# Monash University (Clayton Campus)

# MTH 5510 - Quantitative Risk Management
# Semester 2

## Assignment 3

Lecturer: Hassan Fallahgoul



Group member: Ken SO
SID: 34856323
Coding/Computation method: Python (coding)/LaTeX (formula in this report)

Introduction

The assignment is about a replication of research – Time Series for financial market meltdowns (Young, et al., 2011). In this paper, Young et. al. pinpointed that the classical time series models, particularly those assuming a normal distribution, have significant limitations when analyzing and forecasting financial market meltdowns such as inability to capture real-world data patterns, underestimation of extreme events and unreliable forecast of extreme events. To substantiate their argument, they conducted a series of statistical analysis, goodness-of-fit test, VaR backtesting, Average of relative difference ("ARD") and ended with figures that strongly supported their arguments. Note that the paper is written as of 2011 while approximately 30 years data were used in their research. In this assignment, I expect to replicate the methods used in the research and demonstrate that those classical time series models were still flaw under normal distribution assumption when 10 years passed.

Replication methodology

As requested, the data applied in this assignment covered 20 years daily price of S&P 500 from 10 Jan 2000 to 9 Jan 2020. The Data provider is Market Watch. Below is the list of models to be replicated:

1. CV where innovations are normal and Student-t
2. GARCH where innovations are normal and Student-t
3. ARMA-GARCH where innovations are normal and Student-t

According to Young, et. al, models for test follow the framework below:

$$\begin{cases} y_t = ay_{t-1} + b\sigma_{t-1}\epsilon_{t-1} + \sigma_t\epsilon_t + c, \\ \sigma^2 = \alpha_0 + \alpha_1\sigma_{t-1}^2\epsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2, \end{cases} \qquad (1)$$

where $\epsilon_t$ is the innovation and is i.i.d. real random variable with $\epsilon_0 = 0$. For CV model, $a$, $b$, $\alpha_1$, $\beta_1$ are zeros. For GARCH, $a$, $b$ are zero. For ARMA-GARCH, (1) is its general form. The change in distribution assumption works with the innovation $\epsilon_t$ that

$$\begin{cases} \epsilon_t \sim N(0, \sigma_\epsilon^2) \\ \epsilon_t \sim t(v, 0, \sigma_\epsilon^2), \end{cases} \qquad \text{where } v \text{ is degree of freedom}$$

Here, I note that Young et. al. applied Maximum Likelihood Estimation ("MLE") in their parameterization while applied simply MLE cannot achieve similar result as in Table 3 in their paper. I will elaborate in detail in model calibration section. Below is the framework of this assignment:

1. Model calibration
2. Goodness-of-fit test
3. Historical Crash Probability analysis
4. Dynamic forecasting and VaR backtesting
5. Economic Significance Analysis

In each part, I will discuss the assumptions for the models, data applied, methodologies and analysis. The algorithm will be discussed in another attachment including my plan to develop the algorithm.

*Disclaimer: The replication system is developed with AI including code generation.*

User guidance of my replication system:

1. "Open folder" in VSC, input "3" (custom range)
2. Input the rolling window size (=10)
3. Input start date of calibration period (1998-10-01)
4. Input end date of calibration period (2008-09-26)
5. Input "1". The replication system takes 15-20 minutes to run for the parameterisation
6. Enter the start date and end date for testing period under VaR backtesting (input the list of dates in part 4)

Model Calibration

In the paper, I notice that table 3 provides a relatively achievable standard for calibration. In table 3, they provide a parameterisation result with Kolmogorov-Smirnov ("KS") test, Anderson-Darling ("AD") statistic and Average time to market crash ("Average time") with data coverage – 2505 observations, ending with the 26 September 2008 (just before the date of US Financial Crisis – 28 September 2008). Yet, during the calibration, I notice that Young et. al. may not use the same technique as I in 2011. My system is built on Python language while Python is not yet a very common language in the days in 2011. Hence, discrepancies in figure between my replication model and table 3 appear. My aim is to set my replication system as close as possible to the figure in table 3 with the technology nowadays.

To ensure my replication system generates result as close as table 3 in the paper, I researched on the methodology on parameterisation in the practice of Financial Econometrician and note that 2-stages strategy is the common practice on Time series modelling. The GARCH(1,1) model with student-t distributions has the log-likelihood that creates a high-dimensional non-convex optimization problem with multiple local optima. When degree of freedom $v \to \infty$, student-t tends to normal and $\alpha, \beta$ should approach normal GARCH(1,1) estimates. **Figure 1** shows that the log-likehood surface has multiple peaks and valleys. Hence, standard gradient-based methods can get stuck in local optima. By 2 stages strategy, we can start with fitting normal GARCH(1,1) model as the representative of $v \to \infty$ case. Then, apply Grid Search optimization algorithm to search for a range of $v$. In my replication system, $v \in \{5,6,7,8,9,9.5,9.8,10\}$ and it initializes the parameterisation of model under student-t with adjusted normal estimates. Once optimized $v$ is located, my system starts to optimize $\{\alpha_0, \alpha_1, \beta_1, c\}$ by standard gradient-based method. **Process 1** displays the process of optimization of parameters.

For capturing the stylized fact in the market, I also introduced the near integrated persistence condition when parameterising GARCH(1,1) and ARMA-GARCH(1,1) models with innovation under student-t. When $\alpha_1 + \beta_1 = 1$, mathematically, any shock to volatility is permanent and the variance does not revert to a long-run mean, which is often considered unrealistic for index return in long run. Hence, near integrated persistence set $\alpha_1 + \beta_1 = 1 - \frac{c}{T}$, where $T$ is the sample size and $c$ is a small, non-zero constant, can assist the model calibration by parameterising models more realistically. **Table 1** in this assignment demonstrated a fair and reasonable result in the parameterisation that closely aligned with table 3 in the paper with 2506 observations from 1 October 1998 to 26 September 2008 (note that data before 10 January 2000 is not required in the assignment but I prepared in extra for calibration purpose). Error between **table 1** and table 3 in the paper is believed to be the discrepancy of data provider, statistical tools and data coverage (where I note that a single extra data point is not sufficient to invalidate my replication).

Goodness of fit test

Pursuant to Young et. at., the initiative for this part is to test the goodness-of-fit of normal and student-t distribution assumptions to the residual with empirical data. First of all, introduce the goodness-of-fit tests adopted in Young et. al. research below:

1. Kolmogorov-Smirnov ("KS") test
2. Anderson-Darling ("AD") statistic

These 2 tests share same set of hypotheses:

- $H_0$: The sample of data is drawn from the specified theoretical probability
- $H_a$: Else

KS test provides an overview of the goodness-of-fit of entire distribution to the empirical data while AD test is more sensitive to the goodness-of-fit of tails of distribution which gives more weight to deviations at the extreme end of the distribution that reflected market crash and contingent events. Therefore, KS test is as if the baseline of the goodness-of-fit while AD test works specifically on the failure of KS test and thus constructing a robust test of goodness-of-fit. In the paper, Young et. al. note that those 3 models with innovation follows normal and student-t are all failed in both KS and AD test. In **table 1**, I observe that none of the models with innovation follows normal and student-t distribution pass KS and AD test. This part of result aligned with the paper.

Historical Crash Probability Analysis

In the paper, Young et. al. conducted a stress testing on historical crisis such as Black Monday on 19 October 1987 and US Financial Crisis on 29 September 2008 to analyse the probability and average time of recurrence of the same crisis ("average time"). They note that models with innovation under normal distribution results in astronomical figure in average time, which is considered as unrealistic and "no serious warnings about a market crash". Technically, they take 1-step-ahead estimate to estimate crisis probability and average time of crisis recurrence. The formula to estimate average time of crisis recurrence is attached in appendix.

Based on the result from my replication system **table 2**, I note that the probability and average time is reasonably and fairly replicated. The rank of average time is in the correct order that CV > ARMA-GARCH(1,1) > GARCH(1,1) with innovation follows normal and student-t. The value of average time for models with innovation follows normal distribution in **table 2** are considered as unrealistic and not sufficient to alert the risk manager on crisis recurrence. The errors between table 3 in the paper and **table 2** in this assignment is believed to be discrepancies in statistical tools, source data and practice but the impact is immaterial to the conclusion. This part of result aligned with the paper.

Dynamic forecasting and VaR backtesting

This part of analysis in the paper is to evaluate accuracy of one-day-ahead 1% VaR forcasting during both calm and volatile periods with an aim to verify the explanatory power of models with innovation follow either normal or student-t assumption. Young et. al. adopt 2 hypotheses test in this part.

1. Christoffersen Likelihood Ratio ("CLR")
2. Berkowitz Likelihood Ratio ("BLR")

CLR test is used to evaluate the accuracy of VaR forecasts by analyzing the history of VaR violations. The entire test is built upon the foundational concept of a "violation". BLR test assesses the accuracy of the predicted tail distribution holistically where CLR only test by counting the number of violations.

Under CLR, Young et. al. adopted 3 variations of ratio:

1. $CLR_{uc}$: CLR with unconditional coverage
2. $CLR_{ind}$: CLR with independent VaR violation
3. $CLR_{cc}$: Joint test of 1,2

Under BLR, they adopted 2 variations of ratio:

1. $BLR_{ind}$: BLR with test of independence
2. $BLR_{tail}$: BLR with test of tail distribution

In other words, the CLR is a test for "routine competency on capturing extreme events" and test and BLR is a test for "competency to capture extreme events". As such, the time coverage should combine with both calm and volatile periods.

Young et. al. proposed that models with innovation follow normal distribution is a fair-weather model that it only works when the risk is low while non-normal assumption on innovation provides a more robust and reliable risk assessments in both calm and volatile markets. Note that they performed testing for 1 year over short calm period (14 December 2004 – 15 December 2005), 2 years over longer calm period (14 December 2004 – 20 December 2006) and 4 years over volatile period (14 December 2004 – 31 December 2008) with 10 years rolling windows on approximately 30 years empirical data with different variations in ARMA-GARCH.

In this assignment, I followed Young et. al. and applied the rolling windows with 10 years size. The calibration period will start from 10 January 2000 to 9 January 2010 and the rest is testing period. Note that, even if in 2018, there is no crisis with similar scale as 2008 US Financial Crisis. It is difficult to perfectly replicate the result of the table 4 and 5 in the paper. By fact check, during 2018, there is a recorded loss with around −6% in return series. Therefore, I picked the data ranges below:

1. 10 January 2016 – 9 January 2017 (1 year) (Calm period)
2. 10 January 2016 – 9 January 2018 (2 years) (Longer calm period)
3. 10 January 2016 – 9 January 2020 (4 years) (Covered both calm and volatile period)

with 1% significant level for all hypothesis tests in my backtesting.

By inputting the range **#1-3** into my system, I obtained the results in **Table 3.1 – 3.3**. In 1 year calm period test, all models perform well, excepts CV normal model which failed the CLRuc test. This failure of test suggests that CV normal is structurally inaccurate and probably overestimate or underestimate risk. While comparing number of violations between normal and student-t models, the results suggests all models are generally acceptable for short-term risk measurement and model with innovation follows student-t distribution has a more sensitive measure than normal distribution.

In 2 years test, all models failed the unconditional coverage tests, which is likely due to a large number of violations influx. It is also notably that only GARCH t model passed the conditional coverage test which support that GARCH t model has better ability to capture the violation in either frequency and timing.

In 4 years test, CV models with flaw structure and ability to capturing violation, accumulate fewer number of violations than other models, leading to strong fail results on the coverage tests. Note that the GARCH and ARMA-GARCH normal model slightly pass the CLRuc and CLRcc test with p-values 0.0022 (=0.0122-0.01) and 0.0039 (=0.0139-0.01) slightly higher than significance level. This result ties to the conclusion in the paper that models with innovation follow normal distribution is not realistic to capture violation and provide sufficient alert to risk manager in long run. Moreover, I also note that, even if student-t distribution (or say, non-normal) generally performs better than normal distribution assumption, it may become flaw in long run. A more sophisticated distribution assumption and model structure is required in order to better capture the recurring crisis in long run.

This part of result substantially aligned with the paper.

Economic Significance Analysis

Intuitively, ARD measures the cost of risk management. For example, ARD of 30% means a risk manager using the student-t model requires to hold 30% more capital charge than a risk manager using normal model. It can be comprehended as the indicator of trade-off between capital efficiency and safety as per Young et. al., who proposed in the ARD analysis in their paper, and ultimately, concluded that non-normal model offers the best balance. To verify their finding, I follow the data coverage in VaR backtesting that covered the date range below

1. 10 January 2016 – 9 January 2017 (1 year)
2. 10 January 2016 – 9 January 2018 (2 years)
3. 10 January 2016 – 9 January 2020 (4 years)

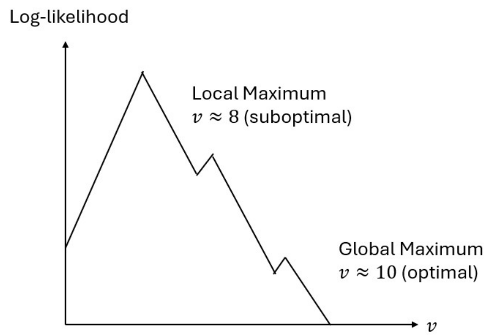
to replicate and perform the ARD analysis.

In 2016, the models showed the highest ARD at 7.78%, indicating the largest static disagreement between the Student-t and Normal distribution assumption, while the GARCH (3.13%) and ARMA-GARCH(2.81%) models showed the lowest initial ARD. In 2017-2018, the GARCH and ARMA-GARCH drastically to 8.53% and 9.04%, reflecting that during periods of rising market volatility (e.g., late 2017-2018), the GARCH and ARMA-GARCH models demand significantly more capital to achieve risk robustness. In 2019-2020, the ARD of GARCH model fell sharply to 3.57% before rising again to 7.32%, while ARMA-GARCH showed a similar volatile pattern but ultimately settled lower in 6.48%. In conclusion, the ARD for CV models is always high and stable, but the ARD for GARCH and ARMA-GARCH models is pro-cyclical, acting as a volatility signal that confirms the models under student-t (a.k.a. non-normal) distribution's primary role is to ensure risk robustness by demanding the most capital precisely when dynamic volatility is at its highest point of stress. This part of results substantially aligned with the paper.

# Appendix

## 1. Model Calibration – Performance analysis KS and AD test

Figure 1 – Likelihood surface




Process 1 - Model fitting

```
======================================================================
FITTING ALL MODELS - GLOBAL OPTIMIZATION FOR STUDENT'S T
======================================================================

Fitting CV Student-t (dedicated implementation)...
Initial: α₀=1.336143e-04, c=8.252527e-05, v=6.83
Final: α₀=1.418999e-04, c=2.048516e-04, v=4.27
Log-likelihood: 7722.3198

Fitting GARCH Normal...
Final: α₀=8.294854e-07, α₁=0.0625, β₁=0.9331, c=3.228326e-04
Log-likelihood: 7914.9637

Fitting GARCH Student-t (theory-driven initialization)...

Fitting GARCH Normal...
Final: α₀=8.294854e-07, α₁=0.0625, β₁=0.9331, c=3.228326e-04
Log-likelihood: 7914.9637
Using Normal estimates as starting point:
  α₀=8.294854e-07, α₁=0.0625, β₁=0.9331
  v≈5.0: LL=7934.6544, α₁+β₁=0.9976 (improved)
  v≈6.0: LL=7938.5276, α₁+β₁=0.9976 (improved)
  v≈7.0: LL=7940.4219, α₁+β₁=0.9976 (improved)
  v≈8.0: LL=7941.2626, α₁+β₁=0.9976 (improved)
  v≈9.0: LL=7941.5099, α₁+β₁=0.9976 (improved)
  v≈9.5: LL=7941.5129, α₁+β₁=0.9976 (improved)
  v≈9.8: LL=7941.5129, α₁+β₁=0.9976 (improved)
Final: α₀=5.589993e-07, α₁=0.0619, β₁=0.9357, c=4.130632e-04, v=10.14
Persistence: α₁+β₁=0.9976
Log-likelihood: 7941.5129
```

```
Fitting ARMA-GARCH Normal...
Final: a=0.4731, b=-0.5000, c=1.724295e-04, α₀=9.520030e-07, α₁=0.0650, β₁=0.9294
Log-likelihood: 7918.3235

Fitting ARMA-GARCH Student-t (theory-driven initialization)...

Fitting ARMA-GARCH Normal...
Final: a=0.4731, b=-0.5000, c=1.724295e-04, α₀=9.520030e-07, α₁=0.0650, β₁=0.9294
Log-likelihood: 7918.3235
Using Normal estimates as starting point:
  a=0.4731, b=-0.5000, c=1.724295e-04
  α₀=9.520030e-07, α₁=0.0650, β₁=0.9294
  v≈5.0: LL=7943.5673, α₁+β₁=0.9977 (improved)
  v≈6.0: LL=7947.1573, α₁+β₁=0.9978 (improved)
  v≈7.0: LL=7948.8526, α₁+β₁=0.9978 (improved)
  v≈8.0: LL=7949.5450, α₁+β₁=0.9978 (improved)
  v≈9.0: LL=7949.6820, α₁+β₁=0.9977 (improved)
  v≈9.5: LL=7949.6820, α₁+β₁=0.9977 (improved)
  v≈10.0: LL=7949.6820, α₁+β₁=0.9977 (improved)
Final: a=0.6999, b=-0.7582, c=1.294643e-04
       α₀=5.412438e-07, α₁=0.0614, β₁=0.9364, v=9.82
Persistence: α₁+β₁=0.9977
Log-likelihood: 7949.6820
Fitted: CV_normal
Fitted: CV_t
Fitted: GARCH_normal
Fitted: GARCH_t
Fitted: ARMA_GARCH_normal
Fitted: ARMA_GARCH_t
```

Table 1 – Parameterisation on 1 October 1998 – 26 September 2008 (size = 2506)

RESULTS TABLE:

| Model | $\alpha_0$ | $\alpha_1$ | $\beta_1$ | a | b | c | v | ε* | KS_stat | KS_pval | KS_Test | AD_stat | AD_Test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CV normal | 1.336143e-04 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 8.252527e-05 | - | -8.6610 | 0.052279 | 0.000002 | Fail | 15.340956 | Fail |
| CV t | 1.418999e-04 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 2.048516e-04 | 4.27 | -8.4146 | 0.078526 | 0.000000 | Fail | 3.343599 | Fail |
| GARCH normal | 8.294854e-07 | 0.062520 | 0.933144 | 0.000000 | 0.000000 | 3.228326e-04 | - | -4.2677 | 0.034837 | 0.004454 | Fail | 4.050898 | Fail |
| GARCH t | 5.589993e-07 | 0.061863 | 0.935703 | 0.000000 | 0.000000 | 4.130632e-04 | 10.14 | -4.2534 | 0.042566 | 0.000221 | Fail | 1.760605 | Fail |
| ARMA GARCH normal | 7.906301e-07 | 0.061564 | 0.934448 | 0.760057 | -0.808929 | 7.887884e-05 | - | -4.3557 | 0.037357 | 0.001786 | Fail | 5.203239 | Fail |
| ARMA GARCH t | 5.412518e-07 | 0.061383 | 0.936351 | 0.699928 | -0.758182 | 1.294660e-04 | 9.82 | -4.3463 | 0.047989 | 0.000019 | Fail | 2.769904 | Fail |

## 2. Crash probability analysis

Table 2 – Crash probability before US financial Crisis (from 1 October 1998 – 26 September 2008)

SUMMARY: CRASH PROBABILITY AND AVERAGE TIME

| Model | ε* | Probability | Days | Years |
|---|---|---|---|---|
| CV normal | -8.66 | 2.34e-18 | 4.28e+17 | 1.71e+15 |
| CV t | -8.41 | 4.08e-04 | 2452.4 | 9.81e+00 |
| GARCH normal | -4.27 | 9.88e-06 | 101244.6 | 4.05e+02 |
| GARCH t | -4.25 | 8.15e-04 | 1226.7 | 4.91e+00 |
| ARMA GARCH normal | -4.36 | 6.63e-06 | 150777.8 | 6.03e+02 |
| ARMA GARCH t | -4.35 | 7.57e-04 | 1320.7 | 5.28e+00 |

## 3. VaR Backtesting

Table 3.1 – Backtesting results table from 10 Jan 2016 to 09 Jan 2017 (1 year)

BACKTESTING RESULTS TABLE (1% SIGNIFICANCE LEVEL)

| Model | Violations | CLRuc | CLRuc p-val | CLRuc Test | CLRind | CLRind p-val | CLRind Test | CLRcc | CLRcc p-val | CLRcc Test | BLRind | BLRind p-val | BLRind Test | BLR_tail | BLR_tail p-val | BLR_tail Test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CV Normal & CV Student-t | 4 (1.59%) | 8.326 | 0.0039 | FAIL | 0.000 | 1.0000 | PASS | 8.326 | 0.0156 | PASS | 0.066 | 0.7971 | PASS | 0.000 | 0.9952 | PASS |
| (ARMA-)GARCH Normal | 5 (1.98%) | 6.196 | 0.0128 | PASS | 0.000 | 1.0000 | PASS | 6.196 | 0.0451 | PASS | 0.104 | 0.7470 | PASS | 0.002 | 0.9641 | PASS |
| (ARMA-)GARCH Student-t | 6 (2.38%) | 4.477 | 0.0344 | PASS | 2.438 | 0.1185 | PASS | 6.915 | 0.0315 | PASS | 5.391 | 0.0202 | PASS | 0.089 | 0.7651 | PASS |

Critical values at 1% significance: $\chi^2(1) = 6.635$, $\chi^2(2) = 9.210$

Table 3.2 – Backtesting results table from 10 Jan 2016 to 09 Jan 2018 (2 years)

BACKTESTING RESULTS TABLE (1% SIGNIFICANCE LEVEL)

| Model | Violations | CLRuc | CLRuc p-val | CLRuc Test | CLRind | CLRind p-val | CLRind Test | CLRcc | CLRcc p-val | CLRcc Test | BLRind | BLRind p-val | BLRind Test | BLR_tail | BLR_tail p-val | BLR_tail Test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CV Normal & CV Student-t | 4 (0.79%) | 28.601 | 0.0000 | FAIL | 0.000 | 1.0000 | PASS | 28.601 | 0.0000 | FAIL | 0.032 | 0.8572 | PASS | 0.000 | 0.9952 | PASS |
| (ARMA-)GARCH Normal | 10 (1.98%) | 12.392 | 0.0004 | FAIL | 0.000 | 1.0000 | PASS | 12.392 | 0.0020 | FAIL | 0.207 | 0.6488 | PASS | 0.000 | 0.9985 | PASS |
| GARCH Student-t | 13 (2.58%) | 7.499 | 0.0062 | FAIL | 0.925 | 0.3361 | PASS | 8.424 | 0.0148 | PASS | 1.386 | 0.2391 | PASS | 0.171 | 0.6796 | PASS |
| ARMA-GARCH Student-t | 12 (2.38%) | 8.954 | 0.0028 | FAIL | 1.164 | 0.2806 | PASS | 10.118 | 0.0064 | FAIL | 1.871 | 0.1714 | PASS | 0.123 | 0.7256 | PASS |

Critical values at 1% significance: $\chi^2(1) = 6.635$, $\chi^2(2) = 9.210$

Table 3.3 – Backtesting results table from 10 Jan 2016 to 09 Jan 2020 (4 years)

BACKTESTING RESULTS TABLE (1% SIGNIFICANCE LEVEL)

| Model | Violations | CLRuc | CLRuc p-val | CLRuc Test | CLRind | CLRind p-val | CLRind Test | CLRcc | CLRcc p-val | CLRcc Test | BLRind | BLRind p-val | BLRind Test | BLR_tail | BLR_tail p-val | BLR_tail Test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CV Normal | 31 (3.08%) | 9.018 | 0.0027 | FAIL | 9.621 | 0.0019 | FAIL | 18.639 | 0.0001 | FAIL | 18.250 | 0.0000 | FAIL | 0.198 | 0.6561 | PASS |
| CV Student-t | 34 (3.38%) | 6.278 | 0.0122 | PASS | 15.852 | 0.0001 | FAIL | 22.130 | 0.0000 | FAIL | 31.943 | 0.0000 | FAIL | 0.825 | 0.3638 | PASS |
| (ARMA-)GARCH Normal | 34 (3.38%) | 6.278 | 0.0122 | PASS | 2.272 | 0.1317 | PASS | 8.551 | 0.0139 | PASS | 3.197 | 0.0738 | PASS | 0.095 | 0.7585 | PASS |
| GARCH Student-t | 41 (4.07%) | 1.946 | 0.1630 | PASS | 2.612 | 0.1060 | PASS | 4.558 | 0.1024 | PASS | 3.531 | 0.0602 | PASS | 0.066 | 0.7974 | PASS |
| ARMA-GARCH Student-t | 37 (3.67%) | 4.087 | 0.0432 | PASS | 3.756 | 0.0526 | PASS | 7.843 | 0.0198 | PASS | 5.522 | 0.0188 | PASS | 0.000 | 0.9831 | PASS |

Critical values at 1% significance: $\chi^2(1) = 6.635$, $\chi^2(2) = 9.210$

Figure 2.1 – VaR Backtesting from 10 Jan 2016 to 09 Jan 2017 (1 year)

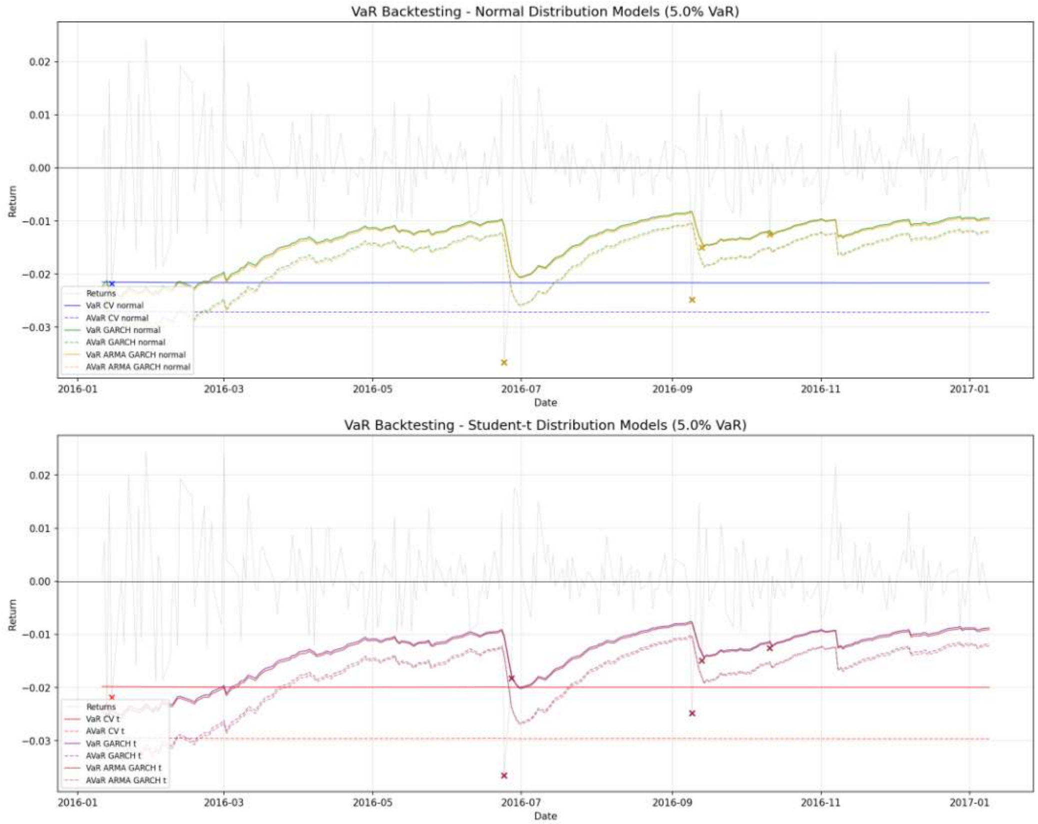



Figure 2.2 – VaR Backtesting from 10 Jan 2016 to 09 Jan 2018 (2 years)

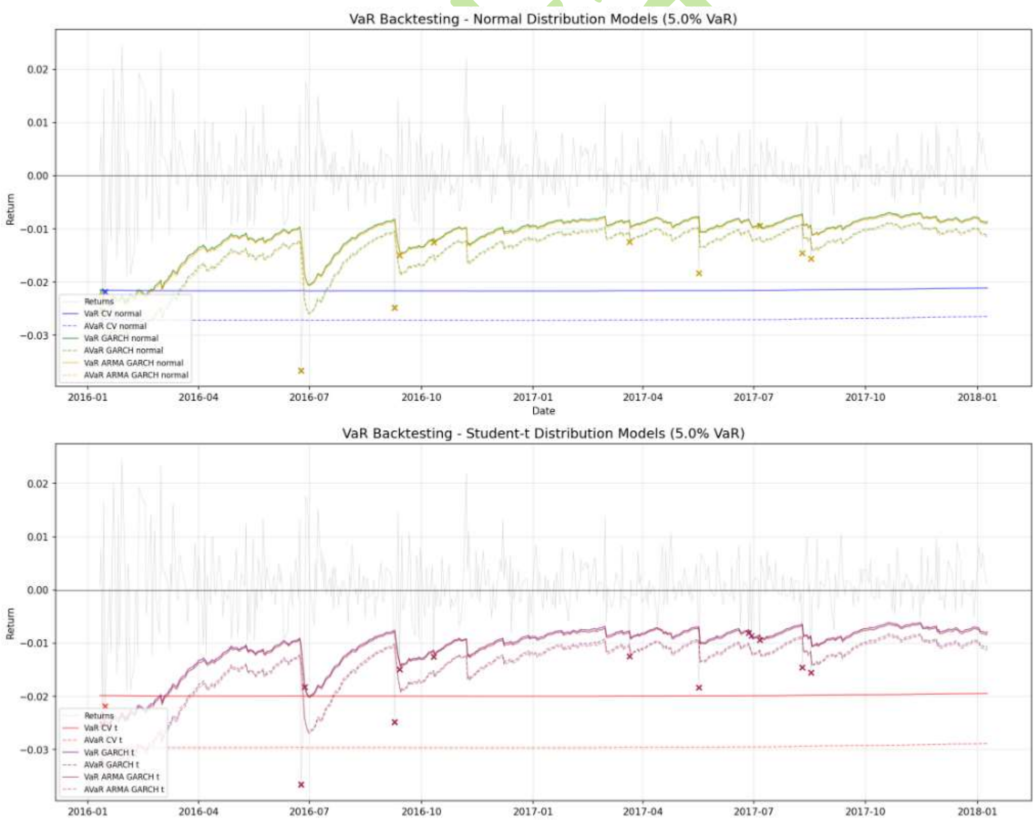
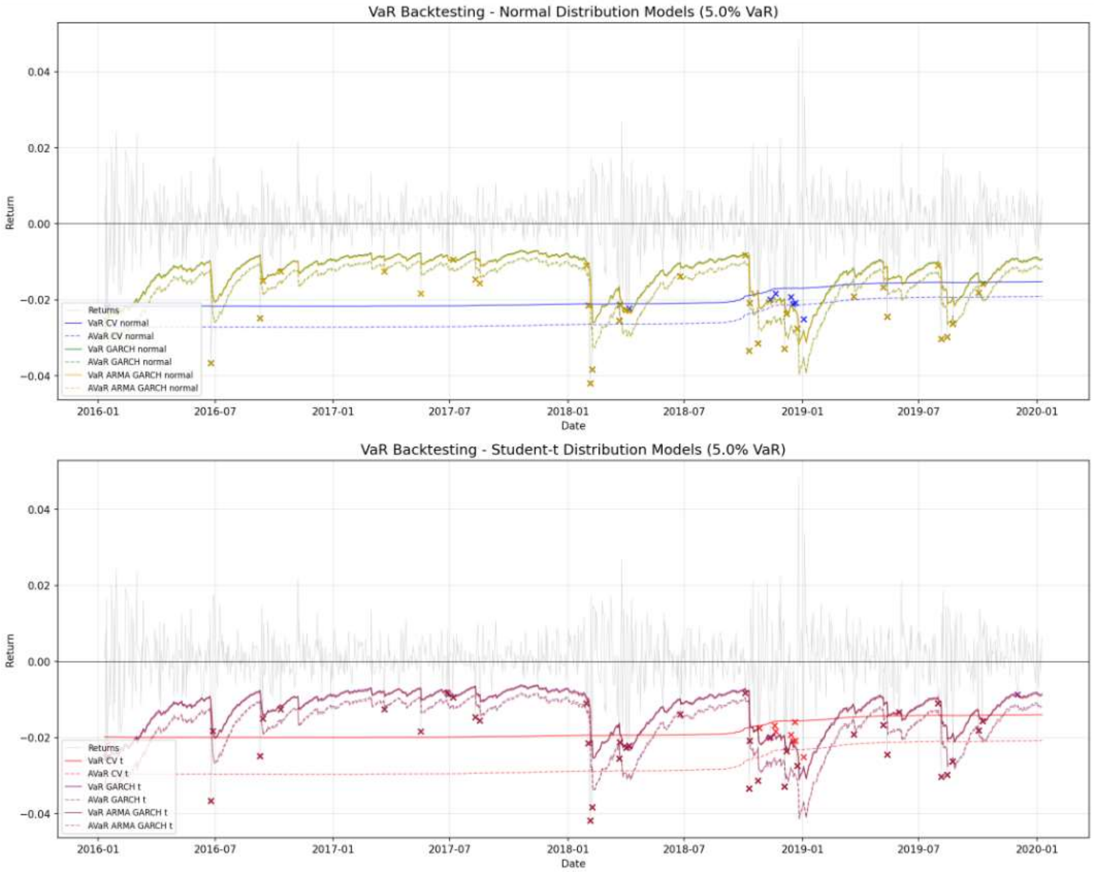


Figure 2.3 – VaR Backtesting from 10 Jan 2016 to 09 Jan 2020 (4 years)




## 4. Economic Significance analysis

Table 4.1 – Average Relative Difference between Student-t and Normal VaR (2 years)

```
======================================================================
ECONOMIC SIGNIFICANCE ANALYSIS
======================================================================

ARD (%) - Average Relative Difference between Student-t and Normal VaR
--------------------------------------------------------------------------------
            Model 2016 2017
        CV normal 7.78 7.78
             CV t 7.78 7.78
     GARCH normal 3.13 6.60
          GARCH t 3.13 6.60
ARMA GARCH normal 2.81 5.96
     ARMA GARCH t 2.81 5.96

--------------------------------------------------------------------------------
AVaR (Average VaR / Expected Shortfall) by Year
--------------------------------------------------------------------------------
            Model    2016    2017
        CV normal -0.0271 -0.0272
             CV t -0.0296 -0.0297
     GARCH normal -0.0176 -0.0121
          GARCH t -0.0182 -0.0120
ARMA GARCH normal -0.0179 -0.0124
     ARMA GARCH t -0.0186 -0.0124

Economic significance results saved to: economic_significance_results.csv
```

Table 4.2 – Average Relative Difference between Student-t and Normal VaR (3 years)

```
ARD (%) - Average Relative Difference between Student-t and Normal VaR
--------------------------------------------------------------------------------
           Model 2016 2017 2018
       CV normal 7.78 7.79 7.80
            CV t 7.78 7.79 7.80
    GARCH normal 3.13 8.53 9.04
         GARCH t 3.13 8.53 9.04
ARMA GARCH normal 2.81 7.68 7.95
    ARMA GARCH t 2.81 7.68 7.95

--------------------------------------------------------------------------------
AVaR (Average VaR / Expected Shortfall) by Year
--------------------------------------------------------------------------------
           Model    2016    2017    2018
       CV normal -0.0271 -0.0270 -0.0265
            CV t -0.0296 -0.0294 -0.0289
    GARCH normal -0.0176 -0.0109 -0.0107
         GARCH t -0.0182 -0.0106 -0.0104
ARMA GARCH normal -0.0179 -0.0111 -0.0110
    ARMA GARCH t -0.0186 -0.0109 -0.0108

Economic significance results saved to: economic_significance_results.csv
```

Table 4.3 – Average Relative Difference between Student-t and Normal VaR (4 years)

```
,Model,2016,2017,2018,2019,2020
ARD Results (%),,,,,,
,CV normal,7.78,7.79,7.83,8.01,8.03
,CV t,7.78,7.79,7.83,8.01,8.03
,GARCH normal,3.13,8.53,4.10,3.57,7.32
,GARCH t,3.13,8.53,4.10,3.57,7.32
,ARMA GARCH normal,2.81,7.68,3.73,3.29,6.48
,ARMA GARCH t,2.81,7.68,3.73,3.29,6.48
,,,,,,
AVaR Results,,,,,,
,CV normal,-0.0271,-0.0270,-0.0252,-0.0197,-0.0191
,CV t,-0.0296,-0.0294,-0.0275,-0.0214,-0.0208
,GARCH normal,-0.0176,-0.0109,-0.0194,-0.0181,-0.0116
,GARCH t,-0.0182,-0.0106,-0.0199,-0.0186,-0.0115
,ARMA GARCH normal,-0.0179,-0.0111,-0.0196,-0.0184,-0.0119
,ARMA GARCH t,-0.0186,-0.0109,-0.0202,-0.0189,-0.0119
```

**Formulae**

Average Time of crisis recurrence (“Average time”):

$$Average\ \ time = \frac{1}{250P(\epsilon_t \leq \epsilon_t^*)}$$

Average Relative Difference (“ARD”):

$$ARD = \boldsymbol{E}\left(\frac{\boldsymbol{VaR_{stud-t}(y_{t+1}) - VaR_{normal}(y_{t+1})}}{\boldsymbol{VaR_{normal}(y_{t+1})}}\right)$$

Average Value at Risk (“AVaR”):

$$AVaR = -E[X|X < -VaR_{\alpha}(X)]$$