ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

The impact of risk model quality on porfolio Expected Shortfall

Econometrics and Management Science

Quantitative Finance

Master Thesis

Supervisors:
Prof. dr. D.J.C. van Dijk (EUR)
C. Meunier Charette (EY)

Author:
Casper Witlox (426233)

Second assessor:
P.A. Opschoor (EUR)

November 30, 2020

Abstract

This paper considers the way financial institutions attempt to cope with market risk, measured in terms of Expected Shortfall (ES). They do so by calibrating their risk models to the market as best as possible. Practically, parameters in the risk model are oftentimes misspecified with respect to the market. In this paper, we contribute to literature by examining the impact these misspecifications have on the ES framework. This is relevant since financial institutions are obliged to report daily ES estimates to comply to regulation of the Basel Committee of Banking Supervision (BCBS), thus they have to be able to accurately examine tail market risk they face. This examination concerns three aspects. First, we construct four portfolios using Monte Carlo simulations under different correlations and/or volatilities. Correlations are modeled using a rotated Gumbel copula, while volatilities follow a Threshold GARCH model. Second, we compare corresponding ES estimations to the market by analysing mean and distribution of excessive losses, together with market ES underestimations. We find that misspecifying volatility can have substantial impact on ES estimations for portfolios including stocks and options. The impact of misspecifying correlations is less notable for these portfolios. When correlation driven instruments, such as a basket option, are included, the margin in misspecifying correlations is also quickly violated. We conclude that a non-derivative portfolio has more margin when misspecifying parameters in the risk model. A portfolio that includes instruments driven by one (or multiple) of the factors in the risk model is heavily impacted by misspecification in (one of) these factors. Oftentimes in practice, both volatilities and correlations are misspecified, leading to an insufficient amount of capital being set aside for stocks only portfolios, and too much for other investigated portfolios[1].

Keywords: *Market risk modeling, Expected Shortfall, Tail risk management, Portfolio properties*

Acknowledgements

This thesis has been conducted under guidance and supervision of Prof. Dr. Dick van Dijk of the Erasmus School of Economics and Christophe Meunier Charette of the Financial Services Risk Management team at EY Amsterdam. Both of them have been vital in shaping the overall direction of the thesis by providing constructive feedback throughout its process.

[1] assuming misspecifications imply lowering volatilities and increasing correlations.

Contents

1 Introduction **1**

1.1 Impact of risk model quality on ES estimates . . . 1

1.2 Background: change in risk measure . . . 2

1.3 ES comparison framework . . . 4

1.4 Research question . . . 5

2 Literature **5**

3 Methods **6**

3.1 Monte Carlo simulation . . . 6

3.1.1 Modeling correlations . . . 7

3.1.2 Modeling volatility . . . 8

3.1.3 Modeling stock price . . . 9

3.1.4 Modeling option price . . . 10

3.2 ES estimation technique . . . 10

3.3 Coverage comparison . . . 11

3.3.1 Coverage test: Moldenhauer-Pitera . . . 11

3.3.2 t-test for mean difference . . . 13

3.4 Distribution comparison: Kolmogorov-Smirnov test . . . 13

3.5 Estimation comparison . . . 14

3.5.1 Size and frequency analysis . . . 14

3.5.2 Independence in underestimations: Ljung-Box test . . . 15

3.6 Traffic light framework . . . 15

4 Data **16**

4.1 Simulation scenarios . . . 16

4.2 Portfolios . . . 17

4.2.1 Portfolio 1: The Stock Portfolio . . . 18

4.2.2 Portfolio 2: The Call Portfolio . . . 18

4.2.3 Portfolio 3: The Downside-protected Portfolio . . . 18

4.2.4 Portfolio 4: The Basket Spread Portfolio . . . 19

5 Results **19**

5.1 Portfolio 1: Stock Portfolio . . . 20

5.1.1 Portfolio price paths . . . 20

5.1.2 Expected Shortfall . . . 21

5.1.3 Coverage comparison . . . 21

5.1.4 Distribution comparison . . . 23

5.1.5 Estimation comparison . . . 25

5.1.6 Traffic light framework 27
5.2 Portfolio 2: The Call Portfolio 28
5.2.1 Portfolio price paths 29
5.2.2 Expected Shortfall 29
5.2.3 Coverage comparison 31
5.2.4 Distribution comparison 32
5.2.5 Estimation comparison 34
5.2.6 Traffic light framework 35
5.3 Portfolio 4: The Basket Spread Portfolio 37
5.3.1 Portfolio price paths 37
5.3.2 Expected Shortfall 38
5.3.3 Coverage comparison 39
5.3.4 Distribution comparison 41
5.3.5 Estimation comparison 41
5.3.6 Traffic light framework 43

6 Conclusion and Discussion **43**

A Appendix **46**
A.1 Risk Axioms 46
A.2 Rotated Gumbel copula 46
A.3 dgp parameter values 46
A.4 TGARCH model: stability 47
A.5 Options: moneyness, payoff and the Greeks 48
A.6 Portfolio price transformation 49
A.7 Return distributions for Portfolio 3 and 4 49
A.8 Comparing simulated with actual market 49
A.9 Distribution comparison for lower volatility persistence for Portfolio 2 with break at $t = 125$. 50
A.10 Results Portfolio 3: The Risk-protected Portfolio 51
A.10.1 Portfolio price paths 51
A.10.2 Expected Shortfall 52
A.10.3 Coverage comparison 52
A.10.4 Distribution comparison 53
A.10.5 Estimation comparison 54
A.10.6 Traffic light framework 55
A.11 Distribution comparison for subscenarios for Portfolio 4 56
A.12 Autocorrelation in ES underestimations for Portfolio 4 57

1 Introduction

1.1 Impact of risk model quality on ES estimates

To be able to report daily Expected Shortfall (ES) estimates, financial institutions attempt to map market tail risk as adequately as possible. In practice, however, their risk models oftentimes are unable to perfectly estimate excessive losses institutions are exposed to, due to the market demanding a trade-off between model accuracy and complexity. If this happens more frequently, parameters in their risk models are misspecified with respect to the market, causing the model to be poorly calibrated to the market. This paper is on the impact these parameter misspecifications have on the ES a financial institution estimates. Besides, we examine this impact under different portfolio properties. Portfolios in this thesis are composed of different instruments, such as stocks, calls and puts, to investigate the impact of poor risk model quality for different portfolio characteristics.

To assess this impact, we first compute ES estimations under correctly specified market conditions and under a risk model that is willingly misspecified with respect to the market (under a scenario). Thereafter, results of the two procedures are compared. This comparison is performed for different portfolios to be able to take their characteristics into account. Our focus is rather on the impact of the model specification quality on ES estimations than on evaluating the performance of the ES estimation technique itself. The goal of this thesis is thus achieved not so much by applying many ES estimation methods to a single portfolio as it is by performing a single ES estimation technique under different model specifications for multiple portfolios. Besides, a more accurate analytical solution becomes hard to find after many iterations.

Our research is practically relevant for two main reasons:

- Financial institutions oftentimes do not take every single risk factor precisely accurately into account due to the market demanding a trade-off between model accuracy and complexity. At the same time, they have to be able to accurately map excessive losses they may face in order to comply to regulation from the BCBS. Our main findings indicate that misspecifying volatility can have substantial impact on corresponding ES estimation for portfolios including stocks and options and thus a financial institution's risk management. When correlation driven instruments are included, the margin in misspecifying correlations is also quickly violated. Our findings thus show that incorrectly specified risk models can result in unreliable ES estimations. A financial institution may no longer be able to sufficiently cope with market tail risk and comply to regulation, leading to reprimands by the regulator or even bankruptcy.

- Financial institutions are likely to hold a wide range of portfolios. Their properties tend to differ depending on the instruments they include. Our main findings show tail risk estimations depend on these portfolio properties, namely that a non-derivative portfolio has more margin when misspecifying parameters in the risk model. This implies risk model calibration should be more accurate for some portfolios to yield satisfactory ES estimations than for others.

Therefore, considering multiple portfolios provides insights that are of practical use for actual investment implementation.

During the financial crisis in 2008, we saw the relevance of being able to adequately capture tail market risk on a global scale. Internal market risk models performance was proven to not always be satisfactory, most notably from halfway 2007 up to the "Lehman year" of 2008 (Gaumert and Kemmer (2015)). During this period, numerous banks found daily loss limits projected by their models to be frequently significantly exceeded. For instance, Deutsche Bank showed losses on certain subportfolios were serious enough to impact the overall performance of the bank's trading unit. This stresses the extremely strong disruption that can follow from external shocks in the market, which resulted in a near collapse of the American financial system wiping out over 11 trillion dollars in household wealth (Nelson and Katzenstein (2014)). Besides, McAleer et al. (2013) found alternative risk models than those used to measure tail risk before and during the 2008 financial crisis to be optimal, indicating market risk forecasts could have been more satisfactory.

Although implications from this research are mostly relevant practically, there is also a scientific component to it. This thesis provides an understanding into how sensitive risk models are to misspecifications in their parameter settings. The sensitivity of the ES framework to model misspecification we obtain may be used to investigate calibration algorithms to asses their goodness of fit with respect to the ES framework. If one identifies the misspecifications that have least margin in terms of reporting ES estimations that are similar to the market, these results can be taken into consideration when setting up algorithms that best calibrate a risk model to the market.

1.2 Background: change in risk measure

In order for financial institutions to stay financially healthy, the Basel Committee on Banking Supervision (BCBS) has put in place regulatory frameworks to which a financial institution's risk management has to comply. The Basel Accords specify the amount of capital financial institutions are required to hold in order to guard against financial and operational losses they face (Hopkin (2018)).

Prior accords Basel I (1988) and Basel II (2004) stipulated that financial institutions are required to report daily Value at Risk (VaR) estimates. VaR is a risk measure that describes how much a portfolio may lose with a given probability in, for instance, a day[2]. Mathematically, we note VaR as follows:

$$\mathrm{VaR}_\alpha(R) := \inf\{x : F_R(x) \geq \alpha\}, \tag{1}$$

in which F_R is the cumulative distribution function of return variable R. Furthermore, $x \in R$ describes a certain return and $\alpha \in (0,1)$ describes the confidence level one wants to test for. Return

[2]Depending on prevailing market conditions.

R is described by a financial institution's profit and losses (P&L) probability density function. Excessive losses are in the left tail of the P&L distribution, which means α takes on small values to investigate tail risk in this thesis. Using Definition (1), one is able to assess potential losses a financial institution is exposed to relative to the amount of capital it possesses.

Nevertheless, VaR has been criticised as a risk measure over the years. This criticism boils down to two major issues. First, Artzner et al. (1999) noted that VaR is not a coherent risk measure, i.e. it does not comply to all 4 risk measure axioms depicted in Section A.1. Specifically, it fails the sub-additivity property. Second, VaR is unable to accurately capture the tail of the P&L distribution. Reason being is that Definition (1) only specifies the loss amount that will not be exceeded $1-\alpha\%$ of the time. It does not, however, indicate the possible magnitude of the loss in the $\alpha\%$ outlier cases, let alone their distribution.

These two issues moved the BCBS to adopt a new way of measuring risk: Expected Shortfall. This new measure has been agreed upon in Basel III and is set to be implemented in Basel IV (expected January 2022), which will require financial institutions to report daily ES estimates for the 2.5% quantile, using a one-tailed confidence level (BCBS (2016)). Patton et al. (2019) describe ES as the average loss, conditional on losses being above a specific quantile of the distribution, namely VaR_α. Mathematically, ES can be described as follows:

$$\text{ES}_\alpha(R) := \mathbb{E}(R|R \leq \text{VaR}_\alpha) = \frac{1}{\alpha}\int_0^\alpha \text{VaR}_u(R)\text{d}u, \quad \text{for } R \leq 0 \tag{2}$$

in which R, α and VaR_α correspond with those defined previously in Definition (1). ES calculates the expected loss given the fact that an institution's loss is above VaR. To be clear, this corresponds with a large negative return and concerns the left tail of the P&L distribution. Definition (2) shows ES can be interpreted as the average of portfolio losses that exceed the VaR_α threshold, as is in line with Patton et al. (2019)'s ES definition. For instance, for a confidence level of $\alpha = 5\%$, ES calculates the average loss of the 5% most extreme cases. This means ES is better capable of capturing tail risk properties. The BCBS stipulates financial institutions to use an ES of 2.5% (BCBS (2013)). Advantage of using this confidence level is that $\text{VaR}_{1\%}$ and $\text{ES}_{2.5\%}$ percentile values for a normal distribution are approximately the same empirically (Danielsson (2018)). Moreover, ES satisfies the properties in Section A.1 and is thus a coherent risk measure. Therefore, around early 2022, BCBS will require financial institutions to implement ES as a risk measure rather than the previously used VaR.

More generally, the implementation of ES as risk measure will have an enormous impact practically. For instance, to meet the soon-to-be implemented Basel IV requirements, McKinsey projects the European banking industry, or 130 European banks, will have to collectively raise additional capital of about €120bn (Schneider et al. (2017)). Besides, the European Banking Authority predicts this capital shortfall across the European banking sector is even greater, around €135bn (Davies (2019)). In short, the difference between VaR or ES as a risk measure is projected to have huge impact, not just for financial institutions separately but also for the financial world as a whole.

1.3 ES comparison framework

Our comparison of ES estimation results under different model specifications with the market is performed by evaluating the framework below.

- **Coverage comparison**: a coverage test and t-test for difference in mean excessive losses, or ES estimates per definition, under a scenario and the simulated market;
- **Distribution comparison**: perform a distribution test to compare the excessive loss distribution under a scenario with that under the simulated market;
- **Estimation comparison**: asses the size and frequency per month of over-/underestimations of the market and examine autocorrelation in market ES underestimations, if they occur.

The framework above helps us to i) investigate ES estimations under different model specifications and ii) compare these tail risk estimations to asses the impact of model quality. This framework extends current literature that focus only on point i) (Holton (2020)). We compare results by adding a t-test to the coverage test and a size and frequency of differences in ES estimation to the independence test. We perform multiple analyses, since they compare excessive loss estimations under the market with those under misspecifications differently. Specifically, they evaluate the difference of excessive losses in terms of mean, distribution and market ES underestimations. Performing multiple analyses thus helps us to gain a better theoretical perspective of the impact on the ES framework.

We expect the VaR threshold to be exceeded $\alpha = 2.5\%$ of total P&L observations, when ES estimations under the market are the same as under a poorly specified model (a scenario). Note that each analysis in our backtesting framework implicitly assumes VaR is correctly specified. Our research concerns the way ES is estimated from a correctly specified VaR, as ES per definition relies on the value for VaR and assumes the same quantity of VaR violations. Sensitivity of the number of violations, or VaR estimates, to changes in the underlying risk model is outside the scope of this thesis. We do not evaluate calibration algorithms ourselves, but rather assume a perfectly calibrated market model to benchmark our misspecified risk model results against. Lindholm (2014) namely states the complexity and implementation difficulties that come with accurately calibrating a risk model, or a local volatility surface in their case. We do touch upon this subject from a more theoretical perspective at the end of Section 2.

Differences in market tail risk are concluded based on statistical tests in our framework. These tests are conducted for multiple P&L distributions over a certain time series, for which we elaborate on the instances where they reject corresponding null hypothesis. This rejection is based on corresponding test statistics and p-values. Prior-going is considered to be common knowledge for the reader of this research, who is assumed to have an econometric background, specifically in the field of quantitative finance. We refer to Graham and Pal (2014) for an in-depth explanation on this statistical background, which will not be discussed further in this thesis.

1.4 Research question

This thesis determines the impact risk model quality has on ES estimations a financial institution reports. This is done by performing our ES comparison framework consisting of the following aspects: coverage, distribution and estimation comparison. This framework is applied under both different risk model specifications and portfolio characteristics. From this, we articulate the following twofold research question:

To what extend do quality of model calibration and portfolio properties impact ES performance by evaluating prior-mentioned ES comparison framework?

2 Literature

On the topic of VaR backtesting, intensive research has been conducted in the past. Most of this research shows the shortcomings of existing backtests for VaR. For instance, Christoffersen and Pelletier (2004) note that most existing VaR backtests have considerably small power. Therefore, they introduce a duration-based approach which has substantially better power properties. Escanciano and Olmo (2010) show that, in absence of estimation risk, unconditional backtests are impacted by model misspecifications. Ziggel et al. (2014) introduce some new VaR backtesting methods that improve upon existing ones. Conclusively, literature shows improvements into VaR backtests can be made, which stresses development of ES backtesting techniques may be relevant.

With respect to ES backtesting, research has been conducted more recently. Some researchers even believed ES could not be backtested (Carver (2013)). Main reason concerns the discovery that ES is not elicitable[3] as opposed to VaR (Patton et al. (2019)). This misconception was disproven by Acerbi and Szekely (2014) and Fissler et al. (2015) even refuted the claim that ES is not elicitable. Since then, more researchers have showed it is possible to successfully backtest ES. For example, Du and Escanciano (2017) perform Monte Carlo simulations to backtest ES on cumulative violations, where they apply the same analogy used for VaR backtests. Besides, Bayer and Dimitriadis (2019) propose regression based backtests. Gelling (2019) investigates numerous backtesting methods for ES and evaluates their size-adjusted power, computation time and implementation difficulty. In this research, Gelling (2019)'s best performing coverage test is applied, thus this research builds upon knowledge obtained from prior literature. Nonetheless, instead of proposing multiple (new) backtesting techniques, this thesis focuses more on the difference in backtested ES estimations. Therefore, more research into existing backtesting techniques is disregarded.

Research into ES estimation dates back to early 21^{st} century. Scaillet (2004) conducts research into ES estimation and sensitivity analysis accordingly, whereas Chen (2008) only considers the former. Furthermore, Yamai et al. (2002) and Taylor (2008) conduct research into both VaR and ES estimates. Nonetheless, we perform ES estimation as in Gelling (2019). Reason being it provides a

[3] A statistics $\psi(Y)$ of random variable Y is elicitable if it is the solution to minimizing the expectation of a scoring function S.

straightforward and easy-to-implement strategy. The conducted ES estimation is provided in Section 3.2.

In a financial context, model calibration is defined as setting parameters in a stochastic process in such a way that prices of financial instruments valued by the process replicate actual market prices (Lindholm (2014)). A wide range of different calibration procedures are used throughout literature. Windrum et al. (2007) discuss the following: indirect calibration, the Werker-Brenner approach, and the history-friendly approach, of which the first is most adopted by literature. Fagiolo et al. (2019) explain indirect calibration consists of four comprehensive steps[4], which makes implementation complex. Lindholm (2014)'s procedure of solving the ill-posed problem of calibrating a local volatility surface concerns two complex algorithms: a numerical quasi-Newton algorithm and optimal control theory techniques, both regularising the problem. Besides this complexity of calibration algorithms, lack of historical market data can cause calibration to be problematic (Gaumert and Kemmer (2015)).

3 Methods

This section is structured as follows. In the first subsection, the simulation of the portfolios is described, in which dependencies between stocks, modeled time-varying volatility and stock and option price simulations are thoroughly explained. The second subsection explains our ES estimations. In the third through fifth subsection, we lay out the ES comparison framework. The last subsection specifies the traffic light framework for ES comparison.

3.1 Monte Carlo simulation

We simulate the portfolios using Monte Carlo techniques. The underlying process of the actual market is unknown and actual stock market information is considered only to get an understanding of actual market prices. We simulate a single, hypothetically realised price path as our (simulated) actual market, so we assume we know the underlying market process. This main price path is thus not calibrated on actual stock market data, but inspired to be realistic by our understanding of the market. It allows us to have a known starting point instead of having to calibrate one, which would add substantially more complexity than is necessary for the purpose of this thesis. From this (simulated) actual market, we construct the simulated market as the next day price distribution based on today's actual market price path. This way, we obtain a range of theoretically possible portfolio price paths. Thereafter, parameters in this dgp are adjusted to investigate the impact that misspecifications in the dgp have on the ES framework compared to the simulated market. We thus regard the simulated market as a benchmark scenario.

[4]The first step concerns identifying the real-world factors of interest that one wants to calibrate the model on. Second, one should specify the following: the model, timeline of events, dynamic equations describing individual agents' behaviour, the parameter set and random disturbances involved. Third, one validates based on hypothesis testing the model to compare its outcome against real-world observations. Fourth, the agent-based model can be employed for policy analyses

The number of simulated stock price observations can be a tricky point. For an ES of 2.5%, as stipulated by the Basel committee, simulating over a period of one trading year (252 days) results in a small sample of about 6 price observations, because only losses that exceed VaR are relevant for our ES comparison. Small sample sizes usually cause statistical tests to have low power (Button et al. (2013)), thus we estimate the next day ES from a large number of simulations of the portfolio price for the next day[5].

In short, Monte Carlo simulations make it easier to test assumptions. For instance, if we simulate the true underlying stock price process to be normal (non-normal), testing the hypothesis for normality should pass (fail). With Monte Carlo simulation, we can sketch specific scenarios and compare tail risk estimations with the market for different risk models. This suits the purpose of this thesis well.

3.1.1 Modeling correlations

Copulas are suitable for generating data from multivariate distributions in presence of complex relationships between variables (Mathworks (2020)). Therefore, we use copulas to model the correlations between stocks. We use a 180 degrees rotated Gumbel copula to model these correlations. Gumbel copulas have an asymmetric dependence structure. A standard Gumbel copula shows high upper tail dependence (Manner (2010)). Its 180 degrees rotated version shows stronger lower tail dependence (Wali et al. (2018)). To illustrate this, a scatter plot of the dependencies between two uniformly distributed variables described by a 180 degrees rotated Gumbel copula is depicted in Section A.2. We investigate returns in the left tail of their distribution. Correlations in this part of the distribution have a larger impact. The market tends to correlate with negative impacts, thus portfolio prices also go down. Therefore, simulating from a copula which models higher left tail dependencies, such as a 180 degrees rotated Gumbel copula, describes actual market behaviour well.

A standard Gumbel copula is illustrated in the following equation:

$$C_{\theta}^{Gu}(u_1, ..., u_m) = \exp\Big(- \Big(\sum_{i=1}^{m} (-\log\, u_i)^{\theta} \Big)^{1/\theta} \Big), \tag{3}$$

where m represents the number of considered uniformly distributed variables u_i. Parameter $\theta \in [1, \infty)$ indicates the dependence level between the considered variables. We assign three groups of stocks and assign different θ values for each. These values are inspired by actual market correlations and are depicted in Section A.3. Dependencies between these buckets of stocks do exist, but are not explicitly modeled.

[5] Specifically, $5,000$ simulations over a one trading year period, which results in $5,000 \cdot 252 = 1,260,000$ observations for the entire P&L distribution per portfolio.

We use the following definition of Kolev et al. (2015) for the 180 degrees rotated Gumbel copula:

$$C_\theta^{RGu}(u_1,...,u_m) = \sum_{i=1}^{m} u_i - m + 1 + C_\theta^{Gu}(1-u_1,...,1-u_m), \tag{4}$$

in which $C_\theta^{Gu}(1-u_1,...,1-u_n)$ is the standard Gumbel copula described in Definition (3), which also defines the other variables in Definition (4). We include simulations from the rotated Gumbel copula in the distribution of variable z_t, which is implemented both in Equation (5) directly and via r_t in Equation (8). This way, it models the dependencies in the noise of both the stock price and volatility simulations.

3.1.2 Modeling volatility

We model volatility using a Threshold GARCH (TGARCH) model. Reason being is a TGARCH model is able to take both time-variation of volatility and volatility clustering into account. The latter phenomenon describes the tendency of big changes in prices to cluster together, which causes persistence of these magnitudes of price changes. Stock price volatility typically adheres to this (Cont (2007)). We use a TGARCH model rather than a standard GARCH model, as a TGARCH model is able to capture certain stylized facts of returns better (Wu (2010)), although admittedly other variations of GARCH models would also be able to explain stock price volatility (Katsiampa (2017)). Specifically, stock price distributions typically exhibit negative skewness. Therefore, we model from an assymmetric distribution for volatility. With respect to the number of lags, we consider TGARCH(1,1) model. No further lags are included as TGARCH(1,1) is less complex while still being sufficiently capable of capturing volatility clustering (Brooks and Burke (2003)). A TGARCH(1,1) model can be interpreted as follows: losses that exceed a certain threshold have a larger impact on volatility persistence than equivalently large positive returns do.

We specify a TGARCH(1,1) model for all m considered stocks as follows:

$$r_t = \sigma_t z_t, \qquad z_t \sim t_\nu^{-1}(u_i) \qquad \text{for } i = 1,...,m \tag{5}$$

$$\sigma_t^2 = \omega + \alpha z_{t-1}^2 + \beta \mathbb{I}_{t-1} z_{t-1}^2 + \delta \sigma_{t-1}^2, \tag{6}$$

$$\mathbb{I}_{t-1} = \begin{cases} 1 & z_{t-1} < -\kappa, \\ 0 & z_{t-1} \geq -\kappa, \end{cases} \tag{7}$$

in which variance and noise at time t are illustrated by σ_t^2 and z_t respectively. The t-distribution from which z_t is drawn has ν degrees of freedom. This implies stock prices are far off from being normally distributed, which is a reasonable assumption (Park and Kim (2016)). Coefficient ω

describes the general volatility level and α the persistence of lagged squared noise z_{t-1}^2 on variance σ_t^2. The dependency of variance σ^2 at time t on its lag $t-1$ is described by δ. Indicator function $\mathbb{I}_{t-1}$ indicates whether or not the return at time $t-1$ drops below a certain threshold. In case this happens, i.e. when a loss of a certain magnitude occurs at time $t-1$, it will impact the volatility at time t by an additional β percent. Whether this volatility is positively or negatively impacted, depends on the sign of coefficient β. In case of a positive, significant coefficient β, this means large losses have a bigger impact on volatility than large positive returns do.

We set parameters α, β and δ such that the volatility model is stable. The condition these parameters should comply to is derived in Section A.4. General volatility level ω is the same for all stocks. For the other TGARCH coefficients, we divide stocks into three buckets and assign a specific coefficient for each. This way, we set different model specifications for different stocks, which helps us to answer our research question more thoroughly. Section A.3 specifies the exact parameter values for the TGARCH model.

We implement correlations between stocks in the noise of the model z_t via variables u_i in Equation (5). These are simulated from the rotated Gumbel copula from Equation (4). After simulating from a TGARCH(1,1) model, we include the obtained volatility series in Equation (8) for the stock price.

3.1.3 Modeling stock price

We simulate stock prices daily, assuming a number of 252 trading days in a year. We assume the simulated stock prices follow a Geometric Brownian motion. Therefore, we simulate the stock price paths from the following formula:

$$S_t = S_{t-1} \exp\big((r - 0.5\sigma_t^2)dt + \sigma_t \sqrt{dt}\big), \quad \text{for } t = 1, ..., T \tag{8}$$

where S_t describes all m stock prices at time t, r the risk-free rate and σ_t the stock volatility. Moreover, dt describes the time intervals over which the portfolios will be simulated, which equals 1/252 for a daily period. Furthermore, r_t corresponds to the variable in Equation (5) and T is the simulation time horizon.

Equation (8) implies the mean stock price grows at a rate of $(r - 0.5\sigma_t^2)dt$. Furthermore, it implies stock prices are log-normally distributed in case volatility is constant and returns are normally distributed (Hull et al. (2009)). Volatility σ_t from Equation (6), however, is time-varying and z_t in Equation (5), is drawn from a t-distribution with only 8 degrees of freedom.

We initialise the m stock prices in S_0 to be 100 and the risk-free rate r equals 1%. Prior-mentioned volatilities and correlations between stocks are incorporated in Equation (8) by adjusting σ_t and z_t's distribution respectively.

3.1.4 Modeling option price

We value the option prices from a Black-Scholes model. This model for call and put options is specified by the following formulas:

$$\text{CALL}_t = S_t \Phi(d_{1,t}) - e^{-r\tau} K \Phi(d_{2,t}), \tag{9}$$

$$\text{PUT}_t = e^{-r\tau} K \Phi(-d_{2,t}) - S_t \Phi(-d_{1,t}), \tag{10}$$

$$d_{1,t} = \frac{\log(S_0/K) + (r + \sigma_t^2/2)\tau}{\sigma_t \sqrt{\tau}}, \tag{11}$$

$$d_{2,t} = d_{1,t} - \sigma_t \sqrt{\tau}, \tag{12}$$

where CALL_t, PUT_t and S_t describe prices for calls, puts and stocks respectively at time t. Strike prices are defined by K. The risk-free rate is described by r and σ_t is TGARCH-modeled stock volatilities. The time to maturity is expressed by $\tau = T - t$, with T the simulation horizon. We include call and put options simulated from this Black-Scholes model in those portfolios that contain options, described in Section 4. We refer to Section A.5 for a detailed explanation on moneyness, payoff and behaviour of call and put options.

3.2 ES estimation technique

Our procedure for calculating ES predictions is described as follows. First, we perform prior-mentioned Monte Carlo simulations to obtain a time series of portfolio prices. We incorporate TGARCH modeled volatilities and dependencies based on a rotated Gumbel copula in these simulations. Second, we transform prices to portfolio returns, described in Section A.6. This way, we simulate 5,000 next day returns based on today's price for each time instance. The price path of the single current price is defined as the main price path.

We have now obtained 5,000 next day portfolio returns for every 252 trading days. This results in 252 daily distributions of 5,000 portfolio returns each. Each of the 5,000 simulated future returns is calculated from the same initial daily return today. Therefore, they follow the same distribution, so we can take the 2.5% quantile over all simulations to set the 2.5% VaR threshold. Note that this concerns the left tail of the P&L distribution. A large loss is in this thesis is defined as a positive number, corresponding to a negative return.

Now, we are able to determine which of the simulated future losses exceed the VaR threshold and take the mean over all simulations of these excessive losses. This results in a daily 2.5% ES time series over 252 days for each portfolio. Per definition, this daily ES series exceeds the daily VaR threshold. We calculate this ES time series for both the simulated market as well as the simulation scenarios. Lastly, the ES comparison framework is evaluated to compare these two series.

Our procedure to obtain an ES time series under a scenario is depicted in the pseudo algorithm below.

Algorithm 1 ES calculation procedure

Require: n: number of simulations.
Require: m: number of stocks.
Require: T: simulation time horizon.
Require: pf: number of portfolios.
Require: nsc: number of scenarios.
Require: ψ_0: instrument parameters.

for each time step t
  V_t^*: simulate a main today's volatility path.
  S_t^*: simulate a main today's stock price path.
  P_t^*: simulate a main today's portfolio price path.
end

for each of nsc scenarios
  set randomness as default.
  η_{sc}: initialize scenario copula parameters.
  θ_{sc}: initialize scenario TGARCH parameters.

  for each time step t
    $Z_t \leftarrow f(\eta_0, n, m)$ (simulate n variables from a rotated Gumbel copula for each of m stocks)
    $V_{t+1} \leftarrow f(\theta_0, Z_t, V_t^*, n, m)$ (simulate n stock volatilities using TGARCH for each of m stocks)
    $S_{t+1} \leftarrow f(\psi_0, S_t^*, V_{t+1}, Z_t, n, m)$ (simulate n next day stock prices following Geometric Brownian motion for each of m stocks)

    for each of pf portfolios
      $q \leftarrow f(m)$ (select a number of m stocks)
      $P_t \leftarrow f(\psi_0, \eta_0, S_t, V_t, q, n, m)$ (simulate n prices depending on instruments in portfolio)
      $R_t \leftarrow (P_t - P_{t-1}^*)/P_{t-1}^*$ (transform n prices into n returns)
      $VaR_t \leftarrow f(R_t)$ (set VaR threshold as 2.5% quantile over n simulated returns)
      $ES_t \leftarrow \text{mean}(R_t \leq VaR_t)$ (take mean over all losses exceeding VaR)
    end
  end
end

return ES_t (ES time series under a scenario).

3.3 Coverage comparison

3.3.1 Coverage test: Moldenhauer-Pitera

Our coverage test gives an indication of whether sufficient capital is reserved to cover the risk of a portfolio. Coverage in this case describes the percentage of observations that are in the 2.5% quantile.

This test is performed for both the market and scenario, thus we can compare their excessive loss quantiles. The coverage test only provides an intuition, a t-test is used to test statistical difference in mean excessive losses (Section 3.3.2).

We conduct the coverage test from Moldenhauer and Pitera (2019), since it has highest size-adjusted power compared to multiple other tests. These include Graham and Pal (2014)'s saddle-point method, Righi and Ceretta (2013)'s truncated distribution dispersion, Löser et al. (2018)'s Irwin-Hall transformation, Bayer and Dimitriadis (2019)'s regression-based backtest and Del Brio et al. (2017)'s t-test. Moreover, this test is easy to implement and performs quickly in terms of computation time (Gelling (2019)).

The Moldenhauer-Pitera test constructs a variable y_t to analyse coverage under both the market and the scenario. It combines portfolio return and ES prediction. We construct this variable as follows:

$$y_t := 1 - \frac{R_t}{\hat{\mathrm{ES}}_t^{\alpha}}, \quad \text{for } t = 1, ..., T, \tag{13}$$

in which R_t describes the portfolio return at time t, as calculated in Section A.6. Estimate $\hat{\mathrm{ES}}_t^{\alpha}$ concerns the daily 2.5% ES of portfolio returns. We define a loss as a negative number, thus $\hat{\mathrm{ES}}_t^{\alpha}$ attains negative values. Variable y_t is constructed for each simulation, $5,000$ times in total.

From this, we define the null hypothesis H_0 and test statistic MP_n as follows:

$$H_0 : \sum_{i=1}^{\alpha T} y_{(i)} = 0, \tag{14}$$

$$MP = \frac{1}{n} \sum_{i=1}^{n} \mathbf{I}_{\{y_{(1)}+...+y_{(i)}<0\}}, \tag{15}$$

in which $y_{(i)}$ represents the i-th order statistic of variables $y_1, ..., y_n$. Variable α denotes the ES confidence level and T the number of simulated time observations. Test statistic MP counts the number of most extreme losses it takes for their sum to be greater than the sum of corresponding ES predictions. Thereafter, we take the average over all simulations to construct the test statistic.

Consequently, we calculate the sum of most excessive losses that still remains more negative than the summed corresponding ES predictions. Under the null hypothesis, the sum of the lowest α% of time observations equals 0. This implies the sum of $\alpha \cdot T$ most excessive losses equals the corresponding ES prediction. The null hypothesis implicitly states that the loss percentile that provides the same ES value as was predicted is smaller than or equals α. The test statistic is Bernoulli distributed, as the sum of most extreme losses is either greater than the sum of ES predictions or not (success or failure). We thus accept the null hypothesis in case: $MP \leq \alpha$.

3.3.2 *t*-test for mean difference

The coverage test is a count of how many excessive losses it takes for their sum to exceed the sum of corresponding ES estimations. This is done both under the simulated market and under the scenarios. Since the ES estimation under the market is based on a different P&L distribution than under a scenario, comparing the two based on our coverage test results does not help to conclude whether the two are significantly different. To solve this issue, we introduce a two-sample t-test to unequivocally say whether or not the excessive losses differ in terms of mean.

The formula below specifies the two-sample t-test for unequal variances, or Welch t-test, of these samples.

$$t = \frac{\bar{R}_m - \bar{R}_1}{\hat{s}}, \quad \text{with } R \leq VaR_\alpha(R). \tag{16}$$

where $\bar{R}_m$ describes the sample mean of $5,000$ simulated returns, as calculated in Section A.6 under the simulated market and $\bar{R}_1$ describes this sample mean under the scenario. Both sample means only consider those losses that exceed the 2.5% VaR threshold of their corresponding P&L distribution. Since the variance of both samples differs, we introduce the following formula for the estimator of the total standard deviation $\hat{s}$:

$$\hat{s} = \sqrt{\frac{(n_m - 1)s^2_{R_m} + (n_1 - 1)s^2_{R_1}}{n_m + n_1 - 2}}, \tag{17}$$

in which n_m and n_1 report sample sizes of losses exceeding VaR under the simulated market and scenario. Their variances are described by $s^2_{R_m}$ and $s^2_{R_1}$. We consider those losses exceeding the 2.5% VaR threshold for each of the $5,000$ simulations. Therefore, we know $n_m \approx n_1 \approx 125$ per simulation. This test statistic follows a t-distribution with 124 degrees of freedom $(n_m - 1)$.

3.4 Distribution comparison: Kolmogorov-Smirnov test

We introduce the distribution test to compare the distribution of losses exceeding VaR under the simulated market with this distribution under each scenario. This way, we add another insight to answer our research question besides the mean difference examined by our t-test.

We apply the two-sample Kolmogorov-Smirnov test as distribution test. It quantifies the difference between two distributions and is widely used throughout literature as a goodness-of-fit test, for instance by the European Banking Authority (EBA (2019)). The test does not make any assumptions about the distributions that are compared. Since losses in the tail do not follow a known distribution, this test is suitable here. We use this test to quantify the difference in tail risk distribution under the scenario and simulated market. This way, we measure the impact different parameter specifications in the dgp have on the distribution of losses exceeding VaR.

Under the null hypothesis, both excessive loss distributions are the same. These may be continuous, discrete or a mix of the former two. Mathematically, we state the null hypothesis H_0 and test statistic D_{R_m,R_1} as follows:

$$H_0 : F_{R_m} = F_{R_1}, \quad \text{with } R \leq VaR_\alpha(R), \tag{18}$$

$$D_{R_m,R_1} = \sup_k |\hat{P}(R_m \leq k) - \hat{P}(R_1 \leq k)|, \quad \text{for } k \in (-\infty, \infty) \tag{19}$$

in which R_m describes $5,000$ simulated returns as calculated in Section A.6 under the simulated market and R_1 describes those under the scenario. Both variables only consider those losses that exceed the 2.5% VaR. Therefore, distributions F_{R_m} and F_{R_1} only concern the left tail of the entire portfolio P&L distributions. Test statistic D_{R_m,R_1} quantifies the discrepancy between samples R_m and R_1 in terms of distribution. Birnbaum (1952) notes that D_{R_m,R_1} is a "distribution-free" statistic, since its behaviour depends on the two distributions that are compared.

We note that this test requires a number of about 50 observations to confidently reject the null hypothesis. As we simulate $5,000$ times, this results in about 125 losses exceeding the 2.5% VaR threshold, which is large enough for the test to have sufficient power.

We do not compare the excessive loss distribution to a known distribution. We note that testing the distribution of a portfolio most likely yields different results than testing the distribution of individual stocks or instruments. For example, if noise is drawn from a normal distribution, this implies future prices of the underlying stock follows a log-normal probability distribution. Nonetheless, an option on this stock is not log-normally distributed in this case as a maximum function is applied to the stock price function. Therefore, the distribution of the portfolio as a whole is only know in a few simple cases.

3.5 Estimation comparison

3.5.1 Size and frequency analysis

The coverage and distribution comparisons support in drawing conclusions about the impact of model misspecifications by investigating the mean excessive losses and their distribution. Nevertheless, we also perform a size and frequency analysis to obtain a better understanding of the over- and underestimations of the simulated market ES estimations. This analysis is performed by investigating a scatterplot of over- and underestimations, together with corresponding histograms. The size is calculated as the difference in ES estimation under the simulated market and the scenario, so underestimations are assigned a negative size. Frequency is visualised as a count of how many over-/underestimations occur within a rolling window of one month (20 trading days) over the whole simulation horizon. This way, we can asses the size and frequency of over-/underestimations to gain insights into how significantly the misspecifications cause the ES framework to be violated.

3.5.2 Independence in underestimations: Ljung-Box test

In case underestimations occur, we are interested in investigating the independence in the events of a simulated market ES underestimation. In case of autocorrelated ES underestimations, this means that if you underestimate the ES at one instance, you are likely to do so at the next. The model does not adapt adequately to changes in the market in this case and one could conclude that the risk model under misspecifications underestimates market risk not coincidentally, but systematically. In practice, we would like these underestimations to be independent from one period to the next.

We conduct the Ljung-Box test to investigate autocorrelation in the daily ES time series of events, i.e. when the market ES is more negative than the scenario ES. We test for each of up to 10 lags over a window of 20 days, or one trading month. We perform this rolling window estimation over the whole trading year for a significance level of 2.5%.

Burns (2002) defines a p lag autocorrelation statistic y_p, which is used to compute the V lag test statistic as follows:

$$y_p = \frac{\sum_{t=p+1}^{n} x_t x_{t-p}}{\sum_{t=1}^{n} x_t^2}, \tag{20}$$

$$LB_V = n(n+2) \sum_{i=1}^{V} \frac{y_i^2}{n-i}, \tag{21}$$

in which x_t describes a time series variable of length n. In this thesis, x_t describes the event of a simulated market ES underestimation with respect to the scenario ES series. In other words, x_t attains a value of 1 if market ES is more negative than the scenario ES at time t. The number of included lags p equals 1,...,10.

We expect ES underestimations to show autocorrelation. Reason being is the portfolio prices exhibit heteroskedasticity, which is modeled using a TGARCH model. This replicates real-life volatility clustering well, which implies ES estimates tend to be autocorrelated. Therefore, we expect that if a simulated market ES underestimate occurs at time t, it is likely to occur at time $t+1$ as well.

3.6 Traffic light framework

Moldenhauer and Pitera (2019) provide a traffic light framework for VaR for the results of their coverage test. Their traffic light evaluates the performance of the test in terms of size-adjusted power, computation time and implementation complexity. Our traffic light does not. Here, we assess to what extend tail risk differs under the simulated market compared to under the scenario. Specifically, it looks at how much losses exceeding VaR differ in terms of mean, distribution and it evaluates autocorrelations in simulated market ES underestimations. Therefore, our traffic light helps determine whether a firm's risk modeling is well able to capture market risk.

The following conditions specify which traffic light is assigned at time t:

	Coverage comparison	Distribution comparison	Estimation comparison
Zone		Condition	
green		$p > 2 * \alpha$	
yellow		$\alpha < p < 2 * \alpha$	
red		$p < \alpha$	

Variable p describes the p-value at time t for the t-test, Kolmogorov-Smirnov and Ljung-Box test. Their p-values are compared against confidence level $\alpha = 2.5\%$.

Table 1: Traffic light framework for similarity between simulated market and scenario.

We compare the p-value of the test to confidence level α for the t-test, Kolmogorov-Smirnov and Ljung-Box test. A red light is shown when the null hypothesis of the test is rejected. Conclusively, our traffic light framework visualises the extend to which the null hypothesis of the three tests is violated. It does not visualise the performance of each test, but the quality of the model specification by its impact on the ES framework.

4 Data

4.1 Simulation scenarios

In this subsection, we explain different simulation scenarios. A scenario implies we set different parameters in the dgp. Randomness in simulation is set the same between scenarios. Reason being we can better filter out the impact that is caused by the parameters specifically. Later on, we compare the results of these scenarios to those of the simulated market. Below, the scenarios are described.

Scenario	Description
actual market	hypothetical main price path
simulated market	same parameters as in actual market
higher correlations	increase a copula parameter
lower volatility persistence	decrease a TGARCH parameter
mixture	a combination of higher correlations and lower volatility persistence.

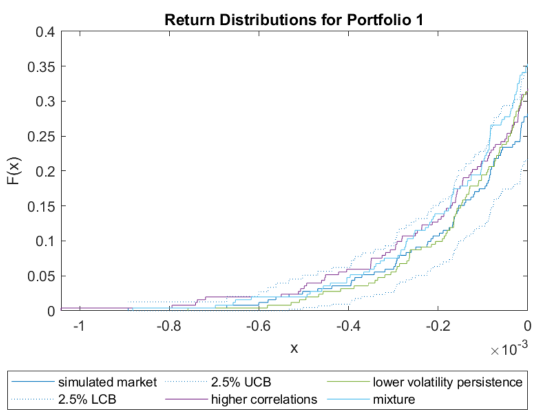
Table 2: Model specifications applied in constructing portfolios

As explained in Section 3.1, the actual market describes a single, hypothetically realised price path, or main price path. We assume we know the underlying price process, so we can construct a simulated market as if it was perfectly calibrated to this actual market. In the simulated market, we simulate the next day price distribution based on today's main price path. This way, we obtain a range of theoretically possible portfolio price paths, which we regard as our benchmark scenario. From there, parameters in this dgp are adjusted to investigate the impact that misspecifications in the dgp have on the ES framework compared to the simulated market. We refer to Section A.3 for the exact parameter values.

Copula parameters are adjusted to examine the impact of higher correlations between stocks. We adjust the TGARCH model to be able to research the impact of lower volatility persistence in stocks. Besides, a mixture between these two scenarios is investigated to capture both impacts.

We view the mixture scenario as the main scenario of interest. In practice, multiple parameters tend to be misspecified in a financial institution's risk model instead of a single one. The mixture scenario thus reflects actual risk management best. Higher correlations and lower volatility persistence are considered as subscenarios, highlighting specific aspects of the mixture scenario to obtain a better understanding.

Below, the return distributions under our misspecification scenarios for Portfolio 1 and 2 are visualised. Returns concern negative returns exceeding the VaR threshold. The dashed lines visualise the 2.5% lower and upper confidence bounds of the simulated market.



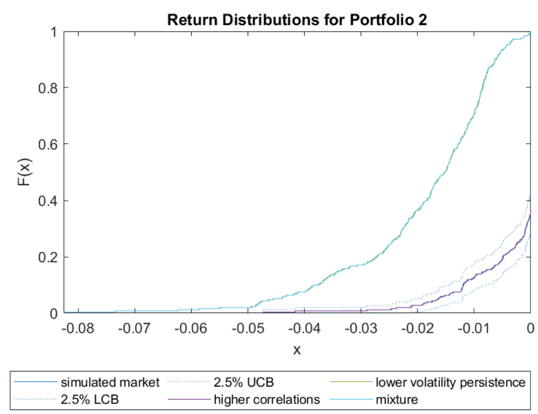


Figure 1: Return distributions for Portfolio 1 (left) and Portfolio 2 (right) for all scenarios.

The excessive loss distribution under lower volatility persistence tends to be similar to the mixture scenario. From this, we suspect that the mixture scenario may be largely dominated by the decrease in volatility persistence. Both scenarios differ substantially from the simulated market, especially for Portfolio 2. We suspect changing volatility persistence has a bigger impact on Portfolio 2. These first intuitions are investigated further in Section 5. Excessive loss distributions for other portfolios are visualised in Section A.7.

4.2 Portfolios

In the following subsections, the composition of the considered portfolios will be explained. They are listed such that Portfolio 1 investigates the impact of a scenario on stocks, Portfolio 2 the impact on options and Portfolio 3 is a combination of the former two. Portfolio 4 investigates the impact on a more complex portfolio composition. Each component in the portfolio is equally weighted and prices are simulated using Monte Carlo techniques, as explained in Section 3.1.

4.2.1 Portfolio 1: The Stock Portfolio

This portfolio is most fundamental, since only stocks are included. Analysing stocks provides us a basic insight into the different price paths and ES series that we can obtain, without having additional protections and/or limitations. In total, 100 stocks are included to investigate the impact of their mutual dependencies.

4.2.2 Portfolio 2: The Call Portfolio

Portfolio 2 is a derivatives only portfolio, solely consisting of call options. Stocks are not included. This portfolio is included to gain basic insights into the price path and ES behavior of option only portfolios. They are of interest since their price paths are driven by fluctuations in volatility. We therefore expect this portfolio to be impacted most by the lower volatility persistence scenario compared to other portfolios.

All call options are on the same underlying assets as those in Portfolio 1. This time, however, 10 of these 100 assets are selected as underlying stocks. The calls mature after 315 trading days, i.e. after an additional 25% of the simulation time horizon, and their strikes equal 100.

4.2.3 Portfolio 3: The Downside-protected Portfolio

This portfolio consists of both put options as well as their underlying assets. The same 100 assets from Portfolio 1 are included. Therefore, Portfolio 3 can be viewed as a combination of Portfolio 1 and 2.

Each underlying asset is accompanied by a put option on these underlying stocks. All put options mature after 504 trading days, i.e. after twice the simulation time horizon, and their strike prices equal 90. In Section 3.1.4, the valuation of these puts is explained. The price path of this portfolio fluctuates, but is floored at the strike price of the puts. This means that if the stock price drops below this strike, the puts ensure that portfolio value remains stable until maturity, if they expire in the money. The value of the puts compensates the loss of value in the stocks. This portfolio is thus floored: losses occur but they are limited, which means the portfolio is protected against downside risk. We include this portfolio to investigate the effect that a downside protection has on the ES framework. In short, we add another portfolio property to answer our research question.

4.2.4 Portfolio 4: The Basket Spread Portfolio

This portfolio consists of a single instrument. Its price depends on the same 10 underlying stocks from Portfolio 2. We describe the portfolio prices as follows:

$$BS_t = N \cdot \frac{1}{m}\sum_{i=1}^{m} \max\big(S_{i,t} - K, 0\big) - \max\big(\frac{1}{m}\sum_{i=1}^{m} S_{i,t} - K, 0\big), \quad \text{for } t = 1, ..., T \tag{22}$$

where m equals the number of underlying stocks considered, 10 in this case. We obtain the first part from Black-Scholes in Section 3.1.4 and obtain the second part by simulation. Both parts have strikes K equal to 100 and we multiply the first part by a factor $N = 80\%$. Both parts in this instrument mature after 315 trading days. Stock price S_t is simulated as in Section 3.1.3. Other variables in this portfolio are the same as those in Equation (9).

We interpret this instrument as the spread between a portfolio of call options and a basket option, or call option on a portfolio of stocks. Having a spread between two similar parts eliminates some factors, like delta or theta, partially or totally. The volatility of the basket is likely to be higher due to correlation than of the portfolio of calls. Reason being the call on a portfolio of stocks gains intrinsic value the more correlated the underlying stocks are. The basket option can thus be interpreted as a long position in correlation in underlying stocks. We therefore expect this portfolio to be impacted most by the higher correlations scenario compared to other portfolios. Since its price is expected to be highly impacted by correlation, this portfolio adds another portfolio property to our research, which we can investigate to answer our research question.

5 Results

The main thread of this section discusses the results for the mixture scenario. This scenario consists of a mixture between higher correlations and lower volatility, whose results are visualised in the outlined text to support the findings in the main thread. Results are discussed per portfolio. We obtain the portfolio price paths and daily ES time series for each portfolio under each scenario. Thereafter, the ES comparison and traffic light frameworks are evaluate for the scenario against the simulated market. This is done both visually through graphs as well as numerically by analysing test statistics. This way, we assess the impact that different dgp parameters have on the ES comparison framework for each portfolio. Section A.8 supports our believe that the simulated market exactly replicates the actual market. Therefore, we can compare results with the simulated market instead of the actual market.

5.1 Portfolio 1: Stock Portfolio

In this subsection, the portfolio price paths, ES time series and the ES comparison and traffic light frameworks for Portfolio 1 are illustrated.

5.1.1 Portfolio price paths

The portfolio price simulation paths for Portfolio 1 under the simulated market and under the mixture scenario are depicted below.

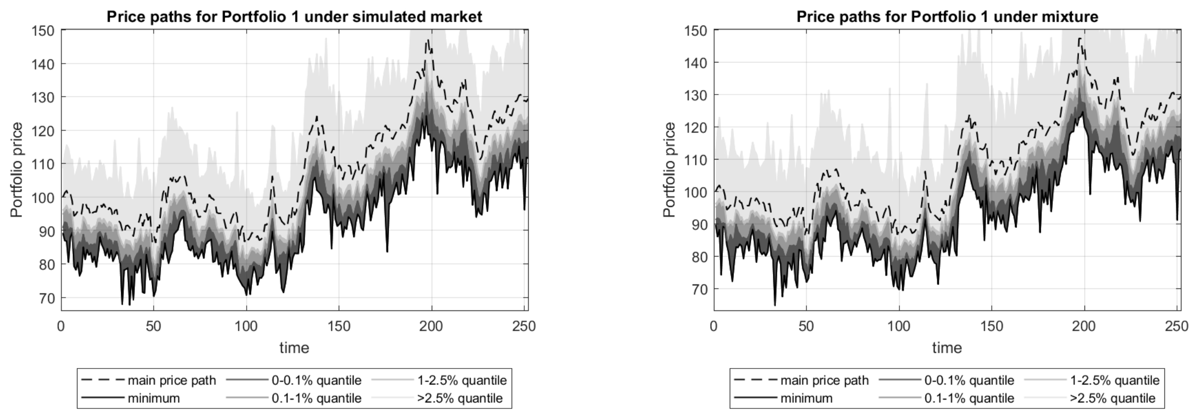


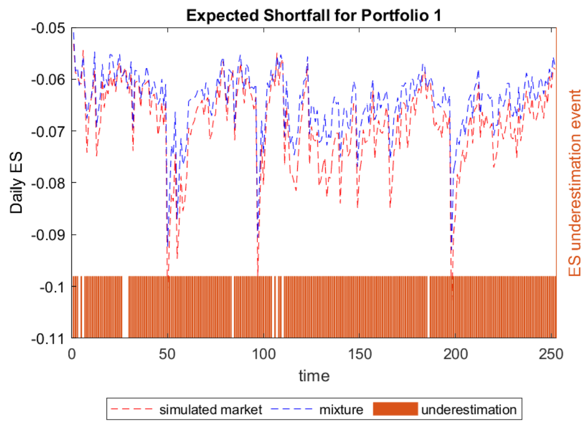
Figure 2: Portfolio price simulations for Portfolio 1 under the simulated market (left) and mixture scenario (right)

We note that the portfolio price paths tend to fluctuate but follow an upwards trend. This trend depends on the seed and randomness of the stocks included. We note it is upwards, however, this is not systematically. We note that highest upward spikes possible tend to be larger under the mixture scenario than under the simulated market. This is caused by higher correlation parameters under the mixture scenario. When these parameters increase, upper tail dependencies modelled by the rotated Gumbel copula increase too. Therefore, when a stock price goes up, the other stocks in the portfolio go up in price as well. Since lower tail dependencies are already fairly high under the rotated Gumbel copula, increasing correlation parameters has less impact on the magnitude of the downward spikes. For a visual representation of these tail dependencies, we refer to Section A.2. In short, under the mixture scenario we do not notice large changes in the minimum price path.

Nevertheless, we are interested in comparing the 2.5% lowest quantile between the market and the scenario, because this quantile concerns our ES estimation. Since the two are hard to compare visually, we calculate the width of the 2.5% quantile and take the average ratio of the scenario against the market, which yields 0.98. Although we do not calculate whether the two are statistically different, it supports our belief from Figure 2 that the quantile is slightly smaller under the mixture scenario. This indicates the possibility of following a price path in this lower quantile is slightly smaller under the mixture scenario. We suspect this smaller 2.5% quantile is caused by the lower volatility persistence under the mixture scenario, since it shortens the impact of volatility spikes and results in a lower volatility level on average.

5.1.2 Expected Shortfall

In this subsection, we obtain the daily Expected Shortfall time series under the mixture scenario for Portfolio 1. Below, the results under the mixture scenario are depicted, together with those under the simulated market:



Underestimations of the simulated market ES are visualised in orange. They are not assigned a numerical value. When both ES series are equal, we do not regard this as an underestimation.

Figure 3: Expected Shortfall for Portfolio 1 under the mixture scenario.

Most ES estimates range between −5.5% and −7.5%. We do observe 3 relatively big downwards spikes around $t = 50, 90$ and 195. At these instances, portfolio prices from Figure 2 drop at such a magnitude that their volatility hits the threshold in the TGARCH model, causing the ES estimation to drop precipitously. Besides, ES series tend to be stable over time, since portfolio prices from Figure 2 follow an upwards trend, resulting in stable returns.

Our previous belief that the 2.5% price quantile is slightly smaller under the mixture scenario is supported by Figure 3, since the simulated market ES almost always attains a more negative value than the mixture ES. Whether this difference is significant will follow from the ES comparison framework. If so, we could conclude that the mixture ES estimation in most instances is less conservative. We would thus anticipate less risk under the mixture scenario than their is in the simulated market. Practically, this would imply a financial institution does not hold enough capital to sufficiently cover excessive losses.

5.1.3 Coverage comparison

The following coverage results are obtained under the mixture scenario for Portfolio 1. Coverage in this case reflects the percentage of observations in the 2.5% quantile under the scenario per time instance. Under the simulated market, this quantile equals the 2.5% VaR threshold.

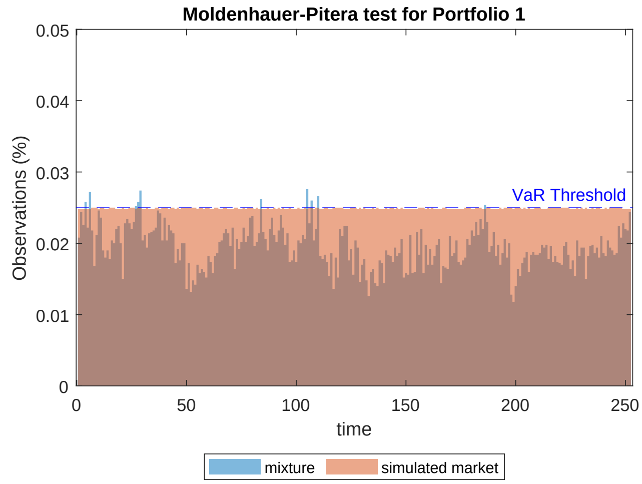


Figure 4: Coverage test for Portfolio 1 under the mixture scenario.

The percentage of observations under the mixture scenario in most instances is less than under the simulated market. This indicates that the distance between ES and VaR, conditional on the underlying return distribution, is almost always smaller under the mixture scenario than under the market. That is also why the instances where the coverage under the mixture is larger than under the market correspond with those where the market ES is less negative than the mixture ES in Figure 3.

To statistically compare these coverage test results, we consult the t-test. It rejects the null hypothesis 88 out of 252 instances, indicating that the mean of excessive losses under the scenario significantly differs from the mean under the market 35% of the time, while they are statistically similar at the other instances. Most of these differences occur around the large negative peaks from Figure 3. We refer to Section 5.1.5 for the size of these differences. Since the mixture scenario combines higher correlations and lower volatility persistence, these mean difference results are a combination of the mean difference results of the subscenarios. These subscenario results are further elaborated on in the text box below.

Opposite shifts in mean

The following coverage test results are obtained for the higher correlations and lower volatility persistence scenarios. They describe the percentage of observations that are in the 2.5% quantile under the scenario. Under the market, this quantile equals the VaR threshold.

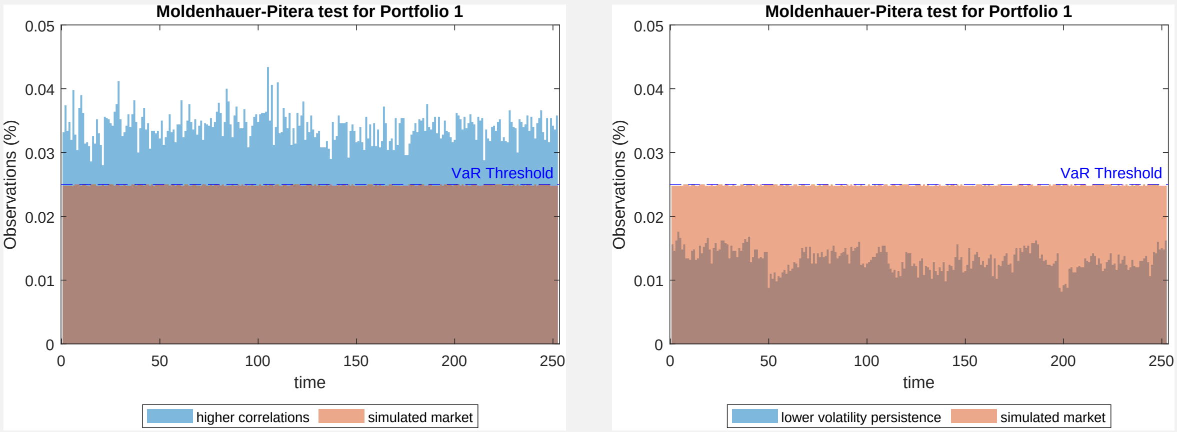


Figure 5: Coverage test for Portfolio 1 under higher correlations (left) and lower volatility persistence (right).

Under higher correlations, the percentage of observations that describe the 2.5% quantile are always higher than the VaR threshold. We therefore expect the ES estimation to be more negative at any instance under higher correlations than under the market, in case their underlying distribution is the same. Under lower volatility persistence, we see the opposite effect: less percentage of observations than under the market. Since both results are based on different underlying distributions, we cannot draw conclusions from this intuition yet.

The t-test rejects the null hypothesis 141 out of 252 instances under higher correlations, indicating the mean excessive losses under higher correlations and this mean under the market differ 56% of the time. Under lower volatility persistence, this is almost 100% of the time. We thus conclude that the mean excessive losses significantly differ from the market more often under lower volatility persistence than under higher correlations. Moreover, we conclude that the subscenarios shift the mean excessive losses in opposite directions and that thus their effect approximately cancels out in the mixture scenario. From Figure 4 and 5, we suspect that the impact of lower volatility persistence overtakes higher correlations, which we analyse further in the distribution comparison.

5.1.4 Distribution comparison

The distribution test results for Portfolio 1 are depicted below. Here, the distribution of excessive losses under the mixture scenario is compared with this distribution under the simulated market. The graph displays the density in the tail combining all 252 daily return distributions.

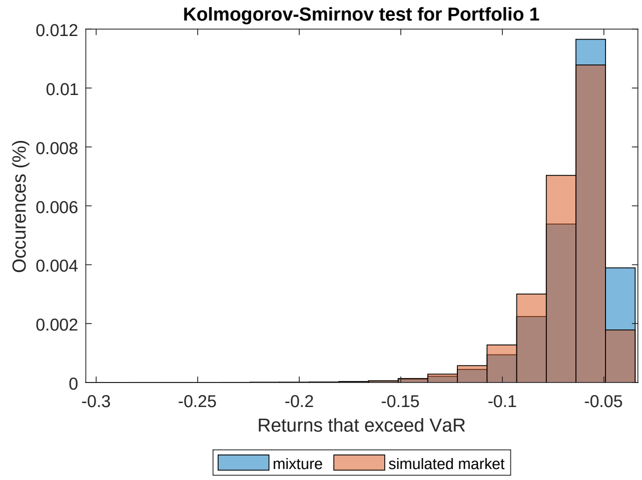


Figure 6: Distribution test for Portfolio 1 under the mixture scenario.

The excessive loss distribution under the mixture scenario tends to be shifted rightwards compared to the market. The shift implies that large losses are less likely to occur under the mixture scenario than under the market. The Kolmogorov-Smirnov test indicates that this shift is significant at most instances. It rejects the null hypothesis 210 out of 252 instances, which implies the distributions differ 83% of the time.

Our explanation from Figure 2 showed that lower volatility persistence results in a smaller 2.5% quantile for portfolio prices. We see from Figure 7 that this lower volatility persistence causes excessive losses to occur less frequently under the mixture scenario. From a more in-depth analysis below, we suspect that the impact of lowering volatility persistence overtakes increasing correlations with respect to the excessive loss distribution for Portfolio 1.

Lower volatility persistence tends to overtake higher correlations

The results for the distribution test under each subscenario are depicted below. Distribution graphs display the density in the tail combining all 252 daily returns distributions.

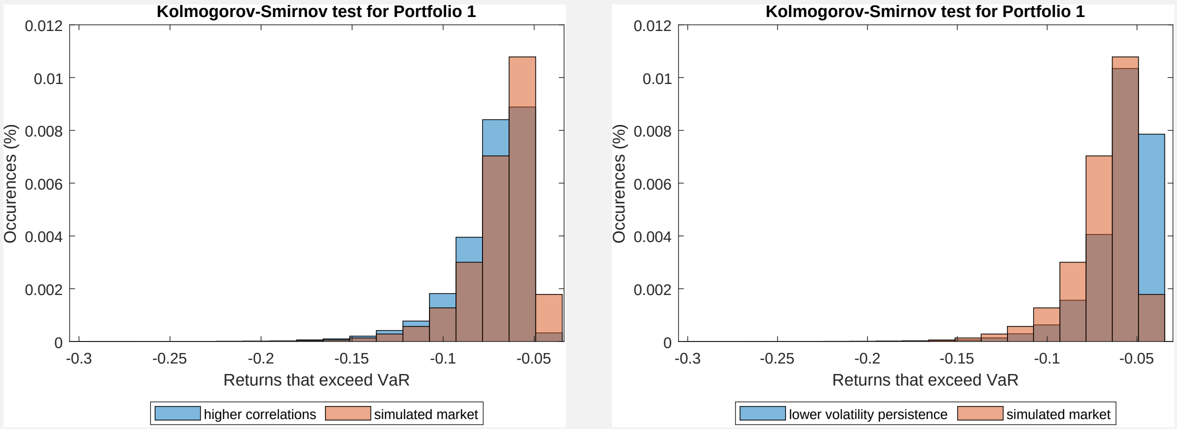


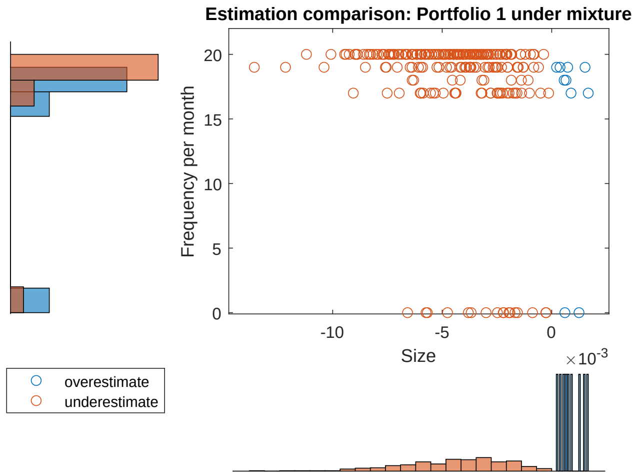
Figure 7: Distribution test for Portfolio 1 under higher correlations (left) and lower volatility persistence (right).

Under higher correlations, the excessive loss distribution tends to be shifted leftwards and the Kolmogorov-Smirnov test rejects the null hypothesis 232 out of 252 instances. In short, a significant leftwards shift occurs 92% of the time. Under the lower volatility persistence, we notice a rightwards shift in the return distribution. The Kolmogorov-Smirnov test rejects the null hypothesis at each instance under this scenario, indicating the shift under this subscenario is always significant.

We can draw two conclusions. First, the shift in excessive return distribution appears to be more significant under lower volatility persistence than under higher correlations. This is supported by a bigger shift in Figure 7 and the fact that 98% of the time the test statistic is larger under lower volatility persistence. Second, a significant shift occurs more often under lower volatility persistence compared to higher correlations. The number of null hypothesis rejections by the Kolmogorov-Smirnov test supports this conclusion. Moreover, these conclusions are reflected in the distribution shift under the mixture, which is rightwards. Furthermore, we observe that each misspecification causes an opposite shift: higher correlations cause excessive losses to occur more frequently, while lower volatility persistence results in these losses occurring less often.

5.1.5 Estimation comparison

Below, the estimation comparison for the mixture scenario for Portfolio 1 is depicted.



Underestimations occur when the market ES attains a more negative value than the scenario ES. They are assigned a negative size since an insufficient amount of capital is set aside in those instances. For overestimations, the reverse is true.

Figure 8: Estimation comparison for Portfolio 1 under the mixture scenario.

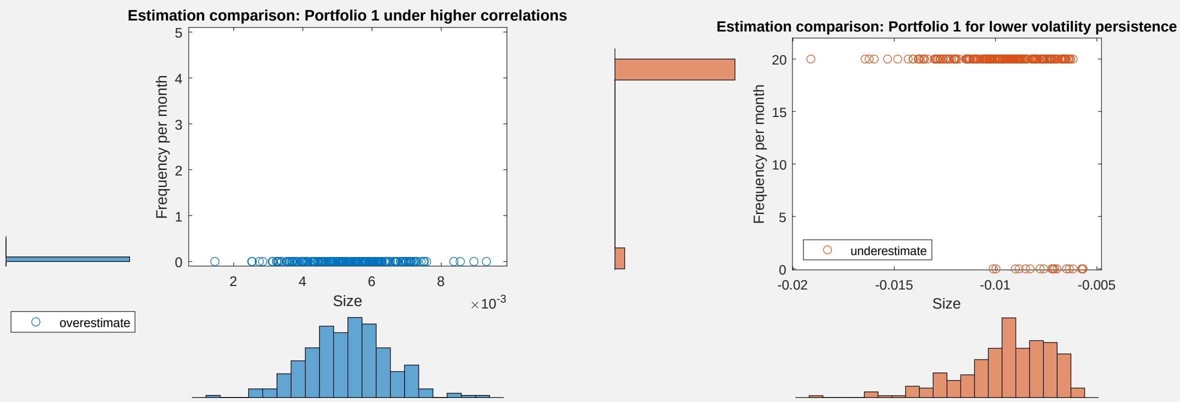
We note that underestimation, i.e. assuming less risk under the scenario than there is in the market, occurs often. This is consistent with Figure 3, which shows many market ES underestimations. Figure 8 shows these underestimations are often much bigger in absolute size than the overestimations. Besides, we note that underestimations often occur frequently, with often at least 15 overestimations occuring in one month, of which most times an overestimate occurs each day of the month. All of these underestimations are caused by lowering volatility persistence. We refer to the grey box below for an in-depth explanation.

To examine autocorrelations in underestimations, we perform the Ljung-Box test for each up to 10 lags. Since it shows similar results for prior lags, we only discuss the test for 10 lags here. It rejects the null hypothesis 160 out of 252 instances, which implies market ES underestimations are 63% of the time autocorrelated at 10 lags over a 20 day rolling window.

Conclusively, we state that autocorrelations are often significant and that the possibility of observing autocorrelations up to 10 lags in a period of one month is about 2/3. Besides, we conclude that these autocorrelations are caused solely by lower volatility persistence.

Underestimations are caused by lower volatility persistence

Results of the estimation comparison for Portfolio 1 for higher correlations and for lower volatility persistence are illustrated below. There are no simulated market ES underestimations for the higher correlations scenario. Underestimations do occur under lower volatility persistence.



Underestimations occur when the market ES is more negative than the scenario ES. They are assigned a negative size since an insufficient amount of capital is set aside in those instances. For overestimations, the reverse is true.

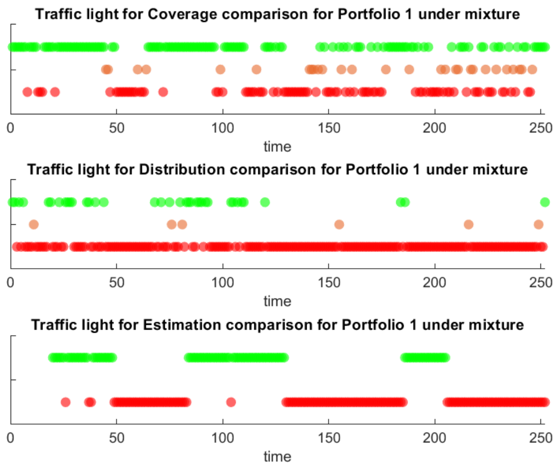
Figure 9: Estimation comparison for Portfolio 1 under higher correlations (left) and lower volatility persistence (right).

Under higher correlations, only overestimations occur and thus (autocorrelation in) underestimations cannot be considered. Under lower volatility persistence, however, only underestimations occur. These results are in line with its ES series always being more negative than the market ES under higher correlations and always less negative under lower volatility persistence (not depicted).

The Ljung-Box test rejects the null hypothesis 233 out of 252 instances under lower volatility persistence for 1, 5 and 10 lags. Therefore, the possibility of observing autocorrelation up to 10 lags in a one month period is 92%. In conclusion, we regard these underestimations to be highly correlated under the lower volatility persistence scenario.

5.1.6 Traffic light framework

In this subsection, the traffic light framework for Portfolio 1 is illustrated. This framework considers all three comparisons under the mixture scenario. We stress that this framework does not assess the performance of the tests performed, but the extend to which the mixture scenario results are similar to the simulated market results.



Coverage comparison		
	Green	Red
Number of instances	136	88
Percentage of total	54%	35%

Distribution comparison		
	Green	Red
Number of instances	36	210
Percentage of total	14%	83%

Estimation comparison		
	Green	Red
Number of instances	73	160
Percentage of total	29%	63%

Figure 10: Traffic light framework (left) and its descriptive statistics (right) for Portfolio 1 under the mixture scenario.

These traffic lights show that the ES comparison framework for Portfolio 1 is significantly violated at many instances. This implies the margin in your risk model is fairly small when misspecifying correlation and volatility persistence parameters. Besides, we saw from Figure 3 that the ES estimation under the mixture scenario is less conservative, i.e. we assume less risk than there is in the market. Moreover, we saw from both the distribution and estimation comparison that lower volatility persistence overtakes the impact of higher correlations for Portfolio 1.

We refer to our research question in stating that these misspecifications lead to a financial institution setting aside an insufficient amount of capital than is necessary in most instances. In other words, a financial institution would often be inadequately able to cope with excessive losses when using our risk model in case it is not well calibrated to the market. This is due to ES estimations under these misspecifications being less conservative. These inadequate results are mainly caused by lowering volatility persistence. Moreover, we note that in many instances these underestimations of market risk are dependent over at least two days. If an underestimation occurs, it is likely to be followed by another underestimation at least two days in a row. These results are undesirable since they imply ES estimations under misspecifications underestimate risk not coincidentally, but systematically. A financial institution would often reserve an insufficient number of capital for multiple days in a row in this case.

5.2 Portfolio 2: The Call Portfolio

In this subsection, the portfolio price paths, daily ES time series, the ES comparison framework and the traffic light framework for Portfolio 2 are depicted. The simulated market analyses, except for the portfolio prices, are not illustrated.

5.2.1 Portfolio price paths

Below, the portfolio price simulations for Portfolio 2 are depicted under the simulated market and the mixture scenario.

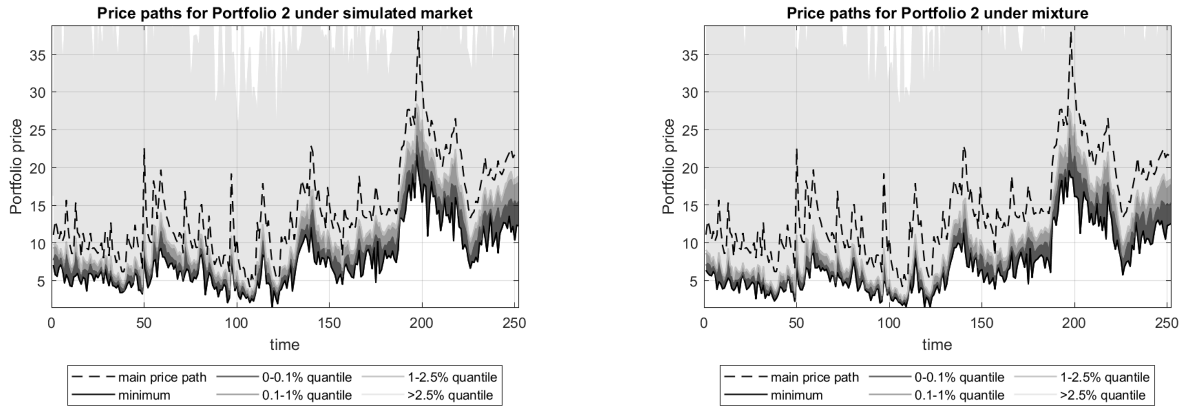


Figure 11: Portfolio price simulations for Portfolio 2 under the simulated market (left) and mixture scenario (right).

Portfolio prices are floored at 0 by construction, due to the payoff of calls in this portfolio (Section A.5). We see that price paths under the mixture scenario tend to be lower than under the market at instances before $t = 125$, together with their 2.5% quantile being smaller. We suspect this is caused by lower volatility persistence, since the lower persistence in volatility is, the lower the general level of volatility becomes, which negatively impacts the option value. Therefore, a price path under lower volatility persistence can obtain lower values than under the market. Moreover, we calculate the same average ratio to compare the lower quantile as for Portfolio 1, which yields 0.90. We observe that the probability of following a price path in the lower quantile is slightly smaller under the mixture than under the market. The consequences for the ES estimations are further elaborated on in Section 5.2.2. Figure 11 does not clearly show other distinctions between market and scenario.

5.2.2 Expected Shortfall

The daily ES time series under the mixture scenario for Portfolio 2 is depicted below.

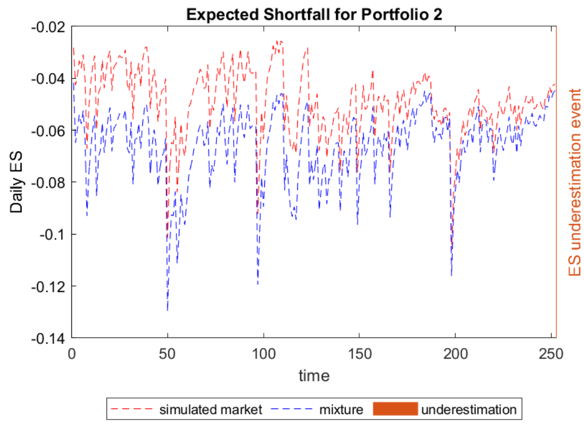


Figure 12: Expected Shortfall for Portfolio 2 under the mixture scenario.

After $t = 125$, we see more similarity between the scenario and the market. We suspect this is due to two reasons. First, this could be explained by an increased delta. We saw from Figure 2 that the underlying stocks tend to increase after $t = 125$ in our case, although they behave randomly depending on their seed. The higher the underlying stock price, the bigger the intrinsic value of the calls, so the more in the money they become. Accordingly, delta increases and the second part of this portfolio behaves more like a stock only portfolio, since it is more impacted by underlying stock prices. Besides, the closer to maturity, the narrower the range of possible option payoffs. This causes the delta of the options to increase[6] , making it more dependent on the underling stock price. For a stock only portfolio, we noted smaller differences between scenario and market visually (Figure 2), which also applies to Portfolio 2 after $t = 125$.

Second, the options payoff becomes more certain towards maturity, so the calls become less impacted by volatility. In other words, the vega of the options decreases towards maturity since volatility has less impact on the time value of the options. Therefore, prices are less impacted by lower volatility persistence and could become more similar under the scenario and the market after $t = 125$. These two conjectures are examined further in Section 5.2.3. For the theory behind moneyness, payoffs and the Greeks, we refer to Section A.5.

Figure 12 does not support our previous belief that the 2.5% price quantile is slightly smaller under the mixture scenario, since the market ES estimation never attains a more negative value than the mixture ES estimation. We conclude that the price quantile as a whole has shifted downwards so much that the smaller size of this quantile is not reflected in the ES series. Moreover, we noted that prices under the scenario can attain lower values than under the market, especially before $t = 125$. Figure 12 does relect this, because it shows the ES estimations are always more negative under the market than under the scenario.

Compared to Portfolio 1, we observe a bigger scale of ES estimations for this portfolio. Reason

[6] Assuming the option is in the money towards maturity and other inputs, except for the passage of time, remain the same.

being is the general price level for Portfolio 2 is smaller, so a change in price has a bigger impact on ES estimations for this portfolio. We still see the drops in ES estimation around $t = 50, 90$ and 195. They are less present, however, since this portfolio is not 1 on 1 impacted by the stock price paths, but depends on the delta of the calls. We note that ES estimations under both the scenario and market tend to be less conservative after $t = 125$. This is due to the fact that the portfolio exhibits characteristics that are similar to a stock portfolio, proportional to delta, given that the calls are in the money following the increased stock prices.

5.2.3 Coverage comparison

The following results are obtained for Portfolio 2 under the mixture scenario. Coverage in this case describes the percentage of observations that are in the 2.5% quantile under the scenario. Under the market, this quantile equals the VaR threshold.

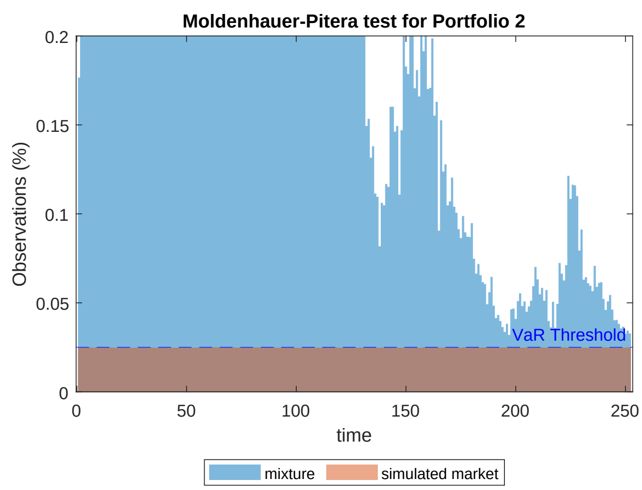


Figure 13: Coverage test for Portfolio 2 under the mixture scenario.

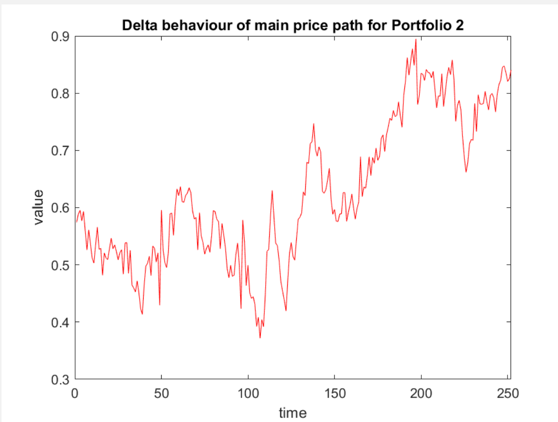
The percentage of observations in the 2.5% quantile under the mixture is always bigger than under the market, in case their underlying distribution is similar. If so, this would imply the difference between ES and VaR is bigger under the mixture scenario than under the market. This is consistent with the mixture scenario ES series always being more negative than the market ES in Figure 12.

The t-test rejects the null hypothesis 249 times out of 252 (99%). We conclude that the difference in mean excessive losses is nearly almost significant, indicating that misspecifications under the scenario have a significant impact on the mean excessive losses.

We observe significantly higher coverage before $t = 125$ for Portfolio 2 compared to Portfolio 1. As detailed below, this is mainly caused by vega behaviour of the calls. Vega attains larger values before $t = 125$, thus the portfolio value is more dependent on volatility at earlier instances. Under the mixture scenario, we lower the general level of volatility by lowering volatility persistence. This subscenario negatively impacts the options payoff, especially when vega is high (before $t = 125$). Therefore, coverage increases and lowering volatility persistence has bigger consequences for an options only portfolio than for a stocks only portfolio.

Delta upwards trend, vega downwards.

We suspect the difference in ES estimations towards maturity are caused by the behaviour of the Greeks of the calls. Therefore, we illustrate the delta and vega behaviour of the options below. The figure describes today's delta and vega behaviour based on the main price path, thus we do not make a distinction between scenarios.



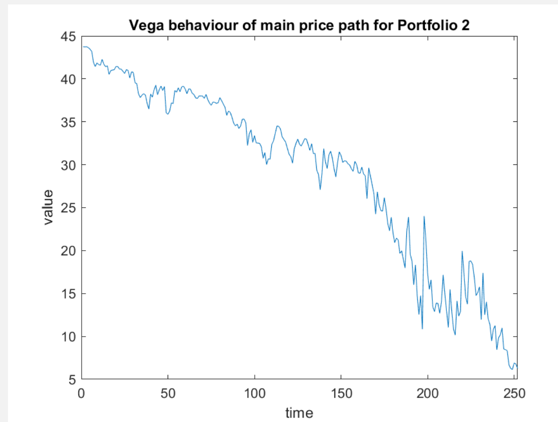


Figure 14: Delta (left) and Vega (right) behaviour of the main price path for Portfolio 2.

We note that the figures above are in line with our explanation from Section 5.2.2. There, we explained that delta tends to increase towards maturity, since we saw from Portfolio 1 that underlying stock prices tend to increase. This is supported by the fact that peaks in Figure 14 correspond with those in Figure 11, implying that delta behaviour is dependent on underlying stock price behaviour. Vega values tend to decrease towards maturity. We refer to our explanation from Section 5.2.2 that the closer towards maturity, the narrower the range of possible payoffs is, which results in the portfolio value being less sensitive to volatility changes. Our main take-away from Figure 14 is that our intuition from Section 5.2.2 is supported by actual delta and vega behaviour.

5.2.4 Distribution comparison

The following results are obtained for Portfolio 2 under the mixture scenario compared to under the simulated market. The graph displays the density in the tail combining all 252 daily return distributions.

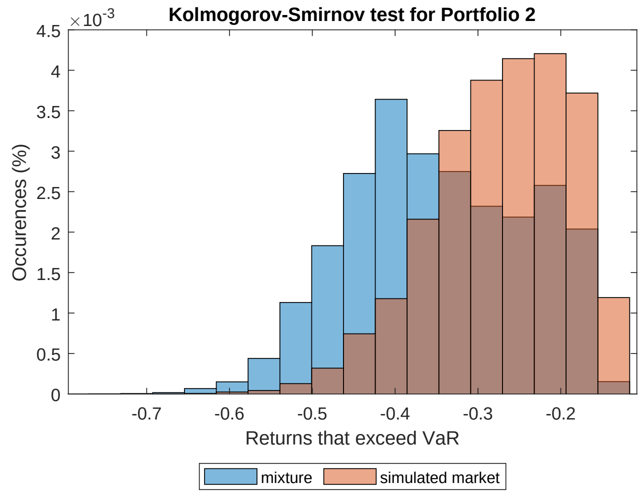


Figure 15: Distribution test for Portfolio 2 under the mixture scenario.

The excessive loss distribution under the mixture scenario tends to be shifted leftwards compared to the market. The shift implies that large losses are more likely to occur under the mixture scenario than under the market. This shift appears to be significant, which is supported by the Kolmogorov-Smirnov test rejecting the null hypothesis at each instance.

Our understanding of the corresponding ES series supports the fact that the return distribution is shifted leftwards under the mixture scenario. In Figure 12, we saw that ES estimates under the mixture are always more negative than those under the market, indicating excessively large losses are more likely to occur under the mixture scenario. As detailed below, we conclude that there is a significant difference between excessive loss distribution before and after $t = 125$ under the mixture. Excessively large losses are more likely to occur before $t = 125$ than afterwards under the mixture (detailed below). This is mainly caused by the impact of lower volatility persistence before $t = 125$.

Break also notable in distribution

Since we noted a break in ES series at $t = 125$ in Figure 12, we are interested in excessive loss distributions before and after this instance. Below, they are depicted for the mixture scenario.

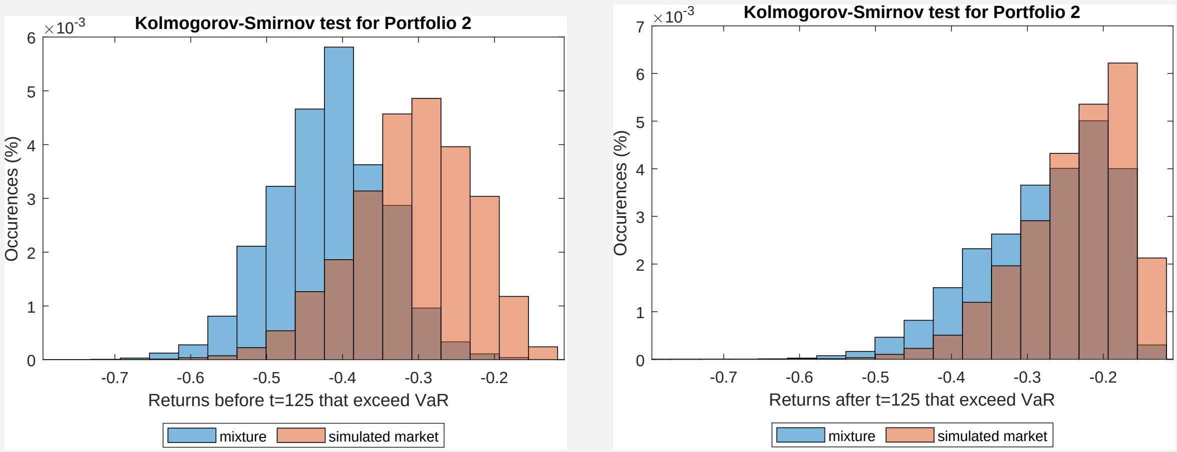
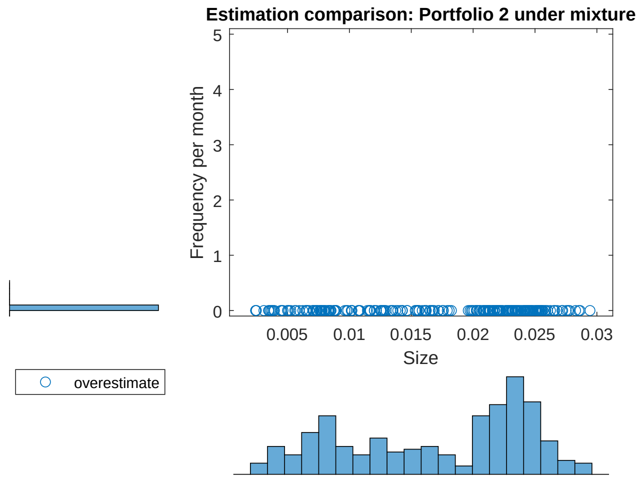


Figure 16: Distribution test for Portfolio 2 under mixture before $t = 125$ (left) and after (right).

From Figure 16, we note that the break in ES estimations (Section 5.2.2) is also present in the distribution of losses above VaR. This is supported by the Kolmogorov-Smirnov test. It shows both shifts before and after $t = 125$ are always significant, but corresponding test statistic before $t = 125$ almost always attains a larger value than afterwards (98%). Section A.9 shows this break is mainly caused by lowering volatility persistence. We noted in Section 5.2.2 and 5.2.3 that this is due to delta and vega behaviour: before $t = 125$ portfolio prices are more impacted by volatility (higher vega), while afterwards prices behave more like a stock only portfolio, proportional to delta. For higher correlations, we do not note a significant difference before and after $t = 125$ (not visualised). In contrast to Portfolio 1, Section A.9 explains that the distribution shift of each subscenario is in the same direction. This is in line with the coverage exceeding the VaR threshold for both subscenarios.

5.2.5 Estimation comparison

Below, the results for the analysis of the simulated market ES underestimations is shown for Portfolio 2 for the mixture scenario. No market ES underestimations occur for both the mixture scenario and the subscenarios. We do provide an interpretation on the analysis of size and frequency of ES over- and underestimations.



Underestimations occur when the market ES is more negative than the scenario ES. They are assigned a negative size since an insufficient amount of capital is set aside in those instances. For overestimations, the reverse is true.

Figure 17: Estimation comparison for Portfolio 2 under the mixture scenario.

We note that only overestimations occur, so for each instance over any one month period, an underestimation does not occur. This implies we assume more risk under our risk model than there is in the market, which is consistent with the absence of market ES underestimations in Figure 12. The histogram in Figure 17 shows that large overestimations do occur frequently. This is undesirable since it indicates the ES estimation under the scenario often is well off from the market. Overestimations are caused by both higher correlations and lower volatility persistence, but under the latter their size is considerably bigger (not depicted). This is in line with the break at $t = 125$ having a bigger impact for this subscenario, as explained by the coverage and distribution comparison.

5.2.6 Traffic light framework

In this subsection, the traffic light framework for Portfolio 2 is visualised. This framework considers the coverage and distribution comparison under the mixture scenario. Since no market ES underestimations occur, we do not asses a traffic light for the estimation comparison.

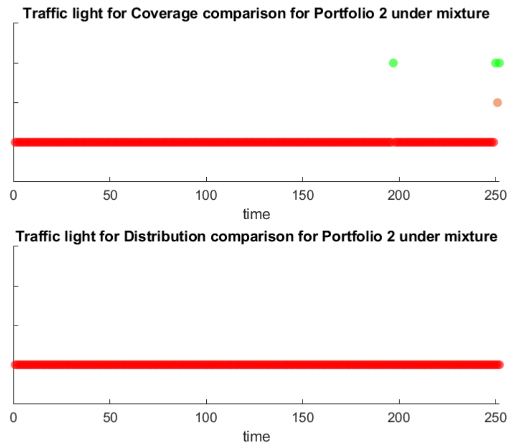


Figure 18: Traffic light framework for Portfolio 2 under the mixture scenario.

At almost each instance a red light is shown. At three instances for the coverage comparison, Figure 18 does not show red lights. They all occur at later instances. This is in line with our previous conclusion that the ES estimations under the scenario are more similar to the market for instances after $t = 125$. We explained this is due to the calls in this portfolio, whose impact vanish towards maturity. Traffic lights for the coverage comparison indicate, however, that also after $t = 125$ the ES framework is often violated. We noted from the distribution comparison that the mixture scenario always causes a significant shift in excessive loss distribution. This shift is more present before $t = 125$, but also afterwards it is significant.

Figure 18 indicates the ES framework is almost always violated. This implies the model specifications under the mixture scenario cause our ES estimations to be far off from the market. We saw that it causes the ES estimations to be more conservative, i.e. the model indicates more risk than their is in the market. This is not undesirable per se, since being too conservative implies you always cover for excessive losses. Nevertheless, we saw that, especially before $t = 125$, the ES estimations are far too conservative, which causes too much capital to be set aside to cover for large losses.

To relate to our research question, we saw that the portfolio construction causes the ES estimations for Portfolio 2 under the mixture scenario to be more conservative than the market, in contrast to Portfolio 1. This is mainly caused by delta and vega behaviour of the calls. We conclude that the misspecification of lowering volatility persistence causes the ES estimations to be even more off and more conservative for this portfolio than for a stock only portfolio.

Portfolio 3 confirms option behaviour from Portfolio 2

Results for Portfolio 3 are similar to those of Portfolio 2. We refer to Section A.10 for an extensive elaboration on Portfolio 3. Our main conclusion applies again: lower volatility persistence impacts ES estimations substantially, causing the mixture ES estimations to be too conservative. Again, this is due to options being heavily impacted by volatility changes. Once more, this impact is most notable before $t = 125$, or when options still substantially impact the portfolio price. This is explained by their vega behaviour (Section 13), which is similar for puts and for calls, and delta, which is opposite for puts compared to calls. Puts lose value when stock prices increase. This occurs especially after $t = 125$, so the options become less in the money (delta decreases) towards maturity. Both vega and delta behaviours result in the options having less impact after $t = 125$ and the price paths thus being more similar under the scenario compared to under the market.

We relate to our research question in stating that a portfolio of both options and stocks (Portfolio 3) does not lead to significantly different ES estimations compared to an option only portfolio (Portfolio 2). Misspecifying volatility has a bigger impact on options than on stocks, especially before $t = 125$. This applies to both calls and puts. Misspecifying correlations seems to have less of an impact on the ES estimations, especially for a portfolio that includes options but we also noted this for a stock only portfolio (Portfolio 1).

5.3 Portfolio 4: The Basket Spread Portfolio

In this subsection, the portfolio price paths, daily ES time series and the ES comparison and traffic light frameworks for Portfolio 4 are depicted. The simulated market analyses, except for the portfolio prices, are not illustrated.

5.3.1 Portfolio price paths

Below, the portfolio price simulations under both the simulated market and the mixture scenario are depicted for Portfolio 4.

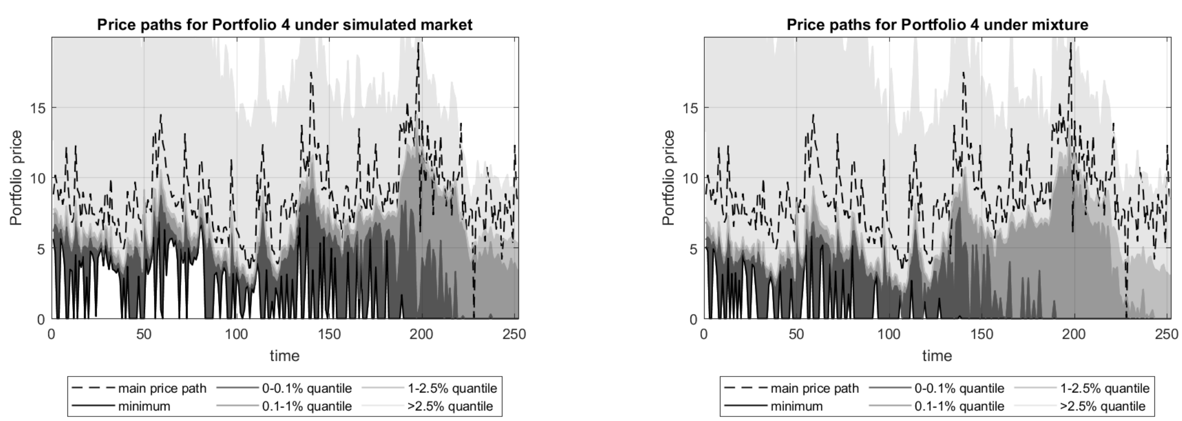


Figure 19: Portfolio price simulations for Portfolio 4 under the simulated market (left) and mixture scenario (right).

The main price path tends to be relatively stable around 7.5. We thus observe that the main price path is not much impacted by changes in underlying stock prices. We therefore suspect that the delta of both components in this portfolio (the portfolio of calls and the basket option) tend to cancel out, i.e. cause a net delta of zero. We suspect the same reasoning applies to vega, since it impacts both components implicitly. Correlations, however, may play a bigger role in the trend of the price paths above. Reason being is the intrinsic value of the basket option depends on correlation (to calculate its volatility) more explicitly, while correlation impacts the portfolio of calls implicitly. Therefore, the spread between one instrument that is implicitly and another that is more explicitly impacted by correlation, is out of sync. This may explain our observation that the lower quantile converges to zero towards maturity, since the basket option is more impacted by correlation.

We note that under the mixture scenario, the 2.5% quantile tends to be bigger towards maturity and we note that minimum price paths converge to zero quicker. We suspect this is largely due to higher correlations under the mixture scenario. For Portfolio 1, we noted that when correlations increase, the possibility of an upwards stock price path grows more than downwards. This means volatility in the portfolio of stocks increases, causing the basket option to gain intrinsic value. This causes the portfolio price as a whole to decrease, due to the second part in Equation (22) becoming bigger.

5.3.2 Expected Shortfall

The daily ES time series for Portfolio 4 under the mixture scenario is depicted below.

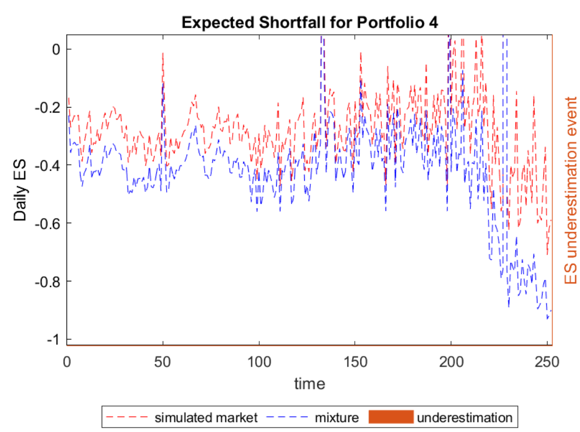


Figure 20: Expected Shortfall for Portfolio 4 under the mixture scenario.

The range of ES is similar to our previous option only portfolio (Portfolio 2). Similar to Portfolio 2 and 3, no market ES underestimations occur. This implies the scenario ES estimation is too conservative, in case the ES series are statistically different. We investigate this further in our ES comparison framework.

The ES series tend to converge to −1 towards maturity of the options. This is in line with the portfolio prices converging towards 0. When portfolio prices converge downwards, but are not equal to zero yet, we expect the ES estimation for the next day to be very negative. We refer to Section 5.3.1 for a detailed explanation. Moreover, when portfolio prices are equal to zero, portfolio prices from Figure 19 hit the lower threshold, which means that at the next instance, the portfolio price can never go down further. This results in very large upwards peaks in ES series at these instances.

We notice two breaks in ES series where the ES estimations differ more between scenario and market. First, we notice a similar break at $t = 125$ as for Portfolio 2 and 3. This in line with our belief from Portfolio 2 (Section 5.2.1) that before $t = 125$, a portfolio including options is much more impacted by volatility than afterwards (vega is higher beforehand). Since we decrease volatility persistence under the mixture scenario, this has a more profound effect on the difference in ES series before $t = 125$ than afterwards. Second, we notice the ES series tend to differ more after $t = 220$, which is in line with our previous belief that portfolio prices tend to differ towards maturity. We refer to the last paragraph of Section 5.3.1 for a detailed explanation.

5.3.3 Coverage comparison

The following results are obtained for Portfolio 4 under the mixture scenario. Coverage in this case reflects the percentage of observations that are in the 2.5% quantile under the scenario. Under the market, this quantile equals the VaR threshold. For those instances where the ES estimation attains a positive value, no coverage test is performed.

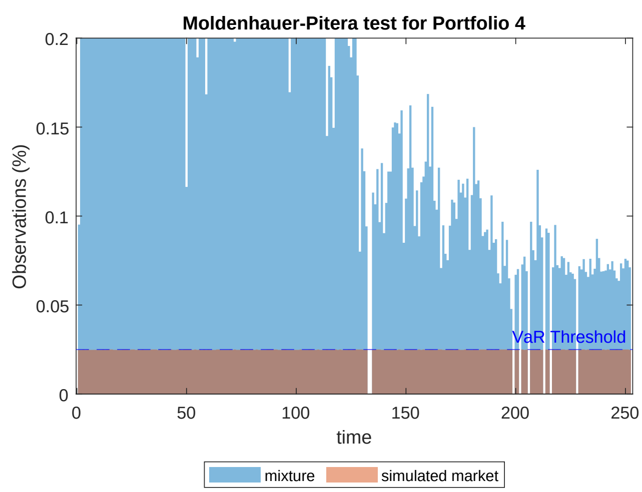


Figure 21: Coverage test for Portfolio 4 under the mixture scenario.

From Figure 21, we note similar coverage test results as for previous portfolios including options. This time, however, the coverage towards the simulation horizon tends to be bigger compared to other portfolios. This is due to the second break in ES series at $t = 220$. As detailed in the text box below, we note that this is caused by higher correlations having more impact towards maturity. For a detailed explanation on the reason why, we refer to Section 5.3.1. We also note from the text box that the coverage under lower volatility persistence drops below the VaR threshold towards maturity. It may suggest we underestimate market risk towards maturity for this subscenario, but this still depends on the underlying distribution. This intuition is further examined in Section 5.3.5.

The t-test rejects the null hypothesis at each instance, implying that the mean excessive returns under the mixture differ significantly from those under the market. We can therefore conclude that our mixture ES estimation from Figure 20 is significantly more conservative. Practical consequences are similar to those explained for Portfolio 2 (Section 5.2.6).

Higher correlations overtake lower volatility persistence towards maturity

The following results are obtained for Portfolio 4 under each subscenario. They describe the percentage of observations that are in the 2.5% quantile under the scenario. Under the market, this quantile equals the VaR threshold. For those instances where the ES estimation attains a positive value, no coverage test is performed.

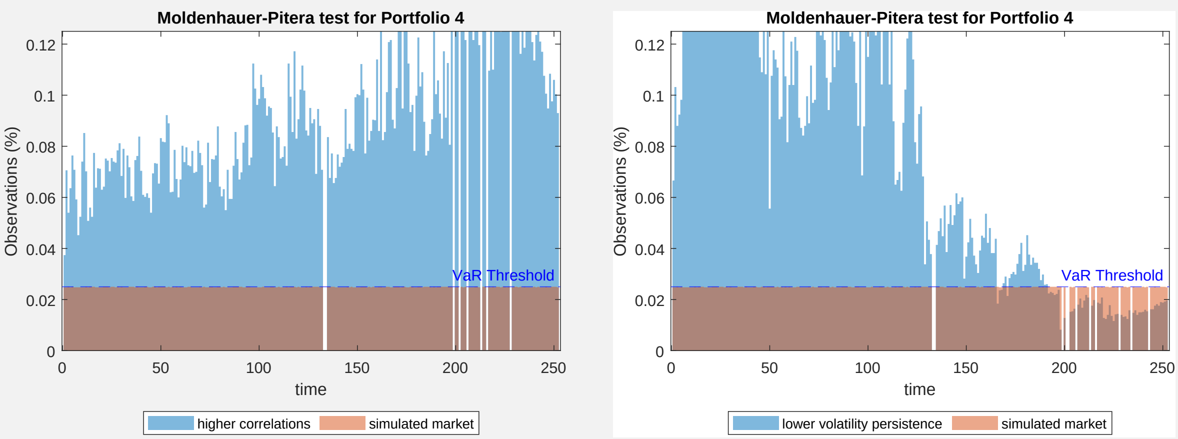


Figure 22: Coverage test for Portfolio 4 for higher correlations (left) and lower volatility persistence (right).

Under higher correlations, we note an increasing coverage towards maturity. As explained in Section 5.3.1, higher correlations have more impact towards maturity due to the basket option being a correlation driven instrument. Under lower volatility persistence, we note similar coverage before $t = 125$ as for Portfolio 2 and 3. We refer to Section 5.2.2 for an in-depth explanation. In Figure 22, we note lower volatility persistence causes less coverage towards maturity. This is due to the portfolio of calls in Portfolio 4 and this result is similar to Portfolio 2. Since the coverage under the mixture exceeds the market towards maturity, we conclude that the impact of higher correlations on mean excessive losses overtakes lower volatility persistence, in case both underlying distributions are the same.

5.3.4 Distribution comparison

The following distribution comparison results are obtained for Portfolio 4. They compare the excessive loss distribution under the mixture with the simulated market. The graph displays the density in the tail combining all 252 daily return distributions.

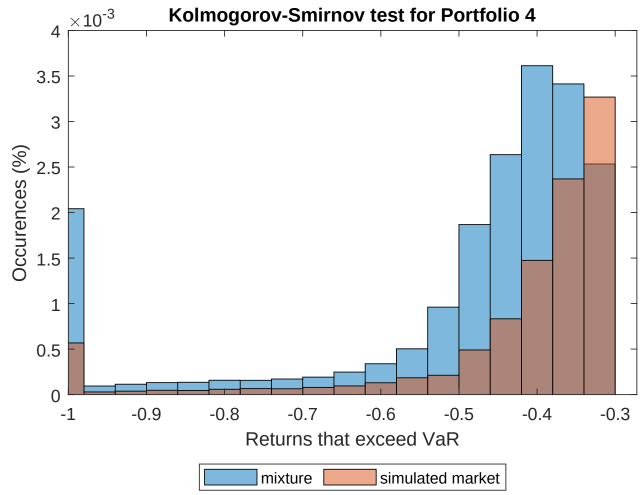
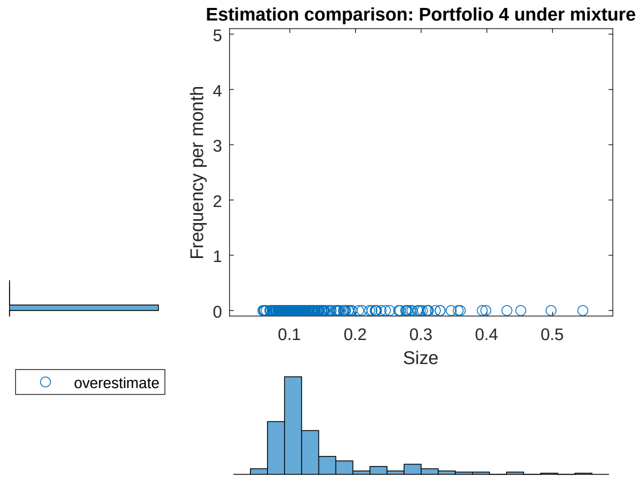


Figure 23: Distribution test for Portfolio 4 for the mixture scenario.

The same direction of shift, Kolmogorov-Smirnov test results and distribution comparison conclusions as for Portfolio 2 and 3 apply here. This time, however, returns are more often almost equal to -1, since the portfolio prices converge to 0. From Section A.11, we note this is mainly caused by higher correlations. This supports our belief that the impact of higher correlations overtakes lower volatility persistence towards maturity. Because the basket option increases in value relative to the portfolio of calls due to higher correlations towards maturity, it causes the portfolio prices from Section 5.3.1 to converge to zero.

5.3.5 Estimation comparison

Below, the results for the estimation comparison under the mixture scenario are depicted:



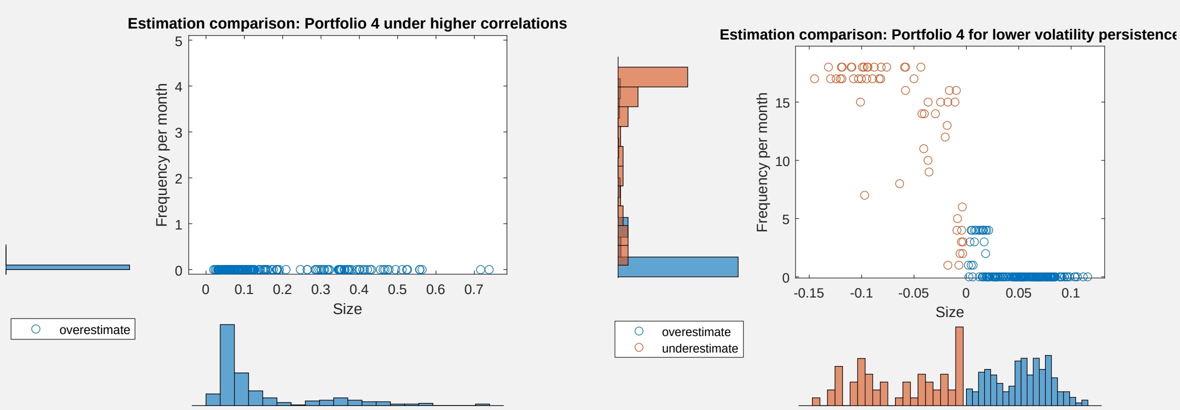
Underestimations occur when the market ES is more negative than the scenario ES. They are assigned a negative size since an insufficient amount of capital is set aside in those instances. For overestimations, the reverse is true.

Figure 24: Estimation comparison for Portfolio 4 under the mixture scenario.

The estimation comparison for the mixture scenario yields approximately similar results for Portfolio 4 as for Portfolio 2. This shows that Portfolio 4 behaves much more like an option only portfolio, because only overestimations occur. In contrast to Portfolio 2, many underestimations occur under lower volatility persistence, as detailed below. The main scenario is similar to Portfolio 2, because higher correlations overtakes the impact of the underestimations caused by lower volatility persistence towards maturity. We previously noted that this is due to the basket option being heavily impacted by increasing correlations towards maturity.

Impact of higher correlations overtakes lower volatility persistence.

The estimation comparison results under the higher correlations and lower volatility persistence scenarios are depicted below.



Underestimations occur when the market ES is more negative than the scenario ES. They are assigned a negative size since an insufficient amount of capital is set aside in those instances.

Figure 25: Estimation comparison for Portfolio 4 under higher correlations (left) and lower volatility persistence (right).

We note from Figure 25 that market ES underestimations do occur under lower volatility persistence. They occur often many times within the same month and seem to be highly autocorrelated. These autocorrelations are often significant, which is supported by the Ljung-Box test. It rejects the null hypothesis 55 out of 252 (22%) instances for 10 lags. From a practical perspective, this implies that under lower volatility persistence one often underestimates risk in the market not coincidentally, but systematically. In case of significant autocorrelations, they are highly significant at many lags (Section A.12). From the coverage comparison (Section 5.3.3), we noted that these underestimations tend to occur towards maturity, which is in line our belief of the basket option being heavily impacted towards maturity. We thus conclude that the underestimates are clustered at instances close to maturity. This also explains the fact that autocorrelations do not occur all the time, as indicated by the percentage of null hypothesis rejections by the Ljung-Box test, but when they do they are highly significant at many lags (Section A.12).

5.3.6 Traffic light framework

In this subsection, we discuss the traffic light framework for Portfolio 4. This framework considers the coverage and distribution comparison for the mixture scenario. It shows red lights at each instance and therefore is not illustrated. We conclude that our ES estimation under the scenario is well off for Portfolio 4. We saw that at each instance the ES estimation is too conservative. We refer to Portfolio 2 for the practical consequences of this conclusion (Section 5.2.6).

To relate to our research question, we conclude that the basket option causes the ES estimation to be heavily impacted by correlation. We noted for the coverage comparison that especially towards maturity this impact causes the mean excessive losses to differ significantly from those of the market. This also caused the break in ES series at $t = 220$ in Figure 20 and the fact that overestimations under higher correlations overtake underestimations under lower volatility persistence. The portfolio construction thus causes the ES estimation to be even more off from the market than for the other portfolios including options. This difference is caused by the basket option being heavily impacted by correlation and is particularly notable towards maturity.

6 Conclusion and Discussion

The core of this thesis concerns the impact different model specifications and portfolio characteristics have on ES estimations. We analyse this by constructing four portfolios using Monte Carlo simulations under different correlations and/or volatilities. Furthermore, we compare corresponding ES estimations to the market by analysing coverage, distribution and underestimations of excessive losses.

We answer our research question in stating that misspecifying volatility can have substantial impact on corresponding ES estimation for portfolios including stocks and options and thus a financial institution's risk management. The impact of misspecifying correlations is less notable for these portfolios. When correlation driven instruments, such as a basket option, are included, the margin in misspecifying correlations is also quickly violated. We conclude that a non-derivative portfolio has more margin when misspecifying parameters in the risk model. A portfolio that includes instruments driven by one (or multiple) of the factors in the risk model is heavily impacted by misspecification of these factors. Oftentimes in practice, both volatilities and correlations are misspecified, leading to an insufficient amount of capital being set aside for stocks only portfolios, and too much for other investigated portfolios[7].

For a stocks only portfolio, we conclude that the ES estimation is too optimistic under both higher correlations and lower volatility persistence. Under these misspecifications, a financial institution would oftentimes set aside an insufficient amount of capital than is necessary in most instances. Besides, underestimations of market risk for our stocks only portfolio tend to be autocorrelated, indicating prior-mentioned misspecifications cause our model to underestimate risk systematically. A financial institution would thus oftentimes be inadequately able to cope with excessive losses for multiple days in a row, when using our risk model in case it is not well calibrated to the market. To relate to our research question, these results are mainly caused by lowering volatility persistence, implying the margin in misspecifying volatilities is smaller than for correlations.

For an options only portfolio, we observed too conservative ES estimations under higher correlations and lower volatility persistence. These misspecifications in our risk model result in too conservative ES estimations, implying a financial institution would set aside too much capital. Being able to cover losses slightly more than necessary is desirable, however, for our calls only portfolio we noted the margin is almost always significantly violated, thus far too much capital is set aside and one loses out on investment opportunities for this reserved capital. These results differ from the stocks only portfolio due to lower volatility persistence having a bigger impact, mostly before $t = 125$. This is caused by delta and vega behaviour of the options, causing ES estimations to be even more off and more conservative than for a stocks only portfolio. These results apply both to calls and puts. We thus saw that including options to protect a stocks only portfolio adds complexity to the ES estimations, since misspecifying volatility has a bigger impact on options. Misspecifying correlations seems to have even less impact on ES estimations for options than for stocks.

For a more complex portfolio, or a spread between a portfolio of calls and a basket option, we conclude that misspecifying correlations has a profound impact on ES estimations. This impact causes overestimations of market risk to overtake underestimations caused by lower volatility persistence, when both misspecifications occur. This is especially notable towards maturity, since the impact of misspecifying volatility drops when vega decreases. To relate to our research question, the portfolio construction thus causes the ES estimation to be even more off compared to other

[7] assuming misspecifications imply lowering volatilities and increasing correlations.

portfolios, since the basket option is heavily impacted by misspecifying correlations.

As an extension of this research, we suggest investigating more portfolios. One could include other derivatives, such as vanilla options, exotic options (other than the basket option), spreads (other than Portfolio 4) and/or combinations of prior-mentioned. Besides, we assumed no rebalancing of portfolios occurred, which does occur in practice often. We suggest extending this research into rebalanced portfolios, for instance mean-variance efficient portfolios with transaction costs. Both other instruments and rebalancing would introduce more portfolio properties and thus would expand the scope of our research.

Our research does have limitations. First, our simulated market results are not calibrated to actual market data. Simulated market results do perfectly replicate the actual market, but this actual market is not calibrated to actual stock price data. Model calibration would add substantially more complexity than is necessary for the purpose of this thesis. Besides, calibration results depend on calibration technique used and the way one assesses its performance, thus more assumptions would be introduced. Now that we theoretically know which misspecification has most impact on a certain portfolio, we can extend our research more practically by calibrating our model to actual market data. Second, our research is based on a number of assumptions. For instance, in our case underlying stock prices increased towards maturity and our results are conditional on this observation. Similarly, we focused on lower volatilities (through persistence) and higher correlations as our misspecifications of interest. We did look into other scenarios, such as adjusting the TGARCH threshold or adjusting the distribution of the noise in the copula, and found results consistent with our reasoning. This is thus not so much a limitation, as it is a point of awareness that alternative scenarios would yield (different) results that are still subject to the same reasoning as conveyed in this thesis.

A Appendix

A.1 Risk Axioms

For a function $\rho : L \longrightarrow \mathbb{R}$ with portfolio losses $X, X_1, X_2 \in L$, Artzner et al. (1999) defines ρ to be a coherent risk measure if the following four properties are satisfied:

1) Monotonicity: If $X_1 \leq X_2$, then $\rho(X_1) \leq \rho(X_2)$.
2) Sub-additivity: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.
3) Positive homogeneity: for constant $\lambda \geq 0$, we have $\rho(\lambda X) = \lambda\rho(X)$.
4) Translation inverse: For every constant function a, then $\rho(X + a) = \rho(X) + a$.

A.2 Rotated Gumbel copula

In the following figure, a scatter plot of a rotated Gumbel copula for two uniformly distributed variables y and x is shown. The sample size equals $1,000$ and a dependence variable θ of six is chosen.

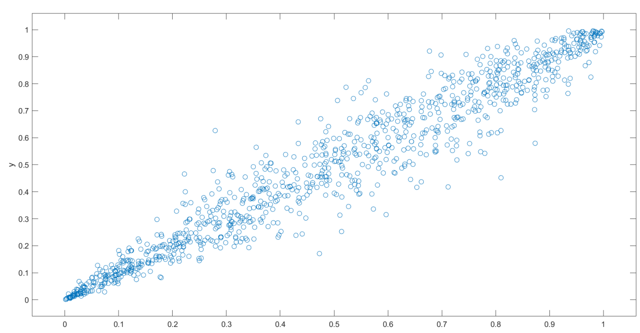


Figure 26: 2D scatter plot of a 180 degrees rotated Gumbel copula

The rotated Gumbel copula models stronger lower tail dependence. This is reflected in the fact that observations around (0,0) tend to be more clustered than around (1,1) for uniformly distributed variables. The 2D figure above is a simplified representation of the clustering of observations modeled by a rotated Gumbel copula, since also interaction between observations in different buckets occurs, as explained in Section 3.1.1.

A.3 dgp parameter values

In the following table, the parameter values for modeling correlations and volatilities are specified.

Bucket parameters		Lower bucket	Middle bucket	Higher bucket
TGARCH	v	-	8	-
	ω	-	0.1	-
	α	0.01	0.03	0.05
	β	0.003	0.0065	0.01
	δ	0.3	0.5 $\rightarrow$ 0.4	0.85 $\rightarrow$ 0.4
	κ	-	-1.5	-
Copula	θ	1.5 $\rightarrow$ 3	3 $\rightarrow$ 5	5 $\rightarrow$ 7

Table 3: dgp parameter values

Table 3 specifies the parameter values that are set in the dgp of our risk model. Variables for the TGARCH model correspond with those in Equations (5) - (7) and copula parameter θ with Equation (3). The values after the arrows describe the parameters applied in each scenario. For GARCH parameters, this concerns the lower volatility persistence scenario, while for copula parameters this concerns the higher correlations scenario. The mixture scenario applies both parameter changes.

Parameters v, ω and κ are not assigned to multiple buckets. Since we differentiate the stocks based on their fluctuations around the general volatility level ω, a single ω value for all stocks is assigned. We analyse the impact a quantile of returns will have on volatility for different thresholds to determine the value of κ. We set threshold κ equal to -1.5, and increase corresponding parameter β. This way, fewer stocks are impacted by the threshold but in case they are, they are impacted heavier. The combined effect is approximately equivalent to the standard $\kappa = 0$ threshold with lower coefficient β.

A.4 TGARCH model: stability

In this subsection, a stable TGARCH(1,1) model is depicted. Specifically, a stable TGARCH(1,1) model is satisfied in case the following inequality holds:

$$\alpha + P[r_{t-1} < \kappa]\beta + \delta < 1, \tag{23}$$

in which the parameters α, β and δ are depicted in Equation (6). Inequality (23) is derived as follows:

$$\begin{aligned}
E[\sigma_t^2] &= E[\omega + \alpha r_{t-1}^2 + \beta \mathbb{I}_{t-1} r_{t-1}^2 + \delta \sigma_{t-1}^2] \\
&= \omega + \alpha E[r_{t-1}^2] + \beta E[\mathbb{I}_{t-1}] E[r_{t-1}^2] + \delta E[\sigma_{t-1}^2] \\
\bar{\sigma}^2 &= \omega + \alpha \bar{\sigma}^2 + \beta P[r_{t-1} < \kappa] \bar{\sigma}^2 + \delta \bar{\sigma}^2 \\
&= \omega + (\alpha + P[r_{t-1} < \kappa]\beta + \delta)\bar{\sigma}^2 \\
\bar{\sigma}^2 &= \frac{\omega}{(1 - \alpha - P[r_{t-1} < \kappa]\beta - \delta)},
\end{aligned}$$

which implies the TGARCH(1,1) model is stable in case $\alpha + P[r_{t-1} < \kappa]\beta + \delta < 1$. The parameters and variables above are specified in Equation (6), except for unconditional variance $\bar{\sigma}^2$. In the derivation above, the following property for a stable volatility process is used: $\bar{\sigma}^2 = E[\sigma_t^2] = E[\sigma_{t-1}^2] = E[r_{t-1}^2]$.

A.5 Options: moneyness, payoff and the Greeks

Call and put options can have either intrinsic value and/or time value. This is represented by their moneyness as follows:

- At the money: the option has no intrinsic value, solely time value;
- In the money: the option has positive intrinsic and time value.
- Out of the money: the option has no intrinsic value.

A call (put) option is in the money if the price of its underlying asset is bigger (smaller) than the corresponding strike. For an out of the money option, the reverse is true. For an at the money option, its underlying price equals the strike. This behaviour is in line with the payoffs of call and put options, denoted as follows:

$$\text{CALL}_t = \max(S_t - K, 0), \tag{24}$$

$$\text{PUT}_t = \max(K - S_t, 0), \tag{25}$$

where CALL_t and PUT_t describe the prices for the call and put options respectively at time t. They are modeled using Black-Scholes from Section 3.1.4. Variable S_t defines stock prices at time t and K strike prices.

Behaviour of these options is captured by the Greeks. They define the value of the option with respect to certain underlying instruments, such as stock price, volatility and time to maturity. The following two measures are discussed in this thesis:

- Delta: the dependency of the option price on changes in underlying stock price. For call options, delta attains a value between 0 (certain to expire out of the money) and 1 (certain to expire in the money). For puts, this value ranges between -1 (certain to expire in the money) and 0 (certain to expire out of the money);
- Vega: the amount the option price will change for a one-point implied volatility change. This concerns changes in expected future volatility. Vega thus only influences the time value of option prices and drops towards maturity.

A.6 Portfolio price transformation

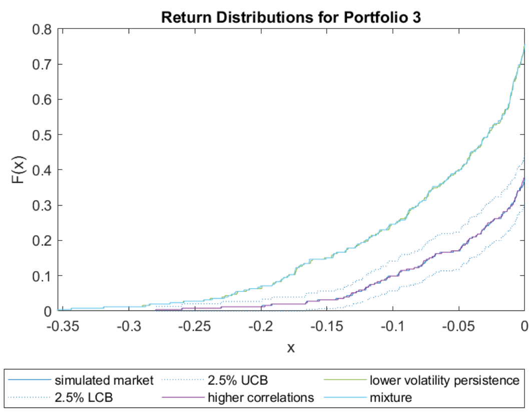
The following formula specifies the transformation of portfolio prices into returns:

$$R_t = \frac{P_t - p_{t-1}}{p_{t-1}}, \quad \text{for } t = 2, ..., T \tag{26}$$

in which P_t describes the 5,000 simulated portfolio prices forecast from a single portfolio price p_{t-1} at $t-1$ for time t. We refer to Section 4.2 for their construction. Using the equation above, we obtain a daily portfolio return time series R_t of 251 trading days.

A.7 Return distributions for Portfolio 3 and 4

Below, the return distributions under our misspecification scenarios for Portfolio 3 and 4 are visualised. Returns concern negative returns exceeding the VaR threshold. The dashed lines visualise the 2.5% lower and upper confidence bounds of the simulated market.



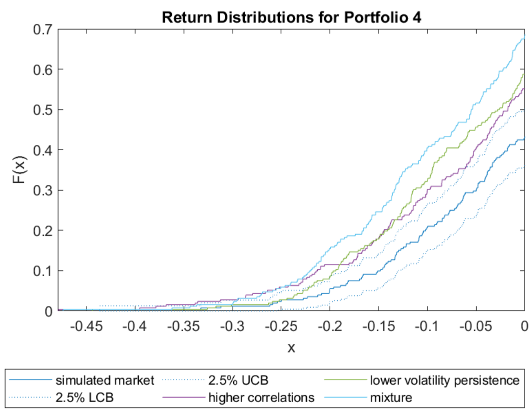


Figure 27: Return distributions under each scenario for Portfolio 3 (left) and Portfolio 4 (right).

We observe the similar patterns for Portfolio 3 as for Portfolio 2, explained in Section 4.1. For Portfolio 4, we observe the higher correlations scenario causes the excessive loss distribution to shift compared to the simulated market. We did not notice this for other portfolios, which suggests changing correlations has a more profound impact on Portfolio 4.

A.8 Comparing simulated with actual market

In this subsection, the results for the simulated market for Portfolio 1 are shown. All figures in this subsection support the fact that the simulated market exactly replicates the actual market. Therefore, we do not consider comparing the two for more portfolios.

Below, the results for Portfolio 1 for the simulated market are shown:

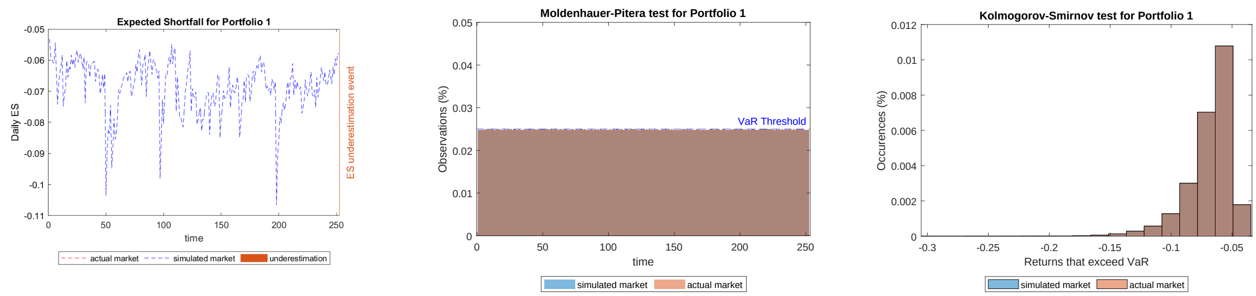


Figure 28: Expected Shortfall (left), Coverage test (middle) and Distribution test (right) for Portfolio 1 for simulated market compared to actual market.

From Figure 28, we notice the ES framework remains exactly the same since the simulated market perfectly replicates the actual market. To summarize, the daily ES for both are exactly equal, which we define as ES underestimations do not occur at any point in time. The mean in losses exceeding VaR for the simulated market is exactly the same as for the actual market, just like their distributions. This implies the model specification does not impact the ES framework at all. The simulated market can thus correctly be viewed as the benchmark scenario.

A.9 Distribution comparison for lower volatility persistence for Portfolio 2 with break at $t = 125$.

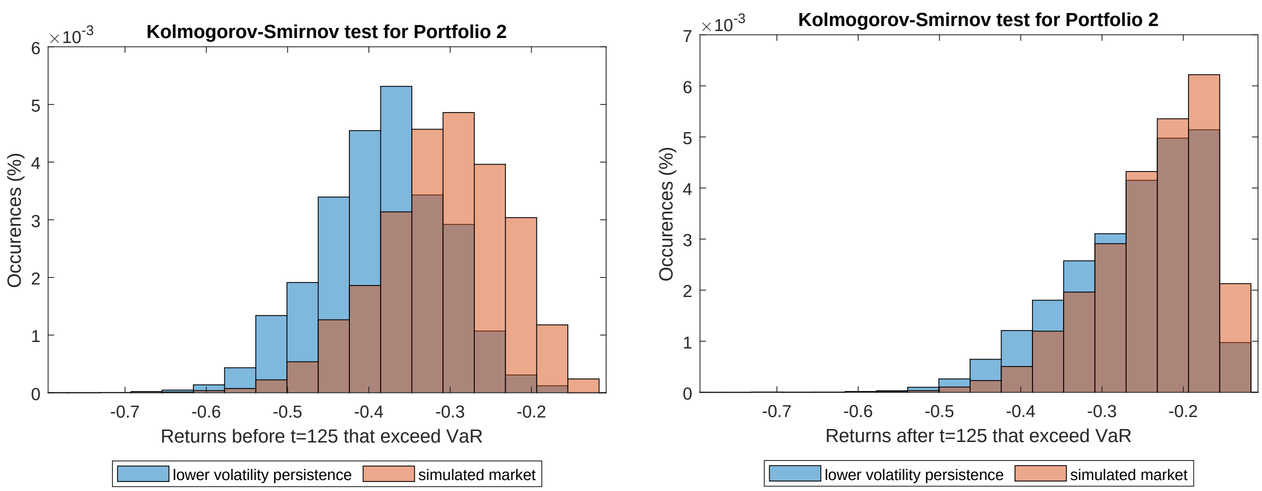


Figure 29: Distribution test for Portfolio 2 under lower volatility persistence before $t = 125$ (left) and after (right).

From Figure 29, we note that the break at $t = 125$ in the distribution of losses above VaR under the mixture is mainly caused by lower volatility peristence. Under higher correlations, the distribution shift is in the same direction but a break is less visible (not depicted). The Kolmogorov-Smirnov test supports this. It indicates the distribution shift is always more significant under lower volatility persistence before $t = 125$, since the test statistic under lower volatility persistence always exceeds the one under higher correlations for the instances up to $t = 125$. At later instances, however, this shift is bigger at 73 out of the remaining 127 instances (57%). Consequently, after t = 125 there does not appear to be one subscenario clearly overtaking another in terms of distribution shift.

A.10 Results Portfolio 3: The Risk-protected Portfolio

In this subsection, the results for Portfolio 3 are illustrated and elaborated on. This time, the simulated market is shown neither here, nor in the appendix, except for the portfolio price paths.

A.10.1 Portfolio price paths

The portfolio price simulation paths for Portfolio 3 under the simulated market and mixture scenario are depicted below:

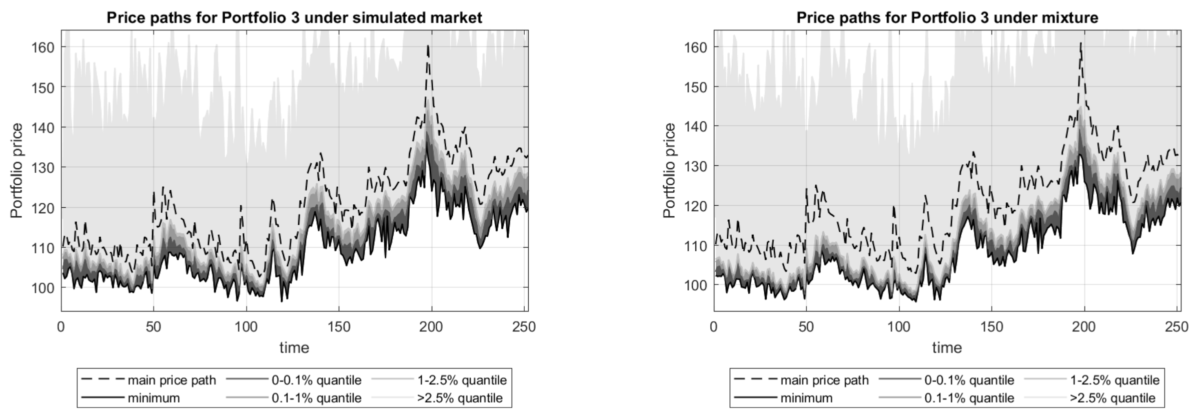


Figure 30: Portfolio price simulations for Portfolio 3 under the simulated market (left) and mixture scenario (right).

Price paths are floored at the lower end, specifically at 90. This corresponds with the strike of the put options in this portfolio. If stock prices drop below this price, the puts gain intrinsic value causing the portfolio price to increase. We note that the stock price level for this portfolio tends to be higher than for other portfolios. Since the puts protect this portfolio against downside risk, one would have to pay a premium to invest in this portfolio, which is reflected in a higher level of portfolio prices.

Based on our analysis of Portfolio 2, we expect the 2.5% quantile to be smaller under the scenario before t = 125. Thereafter, we expect the prices to follow approximately the same paths under the scenario compared to under the market. This is due to similar vega behaviour for the puts in this portfolio as for the calls in Portfolio 2. Specifically, the options lose time value since the closer to maturity they get, the narrower the range of possible payoffs becomes, i.e. the less impact volatility has on the options. For delta, we suspect opposite behaviour for this portfolio. The put options lose value since stock prices increase after t = 125 so the options become less in the money (delta decreases). Both behaviours result in the options having less impact after t = 125 and the price paths thus being more similar under the scenario compared to under the market. Other observations, such as a smaller 2.5% quantile and the minimum path being lower under the mixture than under the market, are explained in Section 5.2.1.

A.10.2 Expected Shortfall

In this subsection, the daily ES time series under the mixture scenario is obtained for Portfolio 3, together with the simulated market ES time series. These results are visualised below.

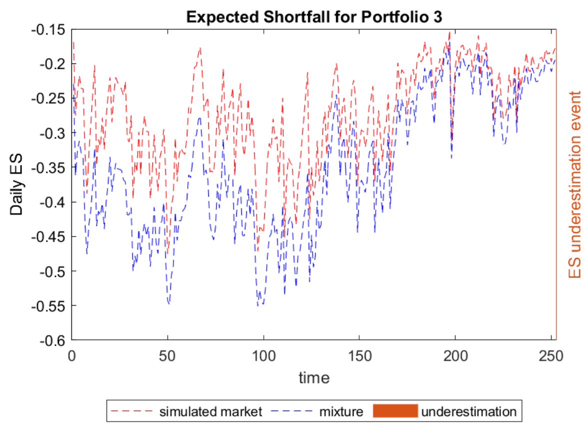


Figure 31: Expected Shortfall for Portfolio 3 under the mixture scenario.

The general level of the market ES series tends to be less negative compared to other portfolios before $t = 125$. Before this instance, market ES ranges mostly between -3% to -5%, while later it ranges between -4% and -7%. When the put options affect the price (before $t = 125$), they protect this portfolio against downside risk. Consequently, excessively large losses are floored, causing the level of ES series to be less negative. This especially the case for the market ES series, since volatility is more present in the market, causing puts to have more impact.

Similar to Portfolio 2, the ES series under the scenario is more negative than under the market. Whether this difference is significant will follow from the ES comparison framework. Again, the ES series tend to differ more before $t = 125$ than at later instances. We refer to Section 5.2.2 for a detailed explanation (of the consequences) of these observations.

A.10.3 Coverage comparison

The following results are obtained under the mixture scenario for Portfolio 3. Coverage in this case describes the percentage of observations that are in the 2.5% quantile under the scenario. Under the market, this quantile equals the VaR threshold.

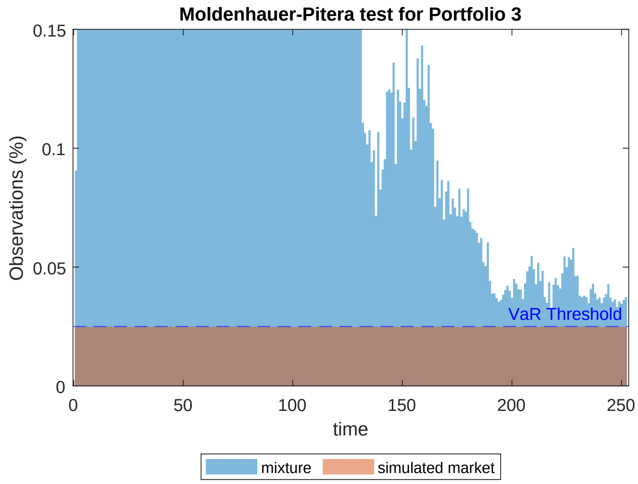


Figure 32: Coverage test for Portfolio 3 under the mixture scenario.

The coverage comparison for the subscenarios for this portfolio is similar to Portfolio 2 (both not visualised). Both subscenarios cause the mean to shift in the same direction and the coverage before $t = 125$ is substantially bigger than at later instances, especially under lower volatility persistence. Calls and puts behave similarly with respect to volatility. The delta of the options behaves in opposite directions, specifically delta increases the closer to maturity for calls, and decreases for puts. From this coverage test, we conclude that opposite delta behaviour is overtaken by similar vega behaviour, as the coverage comparison results are similar for portfolios with calls and puts. We thus conclude that ES estimations are much more impacted by options due to sensitivity to volatility, but the prices behave still like a stocks only portfolio (Portfolio 1). We refer to Section 5.2.3 for the in-depth explanation. Also other observations, such as higher coverage under mixture than under market and conclusion from the t-test, are similar to those for Portfolio 2.

A.10.4 Distribution comparison

The distribution comparison for the mixture scenario for Portfolio 3 is depicted below. The graph displays the density in the tail combining all 252 daily return distributions.

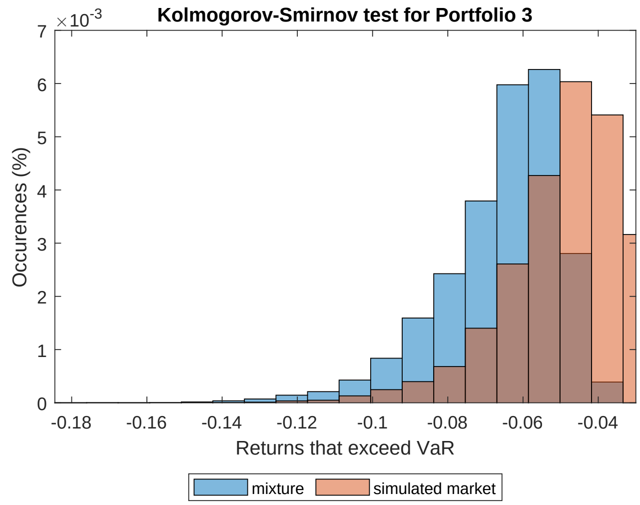


Figure 33: Distribution test for Portfolio 3 under the mixture scenario.

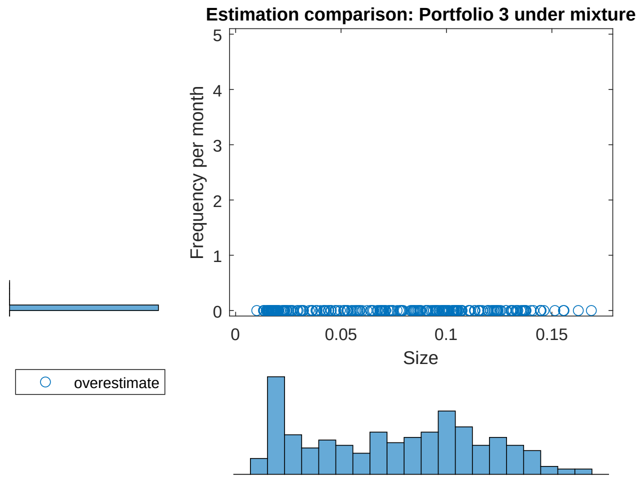
The distribution comparison shows similar results[8] as for Portfolio 2 (Section 5.2.4). Nevertheless, distributions are not exactly the same, since the excessive return distributions for this portfolio are also impacted by including stocks, which causes the returns after $t = 125$ to behave much more like Portfolio 1. Besides, the distributions of Portfolio 2 and Portfolio 3 are not exactly the same since Portfolio 2 only includes 10 of the 100 underlying stocks of Portfolio 3.

Main takeaway from this analysis is that the excessive loss distribution is much more impacted by the options due to lower volatility persistence, while the prices still behave like a stock only portfolio (same for Portfolio 2). Nevertheless, the scale of the losses above VaR is smaller compared to other portfolios. This time, these tail returns range from about -0.04 to -0.14. This is in line with the price paths from Figure 30 being floored by put options, indicating that extremely large losses do not occur.

A.10.5 Estimation comparison

Below, the estimation comparison for the mixture scenario for Portfolio 3 is depicted. We only show both analysis of size and frequency of underestimations, without the autocorrelation in underestimations. Reason being is no underestimations occur for Portfolio 3 under the mixture scenario.

[8] Again, a shift leftwards is notable and the Kolmogorov-Smirnov test indicates this shift is significant at each instance. Also the distribution under the subscenarios is similar to Portfolio 2 (both not illustrated).



Underestimations occur when the market ES is more negative than the scenario ES. They are assigned a negative size since an insufficient amount of capital is set aside in those instances. For overestimations, the reverse is true.

Figure 34: Estimation comparison for Portfolio 3 under the mixture scenario.

The estimation comparison for this portfolio shows similar results as for an only options portfolio (Portfolio 2). Main difference concerns the fact that overestimations tend to be bigger for this portfolio. Again, this is due to the fact that this portfolio is protected against downside risk, which causes overestimations to attain a bigger size, as visualised in the histogram in Figure 34. This bigger size implies that the scenario ES estimation is often well off from and less negative than the market, which implies too much capital is set aside. In short, the estimation comparison supports our belief that ES estimations for Portfolio 3 are much more impacted by options due to sensitivity to volatility.

The estimation comparison for the subscenarios is similar to Portfolio 2 (both not visualised). Only difference consists of the fact that there do occur two underestimations under the lower volatility persistence, which implies the market ES is more negative than the scenario ES at two instances under lower volatility persistence. Since the overwhelming majority of observations are overestimates for each subscenario, these few underestimates do not impact the mixture scenario much and thus are not discussed further.

A.10.6 Traffic light framework

The traffic light framework under the mixture scenario for Portfolio 3 is illustrated below. This framework does not include the independence of underestimations since the scenario ES is always more conservative than the simulated market ES.

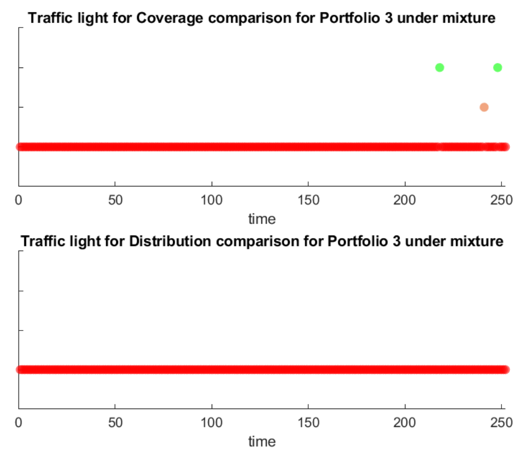


Figure 35: Traffic light framework for Portfolio 3 under the mixture scenario.

The traffic light framework for Portfolio 3 is similar to that of Portfolio 2. The same conclusions apply for this portfolio: a break in ES estimations occurs at $t = 125$ which is largely due to lower volatility persistence. Specifically, the vega of the options cause the portfolio to be heavily impacted by volatility before $t = 125$ and less so afterwards. Also for this portfolio, the ES estimations under the mixture scenario are often well off compared to the market. This leads to more conservative ES estimations and too much capital being set aside in practice. The grey box in Section 5.2.6 summarises the main conclusions and their relation to the research question for this portfolio.

A.11 Distribution comparison for subscenarios for Portfolio 4

The distribution test results for Portfolio 4 the higher correlations and lower volatility persistence scenarios are depicted below. Distribution graphs display the density in the tail combining all 252 daily return distributions.

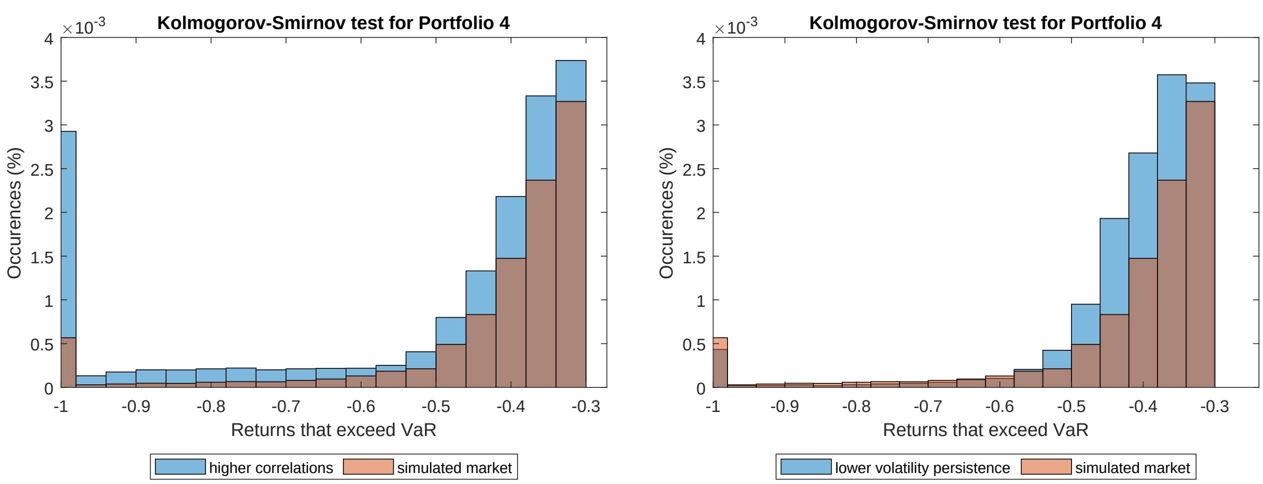


Figure 36: Distribution test for Portfolio 4 under higher correlations (left) and lower volatility persistence (right).

The figure above visually supports our belief from the coverage test that the impact of higher correlations overtakes lower volatility persistence towards maturity. This time, it is also notable in the excessive loss distribution, since observed -1 returns occur more often under higher correlations, indicating this subscenario causes the portfolio price paths from Section 5.3.1 to converge to zero. The Kolmogorov-Smirnov test for both subscenarios oftentimes rejects the null hypothesis (always under higher correlations, 93% under lower volatility persistence). Both subscenarios cause the distribution to (almost) always significantly shift and higher correlations overtakes lower volatility persistence towards maturity.

A.12 Autocorrelation in ES underestimations for Portfolio 4

Below, we illustrate the Ljung-Box test results for autocorrelations in simulated market ES underestimations for Portfolio 4 under lower volatility persistence. The test results are computed over all 252 observations.

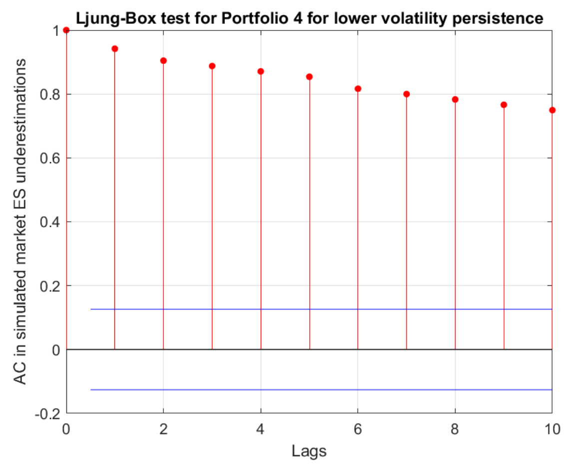


Figure 37: Independence in simulated market ES underestimations for Portfolio 4 under lower volatility persistence.

We note that at each lag, the ES underestimations show significant autocorrelations. This implies that when the market is more negative than the scenario ES estimation, it is likely to be so multiple days in a row, up to at least 10 days. The ES estimation under this scenario is far off from being close to the market when we examine the market ES underestimations.

References

Acerbi, C. and Szekely, B. (2014). Backtesting expected shortfall. *Risk*, 27(11):76–81.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Bayer, S. and Dimitriadis, T. (2019). Regression based expected shortfall backtesting. *arXiv preprint arXiv:1801.04112*.

BCBS (2013). Fundamental review of the trading book: A revised market risk framework.

BCBS (2016). Minimum capital requirements for market risk. Bank for International Settlements Basel.

Birnbaum, Z. W. (1952). Numerical tabulation of the distribution of kolmogorov's statistic for finite sample size. *Journal of the American Statistical Association*, 47(259):425–441.

Brooks, C. and Burke, S. P. (2003). Information criteria for garch model selection. *The European Journal of Finance*, 9(6):557–580.

Burns, P. (2002). Robustness of the ljung-box test and its rank equivalent. *Available at SSRN 443560*.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., and Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5):365–376.

Carver, L. (2013). Mooted var substitute cannot be back-tested, says top quant. *Risk, March*, 8.

Chen, S. X. (2008). Nonparametric estimation of expected shortfall. *Journal of financial econometrics*, 6(1):87–107.

Christoffersen, P. and Pelletier, D. (2004). Backtesting value-at-risk: A duration-based approach. *Journal of Financial Econometrics*, 2(1):84–108.

Cont, R. (2007). Volatility clustering in financial markets: empirical facts and agent-based models. In *Long memory in economics*, pages 289–309. Springer.

Danielsson, J. (2018). An evaluation of the basel iii market risk proposals. `https://www.modelsandrisk.org/basel/1.html`.

Davies, T. (2019). Eba: Europe's banks face €135bn capital shortfall after basel iv. `https://www.globalcapital.com/article/b1g38k1qpg8281/eba-europe39s-banks-face-135bn-capital-shortfall-after-basel-iv`.

Del Brio, E. B., Mora-Valencia, A., and Perote, J. (2017). Risk quantification for commodity etfs: Backtesting value-at-risk and expected shortfall. *International Review of Financial Analysis*.

Du, Z. and Escanciano, J. C. (2017). Backtesting expected shortfall: Accounting for tail risk. *Management Science*, 63(4):940–958.

EBA (2019). Consultation paper on draft rts on back-testing and pla attribution requirements (eba/cp/2019/11). https://eba.europa.eu/calendar%3Fp_p_id%3D8%26_8_struts_action%3D%252Fcalendar%252Fview_event%26_8_eventId%3D2844490.

Escanciano, J. C. and Olmo, J. (2010). Robust Backtesting Tests for Value-at-risk Models. *Journal of Financial Econometrics*, 9(1):132–161.

Fagiolo, G., Guerini, M., Lamperti, F., Moneta, A., and Roventini, A. (2019). Validation of agent-based models in economics and finance. In *Computer Simulation Validation*, pages 763–787. Springer.

Fissler, T., Ziegel, J. F., and Gneiting, T. (2015). Expected shortfall is jointly elicitable with value at risk-implications for backtesting. *arXiv preprint arXiv:1507.00244*.

Gaumert, U. and Kemmer, M. (2015). Regulatory developments in risk management: Restoring confidence in internal models. In *Innovations in Quantitative Risk Management*, pages 19–37. Springer, Cham.

Gelling, C. (2019). Expected shortfall backtesting. Master's thesis.

Graham, A. and Pal, J. (2014). Backtesting value-at-risk tail losses on a dynamic portfolio. *The Journal of Risk Model Validation*, 8(2):59.

Holton, G. A. (2020). Value-at-risk. https://www.value-at-risk.net/motivation-backtesting/.

Hopkin, P. (2018). *Fundamentals of Risk Management: Understanding, Evaluating and Implementing Effective Risk Management*, volume 5. Kogan Page Publishers.

Hull, J. et al. (2009). *Options, futures and other derivatives/John C. Hull.* Upper Saddle River, NJ: Prentice Hall,.

Katsiampa, P. (2017). Volatility estimation for bitcoin: A comparison of garch models. *Economics Letters*, 158:3–6.

Kolev, N., Ferreira, L., and Aguilera, R. (2015). Dependence analysis via copulas. In *Book of abstracts*.

Lindholm, L. (2014). *Calibration and Hedging in Finance.* PhD thesis, KTH Royal Institute of Technology.

Löser, R., Wied, D., and Ziggel, D. (2018). New backtests for unconditional coverage of expected shortfall. *Journal of Risk*, 21(4):1–21.

Manner, H. (2010). *Modeling asymmetric and time-varying dependence*. Datawyse/Universitaire Pers Maastricht.

Mathworks (2020). Simulating dependent random variables using copulas. `https://www.mathworks.com/help/stats/examples/simulating-dependent-random-variables-using-copulas.html`.

McAleer, M., Jiménez-Martín, J.-Á., and Pérez-Amaral, T. (2013). Has the basel accord improved risk management during the global financial crisis? *The North American Journal of Economics and Finance*, 26:250–265.

Moldenhauer, F. and Pitera, M. (2019). Backtesting expected shortfall: a simple recipe? *Journal of Risk*, 22(1).

Nelson, S. C. and Katzenstein, P. J. (2014). Uncertainty, risk, and the financial crisis of 2008. *International Organization*, pages 361–392.

Park, M. H. and Kim, J. H. (2016). Estimating extreme tail risk measures with generalized pareto distribution. *Computational Statistics & Data Analysis*, 98:91–104.

Patton, A. J., Ziegel, J. F., and Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and value-at-risk). *Journal of econometrics*, 211(2):388–413.

Righi, M. and Ceretta, P. S. (2013). Individual and flexible expected shortfall backtesting. *Journal of Risk Model Validation*, 7(3):3–20.

Scaillet, O. (2004). Nonparametric estimation and sensitivity analysis of expected shortfall. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 14(1):115–129.

Schneider, S., Schröck, G., Koch, S., and Schneider, R. (2017). Basel "iv": What's next for banks. *Implications of intermediate results.*

Taylor, J. W. (2008). Estimating value at risk and expected shortfall using expectiles. *Journal of Financial Econometrics*, 6(2):231–252.

Wali, B., Greene, D. L., Khattak, A. J., and Liu, J. (2018). Analyzing within garage fuel economy gaps to support vehicle purchasing decisions–a copula-based modeling & forecasting approach. *Transportation Research Part D: Transport and Environment*, 63:186–208.

Windrum, P., Fagiolo, G., and Moneta, A. (2007). Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation*, 10(2):8.

Wu, J. (2010). Threshold garch model: Theory and application. *The University of Western Ontario.*

Yamai, Y., Yoshiba, T., et al. (2002). Comparative analyses of expected shortfall and value-at-risk: their estimation error, decomposition, and optimization. *Monetary and economic studies*, 20(1):87–121.

Ziggel, D., Berens, T., Weiß, G. N., and Wied, D. (2014). A new set of improved value-at-risk backtests. *Journal of Banking & Finance*, 48:29–41.