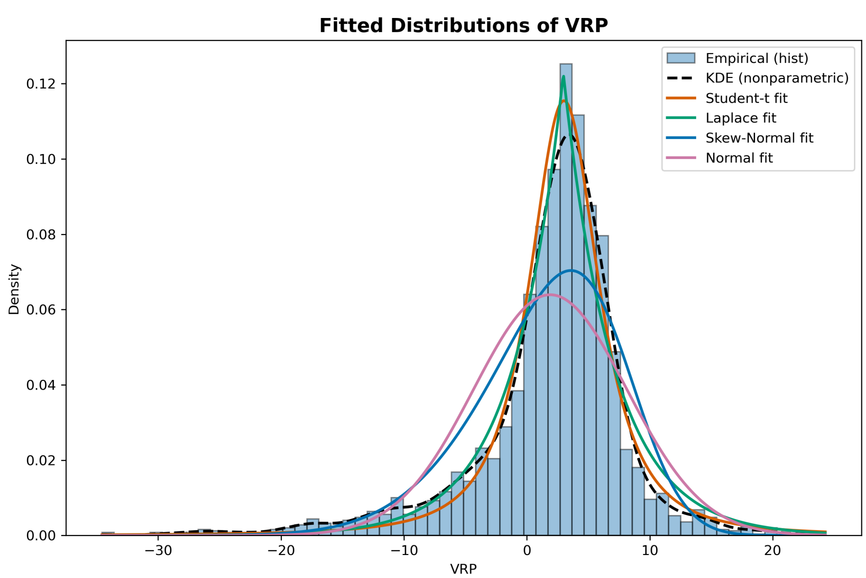

Figure 2: Hypothesis testing result of $\widehat{VRP_t}$

5 Evaluation of Estimator

To assess whether the estimated variance risk premium (VRP) $\widehat{f}_t$ is a reliable estimator of the true VRP f_t^{true}, we evaluate its performance from three complementary aspects: unbiasedness, consistency, and accuracy.

Unbiasedness. An estimator $\widehat{f}_t$ is said to be *unbiased* if its expectation equals the true value:

$$\mathbb{E}[\widehat{f}_t - f_t^{\text{true}}] = 0.$$

In empirical analysis, we test this condition by performing a one-sample t-test on the estimation errors $e_t = \widehat{f}_t - f_t^{\text{true}}$. A statistically insignificant mean error indicates the estimator has no systematic bias.

Consistency. To examine whether $\widehat{f}_t$ varies proportionally with the true value, we estimate the following regression:

$$f_t^{\text{true}} = \alpha + \beta \widehat{f}_t + \varepsilon_t.$$

A consistent estimator should yield coefficients satisfying $\alpha = 0$ and $\beta = 1$, implying both unbiasedness and correct scaling. Deviations from these values indicate the presence of systematic bias or attenuation, where the estimator may underreact or overreact to changes in the true VRP. The goodness-of-fit R^2 further measures how much of the true variation in f_t^{true} is captured by $\widehat{f}_t$.

Accuracy. While unbiasedness and consistency describe the estimator's expectation properties, *accuracy* measures the magnitude of its deviations from the truth. We compute the mean squared error (MSE) and mean absolute error (MAE):

$$\text{MSE} = \mathbb{E}\left[(\widehat{f}_t - f_t^{\text{true}})^2\right], \quad \text{MAE} = \mathbb{E}\left[|\widehat{f}_t - f_t^{\text{true}}|\right].$$

A smaller MSE or MAE indicates that the estimator provides values closer to the true underlying VRP, even if some bias remains.

Distributional Divergence (KL Divergence). To complement the moment-based and regression-based evaluations, we further assess the estimator from a distributional perspective. Specifically, we measure the information discrepancy between the empirical distribution of the estimator $\widehat{\text{VRP}}_t$ and that of the true VRP $\text{VRP}_t^{\text{true}}$ using the Kullback–Leibler (KL) divergence:

$$D_{\text{KL}}\big(p(\widehat{\text{VRP}}) \,\|\, q(\text{VRP}^{\text{true}})\big) = \int p(x) \log \frac{p(x)}{q(x)}\, dx, \tag{5}$$

where $p(x)$ denotes the probability density of the estimated VRP and $q(x)$ the density of the true VRP. The KL divergence quantifies the information loss when the distribution of the true VRP is approximated by that of the estimator.

From an economic standpoint, a nonzero KL divergence between the estimated and true VRP distributions has direct interpretative significance. Since the estimator $\widehat{\text{VRP}}_t$ is constructed using past realized volatility over the historical window $(t-30, t)$, it reflects *backward-looking* information about market uncertainty. In contrast, the true VRP—defined as $\text{CNVIX}_t - \text{RV}_{t,t+30}$—is *forward-looking*, capturing investors' expectations of volatility risk over the next 30 days.

Hence, a positive $D_{\text{KL}}\big(p(\widehat{\text{VRP}}) \,\|\, q(\text{VRP}^{\text{true}})\big)$ quantifies the extent of information loss caused by this temporal mismatch. Economically, this divergence reflects how much past volatility dynamics fail to represent market participants' expectations of future volatility risk. A larger divergence indicates that historical information provides limited predictive content for upcoming volatility compensation, potentially due to shifts in risk sentiment, structural breaks, or regime changes in volatility persistence. In this sense, the KL divergence measures not only a statistical discrepancy but also the *economic distance* between backward-looking and forward-looking assessments of the volatility risk premium.

Visualization. In addition to numerical tests, we visualize (i) the scatter plot of $\widehat{f}_t$ against f_t^{true}, comparing the fitted regression line to the ideal 45° line ($y = x$), and (ii) the empirical distribution of estimation errors e_t. These plots help diagnose potential nonlinearity or heteroskedasticity in the estimation process.

Overall, a good estimator should be approximately unbiased, linearly consistent with the true variable, and display low variance in errors.

6 Empirical Evaluation of the VRP Estimator

We assess the performance of the VRP estimator $\widehat{\text{VRP}}_t = \text{CNVIX}_t - \text{RV}_{t-30,t}$ relative to the true volatility risk premium $\text{VRP}_t^{\text{true}} = \text{CNVIX}_t - \text{RV}_{t,t+30}$. Table 2 summarizes the main diagnostic statistics.

Table 2: Statistical Evaluation of the VRP Estimator

Category	Statistic	Estimate	Std. Error	t-value	p-value
Unbiasedness	Mean(error)	-0.307	9.272	-1.662	0.0966
Consistency	α	1.482	0.174	–	–
	β	0.364	0.027	–	$< 10^{-40}$
	R^2	0.069	–	–	–
Slope $\approx$ 1 test	$(H_0 : \beta = 1)$	–	–	-23.836	2.12×10^{-113}
Accuracy	MSE	86.04	–	–	–
	RMSE	9.28	–	–	–
	MAE	6.31	–	–	–
Distributional Divergence	$D_{\text{KL}}(p\|q)$	0.111	–	–	–

Unbiasedness. The estimator's mean error of -0.31 with a p-value of 0.0966 indicates that the bias is statistically insignificant at the 5% level. Hence, the VRP estimator is approximately unbiased on average.

Consistency. The regression

$$\text{VRP}_t^{\text{true}} = \alpha + \beta\, \widehat{\text{VRP}}_t + \varepsilon_t$$

yields $\hat{\alpha} = 1.48$ and $\hat{\beta} = 0.36$ with $R^2 = 0.069$. The strong rejection of $H_0 : \beta = 1$ implies that the estimated VRP underreacts to true VRP fluctuations—large positive (negative) realizations are systematically compressed toward the mean. The low R^2 suggests that the estimator captures only a small share of the true variation.

Accuracy. With RMSE = 9.28 and MAE = 6.31, the estimator's deviations from the true VRP are moderate but nontrivial. This pattern, together with the heavy-tailed error distribution observed in Figure 4, indicates that the estimator occasionally produces large deviations when market conditions shift abruptly.

Distributional Divergence. The KL divergence between the estimated and true VRP distributions is $D_{\text{KL}}(p\|q) = 0.111$. This small but nonzero value indicates that the empirical distribution of the estimator retains most of the information content of the true VRP, but with mild distortion. From an economic standpoint, the positive divergence reflects the *temporal information loss* arising from the 30-day lag structure of $\widehat{\text{VRP}}_t$: the estimator relies on past realized volatility, whereas the true VRP is inherently forward-looking. Hence, the divergence quantifies how much historical volatility fails to fully represent investors' expectations of future volatility risk. The moderate magnitude ($D_{\text{KL}} = 0.111$) suggests that although the estimator captures the general shape of the true risk-premium distribution, it still omits part of the market's anticipatory component, likely due to time-varying sentiment and volatility-regime shifts.

Overall Assessment. The VRP estimator is approximately unbiased but exhibits attenuation bias ($\beta < 1$), implying that it systematically underestimates the magnitude of true VRP movements. This attenuation likely arises from measurement noise and the temporal mismatch between the model-implied (forward-looking) and realized (backward-looking) volatility components.

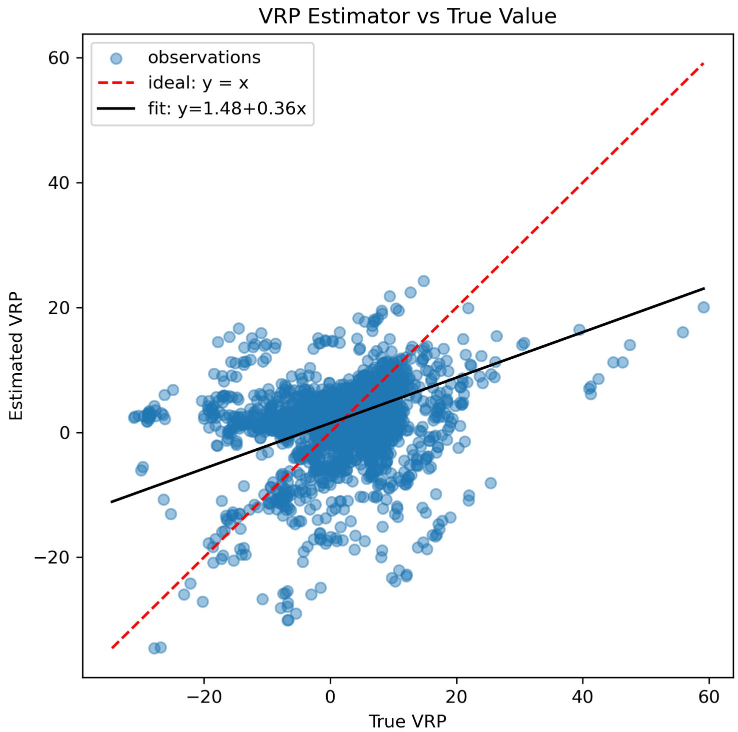

Figure 3: $\widehat{VRP_t}$ v.s. VRP_t

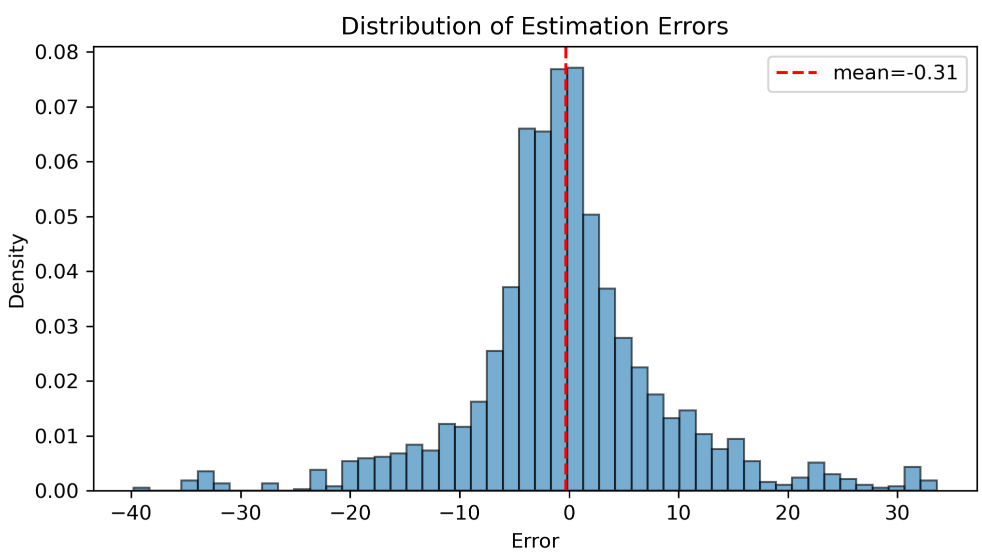

Figure 4: Distribution of VRP estimation error

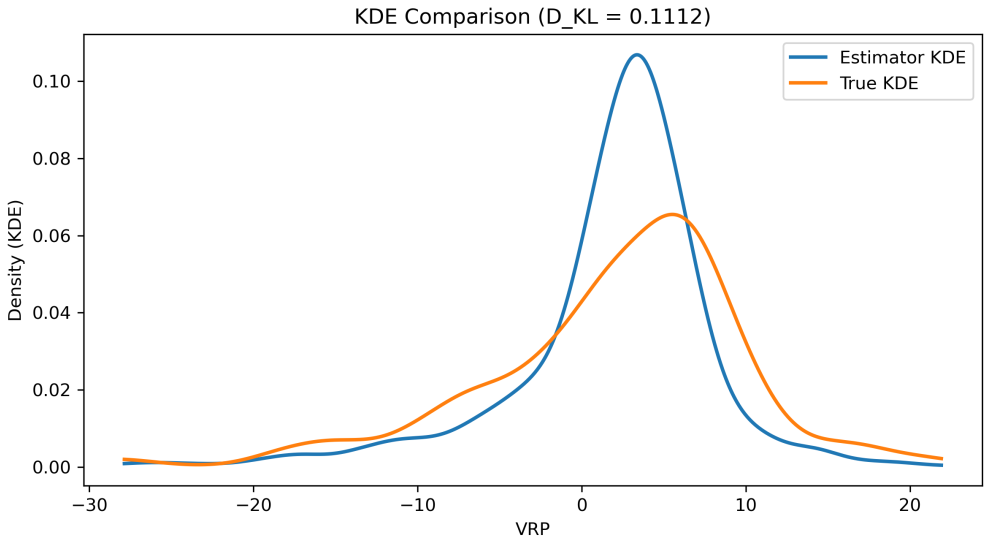

Figure 5: KDE comparison between estimated and true VRP distributions