EIDGENÖSSISCHE TECHNISCHE HOCHSCHULE ZÜRICH

MILLENNIUM CAPITAL PARTNERS

MASTER THESIS - PROJECT DESCRIPTION

263-0800-00L

Enhanced Volatility Forecasting and Regime Detection for Options Trading Using HAR Models and Spectral Clustering

---

Author
Pierre DORGE

ID Number
17-817-404

Professors	*Role*
Dr. Scott JONES, MCP	MAIN EXTERNAL SUPERVISOR
Prof. Roger WATTENHOFFER, ETHZ	PROFESSOR
PhD student BERDOZ, ETHZ	MAIN-INTERNAL SUPERVISOR
PhD student DELACOUR, ETHZ	CO-INTERNAL SUPERVISOR

September 10, 2024

Introduction

Accurately predicting future volatility is essential for volatility options traders, as it directly impacts their ability to price options, manage risk, and optimize returns. It is one of the bottlenecks to any underlying trading strategy. By improving the HAR volatility model [1] with more accurate estimation techniques and applying cutting-edge regime detection methods on predicted volatility time series [2], the project aims to build a successful trading strategy. The ability to accurately forecast volatility and dynamically adjust trading strategies based on identified market regimes can lead to better risk management, more efficient capital allocation, and ultimately higher returns.
This project unfolds into three parts:

1) Estimate future volatility using the HAR volatility model
2) Split the predicted volatility time series into stationary intervals
3) Cluster the intervals into distinct volatility regimes
4) Develop profitable strategies for each volatility regime

I. Estimate future volatility using the HAR volatility model

A. Definition of Volatility

Volatility refers to the degree of variation in the price of a financial instrument over time, typically measured by the standard deviation of returns. It reflects the uncertainty or risk associated with the price movements of assets. High volatility indicates large price movements, while low volatility suggests more stable price behavior.

B. Realized Volatility

Realized volatility is a specific measure of volatility calculated using historical price data within a fixed time window, typically on an intraday or daily basis. It is computed as the square root of the sum of squared returns over the chosen period. The formula for realized volatility V_t over a period of M days is given by:

$$\mathcal{V}_t = \sqrt{\mathcal{RV}_t} = \sqrt{\sum_{i=1}^{N} r_{t,i}^2}$$

where RV_t is called the realized variance and $r_{t,i}$ represents the return on period i within day t.

C. The HAR and HARQ volatility models

The HAR volatility model

The HAR (heterogeneous autoregressive) volatility model is a family of models. The original one estimates return volatility using raw realized variance and ordinary least squares (OLS). However, various transformations can be applied to RV and different estimators can be used to achieve better predictions.
The HAR volatility model approximates future RV as a linear function of past realized variance. According to the past day, week and month values:

$$\mathcal{RV}_t = \beta_0 + \beta_1 \mathcal{RV}_{t-1}^d + \beta_2 \mathcal{RV}_{t-1}^w + \beta_3 \mathcal{RV}_{t-1}^m + u_t,$$

where the components are defined as:

$$\mathcal{RV}_{t-1}^d = \mathcal{RV}_{t-1}, \quad \mathcal{RV}_{t-1}^w = \frac{1}{5}\sum_{i=1}^{5} \mathcal{RV}_{t-i}, \quad \mathcal{RV}_{t-1}^m = \frac{1}{22}\sum_{i=1}^{22} \mathcal{RV}_{t-i}.$$

In this notation: $\mathcal{RV}_t$ is the realized variance on day t, $\mathcal{RV}_{t-1}^d$ is the daily lagged realized variance, $\mathcal{RV}_{t-1}^w$ is the weekly lagged realized variance, and $\mathcal{RV}_{t-1}^m$ is the monthly lagged realized variance. The β_j (for $j = 0, 1, 2, 3$) are the unknown parameters to be estimated, and u_t represents the error term.

The HARQ volatility model

The HARQ model, proposed by [3], is an extension of the standard HAR model that incorporates the estimation error of realized variance ($\mathcal{RV}$) by using realized quarticity ($\mathcal{RQ}$).
The full HARQ model can be expressed as:

$$\mathcal{RV}_t = \beta_0 + \left(\beta_1 + \beta_{1Q}\sqrt{\mathcal{RQ}_{t-1}^d}\right)\mathcal{RV}_{t-1}^d + \left(\beta_2 + \beta_{2Q}\sqrt{\mathcal{RQ}_{t-1}^w}\right)\mathcal{RV}_{t-1}^w + \left(\beta_3 + \beta_{3Q}\sqrt{\mathcal{RQ}_{t-1}^m}\right)\mathcal{RV}_{t-1}^m + u_t,$$

where $\mathcal{RV}_t$ is the realized variance on day t, and $\mathcal{RQ}^d_{t-1}$, $\mathcal{RQ}^w_{t-1}$, and $\mathcal{RQ}^m_{t-1}$ denote the daily, weekly, and monthly lagged realized quarticity, respectively.
A simplified version of the HARQ model, which focuses on short-term forecasting, is given by:

$$\mathcal{RV}_t = \beta_0 + \left(\beta_1 + \beta_{1Q}\sqrt{\mathcal{RQ}^d_{t-1}}\right)\mathcal{RV}^d_{t-1} + \beta_2\mathcal{RV}^w_{t-1} + \beta_3\mathcal{RV}^m_{t-1} + u_t.$$

This model adjusts the weight placed on historical observations of daily realized variance ($\mathcal{RV}^d_{t-1}$) based on the realized quarticity, mitigating the impact of measurement error and improving forecast accuracy.

Transformations and Estimators

In [4], the original HAR and HARQ models with raw realized variance and OLS are used as a benchmark. It is shown that applying transformations such as the qr, log or Box-Cox transformation can lead to better variance-stabilization of the data. As well as using more robust estimators such as WLS or RR combined with transformations can lead to better results.

II. Split the predicted volatility time series into stationary intervals

To split the volatility time series, we will apply the nonparametric Mood test to extract stationary segments with respect to variance. Given the time series of log returns $\{\mathcal{R}_t\}_{t=1}^T$, where $\mathcal{R}_t = \log\left(\frac{P_t}{P_{t-1}}\right)$, P_t being the closing price of an asset on day t, we assume the log returns $\mathcal{R}_t$ are independent random variables with a mean close to zero, without assuming any particular distribution [5].
Using the Mood test, we detect changes in the variance of the time series. Although primarily known as a median test, the Mood test is also suitable for detecting changes in variance between two distributions, as described in Section 4 of Mood [6]. The test identifies a set of change points $\{\tau_1, \tau_2, \ldots, \tau_{m-1}\}$, where for convenience, we define $\tau_0 = 1$ and $\tau_m = T$.
This partitioning yields m stationary segments of the log return series, denoted as $\{\mathcal{R}_t\}_{t\in[\tau_{j-1},\tau_j]}$ for $j = 1, 2, \ldots, m$. Each segment $\mathcal{Y}^{(j)}$ is the restricted time series of $\mathcal{R}_t$ over the interval $[\tau_{j-1}, \tau_j]$. The algorithm determines that each $\mathcal{Y}^{(j)}$ is sampled from a consistent distribution, ensuring stationarity within each segment.

III. Cluster the intervals into distinct volatility regimes

This step consists of identifying patterns in the stationary segments identified earlier.

A. Finding the optimal number of classes

To do so, we construct the empirical cumulative distribution functions for each segment $\mathcal{Y}^{(j)}$ and compute the Wasserstein distance matrix $\mathcal{D}_{ij}$ between each distribution to identify clusters. The Wasserstein distance $\mathcal{W}_p(\mu_i, \mu_j)$, where $p = 1$, is computed as:

$$\mathcal{W}_p(\mu_i, \mu_j) = \left(\int_{\mathbb{R}} |F_i(x) - F_j(x)|^p \, dx\right)^{\frac{1}{p}},$$

where F_i and F_j are the cumulative distribution functions of the distributions μ_i and μ_j. This results in an $m \times m$ distance matrix $\mathcal{D}_{ij} = \mathcal{W}_p(\mu_i, \mu_j)$, capturing the distances between the m distributions of the stationary segments.

Next, we apply spectral clustering directly to the Wasserstein distance matrix $\mathcal{D}$ to identify patterns in the cumulative distribution functions. We use the self-tuning method from [7] to define the affinity matrix $\mathcal{A}$ with:

$$\mathcal{A}_{i,j} = \exp\left(-\frac{\mathcal{D}^2_{ij}}{\sigma_i\sigma_j}\right),$$

where σ_i is chosen as the K-th nearest neighbor distance, with $K = \lceil\sqrt{m}\rceil$, with m being the total number of segments following Hassanat et al. (2014). We then construct the normalized Laplacian matrix $\mathcal{L}_{\text{sym}}$ and determine the number of clusters k that either minimize the alignment cost with the cannonical coordinate system of the matrix spanned by the eigenvectors corresponding to the k largest eigenvalues [7] or that maximize the eigengap between successive largest eigenvalues.

B. Assigning labels to segments with spectral clustering

Having determined the number of clusters k, we compute the normalized eigenvectors $u_1, \ldots, u_k$ corresponding to the k smallest eigenvalues of $\mathcal{L}_{\text{sym}}$. These eigenvectors form the matrix $\mathcal{U} \in \mathbb{R}^{m \times k}$, where the rows $\sqsubseteq_i \in \mathbb{R}^k$ represent the data points $i = 1, \ldots, m$. We then apply k-means clustering to these rows $\sqsubseteq_i$ to group them into k clusters $\mathcal{C}_1, \ldots, \mathcal{C}_k$. Finally, we assign the original m elements (segment distributions) to the corresponding clusters $\mathcal{A}_l = \{i : \sqsubseteq_i \in \mathcal{C}_l\}$, for $l = 1, \ldots, k$.

IV. Develop profitable strategies for each volatility regime

In the final part of this project, we will explore and develop trading strategies that take advantage of the volatility regimes identified and optimize returns by dynamically allocating assets based on the current market environment.

Momentum-Based Basket Allocations:

Rather than merely reallocating between asset classes as in [2], we propose a more nuanced strategy that involves reallocating within the equity asset class itself. Specifically, we will define and monitor "momentum baskets," which consist of stocks that have shown significant upward movement with relatively low volatility (or noise). Which could be done by focusing the Mood Test on mean changes. These stocks, often driven by strong recent performance, are particularly susceptible to downturns when market regimes shift. Which could lead us to re-allocating to different equity asset classes.

Slow Reallocation Back to Risky Assets:

Volatility regimes are not static, and the persistence of a regime can affect asset performance over time. In some cases, risky assets may begin to outperform as a high-volatility regime persists, especially if no recent shocks have occurred. To capture this dynamic, we propose a strategy that gradually reallocates capital back into riskier assets as the regime progresses, using a slow reallocation approach guided by an exponential decay factor. This reallocation strategy allows for a more responsive approach to changing market conditions, where the model slowly increases exposure to risky assets as market stability improves, rather than waiting for a complete regime shift to occur.

Strategy Performance Metrics:

To evaluate the effectiveness of the proposed strategies, we will employ several standard performance measures. These include:

- **Sharpe Ratio:** A measure of risk-adjusted return, calculated by dividing the excess return over the risk-free rate by the standard deviation of the portfolio's returns.
- **Sortino Ratio:** Similar to the Sharpe Ratio but only penalizes downside volatility, providing a clearer view of the risk-adjusted returns when negative returns are a key concern.
- **Maximum Drawdown:** The largest peak-to-trough decline in portfolio value, indicating the most significant loss from a high point to a low point.
- **Annual Return (AR):** The annual growth rate of the portfolio's value over a specified period.
- **Portfolio Standard Deviation:** A measure of the portfolio's return variability over time, highlighting the degree of risk involved.
- **Calmar Ratio (CR):** A measure of risk-adjusted return, calculated by dividing the annualized return of an investment by its Maximum Drawdown over the same period. It evaluates how well an investment compensates for the risk taken, as represented by the maximum drawdown.

These metrics will be used to quantitatively assess the performance of each strategy across different volatility regimes.

Data

For this project, we have at our disposal high-frequency data on all 500 components of the SPX index, along with related ETFs, futures, and indices. The dataset includes price information at 5-second intervals. Given our focus on options trading, we will likely focus on a subset of the stocks, selecting either the top 100 ranked by index weight and / or by option liquidity. In addition to individual stocks, we will also incorporate relevant ETFs into our analysis. The availability of such high-frequency data enables us to conduct our analyses at a higher temporal resolution than is typically done in classical HAR(Q) model studies.

References

[1] Fulvio Corsi. A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2):174–196, 02 2009.

[2] Arjun Prakash, Nick James, Max Menzies, and Gilad Francis. Structural clustering of volatility regimes for dynamic trading strategies. *Applied Mathematical Finance*, 28(3):236–274, May 2021.

[3] Tim Bollerslev, Andrew J. Patton, and Rogier Quaedvlieg. Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192(1):1–18, 2016.

[4] Adam Clements and Daniel P.A. Preve. A practical guide to harnessing the har volatility model. *Journal of Banking Finance*, 133:106285, 2021.

[5] Gordon J. Ross. Modelling financial volatility in the presence of abrupt changes. *Physica A: Statistical Mechanics and its Applications*, 392(2):350–360, January 2013.

[6] A. M. Mood. On the Asymptotic Efficiency of Certain Nonparametric Two-Sample Tests. *The Annals of Mathematical Statistics*, 25(3):514 – 522, 1954.

[7] Lihi Zelnik-manor and Pietro Perona. Self-tuning spectral clustering. In L. Saul, Y. Weiss, and L. Bottou, editors, *Advances in Neural Information Processing Systems*, volume 17. MIT Press, 2004.