Drawdown: from practice to theory and back again

Antoine-Michel Alexeev, Julien Bisch, Benjamin Souane and Ludovic Suchet

Professor: Eric Jondeau

Quantitative Asset Management Project

HEC Lausanne

UNIL | Université de Lausanne
HEC Lausanne

Contents

List of Figures

List of Tables

1 Introduction

In his famous book "The Black Swan", released in 2007, Nassim Nicholas Taleb focused on the extreme impact of rare and unpredictable outlier events. He notably challenged the adequacy of standard tools such as Normal distribution and standard deviation to help in predicting the future. The current state of the world, with the coronavirus outbreak, once more highlights the question: how do we deal with risk, and more specifically, how do we deal with risk when it is related to unknown events?[1]

Risk is a very broad, but crucial, notion in the investment world. A general rule, and very intuitive one, is that risk and expected returns are positive dependant. There are as many definitions as there exist ways to measure risk. However, in the event of large drawdowns in the markets, standard risk measures such as the volatility, the VaR and the Expected Shortfall are less significant. Hence this paper aims to discuss one of the alternative risk metrics: the maximum drawdown (MDD), which is an indicator of downside risk exposure. It consists of a specific measure of drawdown, which aim to capture the largest move from a peak to a trough of a portfolio.

More precisely in their paper, Goldberg and Mahmoud (2016) aim to develop a mathematical methodology in order to form expectations about future potential maximum drawdown since it does not seem to exist at the time. They have developed a measure of drawdown risk, the Conditional Expected Drawdown which can be defined as the expected maximum drawdown. This measure has interesting practical and mathematical features that allows optimization with respect to portfolio weights, and therefore address the risk and also asset management industry. More formally, the Conditional Expected drawdown is the tail mean of the distribution of drawdown. For an investor, it represents the average loss on the $\alpha\%$ worst scenario.

This paper is organized as follows. The second chapter describes the data used in our analysis. The third chapter defines the maximum drawdown process and the conditional expected drawdown (CED) as well as the mathematical properties of the CED. We will typically show that the CED is homogeneous of degree one and a convex risk measure.

[1] You can find the implementation of the project in Matlab and the data used on our Github Repository.

Finally, we present the methods to compute the contribution to CED for an asset and the construction of a CED parity portfolio. The fourth chapter is the empirical analysis of the MDD and the CED. We typically explore the relation between CED and the frequency of observations, the length of the path, and the auto-correlations. We do find interesting and significant results about these relations. We then analyze the relation with the recovery speed and the speed of the drawdown. Finally, we look at the risk contribution of CED over different portfolio allocations. The last chapter will discuss on of the potential extensions of the model.

2 Data

In order to perform this analysis on the maximum drawdown and the conditional expected drawdown, we had to use a large data set, allowing many large drawdowns to happen. To do so, we used the Fama-French Market portfolio data and 10 Industry Portfolio data. These portfolios start in 1927 up until this year, which is almost 100 years of data.

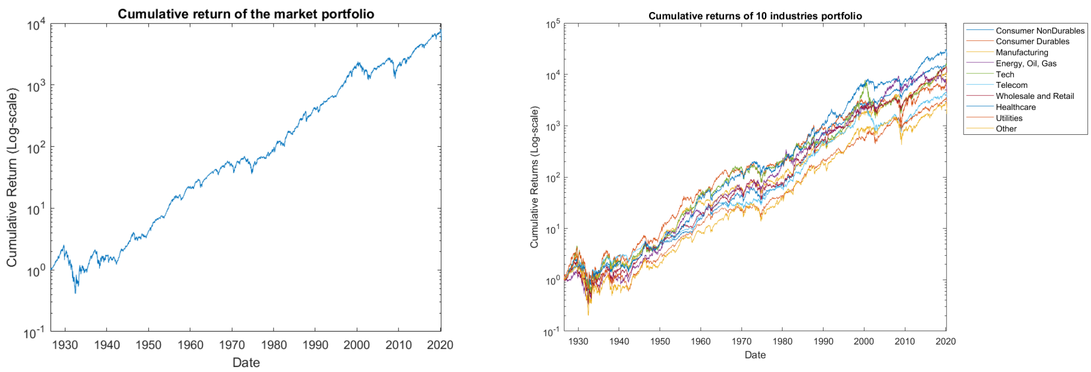

Figure 2.1: Market Portfolio and 10 industries cumulative returns from 1927 to 2020.

As we can see in the previous plots, this long data set are allowing us to experience many large drawdowns. Concerning the industry data, one can see that most of the data have relatively similar dynamics and that they all experienced drawdowns in the same periods. However, the small differences will already allow us to implement some strategies and see the differences in CED and Marginal Contribution to CED.

The market portfolio does also have a risk free rate, which is the one-month treasury bill rate that goes from 1927 to 2020. This will allow us to show the effect of different allocations on the CED.

Moreover, we also use intra-day bitcoin prices from 2010 to 2020 in order to see the effect of high frequencies on the CED. These data are retrieved from Kaggle[2].

[2]You can find the data here : Bitcoin Historical Data

3 MDD, CED and Methodology

3.1 MDD

3.1.1 Intuition behind Maximum Drawdown

When considering a levered portfolio, the worst that could happen would be to face a significant loss in the value of the current positions such that it creates a liquidity trap while reaching the maintenance margin. This would potentially lead to the sale of valuable positions under unfavorable market conditions. Therefore, actively managed funds often use maximum drawdown to have a better idea of what could be the worst case, and thus to manage funds consequently. Investors tend to use the MDD against a better-known risk indicator that is the standard deviation. The latter, which is used to quantify how much returns fluctuate around their means, encounters some limitations when studying the downside aspect of risk. Indeed, as it is influenced by both positive and negative market results, it can offset the risk of loss from a portfolio position. It can therefore lead to completely opposite intuition about the risk of the market. Moreover, since the maximum drawdown focuses on negative events, it will better represent the loss aversion of the investors.

To summarize, the most interesting feature of a maximum drawdown is that it takes into account worst market conditions, as we can consider a long period of time and different frequencies of observations unlike the VaR, which describes losses, simply by indicating the single worst realized return under normal market conditions. One should note that MDD and VaR can be considered two distinct risk measures. In theory, large size of VaR does not necessarily lead to large MDD because one big negative return is realized in the process of positive returns. Alternatively, if several consecutive small negative returns are observed, MDD could be large whereas VaR is small. Even though over a given time horizon, only one maximum drawdown is realized, it is interesting to consider the distribution of drawdowns.

3.1.2 Drawdown Process

The drawdown process is the distribution of drawdowns over a fixed time horizon $T \in (0, \infty)$.

$$D_{t \in [0,T]} = \sup_{u \in [0,t]} X_u - X_t \tag{3.1}$$

3.1.3 Maximum Drawdown

Over the same fixed time horizon T, is therefore the maximum of all Drawdown :

$$MDD(X) = \sup_{t \in [0,T]} D_t \tag{3.2}$$

This process allows us find a distribution of maximum drawdowns over a rolling window of length T. This yield the following random variable.

$$MDD(X)_{t \in [T,N]} = MDD(X)_t \tag{3.3}$$

Where

- $MDD(X)_t$ is the maximum drawdown $\in [t - T, t]$
- N is the number of random variables X observation

This process allows to construct a distribution of $N - T$ maximum drawdowns which is, assuming N is large and T is small, a relatively high number of observations. Clearly, the length of the time horizon has a strong impact on the value of $MDD(X)$.

3.2 CED

This high number of observations allows us to construct a distribution of maximum drawdowns. Therefore, we can derive a distribution of a risk measure, and more precisely the expected mean of the tails and deciding a threshold, which is precisely the CED.

$$CED_\alpha(X) = E[MDD(X) \mid MDD(X) > DT_\alpha(MDD(X))] \tag{3.4}$$

Where :

$$DT_{\alpha}(MDD(x)) = \inf(m \mid P(MDD(X) > m) \leqslant 1 - \alpha) \tag{3.5}$$

i.e the $1 - \alpha$ quantile of the MDD distribution.

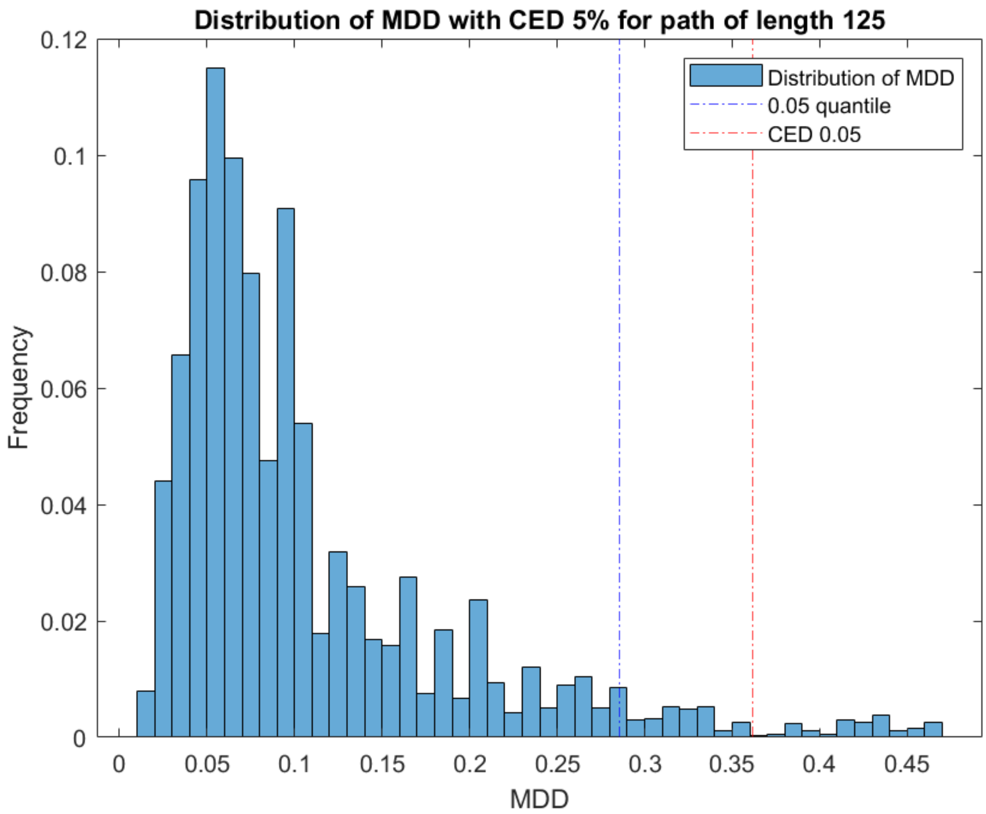

Figure 3.1: Distribution of Maximum Drawdown for a path of 125 days and CED at 5% for the Fama-French market portfolio from 1927 to 2020.

As we can see on the previous plot, the CED will be smaller than the maximum drawdown. Indeed, it is the tail mean of the maximum drawdown that are bigger than the $1 - \alpha$ quantile of the distribution.

By definition, this measure of risk is path dependant but it also has many other properties that we will show. This properties allows CED to be a coherent risk measure [3].

[3] As express by Artzner, Delbaen, Eber, and Heath (1999) in their paper Coherent Measures of Risk

3.2.1 Normalization

For all deterministic constant C, $MDD(C) = 0$

By definition a deterministic portfolio is not exposed to Drawdown risk and therefore $MDD(C) = 0$.

3.2.2 Positivity

For all random variable $X \in R^{\infty}, MDD(X) \geqslant 0$

Again, by definition, $MDD(X) \geqslant 0$ and therefore, it's $1 - \alpha$ quantile is also $\geqslant 0$.

3.2.3 Shift Invariance

For all $X \in R^{\infty}$ and $C \in R^{\infty}, CED_{\alpha}(X + C) = CED_{\alpha}(X)$

From 3.1, one can show that for all random variable $X \in R^{\infty}$ and deterministic constant $C \in R^{\infty}$ the drawdown process:

$$\begin{aligned} D(X+C)_{t\in[0,T]} &= \sup_{u\in[0,t]} (X+C)_u - X_t + C \\ &= \sup_{u\in[0,t]} X_u + C - X_t + C \\ &= \sup_{u\in[0,t]} X_u - X_t \\ &= D(X)_{t\in[0,T]} \end{aligned} \tag{3.6}$$

And therefore $CED_{\alpha}(X + C) = CED_{\alpha}(X)$.

3.2.4 Convexity

For all $X, Y \in R^{\infty}$and $\in [0, 1], CED_{\alpha}(\lambda X + (1-\lambda)Y) \leqslant CED_{\alpha}(\lambda X) + CED_{\alpha}((1-\lambda)Y)$

By definition:

$$D(\lambda X + (1-\lambda)Y)_{t\in[0,T]} = D(\lambda X)_{t\in[0,T]} + D((1-\lambda)Y)_{t\in[0,T]} \tag{3.7}$$

Due to the fact that the relation is linear and we are only doing a linear combination of two random variables. Hence, since $MDD(X) = \sup_{t\in[0,T]} D_t$, we have the following relation:

$$MDD(\lambda X + (1-\lambda)Y) \leqslant MDD(\lambda X) + MDD((1-\lambda)Y) \tag{3.8}$$

Finally, the tail mean is subadditive and positive homogeneous of the underlying distribution and therefore:

$$CED_\alpha(\lambda X + (1-\lambda)Y) \leqslant CED_\alpha(\lambda X) + CED_\alpha((1-\lambda)Y) \tag{3.9}$$

This property is the cornerstone of diversification, this fact is clearly known for the portfolio variance where the efficient frontier is improving with the increasing number of assets. This properties allows to do portfolio optimization based on the CED as a risk measure.

3.2.5 Positive Degree-one homogeneity

For all random variable $X \in R^\infty$ *and* $\lambda > 0$, $CED_\alpha(\lambda X) = \lambda CED_\alpha(X)$ Here again using 3.1, one can show that :

$$\begin{aligned} D(\lambda X)_{t\in[0,T]} &= \sup_{u\in[0,t]} (\lambda X)_u - \lambda X_t \\ &= \lambda \sup_{u\in[0,t]} X_u - \lambda X_t \\ &= \lambda(\sup_{u\in[0,t]} X_u - X_t) \\ &= \lambda D(X)_{t\in[0,T]} \end{aligned} \tag{3.10}$$

All these properties are allowing us to use this risk measure to optimize portfolio and last properties typically allow us to re-scale weights post optimization to lever a position.

3.3 Contribution to risk

For any differentiable risk measure ρ and portfolio $P = \sum_{i=1}^{n} w_i F_i$, the *marginal risk contribution* is the derivative of the risk measure with respect to the portfolio weights :

$$MRC_i^{\rho}(P) = \frac{\partial \rho(P)}{\partial w_i} \tag{3.11}$$

This process provides an economically intuitive way of decomposing risk. Indeed, if ρ is homogeneous of degree one (which is the case for the CED), one can show using Euler's theorem on homogeneous functions :

$$\sum_{i=1}^{n} w_i MRC_i^{\rho}(P) = \sum_{i=1}^{n} RC_i^{\rho}(P) = \rho(P) \tag{3.12}$$

Which allows to define the *fractional risk contribution* which is the percentage of contribution to $\rho(P)$ of asset i :

$$FRC_i^{\rho}(P) = \frac{RC_i^{\rho}(P)}{\rho(P)} \tag{3.13}$$

The challenge to perform this analysis is to find the partial derivative of the risk measure with respect to the weights of each assets. This can be easily done when the risk measure is the volatility of the portfolio. However, with the *CED* it is more complicated. Therefore, in this paper, we will numerically differentiate the risk measure to find the MCR. We will use the symmetric derivative :

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x-h)}{h} \tag{3.14}$$

and more precisely in our case :

$$\begin{aligned} MRC_{i,t}^{\rho}(P) =& \frac{\partial \rho(P)}{\partial w_i} \\ =& \lim_{h \to 0} \frac{\rho(W_{+h} | w_{i,t} = w_{i,t} + h) - \rho(W_{-h} | w_{i,t} = w_{i,t} - h)}{h} \end{aligned} \tag{3.15}$$

Where: $\rho(W_{/-+h} | w_{i,t} = w_{i,t} + / - h)$ means that W is not change expect for $w_{i,t}$ to which we add or take out h.

Hence, to numerically perform this derivation, one need to choose a value for h (a usually accepted value is 0.001) and compute :

$$MRC^{\rho}_{i,t}(P) = \frac{\rho(W_{+h}|w_{i,t} = w_{i,t} + h) - \rho(W_{-h}|w_{i,t} = w_{i,t} - h)}{h} \quad (3.16)$$

This means that to compute the marginal contribution to CED of a portfolio for T period and n assets. We need to perform $T * n$ partial derivatives which can be computationally intensive and introduce some noises into the process.

This measure is very practical and allows us to easily manage the risk of the portfolio. Moreover, since we have shown that our risk measure CED is convex and homogeneous of degree one, we can allocate a portfolio under CED parity with the following procedure.

3.3.1 CED Parity Portfolio

Risk parity aims to distribute the weighs in such a way that each assets contribute equally to the risk :

$$w_t^{RP,i} MRC_t^i(P) = constant, \forall i \quad (3.17)$$

This can be solve using the following constrained optimization problem [4]:

$$L(w_t) = \sum_{i=1}^{N_t} log(w_t^i) - \lambda(CED_{w_t} - CED_{Target}) \quad (3.18)$$

Taking all the partial derivatives :

$$\frac{\partial L(w_t)}{\partial w_t^i} = \frac{1}{w_t^i} - \lambda \frac{\partial CED_{w_t}}{\partial w_t^i}, \forall i, t \quad (3.19)$$

And therefore, setting these partial derivatives to zero :

$$w_t^{RP,i} MRC_t^i(P) = \frac{1}{\lambda} = constant, \forall i \quad (3.20)$$

[4] This approach is inspired from Baltas, 2015 which introduce this procedure for a long-short risk parity portfolio

Therefore, for each re-balancing, whatever the frequency, we can allocate the portfolio such that each assets is entering with the same risk contribution. Since CED is homogeneous of degree one, one can re-scale the weights post optimization without changing the risk parity allocation. Thus, the CED_{Target} can more or less be any number since we will rescale the weights post optimization.

4 Empirical Analysis

4.1 Path Length

The first section of our analysis will consider the length of the path on which the maximum drawdown is computed. More precisely, this means that in equations 3.1 and 3.2, the fixed time horizon will be the length of the path that we will vary. To perform this analysis, we compute the $CED_{95\%}$ for the market portfolio for path lengths going from 10 to 2000 days. One can show that a higher path length will yield a higher or equal CED than a smaller path length. Indeed, if an event occurs in $T \in (0, 1, 2, ...T)$, it also occurs in $T_{+1} \in (0, 1, 2, ...T, T+1)$ and more generally in $T_{+n} \in (0, 1, 2, ...T, T+1, ..., T+n), n \in \mathbb{N}$ and therefore:

$$DD_{SmallerPath} \subset DD_{LongerPath} \tag{4.1}$$

Hence due to properties of the supremum :

$$CED^{\alpha}_{LongerPath} \geqslant CED^{\alpha}_{SmallerPath}, \alpha \in (0, 1) \tag{4.2}$$

To show this, we performed the following regression :

$$CED_i = \beta_0 + \beta_1 PathLength_i + \varepsilon_i \tag{4.3}$$

This regression yielded the following estimation :

	Estimate	SE	tStat	pValue
(Intercept)	0.35951	0.0061154	58.7874	5.2671e-73
Path Length	0.00043372	1.0033e-05	43.2281	1.5898e-61

Table 4.1: Regression of the $CED_{5\%}$ on the length of the path

As we can see, the results are highly significant and the R^2 of the estimation is 95,45 %. This means that the path length increases the CED in a significant way. We can see that the intercept explains a high part of the CED. And then, mechanically, the path length increases the CED. This can explain why we obtain a very high R^2.

This regression shows one of the prime features of the maximum drawdown and conditional expected drawdown, their path dependency. Indeed, if we were computing the volatility or the VaR for the same data and different path lengths, there would be no reason that those measures of risk would significantly increase with a larger path.

The figure 4.1 below is also showing this dependency with respect to the change in confidence level α. In fact, by definition, an increase in the α parameter leads to an increase in the CED. As already mentioned through the regression, we can once more see on the plot that there is a clear positive relationship between CED and path length. However, one must note that this relationship begins to flatten after some times, since drawdown tends to have limits in time and intensity. Therefore, increasing the path lengths indefinitely will not change the CED.

Finally, we have used empirical data to perform this grid analysis, and we notice a lot of "stairs". We would probably have a smoother function with simulated data since we take the tail mean of the MDD distribution and not MDD itself.

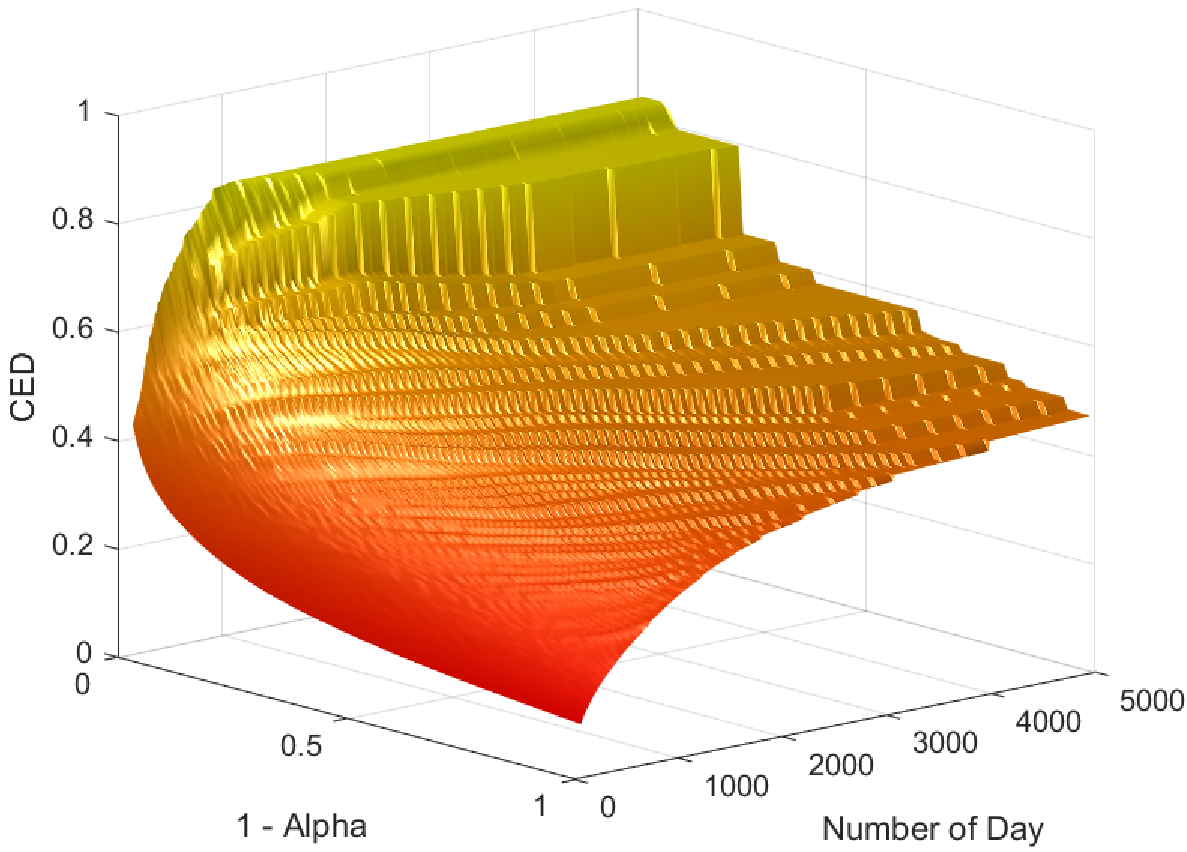

Figure 4.1: Surface plot of the CED w.r.t to the number days and $1 - \alpha$

4.2 Frequency of Observations

The MDD is sensitive to the frequency of data. A higher frequency of observations tends to return larger drawdowns. Indeed, more frequent observations are more likely to capture the true highs and lows. Therefore, using daily or even intra-days data allows the investor to capture shocks and details that one will miss when using only monthly data.

By construction, a higher frequency of observations will automatically yield a CED at least equal to the lower frequency, but that can be higher. Indeed, expressing the distribution of drawdown as mathematical sets, one can show that :

$$DD_{lowerfrequency} \subset DD_{higherfrequency} \tag{4.4}$$

Since every event that happens during a fixed time horizon $T \in (0, 1, 2, ..., T)$ also happens in $T_{higherfrequency} \in (0, t, 2 * t, ..., T), t \in (0, 1)$. However, an event occurring in $T_{higherfrequency} \in (0, t, 2 * t, ..., T)$ does not necessarily occurs in $T \in (0, 1, 2, ..., T)$. Hence, again due to properties of the supremum :

$$CED^{\alpha}_{higherfrequency} \geqslant CED^{\alpha}_{lowerfrequency}, \alpha \in (0, 1) \tag{4.5}$$

To show this relation, we have perform the following regression on our data:

$$CED_{i,5\%} = \beta_0 + \beta_1 * PathLength_i + \beta_2 * BinaryWeekly_i + \beta_3 * BinaryMonthly_i + \varepsilon_i \tag{4.6}$$

This has yielded the following results:

	Estimate	SE	tStat	pValue
(Intercept)	0.38324	0.0051095	75.005	2.3485e-182
Path Length	0.00039057	7.1582e-06	54.5627	1.6945e-147
Weekly	-0.010523	0.0046058	-2.2847	0.023107
Monthly	-0.29917	0.0046058	-64.9557	1.9835e-166

Table 4.2: Regression of the $CED_{5\%}$ on the length of the path and the Frequency of observations. Weekly and Monthly are binary variables.The R^2 is of $96,9\%$

The results demonstrate that, in our sample, the data being observed monthly will yield a smaller CED with a high degree of significance. However, the weekly observations do not seem to yield a relation so strong. Indeed, the estimate of β_2 is negative and significant but the value of the parameter is very close from zero. We have illustrated the following results on the next plot that represent the $CED_{5\%}$ for different length (in days).

On the figure 4.2 below, we can see a visual representation of the relation between CED and both the path length and the frequency of observation. Indeed, we can see that the Monthly curve is way below the two other ones. This also shows the fact that on our sample, the difference between Weekly and Daily observations is significant but almost negligible. This clearly shows the fact that using lower frequency data can be misleading. Indeed, the conditional expected drawdown is lower but the daily and weekly market movements happened anyway. Hence, one might be forced to liquidate without having an important CED and maximum drawdown. However, one can note that this high difference is an empirical one and that it could be different when using other data.

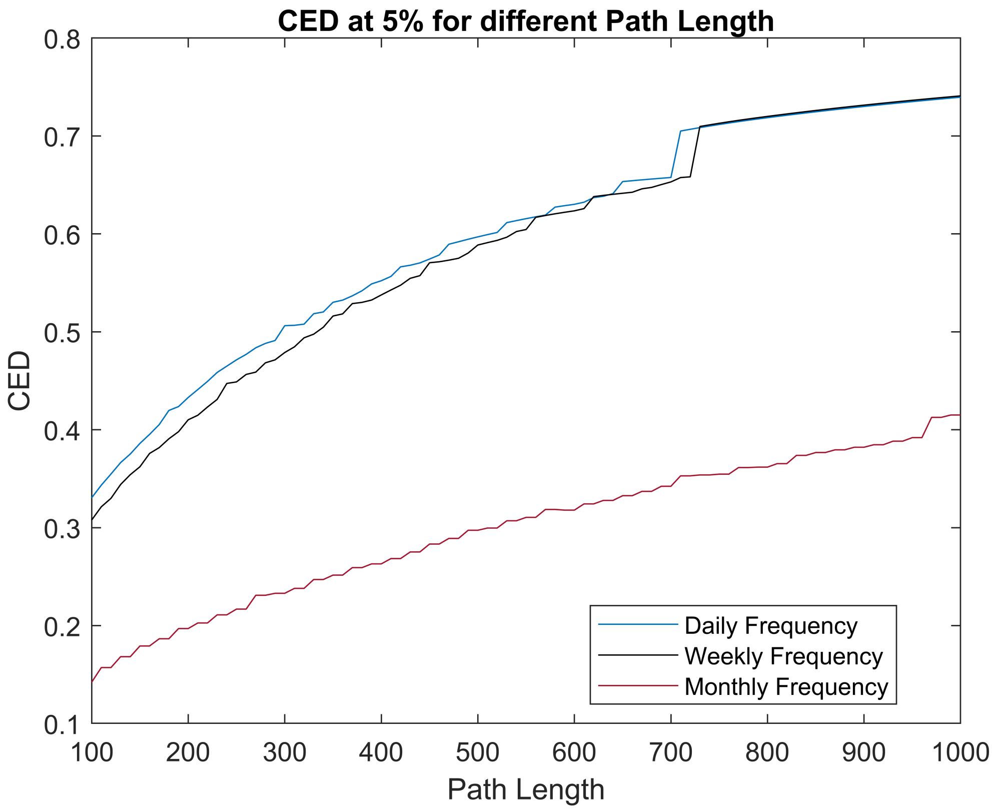

Figure 4.2: $CED_{5\%}$ for different path length and different frequency of observations. The Path Length is in days.

4.3 Auto-Correlations

In order to see the effect of the auto-correlation on the drawdown and CED, we will use a simulated AR(1) process allowing us to change the auto-correlation factors and look at the impact of it on the risk measures. One must note that the absolute level of these risk measures will be of almost no interest since the process is not simulated in order to fit true financial data. The idea is to show mechanically the relation between the auto-correlation and the risk measures. Hence, the dynamic of this risk measure and the comparison between them are interesting. The simulated process is the following AR(1) process:

$$R_{t+1} = R_t * k + \varepsilon_t$$

Where :

- k is the auto-Correlation factors.
- ε_t is normal with mean 0 and volatility 0.01.

To recall, auto-correlation is a measure of the dependence of returns. It is quite similar to correlation, excepting the fact that it calculates the correlation of an asset with itself over time, shifted by a lag period. The auto-correlation model is particularly sought after by the hedge fund industry in order to take advantage of it. Nevertheless, auto-correlation can also come at a cost of losses when poorly managed. On the following plot, one can observe the volatility, the expected shortfall, and the conditional expected drawdown of the process for different levels of auto-correlation.

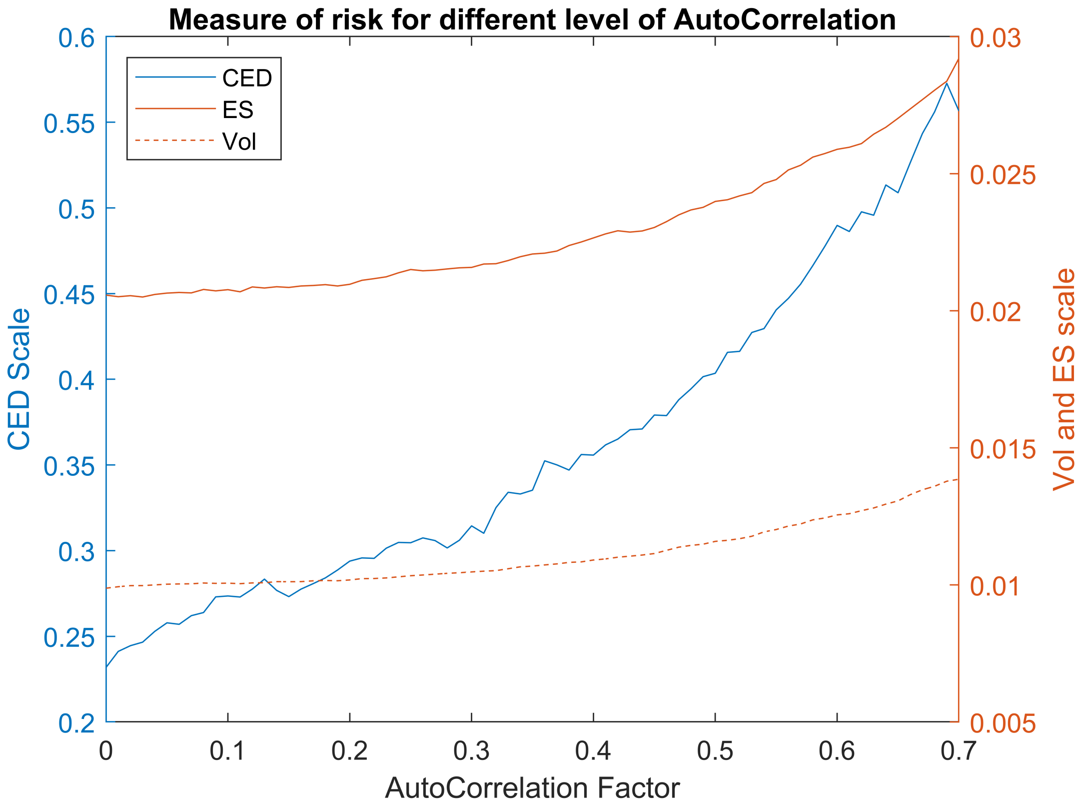

Figure 4.3: Level of volatility and 95% ES and CED for different level of auto-correlation parameters in an AR(1) process.

Here, we can see that every risk measures tend to increase with a higher auto-correlation factor. However, we can see that the dynamic is more pronounced for the CED. This clearly shows one of the advantages of CED, which better captures the auto-correlation of the process. Indeed a process that is more auto-correlated will tend to have a higher drawdown since when it starts losing value, this might last longer. However, small negative changes over-time are not captured as well by the volatility. The following plots show the distribution of drawdown for different levels of auto-correlation.

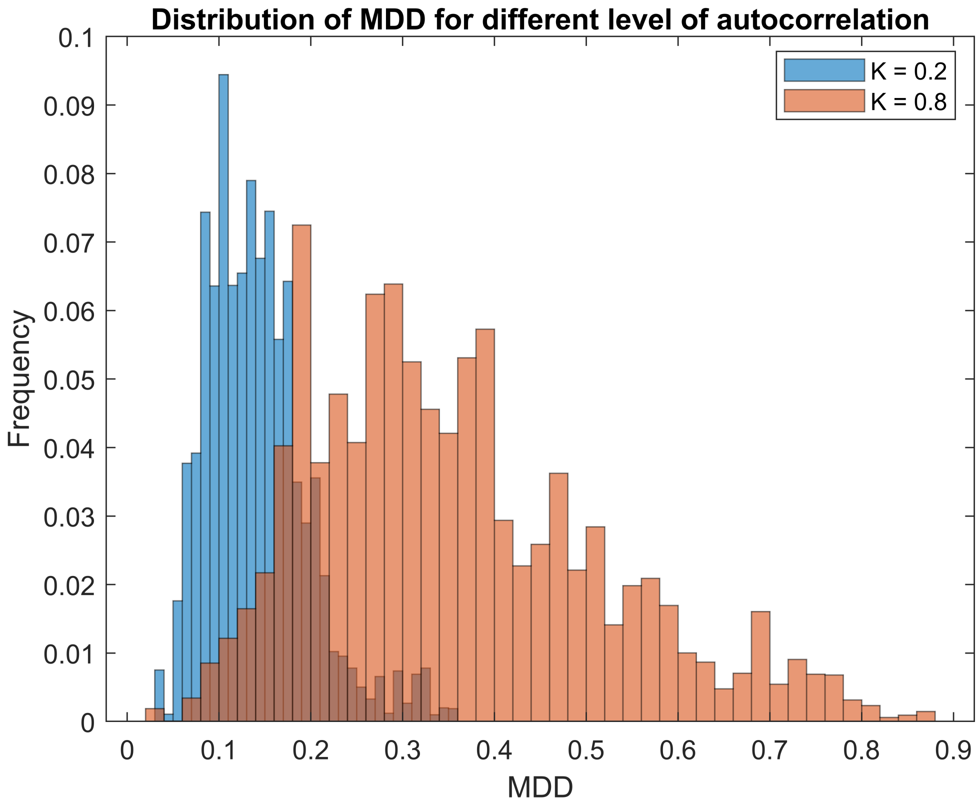

Figure 4.4: Histogram of the distribution of drawdown for different level of auto-correlation in simulated AR(1) Processes.

Here, once more we can point out a drawback of using standard volatility as a measurement of risk as it does not consider auto-correlation. Indeed, the convention that consists in calculating the annualized volatility as a multiplication of daily return's volatility with the square root of the numbers of business day within a year relies on the assumption of a constant variance over time. While this may be valid for uncorrelated returns, namely independent returns, it does not fairly represent the financial world [5].

We performed a regression in order to quantify the impact of the correlation on the CED. The results are shown in the following table:

	Estimate	SE	tStat	pValue
(Intercept)	0.13472	0.022626	5.9544	3.9872e-08
Kappa	0.70508	0.039092	18.0366	4.7051e-33

Table 4.3: Regression of the $CED_{5\%}$ on the Auto-Correlations of simulated AR1 process. The R^2 of the regression is 76.6%

[5] This is typically the case when financial markets are modeled through geometric Brownian motion. The curious reader can go here to show the mathematical proof.

As we can see on the regression table, the higher the estimate of the parameter κ, the higher the CED. Moreover, the results are highly significant.

We can note that this relation of auto-correlations to drawdowns can be incorporated to assets management, by using what can be called a drawdown control. For a positive auto-correlation, it consists in reducing weights of assets when they lose money and increasing the weight of those which make profits. The inverse method is used when returns are negatively auto-correlated. Globally, it's a way of improving one's risk-adjusted returns.[6]

Overall, accounting for auto-correlations allows the investor to obtain a better idea of the riskiness of the various assets in its portfolio.

The evidence of significant auto-correlation in our data leads to the question of its persistence over time. This is what lead us to the following section.

4.4 Relation Between DrawDown and "DrawUp"

After having studied the magnitude of investment loss, we will be interested in another risk characteristic associated with drawdown: the temporal dimension. This particular temporal dimension of risk is a topic that has not really been previously studied in the academic literature, especially in the context of risk measures. Nevertheless, it remains of high importance. Indeed, for example, a fund manager will be very interested in the time its positions will take to regain its previous maximum. Additionally, it can be useful for practitioners as a liquidation stopping time, which captures a subjectively set time threshold beyond which the fund will liquidate its positions if the drawdowns exceed this threshold (Ola Mahmoud, 2016). It allows the fund to take action if the value of its investment remains underwater for too many consecutive periods. To achieve this stopping time, Mahmoud introduced a distribution of the maximum duration time, allowing her to make some expectations about the lengths of drawdowns. We will not cover this point in our paper, but we find it pertinent to mention for the purpose of a curious reader.

In this section, we will discuss the duration as well as the recovery time associated with maximum drawdowns. While duration is a measurement of the length in time from peak to trough, the recovery time describes the amount of time needed from a maximum drawdown

[6] You can find an interesting paper about drawdown control here Dynamic Allocation Strategies for Absolute and Relative Loss Control, Daniel Mantilla-García

to return to its original level (i.e prior peak). We will also aim to give an illustration of the different speeds at which a drawdown occurred compared to its recovery period.

	Estimate	SE	tStat	pValue
(Intercept)	-490.7688	8.4411	-58.1404	0
Intensity	3465.9702	36.3762	95.2812	0
Speed	-0.015251	0.034663	-0.43998	0.65996

Table 4.4: Regression of the recovery speed on the intensity and speed of the DrawDown. The R^2 is of 43.4%.

We can see trough this regression that the clearly the intensity of drawdown has a significant "positive" impact on the recovery period. However, this relation is not the same for the Speed of the drawdown, the longer the drawdown, the smaller the recovery period. However, this results is not significant at all and therefore should not be taken into account. Through illustration 4.5 (see below), we observe that over the last century, the speed of recovery, approximated in number of days is globally slower than the speed of drawdowns. We also note that from 1950 to 2000 those two elements evolved nearly at the same speed.

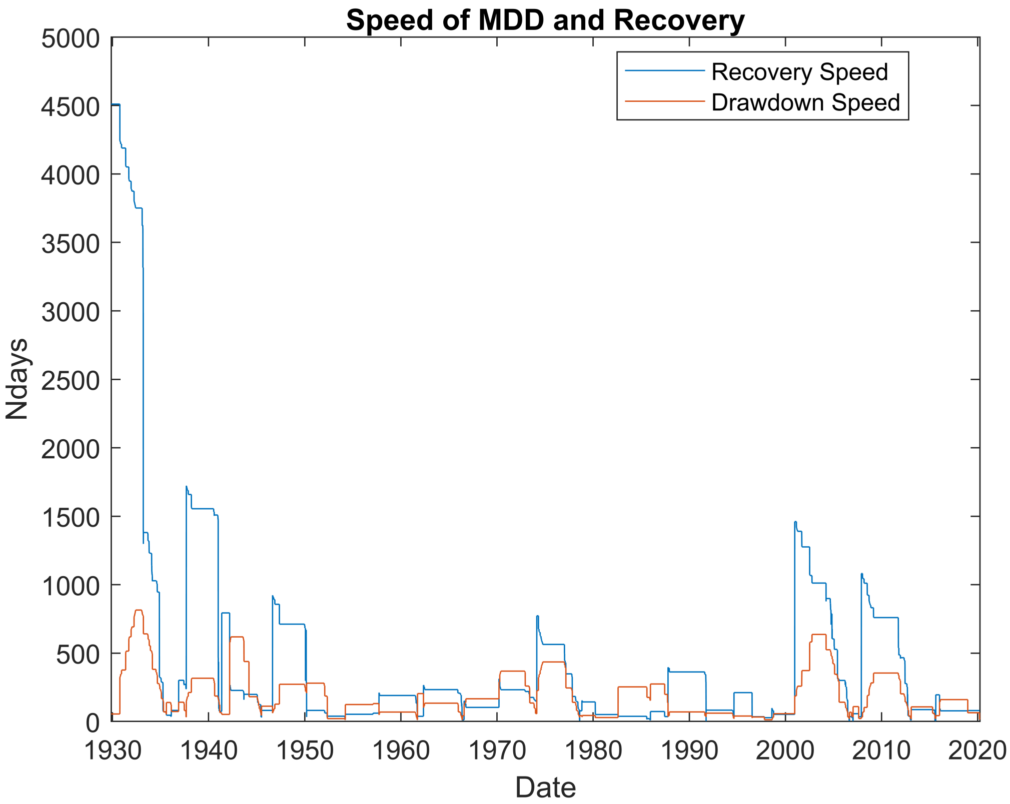

Figure 4.5: DrawDown and Recovery "Process" over the entire period (1926 - 2020). One should note that "Corona Crisis" draw-up are not here since the market did not yet recover.

Also, on the same figure, one can notice that compared to the time to recovery from the crash of 1929, recent drawups appear to last for a shorter time. We could say that it is highly dependant on the type of the crises, in particular on the sector where it occurred (mortgage industries, automotive industries, small retailer...), or whether it is related to a particular event (terrorist attack of the 9/11, the coronavirus pandemic...).

One can also argue that this could be explained by more effective government responses to the crisis. Indeed, the time to the trough and the time to recover from the crash are highly dependent on a government's ability to adopt economic measures such as injecting liquidity into its economy, as well as its ability to maintain the previous unemployment rate[7].

Nevertheless, more recently we can see on the graph that the two big drawdowns in 2000 and 2008, following respectively the Internet Bubble Crisis and the Financial crisis, have larger associated recovery periods. We can think that the high growths of the economy, from the advent of internet, as well as the increasingly speculative market, leads to a bigger and more global impact on the economy. Thus, we can suppose that these severe crises could be harder to recover.

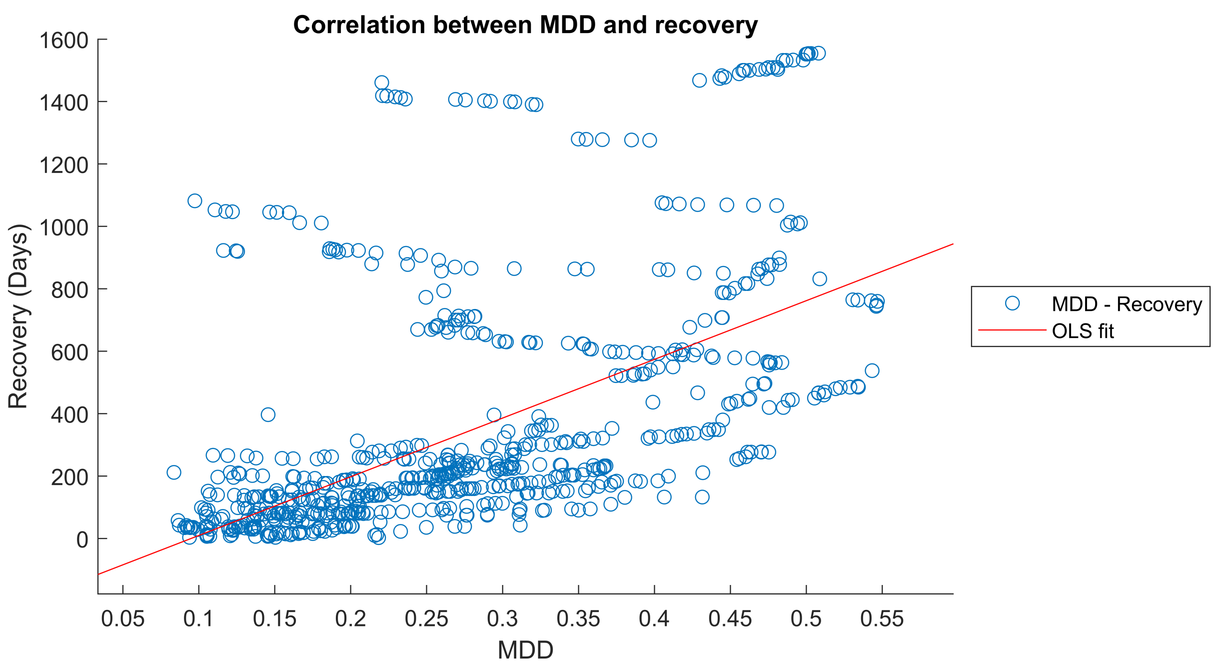

Figure 4.6: Correlation between recovery and MDD with an OLS fit provided my Matlab function "lsline".

[7]This can clearly be extrapolated to the actual corona crisis. Indeed, we can see that the governments and central banks are learning for their mistakes.

Another interesting point to mention is that the speed of the drawdown does not reflect its severity. Indeed, when looking closely at the graph, we observe that the speed of the drawdown related to the financial crisis is less important than the one of the Internet bubble, although it had a larger impact on the global economy. We explained this by an unprecedented interconnection between financial markets, which, therefore, makes crises spread faster and more easily. Typically the spread and speed of the systemic risk is increasing with more interconnected markets [8]

This leads us to our following point. Indeed, after having compared the speed of recovery to one of drawdowns, we are now interested in the correlation between the magnitude of the MDD and the time until recovery. Hereunder, we graphically represent this relation. Not surprisingly, one can observe a positive relation. It implies that an economy will take more time to recover from a drawdown as it becomes more severe.

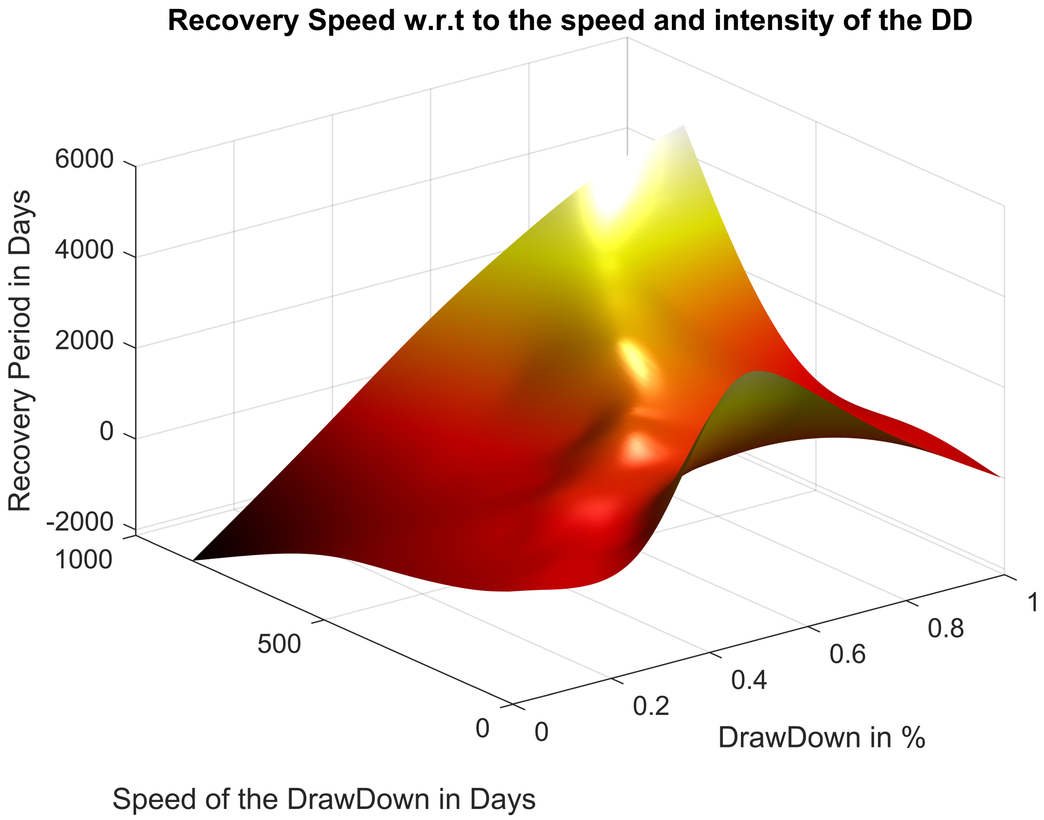

Figure 4.7: Surface of the recovery speed w.r.t the speed and intensity of the MDD. To create this graph, data were interpolated with a relatively high smoothing parameter.

The graph above, which is a summary of the latest two, shows us the relationship between

[8] You can find a lecture from the BIS on the management of the financial crisis in an inter-connected world here : Managing Financial Crisis in an Interconnected World: Anticipating the Mega Tidal Waves

the intensity, the speed of drawdowns, and the recovery speed of drawdowns. We can observe that in the last hundred years, the relationship has a shape of a basin and we can relate it to major events. For instance, as we noticed previously, the peak for fast and median intensity drawdowns can be related to the early years of the 21's century with the internet bubble crisis and those imply a longer recovery.

An attentive reader will notice the presence of some negative recovery periods. The underlying explanation comes from the lack of cleverness of the interpolation algorithm. There were no very small DD that have a duration of more than 500 days, and therefore, the algorithm had no benchmark[9].

4.5 Allocation and Contribution to risk

In this section, we will start by looking at the contribution to CED of the assets in different allocation strategies. Then, we will look at the way the CED reacts to an allocation between the market portfolio and the risk-free rate.

4.5.1 Equally Weighted Portfolio

In this part, we use the ten industry portfolio of Fama-French to construct an equally weighted portfolio which will allow us to look at the contribution to the risk of each industry. The two following graphs represent these contributions to risk.

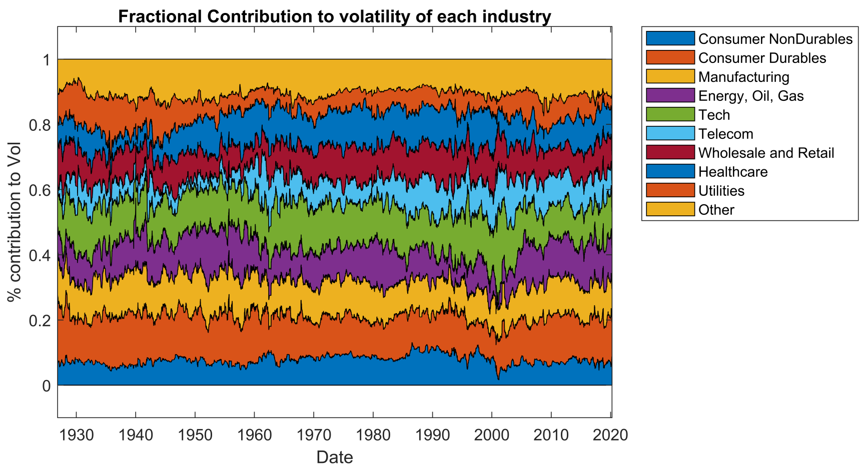

Figure 4.8: Volatility contribution of each industry in an equally weighted portfolio.

[9]The algorithm used is the following : John D'Errico (2020). Surface Fitting using gridfit, MATLAB Central File Exchange. Retrieved May 28, 2020.

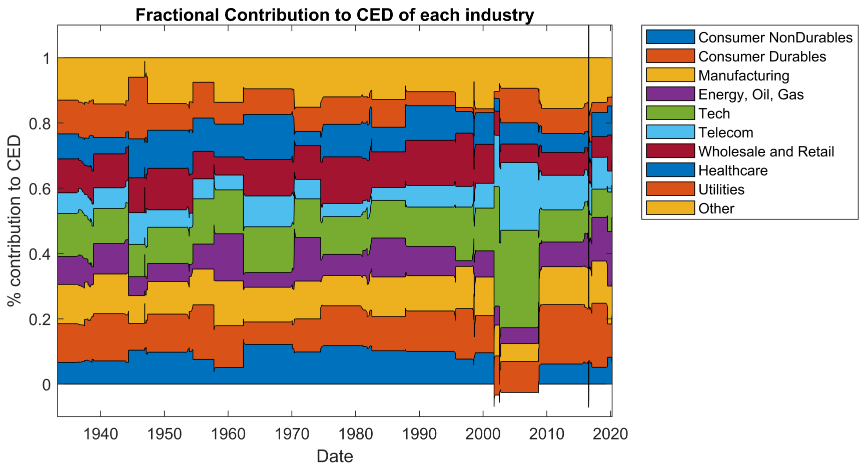

Figure 4.9: CED contribution of each industry in an equally weighted portfolio. The contribution is computed by numerically differentiating the risk measure which introduce noise on the estimation.

On these graphs, we can see some interesting features. Firstly, as we could imagine, the contribution to the risk of each risk measure are not equally weighted. Moreover, we can see that the contribution to the volatility is highly noisy whereas the contribution to CED is way more in "stairs". This is inherent to the way risk measures are calculated. Another interesting feature is to see that when there is high volatility for an industry, the move seems amplified in the CED. Typically, we can see the dotcom crisis has made the "Tech" industry counting for almost 30 % of the total contribution to CED. Finally, we can see that the "utilities" industry had a negative CED contribution after the dotcom crisis. It means that this industry had interesting contra-cyclic properties.

4.5.2 Risk Parity Portfolio

A risk parity portfolio is constructed in a way that makes every assets enter in the portfolio with the same volatility. The portfolio is constructed in the same way as eq. 3.18, 3.19 and 3.20 but with the volatility as a risk measure. Therefore the optimization problem is the following:

$$L(w_t) = \sum_{i=1}^{N_t} log(w_t^i) - \lambda(\sigma_{w_t} - \sigma_{Target}), \forall t \tag{4.7}$$

Since the volatility, as the CED, is homogeneous of degree one, all the weights can be re-scaled post-optimization in order to sum up to one.

On the following graphs, we can observe the contribution to risk for the volatility and the

CED, by construction the contribution to volatility will be equalized.

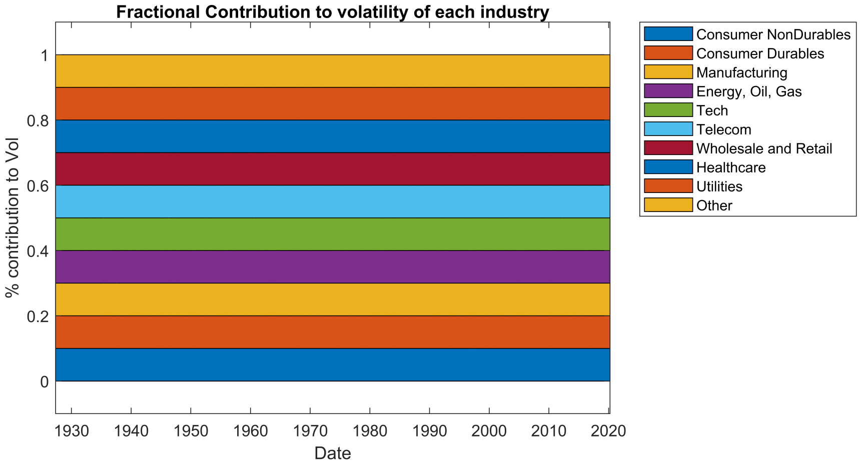

Figure 4.10: Volatility contribution of each industry in the risk parity portfolio. By definition, the portfolio is constructed in a way that make every asset enter the portfolio with the same volatility.

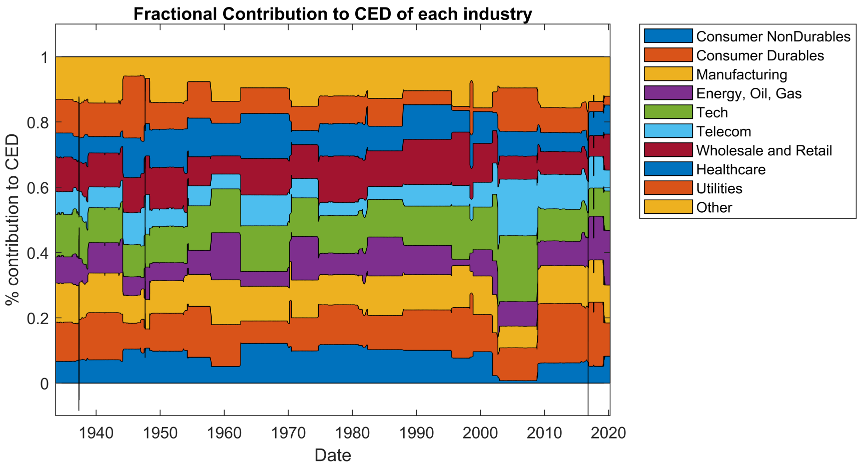

Figure 4.11: CED contribution of each industry in the risk parity portfolio.

The volatility is the most commune measure of the risk in the financial sector. As one can see, risk definition is not universal and there exist a lot of different interpretations. Since we computed a risk parity, we can see in Figure 4.10 that all the industries have equal contributions to the volatility of the portfolio. However, the CED contribution for each of the assets and for different industries, in Figure 4.11, are not equally weighted. This is a good example of what can be the work within a risk management team, which implies that different thoughts can result in very different risk allocations. Indeed, risk managers will have to interpret the risk aversion of their client and choose wisely different risk measures in order to satisfy the client's need. We will look at an example in the next chapter, where we will compute different risk measures under a basic portfolio allocation with two assets, one risky and one risk-free.

4.5.3 Market Portfolio

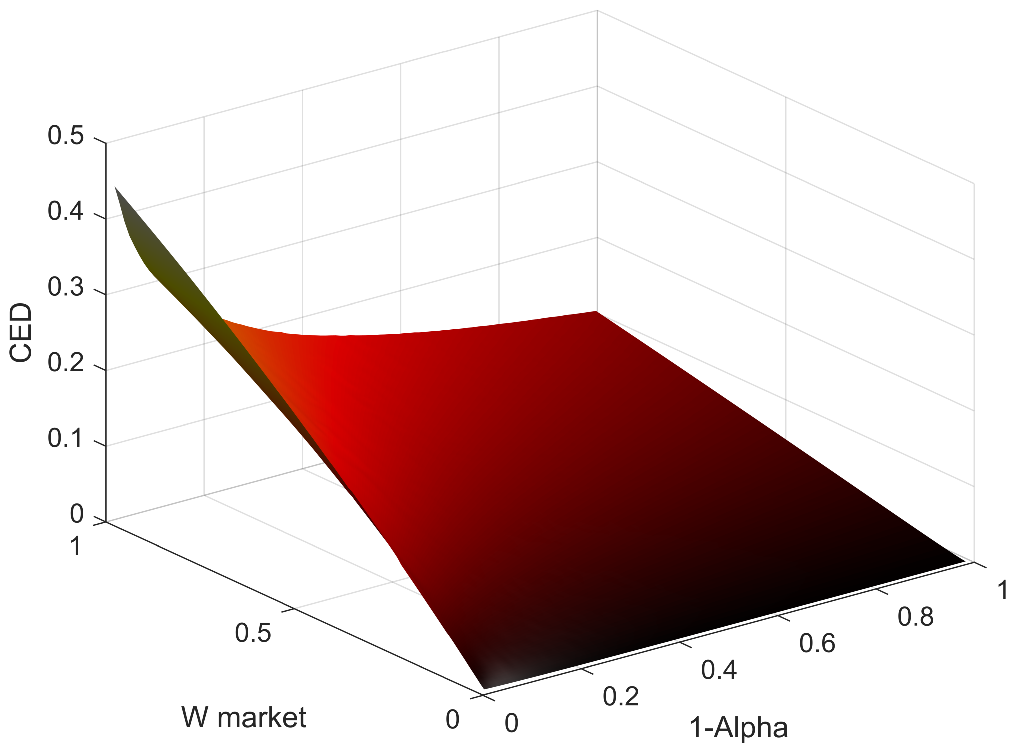

Figure 4.12: Surface of the CED relative to the alpha threshold and the size of the market portfolio.

We have computed two asset allocations with a rolling path of 6 months using the Fama-French market portfolio and the risk-free rates since 1929. Thereafter, we obtained the CED for different weight allocation between the two assets and for different thresholds. The graph above summaries the results of the latter computation.
Naturally, the CED is developing along with the increase in weight in the market portfolio. Obviously, the more risks we take, the more the risk parameter should increase. This is interesting in risk management because it can be used to establish the dependence of personal risk on the economic environment or the liquidity risk of a company. For companies with less liquidity, a smaller alpha would be more appropriate, as it would shift

the CED higher, and as well a lower weighting in the market portfolio would be preferable.

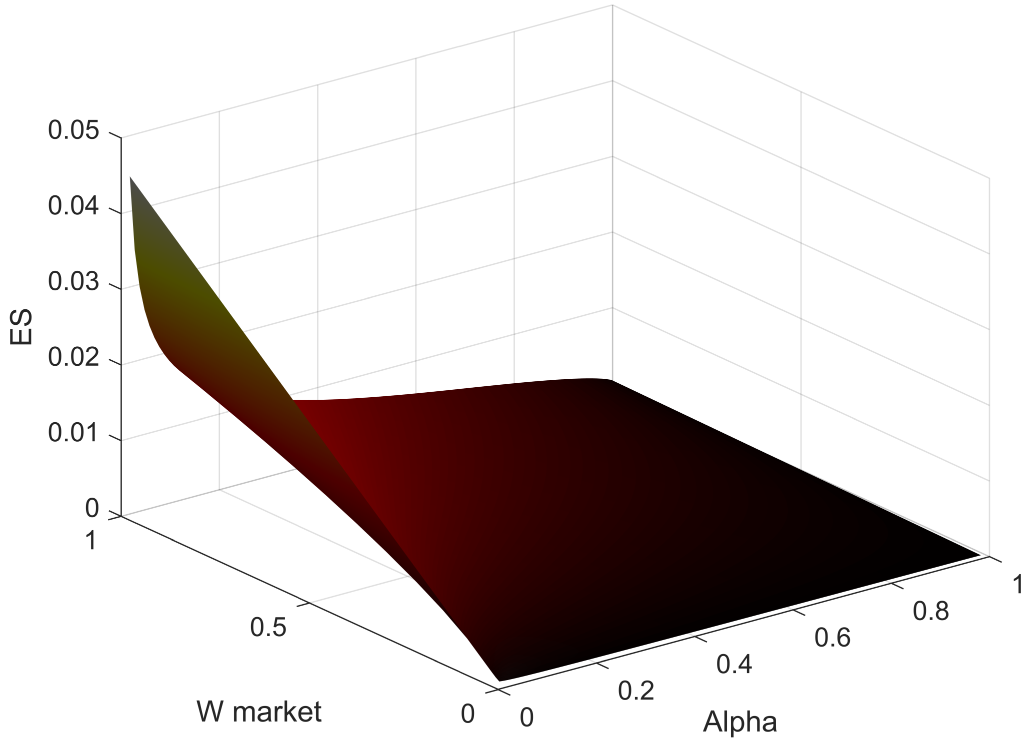

Figure 4.13: Surface of the Expected Shortfall in percentage relative to the alpha threshold and the size of the market portfolio.

In order to compare different risk measures in a portfolio, we computed the same procedure as for CED in Figure 0.13, but this time for the expected shortfall (ES). As we can see, the ES is more dependant on the quantile alpha than on the market weight. While the CED considers that even small proportions of the market in the portfolio is a risk, the ES does not. Indeed, we can observe on the graph above, which shows the ES for different portfolios and alphas, that the smaller the market weights, the smaller the risk. In fact, the risk is large only when alpha is very small.
The difference is the perception of the risk in both cases. For the CED, it will have a better insight of the cumulative of bad returns, and therefore, on the risk of liquidation or the incapacity of holding a position any longer. On the other hand, ES focuses more on the risk of having one single bad return. And thus, it explains more the risk of having a bad day during the time of one's investment (here we have a 6 months rolling returns), and depending on that risk, we can manage the allocation.
Finally, both measures are complementary and can help manage risk depending on the

investment strategy. For instance, a company would be less risk-averse on the ES as they don't care about bad days as they have enough liquidity to resist them. In consequence, they will more pay attention to the final returns. Moreover, they could have a look at the CED, so they have an idea of their risk of being out of business. On the contrary, an individual that has not enough liquidity in the bank will be worried about the ES, as maybe he will not be able to support a large negative return and therefore take less risk and invest in the risk-free rate.

5 Discussion

5.1 A CED parity allocation

As presented with the equations 3.18, 3.19 and 3.20, the properties of the CED allows to construct a *CED Parity Portfolio*. However, the computations are relatively more involved than in a risk parity portfolio. Hence, the computational complexity of the algorithm is higher, and therefore, the optimization takes a lot of time. Moreover, since the estimation of the CED is a highly non-linear process, the numerical derivation are not as stable as for a risk parity portfolio. That being said, we constructed this so-called *CED Parity Portfolio* for five years in order to show his properties.

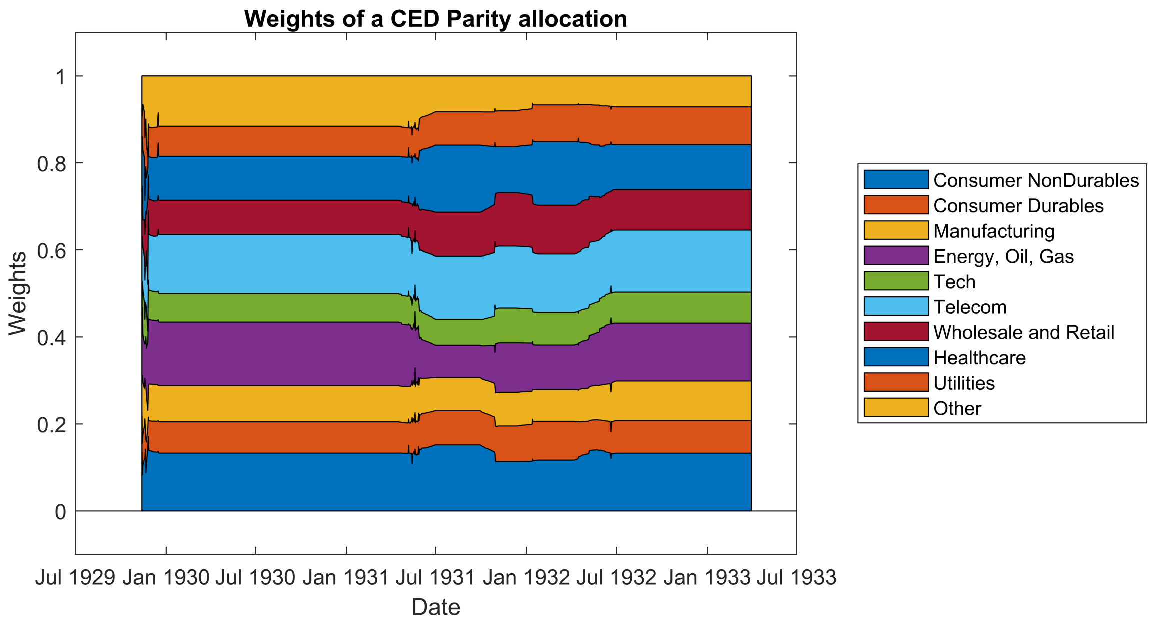

Figure 5.1: Weights of the CED parity allocation for the 10 Industry Data between 1928 and 1933

The figure above shows the weights allocations for this portfolio. As we can see, the weights are relatively stable since the same drawdown has an impact on the CED at different times. Therefore, this strategy would probably be cheaper to implement than a risk parity allocation. The following graphs shows the contribution to CED of this allocation.

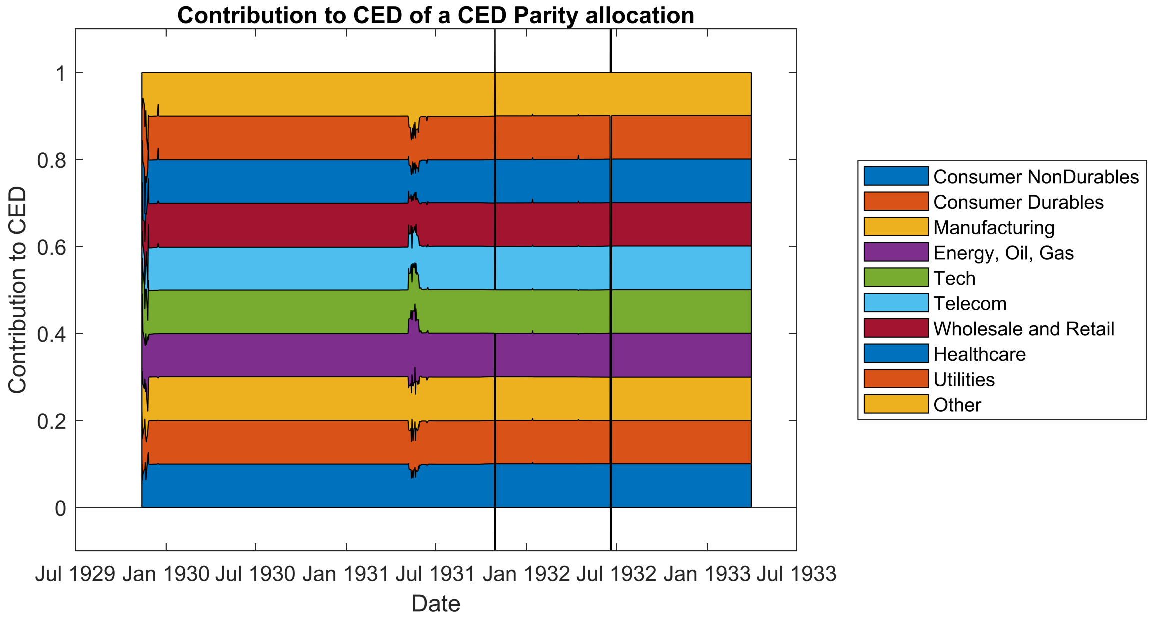

Figure 5.2: Contribution to CED of the CED parity allocation for the 10 Industry Data between 1928 and 1933. The contribution is computed by numerical differentiation which introduce some noise into the estimation.

As we can see, the contribution to CED is not as "perfect" as in the case of the risk parity portfolio (fig. 4.10). This is due to the fact that the computations are done by numerical derivatives. However, one can observe that most of the time, contribution are equalized between the assets. This shows that one can construct such a portfolio, however, we did not test his performance neither compare it to benchmark portfolios, this is out of the scope of this article.

5.2 CED for high frequency Trading

As seen through our empirical analysis on the relation between CED and the frequency of observations, CED and Maximum Drawdown are sensible to this frequency. Hence, this risk measure can be useful for high-frequency trading. Indeed, there can be some massive intra-day moves that even daily data do not capture. To look into, we have used Bitcoin intra-day data (minutes by minutes) in order to see if the CED was impacted by a change of frequency (minutes, hours, and days). The following plots show the distribution of drawdown and the CED for each frequency.

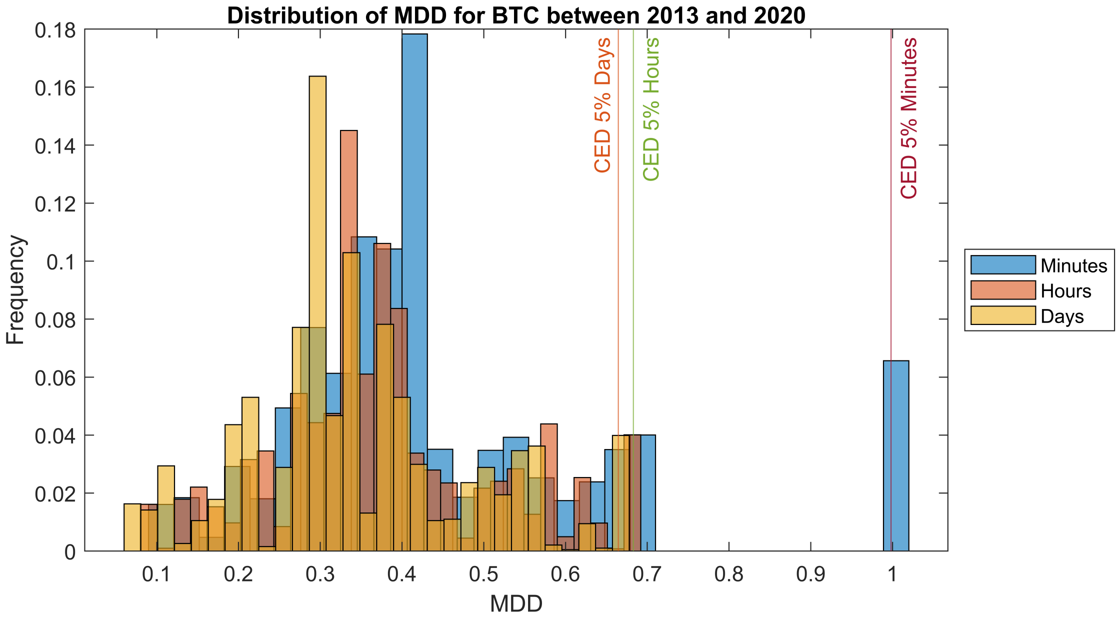

Figure 5.3: Distribution of drawdown for a path length of 125 days and CED at 5% for minutes, hours and daily frequency of Bitcoin

As we can see, when using minutes frequency, the CED captures massive movements of the Bitcoin price that even hours frequency does not capture. One must note that such moves happened especially at the beginning of the life of bitcoins when the later started to be actively traded on the market. This particularly important phenomenon could probably not happen on more standard financial markets. However, we do see that this measure of risk can be useful in higher frequency.

The typical example is the flash crash of 2010, using daily data one would not capture these movements[10]. However, the impact of such a crash can be real with investors being forced to liquidate positions. Therefore using CED for high-frequency trading or/and on high-frequency data can be interesting for the investors.

5.3 Economic soundness of this measure

As Goldberg and Mahmoud explain in their paper, they lay the basis for this risk measure in order to one day incorporate it in the investment process. In our paper, we show the relation between CED and different parameters. The conditional expected drawdown is

[10]Madhavan, A. (2012). Exchange-traded funds, market structure, and the flash crash. Financial Analysts Journal, 68(4).

definitely an interesting risk measure and seems to have interesting properties, not only on the mathematical part but also in a more intuitive way. However, the process to compute this risk measure yields some questions that one could try to answer in a next paper on the subject :

- The CED is the tail mean of the distribution of maximum drawdown. This distribution is computed on a one day rolling window, which allows us to have $N - T$ data (N = Number of observations and T = path length), which can be relatively big. However, is it really sound to take into account T times the same day in the computations ? Indeed, in an extreme case, a one day crash could be the maximum drawdown of T rolling window and therefore appear T times in the distribution.
- Creating a strategy using CED might not be as intuitive as it may seem. Indeed, we could think of any strategy that would weight the assets with respect to their CED or a strategy that would have a constraint on the CED. However, would it really yield a good performance compared to other benchmarks portfolio ? The question is worth asking. Indeed, the CED would be higher after some crashes, and therefore, we would underweight assets that just went down. With random walk hypothesis, this would not be a problem but intuitively, we could think that the probability of entering into another crash after a crisis would be lower. Hence, by doing so, we will end up by underweight assets that would have a lower probability of going down and thus, obtaining lower returns.

Hence, we do think that in another paper, it might be interesting to look at these questions and try to answer them. Obviously, the CED is a mathematically coherent measure of risk. However, it is not necessarily sound from an economic perspective and looking at the performance of this risk measure in portfolio allocation could be an interesting task.

6 Conclusion

Through this paper, we have discussed the usefulness and the importance of the maximum drawdown (MDD) in Risk Management. Offering an illustration of a worst-case scenario to investors, we have seen that the MDD is a better risk indicator than the volatility to capture the downside risk of an investment. We also analyze the Conditional Expected Drawdown (CED), introduced by Goldberg and Mahmoud, which corresponds to the tail mean of maximum drawdowns distributions.

We studied the characteristics of MDD and CED regarding the length of the dataset, the frequency of observations, and its influence on autocorrelations. These two measures of risk go pairwise. When the MDD is higher, it automatically shifts the CED upward. Therefore, the consequences that are true for one hold for the other. We have seen that increasing both, the path lengths and the frequency of observations, result in a larger CED. Then, through simulated data, using an AR1 process, we also found that CED is really sensitive to autocorrelation. This highlights one of the advantages of using MDD against standard deviation, for example.

Besides the risk associated with the magnitude of MDD, we aim to extend the literature, considering another risk factor related to MDD, which temporal dimension of risk, namely it's duration. By looking at both, the duration and the recovery time of MDD, we first found that the speed of a recovery is not directly linked to the speed of the drawdowns. A faster drawdown does not necessarily imply a faster recovery. We rather think of it as a process that varies from an industry to another. In a second time, we showed the relation between the magnitude of the drawdowns and its recovery time. In fact, we found that the more severe the drawdowns are, the longer it will take to return to its previous peak.

Finally, in section 4.5, we discuss the implementation of MDD, more specifically of CED, into an asset management framework. We considered two portfolio allocation, the equally weighted portfolio, and the risk parity portfolio. For an equally weighted, when comparing the risk contribution of volatility to the one of CED, we find that the latter is more stable through time but still coherent with the trend of the market (see Internet Bubble crisis in 2000). While we found an improvement using CED in the equally weighted portfolio, it is no more the case for the risk parity portfolio.

In section 5.3, we illustrated the consequence of asset allocation within the market portfolio and the risk-free rate on the CED and the ES. While the CED takes into account the autocorrelations feature of the returns by considering the risk of having consecutive bad returns, the ES only focuses on the occurrence of one bad return. Therefore, using one or the other risk measures will lead to different asset allocations, each one depending on the risk aversion of the investor.

7 References

[Goldberg Mahmoud 2017] Lisa R. Goldberg Ola Mahmoud. Drawdown: From Practice to Theory and Back Again, Mathematics and Financial Economics, Volume 11, Number 3, 2017.

[Mahmoud 2016], The Temporal Dimension of Risk. Journal of Risk, Volume 19, Number 3, 2017

[R. Tyrrell Rockafellar], Coherent Approaches to Risk in Optimization Under Uncertainty, 2007

[M. López de Prado], How Long Does It Take to Recover from a Drawdown? (2013).

[Choi Jaehyung], Maximum Drawdown, Recovery and Momentum, 2015

[Carl R. Bacon], Practical Risk-Adjusted Performance Measurement, 2013

[Marat Molyboga Christophe L'Ahelec], Portfolio management with drawdown-based measures. The Journal of Alternative Investments Winter, 2017

[Magdon-Ismail, Malik and Atiya, Amir F], Maximum Drawdown. Risk Magazine, Vol. 17, No. 10, 2004

[Alexei Chekhlov Stanislav Uryasev], Drawdown measure in portfolio optimization, International Journal of Theoretical and Applied FinanceVol. 08, No. 01, 2005

Artzner, Delbaen, Eber, and Heath (1997), Coherent Measures of Risk, Mathematical Finance, 9(3), 203–228

Appendix