Optimal Strategies with Tail Correlation Constraints

by

Eduard Ringe

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Quantitative Finance

Waterloo, Ontario, Canada, 2014

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Optimal strategies under worst-case scenarios have been studied in Bernard et al. [2013a]. Bernard et al. utilize copulas to construct cost-efficient strategies with a predefined dependence structure in the tail between the payoff and the market. In their study they show that such strategies with state-dependent copula constraints dominate traditional diversification strategies in terms of the provided protection in the states of market downturns. We derive similar strategies, however using correlation constraints instead of copula constraints in the tail. We found that for an investor seeking negative dependence with the market, it is cheaper to construct a strategy with conditional correlation constraint in the tail. However, the constructed strategies with conditional correlation constraints do not provide sufficient protection in bad states of the economy. Therefore, when analyzing a strategy, negative correlation with the market in the tail is not a sufficient indicator for the protection level in the event of a market crisis.

Acknowledgements

Foremost, I would like to express my deep and sincere gratitude to Prof. Carole Bernard, who has supported me throughout my thesis with her knowledgeable advise and many hours of her time. I would also like to thank my second readers Prof. Adam Kolkiewicz and Prof. David Saunders for their valuable time and constructive comments.

Last but not least, I would like to thank my family and friends. A special gratitude goes to my sister Natalija and my mother Natalja Ringe for the never ending support and the words of encouragement. I also wish to thank my friend Louis Bélisle for his support.

Dedication

To my family and friends, who are always there for me.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

This thesis builds on the earlier work of Dybvig [1988] on the payoff distribution model. The idea is to find the cheapest strategy (also called cost-efficient strategy by Bernard et al. [2013b]) that achieves a given distribution. This objective makes sense to any investor who is only interested in the distributional properties of his portfolio and does not care about the states of the economy when he receives the outcomes/cash-flows of the portfolio. However, in practice investors do also care about the way their investment strategy interacts with the economy. Bernard et al. [2013b] give an explicit construction of cost-efficient strategies and the subsequent paper Bernard et al. [2013a] shows how to construct optimal strategies with constraints in the tail.

The payoff distribution model (PDM) was introduced by Dybvig [1988] as a technique to evaluate portfolio performance. In his paper "Distribution Analysis of Portfolio Choice" Dybvig focuses on an investor who only cares for the final distribution of his wealth, disregarding the state of the market when a certain outcome occurs. In a complete and arbitrage-free market he computes the distributional price of a payoff, which is the smallest price of all payoffs with the same distribution. It turns out that the cheapest payoff is anti-monotonic with the state price process.

However, the PDM also provides a framework for generating portfolios with desired distributional properties. Hocquard et al. [2012] use Dybvig's PDM approach to construct strategies with a predefined distribution of final wealth. They generate funds with a built-in insurance using as the target distribution for log-returns a Left Truncated Gaussian distribution. The desired insurance level can be achieved at lower cost than other traditional portfolio insurance strategies like the Constant Proportion Portfolio Insurance (CPPI). Specifically, they generate a fund with returns characterized by a Left Truncated Gaussian distribution and compare its effectiveness with traditional portfolio insurance strategies like the stop-loss insurance, put-option-based replication insurance, and CPPI.

A multivariate version of the Payoff Distribution Model is implemented in Hocquard et al. [2007]. The authors do not only replicate the marginal distribution of hedge fund returns but also their dependence with other tradable assets by utilizing copulas. They generate the distribution of hedge fund returns and its dependence with the investor's portfolio by trading only in the investors portfolio and the tradable asset. For payoff replication they find that optimal hedging, with expected square hedging error as a measure of quality of replication, dominates delta hedging.

The main goal of this thesis is to construct cost-efficient strategies with conditional correlation constraints and compare those to cost-efficient strategies with state-dependent copula constraints as studied in Bernard et al. [2013b] and Bernard et al. [2013a]. We focus our study on the cost and the provided protection during a crisis by a given strategy. Since correlation is a broader way of describing dependence structure, we anticipate to see a decrease in the cost when focusing at a target tail correlation. The two questions remain: Is enforcing a correlation constraint in the lower tail enough to provide sufficient protection in a crisis? How does the generated protection compare to the one obtained when using a copula constraint in the tail?

Explicit representations of cost-efficient strategies with copula constraints were first derived in Bernard et al. [2013b]. There, Bernard et al. provide a sufficient condition for a payoff to be cost-efficient which is a central tool for construction of cost-efficient payoffs. Further research was done in Bernard et al. [2013a], where the authors compare traditional diversification strategies with cost-efficient strategies satisfying a copula constraint. Bernard et al. pay special attention to how the strategies behave in a market crisis situation. The latter strategies outperform traditional strategies in terms of the protection level provided, though at a higher cost.

This thesis builds mainly on the work of Hocquard et al. [2012] and Bernard et al. [2013a]. We make the following contributions.

First, we extend the work of Chek Hin Choi [2012] on strategies satisfying global correlation constraints to the case of a two-dimensional Black-Scholes market where the benchmark is not the Growth Optimal Portfolio, leading to more realistic strategies.

Second, we derive cost-efficient strategies with conditional correlation constraints in the tail. In contrast to strategies with global correlation constraint this provides more realistic payoff structures for the investor.

Third, we discuss the deficits of strategies with conditional correlation constraints. Although they are cheaper than strategies with state dependent copula constraints they do not provide sufficient protection in a market crisis situation.

The rest of the thesis is organized as follows. In the remaining sections of Chapter 1 the market setting is presented. In Chapter 2 we start with an introduction to the PDM and give later examples of the construction and the replication of payoffs with desired distributional properties. Chapter 3 focuses on strategies with global correlation constraints both in one- and two-dimensional market settings and different benchmarks. In Chapter 4 we derive strategies with conditional correlation constraints in the tail. Chapter 5 deals with strategies satisfying state dependent copula constraints in the tail. Finally, Chapter 6 compares all the strategies especially focusing on the cost and produced protection in case of a market crisis. Appendix A presents useful identities and most of the proofs can be found in Appendix B.

1.2 Market Setting

In this section we describe the financial market and the connection between the Growth Optimal Portfolio (GOP) and pricing. For the ease of exposition we consider a multidimensional Black-Scholes market.

We assume that the market is complete and arbitrage-free. Furthermore, the market is frictionless and the trading can be done continuously. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the corresponding probability space. Then there exists a unique state-price process $(\xi_t)_t$ such that for any traded asset S in this market $(\xi_t S_t)_t$ is a martingale under $\mathbb{P}$. The market is equipped with a risk free bond $\{B_t = B_0 e^{rt}, t \geq 0\}$ and risky assets $S^1, S^2, ..., S^n$ which evolve according to

$$\frac{dS_t^i}{S_t^i} = \mu_i dt + \sigma_i dW_t^i, \quad i = 1, 2, ..., n,$$

where $\{W_t^i, t \geq 0\}$ for $i = 1, 2, ..., n$ are correlated Brownian motions with

$$\rho_{ij} = Corr(W_t^i, W_{t+s}^j), \ \forall t, s \geq 0.$$

Similar as in Bernard et al. [2013b] we define the cost of a payoff in the following way.

Definition 1.2.1. *The initial price or the cost of a strategy with terminal payoff X_T is given by*

$$c(X_T) = E[\xi_T X_T], \tag{1.1}$$

where the expectation is taken under the physical measure $\mathbb{P}$.

However, an alternative is to use the GOP as the numeraire for pricing. To get a better understanding of the GOP, in what is to follow we derive it as the constant-mix portfolio that maximizes the expected growth rate. Now consider a portfolio with weights π_i invested in the risky assets S^i which are kept constant throughout the investment period through rebalancing. The remaining proportion $1 - \sum_{i=1}^{n} \pi_i$ is invested in the risk-free asset. The value process of such investment strategy follows the equation

$$\frac{dS_t^\pi}{S_t^\pi} = \mu_\pi dt + \sigma_\pi dW_t^\pi, \quad \text{with} \quad W_t^\pi = \frac{\sum_{i=1}^{n} \pi_i \sigma_i W_t^i}{\sqrt{\pi^\top \Sigma \pi}},$$

where Σ is the variance-covariance matrix with

$$(\Sigma)_{ij} = \rho_{ij}\sigma_i\sigma_j,\ \mu_\pi = r + \pi^T(\mu - r\mathbf{1}), \text{ and } \sigma_\pi^2 = \pi^T \Sigma \pi$$

where $\mathbf{1}$ denotes a vector of size n of ones. The payoff of this strategy at time t is equal to

$$S_t^\pi = S_0^\pi \exp((\mu_\pi - \tfrac{1}{2}\sigma_\pi^2)t + \sigma_\pi W_t^\pi). \tag{1.2}$$

The portfolio which maximizes the expected growth rate $\mu_\pi - \frac{1}{2}\sigma_\pi^2$ is referred to as the Growth Optimal Portfolio (π^* or S_t^*) with $\pi^* = \Sigma^{-1}(\mu - r\mathbf{1})$. The cdf of S_T^* is given by

$$F_{S_T^*}(x) = \Phi\left(\frac{\log\left(\frac{x}{S_0^*}\right) - \left(\mu_* - \frac{\sigma_*^2}{2}\right)T}{\sigma_*\sqrt{T}}\right). \tag{1.3}$$

The following proposition establishes the relationship between the state-price process and the GOP. Please refer to Appendix A of Bernard et al. [2013a] for the proof of Proposition 1.2.1.

Proposition 1.2.1 (State-price process)**.** *In the multidimensional Black-Scholes market, the state-price process $\{\xi_t, t \geq 0\}$ is given by*

$$\xi_t = \frac{S_0^*}{S_t^*}, \tag{1.4}$$

where S_t^* is the GOP.

We can use the GOP as a numeraire and rewrite the pricing formula as

$$c(X_T) = E_{\mathbb{P}}[\xi_T X_T] = E_{\mathbb{P}}\left[\frac{S_0^*}{S_T^*} X_T\right]. \tag{1.5}$$

The GOP S^* can be interpreted as a major market index, see Platen and Heath [2006]. Similar as in Adrian and Brunnermeier [2010], we define a market crisis as an event when the market goes below its Value-at-Risk, with other words it corresponds to the states

$$\{S_T^* < q_\alpha\}, \tag{1.6}$$

where q_α is such that $\mathbb{P}(S_T^* < q_\alpha) = \alpha$ (e.g. $\alpha = 5\%$).

1.2.1 Special Case: One-Dimensional Market

In a one dimensional Black-Scholes market the unique state-price process is given by $\xi_t = e^{-rt} \frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_t = e^{-rt} e^{-\frac{1}{2}\left(\frac{\mu_1 - r}{\sigma_1}\right)^2 t - \left(\frac{\mu_1 - r}{\sigma_1}\right) W_t}$. Then using Proposition 1.2.1 we can compute the GOP S_t^* as

$$\begin{aligned} S_t^* &= \frac{S_0^*}{\xi_t} \\ &= S_0^* e^{rt + \frac{1}{2}\left(\frac{\mu_1 - r}{\sigma_1}\right)^2 t + \left(\frac{\mu_1 - r}{\sigma_1}\right) W_t} \\ &= S_0^* e^{\left(r + \left(\frac{\mu_1 - r}{\sigma_1}\right)^2 - \frac{1}{2}\left(\frac{\mu_1 - r}{\sigma_1}\right)^2\right) t + \left(\frac{\mu_1 - r}{\sigma_1}\right) W_t} \\ &= S_0^* e^{\left(\mu_* - \frac{\sigma_*^2}{2}\right) t + \sigma_* W_t} \end{aligned}$$

with $\mu_* = r + \left(\frac{\mu_1 - r}{\sigma_1}\right)^2$ and $\sigma_* = \frac{\mu_1 - r}{\sigma_1}$.

The setting in Table 1.1 will be used for the construction of all the payoffs in the one-dimensional market setting. Using the relationship derived in the previous paragraph we calculated the drift and volatility of the GOP to fit the dynamics of the GOP in Bernard et al. [2013a]. This is needed to be able to compare the improvement in cost of the strategy and the change in other metrics to the conditional correlation constraint strategies.

Table 1.1: Parameter values for the one-dimensional Black-Scholes market

Parameter	**Symbol**	**Value**
Initial Underlying Price	S_0	100
Drift Rate of the Underlying S	μ_1	0.08
Volatility of the Underlying S	σ_1	0.2372
Drift Rate of the GOP S^*	μ_*	0.066
Volatility of the GOP S^*	σ_*	$\sqrt{0.016}$
Risk-free Rate	r	0.05

1.2.2 Special Case: Two-Dimensional Market

Consider a two-dimensional Black-Scholes market, where the prices of the risky assets evolve according to

$$\begin{cases} \frac{dS_t^1}{S_t^1} = \mu_1 dt + \sigma_1 dW_t^1 \\ \frac{dS_t^2}{S_t^2} = \mu_2 dt + \sigma_2 dW_t \end{cases}, \tag{1.7}$$

where W_t^1 and W_t are two correlated Brownian motions under the physical measure $\mathbb{P}$ with

$$W_t = \rho_{12} W_t^1 + \sqrt{1-\rho_{12}^2} W_t^2,$$

where W_t^1 and W_t^2 are independent. In Table 1.2 we set the parameters which will be used for the construction of all the payoffs in a two-dimensional market setting. We choose the parameters as in Bernard et al. [2013a] to match the results for strategies in Chapter 5.

1.3 Cost-Efficiency

The notion of cost-efficiency is crucial for this thesis. We use it to derive cheapest payoffs with different types of constraints in subsequent chapters. It turns out that cost-efficiency is intimately connected with anti-monotonicity.

Definition 1.3.1. *A subset A of $\mathbb{R}^2$ is anti-monotonic if for any (x_1, y_1) and $(x_2, y_2) \in A$, $(x_1 - x_2)(y_1 - y_2) \le 0$.*

Definition 1.3.2. *A random pair (X, Y) is anti-monotonic if there exists an anti-monotonic set A of $\mathbb{R}^2$ such that $\mathbb{P}((X, Y) \in A) = 1$.*

Table 1.2: Parameter values for the two-dimensional Black-Scholes market with $i = 1, 2$

Parameter	**Symbol**	**Value**
Initial Underlying Price	S_0^i	100
Drift Rate of the Underlying S^1	μ_1	0.07
Volatility of the Underlying S^1	σ_1	0.20
Drift Rate of the Underlying S^2	μ_2	0.08
Volatility of the Underlying S^2	σ_2	0.30
Drift Rate of the GOP S^*	μ_*	0.066
Volatility of the GOP S^*	σ_*	$\sqrt{0.016}$
Correlation	ρ_{12}	0.25
Risk-free Rate	r	0.05

Lemma 1.3.1. *For given marginal distributions of X and Y, $\mathbb{E}[XY]$ is minimal whenever X and Y are anti-monotonic.*

The bounds for $\mathbb{E}[XY]$ are given by the Fréchet-Hoeffding bounds as in Lemma A.1 of Bernard et al. [2013c]

$$l \triangleq \mathbb{E}[F_X^{-1}(U)F_Y^{-1}(1-U)] \leq \mathbb{E}[XY] \leq \mathbb{E}[F_X^{-1}(U)F_Y^{-1}(U)] \triangleq u, \tag{1.8}$$

where $U \sim \text{Unif}(0,1)$, and $F_X(\cdot)$ and $F_Y(\cdot)$ are the cdf of X and Y respectively.

The connection between cheapest payoffs with given distribution and anti-monotonicity was studied in Bernard et al. [2013b]. Bernard et al. introduce the notion of "cost-efficiency" and a way to generate cost-efficient payoffs with a desired distribution.

Definition 1.3.3. *A strategy is "cost-efficient" if any other strategy that generates the same distribution costs at least as much.*

Proposition 1.3.1. *A payoff is cost-efficient if and only if it is non-increasing in the state-price ξ_T almost surely.*

Although the essential idea of a cheapest payoff with a given distribution was already introduced by Dybvig [1988], the formula for constructing such cost-efficient strategy was first presented by Bernard et al. [2013b] in the form of the following corollary.

Corollary 1.3.1. *Let ξ_T be continuously distributed and F be a given cdf. Define*

$$X_T^* = F^{-1}(1 - F_{\xi_T}(\xi_T)). \tag{1.9}$$

Then, X_T^* is the cheapest way to achieve the distribution F. It is also almost surely unique.

Please refer to the appendix of Bernard et al. [2013b] for the proof of Corollary 1.3.1.

Chapter 2

Payoff Distribution Replication

2.1 Payoff Distribution Model

The Payoff Distribution Model (PDM) was first introduced by Dybvig [1988] and originally the model's aim was to measure the performance of a fund by evaluating the payoff distribution. To this end Dybvig introduces the notion of the distributional price. The following definition is a central result of Dybvig [1988].

Definition 2.1.1. *The "distributional price" of a cdf F is given by*

$$P_D(F) = \min_{\{X_T | X_T \sim F\}} c(X_T) = \int_0^1 F_{\xi_T}^{-1}(u) F^{-1}(1-u) du, \tag{2.1}$$

where F_{ξ_T} is the state price density.

To better understand equation (2.1) please refer to Section 1.3 on cost-efficiency and anti-monotonicity. Note that one can interpret the distributional price of a distribution F as the price of a strategy with distribution F and which is anti-monotonic with the state-price ξ_T.

Dybvig describes a measure for inefficiency of fund payoffs utilizing Definition 2.1.1. The inefficiency is measured as the difference between the initial wealth invested in the fund with payoff X_T and the distributional price of the distribution F_{X_T}:

$$\text{inefficiency} = c(X_T) - P_D(F_{X_T})$$

The PDM approach was new in the sense that it allowed performance measurement to depend on all the moments of a distribution and not only on the first two moments like in the traditional mean-variance approach. It could be used to measure for example the performance of hedge funds as in Amin and Kat [2003] and Hocquard et al. [2007]. In practice, the idea is to evaluate the statistical properties of the ex-post distribution of a fund by generating a fund with the same distributional properties more efficiently using a dynamic trading strategy on liquid assets and compare the costs.

However, in this thesis we focus not on the evaluation of fund performances but rather on generation of funds with desired return distributions. To illustrate the PDM approach, in the following section we present steps to generate a portfolio with desired distribution and derive the payoff function for managing downside risk.

2.2 Constructing Desired Payoffs

In this section we focus on the distribution of log returns and construct two strategies, both in the one-dimensional Black-Scholes market. The goal of the first strategy is to mimic the returns of the Growth Optimal Portfolio (GOP), and the goal of the second strategy is also to mimic the returns of the GOP while providing protection against returns below -5%. Amin and Kat [2003] show that, given an underlying asset S with log returns R_{Under} and a specified target distribution of log returns F_{Target} it is possible to generate the desired distributional properties of the returns at maturity. In the spirit of Corollary 1.3.1 we derive the desired payoff as follows.

Proposition 2.2.1. *The payoff of a cost-efficient strategy with the desired distribution of log returns F_{Target} in the Black-Scholes market with one asset S and $\mu_1 > r$ is given by*

$$X_T = S_0 e^{g(\log(S_T/S_0))} \tag{2.2}$$

where the function $g(x)$ is defined as

$$g(x) = F^{-1}_{Target}(F_{Under}(x)), \ \forall x \in \mathbb{R}^+, \tag{2.3}$$

where F_{Under} is the distribution of the log returns of the underlying asset S.

To illustrate the construction of a desired payoff we first consider a simple example where we aim at a normal cdf for the log returns with different mean and volatility than the underlying S. The cdf of the log returns of the underlying is given by

$$F_{Under}(x) = \Phi\left(\frac{x - (\mu_1 - \sigma_1^2/2)T}{\sigma_1\sqrt{T}}\right), \tag{2.4}$$

The target distribution of log returns with mean $(\mu_* - \sigma_*^2/2)T$ and volatility $\sigma_*\sqrt{T}$ has an inverse cdf given by

$$F_{Target}^{-1}(y) = (\mu_* - \sigma_*^2/2)T + \sigma_*\sqrt{T}\Phi^{-1}(y). \tag{2.5}$$

Proposition 2.2.2. ***(Strategy 1)*** *The cost-efficient payoff with the same distribution of log returns as the Growth Optimal Portfolio in a one-dimensional Black-Scholes market is given by*

$$X_T = S_0 e^{g(\log(S_T/S_0))}$$

where the function $g(x)$ is defined as

$$g(x) = (\mu_* - \sigma_*^2/2)T + \sigma_*\sqrt{T}\left(\frac{x - (\mu_1 - \sigma_1^2/2)T}{\sigma_1\sqrt{T}}\right).$$

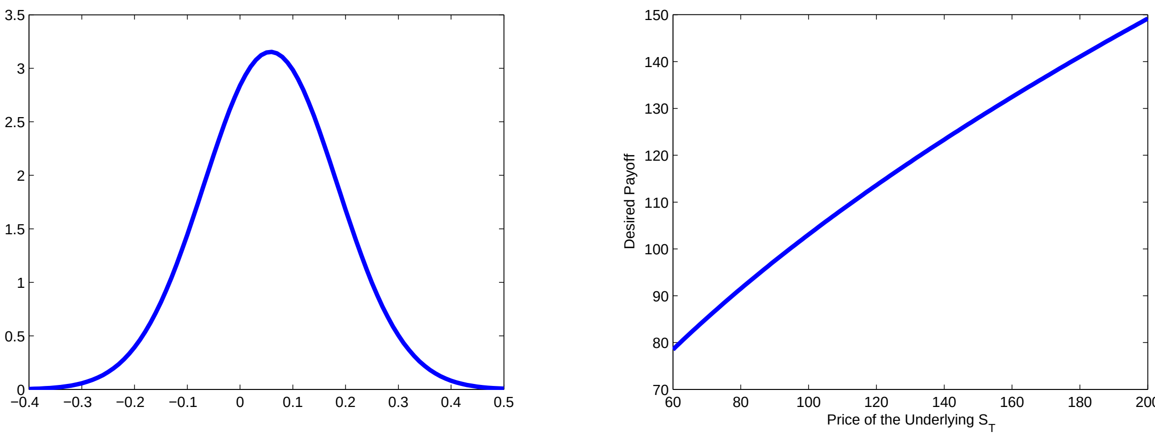
Proof. It is a straight application of Proposition 2.2.1 and results in equations (2.4) and (2.5). □

Figure 2.1 shows the pdf of the log returns of the payoff in Proposition 2.2.2 and the payoff itself, where μ_*, σ_*, μ_1 and σ_1 are chosen as in Table 1.1.

In our next example we focus on a risk management application and seek a distribution with a downside protection. Similar as in Hocquard et al. [2012] we choose for the target log returns a Left Truncated Gaussian distribution with the probability density function described through

$$f(y, \mu, \sigma^2, a) = \frac{(1/\sqrt{2\pi\sigma^2})\exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right)}{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)}\mathbb{1}_{\{a<y\}}, \tag{2.6}$$

with Φ being the standard normal cumulative distribution function. The Left Truncated Gaussian distribution is of a special interest to an investor seeking protection. Indeed, the truncation induces a higher mean, lower volatility and positive skewness when compared to the non-truncated Gaussian. The truncated pdf can be found in Panel A of Figure 2.2.



Panel A Panel B

Figure 2.1: Panel A presents the pdf of $\log(S_T^*/S_0^*)$ for $T = 1$. Panel B presents the payoff X_T from Proposition 2.2.2.

The inverse of the target cdf of log returns has the following form

$$F_{Target}^{-1}(y) = (\mu_* - \sigma_*^2/2)T + \sigma_*\sqrt{T}\Phi^{-1}\Big[\Phi\Big(\frac{a - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\Big) + y\Big[1 - \Phi\Big(\frac{a - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\Big)\Big]\Big]. \tag{2.7}$$

Proposition 2.2.3. ***(Strategy 2)*** *The cost-efficient payoff with the same mean and volatility of the log returns as the Growth Optimal Portfolio, but such that the log returns never go below the insurance level a in a one-dimensional Black-Scholes market is given by*

$$X_T = S_0 e^{g(\log(S_T/S_0))}$$

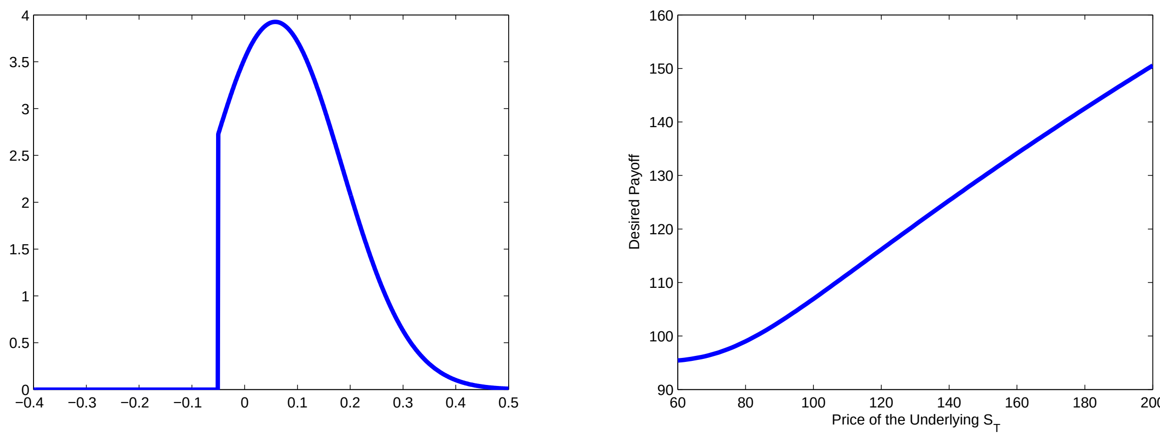
where the function $g(x)$ is defined as

$$g(x) = (\mu_* - \sigma_*^2/2)T + \sigma_*\sqrt{T}\Phi^{-1}\Big[\Phi\Big(\frac{a - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\Big) + \Phi\left(\frac{x - (\mu_1 - \sigma_1^2/2)T}{\sigma_1\sqrt{T}}\right)\Big[1 - \Phi\Big(\frac{a - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\Big)\Big]\Big].$$

Proof. It is a straight application of Proposition 2.2.1 and results in equations (2.4) and

(2.7). □

Panel B in Figure 2.2 presents the payoff of the strategy in Proposition 2.2.3.



Panel A Panel B

Figure 2.2: Panel A presents the pdf of a left truncated Gaussian with mean $(\mu_* - \sigma_*^2/2)T$, volatility $\sigma_*\sqrt{T}$, truncation level $a = -5\%$ and $T = 1$. Panel B presents the payoff X_T from Proposition 2.2.3.

2.3 Replicating the Payoff

In this section we present a method to replicate the payoffs derived in the previous section. In fact, it can be used to replicate all the payoffs we present in this thesis. By the term replication we understand the process of generating the desired payoff by trading in the underlying asset.

Once we have specified the payoff X_T we can replicate it by applying an optimal dynamic trading strategy by selecting the portfolio (V_0, ϕ) so as to minimize the expected squared hedging error

$$E[\beta_T^2\{V_T(V_0, \phi) - X_T\}^2], \tag{2.8}$$

where $X_T = S_0 e^{g(\log(S_T/S_0))}$ is the desired payoff at maturity, β_T is the discount factor, V_0 is the initial value of the replicating portfolio and ϕ is a weight vector.

Suppose that $(\Omega, \mathbb{P}, \mathscr{F})$ is a probability space with filtration $\mathbb{F} = \{\mathscr{F}_0, ..., \mathscr{F}_T\}$ under which the stochastic process S_t is defined. We also introduce the weight vector $\phi = \{\phi_1, \phi_2, ..., \phi_T\}$, where ϕ_t is the number of parts of asset S invested during the period $(t-1, t]$. Using the self-financing condition of the hedging strategy one can get the following representation of V_T

$$\beta_T V_T(V_0, \phi) = V_0 + \Sigma_{t=1}^{T} \phi_t (\beta_t S_t - \beta_{t-1} S_{t-1}). \tag{2.9}$$

The following theorem describes the optimal trading strategy (V_0, ϕ).

Theorem 2.3.1 (Optimal trading strategy)**.** *The expected square hedging error* $E[\beta_T^2 \{V_T(V_0, \phi) - X_T\}^2]$ *is minimized by choosing recursively* ϕ_T, ..., ϕ_1 *satisfying*

$$\phi_t = (Var(S_t \mid \mathscr{F}_{t-1}))^{-1} E[\{S_t - E[S_t \mid \mathscr{F}_{t-1}]\} X_t \mid \mathscr{F}_{t-1}],\ t = T, ..., 1, \tag{2.10}$$

where X_T, ..., X_0 *are defined as*

$$\beta_{t-1} X_{t-1} = \beta_t E[X_t \mid \mathscr{F}_{t-1}] - \phi_t E[\beta_t S_t - \beta_{t-1} S_{t-1} \mid \mathscr{F}_{t-1}],\ t = T, ..., 1, \tag{2.11}$$

and the optimal value of V_0 *is* X_0.

The replication of the payoff is done by implementing a dynamic optimal hedging strategy on the underlying asset as in Section 3.2 of Del Moral et al. [2006]. Therefore, it is essential for the underlying asset, or at least for a proxy, to be liquid. Please refer to Schweizer [1995] for a detailed description of the variance-optimal hedging methodology in discrete time. For the ease of exposition we are going to demonstrate the optimal hedging strategy in the Black-Scholes setting using Monte Carlo simulations. For an example with real market data please refer to Hocquard et al. [2012].

In the case when the price process is Markovian, i.e. the price of the underlying is path independent and only depends on the previous value, and if additionally the target payoff X_T only depends on the terminal price, then $X_t = f_t(S_t)$ and $\phi_t = \psi_t(S_{t-1})$ and the optimal trading strategy (V_0, ϕ) can be derived as

$$
\begin{aligned}
L_{1t}(s) &= E[S_t \mid S_{t-1} = s], \\
L_{2t}(s) &= E[S_t^2 \mid S_{t-1} = s], \\
A_t(s) &= L_{2t}(s) - L_{1t}(s)^2, \\
\psi_t(s) &= A_t(s)^{-1} E[\{S_t - L_{1t}(s)\} f_t(S_t) \mid S_{t-1} = s], \\
U_t(s, x) &= 1 - (L_{1t}(s) - \beta_{t-1} s / \beta_t) A_t(s)^{-1} (x - L_{1t}(s)), \\
f_{t-1}(s) &= \frac{\beta_t}{\beta_{t-1}} E[U_t(s, S_t) f_t(S_t) \mid S_{t-1} = s].
\end{aligned}
$$

Dynamic programming with Monte Carlo simulations and linear interpolation was used to compute X_t for all t as in the algorithm described in Section 3.2 of Del Moral et al. [2006].

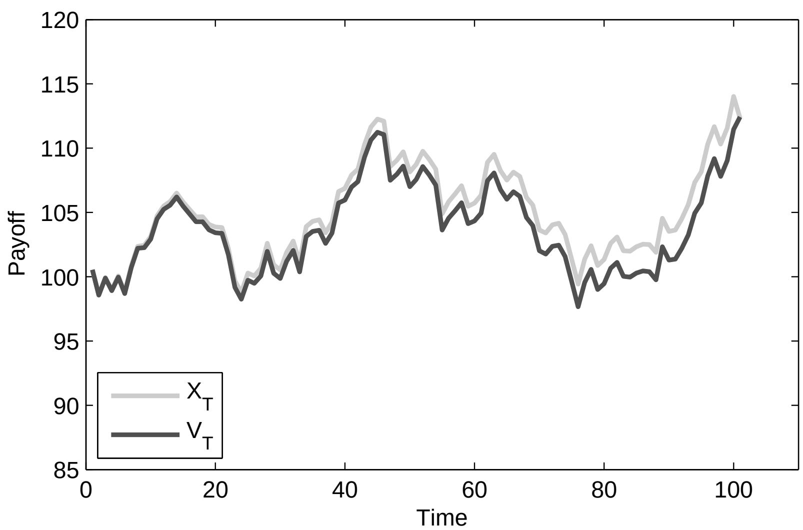


Figure 2.3: Sample path showing the replication of Strategy 1 in Proposition 2.2.2

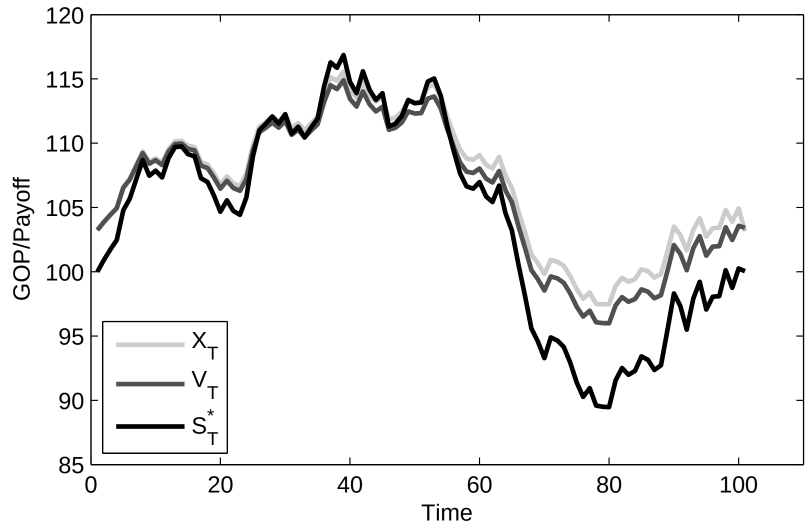


Figure 2.4: Sample path showing the replication of Strategy 2 in Proposition 2.2.3

We define the hedging error relative to the desired payoff X_T as

$$\text{error} = (V_T(V_0, \phi) - X_T)/X_T.$$

Consider the statistic in Table 2.1 of the hedging error when using 500 full path replications. For each replication we used 100,000 simulations and a grid of 150 points for the interpolation.

Table 2.1: Statistics for the relative hedging error

Statistic	Strategy 1	Strategy 2
Mean	0.09%	-0.38%
Std. dev.	0.2%	1%

Chapter 3

Optimal Strategies with Global Correlation Constraint

The motivation for this chapter is to present a general case for a cost-efficient strategy with a correlation constraint with a benchmark and its consequences for the payoff structure before focusing on the tail correlation constraint in Chapter 4.

3.1 Formulation of the Problem

In this section we are going to derive a cost-efficient payoff X_T^* with a desired distribution F and correlation ρ_0 with a benchmark A_T, which has the cdf H. This benchmark can be any process which depends on the market dynamics like a stock or an index. In other words we are looking for X_T^* which solves the following problem

$$\min_{\left\{X_T \middle| \substack{X_T \sim F, \\ \xi_T \sim G, \\ A_T \sim H,\ A_T \geq 0\ a.s., \\ Corr(A_T, X_T) = \rho_0\ \in\ [-1,1]}\right\}} \mathbb{E}[\xi_T X_T]. \tag{3.1}$$

Note that the main results in this chapter have been derived in Section 3 of Chek Hin Choi [2012]. We extend this study to the case of a two-dimensional Black-Scholes market. All important proofs are given in Appendix B. We also derive new applications when the benchmark is not the Growth Optimal Portfolio (GOP), leading to more realistic strategies.

Since the marginal distributions of X_T and A_T are known, the correlation constraint can be rewritten as

$$\mathbb{E}[A_T X_T] = \rho_0 Std(A_T) Std(X_T) + \mathbb{E}[A_T]\mathbb{E}[X_T] \triangleq a_0. \tag{3.2}$$

The bounds for a_0 are given by the Fréchet-Hoeffding bounds as in Section 1.3

$$l \triangleq \mathbb{E}[F^{-1}(U)H^{-1}(1-U)] \leq a_0 \leq \mathbb{E}[F^{-1}(U)H^{-1}(U)] \triangleq u,$$

where $U \sim \text{Unif}(0,1)$. The problem in equation (3.1) can be now rewritten as

$$\min_{\left\{X_T \,\middle|\, \substack{X_T \sim F,\\ \xi_T \sim G,\\ A_T \sim H,\ A_T \geq 0\ a.s.,\\ \mathbb{E}[A_T X_T] = a_0}\right\}} \mathbb{E}[\xi_T X_T]. \tag{3.3}$$

To solve the problem in equation (3.3), consider the following dual problem

$$\min_{\left\{X_T \,\middle|\, \substack{X_T \sim F,\\ \xi_T \sim G,\\ A_T \sim H,\ A_T \geq 0\ a.s.}\right\}} \mathbb{E}[\xi_T X_T] + \lambda \mathbb{E}[A_T X_T], \tag{3.4}$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier. We slightly rewrite the objective function to give an intuition of how to find the optimal solution to this problem for some fixed $\lambda \in \mathbb{R}$

$$\min_{\left\{X_T \,\middle|\, \substack{X_T \sim F,\\ \xi_T \sim G,\\ A_T \sim H,\ A_T \geq 0\ a.s.}\right\}} \mathbb{E}[(\xi_T + \lambda A_T) X_T]. \tag{3.5}$$

Proposition 3.1.1. *For some given $\lambda > 0$, the optimal solution to the problem in equation (3.5) is*

$$X_T^*(\lambda) = F^{-1}(1 - L_{\xi_T + \lambda A_T}(\xi_T + \lambda A_T)), \tag{3.6}$$

where $L_{\xi_T + \lambda A_T}$ is the cdf of $\xi_T + \lambda A_T$.

Note that $X_T^*(\lambda)$ is anti-monotonic with $\xi_T + \lambda A_T$, therefore it is also anti-monotonic with ξ_T and in the spirit of Corollary 1.3.1 it is the cheapest way to achieve the distribution F. Next step is to establish the connection between the original problem in equation (3.1) and the dual problem in equation (3.4). The idea is to find λ^* such that $\mathbb{E}[A_T X_T^*(\lambda^*)] = a_0$. We first show the existence of such λ^*.

Lemma 3.1.1. $\mathbb{E}[A_T X_T^*(\lambda)]$ *is a continuous function in λ for $\lambda \geq 0$. Furthermore, for* *any* $a_0 \in (l, \mathbb{E}[F^{-1}(1-G(\xi_T))A_T]]$, *there exists* $\lambda^* \in [0,+\infty)$ *such that* $\mathbb{E}[A_T X_T^*(\lambda^*)] = a_0$.

Lemma 3.1.1 shows that λ^* is only guaranteed to exist for a certain range of $Corr(X_T, A_T)$. The next corollary provides a sufficient condition for λ^* to exist for all $a_0 \in (l, u]$.

Corollary 3.1.1. *If there exists a strictly decreasing and continuous function h such that* $A_T = h(\xi_T)$, *then for any* $a_0 \in (l, u]$, *there exists* $\lambda^* \in [0,+\infty)$ *such that* $\mathbb{E}[A_T X_T^*(\lambda^*)] = a_0$.

Finally, the following proposition shows that the optimal solution to the dual problem in equation (3.4) is in fact an optimal solution to the problem in equation (3.3).

Proposition 3.1.2. *For* λ^* *chosen such that* $\mathbb{E}[A_T X_T^*(\lambda^*)] = a_0 \in (l, \mathbb{E}[F^{-1}(1-G(\xi_T))A_T]]$, $X_T^*(\lambda^*)$ *is the optimal solution to the problem in equation (3.3).*

Corollary 3.1.2. *If there exists a strictly decreasing and continuous function h such that* $A_T = h(\xi_T)$, *then for* λ^* *chosen such that* $\mathbb{E}[A_T X_T^*(\lambda^*)] = a_0 \in (l, u]$, $X_T^*(\lambda^*)$ *is the optimal solution to the problem in equation (3.1).*

3.2 Growth Optimal Portfolio as Benchmark

In this section we apply Proposition 3.1.1 to find a cost-efficient strategy with the same distribution $F_{S_T^*}$ as the Growth Optimal Portfolio (GOP), which satisfies a global correlation constraint with the GOP. For a detailed description of the GOP please refer to Section 1.2. Thus, our goal is to solve the following problem

$$\min_{\left\{ X_T \,\middle|\, \substack{X_T \sim F_{S_T^*}, \\ \xi_T \sim G, \\ S_T^* \sim F_{S_T^*},\ S_T^* \geq 0 \ a.s., \\ Corr(S_T^*, X_T) = \rho_0 \ \in \ [-1,1]} \right\}} \mathbb{E}[\xi_T X_T].$$

Lemma 3.2.1. *The cdf of* $\xi_T + \lambda S_T^*$ *is given by*

$$L_{\xi_T + \lambda S_T^*}(y) = \begin{cases} 0 & \text{if } y < 2\sqrt{\lambda S_0^*}, \\ \Phi\left(\frac{x_\lambda^{MAX} - (\mu_* - \sigma_*^2/2)T}{\sigma_* \sqrt{T}}\right) - \Phi\left(\frac{x_\lambda^{MIN} - (\mu_* - \sigma_*^2/2)T}{\sigma_* \sqrt{T}}\right) & \text{otherwise}, \end{cases}$$

where x_λ^{MAX} *and* x_λ^{MIN} *are the bigger and the smaller roots of the equation* $e^{-x} + \lambda S_0^* e^x = y$ *respectively and* $\lambda \geq 0$.

Combining the results of Lemma 3.2.1, Corollary 3.1.2 and Proposition 3.1.1 we are getting the following constrained cost-efficient strategy.

Proposition 3.2.1. ***(Strategy 3)*** *A cost-efficient strategy X_T^* with the same distribution of final wealth as the the GOP but such that $Corr(S_T^*, X_T^*) = \rho_0$ is given by*

$$X_T^*(\lambda^*) = S_0^* \exp\left[\left(\mu_* - \frac{\sigma_*^2}{2}\right)T + \sigma_*\sqrt{T}\Phi^{-1}\left(1 - \left(\Phi\left(\frac{x_{\lambda_*}^{MAX} - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right)\right.\right.\right.$$
$$\left.\left.\left. + \Phi\left(\frac{x_{\lambda_*}^{MIN} - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right)\right)\mathbb{1}_{\{\xi_T + \lambda S_T^* \geq 2\sqrt{\lambda S_0^*}\}}\right)\right], \tag{3.7}$$

where λ^* is chosen such that $\mathbb{E}[X_T^*(\lambda^*)S_T^*] = a_0$. Please note that $x_{\lambda_*}^{MAX}$ and $x_{\lambda_*}^{MIN}$ both depend on $\xi_T + \lambda S_T^*$.

Figure 3.1 illustrates that λ^* can be in fact found such that $\mathbb{E}[X_T^*(\lambda^*)S_T^*] = a_0$, as long as a_0 is feasible. Payoffs of Strategy 3 are represented in Figure 3.2 for different global correlation constraints. We can observe that those types of payoffs are not very investor-friendly.

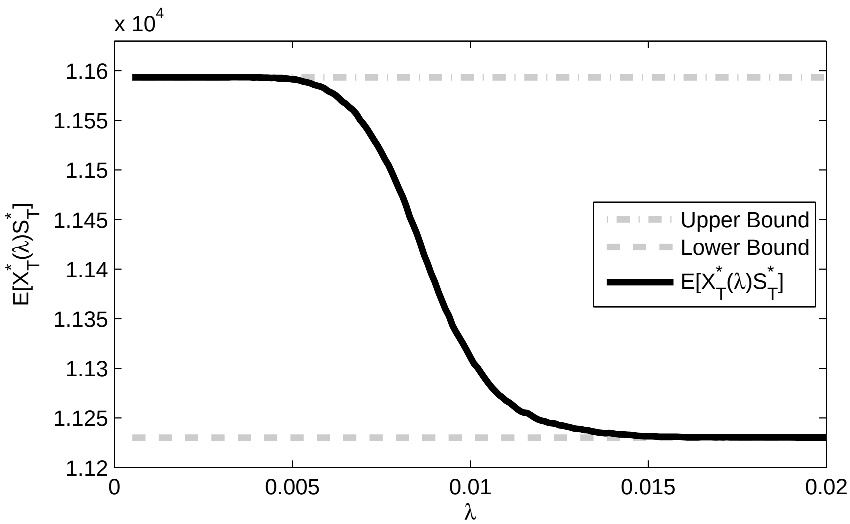


Figure 3.1: Graph of $\mathbb{E}[X_T^*(\lambda)S_T^*]$ against λ

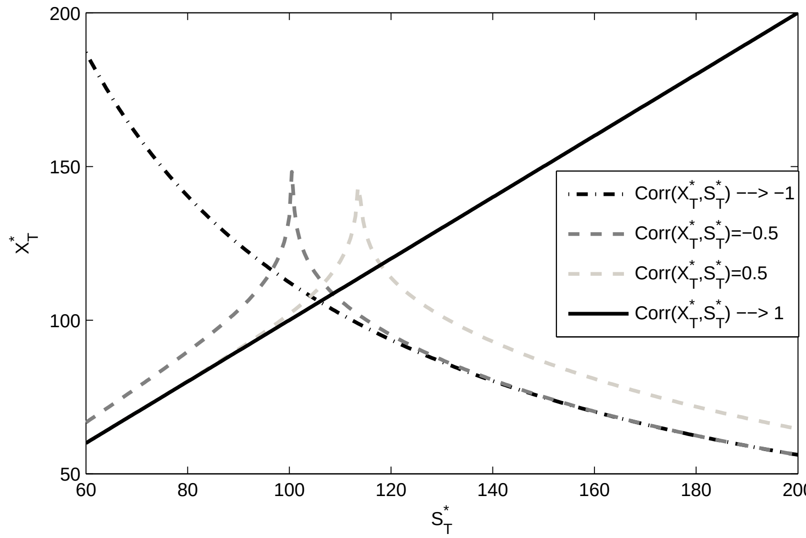


Figure 3.2: Payoffs of Strategy 3 as described in Proposition 3.2.1 with different global correlation constraints

3.3 Stock as Benchmark

In this section we study the case when the benchmark is one of the stocks in a multidimensional market. To this end we consider a two-dimensional Black-Scholes market as described in Section 1.2.2. Our goal is to find a representation of a cost-efficient strategy with the same distribution of final wealth as the GOP S_T^* and a global correlation constraint with the stock S_T^1. Therefore, our aim is to find the solution to the following optimization problem

$$\min_{\left\{X_T \,\middle|\, \substack{X_T \sim F, \\ \xi_T \sim G, \\ S_T^* \sim H,\ S_T^* \geq 0\ a.s., \\ Corr(S_T^1, X_T) = \rho_0\ \in\ [-1,1]}\right\}} \mathbb{E}[\xi_T X_T]. \tag{3.8}$$

Please note that similar as in equation (3.2) the global correlation constraint can be rewritten as $\mathbb{E}[S_T^1 X_T] = a_0$. In order to apply Proposition 3.1.1 we need to find the cdf of $\xi_T + \lambda S_T^1$ first. Since the sum of two log-normal random variables is not log-normal we apply moment matching to find a proxy for the real cdf. The following lemma summarizes the results of moment matching.

Lemma 3.3.1. *A proxy for the cdf of $\xi_T + \lambda S_T^1$ in a two-dimensional Black-Scholes market* *as described in Section 1.2.2 is given by*

$$L_{\xi_T+\lambda S_T^1}(y) = \Phi\left(\frac{\log(y) - \mu_p(\lambda)}{\sigma_p(\lambda)}\right),$$

where

$$\mu_p(\lambda) = -\frac{1}{2}\log\left(\left(Var(\xi_T + \lambda S_T^1) + \mathbb{E}[\xi_T + \lambda S_T^1]^2\right)\mathbb{E}[\xi_T + \lambda S_T^1]^{-4}\right)$$
$$\sigma_p(\lambda) = \sqrt{2\log(\mathbb{E}[\xi_T + \lambda S_T^1]) - 2\mu_p(\lambda)},$$

and $\lambda \geq 0$. *See the proof in Appendix B for more details on* μ_p *and* σ_p*.*

In Figure 3.3 we illustrate how well the approximated cdf resembles the actual cdf by plotting the empirical cdf of $\xi_T + \lambda S_T^1$ based on simulations together with the proxy.

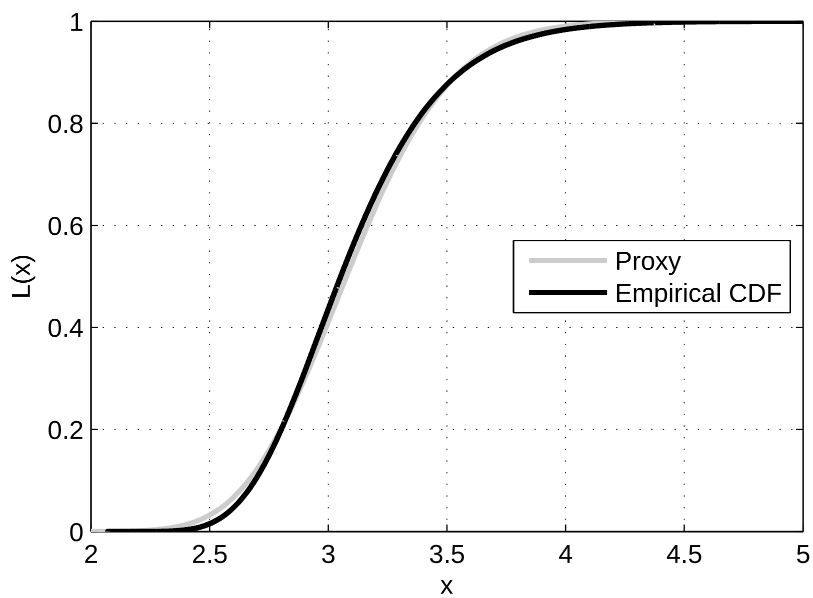


Figure 3.3: Comparison between the empirical cdf of $\xi_T + \lambda S_T^1$ for $\lambda = 0.02$ based on $N = 100,000$ simulations and its proxy based on moment matching

Proposition 3.3.1. *In a two-dimensional Black-Scholes setting, a cost-efficient strategy* X_T^* *with the same distribution of final wealth as the GOP but such that* $Corr(S_T^1, X_T^*) = \rho_0$

is given by

$$X_T^*(\lambda^*) = S_0^* \exp\left(\left(\mu_* - \frac{\sigma_*^2}{2}\right)T + \frac{\sigma_*\sqrt{T}(\mu_p(\lambda^*) - \log(\xi_T + \lambda^* S_T^1))}{\sigma_p(\lambda^*)}\right),$$

with

$$\mu_p(\lambda^*) = -\frac{1}{2}\log\left(\left(Var(\xi_T + \lambda^* S_T^1) + \mathbb{E}[\xi_T + \lambda^* S_T^1]^2\right)\mathbb{E}[\xi_T + \lambda^* S_T^1]^{-4}\right)$$
$$\sigma_p(\lambda^*) = \sqrt{2\log(\mathbb{E}[\xi_T + \lambda^* S_T^1]) - 2\mu_p(\lambda^*)},$$

where λ^* *is chosen such that* $\mathbb{E}[X_T^*(\lambda^*) S_T^1] = a_0$.

Chapter 4

Optimal Strategies with Conditional Correlation Constraints

4.1 Motivation

In the previous chapter we considered strategies with global correlation constraints. As we found out the payoff structure is such that it does not provide protection in a crisis situation. In this chapter we would like to focus on the lower tail of the payoff to achieve a controlled improvement of the strategy according to the needs of the investor. In contrast to Chapter 3 we assume in this chapter that the investor seeks to be negatively correlated with the market in a crisis situation. This can be achieved by considering a conditional correlation constraint, where we aim at achieving a negative correlation with the main market index when the index is below a certain threshold. Note that due to this conditioning we achieve a more investor-friendly payoff profile than with global correlation constraint as in Figure 3.2.

In the view of Chapter 6, where we compare all the discussed strategies, we are specifically interested in the comparison between a cost-efficient strategy with a conditional copula constraint versus a cost-efficient strategy with a conditional correlation constraint. It turns out that although the strategy with the correlation constraint is cheaper when aiming at achieving a certain correlation in the tail with the market, it does not provide the anticipated protection against a crisis. This finding can be also understood backwards: If someone analyzing the correlation in the tail of some strategy finds a negative correlation with the market in the tail it does not necessary mean that this strategy provides enough protection

We follow similar steps as in Chapter 3 where we have shown how to solve the problem with a global correlation constraint. However, in this chapter we assume that the correlation constraint is conditioned on the value of the benchmark A_T. Our goal is to find a cost-efficient strategy X_T with a desired distribution F and satisfying conditional correlation constraint with a benchmark A_T. The correlation constraint can be written as

$$\begin{aligned} Corr(A_T, X_T \mid A_T < a_0) &= \rho_0 \\ &= \frac{Cov(A_T, X_T \mid A_T < a_0)}{\sqrt{Var(A_T \mid A_T < a_0) Var(X_T \mid A_T < a_0)}} \\ &= \frac{\mathbb{E}[A_T X_T \mid A_T < a_0] - \mathbb{E}[A_T \mid A_T < a_0]\mathbb{E}[X_T \mid A_T < a_0]}{\sqrt{Var(A_T \mid A_T < a_0) Var(X_T \mid A_T < a_0)}} \end{aligned}$$

where $\rho_0 \in [-1, 1]$. Hence we can write

$$\begin{aligned} \mathbb{E}[X_T A_T \mid A_T < a_0] &= \rho_0 \sqrt{Var(A_T \mid A_T < a_0) Var(X_T \mid A_T < a_0)} \\ &\quad + \mathbb{E}[A_T \mid A_T < a_0]\mathbb{E}[X_T \mid A_T < a_0] \\ &= \theta_0. \end{aligned} \tag{4.1}$$

Note that in equation (4.1) we know the marginal distributions of X_T and A_T, however we don't know their joint distribution. Therefore, although the values of $Var(A_T \mid A_T < a_0)$ and $\mathbb{E}[A_T \mid A_T < a_0]$ are known, the values of $\mathbb{E}[X_T \mid A_T < a_0]$ and $Var(X_T \mid A_T < a_0)$ are not known from the problem setting. We deal with this issue in more detail in Section 4.3.1.

4.2 Conditional Expectation Constraint

Instead of constraining the correlation in the tail, we first consider the following conditional expectation constraint

$$\mathbb{E}[X_T A_T \mid A_T < a_0] = \theta_0. \tag{4.2}$$

Finding the cost-efficient payoff X_T as described in Section 4.1, but such that it satisfies equation (4.2), is equivalent to solving the following problem

$$\min_{\left\{ X_T \,\middle|\, \substack{X_T \sim F, \\ \xi_T \sim G, \\ A_T \sim H,\ A_T \geq 0 \ a.s., \\ \mathbb{E}[X_T A_T | A_T < a_0] = \theta_0} \right\}} \mathbb{E}[\xi_T X_T] \tag{4.3}$$

where $(\xi_t)_t$ is the state price process. We first rewrite the constraint in equation (4.2) as

$$\mathbb{E}[X_T A_T \mid A_T < a_0] = \frac{\mathbb{E}[X_T A_T \mathbb{1}_{\{A_T < a_0\}}]}{\mathbb{P}(A_T < a_0)} = \theta_0.$$

Note, that since we know the distribution of A_T, we also know $\mathbb{P}(A_T < a_0)$. So we have the constraint

$$\mathbb{E}[X_T A_T \mathbb{1}_{\{A_T < a_0\}}] = \mathbb{P}(A_T < a_0)\theta_0 = b_0 \tag{4.4}$$

for some $b_0 \in \mathbb{R}^+$. So the optimization problem in equation (4.3) becomes

$$\min_{\left\{X_T \middle| \substack{X_T \sim F, \\ \xi_T \sim G, \\ A_T \sim H,\ A_T \geq 0\ a.s., \\ \mathbb{E}[X_T A_T \mathbb{1}_{\{A_T < a_0\}}] = b_0}\right\}} \mathbb{E}[\xi_T X_T]. \tag{4.5}$$

The bounds on the values of b_0 are given in the following lemma.

Lemma 4.2.1. *The bounds on $\mathbb{E}[X_T A_T \mathbb{1}_{\{A_T < a_0\}}] = b_0$ are given by*

$$l \triangleq \mathbb{E}[F^{-1}(U) H_{a_0}^{-1}(1 - U)] \leq b_0 \leq \mathbb{E}[F^{-1}(U) H_{a_0}^{-1}(U)] \triangleq u,$$

where F is the cdf of X_T, U is a uniform random variable and $H_{a_0}^{-1}$ is the quasi-inverse of the cdf of $A_T \mathbb{1}_{\{A_T < a_0\}}$ given by

$$H_{a_0}^{-1}(y) = H^{-1}(y - (1 - H(a_0)))\mathbb{1}_{\{y > 1 - H(a_0)\}}.$$

Consider the dual problem to the problem in equation (4.5)

$$\min_{\left\{X_T \middle| \substack{X_T \sim F, \\ \xi_T \sim G, \\ A_T \sim H,\ A_T \geq 0\ a.s.}\right\}} \mathbb{E}[\xi_T X_T] + \lambda \mathbb{E}[X_T A_T \mathbb{1}_{\{A_T < a_0\}}]. \tag{4.6}$$

The objective function can be rewritten as

$$\mathbb{E}[X_T(\xi_T + \lambda A_T \mathbb{1}_{\{A_T < a_0\}})]. \tag{4.7}$$

Combining equation (4.7) and Lemma 1.3.1 we get the following theorem.

Theorem 4.2.1. *The optimal solution to the problem in equation (4.6) is*

$$X_T^*(\lambda^*) = F^{-1}(1 - L_{\xi_T + \lambda^* A_T \mathbb{1}_{\{A_T < a_0\}}}(\xi_T + \lambda^* A_T \mathbb{1}_{\{A_T < a_0\}})), \tag{4.8}$$

where F is the target cdf, ξ_T is the state-price, A_T is the benchmark and λ^* is chosen such that $\mathbb{E}[X_T^*(\lambda^*)A_T\mathbb{1}_{\{A_T<a_0\}}] = b_0$.

Note that since ξ_T is continuously distributed $\xi_T + \lambda A_T\mathbb{1}_{\{A_T<a_0\}}$ is continuously distributed as well. Therefore, $L_{\xi_T+\lambda A_T\mathbb{1}_{\{A_T<a_0\}}}(\xi_T + \lambda A_T\mathbb{1}_{\{A_T<a_0\}})$ is in fact uniformly distributed.

Lemma 4.2.2. *$\mathbb{E}[X_T^*(\lambda)A_T\mathbb{1}_{\{A_T<a_0\}}]$ is a continuous function in λ for $\lambda \geq 0$. Furthermore, for any $b_0 \in (l, \mathbb{E}[F^{-1}(1-G(\xi_T))A_T\mathbb{1}_{\{A_T<a_0\}}])$, there exists $\lambda^* \in [0, +\infty)$ such that $\mathbb{E}[X_T^*(\lambda^*)A_T\mathbb{1}_{\{A_T<a_0\}}] = b_0$.*

Corollary 4.2.1. *If there exists a strictly decreasing and continuous function h such that $A_T = h(\xi_T)$, then for any $b_0 \in (l, u]$, there exists $\lambda^* \in [0, +\infty)$ such that $\mathbb{E}[X_T^*(\lambda^*)A_T\mathbb{1}_{\{A_T<a_0\}}] = b_0$.*

Finally, the following proposition shows that the optimal solution to the dual problem in equation (4.6) is in fact an optimal solution to the problem in equation (4.5).

Proposition 4.2.1. *For λ^* chosen such that $\mathbb{E}[X_T^*(\lambda^*)A_T\mathbb{1}_{\{A_T<a_0\}}] = b_0 \in (l, \mathbb{E}[F^{-1}(1-G(\xi_T))A_T\mathbb{1}_{\{A_T<a_0\}}]]$, $X_T^*(\lambda^*)$ is the optimal solution to the problem in equation (4.5).*

Corollary 4.2.2. *If there exists a strictly decreasing and continuous function h such that $A_T = h(\xi_T)$, then for λ^* chosen such that $\mathbb{E}[X_T^*(\lambda^*)A_T\mathbb{1}_{\{A_T<a_0\}}] = b_0 \in (l, u]$, $X_T^*(\lambda^*)$ is the optimal solution to the problem in equation (4.3).*

4.3 Growth Optimal Portfolio as Benchmark

In this section we present an application of the derived optimal payoff X_T^* in Theorem 4.2.1. We consider the case where there is only one source of uncertainty in the sense that $\xi_T + \lambda A_T\mathbb{1}_{\{A_T<a_0\}}$ is driven only by one random variable. We set A_T to be the Growth Optimal Portfolio (GOP). The GOP is a portfolio that will outperform any other strictly positive portfolio when given enough time, it can be interpreted as a major market index, see Platen and Heath [2006]. Thus, our goal is to solve the following problem

$$\min_{\left\{X_T \,\middle|\, \substack{X_T\sim F_{S_T^*},\\ \xi_T\sim G,\\ S_T^*\sim F_{S_T^*},\ S_T^*\geq 0 \ a.s.,\\ \mathbb{E}[X_T S_T^* | S_T^*<a_0]=\theta_0}\right\}} \mathbb{E}[\xi_T X_T].$$

Recall that the constraint $\mathbb{E}[X_T S_T^* \mid S_T^* < a_0] = \theta_0$ can be rewritten as $\mathbb{E}[X_T S_T^* \mathbb{1}_{\{S_T^* < a_0\}}] = \theta_0 \mathbb{P}(S_T^* < a_0) = b_0$, where we set a_0 to be the α quantile q_α as in equation (1.6). In order to apply the solution for the optimal payoff as in Theorem 4.2.1 we first need to find the cdf of $\xi_T + \lambda S_T^* \mathbb{1}_{\{S_T^* < a_0\}}$.

Lemma 4.3.1. *The cdf of $\xi_T + \lambda S_T^* \mathbb{1}_{\{S_T^* < a_0\}}$ is given by*

$$L_{\xi_T + \lambda S_T^* \mathbb{1}_{\{S_T^* < a_0\}}}(y) = \begin{cases} \Phi\left(\frac{\log(y) + (\mu_* - \sigma_*^2/2)T}{\sigma_* \sqrt{T}}\right) & \text{if } y < \frac{S_0^*}{a_0}, \\ \Phi\left(\frac{\log(\frac{S_0^*}{a_0}) + (\mu_* - \sigma_*^2/2)T}{\sigma_* \sqrt{T}}\right) & \text{if } \frac{S_0^*}{a_0} \le y < 2\sqrt{\lambda S_0^*}, \\ h(y) & \text{if } 2\sqrt{\lambda S_0^*} \le y < \frac{S_0^*}{a_0} + \lambda a_0, \\ \Phi\left(\frac{-x_\lambda^{MIN} + (\mu_* - \sigma_*^2/2)T}{\sigma_* \sqrt{T}}\right) & \text{if } y \ge \frac{S_0^*}{a_0} + \lambda a_0, \end{cases} \tag{4.9}$$

with

$$h(y) = \left(\Phi\left(\frac{x_\lambda^{MAX} - (\mu_* - \sigma_*^2/2)T}{\sigma_* \sqrt{T}}\right) - \Phi\left(\frac{x_\lambda^{MIN} - (\mu_* - \sigma_*^2/2)T}{\sigma_* \sqrt{T}}\right)\right) \mathbb{1}_{\{\lambda > S_0^*/a_0^2\}} + \Phi\left(\frac{\log(\frac{S_0^*}{a_0}) + (\mu_* - \sigma_*^2/2)T}{\sigma_* \sqrt{T}}\right),$$

where x_λ^{MIN} and x_λ^{MAX} are the smaller and bigger root of the equation $e^{-x} + \lambda S_0^* e^x = y$ respectively and $\lambda \ge 0$.

Figure 4.1 presents the relationship between λ and the conditional expectation constraint $\mathbb{E}[X_T^*(\lambda) S_T^* \mathbb{1}_{\{S_T^* < a_0\}}]$. This is a graphical illustration of Lemma 4.2.2, in other words for any feasible b_0 it is possible to find a λ^* such that $\mathbb{E}[X_T^*(\lambda^*) S_T^* \mathbb{1}_{\{S_T^* < a_0\}}] = b_0$.

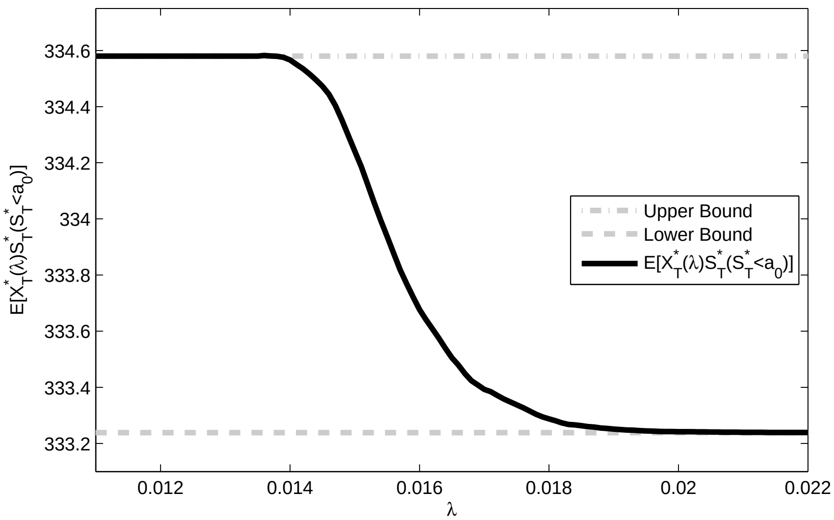


Figure 4.1: Graph of $\mathbb{E}[X_T^*(\lambda)S_T^*\mathbb{1}_{\{S_T^*<a_0\}}]$ against λ. We have used $N = 10{,}000{,}000$ simulations.

4.3.1 Link with Conditional Correlation Constraint

So far we have considered the constraint $\mathbb{E}[X_T^*(\lambda^*)S_T^*\mathbb{1}_{\{S_T^*<a_0\}}] = b_0$. We would like to directly consider a conditional correlation constraint. The correlation constraint can be written as

$$
\begin{aligned}
Corr(S_T^*, X_T \mid S_T^* < a_0) &= \rho_0 && (4.10)\\
&= \frac{Cov(S_T^*, X_T \mid S_T^* < a_0)}{\sqrt{Var(S_T^* \mid S_T^* < a_0)Var(X_T \mid S_T^* < a_0)}} && (4.11)\\
&= \frac{\mathbb{E}(S_T^* X_T \mid S_T^* < a_0) - \mathbb{E}[S_T^* \mid S_T^* < a_0]\mathbb{E}[X_T \mid S_T^* < a_0]}{\sqrt{Var(S_T^* \mid S_T^* < a_0)Var(X_T \mid S_T^* < a_0)}} && (4.12)
\end{aligned}
$$

where $\rho_0 \in [-1, 1]$. In contrast to the global constraint case in Chapter 3 there is no functional relationship between the conditional expectation constraint and the conditional correlation constraint due to the unknown values $\mathbb{E}[X_T \mid S_T^* < a_0]$ and $Var(X_T \mid S_T^* < a_0)$. To overcome this issue one can empirically find a relationship between b_0 and the tail correlation. This relationship is represented in Figure 4.2.

Now we know that it is feasible using a computer to construct a strategy with desired

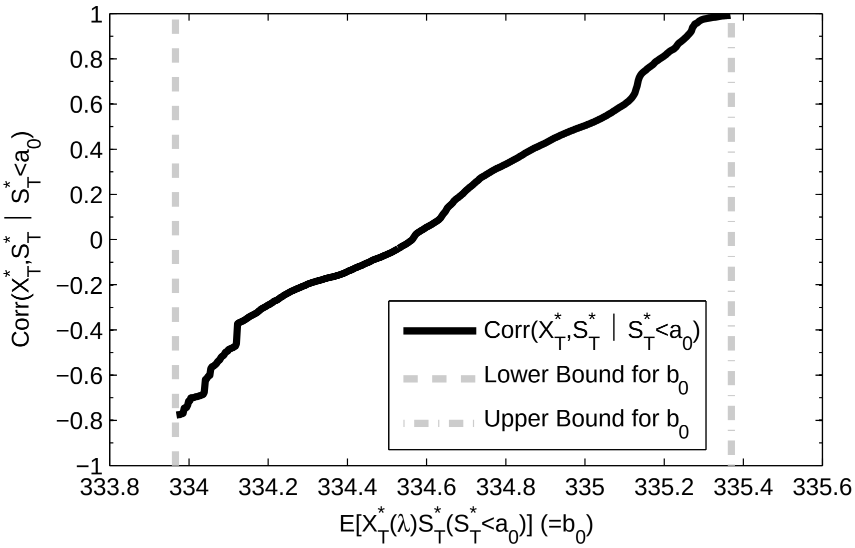


Figure 4.2: Graph of tail correlation $Corr(X_T^*, S_T^* \mid S_T^* < a_0)$ against conditional expectation constraint $\mathbb{E}[X_T^*(\lambda) S_T^* \mathbb{1}_{\{S_T^* < a_0\}}](= b_0)$. We have used $N = 100{,}000$ simulations for $n = 500$ values of b_0 between the two bounds.

conditional correlation with the GOP. The following proposition describes the desired strategy.

Proposition 4.3.1. ***(Strategy 4)*** *A cost-efficient strategy with the same distribution of final wealth as the the GOP but such that $Corr(S_T^*, X_T^* \mid S_T^* < a_0) = \rho_0$ is given by*

$$X_T^*(\lambda^*) = F^{-1}(1 - L_{\xi_T + \lambda^* S_T^* \mathbb{1}_{\{S_T^* < a_0\}}}(\xi_T + \lambda^* S_T^* \mathbb{1}_{\{S_T^* < a_0\}}))$$

where F is the target cdf, ξ_T is the state-price, S_T^* is the GOP, $L_{\xi_T + \lambda^* S_T^* \mathbb{1}_{\{S_T^* < a_0\}}}$ is defined in Lemma 4.3.1 and λ^* is chosen such that $Corr(S_T^*, X_T^*(\lambda^*) \mid S_T^* < a_0) = \rho_0$.

In order to construct such a payoff in practice, one has to follow these steps:

1. Find the bounds for b_0 as in Lemma 4.2.1
2. Construct two vectors: vector of b_0 and the vector with corresponding conditional correlations
3. Use linear interpolation to find the b_0 which produces the desired conditional correlation ρ_0

4. Once b_0 is found, find λ^* such that the conditional expectation constraint is satisfied

This λ^* defines the desired payoff in Proposition 4.3.1. The payoffs constructed by applying Proposition 4.3.1 with different correlation constraint and threshold can be found in Figure 4.3.

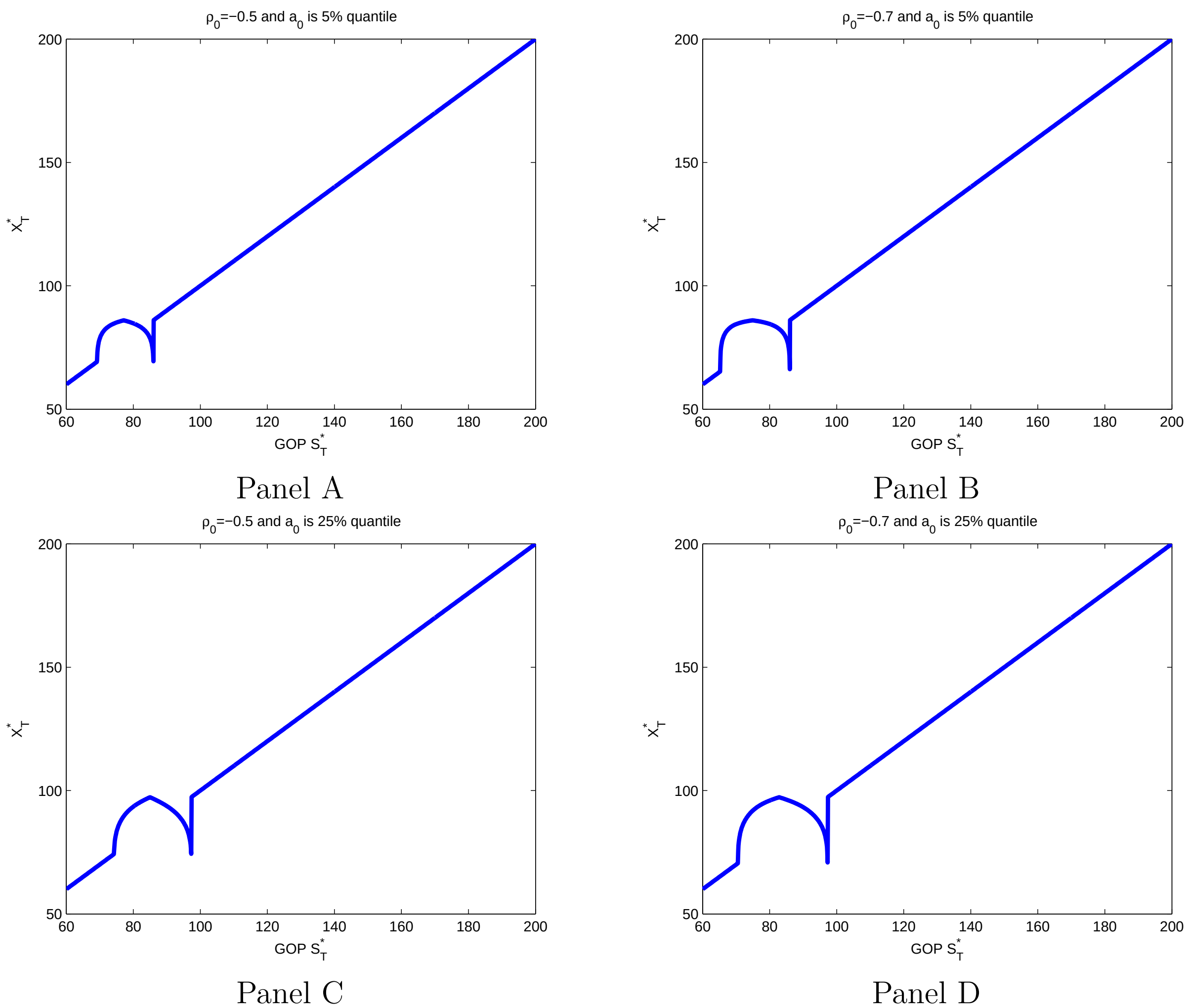
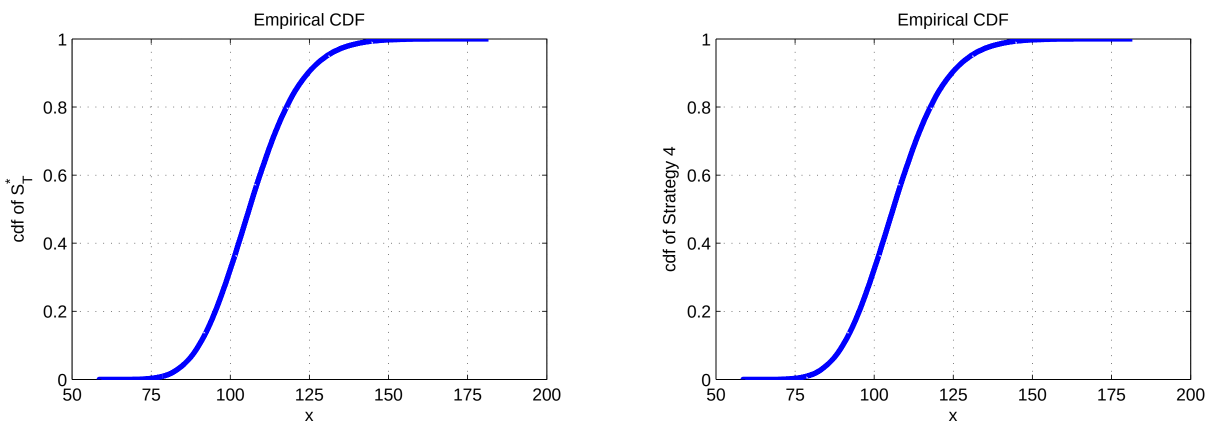


Figure 4.3: Payoffs of Strategy 4 as described in Proposition 4.3.1 for different settings of the conditional correlation $Corr(S_T^*, X_T^* \mid S_T^* < a_0) = \rho_0$ and the threshold a_0 with $T = 1$.

In order to check the correctness of the payoff in Figure 4.3 Panel A we consider its tail correlation and the cdf. The tail correlation is easily calculated to be $\rho_0 = -0.4994$. In

Figure 4.4 the cdf of the GOP and Strategy 4 with $\rho_0 = -0.5$ and $a_0 = q_{0.05}$ are presented. Additionally, in Table 4.1 we provide the corresponding statistical values.



Panel A Panel B

Figure 4.4: Panel A shows the empirical cdf of the GOP. Panel B shows the empirical cdf of Strategy 4 with $\rho_0 = -0.5$, a_0 being the 5% quantile and $T = 1$.

Table 4.1: Statistics for the comparison of the target cdf and the cdf of Strategy 4

Statistic	GOP	Strategy 4
Minimum	58.6013	58.6013
Maximum	181.5824	181.5824
Mean	106.7963	106.7959
Median	105.9733	105.9733
Std. dev.	13.5522	13.5532

Chapter 5

Optimal Strategies with State-Dependent Copula Constraints

5.1 Cost-Efficient Strategies with State-Dependent Copula Constraints

In this chapter we examine a different approach to achieve protection in bad states of the economy. Instead of using correlation to describe the dependence between the payoff of the strategy and the market, one can also describe the dependence using copulas. This type of constraint when constructing cost-efficient strategies has been extensively studied in Bernard et al. [2013b] and Bernard et al. [2013a].

In this section we present a general representation of a cost-efficient strategy with a dependence in the tail, described through a copula, with the Growth Optimal Portfolio (GOP). In the following subsections we cover special cases in order to compare the performance and the cost of those strategies with the strategy in Chapter 4. The comparison can be found in Chapter 6. For the proofs please refer to the appendix of Bernard et al. [2013a].

The following theorem provides the cheapest strategy with the desired distribution of final wealth as well as the desired dependence with the market in a crisis regime. This type of strategy is called constrained cost-efficient strategy, see Bernard et al. [2013b].

Theorem 5.1.1. *(Optimal Strategies under a Crisis Regime) Let F be the desired cdf of terminal payoff X_T. We further assume that Q is the desired joint distribution of the final* *payoff X_T with the GOP S_T^* when there is a financial crisis as specified in equation (1.6), such that*

$$\mathbb{P}(S_T^* \leq s, X_T \leq x \mid S_T^* \leq q_\alpha) = Q(s, x). \tag{5.1}$$

Then, an optimal strategy X_T^* is given by

$$X_T^* = \begin{cases} F^{-1}\left(h(F_{S_T^*}(S_T^*) - \alpha)\right) & \text{when } S_T^* > q_\alpha \\ F^{-1}\left(q\left(1 - F_{S_T^*}(S_T^*), j_{F_{S_T^*}(S_T^*)}(F_{Z_T}(Z_T))\right)\right) & \text{when } S_T^* \leq q_\alpha \end{cases}, \tag{5.2}$$

where Z_T is any random variable s.t. (S_T^*, Z_T) is continuously distributed, hence there exists a unique copula $J(\cdot,\cdot)$ with $\mathbb{P}(S_T^* \leq s, Z_T \leq x) = J(F_{S_T^*}(s), F_{Z_T}(x))$. We denote by $j_u(v)$ its first partial derivative and $h(x)$, $q(u, v)$ are defined as

$$h(x) = \inf\{c \mid c - C^*(\alpha, c) \geq x\},$$

$$q(u, v) = c_u^{-1}(v), \text{ where } c_u(v) = \frac{\partial C^*}{\partial u}(1 - u, v),$$

where the copula C^* is determined through Q.

The strength of Theorem 5.1.1 is the explicit form of payoff in equation (5.2). To better understand the structure of the payoff in equation (5.2) consider first the case $S_T^* > q_\alpha$: It is straightforward to verify that $F^{-1}\left(h(F_{S_T^*}(S_T^*) - \alpha)\right)$ is non-decreasing in the GOP S_T^*, and therefore by Proposition 1.3.1 cost-efficient. For the case $S_T^* \leq q_\alpha$ we wish to have a certain dependence with the market as specified by equation (5.1). To this end recall the conditional distribution method used to generate pairs of random variables with a desired copula, see [Nelsen, 2006, p. 41] for details on the method. First, the function $j_u(v)$ is applied to make $j_{F_{S_T^*}(S_T^*)}(F_{Z_T}(Z_T))$ independent of $1 - F_{S_T^*}(S_T^*)$. Then the function $q(u, v) = c_u^{-1}(v)$ is applied to generate the desired dependence C^* between $q\left(1 - F_{S_T^*}(S_T^*), j_{F_{S_T^*}(S_T^*)}(F_{Z_T}(Z_T))\right)$ and $F_{S_T^*}(S_T^*)$. Therefore the desired constraint (5.1) is satisfied.

In the following sections we apply Theorem 5.1.1 to construct explicit expressions for optimal strategies with different types of dependence in the tail. To simplify the applications we present the strategies in a two-dimensional Black-Scholes market as described in Section 1.2.2. For more general results, please refer to Bernard et al. [2013a].

5.1.1 Gaussian Dependence in the Tail

A Gaussian dependence provides a flexible way of dealing with diversification. It is especially useful when one seeks to achieve a negative correlation with the market when there is a crisis.

Corollary 5.1.1. ***(Strategy 5)*** *In a two-dimensional Black-Scholes market, the cheapest path-independent strategy with a cdf F but such that its tail dependence with S_T^* when $S_T^* \leq q_\alpha$ is prescribed by the Gaussian copula*

$$C^*(u, v) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v), \rho),$$

where Φ_2 is the bivariate cdf, Φ is the cdf of $N(0,1)$ and ρ is the correlation factor, can be constructed as

$$X_T^* = \begin{cases} F^{-1}\left(h(F_{S_T^*}(S_T^*) - \alpha)\right) & \text{when } S_T^* > q_\alpha \\ F^{-1}\left(\Phi(A\sqrt{1-\rho^2} + B\rho)\right) & \text{when } S_T^* \leq q_\alpha \end{cases} \tag{5.3}$$

where

$$A = \frac{\log\left(\frac{S_T^1}{S_0^1}\right) - \left(\mu_1 - \frac{\sigma_1^2}{2}\right)T}{\sigma_1\sqrt{\left(\frac{(1-\rho_{12})}{2}\right)T}} - B\sqrt{\frac{1+\rho_{12}}{1-\rho_{12}}}, \text{ and} \tag{5.4}$$

$$B = \frac{\frac{1}{\sigma_1}\left[\log\left(\frac{S_T^1}{S_0^1}\right) - \left(\mu_1 - \frac{\sigma_1^2}{2}\right)T\right]}{\sqrt{2(1+\rho_{12})T}} + \frac{\frac{1}{\sigma_2}\left[\log\left(\frac{S_T^2}{S_0^2}\right) - \left(\mu_2 - \frac{\sigma_2^2}{2}\right)T\right]}{\sqrt{2(1+\rho_{12})T}}, \tag{5.5}$$

and where h is defined implicitly as

$$h(x) = \inf\{c | c - C^*(\alpha, c) \geq x\}.$$

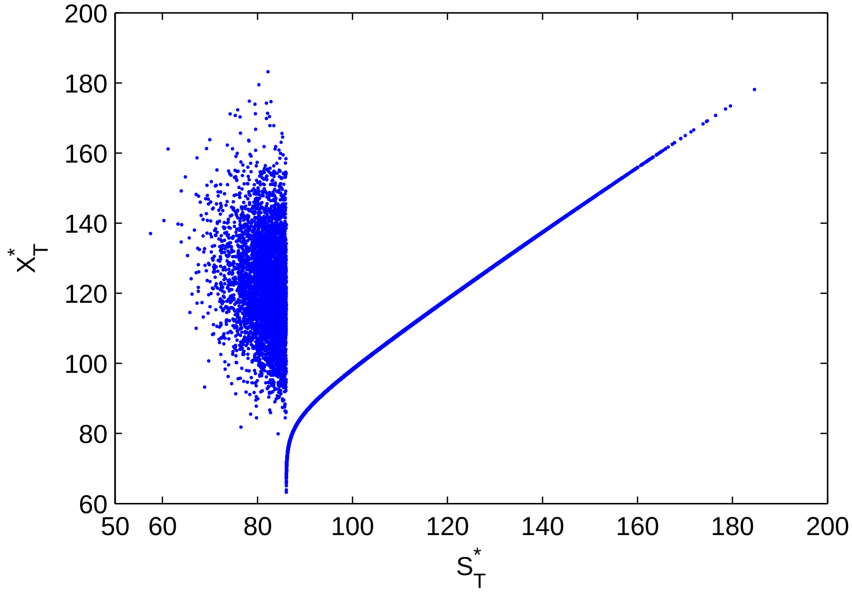


Figure 5.1: Payoff of Strategy 5 as described in Corollary 5.1.1. We choose the threshold to be $q_{0.05}$ and use the parameters in Table 1.2.

5.1.2 Clayton Dependence in the Tail

If one seeks to have a Clayton copula dependence in the tail, it is important to note that it always exhibits positive correlation. To achieve the desired negative correlation in the tail one needs to specify a flipped Clayton copula between the final payoff X_T and the GOP S^*_T. Recall the inverse relationship between ξ_T and S^*_T from equation (1.4). This means that a negative dependence between S^*_T and X_T can be specified via a positive dependence between ξ_T and X_T.

Corollary 5.1.2. ***(Strategy 6)*** *Let C^* be the Clayton copula given as*

$$C^*(u,v) = (u^{-a} + v^{-a} - 1)^{-1/a}.$$

Then the cheapest strategy in a two-dimensional Black-Scholes market with a cdf F but such that its tail dependence with S^*_T when $S^*_T \leq q_\alpha$, is prescribed by the flipped Clayton copula with

$$\forall s \in [0, q_\alpha],\ \ y \in \mathbb{R}, \mathbb{P}(S^*_T \leq s, X_T \leq y) = F(y) - C^*(1 - F_{S^*_T}(s), F(y)),$$

can be constructed as

$$X_T^* = \begin{cases} F^{-1}\left([x^{-a} - (1-\alpha)^{-a} + 1]^{-1/a}\right) & \text{when } S_T^* > q_\alpha \\ F^{-1}\left(\left[u^{-a}(v^{-a/(1+a)} - 1) + 1\right]^{-1/a}\right) & \text{when } S_T^* \leq q_\alpha \end{cases} \tag{5.6}$$

where A and B are defined as in (5.4) and (5.5), $x = \Phi(B) - \alpha$, $u = 1 - x - \alpha$ and $v = \Phi(A)$.

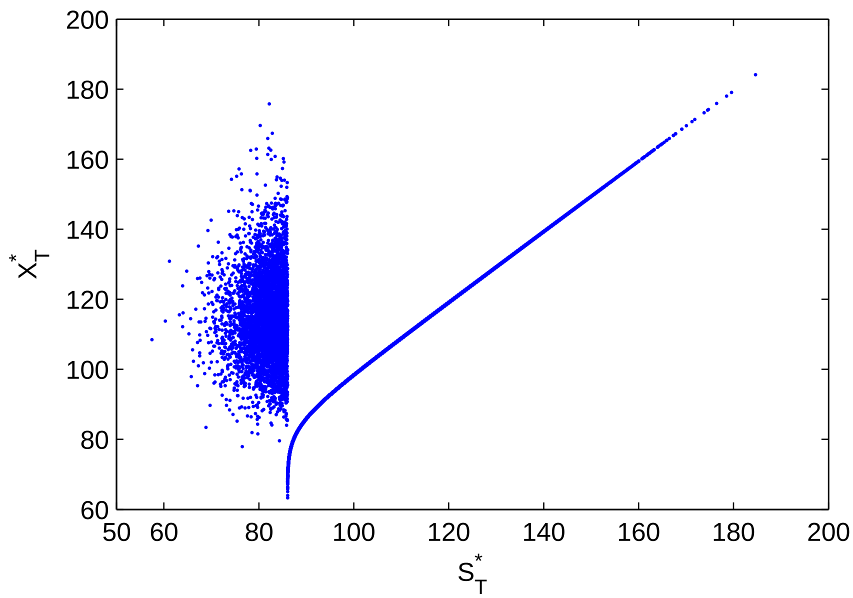


Figure 5.2: Payoff of Strategy 6 as described in Corollary 5.1.2. We choose the threshold to be $q_{0.05}$ and use the parameters in Table 1.2.

Chapter 6

Comparison of the Strategies

6.1 Empirical Results

This chapter presents our empirical study of the comparison of the before mentioned strategies. An overview over the strategies is provided below.

Strategy 1: Investment in the GOP (Proposition 2.2.2). This strategy aims at achieving the same distribution of final wealth as the GOP at minimal cost. The values in Table 6.1 are calculated using the setting in Section 1.2.1.

Strategy 2: Left Truncated Gaussian distribution of log-returns (Proposition 2.2.3). This strategy aims at achieving an insurance level of -5% by specifying the desired distribution of log-returns to be the Left Truncated Gaussian distribution with truncation level -5% and having the same mean and volatility of log-returns as the GOP. The values in Table 6.1 are calculated using the setting in Section 1.2.1.

Strategy 3: Global correlation constraint (Proposition 3.2.1). Strategy 3 is a constraint strategy with the same distribution of final wealth as the GOP but that satisfies a global correlation constraint with the GOP $Corr(S_T^*, X_T^*) = 0.5$. The values in Table 6.1 are calculated using the setting in Section 1.2.1.

Strategy 4 (4a, 4b): Correlation constraint in the tail (Proposition 4.3.1). Strategy 4 is a constraint strategy with the same distribution of final wealth as the GOP but that satisfies a conditional correlation constraint with the GOP $Corr(S_T^*, X_T^* \mid S_T^* < q_{0.05}) = -0.5$ $(-0.17; 0.03)$, where $q_{0.05}$ is a 5% quantile. The values in Table 6.1 are calculated using the setting in Section 1.2.1.

Strategy 5: Gaussian dependence in the tail (Corollary 5.1.1). Strategy 5 is a constraint strategy with the same distribution of final wealth as the GOP but that satisfies a dependence constraint in a crisis, as described in equation (1.6), described through the Gaussian copula with the correlation coefficient $\rho = -0.5$. The values in Table 6.1 are calculated using the setting in Section 1.2.2.

Strategy 6: Clayton dependence in the tail (Corollary 5.1.2). Strategy 6 is a constraint strategy with the same distribution of final wealth as the GOP but that satisfies a dependence constraint in a crisis, as in equation (1.6), described through the flipped Clayton copula with the parameter $a = 1$. The parameter a is chosen such that the Kendall's τ of the Clayton copula is equal to the Kendall's τ of the Gaussian copula with $\rho = 0.5$. The values in Table 6.1 are calculated using the setting in Section 1.2.2.

In Table 6.1 we present the results of our empirical analysis of the discussed strategies. We consider the cost of a strategy, its Sharpe ratio and several conditional probabilities to figure out how the strategies compete in different market situations. To calculate the cost we use Definition 1.2.1. The Sharpe ratio is calculated as

$$\frac{E[X_T] - X_0 e^{rT}}{std(X_T)},$$

where X_0 is the cost of the strategy. As for the conditional probabilities, we consider $\mathbb{P}(A|C)$, $\mathbb{P}(A|D)$ and $\mathbb{P}(B|C)$ where $C = \{S_T^* < q_\alpha\}$ represents a market crisis, $D = \{S_T^* < S_0^* e^{rT}\}$ represents a decrease in the market and $A = \{X_T/X_0 < e^{rT}\}$, $B = \{X_T/X_0 < 75\% e^{rT}\}$ are the two events of interest.

We have used same random numbers for all the strategies from $N = 100,000$ simulations. Furthermore, we have matched the parameters in the one-dimensional and the two-dimensional markets to have the same parameters for the GOP.

6.2 Observations and Analysis

This section focuses on the analysis of Table 6.1. First of all note that we have constructed all the strategies such that they have the same distribution of final wealth. Hence $\mathbb{E}[X_T^1] = \mathbb{E}[X_T^i]$ and $Var(X_T^1) = Var(X_T^i)$ for all $i = 2, ..., 6$, where X_T^i stands for the payoff of strategy i. We observe a decrease in Sharpe ratio for all the strategies X_T^i, $i = 2, ..., 6$ in comparison to Strategy 1 (investment in the GOP). This is due to the efficiency loss when adding constraints to the strategies.

Table 6.1: Cost, Sharpe ratio and conditional probabilities with $A = \{X_T/X_0 < e^{rT}\}$, $B = \{X_T/X_0 < 75\% e^{rT}\}$, $C = \{S_T^* < q_\alpha\}$, and $D = \{S_T^* < S_0^* e^{rT}\}$. By tail correlation we mean $Corr(X_T^*, S_T^* \mid S_T^* < a_0)$.

	Tail Corr	T	Cost	Sharpe	$\mathbb{P}(A\|C)$	$\mathbb{P}(A\|D)$	$\mathbb{P}(B\|C)$
1		1	100.00	0.1287	100%	100%	18.7%
		5	100.00	0.2689	100%	100%	100%
2		1	104.51	0.1249	100%	100%	0%
		5	106.59	0.2588	100%	100%	100%
3		1	100.73	0.0732	100%	83%	20.7%
		5	102.30	0.1935	100%	95%	100%
4	-0.5	1	100.0130	0.1275	100%	100%	19.4%
		5	100.07	0.2666	100%	100%	100%
4*a*	-0.17	1	100.0090	0.1278	100%	100%	19.3%
		5	100.06	0.2672	100%	100%	100%
4*b*	0.03	1	100.0086	0.1280	100%	100%	19.2%
		5	100.05	0.2675	100%	100%	100%
5	-0.17	1	100.54	0.0871	11%	91%	0%
		5	103.23	0.1645	11%	90%	1%
6	0.03	1	100.39	0.0990	23%	92%	0%
		5	102.22	0.1971	23%	91%	2%

Our main concern is the comparison between Strategy 4 ($4a$, $4b$) and Strategies 5 and 6. To match the tail correlation generated by Strategies 5 and 6 we have introduced two more cases: Strategy $4a$ and Strategy $4b$. As was expected, the cost for satisfying a certain correlation constraint in the tail is lower than the cost of satisfying a certain copula constraint in the tail with the same correlation. The reason behind this is that a copula provides a much more precise description of the desired dependence whereas just a correlation constraint may be achieved by many different copulas. Thus it must be cheaper to produce tail correlation with some benchmark rather than utilizing copulas.

We can also observe a loss in Sharpe ratio when utilizing copulas to describe the tail dependence. The Sharpe ratio for the Strategies 4, $4a$, and $4b$ is almost the same as for Strategy 1 suggesting that there is not much inefficiency. The Sharpe ratio is obtained through a stand-alone evaluation and is therefore not a good measure when it comes to the question how a strategy performs in certain states of the economy, e.g. a financial crisis.

To this end, we consider conditional probabilities in the last three columns of Table 6.1 which condition on two events: A crisis and a decrease of the market. In both cases we are interested in the level of protection we can expect from a strategy. The first observation is that we do not get a significantly better protection with Strategy 4 than the straightforward Strategy 1. On the other hand, Strategies 5 and 6 provide significant protection in a crisis situation as seen in $\mathbb{P}(A|C)$ and $\mathbb{P}(B|C)$ columns.

6.3 Conclusions and Future Directions

We conclude that cost-efficient strategies with conditional correlation constraints are in fact cheaper than cost-efficient strategies with state-dependent copula constraints exhibiting the same tail correlation. However, they do not provide adequate protection in bad states of the economy making them not suitable for an investor seeking protection in a market crisis situation. This means also that when one analyzes a strategy and finds out that the strategy exhibits a negative correlation with the market in the lower tail, this does not mean that it provides sufficient protection in a crisis situation.

Further research can be done on the sensitivity to model risk. It would be interesting to know whether correlation constraints are less sensitive to model risk than copula constraints. Copulas are often said to be very hard to measure and to capture. Thus, one might expect a lot of model risk and a difficulty to obtain the desired copula by replication. It might be easier and more practical to target a correlation constraint instead of targeting a copula constraint. Regarding the replication, one could extend Chapter 2 on payoff replication to a more general case with more than one underlying.

APPENDICES

Appendix A

Useful Identities

Lemma A.0.1. *In the one-dimensional Black-Scholes setting as in Section 1.2.1 the state-price ξ_T can be written as an explicit function of the stock price S_T as follows*

$$\xi_T = \alpha_T \left(\frac{S_T}{S_0}\right)^{-\beta}, \tag{A.1}$$

where $\alpha_T = \exp\left(\frac{\theta}{\sigma}\left(\mu_1 - \frac{\sigma_1^2}{2}\right)T - \left(r + \frac{\theta^2}{2}\right)T\right)$, $\beta = \frac{\theta}{\sigma_1}$, *and* $\theta = \frac{\mu_1 - r}{\sigma_1}$.

Proof. See equation (8) on page 8 in Bernard et al. [2013b]. □

Lemma A.0.2. *In the two-dimensional Black-Scholes setting as in Section 1.2.2 the GOP S_T^* can be expressed explicitly in terms of S_T^1 and S_T^2 as*

$$S_T^* = S_0^* \left(\frac{S_T^1}{S_0^1}\right)^{\pi_1^*} \left(\frac{S_T^2}{S_0^2}\right)^{\pi_2^*} e^{(br+M)T},$$

where $b = 1 - \pi_1^* - \pi_2^*$, $M = -\frac{\hat{\lambda}}{1+\rho_{12}}[\hat{\lambda} - \frac{\sigma_1}{2} - \frac{\sigma_2}{2}]$, $\hat{\lambda} = \frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2}$ *and*

$$\pi_1^* = \frac{\hat{\lambda}}{(1+\rho_{12})\sigma_1}, \quad \pi_2^* = \frac{\hat{\lambda}}{(1+\rho_{12})\sigma_2}.$$

Proof. See A.5 on page 15 in Bernard et al. [2013a]. □

Lemma A.0.3. *For a random pair* (X,Y) *the following statement is true*

$$(X,Y) \text{ is anti-monotonic} \Leftrightarrow (X,Y) \sim (F_X^{-1}(U), F_Y^{-1}(1-U)) \tag{A.2}$$

where F_X *and* F_Y *are the respective distributions for* X *and* Y *and* U *is a standard uniform random variable.*

Proof. For the "$\Leftarrow$" direction, let $A = (F_X^{-1}(u), F_Y^{-1}(1-u))$ with $u \in [0,1]$. Consider now $(x_1,y_1),(x_2,y_2) \in A$. Define $(x_1,y_1) = (F_X^{-1}(u_1), F_Y^{-1}(1-u_1))$ and $(x_2,y_2) = (F_X^{-1}(u_2), F_Y^{-1}(1-u_2))$ with $u_1 \neq u_2$ and without loss of generality assume $u_1 < u_2$. Then

$$(x_1-x_2)(y_1-y_2) = (F_X^{-1}(u_1) - F_X^{-1}(u_2))(F_Y^{-1}(1-u_1) - F_Y^{-1}(1-u_2)).$$

Since F_X^{-1} and F_Y^{-1} are increasing functions, it follows that $(x_1-x_2)(y_1-y_2) \leq 0$ for all $(x_1,y_1),(x_2,y_2) \in A$. Hence, A is anti-monotonic and we also have $\mathbb{P}((X,Y) \in A) = 1$, thus (X,Y) is in accordance to Definition 1.3.2 anti-monotonic.

For the "$\Rightarrow$" direction, given $(x,y) \in \mathbb{R}^2$ we have

$$\begin{aligned}\mathbb{P}(F_X^{-1}(U) \leq x, F_Y^{-1}(1-U) \leq y) &= \mathbb{P}(U \leq F_X(x), 1-F_Y(y) \leq U) \\ &= \mathbb{P}(1-F_Y(y) \leq U \leq F_X(x)),\end{aligned}$$

since U is standard uniform distributed, we get

$$\mathbb{P}(F_X^{-1}(U) \leq x, F_Y^{-1}(1-U) \leq y) = \max\{F_X(x) + F_Y(y) - 1, 0\}. \tag{A.3}$$

In the next step we will show that an anti-monotonic pair (X,Y) has the same distribution as $(F_X^{-1}(U), F_Y^{-1}(1-U))$. Denote by A its support.

Define $A_1 = \{(x_1,y_1) \in A \mid x_1 \leq x\}$ and $A_2 = \{(x_2,y_2) \in A \mid y_2 > y\}$. Then $A_1 \subseteq A_2$ or $A_2 \subseteq A_1$ must hold. Otherwise, there would exist (a,b) and (c,d) in A such that $(a,b) \in A_1 \setminus A_2$ and $(c,d) \in A_2 \setminus A_1$, i.e., $a \leq x, b \leq y, c > x$ and $d > y$. Hence $(a-c)(b-d) > 0$, which is a contradiction with the assumed anti-monotonicity for A. Note that $\{X \leq x, Y > y\} = \{(X,Y) \in A_1 \cap A_2\}$.

If $A_1 \subseteq A_2$ then

$$\mathbb{P}(X \leq x, Y > y) = \mathbb{P}((X,Y) \in A_1) = \mathbb{P}(X \leq x) = F_X(x).$$

If $A_2 \subseteq A_1$ then

$$\mathbb{P}(X \leq x, Y > y) = \mathbb{P}((X,Y) \in A_2) = \mathbb{P}(Y > y) = 1 - F_Y(y).$$

Thus, $\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x) - \mathbb{P}(X \leq x, Y > y) = F_X(x) - \mathbb{P}(X \leq x, Y > y)$ is equal to equation (A.3). $\square$

Appendix B

Proofs

Proof of Lemma 1.3.1 on page 7.

Assume that X and Y are anti-monotonic. From the Fréchet-Hoeffding bounds in Section 1.3, one can get $\mathbb{P}(X \leq x, Y \leq y) \geq \max\{F_X(x) + F_Y(y) - 1, 0\}$. Since X and Y are anti-monotonic we know from Lemma A.0.3 that $(X,Y) \sim (F_X^{-1}(U), F_Y^{-1}(1-U))$ and from equation (A.3) that $\mathbb{P}(F_X^{-1}(U) \leq x, F_Y^{-1}(1-U) \leq y) = \max\{F_X(x) + F_Y(y) - 1, 0\}$. Therefore, $\mathbb{P}(X \leq x, Y \leq y)$ is minimal.

Furthermore, for fixed distribution of X and Y, $\mathbb{P}(X \geq x, Y \geq y)$ is minimal if and only if $\mathbb{P}(X \leq x, Y \leq y)$ is minimal. To see this point, consider the following steps

$$\mathbb{P}(X \geq x, Y \geq y) + \mathbb{P}(X \geq x, Y \leq y) = \mathbb{P}(X \geq x) = \text{const}$$

and

$$\mathbb{P}(X \leq x, Y \leq y) + \mathbb{P}(X \geq x, Y \leq y) = \mathbb{P}(Y \leq y) = \text{const}.$$

Therefore, both are equal to $(\text{const} - \mathbb{P}(X \geq x, Y \leq y))$ and if the one is minimal, the other one has to be minimal as well.

Write $X = X^+ - X^-$, where $X^+ = \max(X, 0)$ and $X^- = -\min(X, 0)$. Similarly, we write $Y = Y^+ - Y^-$. Thus,

$$\begin{aligned}
\mathbb{E}[XY] =& \mathbb{E}[(X^+ - X^-)(Y^+ - Y^-)] \\
=& \mathbb{E}[X^+Y^+] + \mathbb{E}[X^-Y^-] - \mathbb{E}[X^+Y^-] - \mathbb{E}[X^-Y^+] \\
=& \int_0^{+\infty}\int_0^{+\infty} \mathbb{P}(X^+ > x, Y^+ > y)\, dx\, dy + \int_0^{+\infty}\int_0^{+\infty} \mathbb{P}(X^- > x, Y^- > y)\, dx\, dy \\
&- \int_0^{+\infty}\int_0^{+\infty} \mathbb{P}(X^+ > x, Y^- > y)\, dx\, dy - \int_0^{+\infty}\int_0^{+\infty} \mathbb{P}(X^- > x, Y^+ > y)\, dx\, dy \\
=& \int_0^{+\infty}\int_0^{+\infty} \mathbb{P}(X > x, Y > y)\, dx\, dy + \int_{-\infty}^{0}\int_{-\infty}^{0} \mathbb{P}(X \le x, Y \le y)\, dx\, dy \\
&+ \int_{-\infty}^{0}\int_0^{+\infty} (\mathbb{P}(X > x, Y > y) - \mathbb{P}(X > x))\, dx\, dy \\
&+ \int_0^{+\infty}\int_{-\infty}^{0} (\mathbb{P}(X > x, Y > y) - \mathbb{P}(Y > y))\, dx\, dy,
\end{aligned}$$

where for the last equality we have used $X^+ > x \Leftrightarrow X > x$ and $X^- > x \Leftrightarrow X < -x$ for any $x > 0$. Since $\mathbb{P}(X \le x, Y \le y)$ and $\mathbb{P}(X > x, Y > y)$ are both minimal when X and Y are anti-monotonic, $\mathbb{E}[XY]$ is also minimal.

□

Proof of Proposition 1.3.1 on page 7.

The proof is a straightforward combination of Definition 1.2.1 and Lemma 1.3.1. □

Proof of Proposition 2.2.1 on page 10.

From Corollary 1.3.1 we already know the form of the unique cost-efficient strategy with the desired distribution F:

$$X_T^* = F^{-1}(1 - F_{\xi_T}(\xi_T)).$$

First we would like to rewrite it in terms of S_T. From Lemma A.0.1 we see that there exists a function f such that $\xi_T = f(S_T)$. We assume $\mu_1 > r$, which is a reasonable assumption, then $\beta > 0$ and thus f is a strictly decreasing continuous function. Note that f is invertible, then for all $x \in \mathbb{R}^+$,

$$\begin{aligned}
F_{\xi_T}(x) &= \mathbb{P}(\xi_T \le x) = \mathbb{P}(f(S_T) \le x) \\
&= \mathbb{P}(S_T > f^{-1}(x)) = 1 - \mathbb{P}(S_T \le f^{-1}(x)) \\
&= 1 - F_{S_T}(f^{-1}(x)).
\end{aligned}$$

Therefore, we have $X_T^* = F^{-1}(F_{S_T}(S_T))$. Since $\log(x/S_0)$ is increasing in x, we have $F_{S_T}(S_T) = F_{\text{Under}}(\log(S_T/S_0))$, where F_{Under} is the cdf of $\log(S_T/S_0)$. If we set

$$S_0 e^{Y_T^*} = X_T^* = F^{-1}(F_{\text{Under}}(\log(S_T/S_0))),$$

then

$$Y_T^* = F_{\text{Target}}^{-1}(F_{\text{Under}}(\log(S_T/S_0))),$$

where F_{Target}^{-1} is the inverse cdf of $\log(S_T^*/S_0^*)$. □

Proof of Proposition 3.1.1 on page 18.

From Lemma 1.3.1, it suffices to show that $(X_T^*(\lambda), \xi_T + \lambda A_T)$ is an anti-monotonic pair. From Lemma A.0.3 we know an equivalent statement for anti-monotonicity:

$$(X, Y) \text{ is anti-monotonic} \Leftrightarrow (X, Y) \sim (F_X^{-1}(U), F_Y^{-1}(1-U)).$$

In our case with $X_T^*(\lambda) = F^{-1}(1 - L_{\xi_T+\lambda A_T}(\xi_T + \lambda A_T))$ we get

$$(X_T^*(\lambda), \xi_T + \lambda A_T) \sim (F^{-1}(1-U), L_{\xi_T+\lambda A_T}^{-1}(U)).$$

Hence, $(X_T^*(\lambda), \xi_T + \lambda A_T)$ is an anti-monotonic pair. □

Proof of Lemma 3.1.1 on page 18.

(i) Continuity of $\mathbb{E}[X_T^*(\lambda)A_T]$ in λ:
From equation (3.6) we see that X_T^* is a random variable in A_T and ξ_T, we can therefore write

$$\mathbb{E}[X_T^*(\lambda)A_T] = \int_0^{+\infty}\int_0^{+\infty} X_T^*(\lambda) A_T f_{A_T,\xi_T}(A_T, \xi_T)\, dA_T\, d\xi_T, \tag{B.1}$$

where f_{A_T,ξ_T} is the joint pdf of A_T and ξ_T. By Theorem 3.5.1 in Cheng S. [2008], it is sufficient to show that $X_T^*(\lambda)$ is continuous and integrable, i.e. that there exists an integrable function g independent of λ such that $|X_T^*(\lambda)| \leq g$. First we note that since F is continuous, $X_T^*(\lambda)$ is continuous in λ. We are left to show the integrability of $X_T^*(\lambda)$.To this end consider the following inequalities

$$\begin{aligned} &\mathbb{P}(\xi_T + \lambda A_T \le \lambda y) \le \mathbb{P}(\xi_T + \lambda A_T \le x + \lambda y) \le \mathbb{P}(\lambda A_T \le x + \lambda y) \\ &\Leftrightarrow \ \mathbb{P}(\frac{\xi_T}{\lambda} + A_T \le y) \le L_{\xi_T + \lambda A_T}(x + \lambda y) \le \mathbb{P}(A_T \le y + \frac{x}{\lambda}). \end{aligned} \tag{B.2}$$

Applying the squeeze theorem on inequality B.2, we have $\lim_{\lambda \to \infty} X_T^*(\lambda) = F^{-1}(1 - H(A_T))$ pointwise. Fix some $\epsilon > 0$, then from the definition of limit $\exists\, \lambda_1$ s.t. $\forall\, \lambda > \lambda_1$, we have

$$|X_T^*(\lambda) - F^{-1}(1 - H(A_T))| < \epsilon,$$

therefore $|X_T^*(\lambda)| < |F^{-1}(1 - H(A_T))| + \epsilon$. For $\lambda \in [0, \lambda_1]$, by extreme value theorem we have that $X_T^*(\lambda_0) \le X_T^*(\lambda) \le X_T^*(\lambda_2)$ for some λ_0, $\lambda_2 \in [0, \lambda_1]$. Therefore, for all $\lambda \ge 0$, we have

$$|X_T^*(\lambda)| \le \max\{|X_T^*(\lambda_0)|, |X_T^*(\lambda_2)|, |F^{-1}(1 - H(A_T))| + \epsilon\}.$$

Hence, $X_T^*(\lambda)$ is bounded by an integrable function independent of λ.

(ii) Existence of λ:

$$\begin{aligned} \lim_{\lambda \searrow 0} \mathbb{E}[X_T^*(\lambda) A_T] &= \mathbb{E}[F^{-1}(1 - G(\xi_T)) A_T] \\ \lim_{\lambda \to +\infty} \mathbb{E}[X_T^*(\lambda) A_T] &= \mathbb{E}[F^{-1}(1 - H(A_T)) A_T] = l. \end{aligned}$$

Since $\mathbb{E}[X_T^*(\lambda) A_T]$ is a continuous function, by the intermediate value theorem, $\exists\, \lambda^* \in [0, +\infty)$ s.t. $\mathbb{E}[X_T^*(\lambda^*) A_T] = a_0$ for any $a_0 \in (l, \mathbb{E}[F^{-1}(1 - G(\xi_T)) A_T]]$.

□

Proof of Corollary 3.1.1 on page 19.

We focus on the upper bound for a_0 and show that

$$E[F^{-1}(1 - G(\xi_T)) A_T] = \mathbb{E}[F^{-1}(H(A_T)) A_T] = u.$$

Recall that the cdf of A_T is H and that of ξ_T is G. If there exists a strictly decreasing and continuous function h such that $A_T = h(\xi_T)$, then we have

$$\begin{aligned} H(x) &= \mathbb{P}(A_T \leq x) \\ &= \mathbb{P}(h(\xi_T) \leq x) \\ &= \mathbb{P}(\xi_T > h^{-1}(x)) \\ &= 1 - G(h^{-1}(x)). \end{aligned}$$

Therefore, $H(A_T) = 1 - G(\xi_T)$. Combining this with Lemma 3.1.1, one gets the desired result. □

Proof of Proposition 3.1.2 on page 19.

We show that if $X_T^*(\lambda^*)$ minimizes the objective function of the problem in equation (3.4) with λ^* chosen such that $\mathbb{E}[A_T X_T^*(\lambda^*)] = a_0$ it also minimizes the objective function of problem in equation (3.3). Proposition 3.1.1 implies

$$\begin{aligned} \mathbb{E}[\xi_T X_T^*(\lambda^*)] + \lambda^* \mathbb{E}[A_T X_T^*(\lambda^*)] &\leq \mathbb{E}[\xi_T X_T(\lambda)] + \lambda \mathbb{E}[A_T X_T(\lambda)] \\ \mathbb{E}[\xi_T X_T^*(\lambda^*)] + \lambda^* a_0 &\leq \mathbb{E}[\xi_T X_T(\lambda)] + \lambda a_0 \\ \mathbb{E}[\xi_T X_T^*(\lambda^*)] &\leq \mathbb{E}[\xi_T X_T(\lambda)] \end{aligned}$$

The second inequality is due to the fact that for all feasible solutions X_T of the problem in equation (3.3), we have $\mathbb{E}[A_T X_T] = a_0$. Thus, $X_T^*(\lambda^*)$ clearly solves the problem in equation (3.3). □

Proof of Corollary 3.1.2 on page 19.

It follows directly from Proposition 3.1.2 and Corollary 3.1.1. □

Proof of Lemma 3.2.1 on page 19.

From Proposition 1.2.1 we have $\xi_T = \frac{S_0^*}{S_T^*}$, therefore

$$\begin{aligned} \xi_T + \lambda S_T^* &= e^{-(\mu_* - \sigma_*^2/2)T - \sigma_* W_T^*} + \lambda S_0^* e^{(\mu_* - \sigma_*^2/2)T + \sigma_* W_T^*} \\ &\stackrel{d}{=} e^{-X} + \lambda S_0^* e^{X}, \end{aligned}$$

with $X \sim N((\mu_* - \sigma_*^2/2)T, \sigma_*^2 T)$. Now set $g(x) = e^{-x} + \lambda S_0^* e^{x}$. The function $g(x)$ is strictly convex and attains its minimum at $x = \frac{1}{2}\log(\frac{1}{\lambda S_0^*})$, with the value $2\sqrt{\lambda S_0^*}$. Therefore, for any $y < 2\sqrt{\lambda S_0^*}$ we get $L_{\xi_T+\lambda S_T^*}(y) = 0$. Otherwise, by the convexity of the function g, we can always find X_λ^{MAX} and X_λ^{MIN} such that

$$\begin{aligned} L_{\xi_T+\lambda S_T^*}(y) &= \mathbb{P}(g(X) \leq y) \\ &= \mathbb{P}(X_\lambda^{MIN} \leq X \leq X_\lambda^{MAX}) \\ &= \Phi\left(\frac{x_\lambda^{MAX} - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) - \Phi\left(\frac{x_\lambda^{MIN} - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right). \end{aligned}$$

□

Proof of Proposition 3.2.1 on page 20.

This follows immediately from Proposition 3.1.1, Lemma 3.2.1 and equation (1.3). □

Proof of Lemma 3.3.1 on page 21.

We use the moment matching technique to find a proxy for the cdf of $\xi_T + \lambda S_T^1$. Although a sum of two log-normal random variables is not log-normal we match the moments with a log-normal random variable. Let Z be standard normally distributed, then we would like to find μ_p and σ_p such that the following is satisfied

$$\begin{aligned} \mathbb{E}[e^{\mu_p+\sigma_p Z}] &= \mathbb{E}[\xi_T + \lambda S_T^1] \\ Var(e^{\mu_p+\sigma_p Z}) &= Var(\xi_T + \lambda S_T^1). \end{aligned}$$

On the left hand side we have $\mathbb{E}[e^{\mu_p+\sigma_p Z}] = e^{\mu_p+\sigma_p^2/2}$ and $Var(e^{\mu_p+\sigma_p Z}) = (e^{\sigma_p^2} - 1)e^{2\mu_p+\sigma_p^2}$. And on the right hand side $\mathbb{E}[\xi_T + \lambda S_T^1] = \mathbb{E}[\xi_T] + \lambda\mathbb{E}[S_T^1]$ and $Var(\xi_T + \lambda S_T^1) = Var(\xi_T) + \lambda^2 Var(S_T^1) + 2\lambda Cov(\xi_T, S_T^1)$. After we match the moments we get the following parameters for the proxy

$$\begin{aligned} \mu_p(\lambda) &= -\frac{1}{2}\log\left(\left(Var(\xi_T + \lambda S_T^1) + \mathbb{E}[\xi_T + \lambda S_T^1]^2\right)\mathbb{E}[\xi_T + \lambda S_T^1]^{-4}\right), \\ \sigma_p(\lambda) &= \sqrt{2\log(\mathbb{E}[\xi_T + \lambda S_T^1]) - 2\mu_p(\lambda)}. \end{aligned}$$

The cdf $L_{\xi_T+\lambda S_T^1}$ is a log-normal cdf with parameters $\mu_p(\lambda)$ and $\sigma_p(\lambda)$.

In what is to follow we present the results of calculations needed to code the proxy parameters in a computer program. We use Proposition 1.2.1 and Lemma A.0.2 to get a representation of ξ_T. After some calculations one gets the following form for ξ_T and S_T^1

$$\begin{aligned}\xi_T =& e^{m_1+s_1Z_1}e^{m_2+s_2Z_2}\\ S_T^1 =& S_0^1 e^{m_3+s_3Z_1}\end{aligned}$$

where $Z_1, Z_2 \sim N(0,1)$ and $Z_1 \perp\!\!\!\perp Z_2$ and

$$\begin{aligned}m_1 &= [-(br+M) - \pi_1^*(\mu_1 - \sigma_1^2/2)]T,\\ s_1 &= [-\pi_1^*\sigma_1 - \pi_2^*\sigma_2\rho_{12}]\sqrt{T},\\ m_2 &= [-\pi_2^*(\mu_2 - \sigma_2^2/2)]T,\\ s_2 &= [-\pi_2^*\sigma_2\sqrt{1-\rho_{12}^2}]\sqrt{T},\\ m_3 &= (\mu_1 - \sigma_1^2/2)T,\\ s_3 &= \sigma_1\sqrt{T}.\end{aligned}$$

With this notation we get

$$\begin{aligned}\mathbb{E}[S_T^1] &= S_0^1 e^{m_3+s_3^2/2},\\ Var(S_T^1) &= (S_0^1)^2(e^{s_3^2}-1)e^{2m_3+s_3^2},\\ \mathbb{E}[\xi_T] &= e^{m_1+m_2+(s_1^2+s_2^2)/2},\\ Var(\xi_T) &= e^{2(m_1+m_2+s_1^2+s_2^2)} - e^{2(m_1+m_2+(s_1^2+s_2^2)/2)},\\ \mathbb{E}[\xi_T S_T^1] &= S_0^1 e^{m_1+m_3+(s_1+s_3)^2/2}e^{m_2+s_2^2/2},\\ Cov(\xi_T, S_T^1) &= \mathbb{E}[\xi_T S_T^1] - \mathbb{E}[\xi_T]\mathbb{E}[S_T^1].\end{aligned}$$

The expressions above can be used for coding. □

Proof of Proposition 3.3.1 on page 22.

The proof is a straight application of Proposition 3.1.1 using the result in Lemma 3.3.1 and equation (1.3). □

Proof of Lemma 4.2.1 on page 26.

The bounds are derived by a straightforward application of the Fréchet-Hoeffding bounds in equation (1.8). What remains to show is the derivation of the quasi-inverse of the cdf of $A_T \mathbb{1}_{\{A_T<a_0\}}$. The cdf of $A_T \mathbb{1}_{\{A_T<a_0\}}$ is given in equation (B.6) as

$$H_{a_0}(x) = \mathbb{P}(A_T \mathbb{1}_{\{A_T<a_0\}} < x) = \begin{cases} H(x) + (1 - H(a_0)) & \text{if } x < a_0, \\ 1 & \text{if } x \geq a_0. \end{cases}$$

Its quasi-inverse is given by

$$\begin{aligned} H_{a_0}^{-1}(y) &= \inf\{x \mid H_{a_0}(x) \geq y\} \\ &= \begin{cases} H^{-1}(y - (1 - H(a_0))) & \text{if } 1 - H(a_0) < y \leq 1, \\ 0 & \text{if } y \leq 1 - H(a_0), \end{cases} \\ &= H^{-1}(y - (1 - H(a_0)))\mathbb{1}_{\{y>1-H(a_0)\}}, \end{aligned}$$

where H is the cdf of A_T. □

Proof of Lemma 4.2.2 on page 27.

(i) Continuity of $\mathbb{E}[A_T X_T^*(\lambda)\mathbb{1}_{\{A_T<a_0\}}]$ in λ:
From equation (4.8) we see that X_T^* is a random variable in A_T and ξ_T, we can therefore write

$$\mathbb{E}[X_T^*(\lambda)A_T\mathbb{1}_{\{A_T<a_0\}}] = \int_0^{+\infty}\int_0^{a_0} X_T^*(\lambda)A_T f_{A_T,\xi_T}(A_T,\xi_T)dA_T d\xi_T, \tag{B.3}$$

where f_{A_T,ξ_T} is the joint pdf of A_T and ξ_T. By Theorem 3.5.1 in Cheng S. [2008], it is sufficient to show that $X_T^*(\lambda)$ is continuous and integrable. The proof is very similar to the proof of Lemma 3.1.1.

(ii) Existence of λ:

$$\lim_{\lambda\searrow 0}\mathbb{E}[X_T^*(\lambda)A_T\mathbb{1}_{\{A_T<a_0\}}] = \mathbb{E}[F^{-1}(1-G(\xi_T))A_T\mathbb{1}_{\{A_T<a_0\}}] \tag{B.4}$$

$$\lim_{\lambda\to+\infty}\mathbb{E}[X_T^*(\lambda)A_T\mathbb{1}_{\{A_T<a_0\}}] = \mathbb{E}[F^{-1}(1-H_{a_0}(A_T\mathbb{1}_{\{A_T<a_0\}}))A_T\mathbb{1}_{\{A_T<a_0\}}] = l, \tag{B.5}$$

where H_{a_0} is the cdf of $A_T\mathbb{1}_{\{A_T<a_0\}}$ given by

$$H_{a_0}(x) = \mathbb{P}(A_T \mathbb{1}_{\{A_T<a_0\}} < x) = \begin{cases} H(x) + (1 - H(a_0)) & \text{if } x < a_0, \\ 1 & \text{if } x \geq a_0, \end{cases} \tag{B.6}$$

where we recall that H is the cdf of A_T. To justify the value of $H_{a_0}(x)$ when $x < a_0$ in equation (B.6), consider the following calculation. By the law of total probability we have

$$\begin{aligned}\mathbb{P}(A_T \mathbb{1}_{\{A_T<a_0\}} < x) &= \mathbb{P}(A_T \mathbb{1}_{\{A_T<a_0\}} < x; A_T < x) \\ &\quad + \mathbb{P}(A_T \mathbb{1}_{\{A_T<a_0\}} < x; x \leq A_T < a_0) \\ &\quad + \mathbb{P}(A_T \mathbb{1}_{\{A_T<a_0\}} < x; A_T \geq a_0) \\ &= \mathbb{P}(A_T < x) + \mathbb{P}(A_T < x; x \leq A_T < a_0) + \mathbb{P}(A_T > a_0) \\ &= H(x) + 0 + (1 - H(a_0)).\end{aligned}$$

Since $\mathbb{E}[X_T^*(\lambda) A_T \mathbb{1}_{\{A_T<a_0\}}]$ is a continuous function, by the intermediate value theorem, $\exists \lambda^* \in [0, +\infty)$ s.t.

$$\mathbb{E}[X_T^*(\lambda) A_T \mathbb{1}_{\{A_T<a_0\}}] = b_0 \tag{B.7}$$

for $b_0 \in (l, \mathbb{E}[F^{-1}(1 - G(\xi_T)) A_T \mathbb{1}_{\{A_T<a_0\}}]) \subset [l, u]$. The upper bound for b_0 is from equation (B.4).

□

Proof of Corollary 4.2.1 on page 27.

We focus on the upper bound for b_0 and show that

$$E[F^{-1}(1 - G(\xi_T)) A_T \mathbb{1}_{\{A_T<a_0\}}] = \mathbb{E}[F^{-1}(H(A_T)) A_T \mathbb{1}_{\{A_T<a_0\}}] = u,$$

where u is given in Lemma 4.2.1. Recall that the cdf of A_T is H and that of ξ_T is G. If there exists a strictly decreasing and continuous function h such that $A_T = h(\xi_T)$, then we have

$$\begin{aligned}H(x) &= \mathbb{P}(A_T \leq x) \\ &= \mathbb{P}(h(\xi_T) \leq x) \\ &= \mathbb{P}(\xi_T > h^{-1}(x)) \\ &= 1 - G(h^{-1}(x)).\end{aligned}$$

Therefore, $H(A_T) = 1 - G(\xi_T)$. Combining this with Lemma 4.2.2 one gets the desired result.

□

Proof of Proposition 4.2.1 on page 27.

We show that if $X_T^*(\lambda^*)$ minimizes the objective function of the problem in equation (4.6) with λ^* chosen such that $\mathbb{E}[X_T^*(\lambda^*)A_T\mathbb{1}_{\{A_T<a_0\}}] = b_0$ it also minimizes the objective function of problem in equation (4.3). Proposition 4.2.1 implies

$$\begin{aligned}\mathbb{E}[\xi_T X_T^*(\lambda^*)] + \lambda^*\mathbb{E}[X_T^*(\lambda^*)A_T\mathbb{1}_{\{A_T<a_0\}}] &\leq \mathbb{E}[\xi_T X_T(\lambda)] + \lambda\mathbb{E}[X_T(\lambda)A_T\mathbb{1}_{\{A_T<a_0\}}]\\ \mathbb{E}[\xi_T X_T^*(\lambda^*)] + \lambda^* b_0 &\leq \mathbb{E}[\xi_T X_T(\lambda)] + \lambda b_0\\ \mathbb{E}[\xi_T X_T^*(\lambda^*)] &\leq \mathbb{E}[\xi_T X_T(\lambda)]\end{aligned}$$

The second inequality is due to the fact that for all feasible solutions X_T of the problem in equation (4.3), we have $\mathbb{E}[X_T A_T\mathbb{1}_{\{A_T<a_0\}}] = b_0$. Thus, $X_T^*(\lambda^*)$ clearly solves the problem in equation (4.3).

□

Proof of Corollary 4.2.2 on page 27.

The statement follows from Proposition 4.2.1 and Corollary 4.2.1.

□

Proof of Lemma 4.3.1 on page 28.

From Proposition 1.2.1 we have $\xi_T = \frac{S_0^*}{S_T^*}$, therefore

$$\begin{aligned}\xi_T + \lambda S_T^*\mathbb{1}_{\{S_T^*<a_0\}} &= e^{-(\mu_*-\sigma_*^2/2)T-\sigma_* W_T^*}\\ &\quad + \lambda S_0^* e^{(\mu_*-\sigma_*^2/2)T+\sigma_* W_T^*}\mathbb{1}_{\{S_0^* e^{(\mu_*-\sigma_*^2/2)T+\sigma_* W_T^*}<a_0\}}\\ &\overset{d}{=} e^{-X} + \lambda S_0^* e^X \mathbb{1}_{S_0^* e^X<a_0},\end{aligned} \tag{B.8}$$

with $X \sim N((\mu_* - \sigma_*^2/2)T, \sigma_*^2 T)$. Now set $g(x) = e^{-x} + \lambda S_0^* e^x \mathbb{1}_{S_0^* e^x<a_0}$, then

$$\begin{aligned}g(x) &= \begin{cases} e^{-x} + \lambda S_0^* e^x & \text{if } S_0^* e^x < a_0,\\ e^{-x} & \text{if } S_0^* e^x \geq a_0,\end{cases}\\ &= \begin{cases} e^{-x} + \lambda S_0^* e^x & \text{if } x < \log(\frac{a_0}{S_0^*}),\\ e^{-x} & \text{if } x \geq \log(\frac{a_0}{S_0^*}).\end{cases}\end{aligned}$$

In order to find the cdf of $\xi_T + \lambda S_T^* \mathbb{1}_{\{S_T^* < a_0\}}$ we need to analyze

$$L_{\xi_T + \lambda S_T^* \mathbb{1}_{\{S_T^* < a_0\}}}(y) = \mathbb{P}(g(X) \leq y), \tag{B.9}$$

where X is distributed as in equation (B.8). Figure B.1 shows the graph of $g(x)$ with $A = \frac{1}{2}\log(\frac{1}{\lambda S_0^*})$, $B = \log(\frac{a_0}{S_0^*})$, $C = \frac{S_0^*}{a_0}$, $D = 2\sqrt{\lambda S_0^*}$ and $E = \frac{S_0^*}{a_0} + \lambda a_0$.

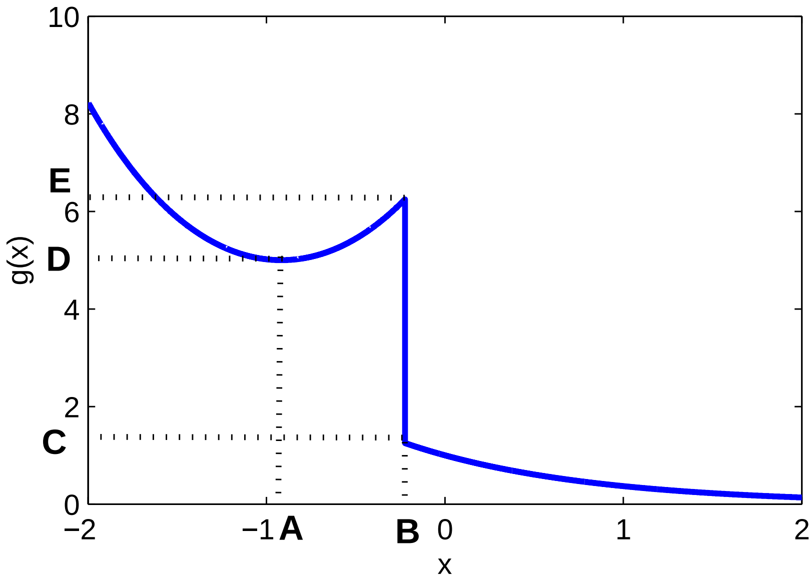


Figure B.1: Function g(x) when $\lambda > \frac{S_0^*}{a_0^2}$

In order to analyze equation (B.9) we separate two cases:

i $A < B \;\Leftrightarrow\; \frac{1}{2}\log(\frac{1}{\lambda S_0^*}) < \log(\frac{a_0}{S_0^*}) \;\Leftrightarrow\; \lambda > \frac{S_0^*}{a_0^2}$

ii $A \geq B \;\Leftrightarrow\; \frac{1}{2}\log(\frac{1}{\lambda S_0^*}) \geq \log(\frac{a_0}{S_0^*}) \;\Leftrightarrow\; \lambda \geq \frac{S_0^*}{a_0^2}$

Consider first the case (i), i.e. we assume that $\lambda > \frac{S_0^*}{a_0^2}$ holds. Recall again that we are looking for $\mathbb{P}(g(X) \leq y)$, therefore

- If $y < C = \frac{S_0^*}{a_0}$, then

$$\begin{aligned}\mathbb{P}(g(X) \le y) &= \mathbb{P}(e^{-X} \le y) = \mathbb{P}(X \ge -\log(y)) \\ &= 1 - \Phi\left(\frac{-\log(y) - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) \\ &= \Phi\left(\frac{\log(y) + (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) \end{aligned} \tag{B.10}$$

- If $\frac{S_0^*}{a_0} = C \le y < D = 2\sqrt{\lambda S_0^*}$, then

$$\begin{aligned}\mathbb{P}(g(X) \le y) &= \mathbb{P}\left(g(X) \le \frac{S_0^*}{a_0}\right) = \mathbb{P}\left(X \ge -\log\left(\frac{S_0^*}{a_0}\right)\right) \\ &= 1 - \Phi\left(\frac{-\log(\frac{S_0^*}{a_0}) - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) \\ &= \Phi\left(\frac{\log(\frac{S_0^*}{a_0}) + (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) \end{aligned} \tag{B.11}$$

- If $2\sqrt{\lambda S_0^*} = D \le y < E = \frac{S_0^*}{a_0} + \lambda a_0$, then by convexity of $e^{-x} + \lambda S_0^* e^x$ we can always find x_λ^{MIN} and x_λ^{MAX}, such that

$$\begin{aligned}\mathbb{P}(g(X) \le y) &= \mathbb{P}(x_\lambda^{MIN} \le X \le x_\lambda^{MAX}) + \mathbb{P}\left(X \ge -\log\left(\frac{S_0^*}{a_0}\right)\right) \\ &= \Phi\left(\frac{x_\lambda^{MAX} - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) - \Phi\left(\frac{x_\lambda^{MIN} - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) \\ &+ 1 - \Phi\left(\frac{-\log(\frac{S_0^*}{a_0}) - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) \\ &= \Phi\left(\frac{x_\lambda^{MAX} - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) - \Phi\left(\frac{x_\lambda^{MIN} - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) \\ &+ \Phi\left(\frac{\log(\frac{S_0^*}{a_0}) + (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) \end{aligned} \tag{B.12}$$

- If $y \geq E = \frac{S_0^*}{a_0} + \lambda a_0$, then there always exists x_λ^{MIN}, such that

$$\mathbb{P}(g(X) \leq y) = \mathbb{P}(X \leq x_\lambda^{MIN}) = 1 - \Phi\left(\frac{x_\lambda^{MIN} - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right)$$
$$= \Phi\left(\frac{-x_\lambda^{MIN} + (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right) \tag{B.13}$$

The case (ii) when $\lambda \geq \frac{S_0^*}{a_0^2}$ is very similar to case (i). The only difference is that for $\frac{S_0^*}{a_0} = C \leq y < E = \frac{S_0^*}{a_0} + \lambda a_0$ the cdf L is equal to

$$\mathbb{P}(g(X) \leq y) = 1 - \Phi\left(\frac{-\log(\frac{S_0^*}{a_0}) - (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right)$$
$$= \Phi\left(\frac{\log(\frac{S_0^*}{a_0}) + (\mu_* - \sigma_*^2/2)T}{\sigma_*\sqrt{T}}\right). \tag{B.14}$$

We therefore omit the full derivation. Combining the results in equations (B.10) to (B.14) we get the desired cdf $L_{\xi_T + \lambda S_T^* \mathbb{1}_{\{S_T^* < a_0\}}}$. □

Proof of Proposition 4.3.1 on page 30.

The statement is a combination of Theorem 4.2.1 and Lemma 4.3.1. □

References

Adrian, T. and Brunnermeier, M., (2010) CoVaR. Working paper available at http://ssrn.com/abstract=1269446

G. Amin, H.M. Kat (2003) Hedge fund performance 1990-2000: Do the "money machines" really add value. *Journal of Financial Quantitative Analysis*, 38(2), pp. 251-275.

Bernard C., Chen J.S., Vanduffel S. (2013) Optimal Portfolios under Worst Case Scenarios *Quantitative Finance*, DOI: 10.1080/14697688.2013.836282

Bernard, C. and Boyle, P. P. and Vanduffel, S. (2013) Explicit Representation of Cost-Efficient Strategies, *Finance, forthcoming*, Available at SSRN: http://ssrn.com/abstract=1561272.

Bernard, C. and Moraux, F. and Rueschendorf, L. and Vanduffel, S. (2013) Optimal Payoffs under State-dependent Constraints. *arXiv preprint arXiv:1308.6465.*

Chek Hin Choi (2012) Optimal Portfolio with Correlation Constraint, Undergraduate research report of a visiting scholar supervised by Carole Bernard.

Cheng S. (2008) A Crash Course on the Lebesgue Integral and Measure Theory, Available at http://www.gold-saucer.org/math/lebesgue/lebesgue.pdf.

P. Del Moral, B. Remillard, S. Rubenthaler (2006) Monte Carlo approximations of American options. *GERAD.*

Dybvig P.H. (1988) Distributional Analysis of Portfolio Choice *The Journal of Business*, Vol. 61, No. 3 (Jul., 1988), pp. 369-393.

Goldstein, D.G., Johnson, E.J. and Sharpe, W.F. (2008) Choosing outcomes versus choosing products: Consumer-focused retirement investment advice. *J. Consum. Res.*, 2008, 35(3), 440-456.

A. Hocquard, N. Papageorgiou and B. Remillard (2007) Replicating the properties of hedge fund returns *DGAM-HEC Alternative Investments Research*, Technical Report 1, 2007.

A. Hocquard, N. Papageorgiou and B. Remillard (2012) The payoff distribution model: an application to dynamic portfolio insurance *Quantitative Finance*, 2012, 1-14.

Nelsen, R. (2006) An Introduction to Copulas. *Springer Series in Statistics*, 2nd edition.

Platen, E. and Heath, D. (2006) A Benchmark Approach to Quantitative Finance *Number 13 in Springer finance*, 2006 (Springer: Berlin).

Schweizer, M. (1995) Variance-optimal hedging in discrete time. *Mathematics of Operations Research*, 20.1 1995: 1-32.