Gold
Univariate and Multivariate Analysis of Risk

Agate Matteo Cardarello Alessio

June 10, 2021

Introduction

Gold has always been a resource of value. Even in the antiquity people believed in the value of gold, it has been for a long time the most important medium of exchange and it is the key element of the first part of the analysis. In this part we model and forecast the conditional volatility of gold and estimate some risk measures such as Value at Risk and Expected Shortfall.

Firstly, there is the need to understand better the asset in consideration: this is done through the interpretation of the plot of the prices, log-returns, QQ-plot and histogram of the log-returns. Then the stylized facts of the returns are checked: non-normality, dependence of squared returns, fat tails and skewness. Once described the series, different GARCH models, either symmetric and asymmetric, are implemented and checked. Gold volatility can be influenced by macroeconomic variables which can be checked with the MIDAS models, that permit to add variables with different frequency with respect to the returns. The variable that we add is the first difference of the Effective Federal Funds Rate. Once checked the significance of the parameters, the parametric VaRs are calculated. Other methodologies are applied to estimate the Value at Risk, namely the semi-parametric and the non-parametric methods. The former directly estimate the quantile of the regression without assumptions, the latter use an historical simulation along a time window, without assumption on the distribution but with the assumption of independence and identical distribution. Then the VaR estimates are checked with the backtesting procedure, that allow us to understand if the estimates are good, so if the model does not produce too many (or to low) exceptions and if produced in an independent way. Finally, the MCS procedure produce a classification of the models, based on the asymmetric loss function. The second risk measure calculated is the Expected Shortfall. This is calculated starting from the GARCH models and evaluated through backtesting procedure.

The second part of the analysis is based on the usage of copulas. We introduce three stocks (Apple, Walmart and Tesla) and a cryptocurrency (Bitcoin). What is the effect of addition of Bitcoin and gold in a portfolio? Adding the two assets in the portfolio jointly help in the diversification and in the improvement of the performances? These are the questions that we try to answer, throughout the implementation of a very simple Markowitz model. In fact, we have found the best portfolio in the case of the three assets, and we have replicated it with the gold and the Bitcoin. In order to do that, we have used a covariance matrix obtained from the estimate of the correlation of the fitting of the copula in the data. So first we have found the best copula for the portfolio and then we have used it to estimate the frontier.

Gold

The plot of the price of gold (Figure 1.a) is very instructive and we can see the nature of this particular type of asset. In fact, during one the most tremendous period of the recent economy,the price of gold rise continuously. This is because it is a safe haven. When things go bad people get scared and want something that has always been valuable. This can be seen also from the plot of the returns (Figure 1.b), where there are particularly high value in that periods. Once arrived at the peak in 2011, gold has never come back to it until 2020, when the uncertainty derived from the global pandemic has troubled the market.

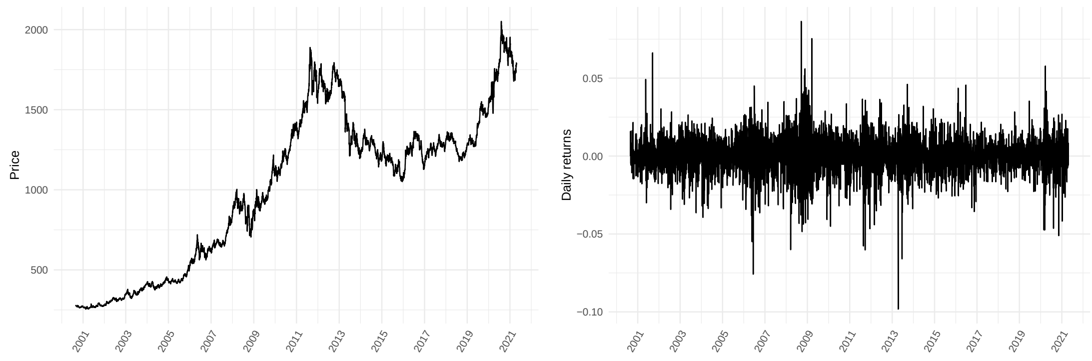

Figure 1: Price and returns for Gold

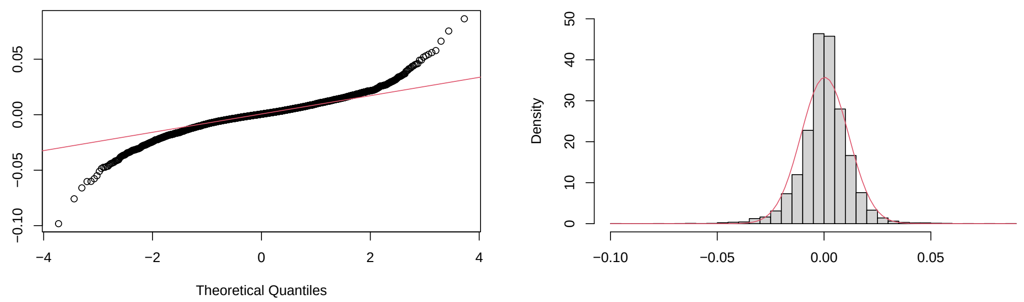

Figure 2: QQ-Plot and Histogram of Gold log returns

The main features of gold returns are summarized in Table 1. The first four items are important statistics to understand the basic behaviour of the series, in particular the mean, the standard deviation, the skewness and the kurtosis. In Table 1 are also included the results of some tests: the Ljung Box and Box Pierce test to assess if the data exhibits serial autocorrelation and the Jarque-Bera test to assess if the data are normally distributed. Ljung Box and Box Pierce are accepted when the returns are tested. If these tests are applied on the squared returns tests are rejected: in this case there is autocorrelation. Jarque Bera test is rejected: the returns are not normally distributed. This feature can be seen also from the QQ-plot. The QQ-plot in Figure 2.a compares the theoretical quantile of a Normal distribution and the sample quantile. As we expected, since it is one of the stylized fact of financial returns, the tail of the distribution are far from the one of a Normal distribution. The kurtosis of the distribution is higher than three, so it is leptokurtic. The series shows also negative skewness: the left tail is longer with respect to the right one.

Table 1: Descriptive analysis for Gold log returns

	Mean	SD	Skew	Kurt	L-B	B-P	L-B^2	B-P^2	J-B
Gold	0.0004	0.0112	-0.271	5.472	1.28 *	1.28 *	293.2	292.9	6491

[a] * denotes fail to reject the null hypothesis at 95% confidence level

Value at Risk Estimation for Gold

The focus of this part of the analysis is the modeling and forecasting the Value at Risk for a risky position in gold. Three different approaches are used: the parametric approach, the semi parametric approach and the non parametric one.

Parametric approach

GARCH models In order to estimate the distribution of the returns the first approach used is the parametric one. Different GARCH models are included. In the description of the data were shown some stylized facts of the returns: the heteroskedasticity and the autocorrelation in the squared returns. This means that the conditional volatility changes through time and it is possible to model and forecast it using its past values.

Standard GARCH models can model the volatility clustering phenomenon but not the so called “leverage effect”, that is the effect that returns tend to be negatively dependent respect to changes in conditional volatility. eGARCH and gjrGARCH model allow to capture the leverage effect. These models are included to understand if there is leverage effect in the behaviour of gold volatility. The csGARCH break down the conditional variance into a permanent and transitory component to investigate the long and short-run movements of volatility. Assumptions about the distribution of the returns are a prerequisite to fit all these models. Normal and student-t distribution assumption (and their skewed version) has been implemented for all GARCH models.

All models, except for the csGARCH model, show the parameter β very close to 1 which indicates a strong persistence of the shocks on the volatility. γ coefficient of eGARCH model shows the possible leverage effect. In particular, if it is negative there is presence of this effect, as larger as the coefficient becomes larger in absolute value. If γ is positive there is the inverse effect. There are many doubts about the presence of leverage effect for gold. Indeed, the γ coefficients for eGARCH are significant and positive: positive shocks increase the volatility by more than negative shocks. This effect could be explained by the fact that the gold is a safe heaven. For this reason, investors could interpret the rise in the gold price as a signal for future uncertainty in the market.For what concern the csGARCH model, ρ is very close to 1. This parameter is the coefficient of the long run component of the variance. It suggests a higher impact of the long run component with respect to the short run one, which coefficient is ϕ.

GARCH MIDAS and DAGM models In the analysis of the volatility of a financial asset could be interesting to include a macroeconomic variable. We choose the Effective Federal Funds Rate. This rate is the interest rate at which a bank with excess cash will lend to another banks that needs liquidity. The effective rate is determined by the two banks but is influenced by the Federal Reserve through open market operations to reach the federal funds rate target. Many investors believe that the federal funds rate is negatively correlated to the gold price. Low interest rates are a sign of an expansive monetary policy which is a bullish signal in the gold market. However, this relationship does not always hold. In Figure 3 it is shown the behavior of the Effective Federal Funds between the 2003/01/01 and the 2021/03/01 and its first difference.

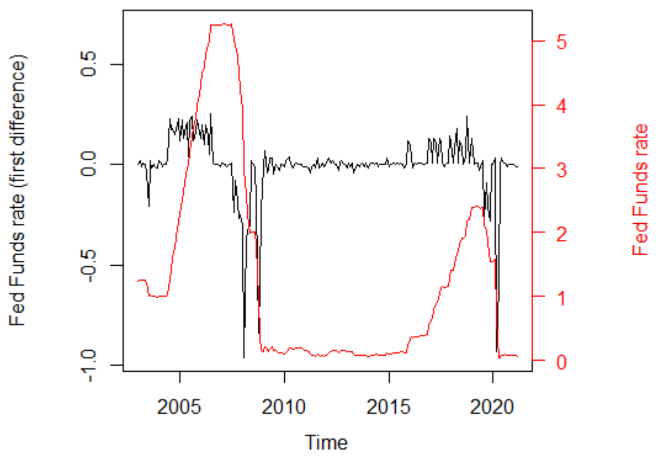

Figure 3: Effective Federal Funds rate and its first difference

From Figure 3 it is clear that starting from 2008 there was an important decrease of the interest rates due to the financial crisis. Many countries (such as USA) try to decrease the interest rates to face the crisis. From 2009 until 2015 the Effective Federal Funds Rate was very close to 0. After 2015 there was an increase which stops in 2020 after the spread of Coronavirus pandemic.

GARCH-MIDAS and DAGM models allow us to use a variable which is observed with different frequency with respect to the returns. So, the Effective Federal Funds Rate observed monthly are implemented and their first difference is used as MIDAS variable. The main difference between GARCH-MIDAS model and DAGM model is that the latter consider the possibility of asymmetric effect of the returns and the MIDAS variable on the conditional volatility. As for the GARCH models, a returns distribution assumption has been made: normal, student-t distribution and their skewed version.

Semi parametric approach

In the estimation of the Value at Risk there is the possibility of using another type of models: the semiparametric models. CaViar are examples of this class of models. Koenker and Bassett (1978) introduced the quantile regression, that directly estimate a certain quantile of the distribution, through the minimization of the quantile loss function. Engle and Manganelli (2004) proposed the CaViar models which does not require any assumption about the errors distribution. Two types of CAViaR are used: CAViaR-Symmetric Absolute Value and CAViaR-Asymmetric Slope. The main difference between the two is that the latter consider the possibility of asymmetric effect on the quantile of the distribution depending on the sign of the past value of the returns, meanwhile the first take the absolute value of the past return.

Non parametric approach

Another possible approach is the Historical Simulation. This is a non parametric approach because it calculates the VaR as the sample quantile over a moving window of data. This approach does not require a distributional assumption for the errors. The problem is that the sample quantile is a consistent and asymptotically normal estimator under the i.i.d. assumption of the returns, which is a strong assumption. For an accurate estimation of the VaR a long window is needed. On the other hand, a long estimation window increases the probability of structural breaks. In this analysis three possible windows are implemented: 50, 100 and 200 observations.

Backtesting procedure

In order to evaluate the accuracy of VaR prediction of each model, different tests have been used. These tests, called usually backtest, are based on comparison between the VaR predicted and the results of the stock market. For each model three different tests have been performed:

- The Unconditional Coverage (UC) test of Kupiec (1995): also called proportion of failure test, performs a likelihood ratio test. It verifies the null hypothesis that the proportion of the exception respect to the VaR predicted and the proportion of the theoretical exception are equal.
- The Conditional Coverage (CC) test of Christoffersen (1998), that also performs a likelihood ratio, and adds to the UC test the null hypothesis of serial independence between the exceptions.
- The Actual over Expected (AE) exceedance ratio, that is the ratio between the actual number of exceptions of the returns respect to the estimated VaR and the theoretical exceptions. The closer the AE ratio is to 1, the more precise is the model; if the ratio is less than 1 the model overestimates the risk, while if it is greater than 1 the model underestimates the risk.

Table 2: Backtesting results

Model	95%			99%		
	LR_{uc}	LR_{cc}	AE	LR_{uc}	LR_{cc}	AE
GARCHn	0.161 *	0.2458 *	1.0016	12.3393	15.449	0.9933
GARCHt	1.5855 *	1.6859 *	0.9949	0.7296 *	3.815 *	0.9985
GARCHsn	0.161 *	0.2458 *	1.0016	11.3227	14.595	0.9936
GARCHst	1.2183 *	1.278 *	0.9955	0.7296 *	3.815 *	0.9985
gjrGARCHn	0.161 *	0.2458 *	1.0016	11.3227	14.595	0.9936
gjrGARCHt	1.9994 *	2.151 *	0.9943	0.7296 *	3.815 *	0.9985
gjrGARCHsn	0.161 *	0.1692 *	1.0016	10.3438	13.7842	0.9939
gjrGARCHst	1.7866 *	1.9114 *	0.9946	0.7296 *	3.815 *	0.9985
eGARCHn	0.9406 *	1.0627 *	1.0038	13.3928	16.3456	0.993
eGARCHt	3.5172 *	4.422 *	0.9923	1.0412 *	7.5447	0.9982
eGARCHsn	0.9406 *	1.0627 *	1.0038	13.3928	16.3456	0.993
eGARCHs	1.2183 *	1.5752 *	0.9955	1.0412 *	7.5447	0.9982
csGARCHn	0.3169 *	0.3649 *	1.0022	15.6083	20.9021	0.9923
csGARCHt	2.2238 *	2.4048 *	0.9939	1.4048 *	4.1602 *	0.9979
csGARCHsn	0.4147 *	0.4482 *	1.0026	13.3928	19.1369	0.993
csGARCHst	1.396 *	1.4748 *	0.9952	0.7296 *	3.815 *	0.9985
CaViaR-SAV	0.0064 *	0.0231 *	0.9997	0 *	0.9047 *	1
CaViaR-AS	0.0254 *	0.0529 *	0.9994	0 *	0.9047 *	1
HS100	0.158 *	0.2333 *	0.9984	1.4048 *	2.3867 *	0.9979
HS50	11.9371	12.986	0.9856	23.0792	25.6116	0.9905
HS200	0.3085 *	0.3086 *	0.9978	0.0303 *	0.7384 *	0.9997
MIDASn	0.1029 *	0.2098 *	1.0013	16.769	21.8473	0.992
MIDASt	43.9467	44.8922	1.0242	17.8563	22.3642	1.0064
MIDASsn	1.0523 *	1.3679 *	0.9959	39.6894	40.5361	0.9871
MIDASst	36.6585	37.2045	1.0223	12.4453	16.0713	1.0055
DAGMn	0.1029 *	0.6272 *	1.0013	19.1934	23.8592	0.9914
DAGMt	57.9713	59.8628	1.0274	19.9783	24.8376	1.0067
DAGMsn	0 *	0.7461 *	1	17.9642	20.3407	0.9917
DAGMst	59.5114	61.5185	1.0278	17.8563	22.3642	1.0064

[a] * denotes the fail to reject the null hypothesis at 95% confidence level

Table 2 shows the results of the backtesting procedure. For what concern the VaR at 95% confidence level few estimations do not pass the test. These are the estimations obtained starting from the DAGM with student-t and skew-student-t distribution model, the MIDAS with student-t and skew-student-t distribution and with the historical simulation with a window of length 50. From these results it seems that if the Federal Funds Rate are included in the model, the student-t assumption for the distribution of the returns is not appropriate. The backtesting results for the Value at Risk at 99% confidence level are different respect to the ones at 95%. For what concern the parametric model the student-t and the skew-student-t assumption are the best choices. Indeed, all the GARCH models which assumes normal and skew-normal distribution do not pass the tests.

The fatter tails of the student-t distribution are more able to produce VaR which can capture more extreme events. The MIDAS models and DAGM models reject all the tests, so the changes in the interest rates do not help to predict the more extreme returns in gold. The semiparametric and the non parametric models pass all tests, with the exception of the historical simulation model with window of length 50. Models which do not pass the tests are not included in the Model Confidence Set procedure which is performed in the next paragraph.

Model Confidence Set

The model confidence set procedure has the aim to find the best model among all considered. The process is based on the EPA, i.e. equal prediction ability, and it means that the final models that pass the procedure have the same ability to predict VaR at a certain confidence level, given by an α. In this work α is set equal to 0.95.

Table 3: MCS results

95%		99%	
Model	Rank	Model	Rank
eGARCH-std	7	HS200	1
eGARCH-sstd	5	-	-
SAV-CAVIAR	4	-	-
AS-CAVIAR	3	-	-
HS100	1	-	-
DAGM-N	2	-	-
DAGM-SN	6	-	-

Table 3 shows the result of the MCS procedure performed both for VaR at 95% and 99% confidence level. For the MCS procedure, the best approach is the non parametric one. The main problem of the historical simulation is that it assumes that the returns are independently identically distributed; but in reality this is not a good assumption, because the empirical evidences shows that financial returns are not i.i.d. Also the semi-parametric one performs well. For what concern eGARCH model with student-t and skew-student-t the inclusion of the asymmetric effect (positive shocks have more impact on the conditional volatility respect to the negative one, as seen in the analysis of the coefficients) is a good way to improve VaR estimation. The other two models which pass the MCS procedure for VaR at 95% confidence level are the DAGM-N and the DAGM-SN: using the interest rates and the asymmetric effects lead to a good VaR estimation.

Expected Shortfall Estimation for Gold

The Expected Shortfall, or Conditional VaR, is a risk measure sensitive to the shape of the tail of the distribution. It is the Expected Loss given that the Value at Risk is exceeded: it is calculated averaging all the possible losses after the quantile representing the Value at Risk. In this analysis we use the parametric models seen before to estimate the conditional variance. From that the conditional distribution of the returns and the quantile of interest are obtained. So, given all these information we can calculate the Expected Shortfall.

Backtesting procedure

After have checked the martingale difference property, that is respected by all the models, a more advanced procedure is proposed. To perform this backtesting procedure, firstly we check the VaR backtesting for the models. Looking at the Table 2 we can see the goodness of VaR estimated and only those that pass all tests are used to estimate the ES. Then, we backtest the ES estimated. We want that the estimated loss distribution is equal to the theoretical one:

$$H_0 : ES_\alpha^{FL} = ES_\alpha^{\hat{FL}}, VaR_\alpha^{FL} = VaR_\alpha^{\hat{FL}}$$

$$H_1 : ES_\alpha^{FL} \geq ES_\alpha^{\hat{FL}}, VaR_\alpha^{FL} \geq VaR_\alpha^{\hat{FL}}$$

In order to verify if H_0 is accepted, the test statistic $Z_1(L)$ must be calculated:

$$Z_1(L) = \frac{\sum_{h=1}^{m}(L_{t+h} - ES_{\alpha}^{t,h}(L))}{1(\alpha)} + 1$$

Under H_0 the test statistic value is expected to be equal to zero. The problem is that the distribution of $Z_1(L)$ is non standard. To perform the backtest we must simulate M paths for the loss, starting from a distribution $\hat{F}_L$. For each path is computed the test statistic $Z_1^i(L)$. After that, the proportion of acceptance $p$of $(Z_1^i < Z_1)$ is computed and compared with the confidence level ϕ: if $p > \phi$ we fail to reject H_0.

Table 4: Backtesting ES results

95%		*99%*	
Model	Null Hypothesis	Model	Null Hypothesis
sGARCH-norm	Rejected	sGARCH-norm	Not pass VaR backtest
sGARCH-std	Accepted	sGARCH-std	Accepted
sGARCH-snorm	Rejected	sGARCH-snorm	Not pass VaR backtest
sGARCH-sstd	Accepted	sGARCH-sstd	Accepted
gjrGARCH-norm	Rejected	gjrGARCH-norm	Not pass VaR backtest
gjrGARCH-std	Accepted	gjrGARCH-std	Accepted
gjrGARCH-snorm	Rejected	gjrGARCH-snorm	Not pass VaR backtest
gjrGARCH-sstd	Accepted	gjrGARCH-sstd	Accepted
eGARCH-norm	Rejected	eGARCH-norm	Not pass VaR backtest
eGARCH-std	Accepted	eGARCH-std	Accepted
eGARCH-snorm	Rejected	eGARCH-snorm	Not pass VaR backtest
eGARCH-sstd	Accepted	eGARCH-sstd	Accepted
csGARCH-norm	Rejected	csGARCH-norm	Not pass VaR backtest
csGARCH-std	Accepted	csGARCH-std	Accepted
csGARCH-snorm	Rejected	csGARCH-snorm	Not pass VaR backtest
csGARCH-sstd	Accepted	csGARCH-sstd	Accepted

Table 4 shows that all the models for ES at 95% confidence level which use the normal assumption do not pass the backtesting procedure. Instead, all models which use the student-t assumption pass the backtesting. The same thing is for the ES at 99% confidence level. Probably, the fatter tails of the student-t distribution are more suited to model extremes events.

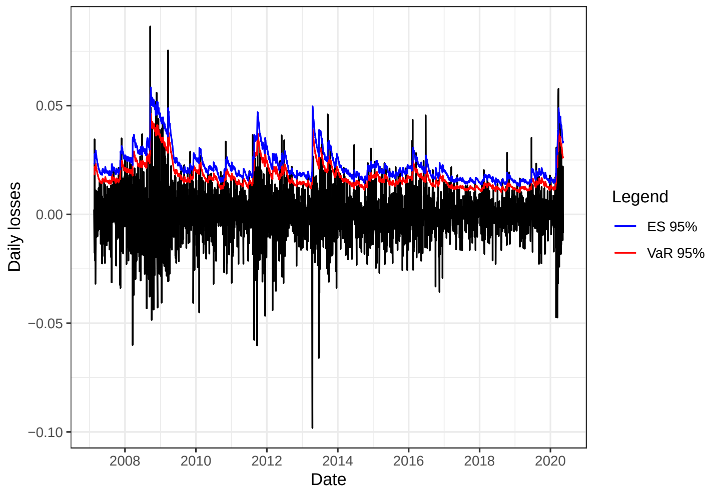

Figure 4: VaR and ES at 0.95 with CaViar-AS

Copula

Until now we have investigated the VaR and the Expected Shortfall in a univariate framework. In this section the two measures of risk are calculated in a multivariate framework, keeping into consideration the dependence structure between the different risk factors changes. Then, we will use the copula to inspect the effect of gold and Bitcoin in a portfolio.

Series Analysis

Table 5: Descriptive analysis for log returns

Series	Mean	SD	Skew	Kurt	L-B	B-P	L-B^2	B-P^2	J-B
Apple	0.0007	0.0290	-1.749	6.455	2.65	2.65	46.4*	46.38*	926503.7*
Walmart	0.0007	0.0187	0.021	5.984	28.21*	28.2*	1258.19*	1257.73*	18357.61*
Tesla	0.0018	0.0355	0.010	5.865	1.76	1.76	212.19*	211.87*	3948.82*
Bitcoin	0.0018	0.0395	-0.843	11.882	2.49	2.48	64.27*	64.16*	14701.73*

[a] * denotes the rejection of the null hypothesis at 95% confidence level

Bitcoin Everybody talks about this cryptocurrency. From Musk to PayPal, from the new paradigm of a new era money to the enormous environmental issue. Bitcoin was born in 2009 with the paper of Satoshi Nakamoto which explained the functioning of this cryptocurrency. In 2018 there was a primal fever that ended with the burst of the bubble. After this, the market for this asset stayed calm for a few years, when in 2020 a new crypto-mania spread of all over the world. The measure of the skewness and the kurtosis reveal the features of this asset, that are very high tails and high degree of asymmetry, as we can see from Table 5 where are summarized the main statistics.

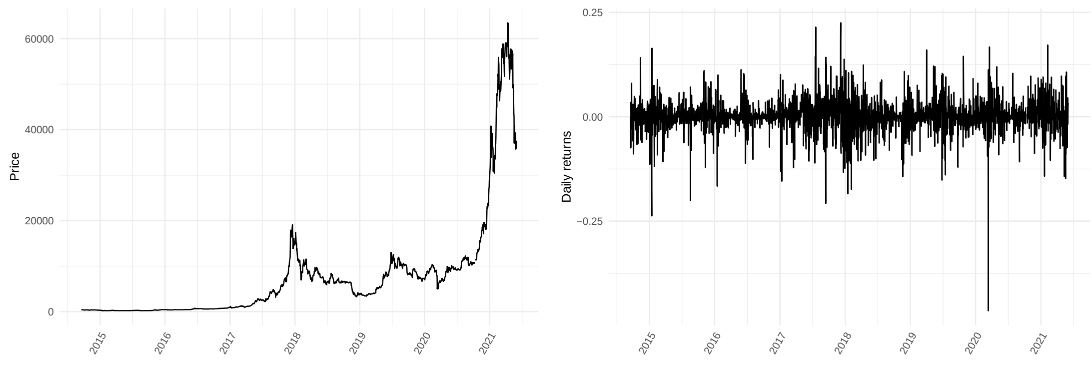

Figure 5: Price and returns for Bitcoin

Apple Apple is the most capitalized firms of the world. With the continuous innovation in the tech industry, it remains one of the most reliable blue chips in the market. In the last years, the job of Tim Cook has brought the company to invest in a broader field of activity, with the expansion in the services sector. The series is the most asymmetric one with a negative skewness equal to -1.749.

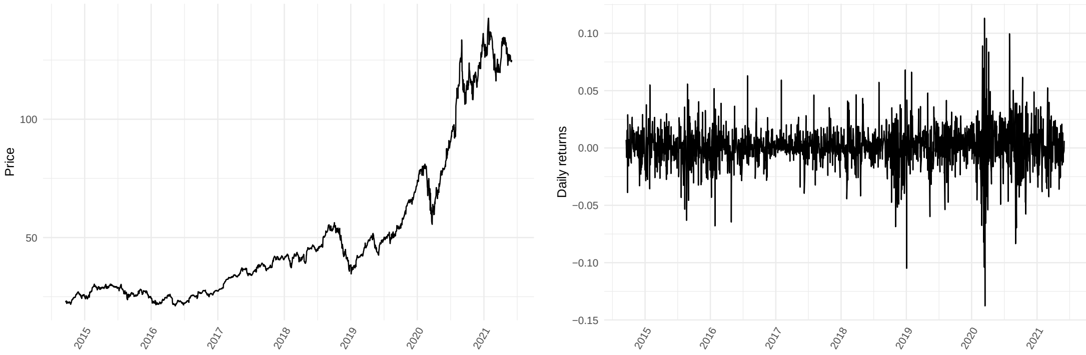

Figure 6: Price and returns for Apple

Walmart Walmart is the company with the highest number of employees in the world. With a PE of 32,62 the investors choose to buy the history of the firms. In this pandemic year the company troubled for a period, but the new politics, the innovations, among which the food delivery, and the strong balance sheet suggests a robust investment. The series of returns has a very low skewness and the prices behave similarly with respect to Apple with a continuous and solid growth.

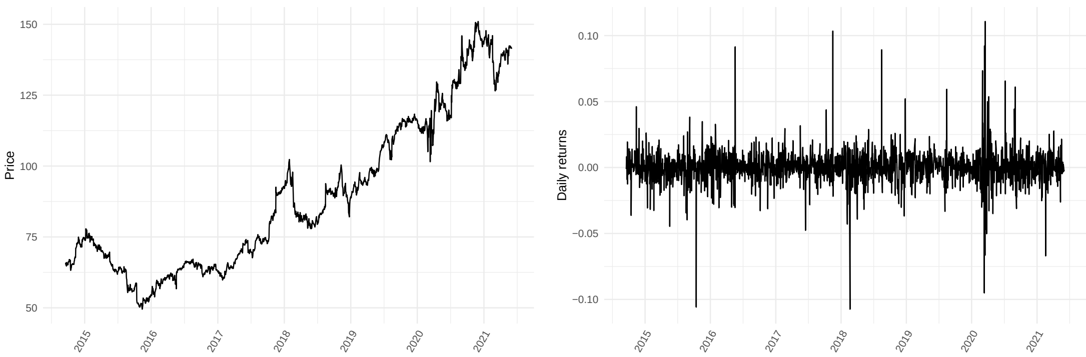

Figure 7: Price and returns for Walmart

Tesla Tesla is a big question; the enormous PE means people are willing to invest their money for earnings that does not still exist. The figure of the charismatic and visioner Elon Musk make think that the future will be bright and prosperous, meanwhile the balance sheet is still pessimistic with respect to the price of the stock. The returns seem symmetric with the kurtosis higher than those of the normal.

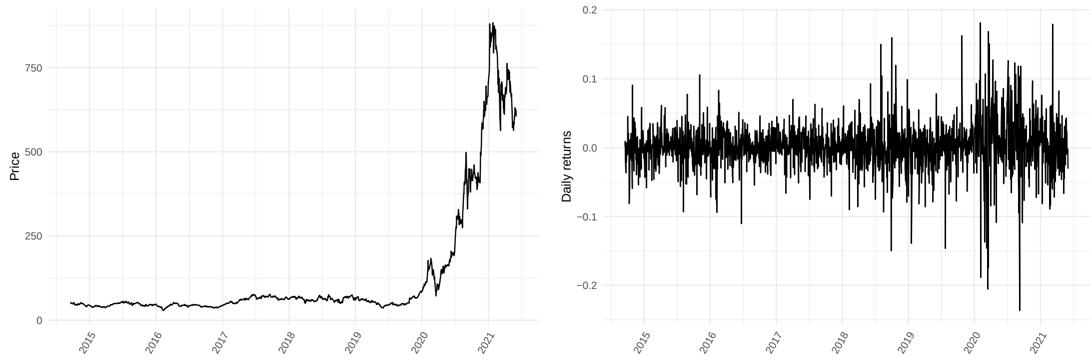

Figure 8: Price and returns for Tesla

Copula estimation

A copula is a statistical tool by which estimate the joint behaviour of risk factor changes. Better, it is a cumulative distribution function of uniforms distribution. Thanks to the fundamental theorem of simulation, other distribution can be transformed into uniforms, so that a multivariate distribution functions can be estimated without affecting the dependence structure of the original random variables.

The strategy is the following: first we find the best marginals through the implementation of different GARCH models. Then, we will use these, the Normal and the t-student specifications in order to estimate 5 different copulas: two elliptical, Normal and t-student; three Archimedean, Gumbel, Clayton and Frank.

Table 6: Log Likelihood GARCH models

	Apple	Walmart	Tesla	Gold	Bitcoin
GARCHn	7414	8479	5440	5423	2966
GARCHt	7558	8795	5677	5512	3282
gjrGARCHn	7452	8480	5441	5424	2966
gjrGARCHt	7583	8795	5677	5512	3287
eGARCHn	7458	8505	5439	5424	2978
eGARCHt	7590*	8801	5682	5512*	3304*
csGARCHn	7420	8507	5471	5426	2966
csGARCHt	7563	8804*	5688*	5511	3285

[a] * denotes the best model

Table 6 shows the value of the LogLik of the GARCH models for the estimation in the different series. The results are that Apple, Bitcoin and Gold have as a good estimator of the variance the e- GARCH with a t-student distributions, meaning that the asymmetric effect improved the estimation. Different the result for Walmart and Tesla with the cs-GARCH model that maximizes the value of the Log Likelihood, so the distinction between long and short run variance help estimating the parameters. The standardized residuals obtained from these fitted models will be used to investigate the dependent structure, and we will be referred to with the name of "Mixed".

Table 7: LogLik and AIC for portfolio with AAPL,WMT,TSLA

	Normal		*Student-t*		*Mixed*	
Copula	LogLik	AIC	LogLik	AIC	LogLik	AIC
Normal	224.6	-443.3	223.8	-441.6	220.0	-434.0
t-student	228.8*	-449.6*	227.6	-447.2	224.1	-440.2
Gumbel	146	-290.1	143.9	-285.9	199.1	-396.1
Clayton	185.5	-369	185.1	-368.3	105.0	-208.0
Frank	182.1	-362.1	181.0	-360.1	178.6	-355.1

[a] * denotes the best model

Table 7 shows the statistical tests LogLik and AIC for all the copulas and marginals fitted. The first things that can be notice is the outperform of the Elliptical copula over the Archimedean ones. In fact, Gumbel, Clayton and Frank have the lowest LogLik and the highest AIC, with a particular bad performance of the Clayton. For what concerns the Elliptical, the t-copula behave better for all the three marginal assumptions. The value of the tests in these three cases for it are not particularly different, but the Normal assumption of the marginals have the better values in the tests proposed.

All the things just said have also a graphical justification. Indeed, we have randomly simulated, with size equal to the length of the series, a multivariate distribution with all the parameters estimated with the different copulas and marginals. It appeared that both t-copula with normal and student-t distributions was similar to the scatter-plot of the original series. For sake of space and conciseness, we propose only the graph of the original series scatter plot and the simulated one of the best fitting copula, that is indeed the t- copula with Normal marginal distributions.

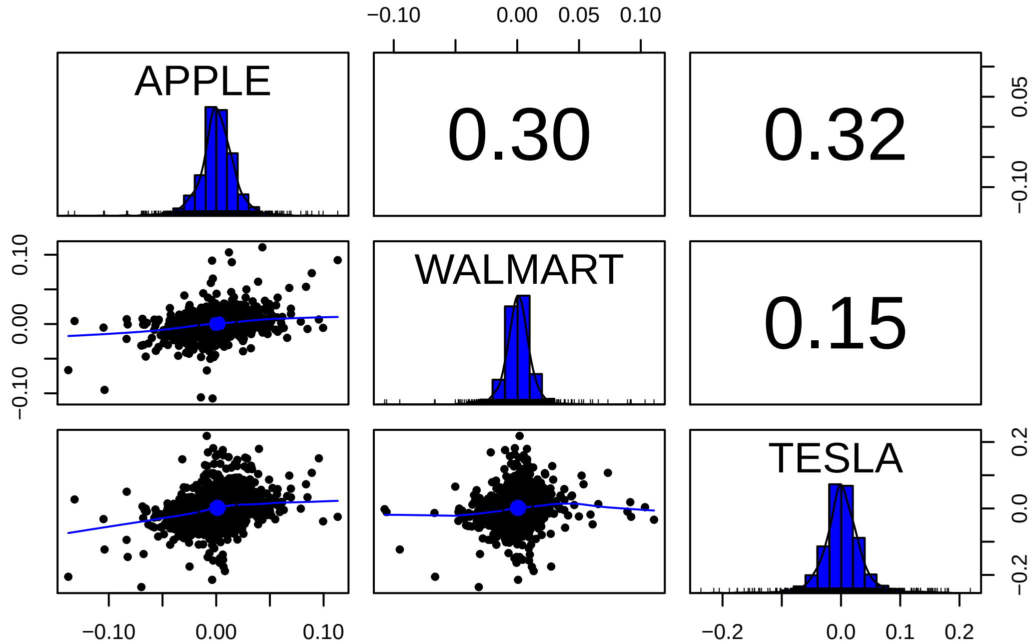

Figure 9: Pairs of portfolio z's and random simulation of t-copula

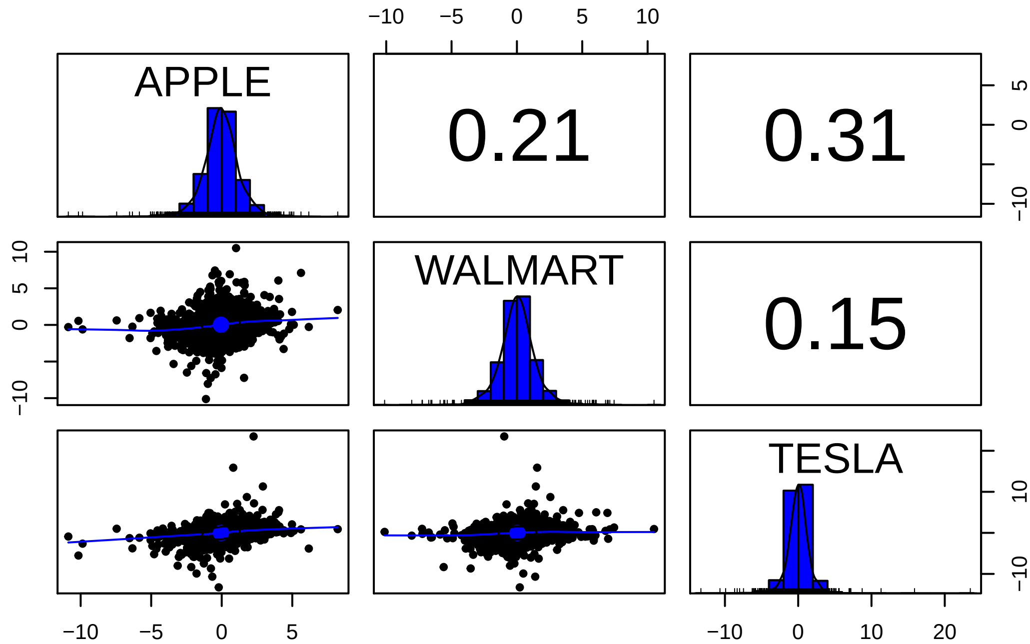

Figure 10: Pairs of portfolio z's and random simulation of t-copula

With a random simulations we can also forecast the Value at Risk and the Expected Shortfall. In Figure 11 are shown the VaR and ES estimated up to 200 steps ahead with the use of t-copula with normal marginals for the portfolio examined before. The black line represent the aggregated loss of the portfolio. The pink shade represent the confidence interval from 10% to 90% of the distribution of the forecasted losses.

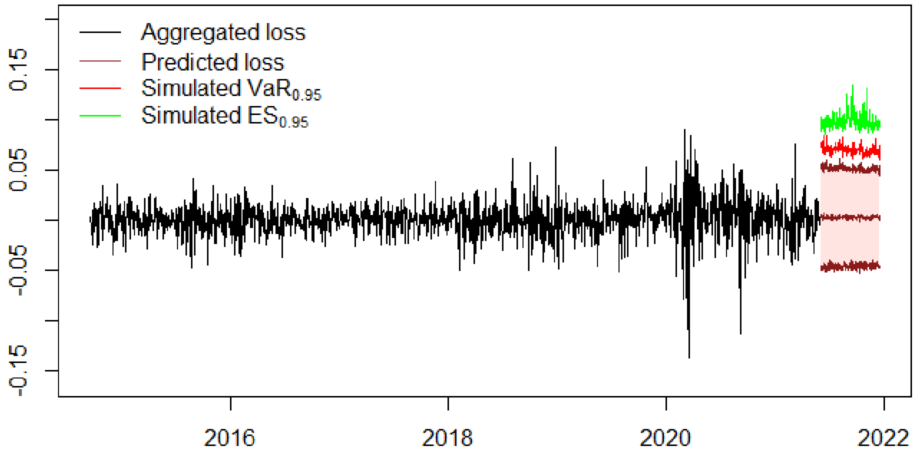

Figure 11: Forecast of Var and ES at 0.95

Efficient frontier

The procedure done for finding the best copula can be repeat with the addition of gold and Bitcoin to the original portfolio, both separately and jointly. Obviously, since the outstanding performance of Tesla last year, the expected returns of the portfolio will be very high and not in line with the average in market.

To study the performances of the portfolios we use the efficient frontier of Markovitz (data), that uses the covariance matrix and the expected returns of the log-returns to build an efficient frontier, where we can find the efficient portfolio.

Table 8: LogLik and AIC for portfolio with AAPL,WMT,TSLA,GOLD

	Normal		*Student-t*		*Mixed*	
Copula	LogLik	AIC	LogLik	AIC	LogLik	AIC
Normal	165.31	-318.62	220.69	-429.37	162.57	-319.13
t-student	169.35	-324.71	230.41*	-446.82*	229.64	-445.29
Gumbel	41.43	-80.86	77.81	-153.63	47.58	-93.16
Clayton	76.85	-151.69	118.14	-234.27	28.94	-55.88
Frank	52.25	-102.50	90.23	-178.47	39.87	-77.73

[a] * denotes the best model

Table 9: LogLik and AIC for portfolio with AAPL,WMT,TSLA,BTC

	Normal		*Student-t*		*Mixed*	
Copula	LogLik	AIC	LogLik	AIC	LogLik	AIC
Normal	175.08	-338.17	172.30	-332.6	164.90	-323.8
t-student	178.28*	-342.56*	175.39	-336.8	167.86	-327.7
Gumbel	57.4	-112.8	53.97	-105.9	82.73	-163.5
Clayton	109.6	-217.21	107.52	-213.0	73.28	-144.6
Frank	76.9	-151.8	74.98	-148.0	74.27	-146.5

[a] * denotes the best model

Table 10: LogLik and AIC for portfolio with AAPL,WMT,TSLA,GOLD,BTC

	Normal		*Student-t*		*Mixed*	
Copula	LogLik	AIC	LogLik	AIC	LogLik	AIC
Normal	177.38	-334.76	175.67	-331.35	175.64	-345.28
t-student	182.29*	-342.57*	180.94	-339.88	180.90	-353.81
Gumbel	35.57	-69.15	34.17	-66.34	60.25	-118.50
Clayton	71.79	-141.58	72.26	-142.52	73.28	-144.57
Frank	40.55	-79.11	40.38	-78.75	50.11	-98.21

[a] * denotes the best model

Table 8, 9 and 10 show the results of the evaluation of LogLik and AIC for the portfolio with the addition of gold, Bitcoin and both gold and Bitcoin, respectively. Looking at the tables, the Archimedean copula once again underperform, letting the Elliptical copulas be better in the fitting of the multivariate distribution. With the addition of gold and Bitcoin, t-copula remains the best choice, but in these cases not the Normal distribution.

In fact, for the portfolio with Bitcoin the best assumptions is the student distribution, meanwhile for gold and Bitcoin and gold the one with mixed distributions. The parameter estimated by the copula will be use for deriving the covariance matrix.

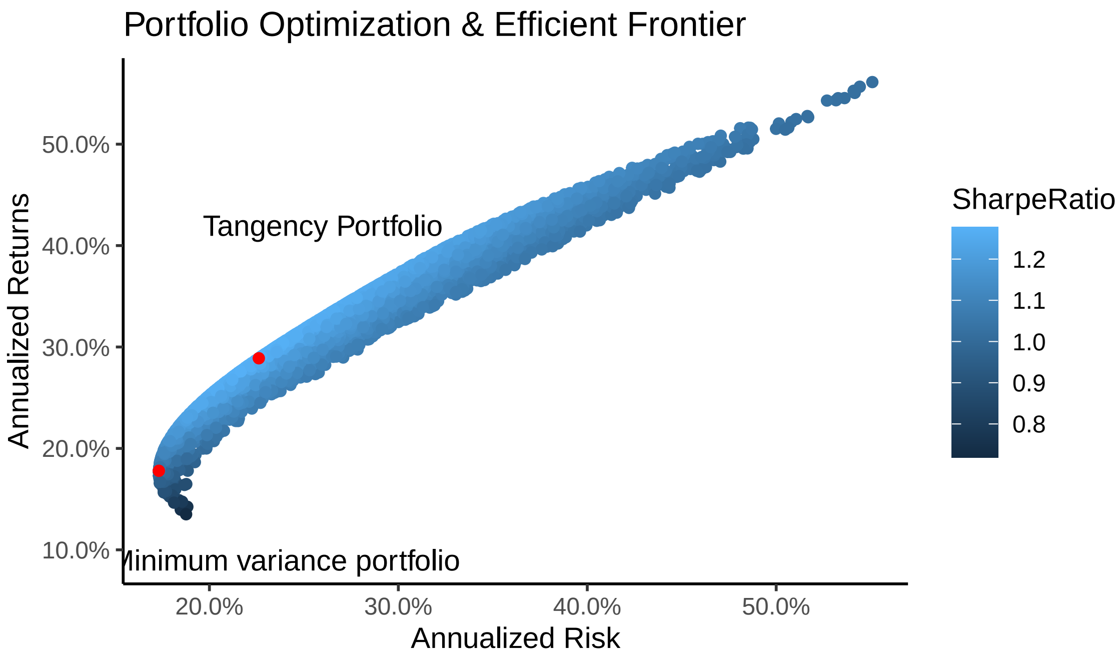

Figure 12: Porfolio with AAPL, WLM, TSLA

In our case, since of the assumption of risk-free rate, the Sharpe ratio is the ratio between the expected returns and the standard deviation of the portfolio. So, the higher the value, the higher the quantity of returns expected over a unit of risk. It is a simplistic model, but it can be useful to inspect the behaviour of the portfolio with the addition of the two assets.

Table 11: Efficient portfolios

Portfolio	Return	Risk	Sharpe.Ratio
Portfolio	0.2889	0.2261	1.277
Portfolio with Gold	0.3180	0.2412	1.319
Portfolio with Bitcoin	0.4364	0.2936	1.487
Portfolio with Gold and Bitcoin	0.4757	0.2932	1.623

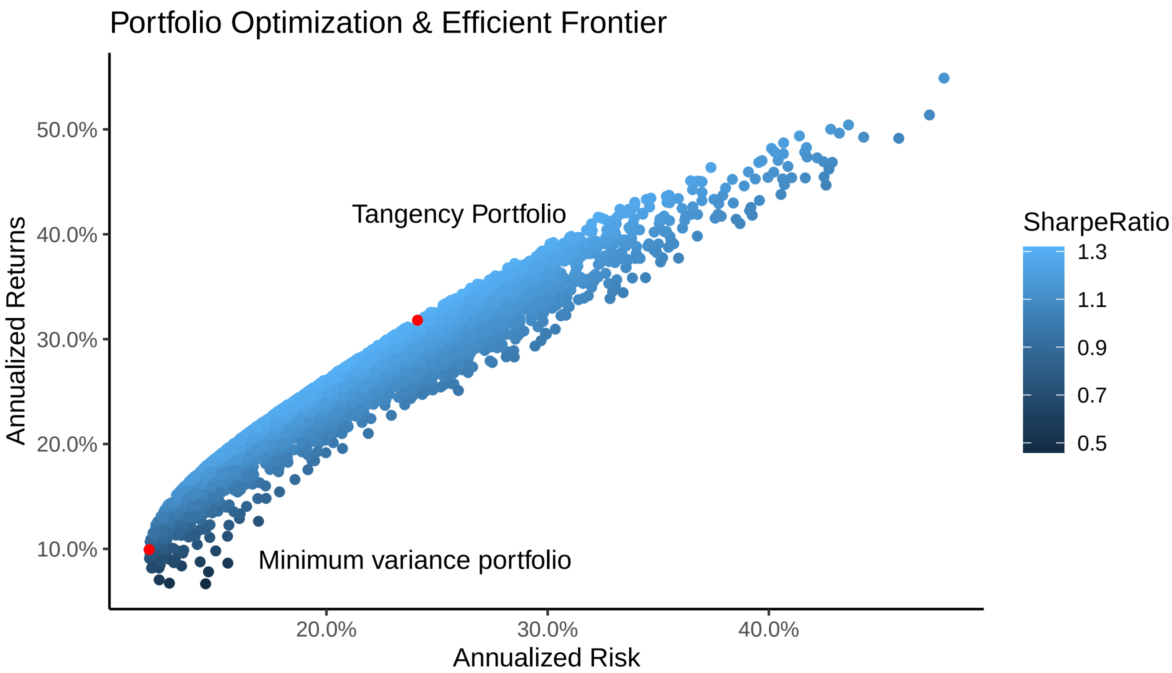

Figure 13: Porfolio with AAPL, WLM, TSLA and GLD

The addition of gold has the effect to raise both returns and risk but the Sharpe ratio is higher than the previous portfolio, meaning that the addition of the asset has improved the convenience with respect to performance and risk.

A different effect is produced by Bitcoin. In our extreme example, the portfolio returns soars above 40%, with an estimation of risk of 29%. An investor who is risk lover could take advantage of the new situation. In fact, the Sharpe ratio increased a lot, 1.6.

The jointly addition of Bitcoin and gold raise the expected return, without affecting so much the risk. This increase the Sharpe ratio to 1.62, the higher value of the four.

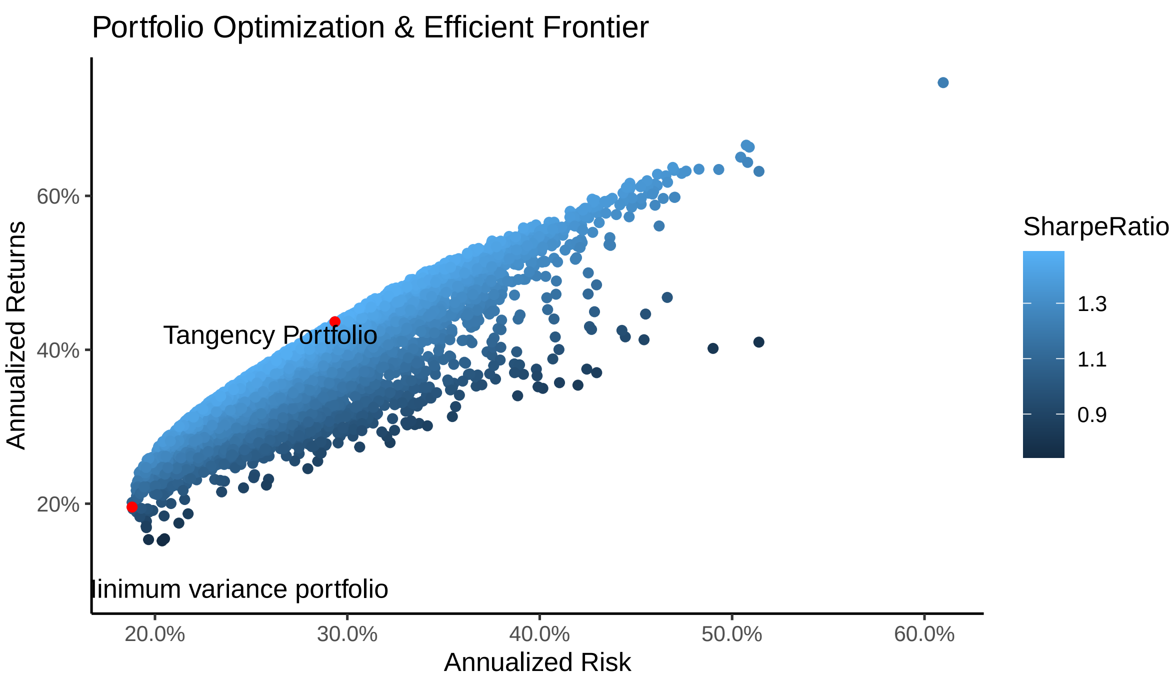

Figure 14: Porfolio with AAPL, WLM, TSLA and BTC

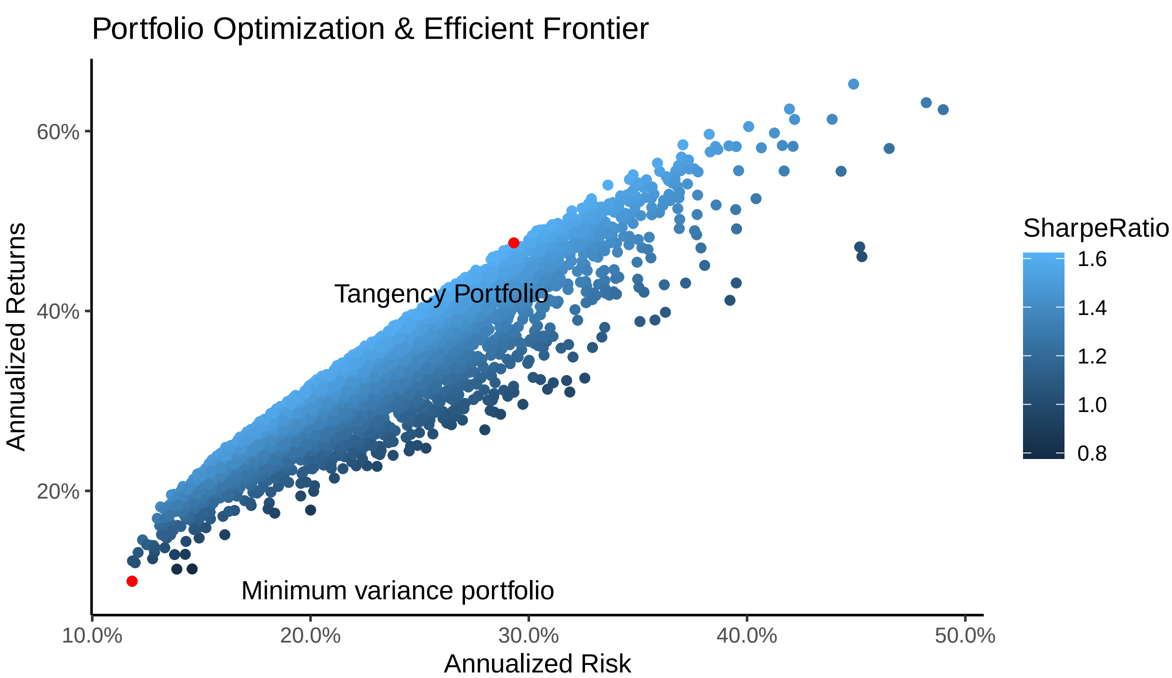

Figure 15: Porfolio with AAPL, WLM, TSLA, GLD and BTC

In conclusion, Historical Simulation is the best model to forecast the Value at Risk for gold returns. For what concern the parametric models, adding a factor that can capture the asymmetry improve the estimation. Also the behaviour of the interest rate helps to have good estimation for VaR. The other risk measure considered, the Expected Shortfall, shows many troubles with the normal distribution assumption, but it passes all the tests with the t-student assumption.

When we move to a multivariate framework, to model the joint behaviour of the series, we need to use copulas. As we have seen, dealing with stock like Tesla or Bitcoin, which observe many extreme events, there is the requirement of use a dependence structure able to capture this feature. In our case, t-copula outperform with respect to the others.

We have also analyzed, through the Markovitz Efficient Frontier, the impact of gold and Bitcoin in a portfolio. Gold, in our example, raise in all cases the Sharpe ratio, making more attractive the portfolio. Bitcoin raise consistently the expected returns, with a consistent growing of the risk, with an improvement of the ratio between the two. This analysis could be a starting point to try to answer at the question made in the introduction. We have no clue to give a complete answer but from the result of our extreme example it appears not a bad idea to keep both Bitcoin and gold in a single portfolio.

References

- Enjoy the Joy of Copulas: With a Package copula, 2007 (Yan)
- Modeling stock market indexes with Copula functions, 2011 (Leśkow,Mokrzycka,Krawiec)
- Portfolio Modeling Using the Theory of Copula in Latvian and American Equity Market, 2005 (Jansons, Lace, Kozlovskis)