

Financial Econometrics I: Course Project

Matias Palmunen

January 30, 2026

This report summarizes results for the Financial Econometrics I course project. The project studies the characteristics of financial time series data focusing on typical characteristics such as volatility, heavy tails, and skewness and general non-normality of returns. The analyses were conducted using Python and relevant econometric libraries are listed in the appendix.

Rest of this report is organized as follows. First, section 1 focuses on the exploratory data analysis of the returns data. Section 2 focuses on AR(1) modeling with GARCH(1,1) volatility. And finally, the third section concludes.

[Short data description here.]

1 Characteristics of Financial Time Series

In this section I will use the stock (S&PCOMP(RI)), commodity (RJEFCRT(TR)), and government bond (SPUTBIX(RI)) return data to analyze typical characteristics of financial time series on daily and weekly frequencies.

1.1 Exploratory analyses on index returns

In this section the analyses is done using daily and weekly log returns of the selected indices. The log returns are computed as

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t) - \ln(P_{t-1}). \quad (1)$$

As we are mostly focused on weekly and daily frequencies, the daily returns are computed using consecutive trading days, while the weekly returns are computed using Sunday-to-Sunday periods.

1.1.1 Main Crashes and Booms on Index Returns

Table 1.1.1 presents the five largest positive returns (booms) and five largest negative returns (crashes) for the S&P 500 stock index. Returns are shown at both daily and weekly frequencies. Notable periods of extreme volatility include the 2008 Global Financial Crisis, the 2020 COVID-19 pandemic, and various other major economic events. Daily returns are calculated as log returns of consecutive trading days, while weekly returns aggregate returns over Sunday-to-Sunday periods.

The table shows that during a financial market distress the moves are typically very extreme. The next section studies if these extreme moves can be explained if we assume that the data generating process is normal.

Table 1: Top 5 Booms and Crashes for S&P 500, Government Bonds, and Commodities (Daily and Weekly)

Daily Returns		
Date	Return (%)	Explanation
Top 5 Crashes		
2020-03-16	-12.76	COVID-19 lockdowns and economic shutdown fears
2020-03-12	-9.97	Pandemic uncertainty and market panic
2008-10-15	-9.46	Lehman Brothers collapse aftermath
2008-12-01	-9.35	Global Financial Crisis deepening
2008-09-29	-9.20	Initial banking crisis panic
Top 5 Booms		
2008-10-13	10.96	Government intervention announcements during GFC
2008-10-28	10.25	Relief rally on bank rescue measures
2025-04-09	9.09	Trump tariff pause relief
2020-03-24	8.98	COVID-19 stimulus package hopes
2020-03-13	8.91	Emergency Fed measures announced

Weekly Returns		
Date	Return (%)	Explanation
Top 5 Crashes		
2008-10-12	-20.02	Peak panic during GFC; bank failures and credit freezes
2020-03-22	-16.20	COVID-19 lockdowns and economic shutdown fears
2001-09-23	-12.29	September 11 terrorist attacks aftermath
2020-03-01	-12.15	Initial COVID-19 global spread shock
2000-04-16	-11.12	Dot-com bubble collapse
Top 5 Booms		
2020-04-12	11.46	COVID-19 crash rebound via fiscal stimulus
2008-11-30	11.41	Government rescue measures during GFC
2009-03-15	10.25	Financial crisis bottom rally; stimulus anticipation
2008-11-02	10.01	Policy intervention expectations during GFC
2020-03-29	9.78	Emergency rate cuts and pandemic relief

1.1.2 Normality Extremes For Large moves

To see how likely the extreme events, such as those in Table 1.1.1 are under a assumed normal distribution, I calculate probabilities of events as extreme or more as

$$P(X \geq x) = 1 - \Phi\left(\frac{x - \mu}{\sigma}\right) \quad \text{for booms}, \quad (2)$$

and

$$P(X \leq x) = \Phi\left(\frac{x - \mu}{\sigma}\right) \quad \text{for crashes}.$$

In the equations (2) Φ is the cumulative distribution function of the standard normal distribution, and μ and σ are the sample mean and standard deviation of the returns series, respectively. The results are summarized in Table 1.1.2.

The results in the table 1.1.2 show that the probabilities of observing such extreme returns under the normality assumption are extremely low, often in the range of 10^{-5} to 10^{-20} or even lower. This indicates that these extreme events that we observe in financial markets are extremely unlikely to occur if the returns were normally distributed.

This observations further supports the well-known fact that financial returns exhibit fat tails and are not well-modeled by a normal distribution. In another words magnitude of these crashes is not in line with the assumed normality.

Now that we have established that the extreme returns are not in line with normality, next I will study more general characteristics of the return distributions and test if these are aligned with normality using the Jarque-Bera test.

1.1.3 Test of Normality using Jarque-Bera Test

The table 1.1.3 displays sample skewness, kurtosis, Jarque-Bera (JB) statistics, and corresponding p-values for daily and weekly returns of the S&P 500 stock index, government bonds, and commodities. The JB test assesses whether the sample data deviates from a normal distribution by examining skewness and kurtosis.

The statistics in table 1.1.3 are computed as

$$S = \frac{1}{n} \sum_{i=1}^n \left(\frac{r_i - \bar{r}}{\sigma} \right)^3 \quad (\text{Sample skewness}) \quad (3)$$

$$K = \frac{1}{n} \sum_{i=1}^n \left(\frac{r_i - \bar{r}}{\sigma} \right)^4 \quad (\text{Sample kurtosis}) \quad (4)$$

$$JB = \frac{n}{6} \left(S^2 + \frac{(K - 3)^2}{4} \right), \quad (5)$$

where n is the sample size, S is the sample skewness, and K is the sample kurtosis. Under the null hypothesis of normality, the JB statistic follows a chi-squared distribution with 2 degrees of freedom, which I use to compute the p-values of the statistic.

Analyzing the results in Table 1.1.3, I note that the JB statistics are significantly large across all asset classes and frequencies, leading to p-values effectively equal to zero. Thus, we reject the null hypothesis of normality for all return series.

This further supports the earlier observations that asset returns are not normally distributed, and sample skewness and kurtosis values are not consistent with normality. Only exception is the skewness of government bonds weekly returns, which is close to zero.

Table 2: Probabilities of Observed Extreme Returns Under Normality Assumption
 Panel A: Weekly Returns Panel B: Daily Returns

Type	Return (%)	Z-score	Probability	Type	Return (%)	Z-score	Probability
S&P 500							
Crash	-20.02	-8.12	2.43e-16	Crash	-12.76	-10.65	8.66e-27
Crash	-16.20	-6.58	2.40e-11	Crash	-9.97	-8.33	4.07e-17
Crash	-12.29	-5.00	2.80e-07	Crash	-9.46	-7.90	1.37e-15
Crash	-12.15	-4.95	3.70e-07	Crash	-9.35	-7.81	2.90e-15
Crash	-11.12	-4.54	2.86e-06	Crash	-9.20	-7.68	7.82e-15
Boom	11.46	4.55	2.65e-06	Boom	10.96	9.10	0.00e+00
Boom	11.41	4.53	2.90e-06	Boom	10.25	8.51	0.00e+00
Boom	10.25	4.06	2.42e-05	Boom	9.09	7.55	2.26e-14
Boom	10.01	3.97	3.62e-05	Boom	8.98	7.45	4.62e-14
Boom	9.78	3.88	5.28e-05	Boom	8.91	7.40	7.03e-14
Government Bonds							
Crash	-1.96	-3.79	7.68e-05	Crash	-1.69	-6.77	6.22e-12
Crash	-1.93	-3.73	9.47e-05	Crash	-1.67	-6.71	9.80e-12
Crash	-1.92	-3.71	1.03e-04	Crash	-1.42	-5.69	6.22e-09
Crash	-1.81	-3.51	2.26e-04	Crash	-1.16	-4.66	1.58e-06
Crash	-1.79	-3.47	2.60e-04	Crash	-1.15	-4.61	1.97e-06
Boom	3.09	5.66	7.55e-09	Boom	1.79	7.07	7.59e-13
Boom	2.44	4.44	4.46e-06	Boom	1.74	6.89	2.79e-12
Boom	2.39	4.35	6.89e-06	Boom	1.55	6.12	4.72e-10
Boom	2.18	3.96	3.76e-05	Boom	1.38	5.46	2.41e-08
Boom	1.86	3.35	3.99e-04	Boom	1.26	4.96	3.51e-07
Commodities							
Crash	-16.06	-6.75	7.48e-12	Crash	-11.09	-10.32	2.92e-25
Crash	-12.82	-5.39	3.44e-08	Crash	-7.94	-7.39	7.27e-14
Crash	-11.88	-5.00	2.82e-07	Crash	-7.33	-6.83	4.33e-12
Crash	-11.00	-4.63	1.79e-06	Crash	-6.88	-6.40	7.62e-11
Crash	-10.11	-4.26	1.01e-05	Crash	-6.02	-5.61	1.02e-08
Boom	12.61	5.24	8.25e-08	Boom	5.88	5.45	2.53e-08
Boom	8.44	3.49	2.40e-04	Boom	5.75	5.33	5.05e-08
Boom	8.31	3.44	2.95e-04	Boom	5.53	5.12	1.54e-07
Boom	6.87	2.84	0.0023	Boom	5.22	4.84	6.64e-07
Boom	6.32	2.61	0.0046	Boom	5.21	4.82	7.01e-07

Table 3: Skewness, Kurtosis and Jarque-Bera Test for Normality of Returns on a daily and weekly frequencies for S&P 500, Government Bonds, and Commodities indices.

Panel A: Weekly Returns

Asset	Skewness	Kurtosis	JB-stat	p-value
S&P 500	-0.867	7.084	2998.44	0.0000
Government Bonds	0.001	2.171	265.67	0.0000
Commodities	-0.737	4.061	1052.19	0.0000

Panel B: Daily Returns

Asset	Skewness	Kurtosis	JB-stat	p-value
S&P 500	-0.350	11.099	34837.71	0.0000
Government Bonds	-0.015	3.273	3017.02	0.0000
Commodities	-0.536	5.339	8352.10	0.0000

1.1.4 Autocorrelation and the Ljung-Box test

Now that we have established that the returns are not normally distributed, next I will study if there is significant autocorrelation in the returns or in the squared returns.

We do this with the Ljung-Box test, which tests the null hypothesis that there is no autocorrelation up to a specified lag h . The test statistic is computed as

$$Q = n(n + 2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n - k}, \quad (6)$$

where n is the sample size, $\hat{\rho}_k$ is the sample autocorrelation at lag k , and h is the number of lags being tested. Under the null hypothesis, the Q statistic follows a chi-squared distribution with h degrees of freedom. The table 1.1.4 shows the computed values for the LB

Figures 1.1.4 and 1.1.4 present the autocorrelation function (ACF) plots for daily and weekly returns as well as for squared daily and weekly returns of the selected assets.

The Figures 1.1.4 clearly display that there is only very small autocorrelation in the returns themselves, as most of the ACF values are within the 95% confidence intervals. However, as expected, the ACF plot of squared returns 1.1.4 are highly positively autocorrelated, indicating volatility clustering and predictability in the squared returns.

The Ljung-Box statistics in the table 1.1.4 confirms that the null hypothesis of no autocorrelation in 10 lags is rejected at 5% confidence level for all asset classes and frequencies except Commodities at the daily frequency.

To conclude this section, there is a significant autocorrelated component in the the returns series. In next section we check are these results robust to different frequencies.

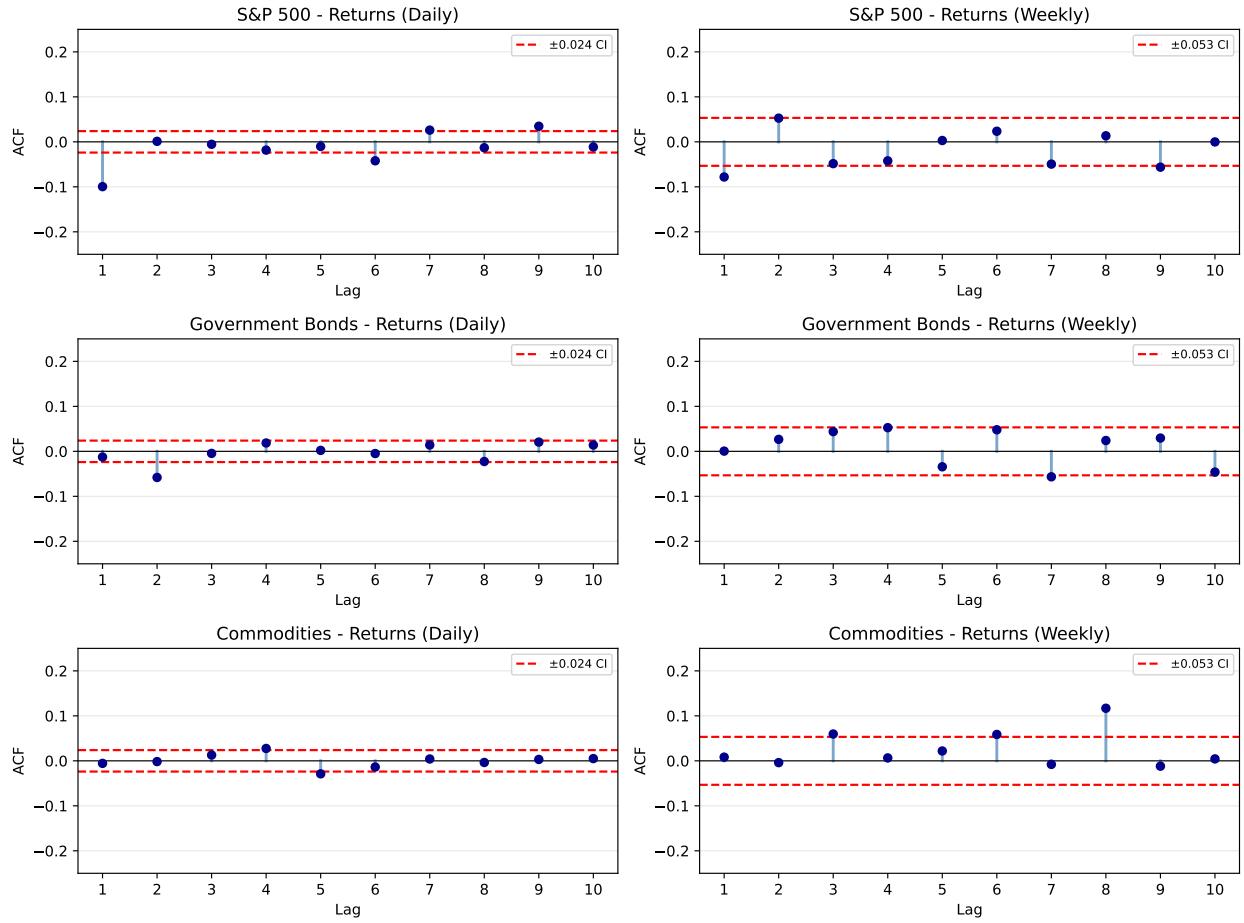


Figure 1: Autocorrelation Function (ACF) and Barlett's formula for individual autocorrelations. The plots show the ACF and the computed 95% confidence intervals for Daily Returns of S&P 500, Government Bonds, and Commodities. The top row shows the ACF plots up to lag 10. Horizontal dashed lines in the ACF plots represent the 95% confidence intervals. The CIs are calculated as $ci = 1.96/\sqrt{n}$.

Table 4: Ljung-Box Test Statistics and p-values for Daily and Weekly Returns and Squared Returns of S&P 500, Government Bonds, and Commodities indices. The table

Panel A: Returns	Panel B: Squared Returns
------------------	--------------------------

Asset	Frequency	Q-stat	p-value	Asset	Frequency	Q-stat	p-value
S&P 500	Daily	97.01	0.0000	S&P 500	Daily	6025.09	0.0000
	Weekly	26.32	0.0033		Weekly	386.75	0.0000
Government Bonds	Daily	35.53	0.0001	Government Bonds	Daily	2012.88	0.0000
	Weekly	21.32	0.0190		Weekly	117.68	0.0000
Commodities	Daily	13.86	0.1794	Commodities	Daily	1845.51	0.0000
	Weekly	29.26	0.0011		Weekly	394.60	0.0000

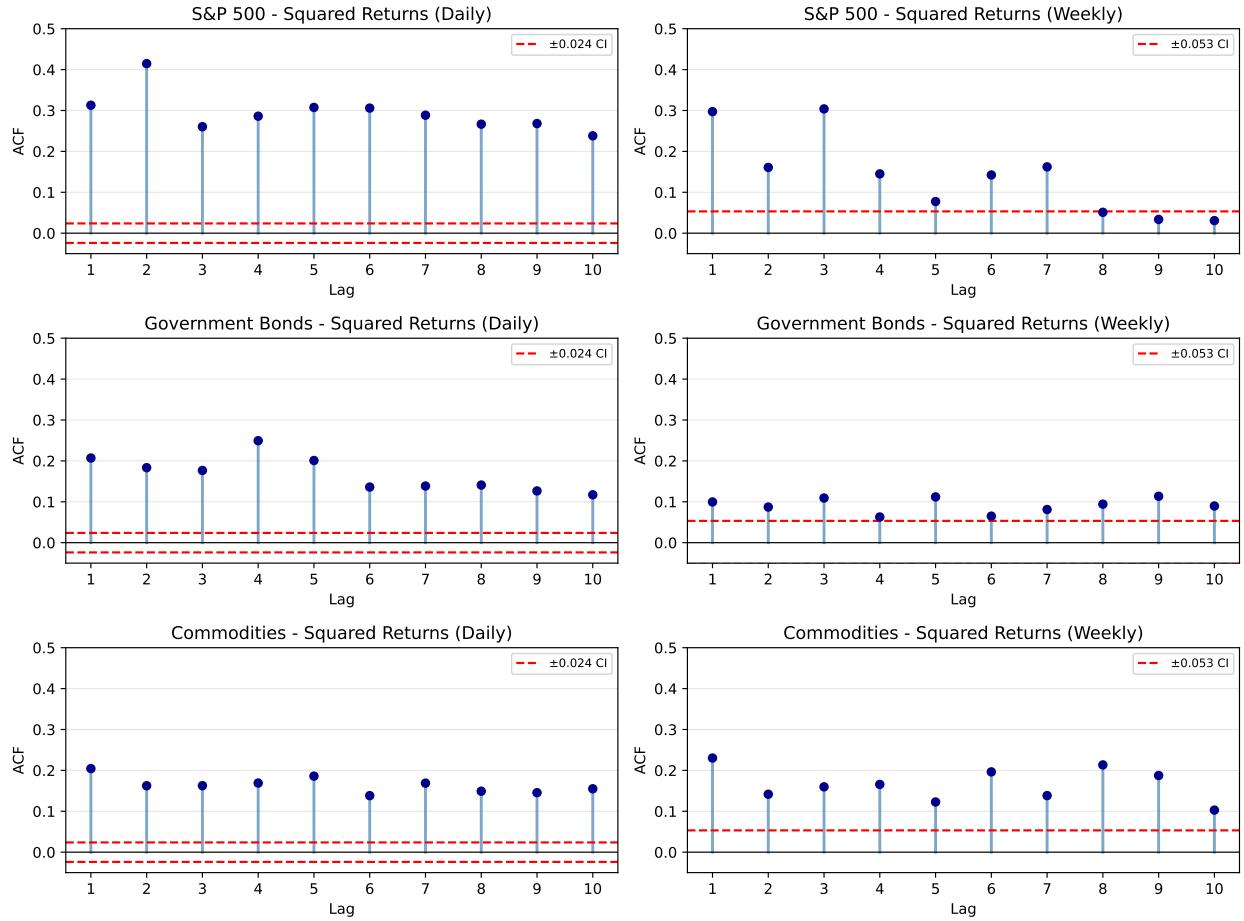


Figure 2: Autocorrelation Function (ACF) and Barlett's formula for individual autocorrelations. The plots show the ACF and the computed 95% confidence intervals for Squared Daily Returns of S&P 500, Government Bonds, and Commodities. The top row shows the ACF plots up to lag 10. Horizontal dashed lines in the ACF plots represent the 95% confidence intervals. The CIs are calculated as $ci = 1.96/\sqrt{n}$

Table 5: Effect of Temporal Aggregation on Normality of S&P 500 Returns

Frequency	N	Skewness	Kurtosis	JB-stat	JB p-val	Q-stat	LB p-val
Daily	6760	-0.350	11.099	34837.71	0.0000	97.01	0.0000
Weekly	1352	-0.867	7.077	2990.39	0.0000	26.36	0.0033
Monthly	311	-0.677	1.144	40.73	0.0000	15.56	0.1130
Yearly	26	-1.347	1.782	11.30	0.0035	25.06	0.0052

1.1.5 Robustness to change in Frequency

As sampling frequency decreases, from daily to weekly and to monthly returns should tend towards normality if they are independent with finite second moment¹ then by CLT the returns should be closer to normality longer sampling periods.

Looking at the tables 1.1.2 of extremes and the table of normality tests 1.1.3 this is exactly what we see. The extreme moves are less extreme (in terms of Z score) on weekly frequencies than in daily. Similarly JB statistics tend to be smaller for longer frequencies.

For the S&P 500 index, table ?? summarizes this result on daily, weekly, monthly, and yearly and monthly frequencies. Looking at the table I find that returns become closer to normal with smaller skewness and kurtosis values and lower JB statistics and higher p-values as frequency decreases. Similarly, the autocorrelations tend to be stronger at lower frequencies as displayed by the Ljung-Box statistics.

1.2 Diagnostics for a Portfolio

In this section I will study the characteristics of portfolios consisting of stocks, government bonds, and commodities. The portfolios are equal-weighted and the returns are computed as *simple* returns as they allow easy cross sectional aggregation based on portfolio weights. The simple returns are computed as

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}. \quad (7)$$

The analyses are done using daily and weekly simple returns of the selected indices.

1.2.1 Summary Statistics of Daily Portfolios

Table 1.2.1 presents summary statistics for daily simple returns of individual index series and an equal-weighted portfolio of these assets. The statistics include mean return, standard deviation, skewness, kurtosis, minimum and maximum returns, and the number of observations.

Comparing the portfolio to holding only stock index, we see that the portfolio has significantly lower standard deviation, and extreme values than the stock index alone.

However, when aggregating cross sectionally we see that the skewness and kurtosis of the portfolio returns remain quite high and skewness is slightly elevated when compared to the single equity index. This indicates non-normality and the fact that with such a small number of likely cross-sectionally dependent assets the diversification benefits are limited and the contemporaneous aggregate toward normality is weak and slow. In another words, we would need more assets that are more independent from one another.

¹Or satisfy some other conditions for CLT to hold

Table 6: Summary statistics for individual index series and a equal weighted portfolio of these assets.

Asset	Mean (x100)	Std (x100)	Skewness	Kurtosis	Min (x100)	Max (x100)
S&P 500	0.037	1.199	-0.117	10.957	-11.98	11.58
Government Bonds	0.013	0.251	0.004	3.285	-1.67	1.80
Commodities	0.023	1.074	-0.424	4.896	-10.50	6.06
Portfolio	0.024	0.617	-0.707	9.832	-6.47	4.64

Table 7: Summary statistics for individual index series and a equal weighted portfolio of these assets.

Asset	Mean (x100)	Std (x100)	Skewness	Kurtosis	Min (x100)	Max (x100)
S&P 500	0.180	2.466	-0.571	5.872	-18.14	12.15
Government Bonds	0.066	0.534	0.035	2.214	-1.94	3.14
Commodities	0.113	2.376	-0.544	3.516	-14.83	13.44
Portfolio	0.120	1.304	-1.016	7.266	-10.30	6.79

1.2.2 Summary Statistics of Weekly Portfolios

Table 1.2.2 presents summary statistics for weekly asset and portfolio returns, similarly as the daily table 1.2.1.

Again, we see that the portfolio has lower standard deviation and extreme values than the stock index alone. Also skewness and kurtosis have similar behaviour and there is no strong convergence toward normality.

In terms of temporal aggregation I note that the weekly returns have naturally higher means and standard deviations which should scale approximately linearly and with square root of time, respectively. However, I also note that the convergence toward normality in temporal dimension is weaker than with log returns, which is as expected since the simple return does not aggregate linearly over time, but rather multiplicatively.

1.2.3 Implications to Risk and Portfolio Management

From a risk and portfolio management perspective the results are interesting. First there are contemporaneous diversification benefits by aggregating across assets, which reduces standard deviation of the portfolio and increases diversification. This diversification is costly in terms of lower expected returns, however for each unit of risk the portfolio return increases as if the assets added to the portfolio are not perfectly correlated.

Second the weak convergence towards normality implies that normal distributions should be used only for large portfolios with multiple independent assets. For small portfolios the non-normality is significant and should be taken into account in risk management. This is especially true in times of market distress when extreme moves are more likely to occur.

2 Modeling Volatility and Non-Normality

This section studies volatility modeling using GARCH(1,1) model for volatility of Corporate Bond and S&P 500 stock index excess log returns.

The returns are calculated as follows: First, I calculate the daily excess returns for corporate bonds and stock index by subtracting the risk-free (proxied by the government bond returns) using simple returns, giving the simple excess returns for stocks and bonds as

$$R_{i,t}^{excess} = R_{i,t} - R_{rf,t} \text{ for } i \in \{\text{stocks, bonds}\}. \quad (8)$$

Then, I convert the simple excess returns to log excess returns as

$$r_{i,t}^{excess} = \log(1 + R_{i,t}^{excess}) \text{ for } i \in \{\text{stocks, bonds}\}. \quad (9)$$

When building the GARCH model in this section I effectively assume that the excess log return process has a autocorrelated component, and the residual process follows a GARCH(1,1) model with non-normal innovations. In another words, the model is specified as

$$r_{i,t}^{excess} = \mu + \phi r_{i,t-1}^{excess} + \epsilon_{i,t}, \quad (10)$$

$$\epsilon_{i,t} = \sigma_{i,t} z_{i,t}, \quad (11)$$

$$\sigma_{i,t}^2 = \omega + \alpha \epsilon_{i,t-1}^2 + \beta \sigma_{i,t-1}^2, \quad (12)$$

where $z_{i,t}$ are i.i.d. innovations following a Normal, a Student-t, or a Skewed Student-t distributions. It is worth to note that the specification of the AR(1) model I use in this project contains the intercept term μ as well as the lagged return term with coefficient ϕ , I keep the intercept term because asset pricing theory predicts that expected excess returns should be positive on average, under the physical measure, due to risk premia associated with the risky assets.

2.1 Estimation of GARCH model

2.1.1 Non-Normality and autocorrelation

Before starting to estimate the GARCH model, I first study the non-normality and autocorrelation characteristics of the excess log returns to confirm the model assumptions. Table 2.1.1 presents the Jarque-Bera and Ljung-Box statistics for daily excess log returns of Corporate Bonds and S&P 500 stock index.

First, the Table 2.1.1 dispalys very large Jarque-Bera statistics for both Corporate Bonds and S&P 500 excess log returns, with p-values effectively equal to zero. This indicates that we reject the null hypothesis of normality for both return series, supporting the evidence of non-normality in financial returns.

Second, the Ljung-Box statistics for both return series are also significant, for both series at the 1% confidence level. The autocorrelation effect tends to be even stronger in in the corporate bonds excess log returns as displayed by the very high Ljung-Box statistic and the p-value effectively equal to zero. These results indicate that there is significant autocorrelation in the excess log returns that we have to filter out before estimating the GARCH model.

2.1.2 AR(1) model to filter autocorrelation

To filter the observed autocorrelation in the excess log returns I use AR(1) model as specified in equation (10). The model is estimated using conditional maximum likelihood estimation (MLE). The estimated coefficients are displayed in table 2.1.2.

Table 8: Jarque-Bera, and Ljung-Box statistics for Corporate Bond and Stock (S&P 500) excess log returns. Excess returns are computed by subtracting government bond returns from the respective simple asset returns.

Asset	JB-stat	JB p-value	Q-stat	LB p-value
Stocks	25446.38	0.0000	14.99	0.0047
Corp Bonds	44381.78	0.0000	788.89	0.0000

Table 9: Estimated AR(1) Model for excess corporate bond and stock returns. The table displays the estimated coefficients along with their standard errors (in parentheses), t-statistics, and p-values. The model is estimated using conditional maximum likelihood estimation (MLE) with Normal distribution. The standard errors are Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors with 6 lags.

Parameter	Coefficient	Std. Error	p-value
Panel A: S&P 500			
Constant	0.000173	0.000142	0.2237
ϕ_1	-0.038332	0.026339	0.1456
Panel B: Corporate Bonds			
Constant	0.000080	0.000049	0.1045
ϕ_1	0.243030	0.033753	0.0000

The coefficients in table 2.1.2 confirm that the autocorrelation component is non-existent in both series. However, the autocorrelation coefficient is larger for corporate bonds than stocks, and when standard errors are HAC adjusted the coefficient is not statistically significant for stocks. Furthermore, the drift term is positive but as expected it is small in magnitude and significance, especially for corporate bonds.

The figure 2.1.2 displays the residuals from the estimated AR(1) models for both excess return series. We observe that there is significant volatility clustering in both residual series, indicating that a time-varying volatility model such as GARCH is appropriate for modeling the residuals.



Figure 3: Residuals from the estimated AR(1) models for Corporate Bond and Stock (S&P 500) excess log returns. The plots show the time series of residuals, highlighting volatility clustering in both series.

2.1.3 ARCH effects in AR(1) residuals

Given the timevarying volatility in the estimated residuals, before going to GARCH estimation, I first test for ARCH effects in the AR(1) residuals using the Engle's ARCH test. The results are summarized in table 2.1.3.

The results in table 2.1.3 show that the squared residuals from the AR(1) model exhibit statistically significant autocorrelation, as the ARCH test statistics are large and the p-values are effectively equal to zero for both series. Thus, we reject the null hypothesis of no ARCH effects in both series.

Table 10: Engle's ARCH Lagrange Multiplier Test for ARCH effects in AR(1) residuals of Corporate Bond and Stock (S&P 500) excess log returns. The table displays the test statistics along with their p-values. The null hypothesis of the test is that there are no ARCH effects up to lag 10. The test statistic tests the null hypothesis of $Var_t(\epsilon_{t+1}) = \sigma^2$, and under null the statistic follows a chi-squared distribution with degrees of freedom equal to the number of lags tested (here 4).

Asset	LM-stat	p-value	Reject H ₀
S&P 500	1489.21	0.0000	Yes
Corporate Bonds	926.17	0.0000	Yes

Table 11: Estimated GARCH(1,1) Model for AR(1) residuals of Corporate Bond and Stock (S&P 500) excess log returns with Normal innovations. The table displays the estimated coefficients along with their standard errors (in parentheses), t-statistics, and p-values. The model is estimated using conditional maximum likelihood estimation (MLE) with Normal distribution. The parameters are estimated from the residuals of the AR(1) model fitted to excess returns and for reasons of numerical stability the residuals are scaled by 100 before estimation. Because of the scaling the estimated parameter ω is in units of 100^2 , but the α and β , and $\alpha + \beta$ are not scaled (as these are scale invariant).

Asset	ω	α	β	Persistence	Log-Lik	AIC
S&P 500	0.0233 (0.0046)	0.1113 (0.0101)	0.8712 (0.0109)	0.9825	-9390.35	18786.70
Corporate Bonds	0.0011 (0.0004)	0.0950 (0.0156)	0.8995 (0.0159)	0.9945	-1428.42	2862.85

2.1.4 GARCH model for AR(1) model residuals

Given the statistically significant ARCH effects in the AR(1) squared residuals, I now proceed to estimate the GARCH(1,1) model to model the time-varying volatility, as specified in equations (11) and (12). The table 2.1.4 displays the estimated GARCH(1,1) model coefficients for both excess return series.

Looking at the Table 2.1.4, I note that the coefficients for both series are statistically significant at 1% confidence level, and economically large in the magnitude.

Furthermore, the sum of the α and β coefficients, also known as the persistence parameter, is close to 1 for both series, indicating high persistence in volatility shocks. This suggests that volatility shocks have a long-lasting effect and spikes in volatility decay slowly over time.

2.2 MLE or QMLE

The previous section showed that there is significant persistence in volatility shocks for both excess return series. In this section I will analyze the adequacy of Normal distribution for modeling the shocks in the GARCH innovation process.

The analysis in this section and in the next section applies Diebold-Gunther-Tay (DGT) test

for goodness-of-fit of the assumed distribution for the GARCH innovations. The test and the way I apply it is described as follows:

DGT Test Implementation Details Let a fitted GARCH model deliver conditional volatility $\hat{\sigma}_t$ and standardized innovations

$$\hat{z}_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t},$$

together with an assumed innovation CDF F for z_t (e.g. Gaussian, Student- t , skew- t). The DGT test evaluates the adequacy of the conditional distribution by forming the probability integral transforms (PITs)

$$u_t = F(\hat{z}_t).$$

Under correct specification of both the volatility dynamics and the innovation distribution, the PITs then satisfy

$$u_t \stackrel{iid}{\sim} U(0, 1).$$

I then implement the test in two parts. First, I check uniformity by partitioning $[0, 1]$ into N ($N = 40$) equal bins, let F_n be the number of u_t in bin n , and compute Pearson's statistic

$$\text{DGT}(N) = \sum_{n=1}^N \frac{(F_n - T/N)^2}{T/N} \sim \chi^2(N-1)$$

under H_0 . Second, to check independence and dynamics I define centered PITs $v_t = u_t - 1/2$ and apply a Ljung–Box test at lag 10 to v_t (and optionally higher moments), where the reported p-value corresponds to the joint null of no autocorrelation up to lag m . In the code, u_t is computed as $F(\hat{z}_t)$ using the chosen F , the chi-square test uses the binned PIT counts, and the Ljung–Box test is applied to v_t^k for selected moments k .

2.2.1 Estimate ML estimates for GARCH with Normal Distribution

I start by estimating the GARCH(1,1) model for the residuals of the AR(1) model fitted in earlier section using Maximum Likelihood Estimation (MLE) assuming that the innovations follow a Normal distribution. Table 2.2.1 summarize the results of this estimation, and the adequacy of the Normal distribution for modeling the standardized residuals from the GARCH model by testing Normality of the GARCH model residuals using JB test. First panel in the table 2.2.1 displays the estimated GARCH(1,1) model parameters assuming Normal innovations. It is worth noting that the parameter estimates are the same as in table 2.1.4 since both estimations use MLE with Normal distribution. However, the standard errors are different as in the previous section they were computed using robust HAC standard errors, while here they are computed using the information matrix, which is valid under the assumption of correct model specification.

However, as we can see in the the second panel of table 2.2.1, these standard errors are likely to be too small as the Jarque-Bera statistics for the standardized residuals from the GARCH model are very large with p-values effectively equal to zero.

Thus, the assumption that the innovations forllow normality can be questioned, and we should consider estimators that weaken this normality assumption.

Table 12: Estimated GARCH(1,1) Model Parameters and Normality Test for Standardized Residuals, using the Maximum Likelihood Estimation (MLE) assuming Normal innovations. The first panel displays the estimated GARCH(1,1) model parameters along with their standard errors (in parentheses), t-statistics, and p-values. The second panel displays the Jarque-Bera statistics and p-values for testing normality of the standardized residuals from the GARCH model.

Panel A: GARCH(1,1) Parameter Estimates

Asset	ω	α	β	Persistence	AIC
S&P 500	0.0233 (0.0028)	0.1113 (0.0079)	0.8712 (0.0084)	0.9825	18786.70
Corporate Bonds	0.0011 (0.0002)	0.0950 (0.0081)	0.8995 (0.0084)	0.9945	2862.85

Panel B: Jarque-Bera Test for Normal Residuals

Asset	JB-stat	p-value	Skewness	Excess Kurtosis
S&P 500	1174.63	0.0000	-0.4830	1.7995
Corporate Bonds	591.88	0.0000	-0.2186	1.3823

Table 13: Estimated GARCH(1,1) Model Parameters with QML and Normal Distribution. The table displays the estimated coefficients along with their distribution robust standard errors (in parentheses), t-statistics, and p-values.

Parameter	Estimate	SE(MLE)	SE(QML)	Ratio
Panel A: S&P 500				
ω	0.0233	0.0028	0.0046	1.662
α	0.1113	0.0079	0.0101	1.275
β	0.8712	0.0084	0.0109	1.303
Persistence	0.9825			
Panel B: Corporate Bonds				
ω	0.0011	0.0002	0.0004	1.528
α	0.0950	0.0081	0.0156	1.927
β	0.8995	0.0084	0.0159	1.904
Persistence	0.9945			

2.2.2 Estimate QML estimates for GARCH with Normal Distribution

To weaken the assumption of normality in the innovation process, I estimate the GARCH(1,1) model using Quasi-Maximum Likelihood Estimation (QMLE) assuming Normal innovations. Table 2.2.2 summarize the results of this estimation.

The estimated parameters from therger than those fomr the MLE estimation assuming normality. QML estimation in table 2.2.2 are same as in the MLE with normal model, as they should be. However, now the standard errors are computed using the robust sandwich estimator and are significantly la Where as the MLE standard errors were computed as

$$SE_{MLE} = \sqrt{\text{diag}(H^{-1})}, \quad (13)$$

where H is the Hessian matrix of the log-likelihood function. For a comparison the sandwich standard errors for the QMLE estimation are computed as

$$SE_{QML} = \sqrt{\text{diag}(H^{-1}GH^{-1})}, \quad (14)$$

where $G = \sum_{t=1}^T s_t s_t'$ (outer product of score vectors), and s_t is the score vector at time t . This form doens't make distributional assumptions similarly as the classical MLE standard errors do.

Under correct specification: $G \approx H$, so $H^{-1}GH^{-1} \approx H^{-1}$ and both standard errors are similar. Thus, the large difference in the standard errors indicates violations for the normality assumption.

2.2.3 Adequacy of Normal Distribution for Standardized Residuals

To further diagnose the Normality problem in the GARCH innovation process, I now apply the Diebold-Gunther-Tay (DGT) test with $N = 40$ cells.

The tests results are presented in Table 2.2.3. First, the uniformity results in Panle A show that the Normal distribution is not adequate for modeling the standardized residuals from the GARCH(1,1) model estimated with the conditional ML. In another words, the empirical distribution of the standardized residuals deviates significantly from the theoretical Normal distribution. Second, the Ljung-Box results in Panel B show that there is significant dependency in the stan-dardized residuals, indicating that our model does not fully capture the dynamics of the process.

Overall, I reject the DGT null hypothesis that standardized innovations follow a Normal distribution. Thus, next, I consider other possible non-normal distributions.

2.3 Estimation of GARCH with a Student-t

As previous sections show the innovations in the volatility process of the GARCH model tend to behave in a non-normal manner. Thus, in this section I will model the GARCH(1,1) process assuming the innovation z_t in the definition (12) follows a Studen-t distribution.

2.3.1 Fitting GARCH(1,1) with Student-t

Table 2.3.1 displays the estimated GARCH(1,1) model coefficients for both excess return series assuming Student-t innovations. The results are shown in the table 2.3.1.

The parameter estimates are highly similar with those estimated earlier using the Normal Dis-tribution, which tells us that the volatility dynamics are robust to the distributional assumption. Furthermore, also the Student-t specification clearly shows that there is strong persistance in the volatility, as we can see from the table 2.3.1.

Table 14: Diebold-Gunther-Tay Test for Adequacy of Normal Distribution for Standardized Residuals from GARCH(1,1) model estimated with QMLE. The table displays the DGT test statistics along with their p-values. The null hypothesis of the test is that the standardized residuals follow a Normal distribution. Under the null hypothesis, the DGT test statistic follows a chi-squared distribution with $N - 1$ degrees of freedom, where N is the number of cells (here 40).

Panel A: DGT Uniformity Test

Asset	χ^2 Statistic	p-value
S&P 500	317.99	0.0000
Corporate Bonds	149.33	0.0000

Panel B: DGT Ljung-Box Test

Asset	Moment	Lag	LB Statistic	p-value
S&P 500	v^1	10	27.40	0.0022
	v^2	10	73.62	0.0000
	v^3	10	21.47	0.0180
	v^4	10	84.21	0.0000
Corporate Bonds	v^1	10	147.94	0.0000
	v^2	10	42.06	0.0000
	v^3	10	108.46	0.0000
	v^4	10	41.54	0.0000

Table 15: Estimated GARCH(1,1) Model for AR(1) residuals of Corporate Bond and Stock (S&P 500) excess log returns with Student-t innovations. The table displays the estimated coefficients along with their standard errors (in parentheses), t-statistics, and p-values. The model is estimated using conditional maximum likelihood estimation (MLE) with Student-t distribution. Because of the scaling the estimated parameter ω is in units of 100^2 , but the α and β , and $\alpha + \beta$ are not scaled (as these are scale invariant).

Asset	ω	α	β	ν	Persistence	AIC
S&P 500	0.0141 (0.0030)	0.1107 (0.0097)	0.8848 (0.0095)	6.1215 (0.4617)	0.9956	18480.30
Corporate Bonds	0.0009 (0.0003)	0.0779 (0.0122)	0.9161 (0.0131)	8.3290 (0.7979)	0.9940	2675.83

Table 16: Delta Method Test for $H_0 : 1/\nu = 0$ in GARCH(1,1) model with Student-t innovations for Corporate Bond and Stock (S&P 500) excess log returns. The table displays the estimated degrees of freedom parameter $\hat{\nu}$, its standard error, the computed test statistic, and the corresponding p-value for the test.

Asset	$\hat{\nu}$	$1/\hat{\nu}$	t-statistic	p-value
S&P 500	6.1215 (0.4617)	0.163359 (0.012320)	13.2600	0.0000
Corporate Bonds	8.3290 (0.7979)	0.120062 (0.011501)	10.4391	0.0000

However, the degrees of freedom parameter, ν , is estimated to be around 6 for stocks and around 8 for corporate bonds, implying fat tails in the innovation distribution. This further supports the evidence of non-normality in the volatility process.

Delta Method for Testing $H_0 : 1/\nu = 0$ Now that we have estimated the GARCH(1,1) model with Student-t innovations, we apply the Delta Method to test the null hypothesis that $1/\nu = 0$, which corresponds to the case of infinite degrees of freedom, in which case the distribution is normal.

The delta method provides the asymptotic distribution of transformations of maximum likelihood estimators. For a differentiable function $g(\cdot)$ and MLE $\hat{\nu}$ with asymptotic normality $\sqrt{n}(\hat{\nu} - \nu) \xrightarrow{d} N(0, \sigma_\nu^2)$, we have $\sqrt{n}(g(\hat{\nu}) - g(\nu)) \xrightarrow{d} N(0, [g'(\nu)]^2 \sigma_\nu^2)$. Applying this to $g(\nu) = 1/\nu$ with derivative $g'(\nu) = -1/\nu^2$, the asymptotic standard error is $SE(1/\hat{\nu}) = |g'(\hat{\nu})| \times SE(\hat{\nu}) = SE(\hat{\nu})/\hat{\nu}^2$, yielding the test statistic

$$t = \frac{1/\hat{\nu} - 0}{SE(1/\hat{\nu})} = \frac{\hat{\nu}}{SE(\hat{\nu})} \stackrel{H_0}{\sim} N(0, 1). \quad (15)$$

The results for the delta method are summarized in table 2.3.1, and it shows that applying the delta method we reject the null hypothesis that $1/\nu = 0$ for both corporate bonds and stocks.

However, it is worth to keep in mind that *testing $H_0 : 1/\nu = 0$ with Delta-Method might be*

problematic because this null corresponds to the boundary point $\nu = \infty$, which lies outside the parameter space $(0, \infty)$. The standard delta method assumes linearization around a finite interior point, but at $\nu = \infty$ the derivative $g'(\nu) \rightarrow 0$, causing the first-order approximation to be not well defined. A more appropriate approach might be to test $H_0 : \nu \geq \nu^*$ for a large finite threshold ν^* , or to employ alternative tests for normality that avoid this boundary issue.

2.3.2 Test of Adequacy of Student-t on GARCH(1,1) residuals

Now that we have estimated the GARCH(1,1) model with Student-t innovations, I will test the adequacy of the Student-t distribution for modeling the standardized residuals from the GARCH model using the DGT test, similarly to what we did with normal distribution in previous section.

Table ?? displays the results of the DGT test for adequacy of Student-t distribution for standardized residuals from the GARCH(1,1) model estimated with Student-t innovations.

Looking first at the uniformity test in panel A of table ???. For stocks we reject the uniformity null hypothesis and note that the statistic values are very large. This implies that the Student-t GARCH density is not calibrated for equities and the realized innovations fall too often in some regions relative to what the model predicts. For corporate bonds, we fail to reject the null hypothesis of uniformity, indicating that the Student-t distribution looks calibrated good for corporate bonds.

However the independence test in panel B of table ?? show that for both stocks and corporate bonds we reject the null hypothesis of independence. This means that the model is not capturing all dynamics in the conditional distribution. Especially, for corporate bonds the test statistics are very large indicating serial correlation / smoothing (stale prices, evaluated pricing), which is typical in non-traded securities.

To conclude, for both asset classes DGT test rejects the null hypothesis that innovations from GARCH(1,1) follow Student-t distribution.

2.4 Estimation of GARCH with a Skewed Student-t

As observed in the previous section, using Student-t distribution for the innovations in the GARCH(1,1) model provided a slight improvement over the Normal distribution, but still failed to adequately capture the characteristics of the standardized residuals. In this section, we expand on this by estimating the GARCH(1,1) model assuming that the innovations follow a Skewed Student-t distribution, which allows for both fat tails and skewness in the distribution of innovations.

Here, to fit the skewed Student-t distribution, we use the parameterization proposed by Hansen [1994].

2.4.1 Fitting GARCH(1,1) with Skewed Student-t

Table 2.4.1 displays the estimated GARCH(1,1) model coefficients for both excess return residual series assuming Skewed Student-t innovations.

The estimation results yield very similar GARCH(1,1) estimates as Normal distribution in Table 2.2.2, and the symmetric Student-t distribution in 2.3.1. Overall, the volatility dynamics appear robust to the distributional assumption, and the persistence in volatility shocks remains high across all specifications.

Interestingly in table 2.4.1 we see that the degrees of freedom parameter, ν , is estimated to be around 7 for stocks and around 9 for corporate bonds, similar to the symmetric Student-t case, indicating fat tails in the innovation distribution. Furthermore, the skewness parameter, λ , is

Table 17: Diebold-Gunther-Tay Test for Adequacy of Student-t Distribution for Standardized Residuals from GARCH(1,1) model estimated with Student-t innovations.

Panel A: DGT Uniformity Test

Asset	χ^2 Statistic	p-value
S&P 500	128.23	0.0000
Corporate Bonds	110.90	0.0000

Panel B: DGT Ljung-Box Test

Asset	Moment	Lag	LB Statistic	p-value
S&P 500	v^1	10	27.65	0.0021
	v^2	10	54.26	0.0000
	v^3	10	22.76	0.0117
	v^4	10	66.36	0.0000
Corporate Bonds	v^1	10	147.31	0.0000
	v^2	10	36.39	0.0001
	v^3	10	120.46	0.0000
	v^4	10	29.56	0.0010

Table 18: Estimated GARCH(1,1) Model for AR(1) residuals of Corporate Bond and Stock (S&P 500) excess log returns with Skewed Student-t innovations. The table displays the estimated coefficients along with their standard errors (in parentheses), t-statistics, and p-values. The model is estimated using conditional maximum likelihood estimation (MLE) with Skewed Student-t distribution. Because of the scaling the estimated parameter ω is in units of 100^2 , but the α and β , and $\alpha + \beta$ are not scaled (as these are scale invariant).

Asset	ω	α	β	η	λ	Persistence	AIC
S&P 500	0.0139 (0.0030)	0.1108 (0.0094)	0.8847 (0.0092)	6.31 (0.49)	-0.1149 (0.0135)	0.9955 (0.0186)	18424.06
Corporate Bonds	0.0009 (0.0003)	0.0762 (0.0118)	0.9178 (0.0127)	8.41 (0.81)	-0.0366 (0.0158)	0.9940 (0.0246)	2672.87

Table 19: Delta method applied to the joint test for $H_0 : 1/\nu = 0$ AND $\lambda = 0$ in GARCH(1,1) model with Skewed Student-t innovations for Corporate Bond and Stock (S&P 500) excess log returns residuals. The table displays the computed test statistic along with its p-value for the test.

Asset	$\hat{\eta}$	$1/\hat{\eta}$	$\hat{\lambda}$	Joint Stat.	p-value
S&P 500	6.31 (0.49)	0.1586 (0.0122)	-0.1149 (0.0135)	247.49	0.0000
Corporate Bonds	8.41 (0.81)	0.1189 (0.0114)	-0.0366 (0.0158)	112.27	0.0000

estimated to be negative for both asset classes, suggesting that the innovation distribution is left-skewed, which is consistent with the common observation of negative skewness in financial return distributions.

The skewness parameter is also statistically significant at 1% confidence level for both asset classes, indicating that the skewness is an important feature of the innovation distribution that should be accounted for in the model.

Joint Wald Test for $H_0 : 1/\eta = 0$ AND $\lambda = 0$ is applied similarly as with earlier with the delta method for testing normality in the Student-t case, however now we consider the joint test for both fat tails and skewness. Taking derivatives of the transformation vector $\mathbf{g}(\boldsymbol{\theta}) = [1/\eta, \lambda]'$ I get the Jacobian $\mathbf{G} = \text{diag}(-1/\eta^2, 1)$. Furthermore, given the asymptotic covariance is $\text{Var}(\mathbf{g}(\hat{\boldsymbol{\theta}})) = \mathbf{G} \text{Var}(\hat{\boldsymbol{\theta}}) \mathbf{G}'$, I get the yielding the test statistic

$$W = \mathbf{g}(\hat{\boldsymbol{\theta}})' [\text{Var}(\mathbf{g}(\hat{\boldsymbol{\theta}}))]^{-1} \mathbf{g}(\hat{\boldsymbol{\theta}}) \stackrel{H_0}{\sim} \chi^2(2). \quad (16)$$

However, again it is worth to keep in mind that testing $H_0 : 1/\nu = 0$, might not be well defined as discussed in the previous section in the case of standard Student t.

Table 2.4.1 summarizes the result of Delta method applied to the parameter vector. The table shows that we reject the null hypothesis that both $1/\nu = 0$ and $\lambda = 0$ for both corporate bonds and stocks.

2.4.2 Test of Adequacy of Skewed Student-t on GARCH(1,1) residuals

Now we test the adequacy of the Skewed Student-t distribution for modeling the standardized residuals from the GARCH model using the DGT test, similarly to what we did with normal and symmetric Student-t distributions in previous sections. Table 2.4.2

Similarly to earlier sections, Table 2.4.2 rejects the null hypothesis that the standardized garch residuals in the GARCH(1,1) model (11) follow a Skewed Student-t distribution.

As earlier, the Panel B of the table shows that there is significant autocorrelation in the residuals for both asset classes, indicating that the model is not capturing all dynamics in the conditional distribution.

In terms of uniformity in panel A of the table 2.4.2, the conclusions are similar to those made with standard Student-t distribution, meaning that for stocks we reject the uniformity null hypothesis and for corporate bonds we fail to reject the null hypothesis of uniformity.

Table 20: Diebold-Gunther-Tay Test for Adequacy of Skewed Student-t Distribution for Standardized Residuals from GARCH(1,1) model estimated with Skewed Student-t innovations.

Panel A: DGT Uniformity Test

Asset	χ^2 Statistic	df	p-value
S&P 500	65.37	39	0.0051
Corporate Bonds	34.31	39	0.6837

Panel B: DGT Ljung-Box Test

Asset	Moment (v^k)	LB Statistic	df	p-value
S&P 500	v^1	28.97	10	0.0013
	v^2	18.51	10	0.0469
	v^3	29.30	10	0.0011
	v^4	16.00	10	0.0997
Corporate Bonds	v^1	147.88	10	0.0000
	v^2	41.43	10	0.0000
	v^3	124.67	10	0.0000
	v^4	32.66	10	0.0003

However, it is worth to emphasize that the uniformity test statistic for stocks is significantly smaller than with standard Student-t distribution, indicating that the Skewed Student-t distribution provides a better fit for stocks and the asymmetry in skewness is an important feature to capture in the innovation distribution.

2.4.3 Samlple Moments of Normal ML GARCH vs Theoretical Moments from Skewed Student-t

The results this far illustrate that even though the Skewed Student-t distriubtion was rejected by the DGT test, it has been the most well fitting when compared to the Normal Distribution, and the Standard Student-t distribution.

To further validate the distribution, in this section I compare the theoretical moments from the skewed student distribution, that are computed based on the estimated parameters in table 2.4.1, to the empirical moments that are calculated from the standarized residuals \hat{z}_t of the Normal model in section 2.2. The comparison of the moments is presented in Table 2.4.3.

The theoretical moments that are used in the table are described in the appendix A.

Overall the results in the table show that the Skewed Student-t distribution moments match well with those in the residuals. Especially we see that the two parameters of the Skewed Student-t that explain skewness and kurtosis of are necessary to capture the sample skewness and kurtosis that are empirically observed in the Normal model.

Table 21:

Asset	Moment	Sample	Theoretical	Difference
S&P 500	Mean	-0.0011	-0.0000	-0.0011
	Variance	0.9941	1.0000	-0.0059
	Skewness	-0.5244	-0.3786	-0.1458
	Kurtosis	5.1071	5.7851	-0.6781
	Excess Kurtosis	2.1071	2.7851	-0.6781
Corporate Bonds	Mean	0.0063	0.0000	0.0063
	Variance	1.0014	1.0000	0.0014
	Skewness	-0.2465	-0.0976	-0.1489
	Kurtosis	4.4987	4.3702	0.1285
	Excess Kurtosis	1.4987	1.3702	0.1285

2.4.4 Volatility Time Series Comparison

The previous results have shown that even though the Skewed Student-t distribution tend to be the best fitting distribution for the GARCH innovations, the parameter estimates for the GARCH(1,1) model have remained quite stable across different distributional assumptions in the MLE. Which indicates that the estimated volatility dynamics are quite robust to the distributional assumption.

These observations are further confirmed by the figure 2.4.4 that plots the conditional volatilities for each of the three models.

The figure 2.4.4 shows through out the sample forecasted volatilities have behaved similarly across all three distributional assumptions.

To further analyze the differences between the different models Figure 2.4.4 plots all the pairwise differences in the conditional volatilities from the three models. For stocks we see that the normal model tends underestimate the volatility relative to the other two models, especially during high volatility periods. Furthermore, the standard Student-t and Skewed Student-t models produce very similar volatility estimates (notice the scale of the difference is very small), but the Skewed Student-t model produces systematically larger volatility estimates than the standard Student-t model.

For corporate bonds we see that the main differences arise during high volatility periods and Normal model actually tends to overestimate the volatility relative to the other two models.

3 Conclusion

In this report, I first analyzed the characteristics of typical financial timeseries, and then proceeded to model the volatility dynamics of corporate bond and stock excess log returns using GARCH(1,1) models with different innovation distributions.

From the comparisons between Normal, Student-t, and Skewed Student-t distributions for the standardized GARCH innovations, I find that the best describinig distribution is the Skewed Student-t distribution, which captures both the fat tails and skewness observed in the financial return series.

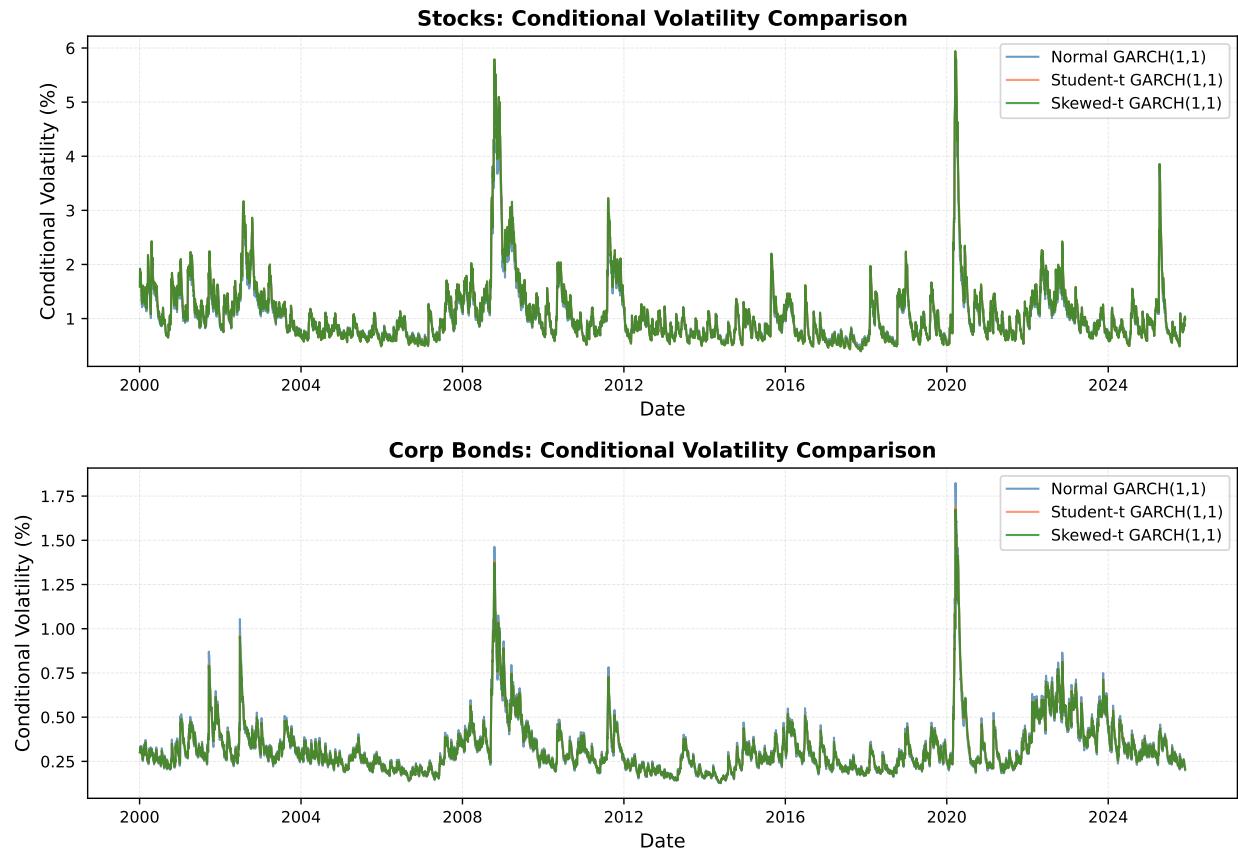


Figure 4: Comparison of Conditional Volatilities from GARCH(1,1) models estimated with Normal, Student-t, and Skewed Student-t innovations for Corporate Bond and Stock (S&P 500) excess log returns. The plots show the time series of conditional volatilities estimated from each model, highlighting the similarities and differences in volatility dynamics across different distributional assumptions.

Conditional Volatility Differences Between GARCH Specifications

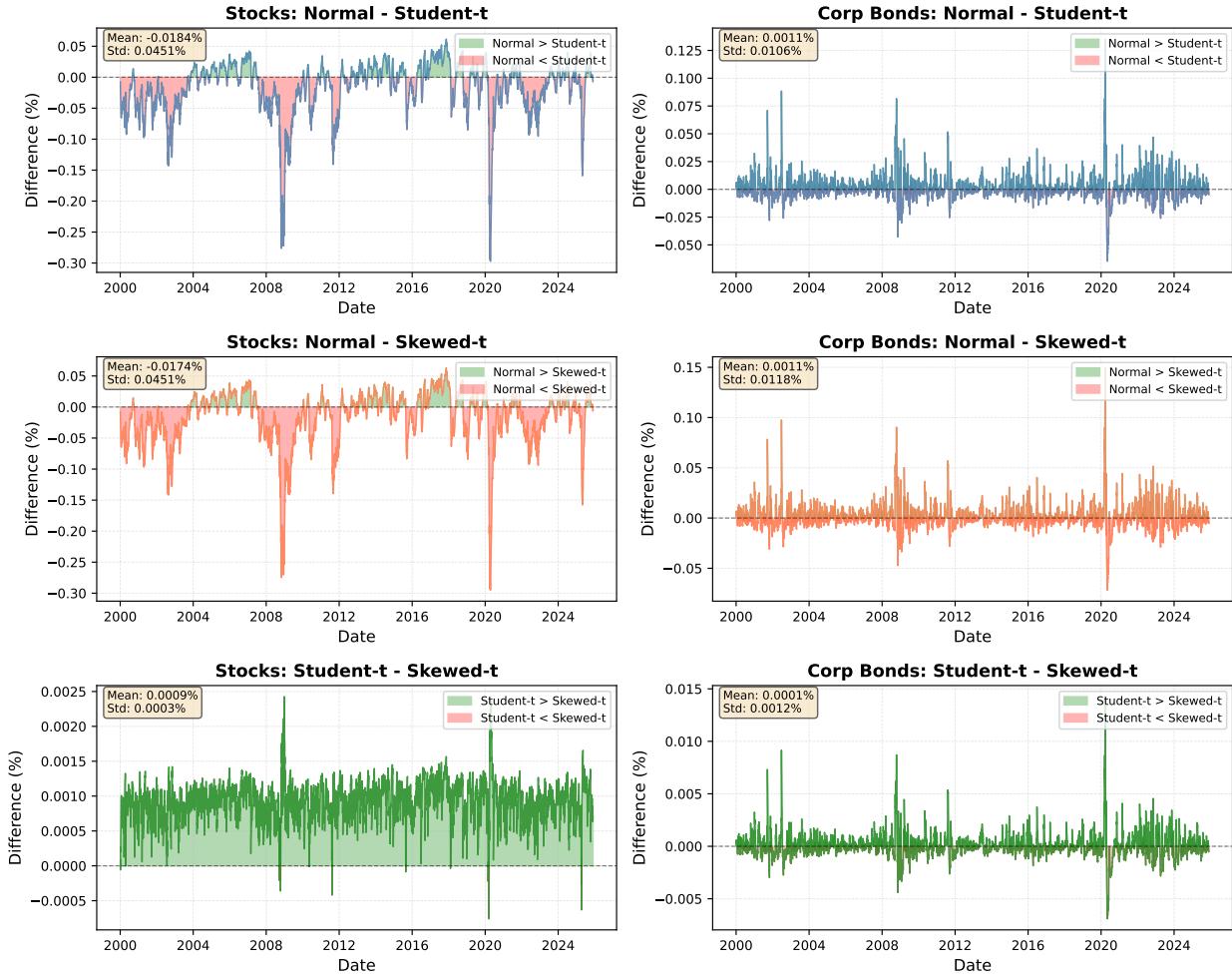


Figure 5: Difference in Conditional Volatilities between GARCH(1,1) models estimated with Normal, Student-t, and Skewed Student-t innovations for Corporate Bond and Stock (S&P 500) excess log returns. The plots show the time series of differences in conditional volatilities between each pair of models, highlighting the magnitude and patterns of differences in volatility estimates across different distributional assumptions.

A Skewed-t Distribution Moments

The following definitions follow Hansen [1994] definition for Skewed Student-t distribution. Let $Z \sim \text{SkewStudent}(\nu, \lambda)$ with $\nu > 2$, $\lambda \in (-1, 1)$, and define following constants and variables:

$$c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)} \Gamma\left(\frac{\nu}{2}\right)},$$

$$a = 4\lambda c \frac{\nu-2}{\nu-1}, \quad b = \sqrt{1 + 3\lambda^2 - a^2}.$$

$$Y := a + bZ.$$

For Y we have obtain

$$\mathbb{E}[Y] = a, \quad \mathbb{E}[Y^2] = 1 + 3\lambda^2,$$

$$\mathbb{E}[Y^3] = 16\lambda(1 + \lambda^2)c \frac{(\nu-2)^2}{(\nu-1)(\nu-3)} \quad (\text{for } \nu > 3),$$

and

$$\mathbb{E}[Y^4] = \frac{3(\nu-2)}{\nu-4} (1 + 10\lambda^2 + 5\lambda^4) \quad (\text{for } \nu > 4).$$

Lastly using the calculated moments for Y and then solving for the moments of Z we obtain the four following moments that are used in the report

$$m_1 := \mathbb{E}[Z] = 0, \quad m_2 := \mathbb{E}[Z^2] = 1.$$

$$m_3 := \mathbb{E}[Z^3] = \frac{\mathbb{E}[(Y-a)^3]}{b^3}$$

$$= \frac{\mathbb{E}[Y^3] - 3a\mathbb{E}[Y^2] + 2a^3}{b^3}$$

$$= \frac{\mathbb{E}[Y^3] - 3a(1 + 3\lambda^2) + 2a^3}{b^3} \quad (\text{for } \nu > 3).$$

$$m_4 := \mathbb{E}[Z^4] = \frac{\mathbb{E}[(Y-a)^4]}{b^4}$$

$$= \frac{\mathbb{E}[Y^4] - 4a\mathbb{E}[Y^3] + 6a^2\mathbb{E}[Y^2] - 3a^4}{b^4}$$

$$= \frac{\mathbb{E}[Y^4] - 4a\mathbb{E}[Y^3] + 6a^2(1 + 3\lambda^2) - 3a^4}{b^4} \quad (\text{for } \nu > 4).$$

B Python Libraries

References

Bruce E Hansen. Autoregressive conditional density estimation. *International Economic Review*, pages 705–730, 1994.