Scedasis in the cross-section and over time in the American and Dutch financial markets

Name student: Fayyaz Sheikh

Student ID number: 484438

Supervisor: Chen Zhou

Second assessor: Marina Khismatullina

Master Thesis Econometrics and Management Science (Quantitative Finance)

Erasmus School of Economics

Erasmus University Rotterdam

The Netherlands

Date final version: April 13, 2023

The content of this thesis is the sole responsibility of the author and does not reflect the view of the supervisor, second assessor, Erasmus School of Economics or Erasmus University.

Abstract

We investigate the tail risk, using the scedasis function, in serially dependent financial data from the American and Dutch stock markets. We use the heteroscedastic extremes theory from Einmahl et al. (2016) and apply the space-time trend model from Einmahl et al. (2022). In order to find which sectors have the riskier or safer stocks, we analyse both stock markets across different time periods. We look particularly at extreme and/or rare events such as the Great Recession or the Covid-19 pandemic and their impact on the large American and small Dutch stock markets. The extreme event increases the tail risk of the safer stocks more than that it increases the tail risk of the riskier stocks. Having homoscedasticity or heteroscedasticity in the tail does not make a distinction between the high and low tail risk stocks. For most sectors, a crisis does not have an impact on the frequency of extremes. There is no clear impact of the crises on the results for the test for a trend over time. In case of non-crisis periods, Dutch stocks are riskier than American stocks. In times of a crisis, the Dutch stocks become relatively safer than the American stocks.

Contents

1 Introduction

Extreme and/or rare events such as a financial crisis with extreme returns, war or pandemics will have an adverse impact on markets. These extreme events can translate into losses and are therefore a concern to market participants and policymakers. In order to minimize the losses caused by these extreme events for both investors and policymakers, it is important to get a better understanding of tail risk. To model the risk of these extreme or rare events, we investigate the tail region of a distribution and analyse its properties based on a data sample.

The risk for financial returns can be time-varying, especially in times of extreme events, for instance, the war in Ukraine or the Covid-19 pandemic. In addition, the impact of these events can be different in magnitude for different markets and sectors. An effect of the Ukraine war, is the disruption of the Dutch energy market as the energy supply from Russia is leading to scarcity, causing a spike in the energy price. The energy sector is directly impacted, however the magnitude of this event on the utility sector is relative low. The sectors that are more prone to these events are considered to be more risky and therefore we expect their tail risk to be higher.

The impact of this event also differs between countries. The Dutch economy compared to the American economy is fairly small and more dependent on energy delivery from Eastern-Europe. This is reflected in the different movements of the stock prices of the same sector in different stock exchanges, e.g. S&P 500 and the AEX. Indicating that the riskiness of similar sectors can be different for large markets compared to small markets, over different time periods. By having more insights and knowledge about the behaviour of the risk, investors will be in a better position to make decisions. We use the space-time trend model from Einmahl et al. (2022) to analyse the riskiness in different markets, across different sectors. This constitutes the research question: *Which sectors have more risky stocks according to the space-time trend model?*

We investigate the tail risk in the cross-section and over time. In particular, what the behaviour is in multiple markets and which sectors are more risky than others. By finding patterns for the behaviour we can answer questions like the consistency of the tail risk or the riskiness of the stocks. Einmahl et al. (2016) assume that the tail risk can be different over time. This is referred to as heteroscedastic extremes. For their application on financial data, they assume a common extreme value index across the stocks. Therefore the difference is in the scale parameter which they name scedasis.

To test the difference in scedasis over time and in the cross-section we use the space-time trend

model from Einmahl et al. (2022). In their paper they use data from rainfall in Germany, but in this paper we use financial data from the S&P 500 and the Dutch markets AEX and AMX. Using financial data and coping with its features makes this paper interesting for readers and researchers in practical applications. There are some different challenges with the financial data compared to the data of the rainfall, which is used in Einmahl et al. (2022). Contrary to rainfall data, financial data is serially correlated and heavy tailed. Other assumptions of the method are scrutinized to observe whether they hold. For example, the model requires a common extreme value index across all stocks. However, we need to test if this is a valid assumption or not. Hence, the space-time trend model needs to be adjusted to cope with financial data. In contrast to Einmahl et al. (2022) we do not need to consider spatial dependence, since stocks are not measured in geographical space.

An extreme event causes the tail risk of the safer stocks to increase more than the tail risk of the riskier stock. Having homoscedasticity or heteroscedasticity in the tail does not make a distinction between the high and low tail risk stocks. The low presence of heteroscedasticity in the tail is caused by the nature of the stocks. For some stocks, their scedasis is always higher but the scedasis does not change much. The impact of the Great Recession is bigger than Covid-19, because of the response time of market participants.

The test for each subsample whether the frequency of extremes is constant in the cross-section shows that the Human Genome Project has a substantial impact on the health research sector that makes us reject the null hypothesis of a constant total integrated scedasis of extreme returns in the cross-section. The frequency of extremes in the agricultural sector is not impacted by a financial crisis or pandemic. We conclude for the consumer discretionary, technology, health consumer materials, industrial and energy sector that the frequency of extremes is similar for each stock, irrespective of the time period. Hence, a crisis does not have an impact on the frequency of extremes for these five sectors. The test for a trend over time shows that for two sectors the frequency of extremes is constant over all the sample periods and for two sectors this is not constant. For the remaining three sectors the frequency of extremes is different per time period. There is no clear impact of the crises on the results of the latter test. The impact of both of the crises is different for each of the sectors and thus also for the frequency of extremes.

For the American and Dutch comparison we find p-values of the Kolmogorov-Smirnov test close to 0.000 for all the Dutch stocks. Hence, the frequency of extremes is not constant over time for all these stocks. In case of non-crisis periods, Dutch stocks are riskier than American stocks. For both extreme events, we conclude that the Dutch stocks become relatively safer.

The paper is structured as follows: Section 2 brings together earlier findings in the field of extreme value theory and tail risk. Section 3 describes the datasets. The space-time trend model is explained in Section 4. Section 5 states the results of this model, the details and evidence why some stocks or sectors are more riskier than others. The conclusion and discussion of this paper are given in Section 6.

2 Literature

In recent years, as a consequence of extreme events such as stock market crashes and high swings of stock prices, the Extreme Value Theory (EVT) attracted more attention in research. Section 2.1 states the EVT framework. An extreme event is defined as an event, that occurs with low probability, but once it occurs has high impact. We try to find patterns in these extreme events, therefore we are searching for a trend. In Section 2.2 we discuss the literature on extreme value statistics with a trend. A case in point is the financial crisis of 2007–2009 and the resulting economic downturn. Another example is the Covid-19 pandemic. Extreme value statistics in finance is discussed in Section 2.3.

2.1 Extreme value theory

Extreme Value Theory is a branch of statistics that deals with the modelling of observations that have a high deviation from their mean. The ultimate aim of EVT is to estimate the probability of an extreme value occurring. For example, Gilli et al. (2006) measure the financial risk with several major stock market indices. McNeil (1997) and McNeil (1998) investigate the estimation of extreme risk in financial time series and Embrechts et al. (1999) shows robustness of EVT in risk estimates. McNeil and Frey (2000) extend the analysis of extreme risk using heteroscedastic financial time series.

The Fisher-Tippett-Gnedenko theorem (Coles et al. (2001)) and the Pickands-Balkema-de Haan theorem (Balkema and De Haan (1974) and Pickands III (1975)) form the basis for EVT. The Fisher-Tippett-Gnedenko theorem states that, when the random variables are independent and identically distributed (iid), after normalization the distribution of the maximum can only converge to the Fréchet, Gumbel or Weibull distribution. The Pickands-Balkema-de Haan theorem models observations that exceed a prespecified threshold. This theorem states that for a large enough threshold, given that the random variables are greater than this threshold, the distribution of its

exceedances converges to a Generalized Pareto Distribution, Coles et al. (2001).

2.2 Extreme value statistics with a trend in the cross-section or over time

EVT is the probability theory which is the foundation of extreme value statistics. Often, there is the assumption that observations are iid. However, in reality this assumption is often violated. Einmahl et al. (2016) acknowledge this and they introduce a theoretical framework to address the issue that observations are drawn from different distributions, also referred to as time-varying frequency of extremes. The objective of their framework is to, via a positive scedasis function, model changes in the tail of a distribution function continuously over time while the extreme value index remains fixed. In addition, they establish the asymptotic behaviour of the scedasis function, extreme quantile and extreme value index. Finally, they develop a test for the scedasis function to determine whether the extreme value index is constant over time. The test for a constant extreme value index, gives us an indication whether we can use the space-time trend model.

In the case of Einmahl et al. (2016), the probability distribution is characterised by a shape and scale parameter. Davison and Smith (1990) do a parametric approach where they model covariates in these parameters. The shape parameter is better known as the extreme value index or tail index. The extreme value index is often considered the most suitable in analysing the tail behaviour. This leads to using a semi-parametric approach. To model a non-parametric trend in the extreme value index, de Haan and Zhou (2021) propose a semi-parametric approach. In this paper, we will only focus on the extreme value index parameter, so a semi-parametric approach will suffice.

Cabral et al. (2020) extend on Einmahl et al. (2016) by using a scedasis function which connects the latent stationary process with marginal tail distributions. This causes the scedasis function to consist of two parameters, instead of a scedasis function with only one parameter as in Einmahl et al. (2016). Ferreira et al. (2017) make another extension on Einmahl et al. (2016), where they allow the extreme value index to be any real value instead of only positives values. For our financial data, we can use the extension from Ferreira et al. (2017). The reason is that in our case the extreme value index can be negative as well.

Einmahl et al. (2022) apply the same test for a constant extreme value index over time as in Einmahl et al. (2016). Einmahl et al. (2022) show that the estimation of the extreme value index can be done via a pseudo-maximum likelihood estimator (MLE). Einmahl et al. (2022) extend the work of Einmahl et al. (2016) to a space-time trend model, where they propose two test statistics. These two test statistics are the Kolmogorov-Smirnov-type test statistic and the Cramér-von Mises-

type test statistic. The first null hypothesis is that there exists homoscedasticity in the tail, so a constant scedasis, over time. The second null hypothesis states that there exists homoscedasticity in the tail throughout space, where space for our financial data can be seen as the cross-section of the loss returns. We will use the Kolmogorov-Smirnov-type test statistic to test the two hypotheses.

Mefleh et al. (2020) do a parametric estimation on the scedasis function, rather than the non-parametric estimation in Einmahl et al. (2016). Mefleh et al. (2020) provide a parametric test, which is used for the detection of a trend. This parametric test is valid in case of a monotonic scedasis function. This parametric test has the possibility to be more powerful than the Kolmogorov-Smirnov test used in Einmahl et al. (2022). This is only the case when the scedasis function is monotonic. We do not apply the parametric estimation from Mefleh et al. (2020), because prior to the analysis, we presume that the scedasis function will not be monotonic.

2.3 Extreme value statistics in finance

In case of an extreme value index that is constantly changing over time, de Haan and Zhou (2021) come up with a nonparametric estimate. This estimate can be used for the functional extreme value index. For the loss returns of the S&P 500 in the data range 1988–2012, de Haan and Zhou (2021) do not reject the null hypothesis which tests for a constant extreme value index. However, for the longer data range 1963–2012 they reject the null hypothesis. This suggests that we need to test for a constant extreme value index over a full sample period, as well as for different subsample periods.

Building upon Einmahl et al. (2016), Bücher and Jennessen (2022) extend upon their findings by proposing and proving three things. First of all, testing for homoscedastic extremes via a bootstrap scheme which they prove is consistent. Secondly, taking care of the dynamics of extreme results, they suggest an extremal index. Lastly, they come up with a kernel estimator which we use for the scedasis function.

Massacci (2017) develops a time-varying POT model to investigate time-varying tail risk in the stock market. The developed approach uses score-based time-varying parameters. The author concludes that the sensitivity of the business cycle is more apparent for large companies compared to small companies. Kelly and Jiang (2014) provide a measurement to estimate time-varying tail risk. This measurement makes use of stock returns in order to find patterns in the stocks behaviour. They take a specific look at extreme events, e.g. stock market crashes and they conclude that the tail risk has significant predictive power for stock returns. Zhang and Zhang (2021) create a framework for time-varying tail risk. In their study they focus on the Chinese banking market and conclude

that for stock market crashes the dependence of these banks is significantly present. We try to add to this literature, by using the space-time trend model for the analysis of time-varying tail risk in serially dependent financial data.

Chaudhry et al. (2022) conclude that the tail risk in technology companies is higher compared to financial companies. Technology companies have a lower probability to negatively be affected by extreme events. A comparison with the other eight investment sectors (e.g. agriculture, health, utilities) is not made, which leaves a gap in the research. We fill this gap by comparing all ten sectors with each other.

3 Data

3.1 Datasets and variables

3.1.1 Datasets

In this paper we use two different datasets. The first one consists of all the stocks with a share code of 10 or 11 on the S&P 500 from CRSP, WRDS. Shares codes 10 and 11 are of common US based stocks. The focus is on the NYSE/AMEX/NASDAQ, with the data starting in January 2000 up and until December 2021. In this data range there are two crisis, namely: The Great Recession of 2007–2009 and the Covid-19 pandemic from 2020 up and until the time of writing. The 24 million daily data points are converted into weekly data, which then gives around 4.8 million observations for 12370 different stocks.

The second dataset is the AEX, consisting of nine stocks, and one stock (Corbion) of the AMX, with the same data range and frequency. At present, no company reflecting the agricultural sector is included in the AEX index. Corbion is selected to overcome this shortcoming. These ten companies have been constantly on the AEX and AMX index since January 2000. This gives a number of observations of 73606 for the daily data. For the weekly data it will be around 15000. This dataset is available on Yahoo Finance.

3.1.2 Variables

For the American data, we use the daily prices of the stocks to create the discrete weekly loss returns. The variable *HSICCD*, which presents the last industrial code, is taken as a proxy for the ten distinct sector groups. The first digit of *HSICCD* identifies the sector.

For the Dutch markets (AEX and AMX), we use the percentage decrease of the adjusted close price as the returns. The adjusted close price is the close price adjusted for stock splits and dividends. The ten companies will not have the complication of delisted returns or prices, because we only choose companies which have a complete history of prices and returns for the entire sample period.

3.2 Summary statistics

Table 1 shows the summary statistics of the declustered American weekly returns, where the declustering process is explained in Section 4.2. Applying the declustering method to each sector individually for the entire datarange leaves between three and 355 different stocks in each sector, except for the consumer staples sector. For the consumer staples sector over the full datarange, there are no stocks which have data points left after the declustering. For the smaller subsamples, we have observations for the consumer staples sector. Hence, we can make a comparison with all sectors, but not for the entire sample. In Table 1 we observe that for nine sectors the mean of the loss returns is between -0.003 and -0.007. The smallest maximum loss return of 61.4% is in the agricultural sector, which on first notice is considerably lower than for the other sectors. One might assume that in this sector there are more relatively safer stocks. However, this sector has the smallest amount of stocks within its sector.

Table 1: Summary statistics of the declustered weekly American returns of the S&P 500 for the entire datarange of 2000–2021

Sector	Mean	Max	Min	St.dev	Kurtosis	Skewness	Stocks
Consumer discretionary	-0.004	0.711	-24.350	0.111	19933.980	-106.021	139
Technology	-0.004	0.800	-12.803	0.111	4126.813	-45.331	124
Health research	-0.007	0.723	-37.203	0.274	9183.666	-84.460	44
Financials	-0.003	0.978	-13.833	0.094	7195.720	-64.220	269
Consumer staples	NA	NA	NA	NA	NA	NA	0
Health consumer materials	-0.005	0.930	-49.800	0.211	20383.440	-116.410	183
Industrials	-0.004	0.894	-20.519	0.110	8851.036	-64.111	355
Utilities	-0.005	0.855	-318.500	0.876	129685.5	-357.080	118
Energy	-0.005	0.672	-39	0.213	18398.34	-114.321	60
Agriculture	-0.003	0.614	-0.544	0.053	20.132	-0.016	3

Note: the values of the mean, maximum and minimum displayed are not in percentages. In order to get percentages, these values need to be multiplied by 100. For example, the maximum loss return in the financial sector is 0.978 or 97.8%. "NA" means not available.

Table 2 shows the summary statistics of the declustered weekly returns of the ten Dutch stocks, where the declustering process is explained in Section 4.2. In Table 2 we observe that the mean of the loss returns is between -0.000 and -0.004. The smallest maximum loss return is in the health research sector.

Table 2: Summary statistics of the declustered weekly Dutch returns of the AEX and AMX for the entire datarange of 2000–2021

Stock	Sector	Mean	Max	Min	St.dev	Kurtosis	Skewness
Ahold	Consumer discretionary	-0.002	0.654	-0.533	0.045	57.490	1.314
ASML	Technology	-0.004	0.226	-0.308	0.057	5.551	-0.103
DSM	Health research	-0.003	0.184	-0.150	0.036	6.145	0.359
ING	Financials	-0.002	0.421	-0.334	0.060	10.472	0.386
Heineken	Consumer staples	-0.002	0.191	-0.149	0.033	5.922	0.030
AkzoNobel	Health consumer materials	-0.002	0.207	-0.252	0.043	7.834	-0.063
Philips	Industrials	-0.002	0.245	-0.204	0.047	5.349	0.181
KPN	Utilities	-0.000	0.289	-0.375	0.053	11.192	-0.398
Shell	Energy	-0.002	0.309	-0.231	0.037	14.132	0.223
Corbion	Agriculture	-0.002	0.274	-0.262	0.040	10.046	0.203

Note: the values of the mean, maximum and minimum displayed are not in percentages. In order to get percentages, these values need to be multiplied by 100. For example, the maximum loss return for ING is 0.421 or 42.1%.

4 Methodology

Our aim is to compare the riskiness of each sector with each other. In this section, we explain the methodology we apply to answer our research question. We apply our methodology to each sector per time period individually. This means that we treat each of these sectors per time period as a separate dataset (subsample).

We try to avoid survivorship bias in our subsamples, by starting the analysis at the declustering step. By doing so, we allow that some of the stocks that are not part of the full sample period (because these stocks have missing values in the entire range), are now part of the subsample. Hence, the subsamples have the possibility to consist off more stocks than the full sample.

As discussed in Section 4.1, we have 100 different subsamples (ten sectors times ten time periods). Therefore, sector X does not influence the declustering of sector Y. This causes the time index (number of different weeks) after declustering, to be different for each subsample. For each subsample, we will only consider stocks which have all observations for a given time period. This causes that each stock has the same amount of observations within the subsample. The amount of observations is different for each of the subsamples that are used, but for the entire sample of 2000–2021 it is 1148 weeks.

The methodology consists of four parts. First of all, the explanation of the sample periods and the market comparison are in Section 4.1. Second, the declustering of the data and its difference from Einmahl et al. (2022) is explained in Section 4.2. Third, Section 4.3 describes the method to test for serial dependence, which is a key issue in financial data. Lastly, the assumptions of the space-time trend model are described in detail in Section 4.4. For the space-time trend model from Einmahl et al. (2022) we identify the riskiest stock as the stock that has the highest average scale parameter scedasis (C).

4.1 Sample periods and market comparison

In order to investigate which sectors have the more riskier stocks we apply this methodology over multiple sample periods. We investigate whether the behaviour of the tail risk is time-varying, so is the behaviour the same over the entire sample period or is the behaviour different in a specific time period. Our sample set consists of observations starting in January 2000 and ending in December 2021. Within this timeframe we have four periods, from which two are classified as extreme events:

1. Normal time period 1 (NT1), January 2000 up and until July 2007
2. The Great Recession (GR), August 2007 up and until November 2009
3. Normal time period 2 (NT2), December 2009 up and until November 2019
4. The Covid-19 pandemic (C19), December 2019 up and until December 2021

The analysis of time-varying tail risk is done over the entire data range, each of the four time periods separately and five combinations of time periods. The five combinations are: NT1 + GR (2000/2009), NT1 + GR + NT2 (2000/2019), GR + NT2 (2007/2019), GR + NT2 + C19 (2007/2021) and NT2 + C19 (2009/2021). Therefore, we can investigate how the addition of a specific time period influences the analysis.

As a final analyse we compare these riskiest and safest stocks from the American stock market with their Dutch counterpart. Since the scedasis is a relative measure, we add the Dutch stock to the American subsample before declustering. We name these subsamples Dutch subsamples, so the reader knows that these subsamples include a Dutch stock. We can then make a fair comparison, how the Dutch stock compares to the American stocks. For example, the American energy sector and its stocks are compared with the energy company Shell. This is done for all the different time periods.

With this approach we can answer and draw a conclusion on our research question. As stated in Section 1 this is: *Which sectors have more risky stocks according to the space-time trend model?*

4.2 Declustering method

After having selected a subsample, we apply the declustering method and then use this declustered data for the space-time trend model. We follow the approach of Einmahl et al. (2022) in the declustering method. The steps of the declustering to remove serial dependence are similar as those described in Einmahl et al. (2022). The five steps are:

1. For every week, we obtain the maximum loss return of that particular week.

2. We make a time series out of these maximum loss returns and name it the rowwise maxima series.

3. We order the rowwise maxima series from high to low and identify the first order statistic with rank one. We do this for every observation in the rowwise maxima series.

4. We then check whether the first and second order statistics are within r weeks of each other. If they are within r weeks, then we remove all observations that are on the same date as the lower order statistic of the two order statistics. Otherwise, both dates are kept.

5. We do this for all the other order statistics in the descending order. So, the k^{th} highest rowwise maxima is compared to all the previous $k-1$ order statistics. If the k^{th} highest rowwise maxima is not within r weeks of all the previous $k-1$ order statistics, we keep that week. We do this procedure till the end, when $k = n$.

There is one last parameter to be tuned, r. We assume that the dependency in extremes lasts for no more than r weeks. In de Haan et al. (2015) the authors explain that for rainfall data the dependency of extreme precipitation is at most two days and the clustering is not more than three days. Einmahl et al. (2022) follow this small window of two days. This window could be higher for financial data. The reason for this is that shocks in stock returns have an impact on the overall market and consequently on the behaviour of market participants.

Before the declustering we check the dataset for serial dependence. To see how long the returns are serially correlated we find that for our datasets the first few Autoregressive terms are significant for most of the stocks. The number of Autoregressive terms is often around four.

High volatility is being triggered by large shocks, hence a GARCH(1,1) is used to validate the existence of volatility clustering, as shown in Equation (1). Where, $\epsilon_{t\text{-}1}$ are the errors from the

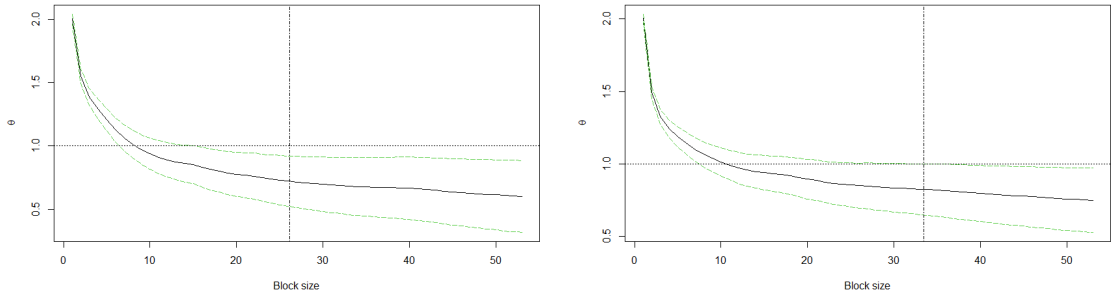
previous period and $\sigma_{\text{t-1}}$ is the volatility of that same period.

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{1}$$

When applying a GARCH(1,1) to check for volatility clustering, which is a sign of serial dependence, we find for every stock significant values of the beta coefficients. Hence, volatility clustering is present and therefore serial dependence. The existence of serial dependence is also confirmed with the Ljung-Box test. GARCH requires normality of the errors. This is generally not the case with financial stocks, as confirmed with the p-values of the test for leptokurtosis. Therefore, we consider the sliding block estimator from Berghaus and Bücher (2018), to investigate the magnitude of the serial dependence.

Leadbetter (1983) defines the extremal index (θ), which explains the serial dependence of extreme events. The range of θ is [0, 1], with a θ equal to one meaning serial independent data. Berghaus and Bücher (2018) state that the best block sizes are around $\sqrt{n}$, with n the number of observations. Figure 1 shows the block sizes and their corresponding extremal index, for two stocks. The vertical line represents $\sqrt{n}$, which reflects the best block sizes and which differs per stock. From Figure 1 we also learn that the serial dependence is at its lowest around $\sqrt{n}$. Therefore, we set our blocksize equal to $\sqrt{n}$. So, r is four is chosen after trial and error, as it proves to be the best value for these datasets.

Figure 1: Block schemes before declustering



(a) Stock 2

(b) Stock 6

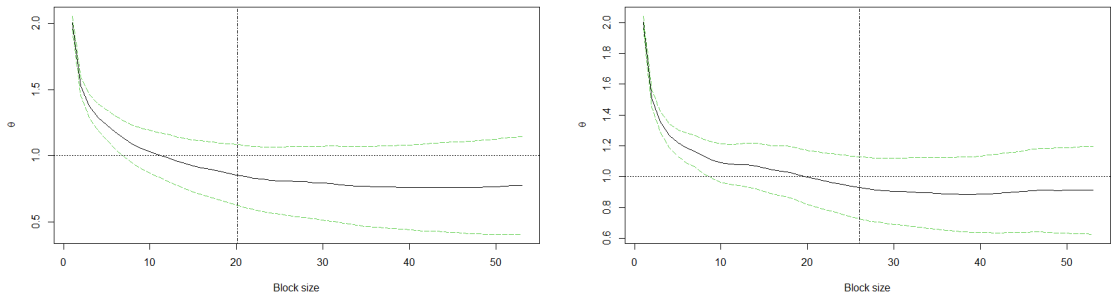
4.3 Serial dependence

The space-time trend model from Einmahl et al. (2022) requires independent data over time. Financial returns most of the time exhibit autocorrelation. In this section we explain how we manage this feature within our data. According to Einmahl et al. (2016) the daily data can suffer from serial dependence like volatility clustering. Two of the ways to tackle this problem are using data of a different frequency or the declustering method. In this paper we apply both approaches. The best frequency to capture the effect of these events is achieved by using data on a daily basis rather than on a monthly basis. The reason for this is that extreme returns on a daily basis are better detectable than on a monthly basis, where all daily observations are combined. For example in a month, there is one relative high return, one relative low return and the other nineteen returns are not considered extreme returns. With a daily frequency, we find two extreme returns, whereas for a monthly frequency we only get the average of all 21 observations. Therefore, there is the possibility that we do not notice an extreme return in the data. Einmahl et al. (2016) explain, that a robust method is to use less frequent data for instance, weekly or monthly returns. In this paper weekly data is used.

Financial returns are known to be heavy tailed and serially dependent. The total removal of the autocorrelation of the data is therefore a challenge. Taking the natural logarithm of the loss returns will not remove the autocorrelation completely. This transformation of the data can only diminish autocorrelation to a certain extent. The declustering method from Einmahl et al. (2022) is used, as it is more advanced and consequently should give better outcomes than a simple transformation. The declustering method is tweaked to a certain extent and explained in full detail in Section 4.2.

After declustering the data, the same four tests are repeated. This in order to see whether the declustering has reduced the serial dependency. According to the GARCH(1,1), volatility clustering remains present. The Ljung-Box test still confirms the existence of serial dependency. However, the test shows a small reduction in serial dependency. The amount of stocks whose p-value become significant for the test for leptokurtosis is marginal. When comparing Figure 1 with Figure 2, we observe for Stock 2 a minor improvement and for Stock 6 a fairly substantial improvement. For Stock 6 the line moves closer to θ is one. This pattern of a mixture of marginal and substantial improvements happens throughout the whole dataset. Implying that for some of the stocks the remaining serial correlation is weak, whereas for others it is still substantially present.

Figure 2: Block schemes after declustering



(a) Stock 2 (b) Stock 6

4.4 Space-time trend model

We explain the space-time trend model in this subsection. In Section 4.4.1, we provide the heteroscedastic extremes framework, together with the scedasis function. We estimate and test for a constant extreme value index across the cross-section and over time in Section 4.4.2. The kernel estimator that is used to estimate the scedasis function is described in Section 4.4.3. The estimation and testing of the heteroscedastic extremes are explained in Section 4.4.4 and Section 4.4.5. Section 4.4.6 and Section 4.4.7 provide the estimation and testing of the space-time trend model.

4.4.1 Heteroscedastic extremes framework and scedasis function

EVT assumes that observations are iid. This assumption is violated in finance, since the returns of stocks are serial dependent and have different distributions. The heteroscedastic extremes framework from Einmahl et al. (2016) relaxes the assumption of independent and identically distributed observations, it allows for possible non-identically distributed observations. For n independent observations over time we have the observations $X_1, X_2, ..., X_n$ that follow different continuous distribution functions $F_{n,1}, F_{n,2}, ..., F_{n,n}$. These continuous distribution functions have a right end point x^* in common, with x^* as shown in Equation (2) and similar to Einmahl et al. (2016).

$$x^* = sup\{x : F_{n,i}(x) < 1\} \in (-\infty, \infty] \tag{2}$$

This continuous distribution function F shares the same right end point. As well as a continuous positive scedasis function c, which is defined on the interval of [0, 1]. Hence, Equation (3) follows from the established setup.

$$\lim_{x \to x*} \frac{1 - F_{n,i}(x)}{1 - F(x)} = c\left(\frac{i}{n}\right), \tag{3}$$

where it is uniformly for all $n \in \mathbb{N}$ and assume $1 \leq i \leq n$. We can define the scedasis function as an integral as in Equation (4). In this form, the function looks similar to a density function.

$$\int_0^1 c(s)\, ds = 1, \tag{4}$$

with c being the frequency of extremes. If c is equal to one we have homoscedastic extremes.

Furthermore, we assume that F belongs to the Domain of Attraction (DA) of a Generalized Extreme Value (GEV) distribution, with a extreme value index γ, which also is assumed in EVT. Meaning, a $\gamma \in \mathbb{R}$ exists, together with a positive scale function. So, for all $x > 0$ we get Equation (5).

$$\lim_{t \to \infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \tag{5}$$

where U is defined as in Equation (6).

$$U := \left(\frac{1}{1 - F}\right)^{\leftarrow}, \tag{6}$$

with the symbol $\leftarrow$ representing the left continuous inverse function.

When combining the condition of the DA in Equation (5) with the condition on proportional tails as in Equation (3), we get the non-parametric model where the extreme value distributions have the same extreme value index γ. In the heavy tailed case, which means a positive γ ($\gamma > 0$), which is common in finance, Equation (5) simplifies to Equation (7). This γ is estimated via the Hill estimator, with more details on this in Section 4.4.2.

$$\lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^\gamma \tag{7}$$

4.4.2 Extreme value index

The models in Einmahl et al. (2016) and Einmahl et al. (2022) require the extreme value index to be constant. Hence, we test whether the extreme value index is constant across the cross-section and over time. For the estimation of the extreme value index (γ), we use the Hill estimator as in Equation (8). Where, X_i for $i = 1, 2, ..., n$, are independent observations. The threshold, so the $(n-k)$-th order statistic is represented by $X_{n,n-k}$. For the prove and conditions we refer to Einmahl et al. (2016).

$$\hat{\gamma}_H := \frac{1}{k}\Sigma_{j=1}^{k} log(X_{n,n+j-1}) - log(X_{n,n-k}) \tag{8}$$

The objective is to estimate the Hill estimator in multiple blocks and compare them. For subsamples we will split the full sample into different blocks. This will give us multiple Hill estimators, for each subsample. These estimators will be combined, as proven in Einmahl et al. (2016) and then the combination of multiple Hill estimators will be compared with the Hill estimator of the full sample.

The asymptotic behaviour of these partial Hill estimators is given in Equation (9).

$$sup_{0\leq s_1<s_2\leq 1, s_2-s_1\geq\delta}\left|\sqrt{k}(\hat{\gamma}_{(s_1,s_2)}-\gamma) - \gamma\frac{W(C(s_2))-W(C(s_1))}{C(s_2)-C(s_1)}\right| \to 0 \qquad \text{almost surely,} \tag{9}$$

with W reflecting a standard Wiener process on the range [0, 1]. Also, s_1 and s_2 are on the range [0, 1]. The extreme value index estimator is represented by $\hat{\gamma}$. Equation (10) shows the test for the comparison of the partial Hill estimator, with the full Hill estimator.

$$T_1 := sup_{0\leq s_1<s_2\leq 1, \hat{C}(s_2)-\hat{C}(s_1)\geq\delta}\left|\frac{\hat{\gamma}_{(s_1,s_2)}}{\hat{\gamma}_H} - 1\right|, \tag{10}$$

with the condition that $\hat{C}_s(2) - \hat{C}_s(1) \geq \delta$, where $\hat{C}$ is the scedasis estimator. For a fixed set of m blocks/subsamples we can use the test as in Equation (11).

$$T_2 := \frac{1}{m}\Sigma_{j=1}^{m}\left(\frac{\hat{\gamma}_{(l_{j-1},l_j)}}{\hat{\gamma}_H} - 1\right)^2 \tag{11}$$

The distribution of the two tests are shown in Equation (12).

$$\sqrt{k}T_1 \xrightarrow{d} sup_{0\leq s_1<s_2\leq 1, s_2-s_1\geq\delta}\left|\frac{W(s_2)-W(s_1))}{s_2-s_1} - W(1)\right|, \qquad kT_2 \xrightarrow{d} \chi^2_{m-1}, \tag{12}$$

where again W is a standard Wiener process.

4.4.3 Kernel Estimator

For the estimation of the scedasis function we start off with a kernel estimator. Estimation of a probability density function can be done with a non-parametric approach, such as kernel density estimation. This kernel estimator is given in Equation (13) and is the same as in Einmahl et al. (2016).

$$\hat{c}(s) = \frac{1}{kh}\Sigma_{i=1}^{n}\mathbf{1}_{\{x_i>x_{n,n-k}\}}K\left(\frac{s-\frac{i}{n}}{h}\right) \tag{13}$$

When $n \to \infty$ and the bandwidth $h := h \times n > 0$, then $h \to 0$ and $k \times h \to \infty$. The function $K\left(\frac{s-\frac{i}{n}}{h}\right)$ is continuous and a symmetric kernel on the interval [-1, 1]. This causes $\int_{-1}^{1} K(s)\, ds = 1$, where we set the condition that for $|s| > 1$, $K(s) = 0$.

For practical purposes Bücher and Jennessen (2022) create a kernel estimator which can be used if the data is serially dependent. Since the interest mostly is on the boundaries of the interval, we use a boundary corrected kernel K_b instead of K, see Bücher and Jennessen (2022) and Jones (1993).

For this boundary corrected kernel to be valid the following conditions from Bücher and Jennessen (2022) need to hold. We follow Bücher and Jennessen (2022), so for beta-mixing and blocking sequences, there are integer sequences for which $1 < l_n < r = r_n < n$. As well as, $l_n = o(r), r = o\left(\sqrt{k} \cup \frac{n}{k}\right)$. Then if n approaches infinity, both sequences converge. Therefore, the assumptions ($\beta(n) = o(1)$ and $\frac{n}{r}\beta(l_n) = o(1)$) for the beta-mixing coefficients can be satisfied.

When these conditions are met, the boundary kernel estimator works as shown in Equation (14) for $s \leq h$ and for $s \geq 1 - h$ as in Equation (15). The integrals of a and b are as in Equation (16). Note, this is similar as Bücher and Jennessen (2022).

$$K_b(x, s) = \frac{a_2(p) - a_1(p)x}{a_0(p)a_2(p) - a_1^2(p)x} K(x), \quad x \in [-1, 1] \tag{14}$$

$$K_b(x, s) = \frac{b_2(p) - b_1(p)x}{b_0(p)b_2(p) - b_1^2(p)x} K(x), \quad x \in [-1, 1] \tag{15}$$

$$a_i(p) = \int_{-1}^{p} x^i K(x)\, dx \quad b_i(p) = \int_{-p}^{1} x^i K(x)\, dx, \tag{16}$$

where $p \in [0, 1]$. Equation (13) is then adjusted to Equation (17).

$$\hat{c}(s) = \frac{1}{kh} \Sigma_{i=1}^{n} \mathbf{1}_{\{X_i > X_{n,n-k}\}} K_b\left(\frac{s - \frac{i}{n}}{h}, s\right) \tag{17}$$

4.4.4 Heteroscedastic extremes estimation

In this section we describe the estimation of the scedasis function as shown in Equation (18). In Section 4.4.5 we provide the tests of the scedasis function.

$$C(s) := \int_{0}^{s} c(u)\, du, \tag{18}$$

for $s \in [0, 1]$. Since we put emphasise on observations that exceed a certain threshold, the integrated scedasis function as in Equation (18) has to be proportional to the amount of threshold exceedances

for the first $[ns]$ observations. Similar to Einmahl et al. (2016), Equation (19) and Equation (20) show two conditions on a intermediate sequence $k = k(n)$.

$$\lim_{n\to\infty} k = \infty \tag{19}$$

$$\lim_{n\to\infty} \frac{k}{n} = 0 \tag{20}$$

Therefore we can define the estimator as shown in Equation (21).

$$\hat{C}(s) := \frac{1}{k}\Sigma_{i=1}^{[ns]}\mathbf{1}_{\{X_i > X_{n,n-k}\}}, \tag{21}$$

where $\mathbf{1}$ is an indicator function that equals one when $X_i > X_{n,n-k}$.

Einmahl et al. (2016) prove the asymptotic normality of the Hill estimator as shown in Equation (8) and also for the estimator for the integrated scedasis function as in Equation (21). Their conclusion, under a Skorokhod construction for a weak convergence to a probability measure, are Equation (22) and Equation (23).

$$sup_{0\le s\le 1}\left|\sqrt{k}\{\hat{C}(s) - C(s)\} - B\{C(s)\}\right| \to 0 \qquad \text{almost surely} \tag{22}$$

$$\sqrt{k}(\hat{\gamma}_H - \gamma) \to \gamma N_0 \qquad \text{almost surely} \tag{23}$$

In these two equations, N_0 means a random standard normal variable and B represents a Brownian bridge.

4.4.5 Heteroscedastic extremes testing

In Section 4.4.4 we provide the estimation technique, similar to Einmahl et al. (2016). In this section, the corresponding tests are explained.

We can test for homoscedasticity in the tail, so a constant frequency of extremes. As in Equation (24), the null hypothesis states that there exists homoscedasticity in the extremes. Whereas when we reject the null hypothesis, the alternative hypothesis states the presence of heteroscedasticity in the tail. A specific example is testing whether $c \equiv 1$ on the interval [0, 1]. If the null hypothesis is not rejected, we can state that C is the identity function.

$$\begin{aligned} H_0 &: C(s) = s \\ H_1 &: C(s) \neq s \end{aligned} \tag{24}$$

As stated in Section 2, Einmahl et al. (2016) propose two test statistics. The Kolmogorov-Smirnov test statistic is shown in Equation (25) and Equation (26) represents the Cramér-von-Mises test statistic.

$$T_3 := sup_{0\leq s\leq 1}|\hat{C}(s) - C_0(s)| \tag{25}$$

$$T_4 := \int_0^1 \{\hat{C}(s) - C_0(s)\}^2 \, dC_0(s) \tag{26}$$

These test statistics are based on the process as in Equation (27).

$$\left\{\sqrt{k}\hat{C}_j(s) - C(s)\right\} \xrightarrow{d} \{B(s)\} \tag{27}$$

The distributions of these test statistics are shown in Equation (28) and Equation (29), respectively. Again B is a standard Brownian bridge.

$$\sqrt{k}T_3 \xrightarrow{d} sup_{0\leq s\leq 1}|B(s)| \tag{28}$$

$$kT_4 \xrightarrow{d} \int_0^1 B^2(s)\, ds \tag{29}$$

4.4.6 Space-time trends estimation

In Section 4.4.4 and Section 4.4.5 we already provide some preliminary tests, for the evaluation of a homoscedastic trend in the tail. Einmahl et al. (2022) extend on Einmahl et al. (2016) for the multivariate case, where the observations are non-stationary over space and time. Space in this case means the cross-section of the stocks. This extension allows us to analyse the data in the situation where observations are dependent in the cross-section. Therefore we can test the dependence of the stocks in the cross-section at specific points in time, as in Section 4.4.7. This allows us to consider the stock returns in a multivariate setting, rather than a univariate setting.

When extending to a multivariate setting, similar to Einmahl et al. (2022) Equation (3) is extended to Equation (30), were we consider m stocks in the cross-section. We assume that the distribution of the random vectors, $(X_{i,1}, X_{i,2}, ..., X_{i,m},)$, of these m stocks is in the DA of a multivariate extreme value distribution.

$$\lim_{x\uparrow x*} \frac{1 - F_{i,j}(x)}{1 - F_0(x)} = c\left(\frac{i}{n}, j\right) \in (0, \infty), \tag{30}$$

where $i = 1, 2, ..., n$ with n the number of time points in weeks. The distribution function of $X_{i,j}$ is given by $F_{i,j}$. The scedasis function (c) is again a positive continuous function on the interval [0, 1], with the condition as in Equation (31).

$$\Sigma_{j=1}^{m} C_j(1) = 1, \tag{31}$$

with $j = 1, 2, ..., m$ and $0 \leq s \leq 1$. To guarantee that the scedasis function is uniquely defined, we use its integral. The integral, so the integrated scedasis function (C_j) is therefore as in Equation (32).

$$C_j(s) := \frac{1}{m} \int_0^s c(u, j)\, du \tag{32}$$

The additional component in the scedasis function $c\left(\frac{i}{n}, j\right)$, causes the function to be able to be interpreted as the relative frequency of extremes at time point i for an stock j.

Observations within the cross-section can be dependent, which can also happen for tail observations. The dependence in the tail is known as tail dependence. In order to estimate the tail dependence, we need to let F_i be the joint distribution function of the vectors $(X_{i,1}, X_{i,2}, ..., X_{i,m},)$ and F_i is continuous. Therefore, we can state $U_{i,j} := F_{i,j}^{\leftarrow}\left(1 - \frac{1}{s}\right)$, with the arrow pointing to the left representing the generalized inverse function. This leads us to Equation (33). This equation does not depend on the time point i and it is in the DA of a multivariate extreme value distribution.

$$\tilde{F}(x_1, x_2, ..., x_m) := F_i(U_{i,1}(x_1), U_{i,2}(x_2), ..., U_{i,m}(x_m)) \tag{33}$$

The multivariate tail dependence structure is also not dependent on time point i. In order to estimate the tail dependence, we can define the tail copula as in Equation (34).

$$R_{j_1,j_2}(x, y) = \lim_{s \downarrow 0} \frac{1}{s} P\left(1 - F_{i,j_1}(x_i, j_1) \leq sx, 1 - F_{i,j_2}(x_i, j_2) \leq sy)\right), \tag{34}$$

here $R_{j_1,j_2}(x, y)$ is the tail copula for the stocks j_1 and j_2 in the cross-section, with $j_1 \neq j_2$. We must also note that $(x, y) \in [0, \infty]^2 \setminus \{(\infty, \infty)\}$.

Similar to Equation (21) and Einmahl et al. (2016), $C_j(s)$ can be seen as the number of exceedances over a high quantile at stock j. However, all stocks should have the same threshold if we want to compare the different $C_j(s)$'s. As in Einmahl et al. (2022) when the observations are put together, the common threshold of the high empirical quantile will be $N = n \times m$. For the observations $\{X_{i,j}\}_{i=1,j=1}^{nm}$, the $(N-k)$-th order statistic is represented by $X_{N-k,N}$. This allows us

to define the estimator as in Equation (35).

$$\hat{C}_j(s) := \frac{1}{k}\Sigma_{i=1}^{[ns]}\mathbf{1}_{\{X_{i,j}>X_{N-k,N}\}}, \tag{35}$$

where we have an intermediate sequence k. Note the conditions as, $k = k(n) \to \infty$, $\frac{k(n)}{n} \to 0$ when $n \to \infty$. When taking the sum over all the integers, we have $1 \leq i \leq ns$.

4.4.7 Space-time trends testing

There are some assumptions and requirements to which the data needs to hold in order for the space-time trend model to give reliable outcomes for its tests. These assumptions are the same as in Einmahl et al. (2022) and in Einmahl et al. (2016). We refer to their paper for the assumptions.

When taking these assumptions into consideration we can start to develop our tests, which test for homoscedasticity in the tail, so a constant scedasis. The null hypothesis in Equation (36) states that the frequency of extremes is constant over the cross-section of the stocks return. Specifically this means, that we test that for all $j = 1, ..., m$ stocks whether $C_j(1) = \frac{1}{m}$. The alternative hypothesis states that the frequency of returns is different per stock.

$$\begin{aligned} H_0 &: C_j(1) = \frac{1}{m}, \qquad \text{for all} \quad j = 1, ..., m \\ H_1 &: C_j(1) \neq \frac{1}{m}, \qquad \text{for some} \quad j = 1, ..., m \end{aligned} \tag{36}$$

Let $M = I_m - \frac{1}{m}\mathbb{1}_m\mathbb{1}_m^{'}$, with I_m the identity matrix and $\mathbb{1}_m$ is a unit vector where both have a dimension of m. Under the null hypothesis of Equation (36), Equation (37) will be m-variate normally distributed with a mean equal to 0 and a covariance matrix of $(M\Sigma_1 M')$.

$$D = \left(\left(\sqrt{k}\hat{C}_1(1) - \frac{1}{m}\right), \left(\sqrt{k}\hat{C}_2(1) - \frac{1}{m}\right), ..., \left(\sqrt{k}\hat{C}_m(1) - \frac{1}{m}\right)\right) \tag{37}$$

$$\sigma_{j_1,j_2} := \frac{1}{m}\int_0^{s_1 \cap s_2} R_{j_1,j_2}(s_1 c(u, j_1), s_2 c(u, j_2))du, \qquad \text{for} \qquad 1 \leq j_1, j_2 \leq m \tag{38}$$

Here Σ_1 is an invertible matrix with elements as in Equation (38), so the rank of $(M\Sigma_1 M') = m - 1$. Hence, similar to Einmahl et al. (2022), we can define the test statistic to be Equation (39) and under the assumption for large n, the corresponding distribution is Equation (40).

$$T_5 := D'_{m-1}((M\hat{\Sigma}_1 M')_{m-1})^{-1} D_{m-1} \tag{39}$$

$$T_5 \xrightarrow{d} \chi^2_{m-1}, \tag{40}$$

where the subscript $m-1$ means the first $m-1$ components.

Lastly, we test for all stock whether the scedasis function is constant. This is presented more formally in Equation (41).

$$\begin{aligned} H_0 &: C_j(s) = sC_j(1) \\ H_1 &: C_j(s) \neq sC_j(1) \end{aligned} \tag{41}$$

This null hypothesis states that the scedasis is constant and thus the existence of homoscedastic extremes. This means that the frequency of extremes is constant over the sample period.

Einmahl et al. (2022) propose a Kolmogorov-Smirnov or a Cramér-von-Mises test statistic based on the process in Equation (42). The Kolmogorov-Smirnov test statistic, where Einmahl et al. (2022) focus on, is similar to the one in Einmahl et al. (2016).

$$\left(\sqrt{k} \frac{(\hat{C}_j(s) - s\hat{C}_j(1))}{\sqrt{\hat{C}_j(1)}} \right), \qquad s \in [0,1] \tag{42}$$

Under this process, the null hypothesis is as in Equation (43).

$$\left\{ \sqrt{k\hat{C}_j(1)} \frac{\hat{C}_j(s)}{\hat{C}_j(1)} - s \right\}_{s\in[0,1]} \xrightarrow{d} \{B(s)\}_{s\in[0,1]}, \tag{43}$$

with B being a Brownian bridge.

This gives a Kolmogorov-Smirnov test statistic as in Equation (44).

$$T_6 := sup_{0\leq s\leq 1} \frac{|\hat{C}_j(s) - s\hat{C}_j(1)|}{\sqrt{\hat{C}_j(1)}} \tag{44}$$

The distribution of the statistic is shown in Equation (45).

$$\sqrt{k}T_6 \xrightarrow{d} sup_{0\leq s\leq 1}|B(s)| \tag{45}$$

5 Results

5.1 American subsamples

5.1.1 Extreme value index

Before we can test whether the scedasis is constant, we first need to validate whether our data adheres to the assumptions. These assumptions require the extreme value index in the cross-section and the extreme value index over time to be constant.

We first test whether the extreme value index is constant in the cross-section. If that is the case for the given subsample, then we can check for that subsample if the extreme value index is constant over time. If a subsample has an extreme value index that is not constant in the cross-section, then that subsample is dropped since the remaining part of the analysis will give unreliable outcomes.

For every stock in the subsample, we estimate their extreme value index. The range of k is 130 up and until 250. For each subsample, the confidence intervals that we obtain are different, because the number of stocks is different in each subsample. Similar to Einmahl et al. (2022), we keep in mind the problem of multiple testing. Hence, we apply a Bonferroni-corrected critical value. We use a significance level of 2.5% (5% double sided).

Examples for two different subsamples are shown in Figure 3 and Figure 4. Figure 3 shows the upper and lower bounds of the extreme value index in the cross-section for the health research sector in the time period 2000/2019. In this subsample we have a total of 50 stocks. The probability level for the 2.5% significance level is $1 - \frac{0.025}{50} = 0.9995$. In this case the distribution for the Bonferroni-corrected critical value is a χ^2_6 with a critical value of 24.103. In Figure 3, we observe a clear gap between the upper bounds in blue and the lower bounds in red for multiple values of k. This means, there is a value that exists in all confidence intervals. In this case for k equal to 200, the value is equal to 0.180. Therefore, this subsample adheres to the assumption of a constant extreme value index in the cross-section.

Figure 4 shows the upper and lower bounds of the extreme value index in the cross-section for the financial sector in the time period NT1. In this subsample we have a total of 593 stocks. The probability level for the 2.5% significance level is $1 - \frac{0.025}{593} = 0.99996$. In this case the distribution for the Bonferroni-corrected critical value is a χ^2_{24} with a critical value of 53.479. For all the k plotted, we do not find a gap between the upper and lower bounds of the 593 stocks. There are many upper bounds of stocks below the lower bounds of other stocks. This means, there does not exist a value that is in all confidence intervals. Therefore, the assumption of a constant extreme

value index in the cross-section is violated for this subsample.

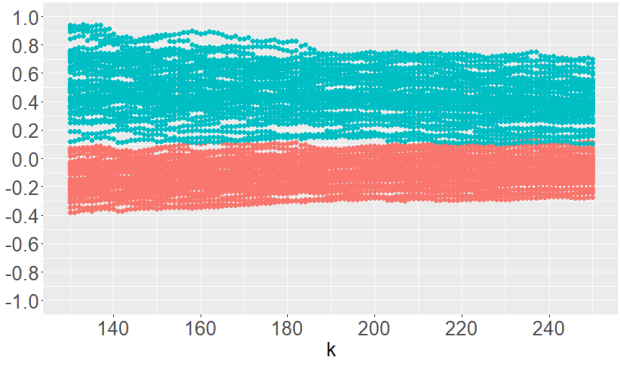


Figure 3: Extreme value index upper and lower bounds for all the stocks in the health research sector in time period 2000/2019

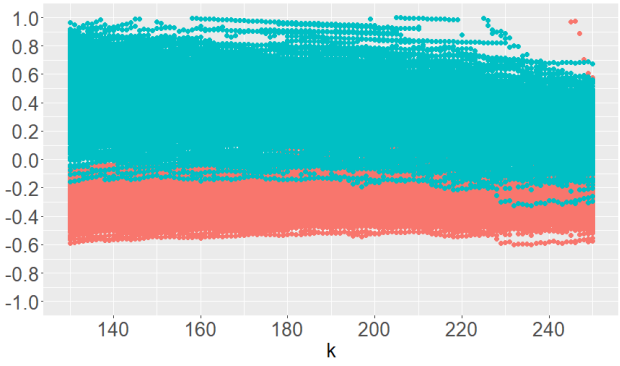


Figure 4: Extreme value index upper en lower bound for all the stocks in the financial sector in time period NT1

Table 3 shows for each subsample whether their extreme value index is constant in the cross-section. If the extreme value index is constant in the cross-section, then the symbol *CS* together with the point estimate in parentheses are put in the cell. Cells with *CS* have a plot for the upper and lower bounds similar to Figure 3. For all of these subsamples, there is a gap between upper and lower bounds at $k = 200$. The symbols *NO* and *NA* mean no constant extreme value index in the cross-section and not applicable, respectively. Cells with *NO* have a plot for the upper and lower bounds similar to Figure 4.

Table 3: Testing constant extreme value index in the cross-section and over time

Sector	Full sample	NT1	GR	NT2	C19	2000/2009	2000/2019	2007/2019	2007/2021	2009/2021
Consumer discretionary (CD)	CS	CS	NO	CS, T	NO	NO	CS	CS, T	CS, T	CS, T
	(0.171)	(0.032)	NO	(0.019)	NO	NO	(0.152)	(0.097)	(0.135)	(0.108)
Technology (Tech.)	CS	CS	NO	CS	NO	CS	CS	CS, T	CS, T	CS, T
	(0.187)	(0.046)	NO	(0.019)	NO	(0.080)	(0.182)	(0.094)	(0.124)	(0.095)
Health research (HR)	CS	CS, T	NO	CS, T	NO	CS, T	CS, T	CS, T	CS, T	CS, T
	(0.201)	(0.043)	NO	(0.045)	NO	(0.087)	(0.180)	(0.095)	(0.133)	(0.102)
Financials (Fin.)	CS	NO	NO	NO	NO	NO	CS	CS	CS	CS
	(0.204)	NO	NO	NO	NO	NO	(0.188)	(0.118)	(0.149)	(0.098)
Consumer staples (Con. S)	NA	NA	NO	CS	NO	NA	NA	CS	CS	CS
	NA	NA	NO	(-0.026)	NO	NA	NA	(0.131)	(0.189)	(0.039)
Health consumer materials (HCM)	CS, T	NO	NO	CS, T	NO	NO	CS, T	CS, T	CS, T	CS, T
	(0.168)	NO	NO	(0.033)	NO	NO	(0.157)	(0.106)	(0.127)	(0.087)
Industrials (Indus.)	CS	CS	NO	CS, T	NO	CS	CS	CS, T	CS, T	CS, T
	(0.158)	(0.029)	NO	(0.001)	NO	(0.065)	(0.144)	(0.075)	(0.095)	(0.062)
Utilities (Util.)	CS	CS	NO	NO	NO	CS	CS	CS	NO	NO
	(0.228)	(0.027)	NO	NO	NO	(0.065)	(0.201)	(0.094)	NO	NO
Energy (Ene.)	CS	NO	NO	CS, T	NO	CS	CS	CS, T	CS, T	CS, T
	(0.168)	NO	NO	(0.040)	NO	(0.068)	(0.131)	(0.051)	(0.104)	(0.060)
Agriculture (Agri.)	CS	CS	NO	CS	NO	CS	CS, T	CS, T	CS	CS, T
	(0.130)	(0.034)	NO	(0.020)	NO	(0.036)	(0.045)	(0.038)	(0.088)	(0.036)

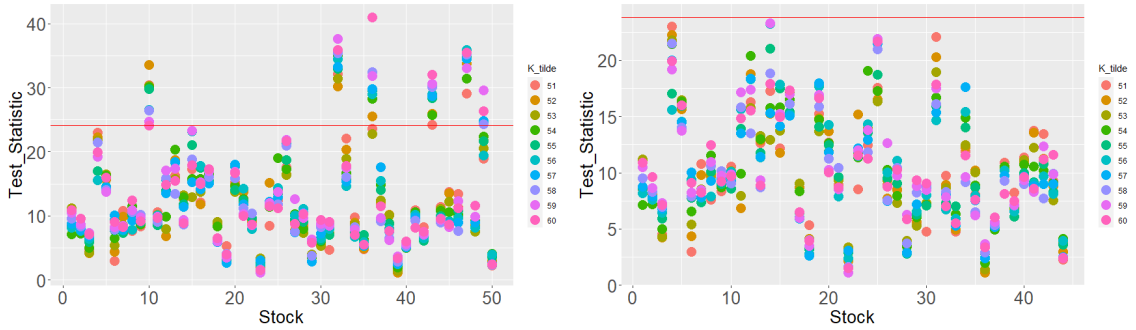
Note: Cells with a "CS" are the ones that have a constant extreme value index in the cross-section for at least one value of k, their point estimates are in parentheses. Their plot of upper and lower bounds looks similar to Figure 3. Whereas for the others ("NO") the plot looks similar to Figure 4. For these subsamples we check whether the extreme value index is constant over time. The cells that have a constant extreme value index over time "T" are next to "CS". Their plots for a constant extreme value index over time are similar to Figure 5. So, the cells with a "CS, T" are the subsamples with a constant extreme value index in the cross-section and over time. These subsamples are used for the rest of the analysis. "NA" means Not Applicable, that there are no stocks in that subsample, so that subsample is dropped.

Second, we test for a constant extreme value index over time and then we keep only the subsamples which have a constant extreme value index. So, we end up with subsamples which have both a constant extreme value index in the cross-section and over time. These subsamples are shown in Table 3 with the symbol *CS, T*. We only test for a constant extreme value index over time for the subsamples which have a constant extreme value index in the cross-section. The reason for this is that both of these assumptions need to hold in order for the rest of the methodology to give reliable results.

Figure 5 shows an example of the test for the extreme value index over time. This example considers the health research sector over the time period 2000/2019. As stated earlier, for this subsample we have 50 stocks. We use the test statistic T_2 for every stock. For all stocks in the subsample, we split the observations into $\sqrt{n}$ blocks. We choose this block size since that gives the lowest serial dependence as described in Section 4.3. Since the number of blocks needs to be an integer, we take the rounded value of $\sqrt{n}$. In the case of the health research sector for 2000/2019 we have $\sqrt{50} \approx 7$ blocks. In all blocks we estimate the next ten test statistics ($\tilde{k}$) for each stock and this is shown in Figure 5a. With the same Bonferroni-corrected critical value as before.

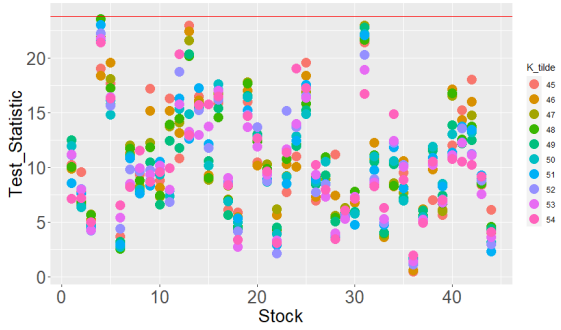
We observe in Figure 5a that there are six stocks which have test statistics that are above the Bonferroni-corrected critical value. Since six out of 50 stocks is considered a small amount we remove them from the subsample. Figure 5b shows the results from the exact same test when we use the same subsample without the six deleted stocks. Since, we have removed six out of 50 stocks we look at new $\tilde{k}$ as in Figure 5c. For both Figure 5b and Figure 5c we observe that now all $\tilde{k}$'s, for all stocks are below the Bonferroni-corrected critical value. Hence, this subsample adheres to the assumption of a constant extreme value index over time.

Figure 5: Extreme value index over time for the health research sector over the time period 2000/2019. The horizontal red line is the Bonferroni-corrected critical value



(a) Before deleting stocks

(b) After deleting six stocks



(c) After deleting six stocks, with new $\tilde{k}$ test statistics

This analysis is done for all subsamples which have a constant extreme value index in the cross-section. If we do not need to remove a lot of stocks, we apply this removal to a subsample. If we have to remove a lot of stocks for the extreme value index to be constant over time then this is not representative for the sector in that time period. In the latter case, we conclude that the extreme value index is not constant over time for that subsample.

We have 31 subsamples which have a constant extreme value index in the cross-section and over time, as shown in Table 3 with a CS, T. Table 4 shows the 28 subsamples that are used for the rest of the analysis. We drop the three subsamples (HCM Full, HR NT1 and HR 2000/2009), since there is no other subsample which is in the same time period to compare it to. From the 28 subsamples in Table 4, we have seven sectors in five time periods.

Even though we do not have the separate periods NT1, Great Recession and Covid-19, we can still detect what their influence is on the scedasis. When we compare the periods 2000/2019

and 2007/2019 with each other we can presume that large parts of the difference in results will be there because of time period NT1, since that period is only present in the former and not the latter. Nevertheless, because of the manner the subsample is structured, there are also some stocks not in the periods 2000/2019 and 2007/2019, which are in NT1. This same idea applies for the Great Recession, comparing NT2 with 2007/2019. As stated, comparing 2000/2019 with 2007/2019 gives NT1's addition. Another comparison, for Covid-19 will be time periods 2007/2019 with 2007/2021. The addition of the Great Recession and Covid-19 to NT2 constructs 2007/2021. Comparing 2007/2021 with NT2, shows the impact of the two crisis. Comparing 2007/2019 with 2009/2021 allows us to investigate what the impact of the two crisis are when we add their respective time period to NT2. All these comparisons are only done for the subsamples which have a constant extreme value index as shown in Table 4 with a CS, T.

Table 4: Subsamples used for further analysis

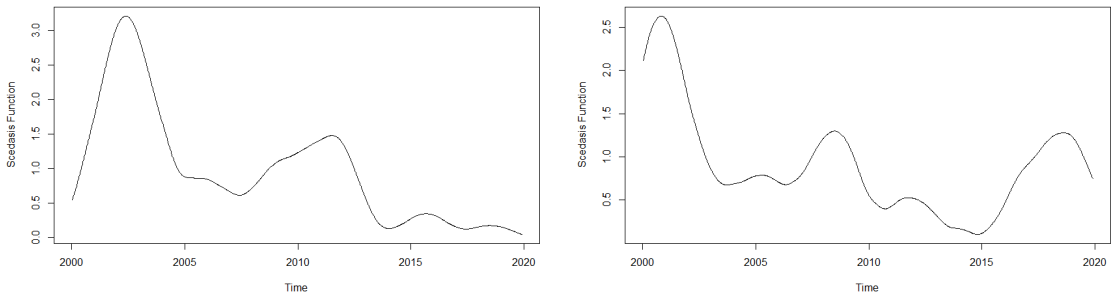
Sector	NT2	2000/2019	2007/2019	2007/2021	2009/2021
CD	CS, T	NA	CS, T	CS, T	CS, T
Tech.	NA	NA	CS, T	CS, T	CS, T
HR	CS, T	CS, T	CS, T	CS, T	CS, T
HCM	CS, T	CS, T	CS, T	CS, T	CS, T
Indus.	CS, T	NA	CS, T	CS, T	CS, T
Ene.	CS, T	NA	CS, T	CS, T	CS, T
Agri.	NA	CS, T	CS, T	NA	CS, T

Note: The subsamples with "CS, T" are used for further analysis. The subsamples with "NA" are not analysed further.

5.1.2 Scedasis function and trend over time per stock

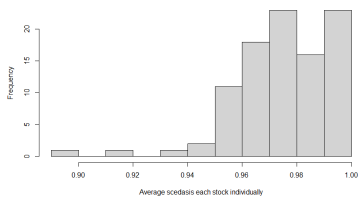
We intend to compare the subsamples with each other and find where the more riskier and safer stocks are. This means we compare their average scedasis. Figure 6 shows the plot of two stocks in the health research 2000/2019 subsample. In Figure 6a we observe the scedasis function for *Franklin Covey*, which is the stock with the highest average scedasis for this subsample. Figure 6b shows the plot for the scedasis function for *Total Renal Care*, which has the lowest average scedasis. We observe in both plots a high spike around 2001–2002, where both stocks have their highest scedasis. There is a gradual decline in scedasis for both stocks after the spike till the middle of 2007, where the Great Recession begins. For *Total Renal Care*, the concave shape occurs around the Great Recession and stops after the end of 2009, where the Great Recession ends. *Total Renal Care's* scedasis then gradually decreases again till 2015 where it hits it lowest point of around 0.2. For the risky stock *Franklin Covey*, we observe that the magnitude of the scedasis in the Great Recession is much higher than that for *Total Renal Care*. The length of the concave period for the riskier stock *Franklin Covey* lasts longer than for the safer *Total Renal Care*. Even though in both concave shapes, their respective maximum is both at around 1.5, *Franklin Covey* has reached that local maximum around the beginning of 2012. We interpret this as that the impact of the Great Recession has an extended effect on the scedasis of a riskier stock compared to a safer stock. From 2014 onwards, *Franklin Covey* has a low and stable scedasis between 0.1 and 0.4. Between 2015 and 2019 the scedasis has again a concave shape for *Total Renal Care*, with a peak around the same level as in the Great Recession.

Figure 6: Scedasis function for the two stocks with the highest and lowest average scedasis in the subsample health research 2000/2019

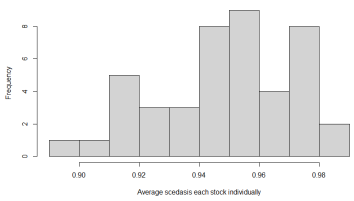


(a) Highest average scedasis (Franklin Covey) (b) Lowest average scedasis (Total Renal Care)

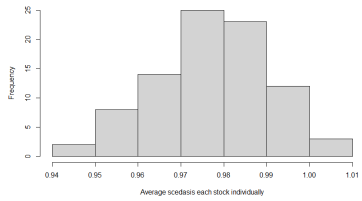
Figure 7: Average scedasis health research



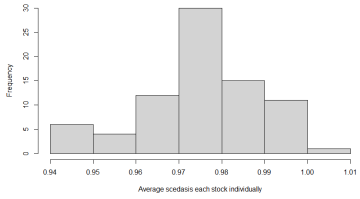
(a) Health research NT2



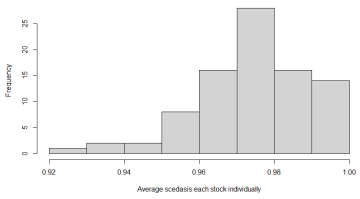
(b) Health research 2000/2019



(c) Health research 2007/2019



(d) Health research 2007/2021



(e) Health research 2009/2021

Figure 7b shows for the health research 2000/2019 subsample, how the average scedasis is distributed within the subsample. The histograms of all 28 subsamples are shown in Table A1 in Section A.1. We choose the average scedasis instead of the total scedasis, therefore the length of the time period does not influence the comparison. This allows for a fair comparison. Therefore, we can investigate two aspects. First of all, whether within each subsample there is or are a few outliers which have a relatively high or low scedasis. Second, comparing the subsamples with each other allows us to understand if a particular sector is riskier than another sector. We can then understand if this difference in riskiness is caused by a different time period or that the sector in general is riskier, despite the time period we are in.

For most of the sectors the figures look similar to Figure 7. There are a few stocks on the left, so a few stocks are relatively safer than most of the other stocks as in Figure 7a and Figure 7b. In case of Figure 7c, we observe that the average scedasis is more evenly distributed. So, the tail risk of the safer stocks increases more than the tail risk of the riskier stocks increases. This means that in the expansion period NT1, there are relatively less riskier stocks, which is in line with what we expect. The riskier stocks in an expansion period are less impacted by a crisis than other stocks. The average scedasis of the riskier stocks increases less than the average scedasis of the safer stocks. Figure 7c and Figure 7e shows the impact of the Great Recession and Covid-19, respectively. We conclude that the impact of the Great Recession is bigger than Covid-19. The reason for this is that

the Great Recession has an immediate impact, since market participants can speculate better on what the impact of a financial crisis is, on a stock. Whereas for a pandemic, market participants at first think that the virus might not be as important until it is classified as a pandemic, so the shock comes later in the time period. This causes all stocks to be impacted in a slower manner than what is the case for the Great Recession, so the impact of Covid-19 is more gradual.

Table 5 shows the stocks with the highest (high) and lowest (low) average scedasis for all 28 subsamples. We report the name, the average scedasis in round parentheses and the Kolmogorov-Smirnov p-value in squared parentheses. The Kolmogorov-Smirnov p-values are from the test of a constant scedasis for that particular stock. The final statistic that is reported is the average scedasis of the entire subsample in the row "ave".

We start with our main example, health research 2000/2019. We find in Table 5 that the average scedasis of the riskiest stock of the subsample *Franklin Covey* is almost 0.1 higher than that of the safest stock in the subsample *Total Renal Care*. The p-value of the Kolmogorov-Smirnov test for *Franklin Covey* is 0.000, which causes us to reject the null hypothesis of a constant scedasis for this stock in the period 2000/2019. This coincides with the plot given in Figure 6a. The exact opposite holds for *Total Renal Care* with a p-value of 0.984, also see Figure 6b. A pattern of a non-constant scedasis for the riskier stocks and a constant scedasis for the safer stocks is not present in the subsample. We conclude that the homoscedasticity or heteroscedasticity in the tail does not make a distinction between the average scedasis of a risky or safe stock. A crisis does not have an impact on the conclusion stated.

Table 6 shows the fraction of stocks out of the total number of stocks for that subsample, who have a p-value below 0.05 for the Kolmogorov-Smirnov test, T_3. The null hypothesis states that there does not exists a trend in the scedasis over time for each stock, so homoscedasticity in the tail. For the health research sector in 2000/2019 we observe that 14 out of 44 (31.8%) stocks have a p-value below 0.05. So often, we do not reject the null hypothesis of homoscedasticity in the tail for a stock in this subsample.

The health research 2000/2019 subsample, is the subsample with the third highest fraction of stocks with heteroscedasticity in the tail. The two highest are the agricultural sector in 2007/2019 and 2007/2021. The number of stocks that exhibit heteroscedasticity in the tail for all subsamples is relatively small. For the stocks in the health consumer materials subsamples, there is almost no heteroscedasticity in the tail present. For 23 out of 28 subsamples, the number of stocks with heteroscedasticity in the tail is less than 10% of the stocks in the subsample. Since the scedasis is a

Table 5: Stocks with highest and lowest scedasis and the average of the entire subsample

Sector	Value	NT2	2000/2019	2007/2019	2007/2021	2009/2021
	high	Dryclean USA	NA	Adams (R & E)	Adams (R & E)	MCKesson HBOC
		(1.000) [0.457]	NA	(1.000) [0.375]	(0.999) [0.321]	(1.000) [0.631]
CD	low	Bluelinx Hold.	NA	Kirklands	Haverty Furniture	1 800 Flowers
		(0.918) [0.431]	NA	(0.910) [0.314]	(0.926) [0.416]	(0.907) [0.447]
	ave.	(0.978)	NA	(0.973)	(0.974)	(0.980)
	high	NA	NA	Gannet	Hertz Global Hold.	Barrett Bus. Ser.
		NA	NA	(1.000) [0.939]	(1.000) [0.889]	(1.000) [0.894]
Tech.	low	NA	NA	F 5 Networks	Healthcare Ser. Gr.	Mercadolibre
		NA	NA	(0.916) [0.266]	(0.920) [0.596]	(0.921) [0.380]
	ave.	NA	NA	(0.976)	(0.974)	(0.977)
	high	Radnet	Franklin Covey	Krug Int. Corp.	F.T.I. Consulting	Ensign Group
		(0.999) [0.173]	(0.986) [0.000]	(1.000) [0.904]	(1.000) [0.323]	(0.998) [0.795]
HR	low	Bridgepoint Ed.	Total Renal Care	Profit Recovery Gr.	Adcare Health Sys.	Regeneration Tech.
		(0.898) [0.626]	(0.898) [0.984]	(0.943) [0.071]	(0.941) [0.025]	(0.927) [0.446]
	ave.	(0.975)	(0.949)	(0.978)	(0.975)	(0.974)
	high	Skyline Corp.	Carlisle Comp.	Hunstman Corp.	Durect Corp.	Cytotherapeutics
		(1.000) [0.717]	(0.997) [0.808]	(1.000) [0.617]	(1.000) [0.692]	(1.000) [0.777]
HCM	low	Reeds	Church & Dwight	Protalix Biotherapeutics	Willamette Valley Vineyards	Willamette Valley Vineyards
		(0.918) [0.292]	(0.919) [0.516]	(0.922) [0.043]	(0.923) [0.738]	(0.909) [0.557]
	ave.	(0.981)	(0.965)	(0.982)	(0.982)	(0.982)
	high	Mine Safety App.	NA	Veri Telc Int. Corp.	Mohawk Ind.	China BAK Battery
		(1.000) [0.502]	NA	(1.000) [0.423]	(1.000) [0.699]	(1.000) [0.606]
Indus.	low	APA Optics	NA	Superconductor Tech.	Emerson Radio Corp.	Friedman Ind.
		(0.936) [0.396]	NA	(0.925) [0.284]	(0.940) [0.564]	(0.939) [0.217]
	ave.	(0.985)	NA	(0.984)	(0.984)	(0.984)
	high	Comstock Resou.	NA	Stone Energy Corp.	Comstock Resou.	Comstock Resou.
		(1.001) [0.095]	NA	(1.000) [0.283]	(1.000) [0.556]	(1.000) [0.156]
Ene.	low	Unit Corp	NA	Unit Corp	MI Schottenstein	Panhandle Royalty
		(0.879) [0.178]	NA	(0.903) [0.281]	(0.933) [0.331]	(0.933) [0.868]
	ave.	(0.973)	NA	(0.974)	(0.976)	(0.975)
	high	NA	Alico	Griffin (L & N)	NA	Cal Maine Foods
		NA	(0.993) [0.024]	(0.953) [0.002]	NA	(0.966) [0.608]
Agri.	low	NA	Scotts Comp.	Scotts Comp.	NA	Cadiz
		NA	(0.895) [0.010]	(0.892) [0.005]	NA	(0.895) [0.163]
	ave.	NA	(0.945)	(0.919)	NA	(0.928)

Note: Stocks that are listed have the highest and lowest scedasis for the given subsample, with their average scedasis given in round parentheses. The p-value of the Kolmogorov-Smirnov test is given in squared parentheses. The average of each subsample is stated in the rows "ave." between round parentheses. "NA" means Not Applicable.

relative measure, we believe that this low presence of heteroscedasticity in the tail is caused by the nature of the stocks. In general some stocks are more riskier than others but at the same time they

Table 6: Fraction of Kolmogorov-Smirnov p-values for a constant scedasis

Sector	NT2	2000/2019	2007/2019	2007/2021	2009/2021
CD	$\frac{3}{204}$ (0.015)	NA	$\frac{9}{197}$ (0.046)	$\frac{6}{179}$ (0.034)	$\frac{7}{184}$ (0.038)
Tech.	NA	NA	$\frac{6}{199}$ (0.030)	$\frac{6}{180}$ (0.033)	$\frac{4}{193}$ (0.021)
HR	$\frac{2}{96}$ (0.021)	$\frac{14}{44}$ (0.318)	$\frac{6}{87}$ (0.069)	$\frac{5}{79}$ (0.063)	$\frac{4}{87}$ (0.046)
HCM	$\frac{1}{309}$ (0.003)	$\frac{1}{200}$ (0.005)	$\frac{1}{288}$ (0.003)	$\frac{0}{263}$ (0)	$\frac{0}{282}$ (0)
Indus.	$\frac{3}{506}$ (0.006)	NA	$\frac{6}{485}$ (0.012)	$\frac{4}{453}$ (0.009)	$\frac{1}{471}$ (0.002)
Ene.	$\frac{8}{109}$ (0.073)	NA	$\frac{10}{100}$ (0.100)	$\frac{2}{81}$ (0.025)	$\frac{6}{89}$ (0.067)
Agri.	NA	$\frac{4}{4}$ (1)	$\frac{6}{6}$ (1)	NA	$\frac{1}{5}$ (0.200)

Note: Shows the fraction of how many p-values are below 0.05. P-values below 0.05, mean heteroscedasticity in the tail. So, $\frac{3}{204}$ means 3 out of 204. In parentheses are the values in decimal format.

have a higher return. Therefore, for some stocks their scedasis is always higher but the scedasis does not change much.

When we again compare the two crises with NT2, we find that the number of stocks with heteroscedasticity in the tail increase in the crises periods. This increase is relatively small, but still visible. The Great Recession causes a higher increase in the number of stocks with heteroscedasticity in the tail than Covid-19. During the Great Recession, the number of stocks with heteroscedasticity in the tail has increased for several reasons. First, the recession has affected different sectors of the economy in different ways, leading to increased variability in the performance of stocks in those sectors. Second, the recession has affected investors sentiment, leading to greater variability in the prices of individual stocks. Covid-19 on the other hand is a pandemic, which at first notice is thought to not impact the financial market in a substantial manner. The effects are first visible in the health research sector and then gradually spills over to other sectors.

Whether the stocks that have homoscedasticity in the tail, are the risky stocks can not be stated. The reason for this is that we have homoscedasticity in the tail for the stocks with the highest average scedasis in some subsamples. Whereas for other subsamples this is not the case. This same pattern exists for the lowest average scedasis. So, having a homoscedastic or heteroscedastic trend in the tail does not make a distinction between the high and low tail risk stocks.

5.1.3 Frequency of extremes in the cross-section

We test for each subsample whether the frequency of extremes is constant in the cross-section. Table 7 shows the p-values of the Kolmogorov-Smirnov test, for the test whether the total integrated scedasis of extreme returns is constant in the cross-section, T_5. We report the p-values that are obtained for the value 1000 for k, which is similar to Einmahl et al. (2022). The p-values for other values of k are rather close to the reported p-values.

For the health research 2000/2019 subsample, we find a p-value of 0.000. Therefore we reject the null hypothesis. This means that the total integrated scedasis of the extreme stock returns is not constant across all stocks in this subsample. Thus, the frequency of extremes is different for each stock within the health research 2000/2019 subsample. The stocks in the health research sector, mainly focus on the research and development of cures and treatments for deceases as cancer or diabetes. Starting in the 1990's and continuing in early 2000's the health research companies are discovering the Genome, formally classified as the Human Genome Project. Since this project is a relatively big foundation for the cures and treatments of illnesses, the entire health research sector is influenced by each announcement consisting of developments of the project. In 2003/2004 92% of the Genome is discovered hence, after 2004 the impact of this project is less present. Hence, the period before the Great Recession has a substantial impact on the health research sector, because of the Human Genome Project.

Besides the agricultural subsamples, the remaining subsamples all have a p-value close to one. We conclude for the consumer discretionary, technology, health consumer materials, industrial and energy sector that the frequency of extremes is similar for each stock, irrespective of the time period. We believe that the frequency of extremes is the same for all stocks in their subsample, as the companies in the related sector are perceived to be competitors. So, if there is a shock in returns for the stock of one company, then this has an impact on the stocks of the other companies as well. This is confirmed by the values of tail dependence which are close to zero, indicating high tail dependency. Hence, a crisis does not have an impact on the frequency of extremes for these five sectors.

The last conclusion is that for the time periods NT2, 2007/2019, 2007/2021 and 2009/2021 for all but the agricultural sector, the frequency of extremes is similar for each stock in its respective sector. This means that no matter the time period, the agricultural sector has different frequency of extremes for all its stocks. This could be due to the fact that there are only a handful of stocks

in those subsamples or with the nature of the agricultural stocks. For the agricultural sector we can state that this probably has to do with the fact that different agricultural companies, grow different types of crops. For instance, if there is a bad harvest for corn, this does not impact the harvest for apples and production of apple juice. Since, the aim of the agricultural sector is to provide food and drinks and these are necessary goods, the time period does not have an impact on the test results. Hence, the frequency of extremes in the agricultural sector is not impacted by a financial crisis or pandemic.

Table 7: Kolmogorov-Smirnov p-values, for the test whether the total integrated scedasis of extreme returns is constant in the cross-section, for k equal to 1000

Sector	NT2	2000/2019	2007/2019	2007/2021	2009/2021
CD	0.999	NA	0.999	0.999	0.999
Tech.	NA	NA	0.999	0.999	0.999
HR	0.999	0.000	0.999	0.999	0.999
HCM	0.999	0.999	0.999	0.999	0.999
Indus.	0.999	NA	0.999	0.999	0.999
Ene.	0.999	NA	0.999	0.999	0.999
Agri.	NA	0.004	0.000	NA	0.001

Note: The p-values that are given are from the Kolmogorov-Smirnov test with $k = 1000$.

5.1.4 Frequency of extremes over time

We test for each subsample whether there exists a trend over time. Table 8 shows the p-values of the Kolmogorov-Smirnov test for a trend over time, T_6. We again report the p-values that are obtained for the value 1000 for k. Again the p-values for other values of k are rather close to the reported p-values. For this test we need to keep in mind the problem of multiple testing, which also occurs in Einmahl et al. (2022). We apply the same Bonferroni correction, so in our case this means $\frac{5}{\text{number of stocks}}\%$, where the number of stocks is different for each subsample. The number of stocks are shown in round parentheses in Table 8. In squared parentheses are the Bonferroni-corrected critical values, where 0.000245 means 0.024510%.

For the health research sector in 2000/2019, we find a p-value of 0.115399 with a Bonferroni-corrected critical value of 0.001136. Therefore, we do not reject the null hypothesis of a trend over time in this subsample. Thus, the frequency of extremes is constant over the sample period. This

conclusion holds for all health research subsamples.

We reject the null hypothesis for all health consumer materials and industrial subsamples. In these two sectors the frequency of extremes is not constant over the sample period. The opposite holds for the agricultural sector, where in none of the three cases we reject the null hypothesis. For the agricultural sector the frequency of extremes is constant over the sample period. For the consumer discretionary (3 out of 4), technology (2 out of 3) and energy (1 out of 4) sector we reject the null hypothesis.

For the time period between the two crisis, NT2, we find that for three out of the five subsample we reject the null hypothesis. When we add the Great Recession then this becomes tree out of seven and when we add Covid-19 to NT2 this is five out of seven. When we add both crises, so 2007/2021, we have that we reject the null hypothesis in three out of the six cases. We conclude that there is no clear impact of a crisis on this test. The impact of both of the crises is different for each of the sectors and thus also for the frequency of extremes.

Table 8: Kolmogorov-Smirnov p-values for a trend over time for k equal to 1000

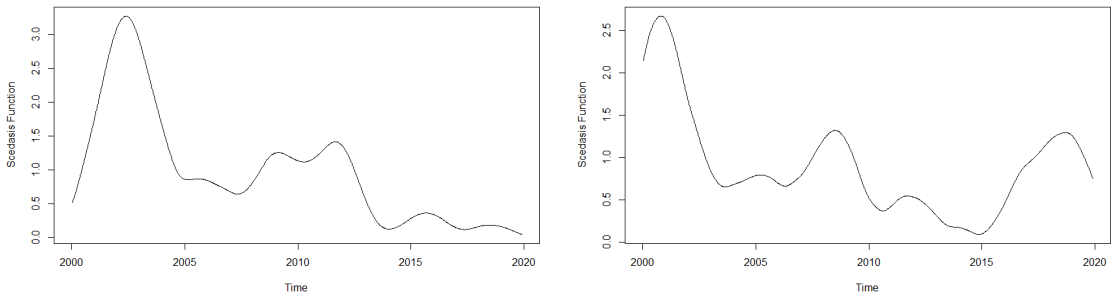
Sector	NT2	2000/2019	2007/2019	2007/2021	2009/2021
CD	0.000000 (204)	NA	0.000281 (197)	0.000224 (179)	0.000021 (184)
	[0.000245]	NA	[0.000254]	[0.000279]	[0.000272]
Tech.	NA	NA	0.000037 (199)	0.001210 (180)	0.000001 (193)
	NA	NA	[0.000251]	[0.000278]	[0.000259]
HR	0.207286 (96)	0.115399 (44)	0.323854 (87)	0.154433 (79)	0.007225 (87)
	[0.000521]	[0.001136]	[0.000575]	[0.000633]	[0.000575]
HCM	0.000000 (309)	0.000016 (200)	0.000000 (288)	0.000000 (263)	0.000000 (282)
	[0.000162]	[0.00025]	[0.000174]	[0.000190]	[0.000177]
Indus.	0.000000 (506)	NA	0.000000 (485)	0.000000 (453)	0.000000 (471)
	[0.000099]	NA	[0.000103]	[0.000110]	[0.000106]
Ene.	0.011109 (109)	NA	0.043746 (100)	0.015147 (81)	0.000034 (89)
	[0.000459]	NA	[0.0005]	[0.000617]	[0.000562]
Agri.	NA	0.085323 (4)	0.237790 (6)	NA	0.392243 (5)
	NA	[0.0125]	[0.008333]	NA	[0.01]

Note: The p-values that are given are from the Kolmogorov-Smirnov test with k = 1000. In round parentheses are the number of stocks in that subsample. In squared parentheses are the Bonferroni-corrected critical values, where we divide the 5% significance level by the number of stocks.

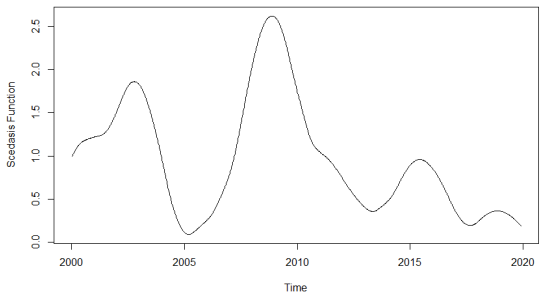
5.2 American versus Dutch counterpart

We want to compare the Dutch market with the American market. Thus, we add the Dutch stocks to the American subsamples and name them "Dutch subsamples". We look at the 28 subsamples that have a constant extreme value index for the American data. We assume that the extreme value index is constant over time and in the cross-section for the 28 Dutch subsamples. We skip the tests for a constant extreme value index in the cross-section and over time, since the AEX and AMX have so few stocks that we only have one stock for each of the seven sectors. We estimate the scedasis function and compare the results of the Dutch stocks with the American stocks and try to find similarities.

Figure 8: Scedasis function for the stock with the highest, lowest average scedasis and Dutch stock in the subsample health research 2000/2019



(a) Highest average scedasis (Franklin Covey) (b) Lowest average scedasis (Total Renal Care)



(c) DSM

Figure 8c shows the plot of the scedasis function for *DSM* (health research) in 2000/2019. We observe that the peak of the scedasis function is at the end of the Great Recession and from there on onwards, it is decreasing. When we compare *DSM* with the highest and lowest American counterparts, see Figure 8a and Figure 8b, we observe that the shape of the scedasis function of

DSM is somewhat similar to the riskiest stock *Franklin Covey* besides the Great Recession. Note, the plots in Figure 6 and Figure 8 are not exactly the same because of the addition of the Dutch stock.

The p-value of the Kolmogorov-Smirnov test for all the Dutch stocks within their respective subsamples are around 0.000, so below 0.05. This means we reject the null hypothesis of a constant scedasis for all these stocks. Hence, the frequency of extremes is not constant over time for all these stocks. A possible reason for this could be that the Dutch stock market is relatively small and sensitive to extreme global events. These events have a fairly large impact.

ASML, *AkzoNobel* and *Ahold* for example are internationally operating companies and therefore their stocks are sensitive to events outside their country of listing. Contrary to the Netherlands, the United States of America is a larger economy and the listed stock in the United States of America are less sensitive to extreme events outside their own country, for instance the default of Greece.

Table 9 shows the comparison of the American and Dutch stocks in the Dutch subsamples. Here, the stocks are ranked according to their relative riskiness and we report the rank of the Dutch stock. So, a stock with rank $\frac{19}{45}$ is the nineteenth riskiest stock out of 45 stocks. From Table 9 we conclude that in case of non-crisis periods, Dutch stocks are riskier than American stocks. The Great Recession has a larger impact on the American economy than on the Dutch economy, consequently on the riskiness of the different stocks and sectors. For the Great Recession, this is reflected in a higher rank for the Dutch stocks. This same pattern also exists for Covid-19. Hence, we conclude that in times of a crisis the rank of the Dutch stocks decrease, so they become relatively safer.

Table 9: Average scedasis comparison American and Dutch stocks in Dutch subsample

Sector	Stock	NT2	2000/2019	2007/2019	2007/2021	2009/2021
CD	Ahold	$\frac{109}{205}$	NA	$\frac{77}{198}$	$\frac{69}{180}$	$\frac{98}{185}$
Tech.	ASML	NA	NA	$\frac{27}{200}$	$\frac{20}{181}$	$\frac{16}{194}$
HR	DSM	$\frac{29}{97}$	$\frac{19}{45}$	$\frac{44}{88}$	$\frac{36}{80}$	$\frac{45}{88}$
HCM	AkzoNobel	$\frac{135}{310}$	$\frac{54}{201}$	$\frac{207}{289}$	$\frac{235}{264}$	$\frac{182}{283}$
Indus.	Philips	$\frac{25}{507}$	NA	$\frac{149}{486}$	$\frac{418}{454}$	$\frac{436}{472}$
Ene.	Shell	$\frac{20}{110}$	NA	$\frac{22}{101}$	$\frac{12}{82}$	$\frac{10}{90}$
Agri.	Corbion	NA	$\frac{3}{5}$	$\frac{3}{7}$	NA	$\frac{2}{6}$

Note: shows the rank of the Dutch stocks average scedasis compared to the American stocks average scedasis in the "Dutch subsample". So, $\frac{19}{45}$ means from the 45 stocks the Dutch stock is number nineteen when sorted from riskiest to safest.

6 Conclusion and Discussion

6.1 Conclusion

In this paper we use the space-time trend model from Einmahl et al. (2022) to find patterns in the American and Dutch stock markets. With these patterns we try to find which sectors have the risky stocks and which have the safer stocks. This gives our research question: *Which sectors have more risky stocks according to the space-time trend model?* To answer this question we create 100 subsamples for the analysis. For each of the 100 subsamples we decluster the data to reduce the serial dependence of the data. Then we check the extreme value index of each subsample. We find that for 31 of the 100 subsamples, that the extreme value index is constant over time and in the cross-section. For the comparison, we analyse 28 subsample of these subsamples in seven sectors over five different time periods.

We intend to compare the subsamples with each other and find where the more riskier and safer stocks are. This means we compare their average scedasis. For each subsample we take a closer look at the stock with the highest average scedasis, which is the riskiest stock in the subsample. We also look at the stock with the lowest average scedasis, so the safest stock and the average of the entire subsample.

An extreme event causes the tail risk of the safer stocks to increase more than the tail risk of the riskier stock. The riskier stocks in an expansion period, are less impacted by a crisis than other stocks. The impact of the Great Recession has an extended effect on the scedasis of a riskier stock compared to a safer stock. Having homoscedasticity or heteroscedasticity in the tail does not make a distinction between the the high and low tail risk stocks. The number of stocks that exhibit heteroscedasticity in the tail for all subsamples is relatively small. The low presence of heteroscedasticity in the tail is caused by the nature of the stocks. For some stocks, their scedasis is always higher but the scedasis does not change much. The number of stocks with heteroscedasticity in the tail increase in the crises periods. The Great Recession causes a higher increase in the number of stocks with heteroscedasticity in the tail than Covid-19. We conclude that the impact of the Great Recession is bigger than Covid-19, because of the response time of market participants.

When we test for each subsample whether the frequency of extremes is constant in the cross-section, we state that the Human Genome Project has a substantial impact on the health research sector that makes us reject the null hypothesis of a constant total integrated scedasis of extreme returns in the cross-section. No matter the time period, the agricultural sector has different fre-

quency of extremes for all its stocks. Since, the aim of the agricultural sector is to provide food and drinks and these are necessary goods, the time period does not have an impact on the test results. Hence, the frequency of extremes in the agricultural sector is not impacted by a financial crisis or pandemic. We conclude for the consumer discretionary, technology, health consumer materials, industrial and energy sector that the frequency of extremes is similar for each stock, irrespective of the time period. We believe that the frequency of extremes is the same for all stocks in their subsample, as the companies in the related sector are perceived to be competitors. This is confirmed by the values of tail dependence which indicate high tail dependency. Hence, a crisis does not have an impact on the frequency of extremes for these five sectors.

When we test for each subsample whether there exists a trend over time, we find for all health research and agricultural subsamples that the frequency of extremes is constant over the sample period. In the health consumer materials and industrial subsamples the frequency of extremes is not constant over the sample period. For the consumer discretionary, technology and energy subsamples the conclusion differs per time period. There is no clear impact of a crisis on this test. The impact of both of the crises is different for each of the sectors and thus also for the frequency of extremes.

We compare the Dutch market with the American market, by adding the Dutch stock to the American subsample. The p-value of the Kolmogorov-Smirnov test for all the Dutch stocks within their respective subsamples are around 0.000. Hence, the frequency of extremes is not constant over time for all these stocks. The reason is that the Dutch stock market is relatively small and sensitive to extreme global events. Contrary to the Netherlands, the United States of America is a larger economy and the listed stock in the United States of America are less sensitive to extreme events outside their own country. We conclude that in case of non-crisis periods, Dutch stocks are riskier than American stocks. The Great Recession has a larger impact on the American economy than on the Dutch economy, consequently on the riskiness of the different stocks and sectors. For both extreme events, we conclude that the Dutch stocks become relatively safer.

6.2 Discussion

If there exists a new or better method compared to the declustering, where one can reduce the serial dependence even further, to weak serial dependence or no dependence at all, one can use that method instead of this declustering. For declustering, more research on what value to choose, not just arbitrarily choose one month, for rainfall there is literature which confirms the two days from Einmahl et al. (2022). More research on the bandwidth which is 0.1, this might not be the most optimal value. When it is possible to find a scedasis function that is monotonic, use the parametric estimation from Mefleh et al. (2020) on the scedasis function.

Since we have analysed the impact of the two past crises, applying this knowledge to predict what the impact will be from a future crisis. For example, with a potential re-emergence of a Covid-19 pandemic as already happening in China. We might be able to a certain extent to predict what the impact will be on the scedasis function. What can also play a role is that it again is a pandemic. It is also not long after the first pandemic which might have similar impacts. Another potential crisis can emerge, because of the collapse of the Silicon Valley Bank in March 2023. Consequently, the Credit Suisse takeover also causes some turmoil.

An idea for the analysis of the sectors can be the creation of ten new time series which are the means of each sector at a particular point in time. This small extension might add extra insight into the riskiness of each sector compared to one another. Another small extension can be the comparison of more small (Dutch) and large (American) markets. In such an extension, the researcher might find patterns for more small and large markets and given the countries the researchers chooses, maybe even differences between continents.

References

Balkema, A. A., & De Haan, L. (1974). Residual life time at great age. *Annals of probability*, *2*(5), 792–804.

Berghaus, B., & Bücher, A. (2018). Weak convergence of a pseudo maximum likelihood estimator for the extremal index. *Annals of Statistics*, *46*(5), 2307–2335.

Bücher, A., & Jennessen, T. (2022). Statistics for heteroscedastic time series extremes. *arXiv preprint arXiv:2204.09534*.

Cabral, R., Ferreira, A., & Friederichs, P. (2020). Space–time trends and dependence of precipitation extremes in north-western germany. *Environmetrics*, *31*(3), e2605.

Chaudhry, S. M., Ahmed, R., Huynh, T. L. D., & Benjasak, C. (2022). Tail risk and systemic risk of finance and technology (fintech) firms. *Technological Forecasting and Social Change*, *174*, 121191.

Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). *An introduction to statistical modeling of extreme values* (Vol. 208). Springer.

Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B (Methodological)*, *52*(3), 393–425.

de Haan, L., Tank, A. K., & Neves, C. (2015). On tail trend detection: Modeling relative risk. *Extremes*, *18*(2), 141–178.

de Haan, L., & Zhou, C. (2021). Trends in extreme value indices. *Journal of the American Statistical Association*, *116*(535), 1265–1279.

Einmahl, J. H., de Haan, L., & Zhou, C. (2016). Statistics of heteroscedastic extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *78*(1), 31–51.

Einmahl, J. H., Ferreira, A., de Haan, L., Neves, C., & Zhou, C. (2022). Spatial dependence and space–time trend in extreme events. *Annals of Statistics*, *50*(1), 30–52.

Embrechts, P., Resnick, S. I., & Samorodnitsky, G. (1999). Extreme value theory as a risk management tool. *North American Actuarial Journal*, *3*(2), 30–41.

Ferreira, A., Friederichs, P., de Haan, L., Neves, C., & Schlather, M. (2017). Estimating space-time trend and dependence of heavy rainfall. *arXiv preprint arXiv:1707.04434*.

Gilli, M., Këllezi, E., & Hysi, H. (2006). A data-driven optimization heuristic for downside risk minimization. *Swiss Finance Institute Research Paper*, (06-2).

Jones, M. C. (1993). Simple boundary correction for kernel density estimation. *Statistics and computing*, *3*(3), 135–146.

Kelly, B., & Jiang, H. (2014). Tail risk and asset prices. *Review of Financial Studies*, *27*(10), 2841–2871.

Leadbetter, M. (1983). Extremes and local dependence in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, *65*, 291–306.

Massacci, D. (2017). Tail risk dynamics in stock returns: Links to the macroeconomy and global markets connectedness. *Management Science*, *63*(9), 3072–3089.

McNeil, A. J. (1997). Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin: The Journal of the IAA*, *27*(1), 117–137.

McNeil, A. J. (1998). *Calculating quantile risk measures for financial return series using extreme value theory* (tech. rep.). ETH Zurich.

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of empirical finance*, *7*(3-4), 271–300.

Mefleh, A., Biard, R., Dombry, C., & Khraibani, Z. (2020). Trend detection for heteroscedastic extremes. *Extremes*, *23*(1), 85–115.

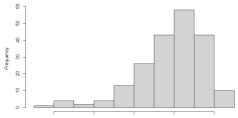
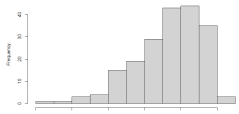
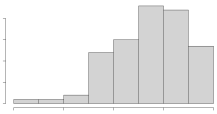
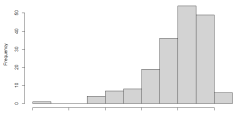
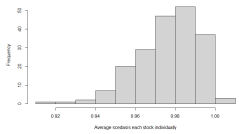
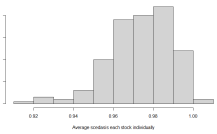
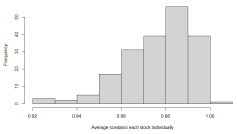
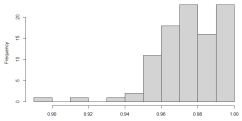
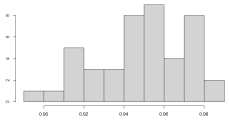
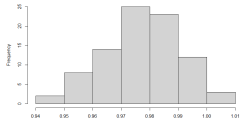
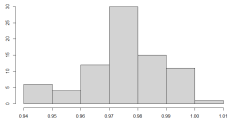
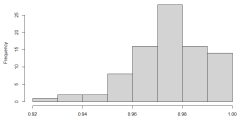
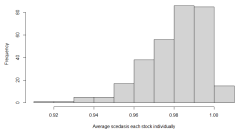
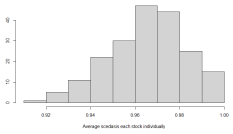
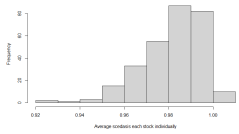
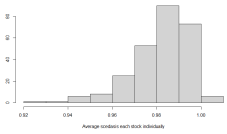
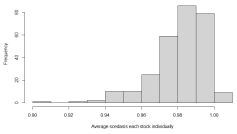
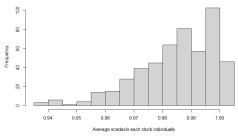
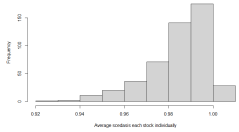
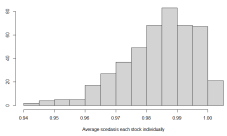
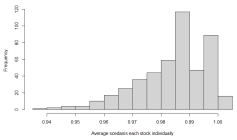
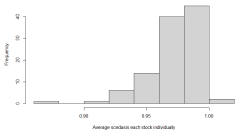
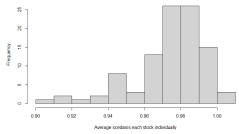
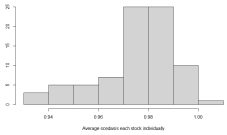
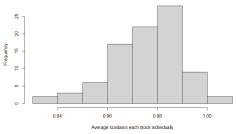
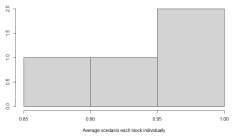
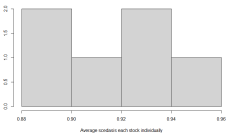
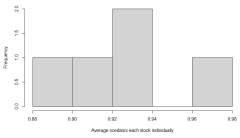
Pickands III, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 119–131.

Zhang, X., & Zhang, S. (2021). Optimal time-varying tail risk network with a rolling window approach. *Physica A: Statistical Mechanics and its Applications*, *580*, 126127.

A Appendix

A.1 Extra graphs

Table A1: Average scedasis for each stock individually

Sector	NT2	2000/2019	2007/2019	2007/2021	2009/2021
CD		NA			
Tech.	NA	NA			
HR					
HCM					
Indus.		NA			
Ene.		NA			
Agri.	NA			NA	

Note: For each stock in each subsample the average scedasis is plotted in a histogram.

A.2 Codes

A.2.1 American_cleaning

We load in daily data from the S&P 500 and covert it to weekly data. We choose a specific time period and sector. We test for the existence of serial dependence via a GARCH(1,1) and test for leptokurtosis. We search for the best block sizes. The most important step is the declustering of the data. After we have declustered the data, we test again for serial dependence and the best block sizes. We calculate the summary statistics of the declustered data. This declustered data is saved in a csv that is then used in the script of Section A.2.2 for the space-time trend model.

A.2.2 American_space_time_trend_model

For the coding of the space-time trend model we apply the code that is made accessible by Einmahl et al. (2022). This code is readjusted to fit the financial data. We load in the declustered data. First, we check for a constant extreme value index in the cross-section and then for a constant extreme value index over time. We apply the Kolmogorov-Smirnov test for a trend over time for each stock. Calculate the average scedasis of each stock and plot them in a histogram. Use these average scedasis to find the safest and riskiest stocks. Finally, we do two tests for the frequency of extremes in the cross-section and over time.

A.2.3 Dutch_cleaning

This code applies the same methodology as Section A.2.1. The only difference is, that in this code the Dutch data is added to the American data for comparison.

A.2.4 Dutch_space_time_trend_model

This code applies the same methodology as Section A.2.2. The only difference is, that in this code the Dutch data is added to the American data for comparison.