Analysis of GS1: 1-Year Treasury Constant Maturity Rate

Stefano Mauloni, C05503

Spring 2024

1 Introduction

In this report, we will perform an analysis of GS1: `1-Year Treasury Constant Maturity Rate (Percent)`. This is a series concerning the movement of the interest rate on constant-maturity-rate (1-year) US bonds.

The following is the plot of the raw data for the series.

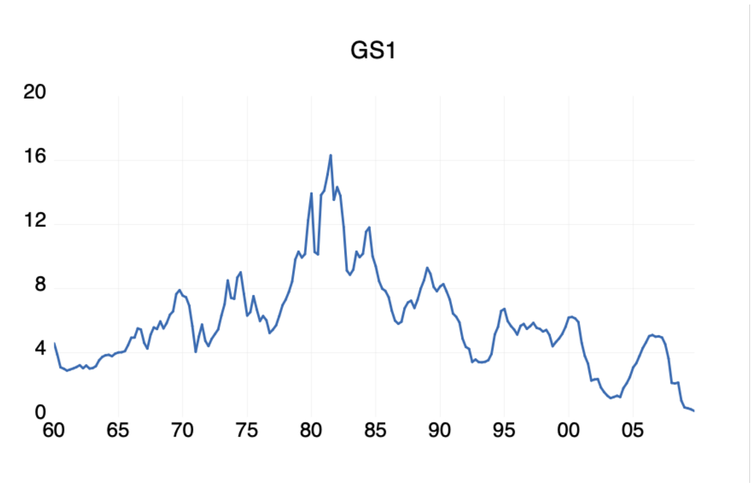
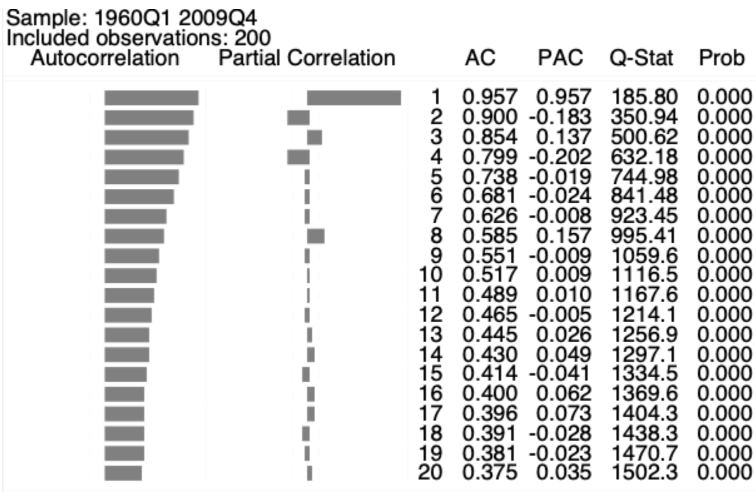


Figure 1: Plot of the raw data for GS1

Now we will plot the sample ACF and PACF to understand the structure of the series better.



Sample: 1960Q1 2009Q4
Included observations: 200

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.957	0.957	185.80	0.000
		2	0.900	-0.183	350.94	0.000
		3	0.854	0.137	500.62	0.000
		4	0.799	-0.202	632.18	0.000
		5	0.738	-0.019	744.98	0.000
		6	0.681	-0.024	841.48	0.000
		7	0.626	-0.008	923.45	0.000
		8	0.585	0.157	995.41	0.000
		9	0.551	-0.009	1059.6	0.000
		10	0.517	0.009	1116.5	0.000
		11	0.489	0.010	1167.6	0.000
		12	0.465	-0.005	1214.1	0.000
		13	0.445	0.026	1256.9	0.000
		14	0.430	0.049	1297.1	0.000
		15	0.414	-0.041	1334.5	0.000
		16	0.400	0.062	1369.6	0.000
		17	0.396	0.073	1404.3	0.000
		18	0.391	-0.028	1438.3	0.000
		19	0.381	-0.023	1470.7	0.000
		20	0.375	0.035	1502.3	0.000

Figure 2: Plot ACF and PACF

As we can see, there is a single major spike at lag 1 for the PACF, while the ACF is sustained for several lags, but always decaying. Also, no seasonality is evident from the two plots. This would suggest the use of an **AR(1)** model. However, we first need to check for stationarity.

2 Augmented Dickey-Fuller test

Now we will perform a test to check whether the series has a unit root, meaning that the data is non-stationary. If that is the case, we will apply some transformations to model the series (e.g. differencing).

By looking at 1, we can assume that there exists an intercept in the series. Thus, we perform the ADF test including it.

The test equation is:

$$Y_t = \alpha + \rho Y_{t-1} + \sum_{j=1}^{p-1} \zeta_j \Delta Y_{t-j} + \epsilon_t$$

Which derives from $Y_t = Y_{t-1} + u_t \implies Y_t = Y_{t-1} + \sum_{j=1}^{p-1} \zeta_j u_{t-j} + \epsilon_t \implies Y_t = Y_{t-1} + \sum_{j=1}^{p-1} \zeta_j \Delta Y_{t-j} + \epsilon_t$.
Note that the assumption is that ΔY_{t-j} represents u_t, a stationary `AR(p-1)` model.
We want to test $\mathrm{H}_0 : \{\rho = 1\}$ vs $\mathrm{H}_A : \{|\rho| \leq 1\}$. When $\alpha \neq 0$, $\rho = 1$, the t-statistic has a normal limit distribution.

Null Hypothesis: GS1 has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.142823	0.2283
Test critical values:	1% level	-3.463749	
	5% level	-2.876123	
	10% level	-2.574622	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GS1)
Method: Least Squares
Date: 06/21/24 Time: 23:12
Sample (adjusted): 1961Q1 2009Q4
Included observations: 196 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GS1(-1)	-0.038943	0.018174	-2.142823	0.0334
D(GS1(-1))	0.317928	0.069942	4.545600	0.0000
D(GS1(-2))	-0.263703	0.069923	-3.771346	0.0002
D(GS1(-3))	0.280871	0.069782	4.024963	0.0001
C	0.224097	0.120469	1.860208	0.0644

R-squared	0.166457	Mean dependent var	-0.013486
Adjusted R-squared	0.149001	S.D. dependent var	0.782386
S.E. of regression	0.721749	Akaike info criterion	2.210899
Sum squared resid	99.49590	Schwarz criterion	2.294525
Log likelihood	-211.6681	Hannan-Quinn criter.	2.244755
F-statistic	9.535583	Durbin-Watson stat	1.987125
Prob(F-statistic)	0.000000		

Figure 3: Report of the ADF test

Since the p-value is greater than 0.05, we retain the null hypothesis, as there isn't statistical evidence of the absence of a unit root.

Note that a similar result holds if we also include the hypothesis of a linear trend in the model, even if from 1 a linear trend does not seem to be present. For this reason, we proceed to take the first differences of the series and analyse them. We expect that the process will be stationary at this point.

3 First Difference

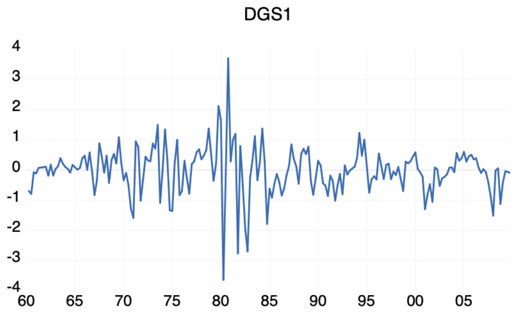
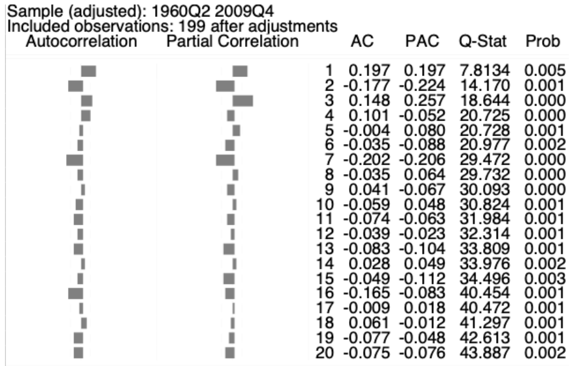


Figure 4: Data after first difference



Sample (adjusted): 1960Q2 2009Q4
Included observations: 199 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.197	0.197	7.8134	0.005
		2	-0.177	-0.224	14.170	0.001
		3	0.148	0.257	18.644	0.000
		4	0.101	-0.052	20.725	0.000
		5	-0.004	0.080	20.728	0.001
		6	-0.035	-0.088	20.977	0.002
		7	-0.202	-0.206	29.472	0.000
		8	-0.035	0.064	29.732	0.000
		9	0.041	-0.067	30.093	0.000
		10	-0.059	0.048	30.824	0.001
		11	-0.074	-0.063	31.984	0.001
		12	-0.039	-0.023	32.314	0.001
		13	-0.083	-0.104	33.809	0.001
		14	0.028	0.049	33.976	0.002
		15	-0.049	-0.112	34.496	0.003
		16	-0.165	-0.083	40.454	0.001
		17	-0.009	0.018	40.472	0.001
		18	0.061	-0.012	41.297	0.001
		19	-0.077	-0.048	42.613	0.001
		20	-0.075	-0.076	43.887	0.002

Figure 5: ACF, PACF of the first difference

We can now see that the data lay around a zero mean. However, we can spot some differences in volatility, especially for the period 1978-1985 in which it seems that interest rates oscillated more than usual.

We plot again the ACF and PACF to check now the possible model we can apply to the transformed series, to remove serial correlations in the residuals. Both the ACF and the PACF are pretty much sinusoidal. This suggests that an **ARMA** model could be the right choice.

Since we want a parsimonious model that captures the serial correlation in the data but at the same time remains slim, we fit an **ARMA(2,2)**, and next, we plot the residual ACF, PACF. Even with model selection, this model turns out to be one of the best among several ICs.

Dependent Variable: DGS1
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 06/21/24 Time: 23:18
Sample: 1960Q2 2009Q4
Included observations: 199
Convergence achieved after 69 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.021152	0.064886	-0.325987	0.7448
AR(1)	-1.022404	0.146878	-6.960920	0.0000
AR(2)	-0.511444	0.069306	-7.379538	0.0000
MA(1)	1.366501	0.163573	8.354061	0.0000
MA(2)	0.630057	0.124426	5.063700	0.0000
SIGMASQ	0.506110	0.029960	16.89267	0.0000

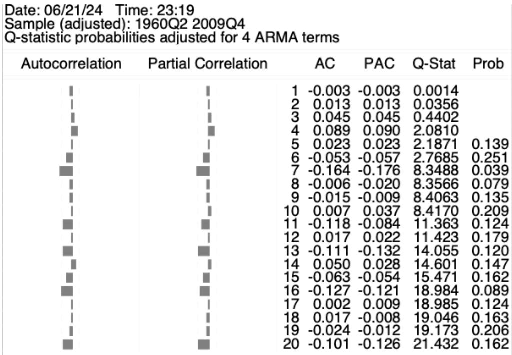
R-squared	0.163783	Mean dependent var	-0.021206
Adjusted R-squared	0.142119	S.D. dependent var	0.779932
S.E. of regression	0.722388	Akaike info criterion	2.219518
Sum squared resid	100.7160	Schwarz criterion	2.318814
Log likelihood	-214.8421	Hannan-Quinn criter.	2.259706
F-statistic	7.560243	Durbin-Watson stat	2.000978
Prob(F-statistic)	0.000002		

Inverted AR Roots	-.51-.50i	-.51+.50i
Inverted MA Roots	-.68+.40i	-.68-.40i

Figure 6: Estimation of an ARMA(2,2) model

As we can see from 6, all `AR` and `MA` terms are statistically significant to model the series. The R^2 is quite low, but we will explore this later. Now we focus on the analysis of the residuals.

3.1 Analysis of residuals of ARMA(2,2)



Date: 06/21/24 Time: 23:19
Sample (adjusted): 1960Q2 2009Q4
Q-statistic probabilities adjusted for 4 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.003	-0.003	0.0014	
		2	0.013	0.013	0.0356	
		3	0.045	0.045	0.4402	
		4	0.089	0.090	2.0810	
		5	0.023	0.023	2.1871	0.139
		6	-0.053	-0.057	2.7685	0.251
		7	-0.164	-0.176	8.3488	0.039
		8	-0.006	-0.020	8.3566	0.079
		9	-0.015	-0.009	8.4063	0.135
		10	0.007	0.037	8.4170	0.209
		11	-0.118	-0.084	11.363	0.124
		12	0.017	0.022	11.423	0.179
		13	-0.111	-0.132	14.055	0.120
		14	0.050	0.028	14.601	0.147
		15	-0.063	-0.054	15.471	0.162
		16	-0.127	-0.121	18.984	0.089
		17	0.002	0.009	18.985	0.124
		18	0.017	-0.008	19.046	0.163
		19	-0.024	-0.012	19.173	0.206
		20	-0.101	-0.126	21.432	0.162

Figure 7: ACF, PACF of residuals, ARMA(2,2)

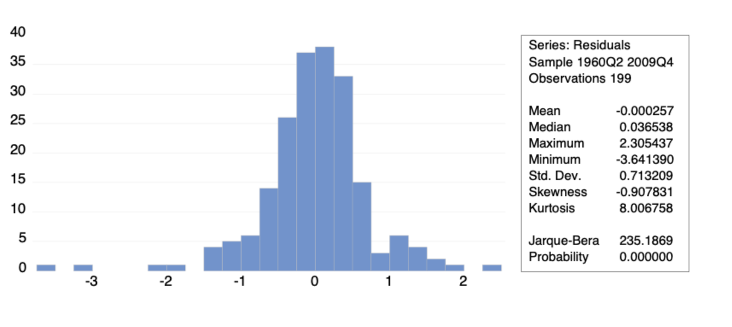


Figure 8: Histogram of residuals, ARMA(2,2)

We successfully removed a major part of the serial correlation in the series. (Note: repeating the same process with an ARMA(4,4) brings a slightly better result, but we believe that a more compact model is preferable in our case, so as to not "overfit" the data).

Then, the histogram above shows the distribution of the residuals. Skewness and kurtosis are, respectively, -0.9 and 8. The Jarque-Bera test for normality (the null hypothesis is a joint hypothesis of the skewness being zero and the excess kurtosis being zero, thus kurtosis = 3) is strongly rejected.

As a side note, the sample size here ($n = 199$) is relatively "small" for the test, making us reject the null hypothesis more often, increasing the type 1 error. MATLAB, for instance, runs a Montecarlo-simulation variant for the test if $n < 2000$. However, in our case, we can still trust the result of the test, since the deviation of the kurtosis from a normal distribution is large.

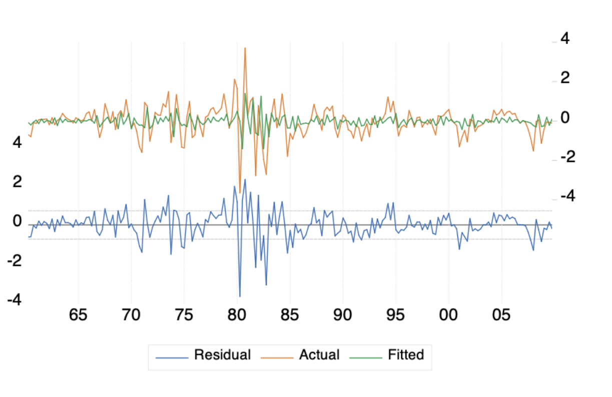


Figure 9: Graph of actual, fitted, residual, ARMA(2,2)

From the last graph, we can see that the model may not capture the underlying patterns of the time series effectively. Additionally, the residual plot shows significant volatility and deviations from zero, indicating potential issues with the model specification. It may be beneficial at the end to explore alternative models to improve the fit.

4 Forecasting

Now we want to make one-step-ahead forecasts for the period `2010Q1-2019Q4`. It means that for every t, we will assume that all times up to $t-1$ are known. We will use our **ARMA(2,2)** model and we will compare the results with a model that has as a unique parameter the intercept (benchmark model).

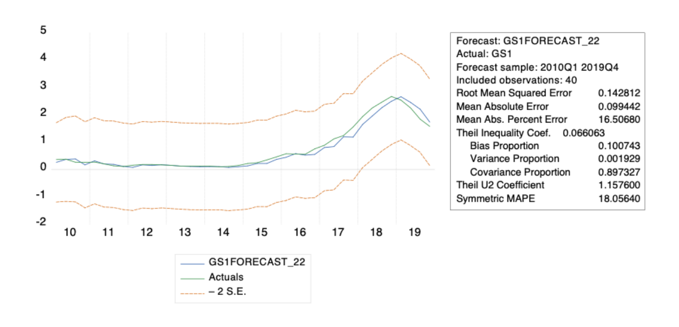


Figure 10: forecasts, ARMA(2,2)

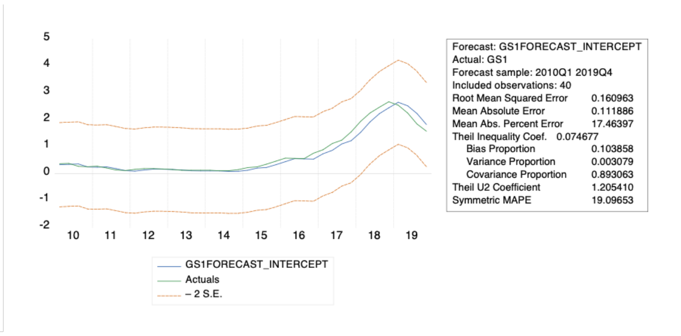


Figure 11: forecasts, benchmark model

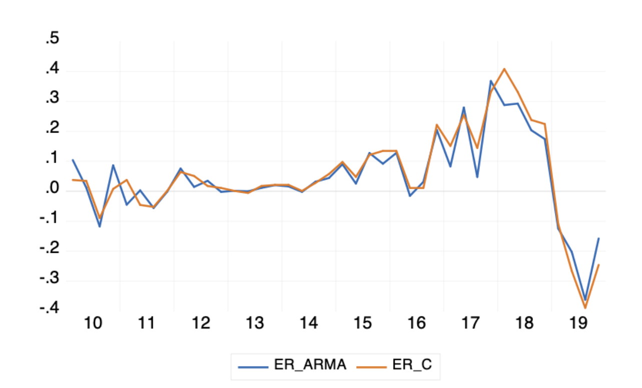


Figure 12: errors

We see that the benchmark model is basically a shift in the actual value at $t-1$ since the value is forecasted as the last available observation. Instead, the **ARMA(2,2)** model behaves differently. The RMSFE is approximately 0.14 for the **ARMA(2,2)** model and 0.16 for the intercept-only one. We conclude that our model is slightly better than the benchmark, but it explains very little variance in the data and the forecasts are only moderately more accurate than following a naive approach.

5 An improved model (Extra)

Above we saw that there is a period (1978-1985) in which the volatility seems to be higher. This gives us a basis to try to add ARCH effects in our analysis. First, we test for heteroskedasticity in the residuals for the **ARMA(2,2)** model we estimated before.

Heteroskedasticity Test: ARCH

F-statistic	3.717119	Prob. F(1,196)	0.0553
Obs*R-squared	3.685160	Prob. Chi-Square(1)	0.0549

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 06/22/24 Time: 15:15
Sample (adjusted): 1960Q3 2009Q4
Included observations: 198 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.437325	0.101630	4.303123	0.0000
RESID^2(-1)	0.136465	0.070781	1.927983	0.0553

Figure 13: Heteroskedasticity test on residuals, 1 lag

Interestingly, here the p-value is slightly above the 0.05 threshold. However, if we include 2 lags in the test, the result changes a lot.

Heteroskedasticity Test: ARCH

F-statistic	28.44952	Prob. F(2,194)	0.0000
Obs*R-squared	44.67578	Prob. Chi-Square(2)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 06/22/24 Time: 15:23
Sample (adjusted): 1960Q4 2009Q4
Included observations: 197 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.234436	0.095151	2.463829	0.0146
RESID^2(-1)	0.073829	0.063741	1.158273	0.2482
RESID^2(-2)	0.460812	0.063761	7.227196	0.0000

Figure 14: Heteroskedasticity test on residuals, 2 lags

Now the null hypothesis of homoskedasticity is strongly rejected, meaning that heteroskedasticity is spotted with 2 lag regressors and an intercept.

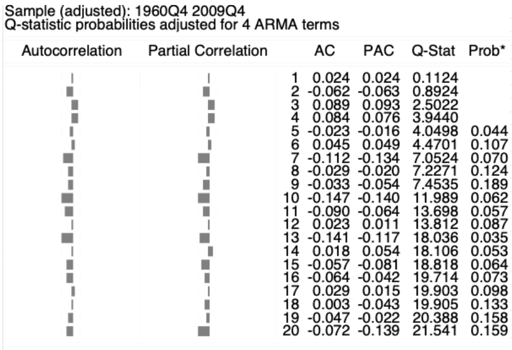
The reason behind this behavior can be explained by the plot of the ACF/PACF of squared residuals, which allows to better understand where the ARCH/GARCH effects are present.

Sample (adjusted): 1960Q2 2009Q4
Included observations: 199 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.136	0.136	3.7574	0.053
		2	0.470	0.460	48.655	0.000
		3	0.074	-0.031	49.765	0.000
		4	0.305	0.110	68.863	0.000
		5	0.005	-0.064	68.869	0.000
		6	0.342	0.226	93.131	0.000
		7	0.065	0.049	94.007	0.000
		8	0.328	0.109	116.48	0.000
		9	0.032	-0.052	116.70	0.000
		10	0.393	0.225	149.42	0.000
		11	0.010	-0.047	149.44	0.000
		12	0.100	-0.280	151.56	0.000
		13	-0.015	0.013	151.62	0.000
		14	0.009	-0.116	151.63	0.000
		15	-0.040	0.030	151.97	0.000
		16	0.069	-0.035	153.02	0.000
		17	-0.058	-0.093	153.75	0.000
		18	0.072	0.057	154.89	0.000
		19	-0.041	0.031	155.27	0.000
		20	0.025	-0.045	155.41	0.000

Figure 15: correlogram of squared residual

We notice that spikes are present at several indices for the ACF and at lag 2 for the PACF. To not make a too large model, we opt for an **ARCH(2)**.



Sample (adjusted): 1960Q4 2009Q4
Q-statistic probabilities adjusted for 4 ARMA terms

	AC	PAC	Q-Stat	Prob*
1	0.024	0.024	0.1124	
2	-0.062	-0.063	0.8924	
3	0.089	0.093	2.5022	
4	0.084	0.076	3.9440	
5	-0.023	-0.016	4.0498	0.044
6	0.045	0.049	4.4701	0.107
7	-0.112	-0.134	7.0524	0.070
8	-0.029	-0.020	7.2271	0.124
9	-0.033	-0.054	7.4535	0.189
10	-0.147	-0.140	11.989	0.062
11	-0.090	-0.064	13.698	0.057
12	0.023	0.011	13.812	0.087
13	-0.141	-0.117	18.036	0.035
14	0.018	0.054	18.106	0.053
15	-0.057	-0.081	18.818	0.064
16	-0.064	-0.042	19.714	0.073
17	0.029	0.015	19.903	0.098
18	0.003	-0.043	19.905	0.133
19	-0.047	-0.022	20.388	0.158
20	-0.072	-0.139	21.541	0.159

Figure 16: Plot ACF and PACF, ARCH

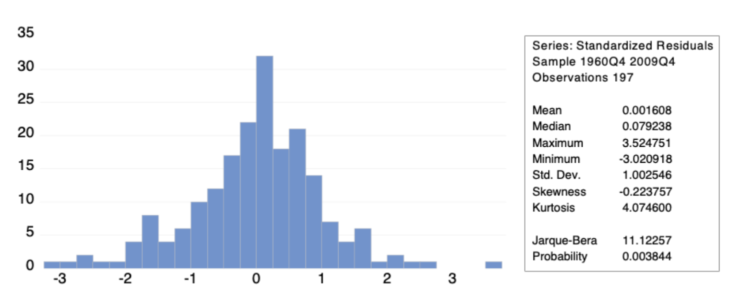


Figure 17: Histogram of residuals, ARCH

We can see that the residuals are slightly worse than the model without the ARCH component, but still, serial correlations are practically removed.

In addition, even though the JB test still rejects the null hypothesis, the kurtosis halved from before, and the test statistic takes a much "better" result (in the sense of retaining the null as our desire).

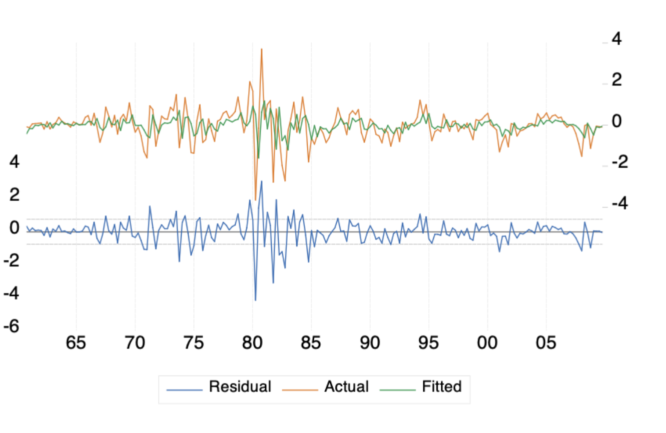
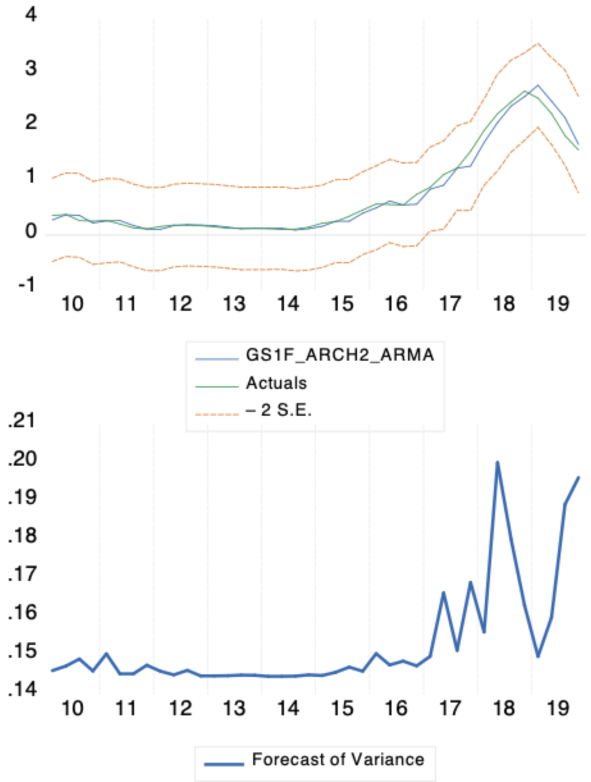


Figure 18: Graph of actual, fitted, residual, ARCH

The residuals graph is very similar to the one of the previous model. However, the adjusted R^2 is smaller than before, thus we presume that the fitted values are worsened.

5.1 Forecasting again

Now we produce forecasts to see if having considered ARCH effects made our last model more accurate at predicting one step ahead.



Forecast: GS1F_ARCH2_ARMA	
Actual: GS1	
Forecast sample: 2010Q1 2019Q4	
Included observations: 40	
Root Mean Squared Error	0.114823
Mean Absolute Error	0.076482
Mean Abs. Percent Error	12.76647
Theil Inequality Coef.	0.052259
Bias Proportion	0.011523
Variance Proportion	0.010646
Covariance Proportion	0.977831
Theil U2 Coefficient	0.820052
Symmetric MAPE	12.96459

As we can see, the error improved from our initial model. We now have a value of roughly **0.11**, better than the ARMA(2,2) and the benchmark forecasts' results.

Finally, The last plot is the plot of the forecasted conditional variance in the model. We see that it oscillates back and forth in the period 2018-2019. This can be because we used a relatively "simple" ARCH model with 2 lags that are always equally weighted. Instead, one could have used an exponential smoothing technique (even a GARCH(1,1) model, which brings similar results to some EMA models) to "smooth out" some excessively large variations, but this is out of the scope of this report.

To sum up, the **ARCH(2)** model explains slightly less variance in the data but has overall more predictive power than the sole **ARMA(2,2)** model.