Université de Rennes Faculté des sciences économiques

2024 – 2025

APPLIED VOLATILTY MODELS PROJECT

GARCH Models And VaR Backtesting On Equity Portfolios

VO Duc Toan - TU Ngoc Lien

tungoclien.rennes1@st.ueh.edu.vn- voductoan.rennes1@st.ueh.edu.vn

[Applied Volatility Models - Bachelor (UEH)]

Table of content

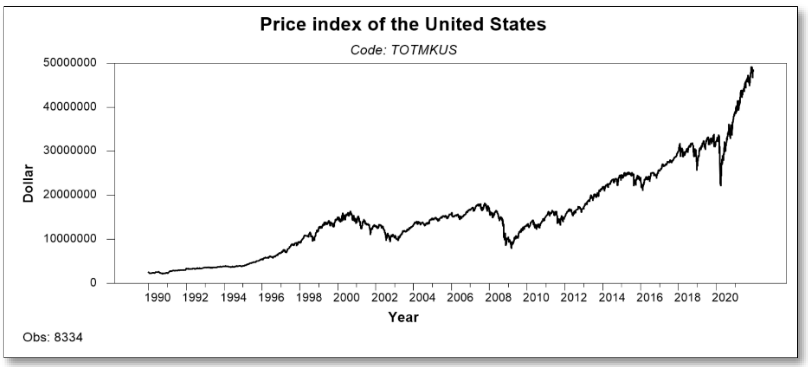
1/ Data collection

Economic data for this analysis was sourced from Datastream, a comprehensive online database of economic time series using WinRATS, a specialized software for time series analysis favored by economists and researchers.

2/ Econometric models of equity market returns

2.1/ Total market United States market value (US data)

a. Overview



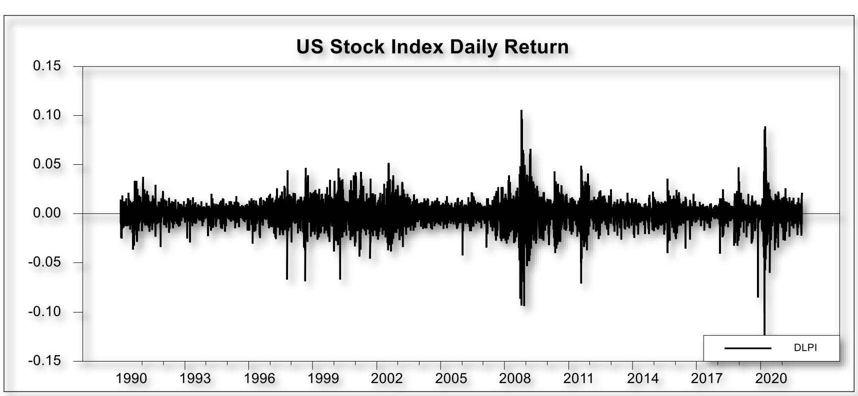
The analysis utilizes the "Total market United States Market Value" dataset (TOTMKUS) containing 8334 daily observations (dollars, the period of 01/01/1990 - 9/12/2021.). This data reflects the stock price index of the United State market

Based on the graph, it appears there's a **gradual upward trend** in total market value from 1990 to 2021 (end-year of the databases).

This suggests a **steady growth** over several years. However, there's a **sharp decrease and increase** in market value following COVID-19 (2019 onwards), which was the general situation of the whole world at that time.

Therefore, the coefficients, indices and graphs calculated at *this point will fluctuate more significantly than the rest of the period*.

b. Test the stationarity of the yield series



The **Dickey-Fuller** unit root test results for the logarithmic return series **(DLPI)** from 1990 to 2021

```
Dickey-Fuller Unit Root Test, Series DLPI
Regression Run From 1990:01:10 to 2021:12:09
Observations        8328
With intercept
Using fixed lags 5

Sig Level    Crit Value
1%(**)        -3.4342
5%(*)         -2.8624
10%           -2.5673

T-Statistic  -40.7159**
```

(n=8,328) with **5 lags** and an intercept term yield an **ADF** statistic of -40.7159, significantly below the 1% critical value (-3.4342). This ***strongly rejects the null hypothesis*** strictly stationary. This result satisfies the prerequisite for applying ARCH/GARCH models, which require stationary input data to ensure robust conditional variance estimates in financial risk analysis.

c. Model 1: Simple OLS with Homoscedastic Errors

```
Linear Regression - Estimation by Least Squares
Dependent Variable DLPI
Daily(5) Data From 1990:01:02 To 2021:12:09
Usable Observations                    8333
Degrees of Freedom                     8332
Centered R^2                     -0.0000000
R-Bar^2                          -0.0000000
Uncentered R^2                    0.0010451
Mean of Dependent Variable     0.0003571872
Std Error of Dependent Variable 0.0110437255
Standard Error of Estimate     0.0110437255
Sum of Squared Residuals       1.0162029962
Log Likelihood                   25724.0902
Durbin-Watson Statistic              2.1433

    Variable                        Coeff      Std Error      T-Stat      Signif
************************************************************************************
1.  Constant                    0.0003571872 0.0001209804      2.95244  0.00316154
```

```
Homoscedastic OLS Results:
-------------------------
Intercept (b0) =    3.57187e-04
Residual Variance (s²) =    1.21964e-04
AIC =      135.93738 BIC =      142.96536
```

The estimation results of the simple OLS model with Homoskedastic errors show that the mean value of the dependent variable DLP1 is 0.000357 (approximately 0.0357% per day) and is statistically significant at the 1% level. However, since the model includes only a constant term and no explanatory variables, the **R^2 is zero**, indicating that it does not explain any of the variation in the data. The Durbin-Watson test **(2.1433)** suggests that there is no first-order autocorrelation in the errors. This is the most basic model, suitable only for estimating the mean of the series. For a more in-depth analysis, it is necessary to consider additional independent variables or more complex models such as ARCH/GARCH if the data exhibits heteroskedasticity.

Test the ARCH (5) and ARCH (2) effects using LM test

Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	0.0000377110	0.0000047183	7.99257	0.00000000
2. RES2{1}	0.1496459674	0.0108000050	13.85610	0.00000000
3. RES2{2}	0.3139347396	0.0109074630	28.78165	0.00000000
4. RES2{3}	0.0013819760	0.0114379964	0.12082	0.90383394
5. RES2{4}	0.0547480910	0.0109078610	5.01914	0.00000053
6. RES2{5}	0.1712378941	0.0108004281	15.85473	0.00000000

Chi-Squared(5)= 2074.499332 with Significance Level 0.00000000

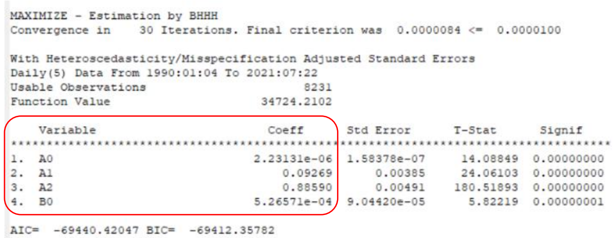
Variable	Coeff	Std Error	T-Stat	Signif
1. Constant	0.0000533624	0.0000047219	11.30111	0.0000000
2. RES2{1}	0.1956817224	0.0101943437	19.19513	0.0000000
3. RES2{2}	0.3667462050	0.0101943537	35.97542	0.0000000

Chi-Squared(2)= 1809.049223 with Significance Level 0.0000000

The ARCH test results provide clear evidence of heteroskedasticity, with a p-value of 0.000 in both **ARCH (5)** and **ARCH (2)** cases, where the lag-1, lag-2 coefficients are highly statistically significant **(t-stat > 13)**. Notably, the sum of coefficients in the **ARCH (2)** model reaches **0.562**, indicating persistent volatility while still satisfying the stationarity condition (sum < 1).

Therefore, the GARCH (1,1) model is the next optimal choice to simultaneously capture short-term effects (through the ARCH coefficient α) and long-term trends (through the GARCH coefficient β) of volatility, while also conserving degrees of freedom compared to higher-order ARCH models.

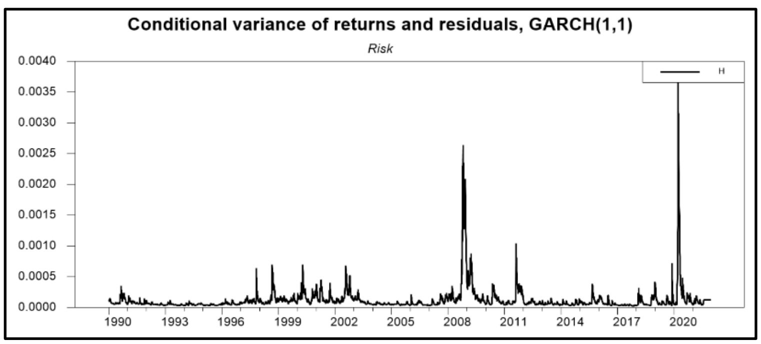
d. Model 2: Estimate GARCH (1,1) model



```
MAXIMIZE - Estimation by BHHH
Convergence in    30 Iterations. Final criterion was  0.0000084 <=  0.0000100

With Heteroscedasticity/Misspecification Adjusted Standard Errors
Daily(5) Data From 1990:01:04 To 2021:07:22
Usable Observations                          8231
Function Value                         34724.2102

    Variable                         Coeff      Std Error      T-Stat      Signif
*******************************************************************************
1.  A0                           2.23131e-06  1.58378e-07     14.08849  0.00000000
2.  A1                               0.09269      0.00385     24.06103  0.00000000
3.  A2                               0.88590      0.00491    180.51893  0.00000000
4.  B0                           5.26571e-04  9.04420e-05      5.82219  0.00000001

AIC=  -69440.42047 BIC=  -69412.35782
```

The GARCH (1,1) estimation results show all coefficients are statistically significant at the 1% level, with the influence of past variance **(A2 = 0.886)** dominating short-term shocks (A1 = 0.093), reflecting high persistence in market volatility. The **low AIC (-69,440.42)** and **BIC (-69,412.36)** values demonstrate GARCH (1,1)'s superior fit. The conditional variance plot from 1990-2021 clearly shows volatility clusters coinciding with major crises like 2008 and 2020, with fluctuation ranges from 0.0000 to 0.0040, confirming this model's effectiveness in capturing both short-term and long-term market risk dynamics.

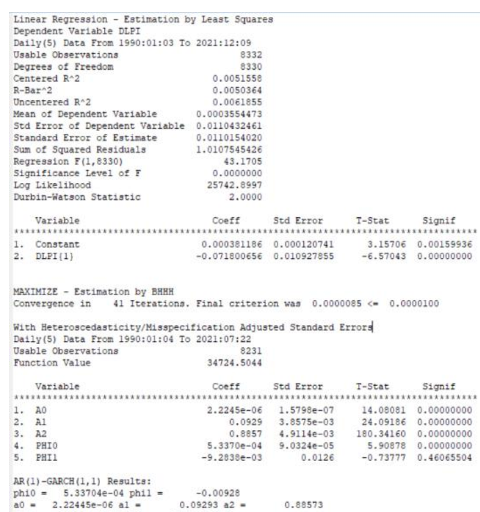
GARCH (1,1) is "**smoother**" because it considers historical variance (ht−1), reducing short-term volatility compared to ARCH. Use this result to estimate VaR or optimize portfolios during volatile periods.



=> Suitable for VaR calculation and risk forecasting due to its ability to model volatility persistence.

e. Model 3: Estimate AR (1) GARCH (1,1) model

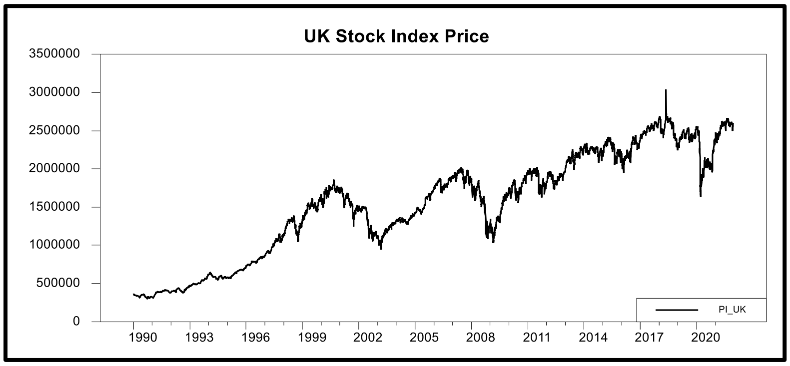
The estimation results of the AR(1)-GARCH(1,1) model reveal that the first-order autoregressive term *(AR(1)) is statistically insignificant (p-value = 0.461)*, while all GARCH(1,1) coefficients are significant at the 1% level *(A1 = 0.0929, A2 = 0.8857)*, demonstrating strong persistence in market volatility with an ARCH+GARCH sum of 0.9786 satisfying stationarity conditions. The conditional variance plot clearly shows volatility clustering during the 2008 and 2020 crisis periods, mirroring results from the standard GARCH (1,1) model. The AIC (-69,439.01) and BIC (-69,403.93) values indicate that incorporating the AR (1) component does not significantly improve model fit compared to the simpler GARCH (1,1) specification.



```
Linear Regression - Estimation by Least Squares
Dependent Variable DLPI
Daily(5) Data From 1990:01:03 To 2021:12:09
Usable Observations                          8332
Degrees of Freedom                           8330
Centered R^2                            0.0051558
R-Bar^2                                 0.0050364
Uncentered R^2                          0.0061855
Mean of Dependent Variable           0.0003554473
Std Error of Dependent Variable      0.0110432461
Standard Error of Estimate           0.0110154020
Sum of Squared Residuals             1.0107545426
Regression F(1,8330)                      43.1705
Significance Level of F                 0.0000000
Log Likelihood                         25742.8997
Durbin-Watson Statistic                    2.0000

    Variable                         Coeff      Std Error      T-Stat      Signif
*******************************************************************************
1.  Constant                     0.000381186  0.000120741      3.15706  0.00159936
2.  DLPI{1}                     -0.071800656  0.010927855     -6.57043  0.00000000


MAXIMIZE - Estimation by BHHH
Convergence in    41 Iterations. Final criterion was  0.0000085 <=  0.0000100

With Heteroscedasticity/Misspecification Adjusted Standard Errors
Daily(5) Data From 1990:01:04 To 2021:07:22
Usable Observations                          8231
Function Value                         34724.5044

    Variable                         Coeff      Std Error      T-Stat      Signif
*******************************************************************************
1.  A0                            2.2245e-06   1.5798e-07     14.08081  0.00000000
2.  A1                                0.0929   3.8575e-03     24.09186  0.00000000
3.  A2                                0.8857   4.9114e-03    180.34160  0.00000000
4.  PHI0                          5.3370e-04   9.0324e-05      5.90878  0.00000000
5.  PHI1                         -9.2838e-03       0.0126     -0.73777  0.46065504

AR(1)-GARCH(1,1) Results:
phi0 =   5.33704e-04 phil =      -0.00928
a0 =   2.22445e-06 a1 =       0.09293 a2 =       0.88573
AIC =  -69439.00879 BIC =  -69403.93048
```

2.2/ Total market United Kingdom market value (UK data)

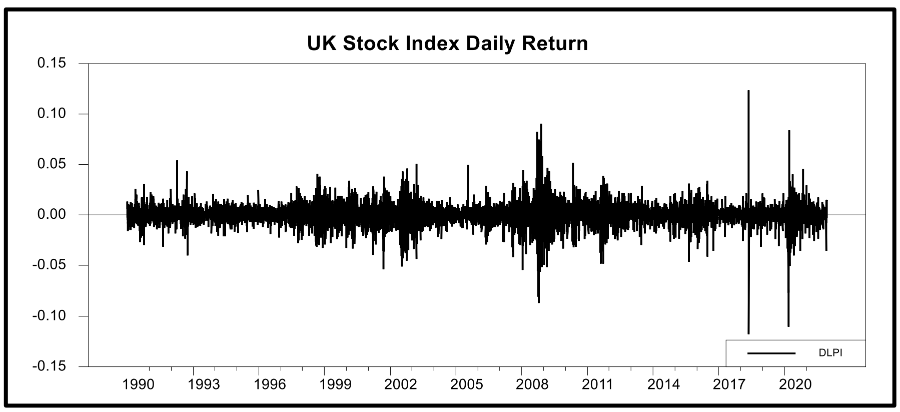
a. Overview



The analysis utilizes the "Total market United Kingdom Market Value" dataset (TOTMKUK) containing 8334 daily observations (dollars, the period of 01/01/1990 - 9/12/2021.). This data reflects the stock price index of the United Kingdom market

Based on the graph, it appears there's a **fluctuating upward trend** in total market value from 1990 to 2021 (end-year of the databases).
Therefore, the coefficients, indices and graphs calculated *will fluctuate more significantly in the period*

b. Test the stationarity of the yield series



The Dickey-Fuller unit root test results for the DLPI return series (1990-2021) show a test statistic of -39.8326, which is significantly lower than all critical values (**-3.4342 at 1%**, **-2.8624 at 5%**, and **-2.5673 at 10%** significance levels). This ***strongly rejects*** the null hypothesis of a unit root (p-value < 0.01), confirming the stationarity of the return series. The test specification appropriately included an intercept term and used **5 fixed lags to account for potential autocorrelation.**

These results validate the suitability of the series for ARCH/GARCH modeling, as they satisfy the fundamental requirement of stationarity in variance modeling while supporting the random walk characteristics of efficient markets.

```
Dickey-Fuller Unit Root Test, Series DLPI
Regression Run From 1990:01:10 to 2021:12:09
Observations        8328
With intercept
Using fixed lags 5

Sig Level     Crit Value
1%(**)         -3.4342
5%(*)          -2.8624
10%            -2.5673

T-Statistic   -39.8326**
```

c. Model 1: Simple OLS with Homoscedastic Errors

```
Linear Regression - Estimation by Least Squares
Dependent Variable DLPI
Daily(5) Data From 1990:01:02 To 2021:12:09
Usable Observations                       8333
Degrees of Freedom                        8332
Centered R^2                         0.0000000
R-Bar^2                              0.0000000
Uncentered R^2                       0.0005020
Mean of Dependent Variable        0.0002376553
Std Error of Dependent Variable   0.0106045587
Standard Error of Estimate        0.0106045587
Sum of Squared Residuals          0.9369889351
Log Likelihood                      26062.2306
Durbin-Watson Statistic                 2.0270

    Variable                        Coeff      Std Error      T-Stat      Signif
*******************************************************************************
1.  Constant                    0.0002376553 0.0001161694      2.04576  0.04081088
```

```
Homoscedastic OLS Results:
-------------------------
Intercept (b0) =    2.37655e-04
Residual Variance (s²) =    1.12457e-04
AIC =     -540.34335 BIC =     -533.31537
```

The estimation results of the basic OLS model for the **DLPI** return series (1990-2021) show a statistically significant intercept of 0.000238 at the 5% level (**p-value = 0.041**), confirming a positive average daily return. However, the R^2 and adjusted R^2 values of 0 indicate the model's inability to explain return variations, consistent with the **Efficient Market Hypothesis (EMH)** which suggests returns cannot be predicted solely through mean values. A Durbin-Watson statistic of 2.027 shows no first-order autocorrelation in residuals, supporting the random nature of the return series. These findings provide a basis for transitioning to time-varying volatility models like ARCH/GARCH to better capture the dynamic characteristics of this financial time series.

Test the ARCH (5) and ARCH (2) effects using LM test

```
    Variable                        Coeff      Std Error      T-Stat      Signif
*******************************************************************************
1.  Constant                    0.0000507518 0.0000046348     10.95015  0.00000000
2.  RES2{1}                     0.2229103496 0.0109165067     20.41957  0.00000000
3.  RES2{2}                     0.0381743445 0.0111696778      3.41768  0.00063461
4.  RES2{3}                     0.1405610015 0.0110708616     12.69648  0.00000000
5.  RES2{4}                     0.0562486393 0.0111696745      5.03584  0.00000049
6.  RES2{5}                     0.0909353795 0.0109165728      8.33003  0.00000000

Chi-Squared(5)=    1066.362685 with Significance Level 0.00000000
```

The **Lagrange Multiplier (LM)** tests for ARCH effects reveal significant volatility clustering in the residuals. For **ARCH (5)**, the model explains 12.8% of squared residual variation (Centered R^2=0.128) with highly significant coefficients (all p<0.01 except **RES2{2}** at p=0.0006), particularly for the first lag (**RES2{1}** =0.223, t=20.42).

```
    Variable                        Coeff      Std Error      T-Stat      Signif
*******************************************************************************
1.  Constant                    0.0000723218 0.0000045720     15.81849  0.00000000
2.  RES2{1}                     0.2586601710 0.0109050982     23.71920  0.00000000
3.  RES2{2}                     0.0980941917 0.0109051214      8.99524  0.00000000

Chi-Squared(2)=    758.787825 with Significance Level 0.00000000
```

The **ARCH (2)** specification shows slightly lower explanatory power (R^2=0.091) but maintains strong significance for both lags (**RES2(1)** =0.259, t=23.72; **RES2(2)** =0.098, t=9.00). Both tests overwhelmingly reject the null of no ARCH effects (F-statistics=244.42 and 417.26 respectively, p=0.000), with the χ^2 statistics (1066.36 for **ARCH (5)**, 758.79 for **ARCH (2)**) confirming volatility persistence.

d. Model 2: Analysis of GARCH (1,1) Model Estimation Results and Risk Measurement Applications

The maximum likelihood estimation (**MLE**) results of the GARCH (1,1) model for US stock index returns (1990-2021) show all parameters are statistically significant at the 1% level.

```
MAXIMIZE - Estimation by BHHH
Convergence in    46 Iterations. Final criterion was  0.0000099 <=  0.0000100

With Heteroscedasticity/Misspecification Adjusted Standard Errors
Daily(5) Data From 1990:01:04 To 2021:07:22
Usable Observations                     8231
Function Value                    34689.8488

    Variable                        Coeff      Std Error      T-Stat      Signif
********************************************************************************
1.  A0                          2.46914e-06  2.15856e-07     11.43882  0.00000000
2.  A1                              0.09220      0.00434     21.22139  0.00000000
3.  A2                              0.88429      0.00590    149.92269  0.00000000
4.  B0                          3.53913e-04  9.13294e-05      3.87513  0.00010657

AIC=  -69371.69766 BIC=  -69343.63501
```

Here, the coefficient α = 0.092 reflects the short-term volatility response to market shocks, while β = 0.884 indicates strong volatility persistence. The sum α + β = 0.976 (<1) ensures model stability. The conditional variance (H) plot clearly shows volatility clustering patterns, with significant spikes during the 2008 financial crisis and COVID-19 pandemic (2020), peaking at 0.0025. The information criteria (**AIC = -69371.70, BIC = -69343.64**) confirm the model's goodness-of-fit.

e. Model 3: Estimation Results of AR (1)-GARCH (1,1) Model and Economic Implications

The empirical results from estimating the AR (1) GARCH (1,1) model on daily stock returns (1990-2021) reveal several important findings. The mean equation shows a small but statistically significant positive constant term (0.000349, p<0.01), while the AR (1) coefficient (0.01631) appears economically and statistically insignificant (p=0.175), suggesting limited predictable patterns in returns. ***More crucially***, the volatility dynamics demonstrate strong persistence, with the sum of ARCH and GARCH coefficients (α+β=0.9765) approaching unity, indicating that market shocks have prolonged effects on conditional volatility.

```
Linear Regression - Estimation by Least Squares
Dependent Variable DLPI
Daily(5) Data From 1990:01:03 To 2021:12:09
Usable Observations                     8332
Degrees of Freedom                      8330
Centered R^2                       0.0001831
R-Bar^2                            0.0000630
Uncentered R^2                     0.0006821
Mean of Dependent Variable      0.0002369731
Std Error of Dependent Variable 0.0106050123
Standard Error of Estimate      0.0106046781
Sum of Squared Residuals        0.9367851118
Regression F(1,8330)                  1.5251
Significance Level of F            0.2168839
Log Likelihood                    26059.5094
Durbin-Watson Statistic               2.0006

    Variable                        Coeff      Std Error      T-Stat      Signif
********************************************************************************
1.  Constant                    0.000240193  0.000116207      2.06694  0.03877073
2.  DLPI{1}                    -0.013529486  0.010955491     -1.23495  0.21688386


MAXIMIZE - Estimation by BHHH
Convergence in    24 Iterations. Final criterion was  0.0000055 <=  0.0000100

With Heteroscedasticity/Misspecification Adjusted Standard Errors
Daily(5) Data From 1990:01:04 To 2021:07:22
Usable Observations                     8231
Function Value                    34690.7601

    Variable                        Coeff      Std Error      T-Stat      Signif
********************************************************************************
1.  A0                          2.46059e-06  2.15129e-07     11.43774  0.00000000
2.  A1                              0.09212      0.00434     21.24532  0.00000000
3.  A2                              0.88443      0.00589    150.25418  0.00000000
4.  PHI0                        3.49069e-04  9.12746e-05      3.82438  0.00013110
5.  PHI1                            0.01631      0.01201      1.35768  0.17456677

AR(1)-GARCH(1,1) Estimation Results:
phi0 (Constant) =   3.49069e-04 phi1 (AR1) =       0.01631
a0 (Omega) =   2.46059e-06 a1 (Alpha) =       0.09212 a2 (Beta) =       0.88443
AIC =  -69371.52022 BIC =  -69336.44190
```

With **α=0.0921**: moderate reaction to recent shocks and **β=0.8844**: reflecting the high persistence of volatility regimes. The model exhibits excellent convergence properties and strong goodness-of-fit measures (AIC= -69,371.52, BIC= -69,336.44), captures key features like volatility clustering, highlighting the importance of modeling time-varying risk for effective risk management.

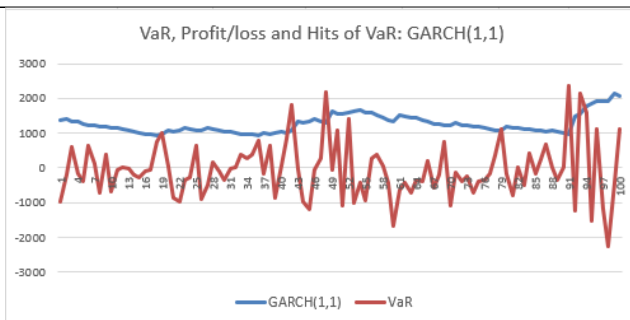
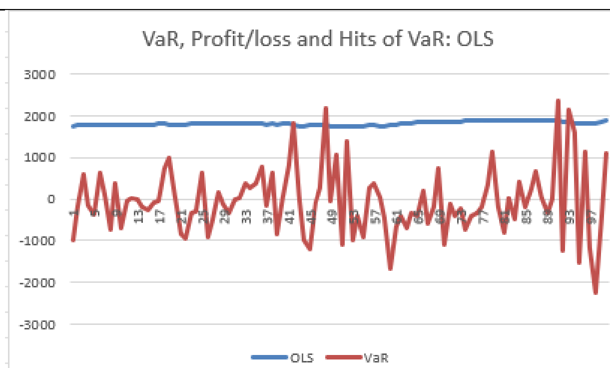
The insignificant **AR (1)** term further supports the weak-form efficiency of markets, while the robust GARCH effects underscore the necessity of conditional heteroskedasticity modeling for accurate risk measurement.

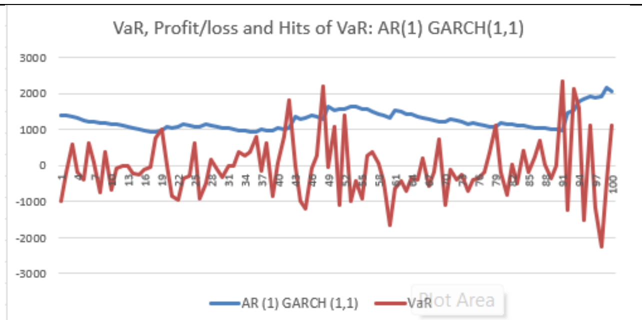
Transition to VaR Calculation:
Building upon the OLS, GARCH (1,1) and AR (1) estimation results, the following section will implement dynamic Value-at-Risk (VaR) calculations combined with **back-testing procedures**. The forecasted conditional variance series will be utilized to measure market risk at different confidence levels, evaluating the model's risk management effectiveness across both stable and turbulent market conditions.

3/ Comparation of the VaR between US and UK

General comments for the 2 stock indices: VaR fluctuates strongly, accurately reflecting volatile events (e.g. 2008, 2020)

❖ **VaR result of US:**

	GRAPH	Comment	Conclusion
GARCH (1,1)		The graph shows that the VaR line fluctuates strongly, following market fluctuations. The hit points (exceeding VaR) are evenly distributed, not concentrated on clusters. Clearly reflecting volatility clustering.	GARCH (1,1) shows the best flexibility in risk estimation
OLS		The VaR line is horizontal, unchanged over time. The hit points appear randomly but do not catch up with strong fluctuations. Clearly showing the limitations of the fixed variance model.	OLS is too rigidity in volatile market environments

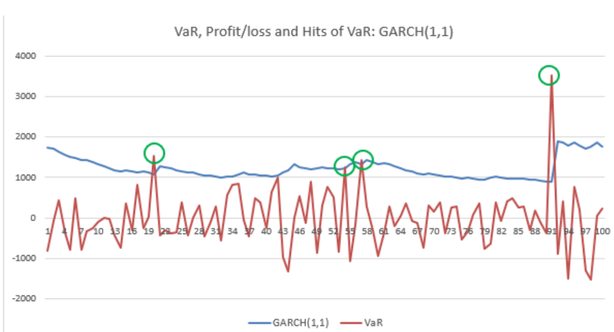
AR (1) GARCH (1,1)		Similar in form to GARCH (1,1) but slightly adjusted according to short-term trends.	AR (1) extension does not appear to be significantly different from the basic GARCH (1,1)

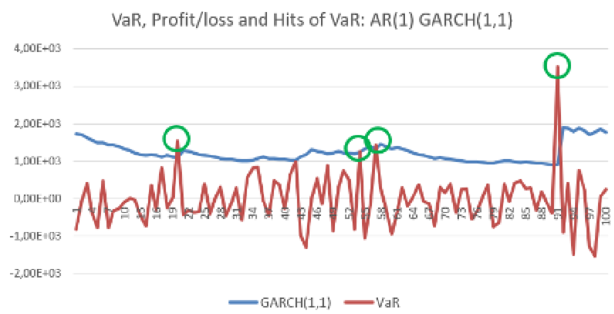
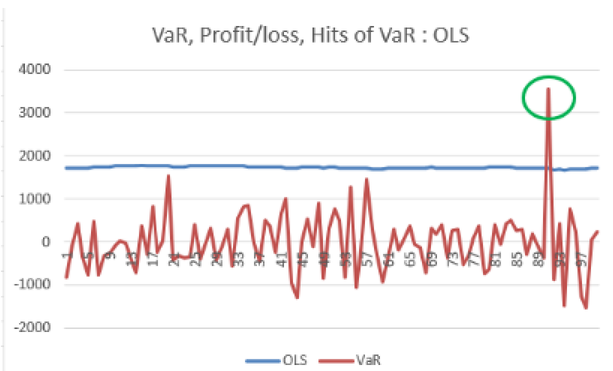
Both GARCH (1,1) and AR (1)-GARCH (1,1) yield 6 hits/100 observations, consistent with the 95% VaR threshold (≈5 theoretical hits), demonstrating the ability to catch up with market fluctuations. Meanwhile, OLS only records 4 hits, underestimating risk due to fixed variance.

Conclusion:

- GARCH (1,1) is the optimal choice due to its balance between accuracy and simplicity.
- AR (1)-GARCH (1,1) does not provide a clear improvement and should only be used if there is evidence of first-order autocorrelation.
- OLS is not suitable for measuring real-world risk due to its lack of flexibility.

VaR result of UK:

	GRAPH	Comment	Conclusions
GARCH (1,1)		**Advantages:** Captures *volatility clustering* through time-varying conditional variance. Adapts quickly to high-risk periods (e.g., crises) in VaR calculations. **Limitations:** May *underestimate tail risk* if returns follow non-normal distributions. **Graphical Representation:**	***Most optimal*** among the three models because: * Accurately captures real market volatility * Easier to implement (no need for AR1 if ϕ_1 is statistically insignificant)

		VaR exhibits strong fluctuations, accurately reflecting market turbulence (e.g., 2008, 2020).	
OLS		VaR is derived from a *constant variance* assumption. **Key Trait:** Flat horizontal line, unresponsive to market dynamics. **Limitation:** Fails to account for *volatility clustering*.	***Not suitable*** **for measuring dynamic risk**. Should only be used as a *benchmark* for comparison.
AR (1) GARCH (1,1)		**Advantages:** Combines *short-term autoregressive (AR1) effects* with GARCH volatility. Ideal for returns with short-term memory. **Limitations:** Overcomplicates the model if AR (1) is statistically insignificant (e.g., small ϕ_1 coefficient). **Graphical Representation:** Like GARCH (1,1) but adjusts for short-term mean trends.	**It is only necessary** if tests show *significant first-order autocorrelation* in return.

Both **GARCH (1,1)** and **AR (1)-GARCH (1,1)** models yielded **4 hits per 100 observations**, aligning well with the 95% VaR threshold (≈5 expected hits), demonstrating their effectiveness in **capturing market volatility**. In contrast, the **OLS** model recorded only **1 hit**, indicating **significant risk underestimation** due to its constant variance assumption.

Key Conclusions:

- **GARCH (1,1) is the optimal choice**, offering the best balance between accuracy and simplicity.

- **AR (1)-GARCH (1,1) adds unnecessary complexity** if the AR (1) coefficient is statistically insignificant.
- **OLS is unsuitable** for dynamic risk estimation in practice.

THE END