Advanced Course in Asset Management

Thierry Roncalli*

*University of Paris-Saclay

January 2021

Part 1. Portfolio Optimization

Part 2. Risk Budgeting

Part 3. Smart Beta, Factor Investing and ARP

Part 4. Green Finance, ESG and Climate Risk

Part 5. Machine Learning in Asset Management

General information

1. **Overview**
   The objective of this course is to understand the theoretical and practical aspects of asset management
2. **Prerequisites**
   M1 Finance or equivalent
3. **ECTS**
   3
4. **Keywords**
   Finance, Asset Management, Optimization, Statistics
5. **Hours**
   Lectures: 24h, HomeWork: 30h
6. **Evaluation**
   Project + oral examination
7. **Course website**
   `http://www.thierry-roncalli.com/RiskBasedAM.html`

Objective of the course

The objective of the course is twofold:

1. having a financial culture on asset management
2. being proficient in quantitative portfolio management

Class schedule

Course sessions

- January 8 (6 hours, AM+PM)
- January 15 (6 hours, AM+PM)
- January 22 (6 hours, AM+PM)
- January 29 (6 hours, AM+PM)

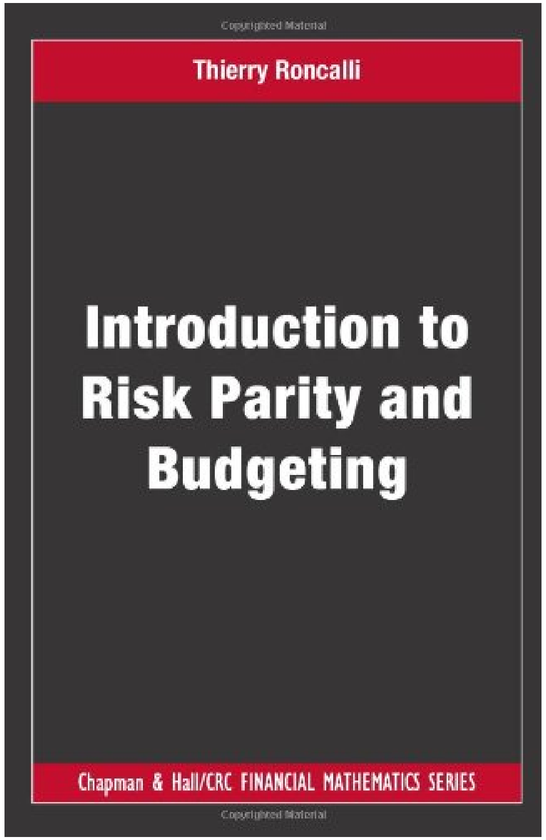
Class times: Fridays 9:00am-12:00pm, 1:00pm–4:00pm, University of Evry

Agenda

- Lecture 1: Portfolio Optimization
- Lecture 2: Risk Budgeting
- Lecture 3: Smart Beta, Factor Investing and Alternative Risk Premia
- Lecture 4: Green and Sustainable Finance, ESG Investing and Climate Risk
- Lecture 5: Machine Learning in Asset Management

Textbook

- Roncalli, T. (2013), *Introduction to Risk Parity and Budgeting*, Chapman & Hall/CRC Financial Mathematics Series.



Additional materials

- Slides, tutorial exercises and past exams can be downloaded at the following address:

  `http://www.thierry-roncalli.com/RiskBasedAM.html`

- Solutions of exercises can be found in the companion book, which can be downloaded at the following address:

  `http://www.thierry-roncalli.com/RiskParityBook.html`

Asset Management
Lecture 1. Portfolio Optimization

Thierry Roncalli*

*University of Paris-Saclay

January 2021

Agenda

- **Lecture 1: Portfolio Optimization**
- Lecture 2: Risk Budgeting
- Lecture 3: Smart Beta, Factor Investing and Alternative Risk Premia
- Lecture 4: Green and Sustainable Finance, ESG Investing and Climate Risk
- Lecture 5: Machine Learning in Asset Management

Notations

- We consider a universe of n assets
- $x = (x_1, \ldots, x_n)$ is the vector of weights in the portfolio
- The portfolio is fully invested:

$$\sum_{i=1}^{n} x_i = \mathbf{1}_n^\top x = 1$$

- $R = (R_1, \ldots, R_n)$ is the vector of asset returns where R_i is the return of asset i
- The return of the portfolio is equal to:

$$R(x) = \sum_{i=1}^{n} x_i R_i = x^\top R$$

- $\mu = \mathbb{E}[R]$ and $\Sigma = \mathbb{E}\left[(R-\mu)(R-\mu)^\top\right]$ are the vector of expected returns and the covariance matrix of asset returns

Computation of the first two moments

The expected return of the portfolio is:

$$\mu(x) = \mathbb{E}[R(x)] = \mathbb{E}\left[x^\top R\right] = x^\top \mathbb{E}[R] = x^\top \mu$$

whereas its variance is equal to:

$$\begin{aligned}
\sigma^2(x) &= \mathbb{E}\left[(R(x) - \mu(x))(R(x) - \mu(x))^\top\right] \\
&= \mathbb{E}\left[\left(x^\top R - x^\top \mu\right)\left(x^\top R - x^\top \mu\right)^\top\right] \\
&= \mathbb{E}\left[x^\top (R - \mu)(R - \mu)^\top x\right] \\
&= x^\top \mathbb{E}\left[(R - \mu)(R - \mu)^\top\right] x \\
&= x^\top \Sigma x
\end{aligned}$$

Efficient frontier

Two equivalent optimization problems

1. Maximizing the expected return of the portfolio under a volatility constraint (σ-**problem**):

$$\max \mu(x) \quad \text{u.c.} \quad \sigma(x) \leq \sigma^{\star}$$

2. Or minimizing the volatility of the portfolio under a return constraint (μ-**problem**):

$$\min \sigma(x) \quad \text{u.c.} \quad \mu(x) \geq \mu^{\star}$$

Efficient frontier

Example 1

We consider four assets. Their expected returns are equal to 5%, 6%, 8% and 6% while their volatilities are equal to 15%, 20%, 25% and 30%. The correlation matrix of asset returns is given by the following matrix:

$$C = \begin{pmatrix} 1.00 & & & \\ 0.10 & 1.00 & & \\ 0.40 & 0.70 & 1.00 & \\ 0.50 & 0.40 & 0.80 & 1.00 \end{pmatrix}$$

Efficient frontier

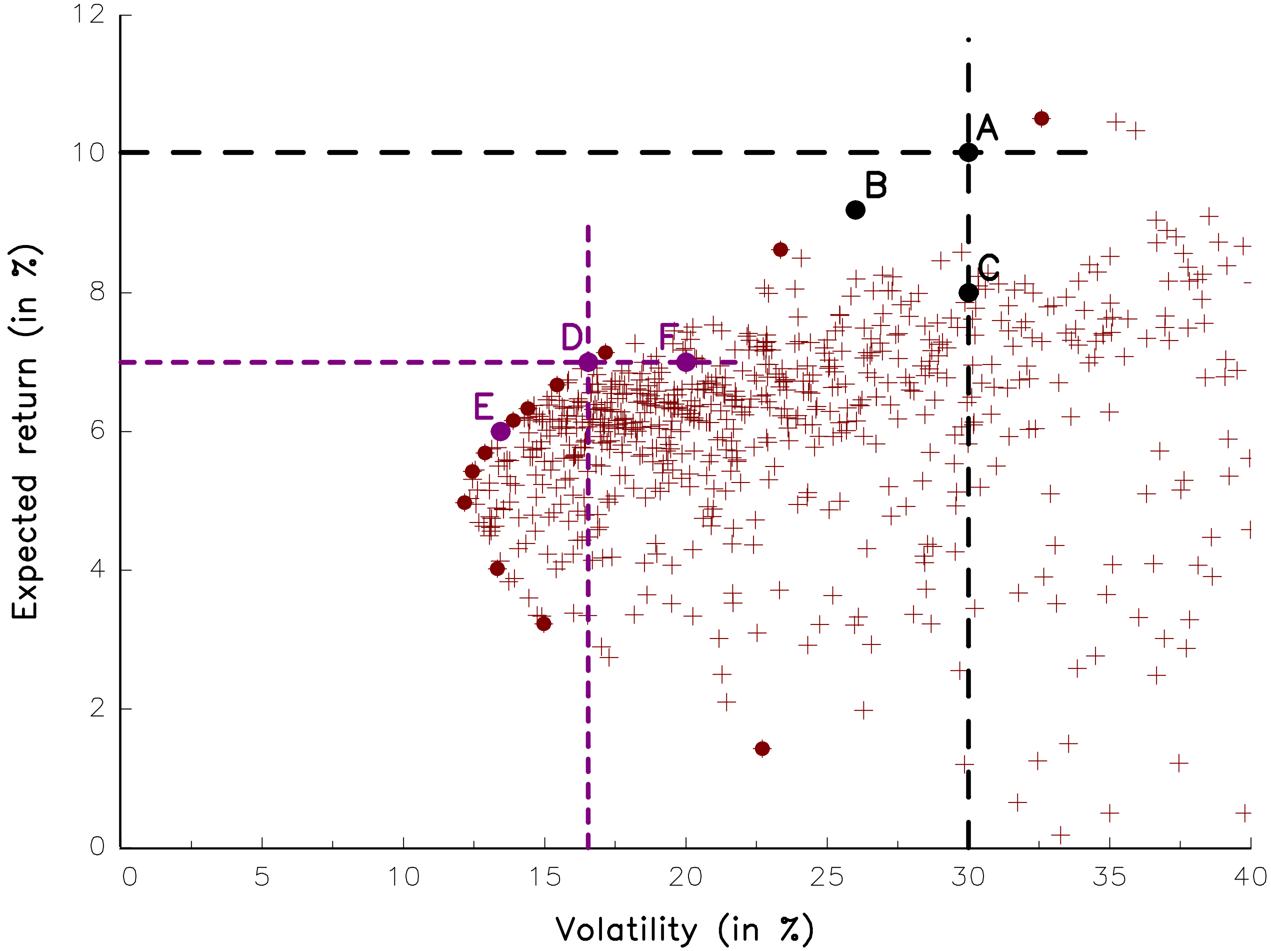


Figure 1: Optimized Markowitz portfolios (1 000 simulations)

Markowitz trick

Markowitz transforms the two original non-linear optimization problems into a quadratic optimization problem:

$$\begin{aligned} x^{\star}(\phi) &= \arg\max x^{\top}\mu - \frac{\phi}{2}x^{\top}\Sigma x \\ &\text{u.c.} \quad \mathbf{1}_n^{\top}x = 1 \end{aligned}$$

where ϕ is a risk-aversion parameter:

- $\phi = 0 \Rightarrow$ we have $\mu(x^{\star}(0)) = \mu^{+}$
- If $\phi = \infty$, the optimization problem becomes:

$$\begin{aligned} x^{\star}(\infty) &= \arg\min \frac{1}{2}x^{\top}\Sigma x \\ &\text{u.c.} \quad \mathbf{1}_n^{\top}x = 1 \end{aligned}$$

$\Rightarrow$ we have $\sigma(x^{\star}(\infty)) = \sigma^{-}$. This is the minimum variance (or MV) portfolio

The γ-problem

The previous problem can also be written as follows:

$$\begin{aligned} x^{\star}(\gamma) &= \arg\min \frac{1}{2} x^{\top} \Sigma x - \gamma x^{\top} \mu \\ \text{u.c.} & \quad \mathbf{1}_n^{\top} x = 1 \end{aligned}$$

with $\gamma = \phi^{-1}$

$\Rightarrow$ This is a standard QP problem

- The minimum variance portfolio corresponds to $\gamma = 0$
- Generally, we use the γ-problem, not the ϕ-problem

Quadratic programming problem

Definition

This is an optimization problem with a quadratic objective function and linear inequality constraints:

$$\begin{aligned} x^\star &= \arg\min \frac{1}{2} x^\top Q x - x^\top R \\ \text{u.c.} &\quad Sx \leq T \end{aligned}$$

where x is a $n \times 1$ vector, Q is a $n \times n$ matrix and R is a $n \times 1$ vector

$\Rightarrow Sx \leq T$ allows specifying linear equality constraints $Ax = B$ ($Ax \geq B$ and $Ax \leq B$) or weight constraints $x^- \leq x \leq x^+$

Quadratic programming problem

Mathematical softwares consider the following formulation:

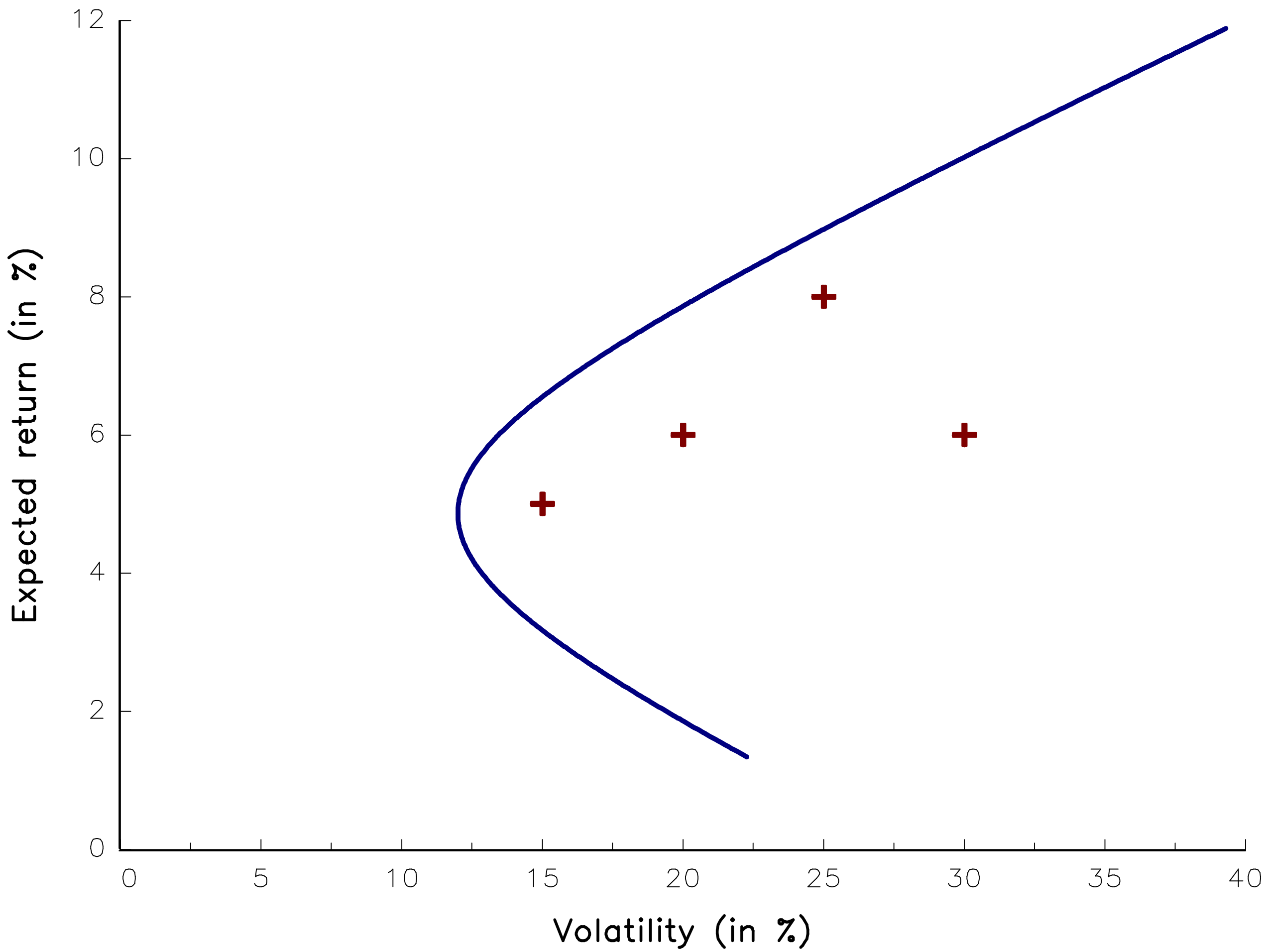
$$
\begin{aligned}
x^\star &= \arg\min \frac{1}{2} x^\top Q x - x^\top R \\
\text{u.c.} & \left\{ \begin{array}{l} Ax = B \\ Cx \leq D \\ x^- \leq x \leq x^+ \end{array} \right.
\end{aligned}
$$

because:

$$
Sx \leq T \Leftrightarrow \left[\begin{array}{c} -A \\ A \\ C \\ -I_n \\ I_n \end{array} \right] x \leq \left[\begin{array}{c} -B \\ B \\ D \\ -x^- \\ x^+ \end{array} \right]
$$

Efficient frontier

The efficient frontier is the parametric function $(\sigma(x^\star(\phi)), \mu(x^\star(\phi)))$ with $\phi \in \mathbb{R}_+$



Optimized portfolios

Table 1: Solving the ϕ-problem

ϕ	$+\infty$	5.00	2.00	1.00	0.50	0.20
$x_1^\star$	72.74	68.48	62.09	51.44	30.15	−33.75
$x_2^\star$	49.46	35.35	14.17	−21.13	−91.72	−303.49
$x_3^\star$	−20.45	12.61	62.21	144.88	310.22	806.22
$x_4^\star$	−1.75	−16.44	−38.48	−75.20	−148.65	−368.99
$\mu(x^\star)$	4.86	5.57	6.62	8.38	11.90	22.46
$\sigma(x^\star)$	12.00	12.57	15.23	22.27	39.39	94.57

Solving μ- and σ-problems

This is equivalent to finding the optimal value of γ such that:

$$\mu\left(x^\star\left(\gamma\right)\right) = \mu^\star$$

or:

$$\sigma\left(x^\star\left(\gamma\right)\right) = \sigma^\star$$

We know that:

- the functions $\mu\left(x^\star\left(\gamma\right)\right)$ and $\sigma\left(x^\star\left(\gamma\right)\right)$ are increasing with respect to γ
- the functions $\mu\left(x^\star\left(\gamma\right)\right)$ and $\sigma\left(x^\star\left(\gamma\right)\right)$ are bounded:

$$\begin{aligned} \mu^- &\leq \mu\left(x^\star\left(\gamma\right)\right) \leq \mu^+ \\ \sigma^- &\leq \sigma\left(x^\star\left(\gamma\right)\right) \leq \sigma^+ \end{aligned}$$

$\Rightarrow$ The optimal value of γ can then be easily computed using the bisection algorithm

Solving μ- and σ-problems

We want to solve $f(\gamma) = c$ where:

- $f(\gamma) = \mu(x^{\star}(\gamma))$ and $c = \mu^{\star}$
- or $f(\gamma) = \sigma(x^{\star}(\gamma))$ and $c = \sigma^{\star}$

Bisection algorithm

1. We assume that $\gamma^{\star} \in [\gamma_1, \gamma_2]$
2. If $\gamma_2 - \gamma_1 \leq \varepsilon$, then stop
3. We compute:
$$\bar{\gamma} = \frac{\gamma_1 + \gamma_2}{2}$$
and $f(\bar{\gamma})$
4. We update γ_1 and γ_2 as follows:
    1. If $f(\bar{\gamma}) < c$, then $\gamma^{\star} \in [\gamma_c, \gamma_2]$ and $\gamma_1 \leftarrow \gamma_c$
    2. If $f(\bar{\gamma}) > c$, then $\gamma^{\star} \in [\gamma_1, \gamma_c]$ and $\gamma_2 \leftarrow \gamma_c$
5. Go to Step 2

Solving μ- and σ-problems

Table 2: Solving the unconstrained μ-problem

$\mu^\star$	5.00	6.00	7.00	8.00	9.00
$x_1^\star$	71.92	65.87	59.81	53.76	47.71
$x_2^\star$	46.73	26.67	6.62	−13.44	−33.50
$x_3^\star$	−14.04	32.93	79.91	126.88	173.86
$x_4^\star$	−4.60	−25.47	−46.34	−67.20	−88.07
$\sigma(x^\star)$	12.02	13.44	16.54	20.58	25.10
ϕ	25.79	3.10	1.65	1.12	0.85

Table 3: Solving the unconstrained σ-problem

$\sigma^\star$	15.00	20.00	25.00	30.00	35.00
$x_1^\star$	62.52	54.57	47.84	41.53	35.42
$x_2^\star$	15.58	−10.75	−33.07	−54.00	−74.25
$x_3^\star$	58.92	120.58	172.85	221.88	269.31
$x_4^\star$	−37.01	−64.41	−87.62	−109.40	−130.48
$\mu(x^\star)$	6.55	7.87	8.98	10.02	11.03
ϕ	2.08	1.17	0.86	0.68	0.57

Adding some constraints

We have:

$$\begin{aligned} x^{\star}(\gamma) &= \arg\min \frac{1}{2} x^\top \Sigma x - \gamma x^\top \mu \\ & \text{u.c. } \left\{ \begin{array}{l} \mathbf{1}_n^\top x = 1 \\ x \in \Omega \end{array} \right. \end{aligned}$$

where $x \in \Omega$ corresponds to the set of restrictions

Two classical constraints:

- no short-selling restriction

$$x_i \geq 0$$

- upper bound

$$x_i \leq c$$

Adding some constraints

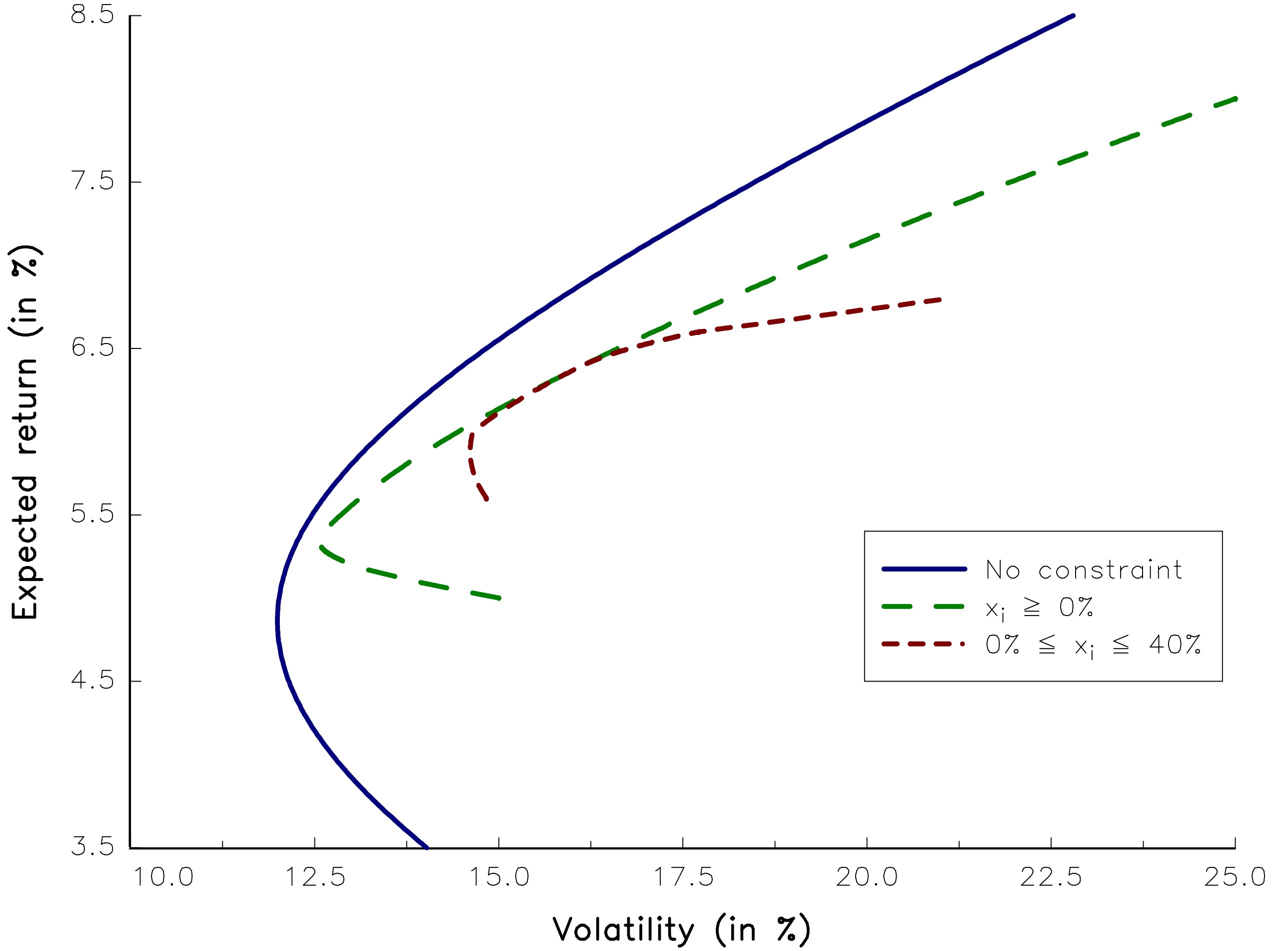


Figure 2: The efficient frontier with some weight constraints

Adding some constraints

Table 4: Solving the σ-problem with weight constraints

	$x_i \in \mathbb{R}$		$x_i \geq 0$		$0 \leq x_i \leq 40\%$	
$\sigma^\star$	15.00	20.00	15.00	20.00	15.00	20.00
$x_1^\star$	62.52	54.57	45.59	24.88	40.00	6.13
$x_2^\star$	15.58	−10.75	24.74	4.96	34.36	40.00
$x_3^\star$	58.92	120.58	29.67	70.15	25.64	40.00
$x_4^\star$	−37.01	−64.41	0.00	0.00	0.00	13.87
$\mu(x^\star)$	6.55	7.87	6.14	7.15	6.11	6.74
ϕ	2.08	1.17	1.61	0.91	1.97	0.28

Analytical solution

The Lagrange function is:

$$\mathcal{L}(x;\lambda_0) = x^\top \mu - \frac{\phi}{2} x^\top \Sigma x + \lambda_0 \left(\mathbf{1}_n^\top x - 1\right)$$

The first-order conditions are:

$$\left\{\begin{array}{l} \partial_x \mathcal{L}(x;\lambda_0) = \mu - \phi \Sigma x + \lambda_0 \mathbf{1}_n = \mathbf{0}_n \\ \partial_{\lambda_0} \mathcal{L}(x;\lambda_0) = \mathbf{1}_n^\top x - 1 = 0 \end{array}\right.$$

We obtain:

$$x = \phi^{-1}\Sigma^{-1}\left(\mu + \lambda_0 \mathbf{1}_n\right)$$

Because $\mathbf{1}_n^\top x - 1 = 0$, we have:

$$\mathbf{1}_n^\top \phi^{-1}\Sigma^{-1}\mu + \lambda_0 \left(\mathbf{1}_n^\top \phi^{-1}\Sigma^{-1}\mathbf{1}_n\right) = 1$$

It follows that:

$$\lambda_0 = \frac{1 - \mathbf{1}_n^\top \phi^{-1}\Sigma^{-1}\mu}{\mathbf{1}_n^\top \phi^{-1}\Sigma^{-1}\mathbf{1}_n}$$

Analytical solution

The solution is then:

$$x^{\star}(\phi) = \frac{\Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^{\top}\Sigma^{-1}\mathbf{1}_n} + \frac{1}{\phi} \cdot \frac{\left(\mathbf{1}_n^{\top}\Sigma^{-1}\mathbf{1}_n\right)\Sigma^{-1}\mu - \left(\mathbf{1}_n^{\top}\Sigma^{-1}\mu\right)\Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^{\top}\Sigma^{-1}\mathbf{1}_n}$$

Remark

The global minimum variance portfolio is:

$$x_{\mathrm{mv}} = x^{\star}(\infty) = \frac{\Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^{\top}\Sigma^{-1}\mathbf{1}_n}$$

Analytical solution

In the case of no short-selling, the Lagrange function becomes:

$$\mathcal{L}(x; \lambda_0, \lambda) = x^\top \mu - \frac{\phi}{2} x^\top \Sigma x + \lambda_0 \left(\mathbf{1}_n^\top x - 1\right) + \lambda^\top x$$

where $\lambda = (\lambda_1, \ldots, \lambda_n) \geq \mathbf{0}_n$ is the vector of Lagrange coefficients associated with the constraints $x_i \geq 0$

- The first-order condition is:

$$\mu - \phi \Sigma x + \lambda_0 \mathbf{1} + \lambda = \mathbf{0}_n$$

- The Kuhn-Tucker conditions are:

$$\min (\lambda_i, x_i) = 0$$

The tangency portfolio

Markowitz

There are many optimized portfolios
$\Rightarrow$ there are many optimal portfolios

Tobin

One optimized portfolio dominates all the others if there is a risk-free asset

The tangency portfolio

We consider a combination of the risk-free asset and a portfolio x:

$$R(y) = (1-\alpha) r + \alpha R(x)$$

where:

- r is the return of the risk-free asset
- $y = \begin{pmatrix} \alpha x \\ 1-\alpha \end{pmatrix}$ is a vector of dimension $(n+1)$
- $\alpha \geq 0$ is the proportion of the wealth invested in the risky portfolio

It follows that:

$$\mu(y) = (1-\alpha) r + \alpha \mu(x) = r + \alpha(\mu(x) - r)$$

and:

$$\sigma^2(y) = \alpha^2 \sigma^2(x)$$

We deduce that:

$$\mu(y) = r + \frac{(\mu(x) - r)}{\sigma(x)} \sigma(y)$$

The tangency portfolio

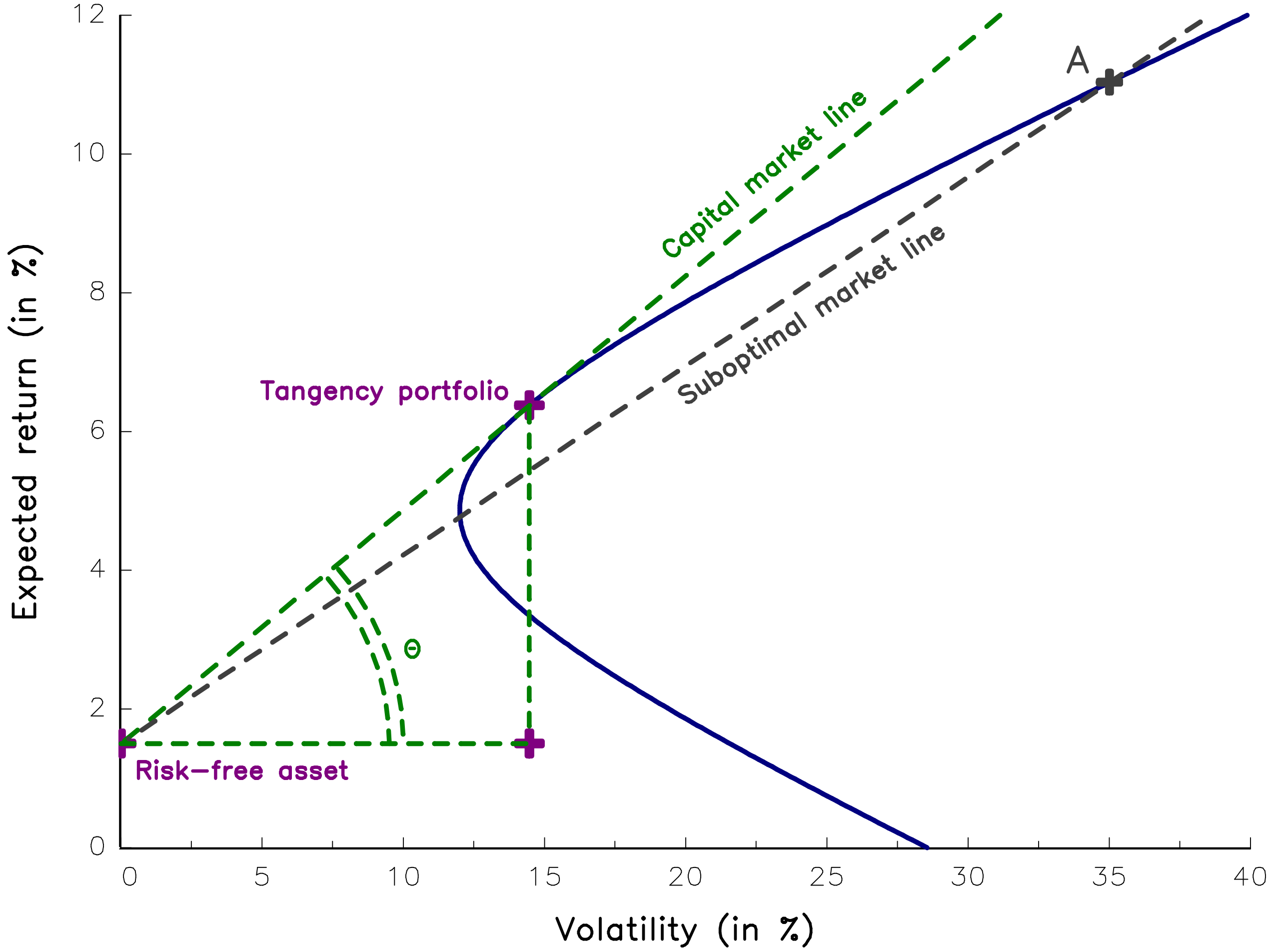


Figure 3: The capital market line ($r = 1.5\%$)

The tangency portfolio

Let $\mathrm{SR}\left(x \mid r\right)$ be the Sharpe ratio of portfolio x:

$$\mathrm{SR}\left(x \mid r\right) = \frac{\mu\left(x\right) - r}{\sigma\left(x\right)}$$

We obtain:

$$\frac{\mu\left(y\right) - r}{\sigma\left(y\right)} = \frac{\mu\left(x\right) - r}{\sigma\left(x\right)} \Leftrightarrow \mathrm{SR}\left(y \mid r\right) = \mathrm{SR}\left(x \mid r\right)$$

The tangency portfolio is the one that maximizes the angle θ or equivalently $\tan\theta$:

$$\tan\theta = \mathrm{SR}\left(x \mid r\right) = \frac{\mu\left(x\right) - r}{\sigma\left(x\right)}$$

The tangency portfolio is the risky portfolio corresponding to the maximum Sharpe ratio

The tangency portfolio

Example 2

We consider Example 1 and $r = 1.5\%$

The composition of the tangency portfolio $x^{\star}$ is:

$$x^{\star} = \begin{pmatrix} 63.63\% \\ 19.27\% \\ 50.28\% \\ -33.17\% \end{pmatrix}$$

We have:

$$\begin{aligned} \mu\left(x^{\star}\right) &= 6.37\% \\ \sigma\left(x^{\star}\right) &= 14.43\% \\ \mathrm{SR}\left(x^{\star} \mid r\right) &= 0.34 \\ \theta\left(x^{\star}\right) &= 18.64 \text{ degrees} \end{aligned}$$

The tangency portfolio

Let us consider a portfolio x of risky assets and a risk-free asset r. We denote by $\tilde{x}$ the augmented vector of dimension $n+1$ such that:

$$\tilde{x} = \begin{pmatrix} x \\ x_r \end{pmatrix} \quad \text{and} \quad \tilde{\Sigma} = \begin{pmatrix} \Sigma & \mathbf{0}_n \\ \mathbf{0}_n^\top & 0 \end{pmatrix} \quad \text{and} \quad \tilde{\mu} = \begin{pmatrix} \mu \\ r \end{pmatrix}$$

If we include the risk-free asset, the Markowitz γ-problem becomes:

$$\begin{aligned} \tilde{x}^\star(\gamma) &= \arg\min \frac{1}{2}\tilde{x}^\top \tilde{\Sigma}\tilde{x} - \gamma \tilde{x}^\top \tilde{\mu} \\ \text{u.c.} \quad & \mathbf{1}_n^\top \tilde{x} = 1 \end{aligned}$$

Two-fund separation theorem

We can show that (RPB, pages 13-14):

$$\tilde{x}^\star = \underbrace{\alpha \cdot \begin{pmatrix} x_0^\star \\ 0 \end{pmatrix}}_{\text{risky assets}} + \underbrace{(1-\alpha) \cdot \begin{pmatrix} \mathbf{0}_n \\ 1 \end{pmatrix}}_{\text{risk-free asset}}$$

The tangency portfolio

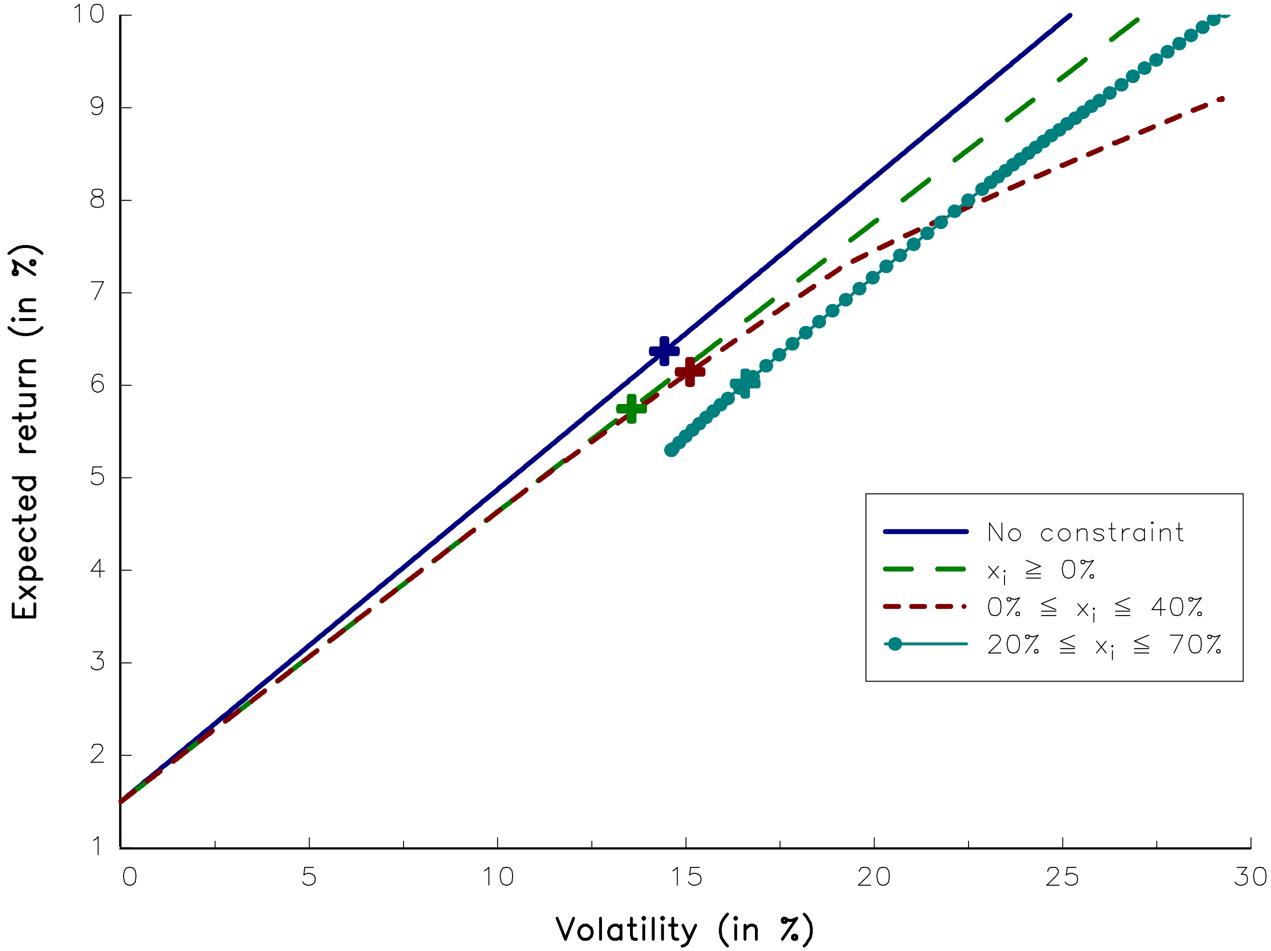


Figure 4: The efficient frontier with a risk-free asset

Market equilibrium and CAPM

- $x^\star$ is the tangency portfolio
- On the efficient frontier, we have:

$$\mu(y) = r + \frac{\sigma(y)}{\sigma(x^\star)}\left(\mu(x^\star) - r\right)$$

- We consider a portfolio z with a proportion w invested in the asset i and a proportion $(1-w)$ invested in the tangency portfolio $x^\star$:

$$\begin{aligned}\mu(z) &= w\mu_i + (1-w)\,\mu(x^\star) \\ \sigma^2(z) &= w^2\sigma_i^2 + (1-w)^2\,\sigma^2(x^\star) + 2w(1-w)\,\rho(\mathbf{e}_i, x^\star)\,\sigma_i\sigma(x^\star)\end{aligned}$$

It follows that:

$$\frac{\partial\,\mu(z)}{\partial\,\sigma(z)} = \frac{\mu_i - \mu(x^\star)}{\left(w\sigma_i^2 + (w-1)\,\sigma^2(x^\star) + (1-2w)\,\rho(\mathbf{e}_i, x^\star)\,\sigma_i\sigma(x^\star)\right)\sigma^{-1}(z)}$$

Market equilibrium and CAPM

1. When $w = 0$, we have:

$$\frac{\partial\,\mu(z)}{\partial\,\sigma(z)} = \frac{\mu_i - \mu(x^\star)}{\left(-\sigma^2(x^\star) + \rho(\mathbf{e}_i, x^\star)\,\sigma_i\sigma(x^\star)\right)\sigma^{-1}(x^\star)}$$

2. When $w = 0$, the portfolio z is the tangency portfolio $x^\star$ and the previous derivative is equal to the Sharpe ratio $\mathrm{SR}(x^\star \mid r)$

We deduce that:

$$\frac{(\mu_i - \mu(x^\star))\,\sigma(x^\star)}{\rho(\mathbf{e}_i, x^\star)\,\sigma_i\sigma(x^\star) - \sigma^2(x^\star)} = \frac{\mu(x^\star) - r}{\sigma(x^\star)}$$

which is equivalent to:

$$\pi_i = \mu_i - r = \beta_i\left(\mu(x^\star) - r\right)$$

with π_i the risk premium of the asset i and:

$$\beta_i = \frac{\rho(\mathbf{e}_i, x^\star)\,\sigma_i}{\sigma(x^\star)} = \frac{\mathrm{cov}(R_i, R(x^\star))}{\mathrm{var}(R(x^\star))}$$

Market equilibrium and CAPM

CAPM

The risk premium of the asset i is equal to its beta times the excess return of the tangency portfolio

$\Rightarrow$ We can extend the previous result to the case of a portfolio x (and not only to the asset i):

$$z = wx + (1 - w)\, x^{\star}$$

In this case, we have:

$$\pi(x) = \mu(x) - r = \beta(x \mid x^{\star})(\mu(x^{\star}) - r)$$

Computation of the beta

The least squares method

- $R_{i,t}$ and $R_t(x)$ be the returns of asset i and portfolio x at time t
- β_i is estimated with the linear regression:

$$R_{i,t} = \alpha_i + \beta_i R_t(x) + \varepsilon_{i,t}$$

- For a portfolio y, we have:

$$R_t(y) = \alpha + \beta R_t(x) + \varepsilon_t$$

Computation of the beta

The covariance method

Another way to compute the beta of portfolio y is to use the following relationship:

$$\beta\left(y \mid x\right) = \frac{\sigma\left(y, x\right)}{\sigma^2\left(x\right)} = \frac{y^\top \Sigma x}{x^\top \Sigma x}$$

We deduce that the expression of the beta of asset i is also:

$$\beta_i = \beta\left(\mathbf{e}_i \mid x\right) = \frac{\mathbf{e}_i^\top \Sigma x}{x^\top \Sigma x} = \frac{\left(\Sigma x\right)_i}{x^\top \Sigma x}$$

The beta of a portfolio is the weighted average of the beta of the assets that compose the portfolio:

$$\beta\left(y \mid x\right) = \frac{y^\top \Sigma x}{x^\top \Sigma x} = y^\top \frac{\Sigma x}{x^\top \Sigma x} = \sum_{i=1}^{n} y_i \beta_i$$

Market equilibrium and CAPM

We have $x^{\star} = (63.63\%, 19.27\%, 50.28\%, -33.17\%)$ and $\mu(x^{\star}) = 6.37\%$

Table 5: Computation of the beta and the risk premium (Example 2)

Portfolio y	$\mu(y)$	$\mu(y) - r$	$\beta(y \mid x^{\star})$	$\pi(y \mid x^{\star})$
$\mathbf{e}_1$	5.00	3.50	0.72	3.50
$\mathbf{e}_2$	6.00	4.50	0.92	4.50
$\mathbf{e}_3$	8.00	6.50	1.33	6.50
$\mathbf{e}_4$	6.00	4.50	0.92	4.50
x_{ew}	6.25	4.75	0.98	4.75

Example 2

We consider four assets. Their expected returns are equal to 5%, 6%, 8% and 6% while their volatilities are equal to 15%, 20%, 25% and 30%. The correlation matrix of asset returns is given by the following matrix:

$$C = \begin{pmatrix} 1.00 & & & \\ 0.10 & 1.00 & & \\ 0.40 & 0.70 & 1.00 & \\ 0.50 & 0.40 & 0.80 & 1.00 \end{pmatrix}$$

The risk free rate is equal to $r = 1.5\%$

From active management to passive management

- Active management
- Sharpe (1964)

$$\pi(x) = \beta(x \mid x^\star)\,\pi(x^\star)$$

- Jensen (1969)

$$R_t(x) = \alpha + \beta R_t(b) + \varepsilon_t$$

  where $R_t(x)$ is the fund return and $R_t(b)$ is the benchmark return
- Passive management (John McQuown, WFIA, 1971)

Active management = Alpha

Passive management = Beta

Impact of the constraints

If we impose a lower bound $x_i \geq 0$, the tangency portfolio becomes $x^\star = (53.64\%, 32.42\%, 13.93\%, 0.00\%)$ and we have $\mu(x^\star) = 5.74\%$

Table 6: Computation of the beta with a constrained tangency portfolio

Portfolio	$\mu(y) - r$	$\beta(y \mid x^\star)$	$\pi(y \mid x^\star)$
$\mathbf{e}_1$	3.50	0.83	3.50
$\mathbf{e}_2$	4.50	1.06	4.50
$\mathbf{e}_3$	6.50	1.53	6.50
$\mathbf{e}_4$	4.50	1.54	6.53
x_{ew}	4.75	1.24	5.26

$\Rightarrow \mu_4 - r = \beta_4(\mu(x^\star) - r) + \pi_4^-$ where $\pi_4^- \leq 0$ represents a negative premium due to a lack of arbitrage on the fourth asset

Tracking error

- Portfolio $x = (x_1, \ldots, x_n)$
- Benchmark $b = (b_1, \ldots, b_n)$
- The tracking error between the active portfolio x and its benchmark b is the difference between the return of the portfolio and the return of the benchmark:

$$e = R(x) - R(b) = \sum_{i=1}^{n} x_i R_i - \sum_{i=1}^{n} b_i R_i = x^\top R - b^\top R = (x - b)^\top R$$

- The expected excess return is:

$$\mu(x \mid b) = \mathbb{E}[e] = (x - b)^\top \mu$$

- The volatility of the tracking error is:

$$\sigma(x \mid b) = \sigma(e) = \sqrt{(x - b)^\top \Sigma (x - b)}$$

Markowitz optimization problem

The expected return of the portfolio is replaced by the expected excess return and the volatility of the portfolio is replaced by the volatility of the tracking error

σ-problem

The objective of the investor is to maximize the expected tracking error with a constraint on the tracking error volatility:

$$
\begin{aligned}
x^{\star} &= \arg\max \mu\left(x \mid b\right) \\
\text{u.c.} & \left\{\begin{array}{l} \mathbf{1}_n^{\top} x = 1 \\ \sigma\left(x \mid b\right) \leq \sigma^{\star} \end{array}\right.
\end{aligned}
$$

Equivalent QP problem

We transform the σ-problem into a γ-problem:

$$x^{\star}(\gamma) = \arg\min f(x \mid b)$$

with:

$$\begin{aligned} f(x \mid b) &= \frac{1}{2}(x-b)^{\top}\Sigma(x-b) - \gamma(x-b)^{\top}\mu \\ &= \frac{1}{2}x^{\top}\Sigma x - x^{\top}(\gamma\mu + \Sigma b) + \left(\frac{1}{2}b^{\top}\Sigma b + \gamma b^{\top}\mu\right) \\ &= \frac{1}{2}x^{\top}\Sigma x - x^{\top}(\gamma\mu + \Sigma b) + c \end{aligned}$$

where c is a constant which does not depend on Portfolio x

QP problem with $Q = \Sigma$ and $R = \gamma\mu + \Sigma b$

Remark

The efficient frontier is the parametric curve $(\sigma(x^{\star}(\gamma) \mid b), \mu(x^{\star}(\gamma) \mid b))$ with $\gamma \in \mathbb{R}_{+}$

Efficient frontier with a benchmark

Example 3

We consider four assets. Their expected returns are equal to 5%, 6%, 8% and 6% while their volatilities are equal to 15%, 20%, 25% and 30%. The correlation matrix of asset returns is given by the following matrix:

$$C = \begin{pmatrix} 1.00 & & & \\ 0.10 & 1.00 & & \\ 0.40 & 0.70 & 1.00 & \\ 0.50 & 0.40 & 0.80 & 1.00 \end{pmatrix}$$

The benchmark of the portfolio manager is equal to $b = (60\%, 40\%, 20\%, -20\%)$

- 1st case: No constraint
- 2nd case: $x_i^- \leq x_i$ with $x_i^- = -10\%$
- 3rd case: $x_i^- \leq x_i \leq x_i^+$ with $x_1^- = x_2^- = x_3^- = 0\%$, $x_4^- = -20\%$ and $x_i^+ = 50\%$

Efficient frontier with a benchmark

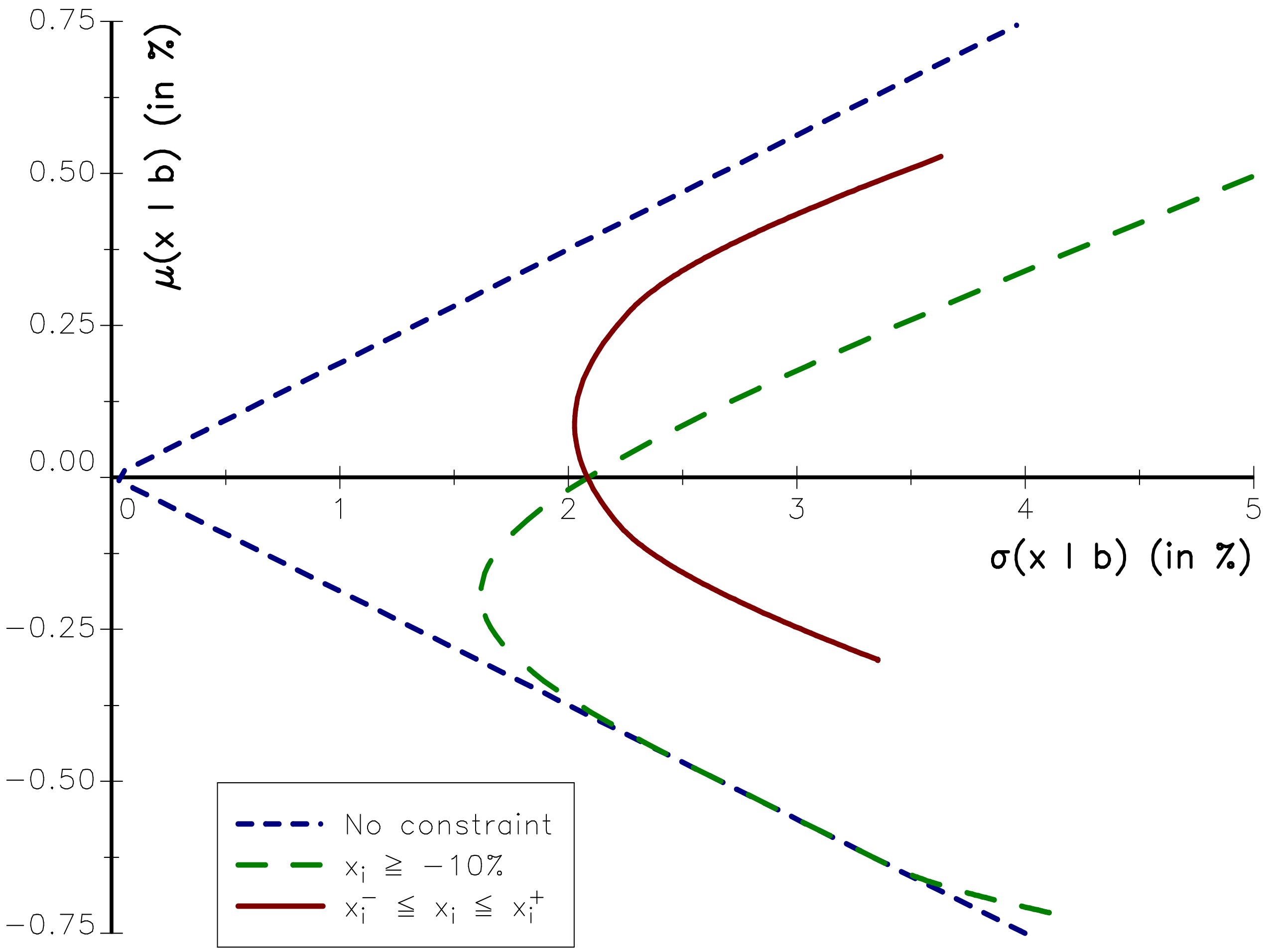


Figure 5: The efficient frontier with a benchmark (Example 3)

Information ratio

Definition

The information ratio is defined as follows:

$$\mathrm{IR}\left(x \mid b\right) = \frac{\mu\left(x \mid b\right)}{\sigma\left(x \mid b\right)} = \frac{\left(x - b\right)^\top \mu}{\sqrt{\left(x - b\right)^\top \Sigma \left(x - b\right)}}$$

Information ratio

If we consider a combination of the benchmark b and the active portfolio x, the composition of the portfolio is:

$$y = (1 - \alpha)\, b + \alpha x$$

with $\alpha \geq 0$ the proportion of wealth invested in the portfolio x. It follows that:

$$\mu (y \mid b) = (y - b)^\top \mu = \alpha \mu (x \mid b)$$

and:

$$\sigma^2 (y \mid b) = (y - b)^\top \Sigma (y - b) = \alpha^2 \sigma^2 (x \mid b)$$

We deduce that:

$$\mu (y \mid b) = \mathrm{IR} (x \mid b) \cdot \sigma (y \mid b)$$

The efficient frontier is a straight line

Tangency portfolio

If we add some constraints, the portfolio optimization problem becomes:

$$\begin{aligned} x^{\star}(\gamma) &= \arg\min \frac{1}{2} x^{\top} \Sigma x - x^{\top} (\gamma \mu + \Sigma b) \\ & \text{u.c.} \quad \left\{ \begin{array}{l} \mathbf{1}_n^{\top} x = 1 \\ x \in \Omega \end{array} \right. \end{aligned}$$

The efficient frontier is no longer a straight line

Tangency portfolio

One optimized portfolio dominates all the other portfolios. It is the portfolio which belongs to the efficient frontier and the straight line which is tangent to the efficient frontier. It is also the portfolio which maximizes the information ratio

Constrained efficient frontier with a benchmark

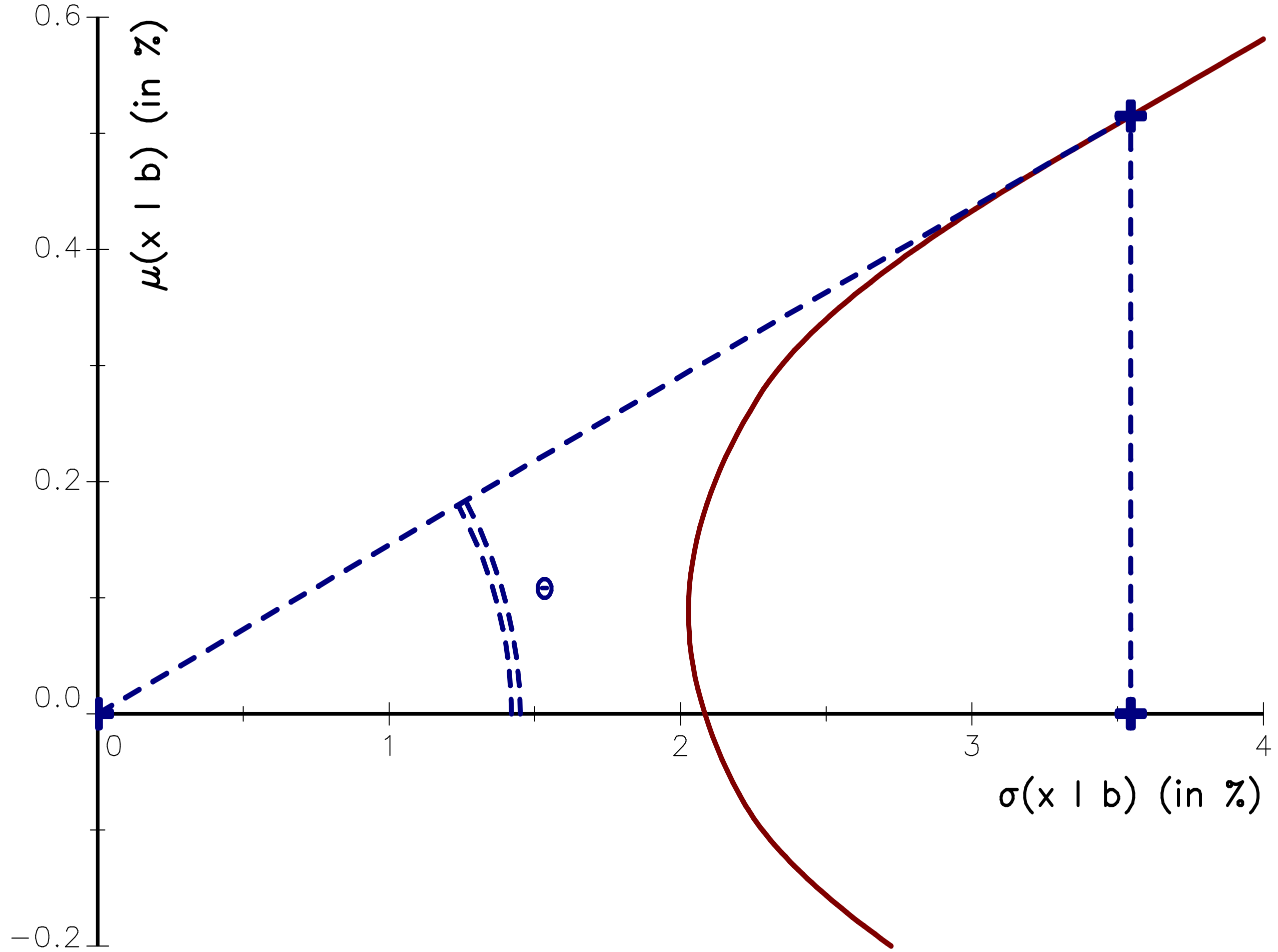


Figure 6: The tangency portfolio with respect to a benchmark (Example 3, 3rd case)

Tangency portfolio

If $x_i^- \leq x_i \leq x_i^+$ with $x_1^- = x_2^- = x_3^- = 0\%$, $x_4^- = -20\%$ and $x_i^+ = 50\%$, the tangency portfolio is equal to:

$$x^\star = \begin{pmatrix} 49.51\% \\ 29.99\% \\ 40.50\% \\ -20.00\% \end{pmatrix}$$

If $r = 1.5\%$, we recall that the MSR (maximum Sharpe ratio) portfolio is equal to:

$$x^\star = \begin{pmatrix} 63.63\% \\ 19.27\% \\ 50.28\% \\ -33.17\% \end{pmatrix}$$

When the benchmark is the risk-free rate

The Markowitz-Tobin-Sharpe approach is obtained when the benchmark is the risk-free asset r. We have:

$$\tilde{x} = \begin{pmatrix} x \\ 0 \end{pmatrix} \quad \text{and} \quad \tilde{b} = \begin{pmatrix} \mathbf{0}_n \\ 1 \end{pmatrix}$$

It follows that:

$$\tilde{\Sigma} = \begin{pmatrix} \Sigma & \mathbf{0}_n \\ \mathbf{0}_n^\top & 0 \end{pmatrix} \quad \text{and} \quad \tilde{\mu} = \begin{pmatrix} \mu \\ r \end{pmatrix}$$

When the benchmark is the risk-free rate

The objective function is then defined as follows:

$$
\begin{aligned}
f\left(\tilde{x} \mid \tilde{b}\right) &= \frac{1}{2}\left(\tilde{x}-\tilde{b}\right)^\top \Sigma \left(\tilde{x}-\tilde{b}\right) - \gamma\left(\tilde{x}-\tilde{b}\right)^\top \mu \\
&= \frac{1}{2}\tilde{x}^\top \tilde{\Sigma}\tilde{x} - \tilde{x}^\top\left(\gamma\tilde{\mu}+\tilde{\Sigma}\tilde{b}\right) + \left(\frac{1}{2}\tilde{b}^\top\tilde{\Sigma}\tilde{b}+\gamma\tilde{b}^\top\tilde{\mu}\right) \\
&= \frac{1}{2}x^\top \Sigma x - \gamma\left(x^\top\mu - r\right) \\
&= \frac{1}{2}x^\top \Sigma x - \gamma x^\top\left(\mu - r\mathbf{1}_n\right)
\end{aligned}
$$

When the benchmark is the risk-free rate

The solution of the QP problem $\tilde{x}^{\star}(\gamma) = \arg\min f\left(\tilde{x} \mid \tilde{b}\right)$ is related to the solution $x^{\star}(\gamma)$ of the Markowitz γ-problem in the following way:

$$\tilde{x}^{\star}(\gamma) = \left(\begin{array}{c} x^{\star}(\gamma) \\ 0 \end{array}\right)$$

We have $\sigma\left(\tilde{x}^{\star}(\gamma) \mid \tilde{b}\right) = \sigma\left(x^{\star}(\phi)\right)$

Remark

$\Rightarrow$ *The MSR portfolio is obtained by replacing the vector μ of expected returns by the vector $\mu - r\mathbf{1}_n$ of expected excess returns. We have:*

$$\mathrm{SR}\left(x^{\star}(\gamma) \mid r\right) = \mathrm{IR}\left(\tilde{x}^{\star}(\gamma) \mid \tilde{b}\right)$$

Black-Litterman model

Tactical asset allocation (TAA) model

How to incorporate portfolio manager's views in a strategic asset allocation (SAA)?

Two-step approach:

1. Initial allocation $\Rightarrow$ implied risk premia (Sharpe)
2. Portfolio optimization $\Rightarrow$ coherent with the bets of the portfolio manager (Markowitz)

Implied risk premium

$$\begin{aligned} x^\star &= \arg\min \frac{1}{2} x^\top \Sigma x - \gamma x^\top \left(\mu - r\mathbf{1}_n\right) \\ \text{u.c.} & \left\{ \begin{array}{l} \mathbf{1}_n^\top x = 1 \\ x \in \Omega \end{array} \right. \end{aligned}$$

If the constraints are satisfied, the first-order condition is:

$$\Sigma x - \gamma \left(\mu - r\mathbf{1}_n\right) = \mathbf{0}_n$$

The solution is:

$$x^\star = \gamma \Sigma^{-1} \left(\mu - r\mathbf{1}_n\right)$$

- In the Markowitz model, the unknown variable is the vector x
- If the initial allocation x_0 is given, it must be optimal for the investor, implying that:

$$\tilde{\mu} = r\mathbf{1}_n + \frac{1}{\gamma} \Sigma x_0$$

- $\tilde{\mu}$ is the vector of expected returns which is coherent with x_0

Implied risk premium

We deduce that:

$$\begin{aligned}\tilde{\pi} &= \tilde{\mu} - r \\ &= \frac{1}{\gamma}\Sigma x_0\end{aligned}$$

The variable $\tilde{\pi}$ is:

- the *risk premium priced* by the portfolio manager
- the '*implied risk premium*' of the portfolio manager
- the '*market risk premium*' when x_0 is the market portfolio

Implied risk aversion

The computation of $\tilde{\mu}$ needs to the value of the parameter γ or the risk aversion $\phi = \gamma^{-1}$

Since we have $\Sigma x_0 - \gamma\left(\tilde{\mu} - r\mathbf{1}_n\right) = \mathbf{0}_n$, we deduce that:

$$\begin{aligned}
(*) \quad & \Leftrightarrow \quad \gamma\left(\tilde{\mu} - r\mathbf{1}_n\right) = \Sigma x_0 \\
& \Leftrightarrow \quad \gamma\left(x_0^\top \tilde{\mu} - r x_0^\top \mathbf{1}_n\right) = x_0^\top \Sigma x_0 \\
& \Leftrightarrow \quad \gamma\left(x_0^\top \tilde{\mu} - r\right) = x_0^\top \Sigma x_0 \\
& \Leftrightarrow \quad \gamma = \frac{x_0^\top \Sigma x_0}{x_0^\top \tilde{\mu} - r}
\end{aligned}$$

It follows that

$$\phi = \frac{x_0^\top \tilde{\mu} - r}{x_0^\top \Sigma x_0} = \frac{\mathrm{SR}\left(x_0 \mid r\right)}{\sqrt{x_0^\top \Sigma x_0}} = \frac{\mathrm{SR}\left(x_0 \mid r\right)}{\sigma\left(x_0\right)}$$

where $\mathrm{SR}\left(x_0 \mid r\right)$ is the portfolio's expected Sharpe ratio

Implied risk aversion

We have:

$$\tilde{\mu} = r + \mathrm{SR}\left(x_0 \mid r\right) \frac{\Sigma x_0}{\sqrt{x_0^\top \Sigma x_0}}$$

and:

$$\tilde{\pi} = \mathrm{SR}\left(x_0 \mid r\right) \frac{\Sigma x_0}{\sqrt{x_0^\top \Sigma x_0}}$$

Implied risk premium

Example 4

We consider Example 1 and we suppose that the initial allocation x_0 is $(40\%, 30\%, 20\%, 10\%)$

- The volatility of the portfolio is equal to:

$$\sigma(x_0) = 15.35\%$$

- The objective of the portfolio manager is to target a Sharpe ratio equal to 0.25
- We obtain $\phi = 1.63$
- If $r = 3\%$, the implied expected returns are:

$$\tilde{\mu} = \begin{pmatrix} 5.47\% \\ 6.68\% \\ 8.70\% \\ 9.06\% \end{pmatrix}$$

Specification of the bets

Black and Litterman assume that μ is a Gaussian vector with expected returns $\tilde{\mu}$ and covariance matrix Γ:

$$\mu \sim \mathcal{N}\left(\tilde{\mu}, \Gamma\right)$$

The portfolio manager's views are given by this relationship:

$$P\mu = Q + \varepsilon$$

where P is a $(k \times n)$ matrix, Q is a $(k \times 1)$ vector and $\varepsilon \sim \mathcal{N}\left(0, \Omega\right)$ is a Gaussian vector of dimension k

- If the portfolio manager has two views, the matrix P has two rows $\Rightarrow$ k is then the number of views
- Ω is the covariance matrix of $P\mu - Q$, therefore it measures the uncertainty of the views

Absolute views

- We consider the three-asset case:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}$$

- The portfolio manager has an absolute view on the expected return of the first asset:

$$\mu_1 = q_1 + \varepsilon_1$$

We have:

$$P = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix},\ Q = q_1,\ \varepsilon = \varepsilon_1 \text{ and } \Omega = \omega_1^2$$

If $\omega_1 = 0$, the portfolio manager has a very high level of confidence. If $\omega_1 \neq 0$, his view is uncertain

Absolute views

- The portfolio manager has an absolute view on the expected return of the second asset:

$$\mu_2 = q_2 + \varepsilon_2$$

We have:

$$P = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix},\ Q = q_2,\ \varepsilon = \varepsilon_2 \text{ and } \Omega = \omega_2^2$$

- The portfolio manager has two absolute views:

$$\begin{aligned} \mu_1 &= q_1 + \varepsilon_1 \\ \mu_2 &= q_2 + \varepsilon_2 \end{aligned}$$

We have:

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},\ Q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix},\ \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \text{ and } \Omega = \begin{pmatrix} \omega_1^2 & 0 \\ 0 & \omega_2^2 \end{pmatrix}$$

Relative views

- The portfolio manager thinks that the outperformance of the first asset with respect to the second asset is q:

$$\mu_1 - \mu_2 = q_{1|2} + \varepsilon_{1|2}$$

We have:

$$P = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}, Q = q_{1|2}, \varepsilon = \varepsilon_{1|2} \text{ and } \Omega = \omega_{1|2}^2$$

Portfolio optimization

The Markowitz optimization problem becomes:

$$\begin{aligned} x^\star(\gamma) &= \arg\min \frac{1}{2} x^\top \Sigma x - \gamma x^\top (\bar{\mu} - r\mathbf{1}_n) \\ \text{u.c.} \quad & \mathbf{1}_n^\top x = 1 \end{aligned}$$

where $\bar{\mu}$ is the vector of expected returns conditional to the views:

$$\begin{aligned} \bar{\mu} &= \mathbb{E}[\mu \mid \text{views}] \\ &= \mathbb{E}[\mu \mid P\mu = Q + \varepsilon] \\ &= \mathbb{E}[\mu \mid P\mu - \varepsilon = Q] \end{aligned}$$

To compute $\bar{\mu}$, we consider the random vector:

$$\begin{pmatrix} \mu \\ \nu = P\mu - \varepsilon \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \tilde{\mu} \\ P\tilde{\mu} \end{pmatrix}, \begin{pmatrix} \Gamma & \Gamma P^\top \\ P\Gamma & P\Gamma P^\top + \Omega \end{pmatrix} \right)$$

Conditional distribution in the case of the normal distribution

Let us consider a Gaussian random vector defined as follows:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{x,x} & \Sigma_{x,y} \\ \Sigma_{y,x} & \Sigma_{y,y} \end{pmatrix}\right)$$

We have:

$$Y \mid X = x \sim \mathcal{N}\left(\mu_{y|x}, \Sigma_{y,y|x}\right)$$

where:

$$\mu_{y|x} = \mathbb{E}\left[Y \mid X = x\right] = \mu_y + \Sigma_{y,x}\Sigma_{x,x}^{-1}\left(x - \mu_x\right)$$

and:

$$\Sigma_{y,y|x} = \operatorname{cov}\left(Y \mid X = x\right) = \Sigma_{y,y} - \Sigma_{y,x}\Sigma_{x,x}^{-1}\Sigma_{x,y}$$

Computation of the conditional expectation

We apply the conditional expectation formula:

$$\begin{aligned}\bar{\mu} &= \mathbb{E}\left[\mu \mid \nu = Q\right] \\ &= \mathbb{E}\left[\mu\right] + \operatorname{cov}\left(\mu, v\right) \operatorname{var}\left(v\right)^{-1}\left(Q - \mathbb{E}\left[v\right]\right) \\ &= \tilde{\mu} + \Gamma P^{\top}\left(P\Gamma P^{\top} + \Omega\right)^{-1}\left(Q - P\tilde{\mu}\right)\end{aligned}$$

The conditional expectation $\bar{\mu}$ has two components:

1. The first component corresponds to the vector of implied expected returns $\tilde{\mu}$
2. The second component is a correction term which takes into account the *disequilibrium* $(Q - P\tilde{\mu})$ between the manager views and the market views

Computation of the conditional covariance matrix

The condition covariance matrix is equal to:

$$\begin{aligned}\bar{\Sigma} &= \operatorname{var}\left(\mu \mid \nu = Q\right) \\ &= \Gamma - \Gamma P^\top \left(P \Gamma P^\top + \Omega\right)^{-1} P \Gamma\end{aligned}$$

Another expression is:

$$\begin{aligned}\bar{\Sigma} &= \left(I_n + \Gamma P^\top \Omega^{-1} P\right)^{-1} \Gamma \\ &= \left(\Gamma^{-1} + P^\top \Omega^{-1} P\right)^{-1}\end{aligned}$$

The conditional covariance matrix is a weighted average of the covariance matrix Γ and the covariance matrix Ω of the manager views.

Choice of covariance matrices

Choice of Σ

From a theoretical point of view, we have:

$$\Sigma = \bar{\Sigma} = \left(\Gamma^{-1} + P^\top \Omega^{-1} P\right)^{-1}$$

In practice, we use:

$$\Sigma = \hat{\Sigma}$$

Choice of Γ

We assume that:

$$\Gamma = \tau \Sigma$$

We can also target a tracking error volatility and deduce τ

Numerical implementation of the model

The five-step approach to implement the Black-Litterman model is:

1. We estimate the empirical covariance matrix $\hat{\Sigma}$ and set $\Sigma = \hat{\Sigma}$
2. Given the current portfolio, we compute the implied risk aversion $\phi = \gamma^{-1}$ and we deduce the vector $\tilde{\mu}$ of implied expected returns
3. We specify the views by defining the P, Q and Ω matrices
4. Given a matrix Γ, we compute the conditional expectation $\bar{\mu}$
5. We finally perform the portfolio optimization with $\hat{\Sigma}$, $\bar{\mu}$ and γ

Illustration

- We use Example 4 and impose that the optimized weights are positive
- The portfolio manager has an absolute view on the first asset and a relative view on the second and third assets:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}, \; Q = \begin{pmatrix} q_1 \\ q_{2-3} \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} \varpi_1^2 & 0 \\ 0 & \varpi_{2-3}^2 \end{pmatrix}$$

- $q_1 = 4\%$, $q_{2-3} = -1\%$, $\varpi_1 = 10\%$ and $\varpi_{2-3} = 5\%$

Illustration

- Case #1: $\tau = 1$
- Case #2: $\tau = 1$ and $q_1 = 7\%$
- Case #3: $\tau = 1$ and $\varpi_1 = \varpi_{2-3} = 20\%$
- Case #4: $\tau = 10\%$
- Case #5: $\tau = 1\%$

Illustration

Table 7: Black-Litterman portfolios

	#0	#1	#2	#3	#4	#5
$x_1^\star$	40.00	33.41	51.16	36.41	38.25	39.77
$x_2^\star$	30.00	51.56	39.91	42.97	42.72	32.60
$x_3^\star$	20.00	5.46	0.00	10.85	9.14	17.65
$x_4^\star$	10.00	9.58	8.93	9.77	9.89	9.98
$\sigma\left(x^\star \mid x_0\right)$	0.00	3.65	3.67	2.19	2.18	0.45

Illustration

To calibrate the parameter τ, we could target a tracking error volatility $\sigma^\star$:

- If $\sigma^\star = 2\%$, the optimized portfolio is between portfolios $\#4$ $(\sigma(x^\star \mid x_0) = 2.18\%)$ and $\#5$ $(\sigma(x^\star \mid x_0) = 0.45\%)$
- The optimal value of τ is between 10% and 1%
- Using a bisection algorithm, we obtain $\tau = 5.2\%$

The optimal portfolio is:

$$x^\star = \begin{pmatrix} 36.80\% \\ 41.83\% \\ 11.58\% \\ 9.79\% \end{pmatrix}$$

Empirical estimator

We have:

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} \left(R_t - \bar{R} \right) \left(R_t - \bar{R} \right)^\top$$

Asynchronous markets

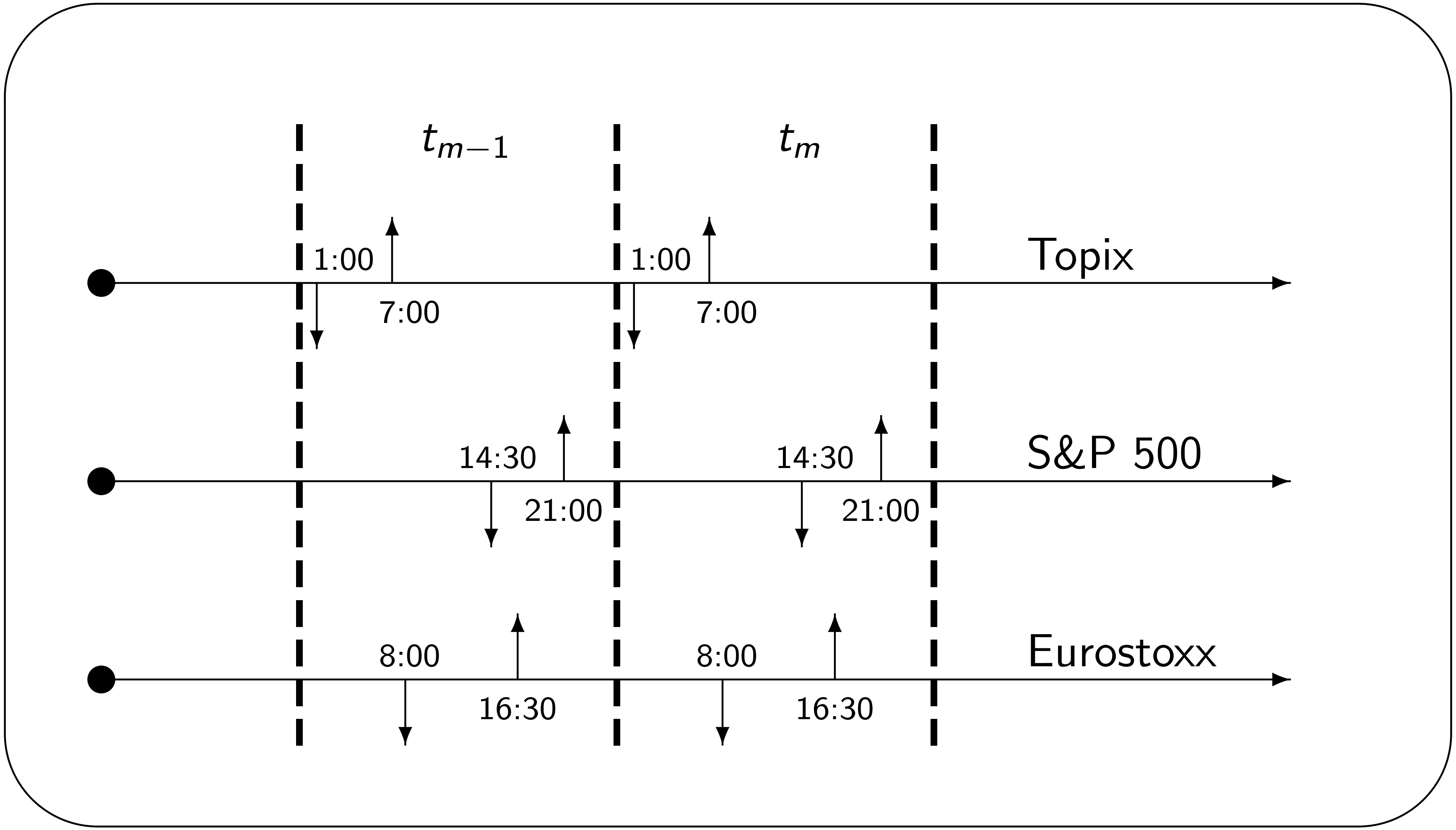


Figure 7: Trading hours of asynchronous markets (UTC time)

Asynchronous markets

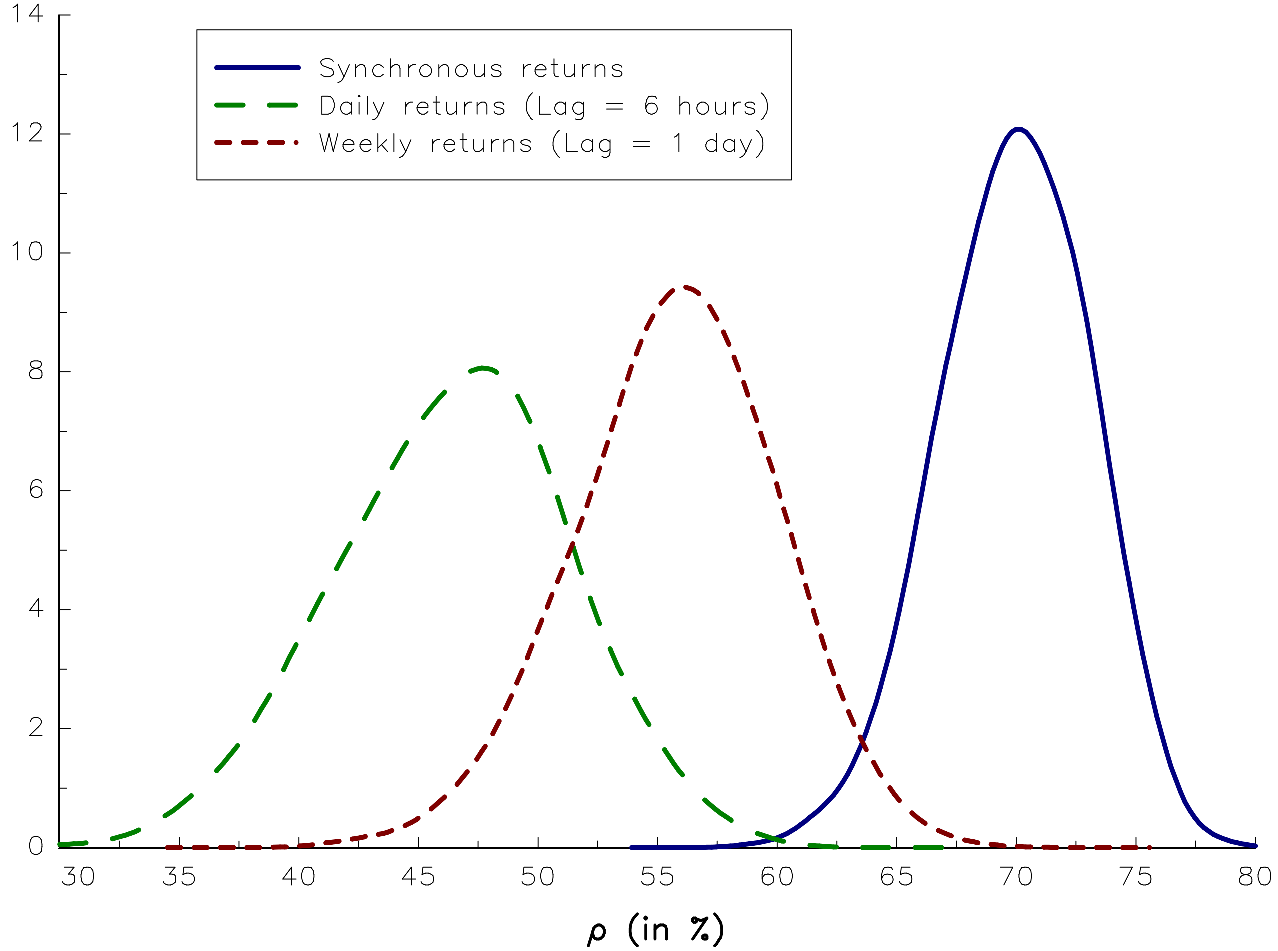


Figure 8: Density of the estimator $\hat{\rho}$ with asynchronous returns ($\rho = 70\%$)

Asynchronous markets

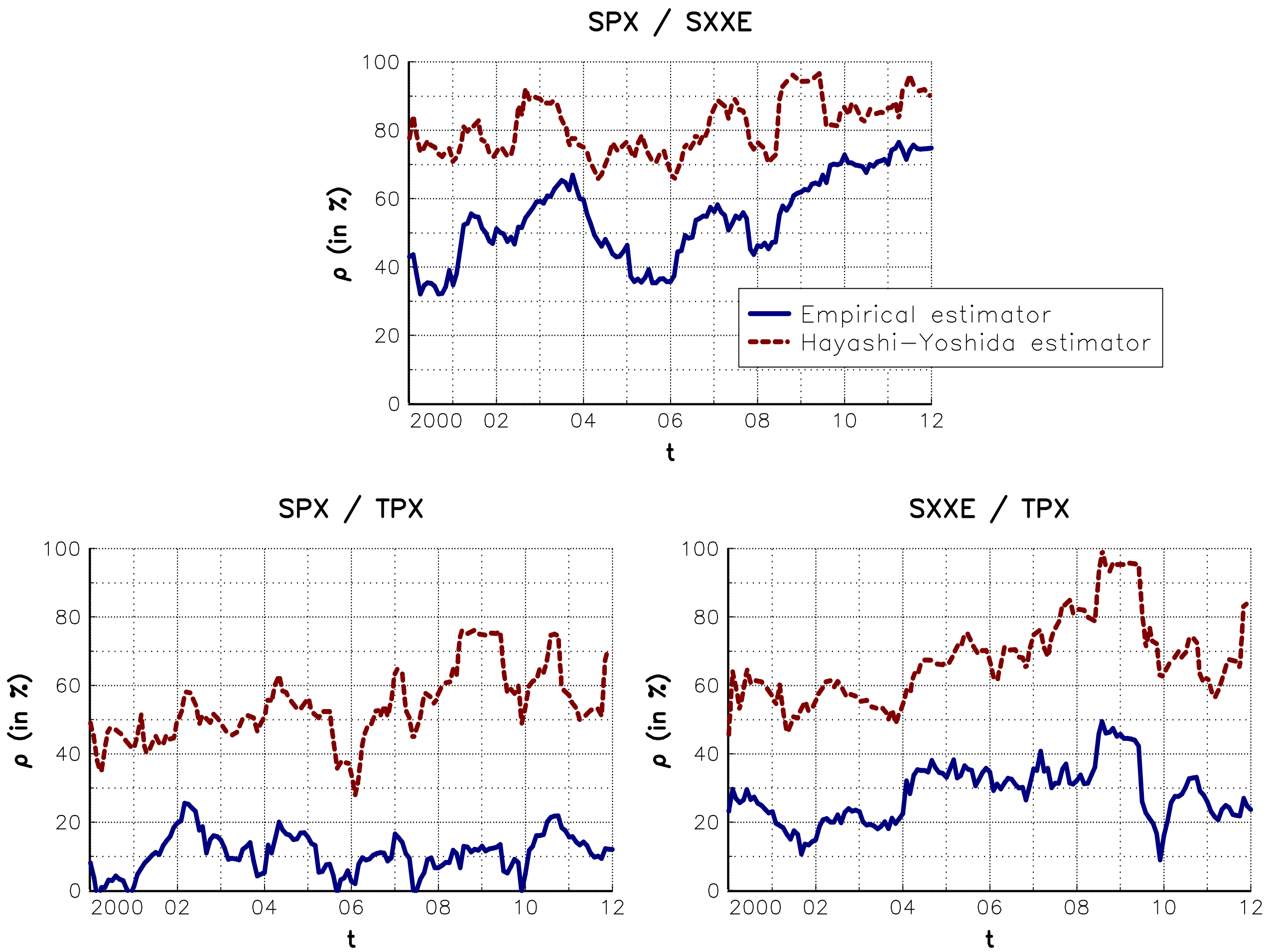


Figure 9: Hayashi-Yoshida estimator

Hayashi-Yoshida estimator

We have:

$$\tilde{\Sigma}_{i,j} = \frac{1}{T}\sum_{t=1}^{T}\left(R_{i,t} - \bar{R}_i\right)\left(R_{j,t} - \bar{R}_j\right) + \frac{1}{T}\sum_{t=1}^{T}\left(R_{i,t} - \bar{R}_i\right)\left(R_{j,t-1} - \bar{R}_j\right)$$

where j is the equity index which has a closing time after the equity index i. In our case, j is necessarily the S&P 500 index whereas i can be the Topix index or the Eurostoxx index. This estimator has two components:

1. The first component is the classical covariance estimator $\hat{\Sigma}_{i,j}$
2. The second component is a correction to take into account the lag between the two closing times

Other statistical methods

- EWMA methods
- GARCH models
- Factor models
  - Uniform correlation
  $$\rho_{i,j} = \rho$$
  - Sector approach (inter-correlation and intra-correlation)
  - Linear factor models:
  $$R_{i,t} = A_i^\top \mathcal{F}_t + \varepsilon_{i,t}$$

Economic/econometric approach

- Market timing (MT)
- Tactical asset allocation (TAA)
- Strategic asset allocation (SAA)

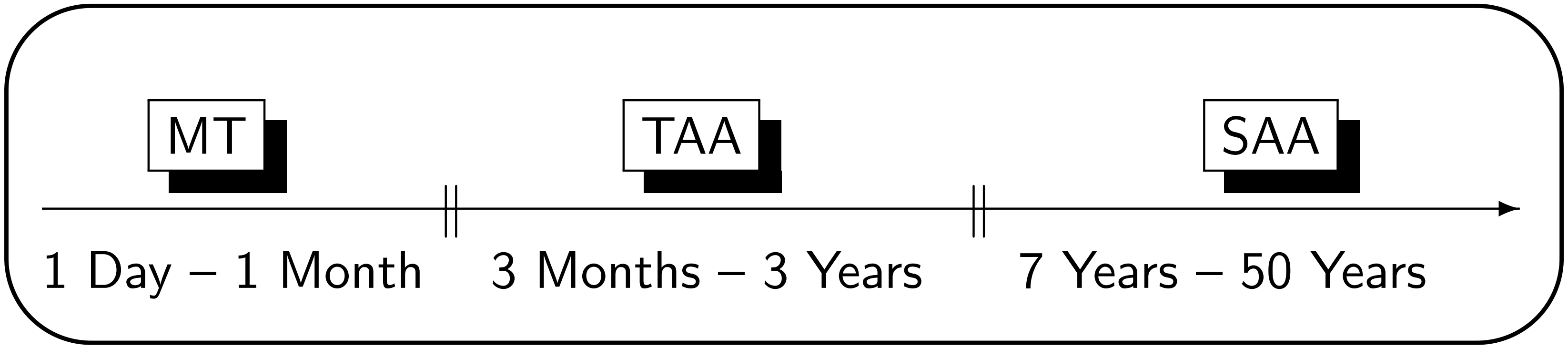


Figure 10: Time horizon of MT, TAA and SAA

Statistical/scoring approach

- Stock picking models: fundamental scoring, value, quality, sector analysis, etc.
- Bond picking models: fundamental scoring, structural model, credit arbitrage model, etc.
- Statistical models: mean-reverting, trend-following, cointegration, etc.
- Machine learning: return forecasting, scoring model, etc.

Stability issues

Example 5

We consider a universe of 3 assets. The parameters are: $\mu_1 = \mu_2 = 8\%$, $\mu_3 = 5\%$, $\sigma_1 = 20\%$, $\sigma_2 = 21\%$, $\sigma_3 = 10\%$ and $\rho_{i,j} = 80\%$. The objective is to maximize the expected return for a 15% volatility target. The optimal portfolio is $(38.3\%, 20.2\%, 41.5\%)$.

Table 8: Sensitivity of the MVO portfolio to input parameters

ρ		70%	90%		90%	
σ_2				18%	18%	
μ_1						9%
x_1	38.3	38.3	44.6	13.7	−8.0	60.6
x_2	20.2	25.9	8.9	56.1	74.1	−5.4
x_3	41.5	35.8	46.5	30.2	34.0	44.8

Solutions

In order to stabilize the optimal portfolio, we have to introduce some regularization techniques:

- Resampling techniques
- Factor analysis
- Shrinkage methods
- Random matrix theory
- Norm penalization
- Etc.

Resampling techniques

- Jacknife
- Cross validation
  - Hold-out
  - K-fold
- Bootstrap
  - Resubstitution
  - Out of the bag
  - .632

Resampling techniques

Example 6

We consider a universe of four assets. The expected returns are $\hat{\mu}_1 = 5\%$, $\hat{\mu}_2 = 9\%$, $\hat{\mu}_3 = 7\%$ and $\hat{\mu}_4 = 6\%$ whereas the volatilities are equal to $\hat{\sigma}_1 = 4\%$, $\hat{\sigma}_2 = 15\%$, $\hat{\sigma}_3 = 5\%$ and $\hat{\sigma}_4 = 10\%$. The correlation matrix is the following:

$$\hat{C} = \begin{pmatrix} 1.00 & & & \\ 0.10 & 1.00 & & \\ 0.40 & 0.20 & 1.00 & \\ -0.10 & -0.10 & -0.20 & 1.00 \end{pmatrix}$$

Resampling techniques

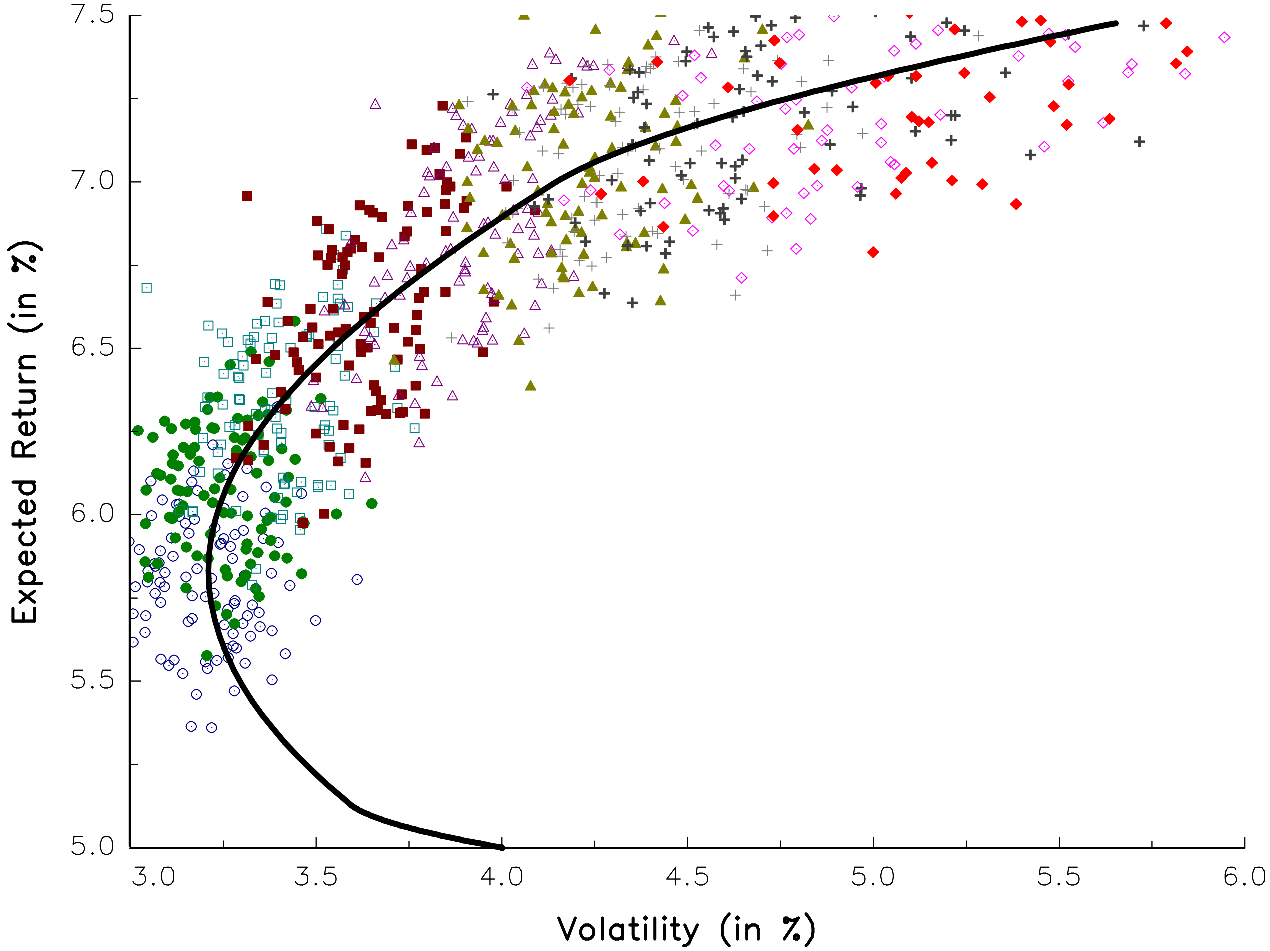


Figure 11: Uncertainty of the efficient frontier

Resampling techniques

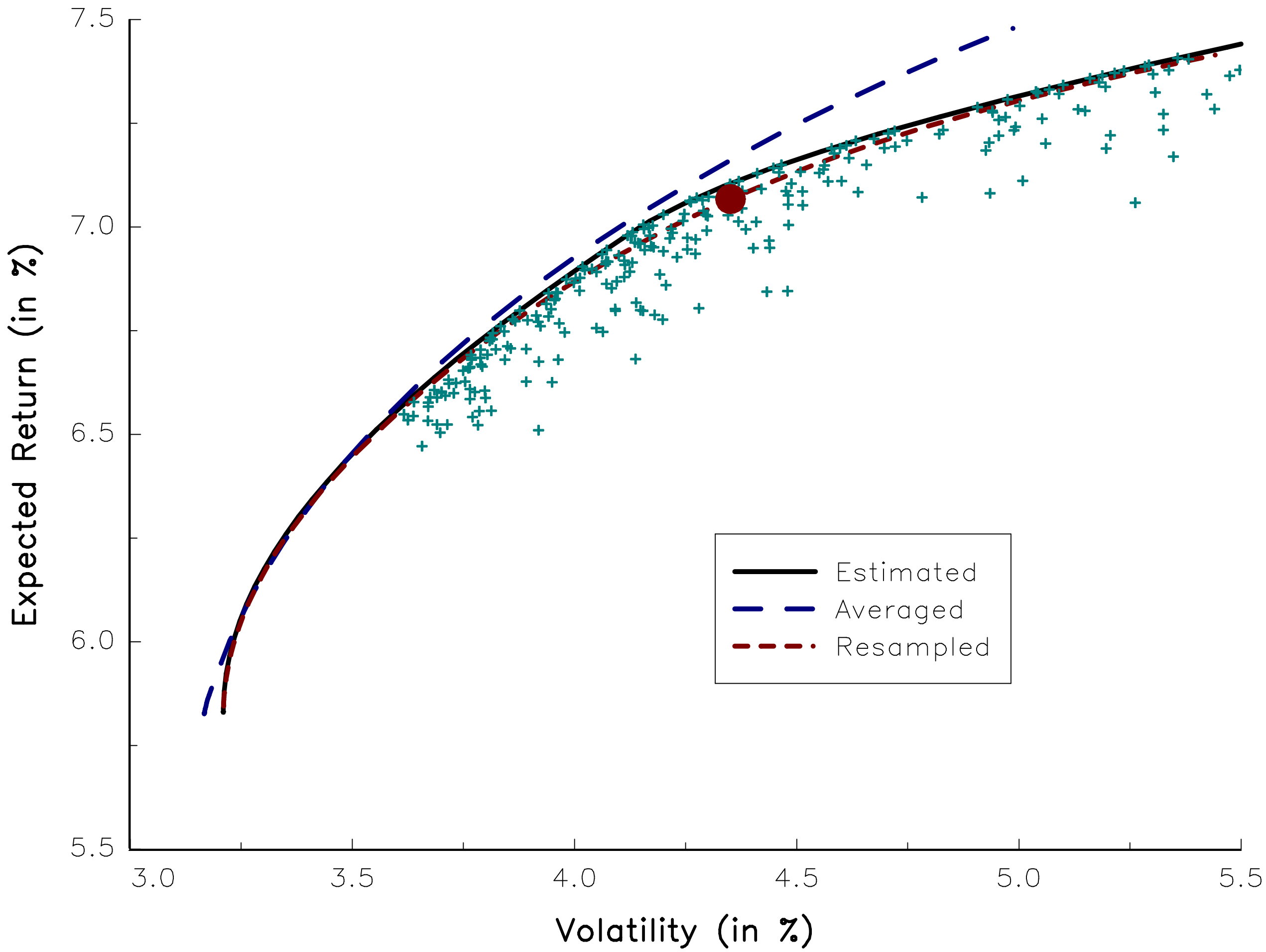


Figure 12: Resampled efficient frontier

Resampling techniques

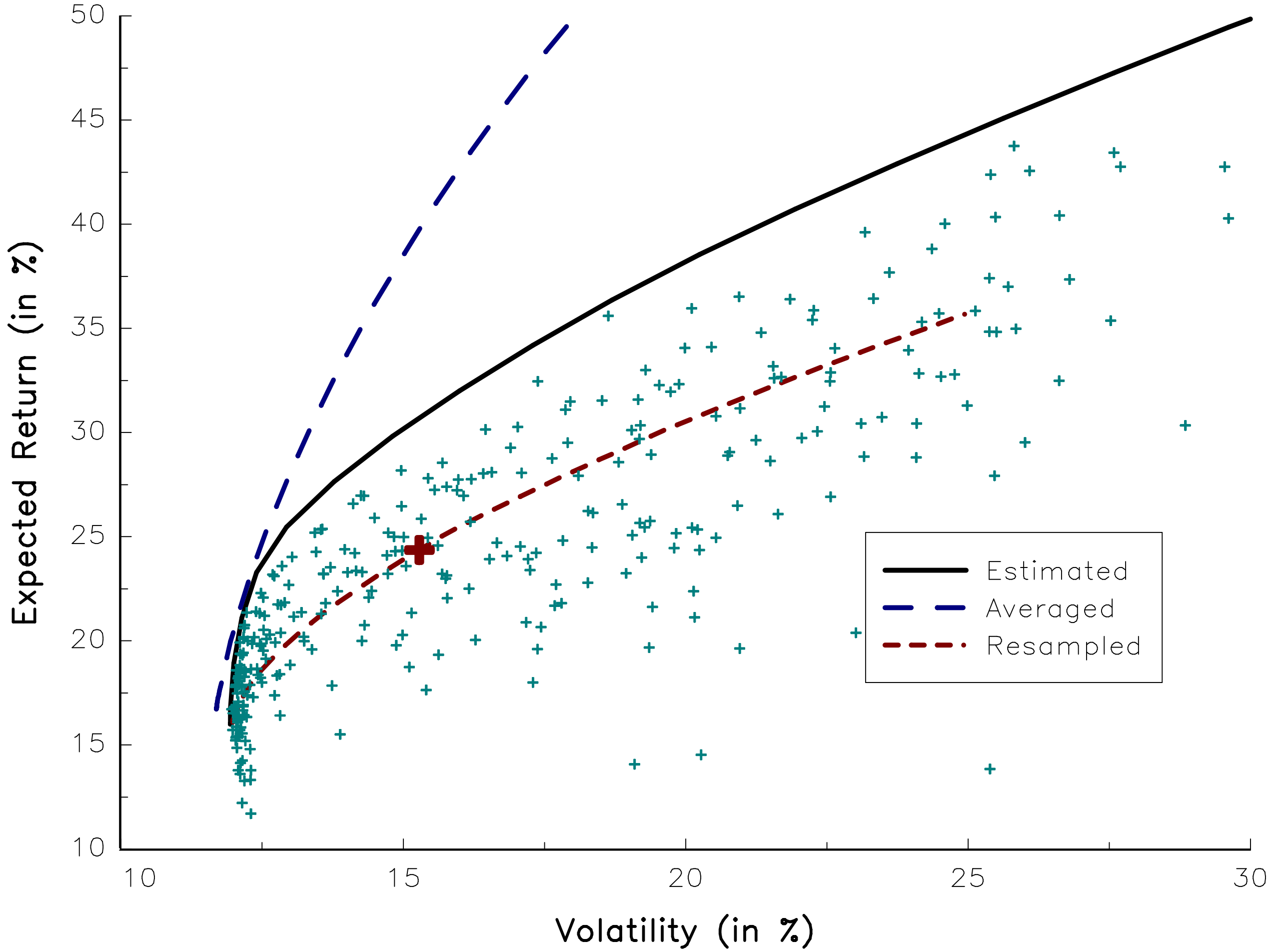


Figure 13: S&P 100 resampled efficient frontier (Bootstrap approach)

Source: Bruder *et al.* (2013)

How to denoise the covariance matrix?

1. Factor analysis by imposing a correlation structure (MSCI Barra)
2. Factor analysis by filtering the correlation structure (APT)
3. Principal component analysis
4. Random matrix theory
5. Shrinkage methods

How to denoise the covariance matrix?

- The eigendecomposition $\hat{\Sigma}$ of is

$$\hat{\Sigma} = V \Lambda V^{\top}$$

  where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues with $\lambda_1 > \lambda_2 > \ldots > \lambda_n$ and V is an orthonormal matrix
- The endogenous factors are $\mathcal{F}_t = \Lambda^{-1/2} V^{\top} R_t$
- By considering only the m first components, we can build an estimation of Σ with less noise

How to denoise the covariance matrix?

Choice of m

1. We keep factors that explain more than 1/n of asset variance:

$$m = \sup \left\{ i : \lambda_i \geq \left(\lambda_1 + \ldots + \lambda_n \right) / n \right\}$$

2. Laloux *et al.* (1999) propose to use the random matrix theory (RMT)
   1. The maximum eigenvalue of a random matrix M is equal to:

   $$\lambda_{\max} = \sigma^2 \left(1 + n/T + 2\sqrt{n/T} \right)$$

   where T is the sample size
   2. We keep the first m factors such that:

   $$m = \sup \left\{ i : \lambda_i > \lambda_{\max} \right\}$$

How to denoise the covariance matrix?

Shrinkage methods

- $\hat{\Sigma}$ is an unbiased estimator, but its convergence is very slow
- $\hat{\Phi}$ is a biased estimator that converges more quickly

Ledoit and Wolf (2003) propose to combine $\hat{\Sigma}$ and $\hat{\Phi}$:

$$\hat{\Sigma}_\alpha = \alpha \hat{\Phi} + (1 - \alpha) \hat{\Sigma}$$

The value of α is estimated by minimizing a quadratic loss:

$$\alpha^\star = \arg\min \mathbb{E}\left[\left\| \alpha \hat{\Phi} + (1 - \alpha) \hat{\Sigma} - \Sigma \right\|^2\right]$$

They find an analytical expression of $\alpha^\star$ when:

- $\hat{\Phi}$ has a constant correlation structure
- $\hat{\Phi}$ corresponds to a factor model or is deduced from PCA

How to denoise the covariance matrix?

Example 7 (equity correlation matrix)

We consider a universe with eight equity indices: S&P 500, Eurostoxx, FTSE 100, Topix, Bovespa, RTS, Nifty and HSI. The study period is January 2005–December 2011 and we use weekly returns.

The empirical correlation matrix is:

$$\hat{C} = \left(\begin{array}{cccc|cccc}
1.00 & & & & & & & \\
0.88 & 1.00 & & & & & & \\
0.88 & 0.94 & 1.00 & & & & & \\
0.64 & 0.68 & 0.65 & 1.00 & & & & \\
\hline
0.77 & 0.76 & 0.78 & 0.61 & 1.00 & & & \\
0.56 & 0.61 & 0.61 & 0.50 & 0.64 & 1.00 & & \\
0.53 & 0.61 & 0.57 & 0.53 & 0.60 & 0.57 & 1.00 & \\
0.64 & 0.68 & 0.67 & 0.68 & 0.68 & 0.60 & 0.66 & 1.00
\end{array}\right)$$

How to denoise the covariance matrix?

- Uniform correlation

$$\hat{\rho} = 66.24\%$$

- One common factor + two specific factors

$$\hat{C} = \left(\begin{array}{cccc|cccc} 1.00 & & & & & & & \\ 0.77 & 1.00 & & & & & & \\ 0.77 & 0.77 & 1.00 & & & & & \\ 0.77 & 0.77 & 0.77 & 1.00 & & & & \\ \hline 0.50 & 0.50 & 0.50 & 0.50 & 1.00 & & & \\ 0.50 & 0.50 & 0.50 & 0.50 & 0.59 & 1.00 & & \\ 0.50 & 0.50 & 0.50 & 0.50 & 0.59 & 0.59 & 1.00 & \\ 0.50 & 0.50 & 0.50 & 0.50 & 0.59 & 0.59 & 0.59 & 1.00 \end{array}\right)$$

How to denoise the covariance matrix?

- Two-linear factor model

$$\hat{C} = \left(\begin{array}{cccc|cccc}
1.00 & & & & & & & \\
0.88 & 1.00 & & & & & & \\
0.88 & 0.94 & 1.00 & & & & & \\
0.63 & 0.67 & 0.66 & 1.00 & & & & \\
\hline
0.73 & 0.78 & 0.78 & 0.63 & 1.00 & & & \\
0.58 & 0.62 & 0.60 & 0.54 & 0.59 & 1.00 & & \\
0.56 & 0.59 & 0.58 & 0.56 & 0.60 & 0.54 & 1.00 & \\
0.64 & 0.68 & 0.66 & 0.65 & 0.69 & 0.62 & 0.67 & 1.00
\end{array}\right)$$

How to denoise the covariance matrix?

- RMT estimation

$$\hat{C} = \left(\begin{array}{cccc|cccc}
1.00 & & & & & & & \\
0.73 & 1.00 & & & & & & \\
0.72 & 0.76 & 1.00 & & & & & \\
0.61 & 0.64 & 0.64 & 1.00 & & & & \\
\hline
0.72 & 0.76 & 0.75 & 0.64 & 1.00 & & & \\
0.71 & 0.75 & 0.74 & 0.63 & 0.74 & 1.00 & & \\
0.63 & 0.66 & 0.65 & 0.56 & 0.66 & 0.65 & 1.00 & \\
0.68 & 0.72 & 0.71 & 0.60 & 0.71 & 0.70 & 0.62 & 1.00
\end{array}\right)$$

How to denoise the covariance matrix?

- Ledoit-Wolf shrinkage estimation (constant correlation matrix)

$$\hat{C} = \left(\begin{array}{cccc|cccc}
1.00 & & & & & & & \\
0.77 & 1.00 & & & & & & \\
0.77 & 0.80 & 1.00 & & & & & \\
0.65 & 0.67 & 0.65 & 1.00 & & & & \\
\hline
0.72 & 0.71 & 0.72 & 0.63 & 1.00 & & & \\
0.61 & 0.64 & 0.63 & 0.58 & 0.65 & 1.00 & & \\
0.60 & 0.64 & 0.62 & 0.60 & 0.63 & 0.62 & 1.00 & \\
0.65 & 0.67 & 0.67 & 0.67 & 0.67 & 0.63 & 0.66 & 1.00
\end{array}\right)$$

- We obtain:

$$\alpha^{\star} = 51.2\%$$

- What does this result become in the case of a multi-asset-class universe?

$$\alpha^{\star} \simeq 0$$

Why standard regularization techniques are not sufficient

Optimized portfolios are solutions of the following quadratic program:

$$
\begin{aligned}
x^{\star}(\gamma) &= \arg\min \frac{1}{2} x^{\top} \Sigma x - \gamma x^{\top} \mu \\
\text{u.c.} &\quad \left\{ \begin{array}{l} \mathbf{1}_n^{\top} x = 1 \\ x \in \mathbb{R}^n \end{array} \right.
\end{aligned}
$$

We have:

$$
x^{\star}(\gamma) = \frac{\Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^{\top}\Sigma^{-1}\mathbf{1}_n} + \gamma \cdot \frac{\left(\mathbf{1}_n^{\top}\Sigma^{-1}\mathbf{1}_n\right)\Sigma^{-1}\mu - \left(\mathbf{1}_n^{\top}\Sigma^{-1}\mu\right)\Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^{\top}\Sigma^{-1}\mathbf{1}_n}
$$

Why standard regularization techniques are not sufficient

Optimal solutions are of the following form:

$$x^\star \propto f\left(\Sigma^{-1}\right)$$

The important quantity is then the precision matrix $\mathcal{I} = \Sigma^{-1}$, not the covariance matrix Σ

Why standard regularization techniques are not sufficient

- For the covariance matrix Σ, we have:

$$\Sigma = V \Lambda V^\top$$

where $V^{-1} = V^\top$ and $\Lambda = (\lambda_1, \ldots, \lambda_n)$ with $\lambda_1 \geq \ldots \geq \lambda_n$ the ordered eigenvalues

- The decomposition for the precisions matrix is

$$\mathcal{I} = U \Delta U^\top$$

- We have:

$$\begin{aligned} \Sigma^{-1} &= \left(V \Lambda V^\top\right)^{-1} \\ &= \left(V^\top\right)^{-1} \Lambda^{-1} V^{-1} \\ &= V \Lambda^{-1} V^\top \end{aligned}$$

- We deduce that $U = V$ and $\delta_i = 1/\lambda_{n-i+1}$

Why standard regularization techniques are not sufficient

Remark

The eigenvectors of the precision matrix are the same as those of the covariance matrix, but the eigenvalues of the precision matrix are the inverse of the eigenvalues of the covariance matrix. This means that the risk factors are the same, but they are in the reverse order

Why standard regularization techniques are not sufficient

Example 8

We consider a universe of 3 assets, where $\mu_1 = \mu_2 = 8\%$, $\mu_3 = 5\%$, $\sigma_1 = 20\%$, $\sigma_2 = 21\%$, $\sigma_3 = 10\%$ and $\rho_{i,j} = 80\%$.

The eigendecomposition of the covariance and precision matrices is:

	Covariance matrix Σ			Information matrix $\mathcal{I}$		
Asset / Factor	1	2	3	1	2	3
1	65.35%	−72.29%	−22.43%	−22.43%	−72.29%	65.35%
2	69.38%	69.06%	−20.43%	−20.43%	69.06%	69.38%
3	30.26%	−2.21%	95.29%	95.29%	−2.21%	30.26%
Eigenvalue	8.31%	0.84%	0.26%	379.97	119.18	12.04
% cumulated	88.29%	97.20%	100.00%	74.33%	97.65%	100.00%

$\Rightarrow$ It means that the first factor of the information matrix corresponds to the last factor of the covariance matrix and that the last factor of the information matrix corresponds to the first factor.

$\Rightarrow$ Optimization on arbitrage risk factors, idiosyncratic risk factors and (certainly) noise factors!

Why standard regularization techniques are not sufficient

Example 9

We consider a universe of 6 assets. The volatilities are respectively equal to 20%, 21%, 17%, 24%, 20% and 16%. For the correlation matrix, we have:

$$\rho = \begin{pmatrix} 1.00 & & & & & \\ 0.40 & 1.00 & & & & \\ 0.40 & 0.40 & 1.00 & & & \\ 0.50 & 0.50 & 0.50 & 1.00 & & \\ 0.50 & 0.50 & 0.50 & 0.60 & 1.00 & \\ 0.50 & 0.50 & 0.50 & 0.60 & 0.60 & 1.00 \end{pmatrix}$$

$\Rightarrow$ We compute the minimum variance (MV) portfolio with a shortsale constraint

Why standard regularization techniques are not sufficient

Table 9: Effect of deleting a PCA factor

$x^\star$	MV	$\lambda_1 = 0$	$\lambda_2 = 0$	$\lambda_3 = 0$	$\lambda_4 = 0$	$\lambda_5 = 0$	$\lambda_6 = 0$
$x_1^\star$	15.29	15.77	20.79	27.98	0.00	13.40	0.00
$x_2^\star$	10.98	16.92	1.46	12.31	0.00	8.86	0.00
$x_3^\star$	34.40	12.68	35.76	28.24	52.73	53.38	2.58
$x_4^\star$	0.00	22.88	0.00	0.00	0.00	0.00	0.00
$x_5^\star$	1.01	17.99	2.42	0.00	15.93	0.00	0.00
$x_6^\star$	38.32	13.76	39.57	31.48	31.34	24.36	97.42

Why standard regularization techniques are not sufficient

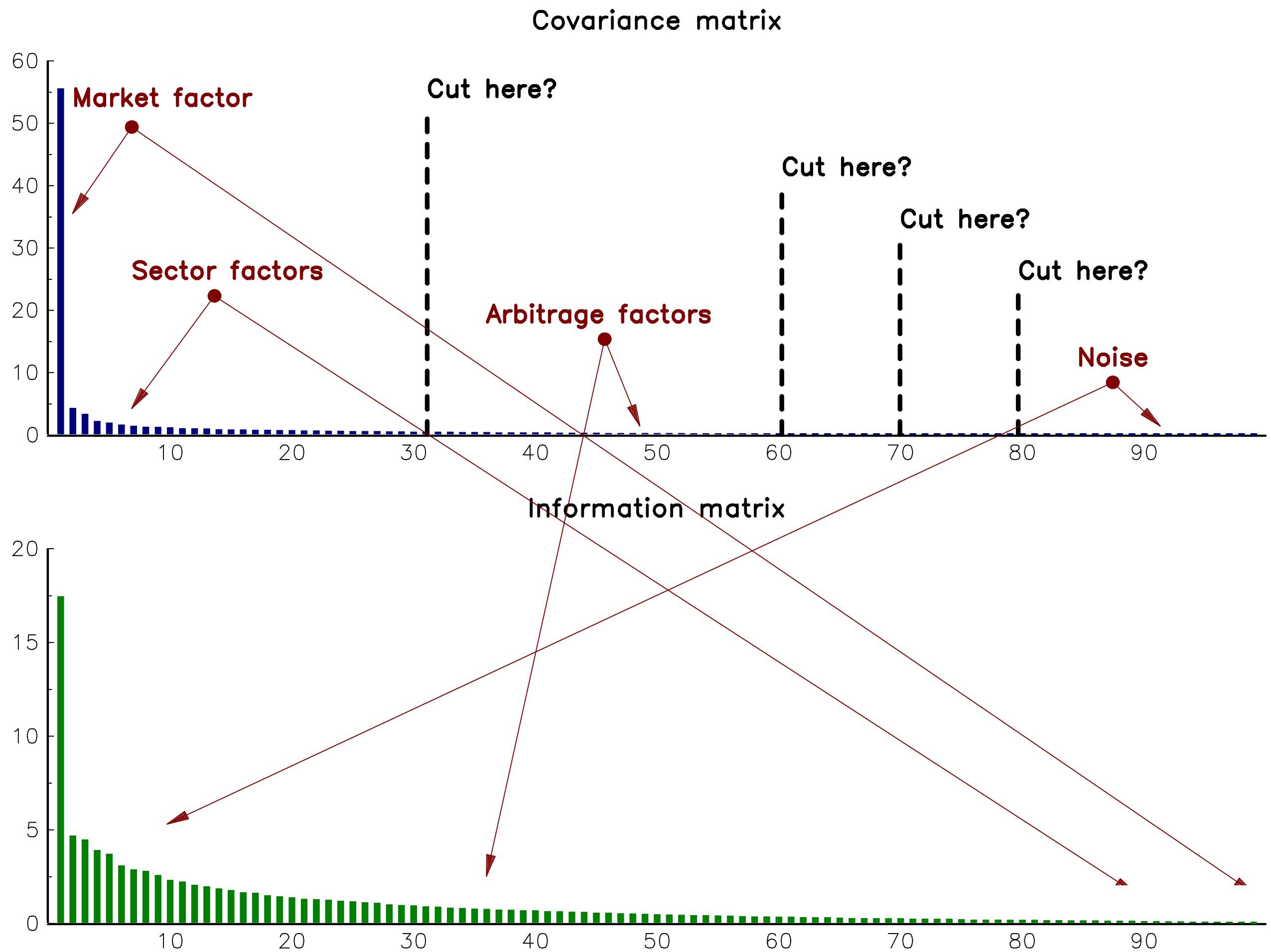


Figure 14: PCA applied to the stocks of the FTSE index (June 2012)

Arbitrage factors, hedging factors or risk factors

We consider the following linear regression model:

$$R_{i,t} = \beta_0 + \beta_i^\top R_t^{(-i)} + \varepsilon_{i,t}$$

- $R_t^{(-i)}$ denotes the vector of asset returns R_t excluding the i^{th} asset
- $\varepsilon_{i,t} \sim \mathcal{N}(0, s_i^2)$
- $\mathcal{R}_i^2$ is the R-squared of the linear regression

Precision matrix

Stevens (1998) shows that the precision matrix is given by:

$$\mathcal{I}_{i,i} = \frac{1}{\hat{\sigma}_i^2 \left(1 - \mathcal{R}_i^2\right)} \text{ and } \mathcal{I}_{i,j} = -\frac{\hat{\beta}_{i,j}}{\hat{\sigma}_i^2 \left(1 - \mathcal{R}_i^2\right)} = -\frac{\hat{\beta}_{j,i}}{\hat{\sigma}_j^2 \left(1 - \mathcal{R}_j^2\right)}$$

Arbitrage factors, hedging factors or risk factors

Example 10

We consider a universe of four assets. The expected returns are $\hat{\mu}_1 = 7\%$, $\hat{\mu}_2 = 8\%$, $\hat{\mu}_3 = 9\%$ and $\hat{\mu}_4 = 10\%$ whereas the volatilities are equal to $\hat{\sigma}_1 = 15\%$, $\hat{\sigma}_2 = 18\%$, $\hat{\sigma}_3 = 20\%$ and $\hat{\sigma}_4 = 25\%$. The correlation matrix is the following:

$$\hat{C} = \begin{pmatrix} 1.00 & & & \\ 0.50 & 1.00 & & \\ 0.50 & 0.50 & 1.00 & \\ 0.60 & 0.50 & 0.40 & 1.00 \end{pmatrix}$$

We do not impose that the sum of weights are equal to 100%

Arbitrage factors, hedging factors or risk factors

Table 10: Hedging portfolios when $\rho_{3,4} = 40\%$

Asset	$\hat{\beta}_i$				$\mathcal{R}_i^2$	$\hat{s}_i$	$\bar{\mu}_i$	$x^\star$
1		0.139	0.187	0.250	45.83%	11.04%	1.70%	69.80%
2	0.230		0.268	0.191	37.77%	14.20%	2.06%	51.18%
3	0.409	0.354		0.045	33.52%	16.31%	2.85%	53.66%
4	0.750	0.347	0.063		41.50%	19.12%	1.41%	19.28%

Table 11: Hedging portfolios when $\rho_{3,4} = 95\%$

Asset	$\hat{\beta}_i$				$\mathcal{R}_i^2$	$\hat{s}_i$	$\bar{\mu}_i$	$x^\star$
1		0.244	−0.595	0.724	47.41%	10.88%	3.16%	133.45%
2	0.443		0.470	−0.157	33.70%	14.66%	2.23%	52.01%
3	−0.174	0.076		0.795	91.34%	5.89%	1.66%	239.34%
4	0.292	−0.035	1.094		92.38%	6.90%	−1.61%	−168.67%

Arbitrage factors, hedging factors or risk factors

Table 12: Hedging portfolios (in %) at the end of 2006

	SPX	SX5E	TPX	RTY	EM	US HY	EMBI	EUR	JPY	GSCI
SPX		58.6	6.0	150.3	-30.8	-0.5	5.0	-7.3	15.3	-25.5
SX5E	9.0		-1.2	-1.3	35.2	0.8	3.2	-4.5	-5.0	-1.5
TPX	0.4	-0.6		-2.4	38.1	1.1	-3.5	-4.9	-0.8	-0.3
RTY	48.6	-2.7	-10.4		26.2	-0.6	1.9	0.2	-6.4	5.6
EM	-4.1	30.9	69.2	10.9		0.9	4.6	9.1	3.9	33.1
US HY	-5.0	53.5	160.0	-18.8	69.5		95.6	48.4	31.4	-211.7
EMBI	10.8	44.2	-102.1	12.3	73.4	19.4		-5.8	40.5	86.2
EUR	-3.6	-14.7	-33.4	0.3	33.8	2.3	-1.4		56.7	48.2
JPY	6.8	-14.5	-4.8	-8.8	12.7	1.3	8.4	50.4		-33.2
GSCI	-1.1	-0.4	-0.2	0.8	10.7	-0.9	1.8	4.2	-3.3	
$\hat{s}_i$	0.3	0.7	0.9	0.5	0.7	0.1	0.2	0.4	0.4	1.2
$\mathcal{R}_i^2$	83.0	47.7	34.9	82.4	60.9	39.8	51.6	42.3	43.7	12.1

Source: Bruder *et al.* (2013)

Arbitrage factors, hedging factors or risk factors

We finally obtain:

$$x_i^\star(\gamma) = \gamma \frac{\mu_i - \hat{\beta}_i^\top \mu^{(-i)}}{\hat{s}_i^2}$$

From this equation, we deduce the following conclusions:

1. The better the hedge, the higher the exposure. This is why highly correlated assets produces unstable MVO portfolios
2. The long/short position is defined by the sign of $\mu_i - \hat{\beta}_i^\top \mu^{(-i)}$. If the expected return of the asset is lower than the conditional expected return of the hedging portfolio, the weight is negative

Markowitz diversification	$\neq$	**Diversification of risk factors**
	$=$	**Concentration on arbitrage factors**

QP problem

We use the following formulation of the QP problem:

$$
\begin{aligned}
x^{\star} &= \arg\min \frac{1}{2} x^{\top} Q x - x^{\top} R \\
\text{u.c.} & \left\{
\begin{array}{l}
Ax = B \\
Cx \leq D \\
x^{-} \leq x \leq x^{+}
\end{array}
\right.
\end{aligned}
$$

Standard constraints

- γ-problem

$$\arg\min \frac{1}{2} x^\top \Sigma x - \gamma x^\top (\mu - r\mathbf{1}_n) \Rightarrow \left\{ \begin{array}{l} Q = \Sigma \\ R = \gamma \mu \end{array} \right.$$

- Full allocation

$$\mathbf{1}_n^\top x = 1 \Rightarrow \left\{ \begin{array}{l} A = \mathbf{1}_n^\top \\ B = 1 \end{array} \right.$$

- No short selling

$$x_i \geq 0 \Rightarrow x^- = \mathbf{0}_n$$

- Cash neutral (and portfolio optimization with unfunded strategies)

$$\mathbf{1}_n^\top x = 0 \Rightarrow \left\{ \begin{array}{l} A = \mathbf{1}_n^\top \\ B = 0 \end{array} \right.$$

Asset class constraints

Example 11

We consider a multi-asset universe of eight asset classes represented by the following indices:

- four equity indices: S&P 500, Eurostoxx, Topix, MSCI EM
- two bond indices: EGBI, US BIG
- two alternatives indices: GSCI, EPRA

The portfolio manager wants the following exposures:

- at least 50% bonds
- less than 10% commodities
- Emerging market equities cannot represent more than one third of the total exposure on equities

Asset class constraints

The constraints are then expressed as follows:

$$\left\{\begin{array}{l} x_5 + x_6 \geq 50\% \\ x_7 \leq 10\% \\ x_4 \leq \frac{1}{3}\left(x_1 + x_2 + x_3 + x_4\right) \end{array}\right.$$

The corresponding formulation $Cx \leq D$ of the QP problem is:

$$\left(\begin{array}{cccccccc} 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1/3 & -1/3 & -1/3 & 2/3 & 0 & 0 & 0 & 0 \end{array}\right)\left(\begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{array}\right) \leq \left(\begin{array}{c} -0.50 \\ 0.10 \\ 0.00 \end{array}\right)$$

Non-standard constraints (turnover management)

- We want to limit the turnover of the long-only optimized portfolio with respect to a current portfolio x^0:

$$\Omega = \left\{ x \in [0,1]^n : \sum_{i=1}^{n} \left| x_i - x_i^0 \right| \leq \tau^+ \right\}$$

  where τ^+ is the maximum turnover
- Scherer (2007) proposes to introduce some additional variables x_i^- and x_i^+ such that:

$$x_i = x_i^0 + \Delta x_i^+ - \Delta x_i^-$$

  with $\Delta x_i^- \geq 0$ and $\Delta x_i^+ \geq 0$
- Δx_i^+ indicates a positive weight change with respect to the initial weight x_i^0
- Δx_i^- indicates a negative weight change with respect to the initial weight x_i^0

Non-standard constraints (turnover management)

- The expression of the turnover becomes:

$$\sum_{i=1}^{n} \left| x_i - x_i^0 \right| = \sum_{i=1}^{n} \left| \Delta x_i^+ - \Delta x_i^- \right| = \sum_{i=1}^{n} \Delta x_i^+ + \sum_{i=1}^{n} \Delta x_i^-$$

- We obtain the following γ-problem:

$$\begin{aligned} x^\star & = & \arg\min \frac{1}{2} x^\top \Sigma x - \gamma x^\top \mu \\ & & \text{u.c.} \quad \left\{ \begin{array}{l} \sum_{i=1}^{n} x_i = 1 \\ x_i = x_i^0 + \Delta x_i^+ - \Delta x_i^- \\ \sum_{i=1}^{n} \Delta x_i^+ + \sum_{i=1}^{n} \Delta x_i^- \leq \tau^+ \\ 0 \leq x_i \leq 1 \\ 0 \leq \Delta x_i^- \leq 1 \\ 0 \leq \Delta x_i^+ \leq 1 \end{array} \right. \end{aligned}$$

Non-standard constraints (turnover management)

We obtain an augmented QP problem of dimension $3n$ instead of n:

$$
\begin{aligned}
X^\star & = \arg\min \frac{1}{2} X^\top Q X - X^\top R \\
\text{u.c.} & \left\{
\begin{array}{l}
AX = B \\
CX \leq D \\
\mathbf{0}_{3n} \leq X \leq \mathbf{1}_{3n}
\end{array}
\right.
\end{aligned}
$$

where X is a $3n \times 1$ vector:

$$
X = \left(x_1, \ldots, x_n, \Delta x_1^-, \ldots, \Delta x_n^-, \Delta x_1^+, \ldots, \Delta x_n^+ \right)
$$

Non-standard constraints (turnover management)

The augmented QP matrices are:

$$Q_{3n\times 3n} = \begin{pmatrix} \Sigma & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \end{pmatrix}, \quad R_{3n\times 1} = \begin{pmatrix} \gamma\mu \\ \mathbf{0}_n \\ \mathbf{0}_n \end{pmatrix},$$

$$A_{(n+1)\times 3n} = \begin{pmatrix} \mathbf{1}_n^\top & \mathbf{0}_n^\top & \mathbf{0}_n^\top \\ I_n & I_n & -I_n \end{pmatrix}, \quad B_{(n+1)\times 1} = \begin{pmatrix} 1 \\ x^0 \end{pmatrix},$$

$$C_{1\times 3n} = \begin{pmatrix} \mathbf{0}_n^\top & \mathbf{1}_n^\top & \mathbf{1}_n^\top \end{pmatrix} \quad \text{and} \quad D_{1\times 1} = \tau^+$$

Non-standard constraints (turnover management)

Example 12

We consider four assets. Their expected returns are equal to 5%, 6%, 8% and 6% while their volatilities are equal to 15%, 20%, 25% and 30%. The correlation matrix of asset returns is given by the following matrix:

$$\rho = \begin{pmatrix} 1.00 & & & \\ 0.10 & 1.00 & & \\ 0.40 & 0.70 & 1.00 & \\ 0.50 & 0.40 & 0.80 & 1.00 \end{pmatrix}$$

We impose that the weights are positive

- The optimal portfolio $x^\star$ for a 15% volatility target is $(45.59\%, 24.74\%, 29.67\%, 0.00\%)$
- We assume that the current portfolio x^0 is $(30\%, 45\%, 15\%, 10\%)$
- If we move directly from portfolio x^0 to portfolio $x^\star$, the turnover is equal to 60.53%

Non-standard constraints (turnover management)

Table 13: Limiting the turnover of MVO portfolios

τ^{+}	5.00	10.00	25.00	50.00	75.00	x^{0}
$x_1^{\star}$		35.00	36.40	42.34	45.59	30.00
$x_2^{\star}$		45.00	42.50	30.00	24.74	45.00
$x_3^{\star}$		15.00	21.10	27.66	29.67	15.00
$x_4^{\star}$		5.00	0.00	0.00	0.00	10.00
$\mu(x^{\star})$		5.95	6.06	6.13	6.14	6.00
$\sigma(x^{\star})$		15.00	15.00	15.00	15.00	15.69
$\tau(x^{\star} \mid x^{0})$		10.00	25.00	50.00	60.53	

Non-standard constraints (transaction cost management)

Let c_i^- and c_i^+ be the bid and ask transactions costs. The net expected return is equal to:

$$\mu(x) = \sum_{i=1}^{n} x_i \mu_i - \sum_{i=1}^{n} \Delta x_i^- c_i^- - \sum_{i=1}^{n} \Delta x_i^+ c_i^+$$

The γ-problem becomes:

$$\begin{aligned} x^\star &= \arg\min \frac{1}{2} x^\top \Sigma x - \gamma \left(\sum_{i=1}^{n} x_i \mu_i - \sum_{i=1}^{n} \Delta x_i^- c_i^- - \sum_{i=1}^{n} \Delta x_i^+ c_i^+ \right) \\ \text{u.c.} & \left\{ \begin{array}{l} \sum_{i=1}^{n} \left(x_i + \Delta x_i^- c_i^- + \Delta x_i^+ c_i^+ \right) = 1 \\ x_i = x_i^0 + \Delta x_i^+ - \Delta x_i^- \\ 0 \leq x_i \leq 1 \\ 0 \leq \Delta x_i^- \leq 1 \\ 0 \leq \Delta x_i^+ \leq 1 \end{array} \right. \end{aligned}$$

Non-standard constraints (transaction cost management)

The augmented QP problem becomes:

$$\begin{aligned} X^\star &= \arg\min \frac{1}{2} X^\top Q X - X^\top R \\ &\text{u.c.} \quad \left\{ \begin{array}{l} AX = B \\ \mathbf{0}_{3n} \leq X \leq \mathbf{1}_{3n} \end{array} \right. \end{aligned}$$

where X is a $3n \times 1$ vector:

$$X = \left(x_1, \ldots, x_n, \Delta x_1^-, \ldots, \Delta x_n^-, \Delta x_1^+, \ldots, \Delta x_n^+\right)$$

and:

$$\begin{aligned} Q_{3n\times 3n} &= \left(\begin{array}{ccc} \Sigma & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \end{array} \right), \quad R_{3n\times 1} = \left(\begin{array}{c} \gamma\mu \\ -c^- \\ -c^+ \end{array} \right), \\ A_{(n+1)\times 3n} &= \left(\begin{array}{ccc} \mathbf{1}_n^\top & \left(c^-\right)^\top & \left(c^+\right)^\top \\ I_n & I_n & -I_n \end{array} \right) \quad \text{and} \quad B_{(n+1)\times 1} = \left(\begin{array}{c} 1 \\ x^0 \end{array} \right) \end{aligned}$$

Index sampling

Index sampling

The underlying idea is to replicate an index b with n stocks by a portfolio x with n_x stocks and $n_x \ll n$

From a mathematical point of view, index sampling can be written as a portfolio optimization problem with a benchmark:

$$
\begin{aligned}
x^{\star} & = \arg\min \frac{1}{2} (x - b)^{\top} \Sigma (x - b) \\
\text{u.c.} & \left\{ \begin{array}{l} \mathbf{1}_n^{\top} x = 1 \\ x \geq \mathbf{0}_n \\ \sum_{i=1}^{n} \mathbb{1} \{ x_i > 0 \} \leq n_x \end{array} \right.
\end{aligned}
$$

where b is the vector of index weights

We obtain a mixed integer non-linear optimization problem

Index sampling

Three stepwise algorithms:

1. The backward elimination algorithm starts with all the stocks, computes the optimized portfolio, deletes the stock which presents the highest tracking error variance, and repeats this process until the number of stocks in the optimized portfolio reaches the target value n_x
2. The forward selection algorithm starts with no stocks in the portfolio, adds the stock which presents the smallest tracking error variance, and repeats this process until the number of stocks in the optimized portfolio reaches the target value n_x
3. The heuristic algorithm is a variant of the backward elimination algorithm, but the elimination process of the heuristic algorithm uses the criterion of the smallest weight

Heuristic algorithm

1. The algorithm is initialized with $\mathcal{N}_{(0)} = \emptyset$ and $x^{\star}_{(0)} = b$.
2. At the iteration k, we define a set $\mathcal{I}_{(k)}$ of stocks having the smallest positive weights in the portfolio $x^{\star}_{(k-1)}$. We then update the set $\mathcal{N}_{(k)}$ with $\mathcal{N}_{(k)} = \mathcal{N}_{(k-1)} \cup \mathcal{I}_{(k)}$ and define the upper bounds $x^{+}_{(k)}$:

$$x^{+}_{(k),i} = \left\{ \begin{array}{lll} 0 & \text{if} & i \in \mathcal{N}_{(k)} \\ 1 & \text{if} & i \notin \mathcal{N}_{(k)} \end{array} \right.$$

3. We solve the QP problem by using the new upper bounds $x^{+}_{(k)}$:

$$\begin{aligned} x^{\star}_{(k)} &= \arg\min \frac{1}{2} \left(x_{(k)} - b\right)^{\top} \Sigma \left(x_{(k)} - b\right) \\ & \text{u.c.} \quad \left\{ \begin{array}{l} \mathbf{1}_n^{\top} x_{(k)} = 1 \\ \mathbf{0}_n \leq x_{(k)} \leq x^{+}_{(k)} \end{array} \right. \end{aligned}$$

4. We iterate steps 2 and 3 until the convergence criterion:

$$\sum_{i=1}^{n} \mathbb{1}\left\{ x^{*}_{(k),i} > 0 \right\} \leq n_x$$

Complexity of the three numerical algorithms

The number of solved QP problems is respectively equal to:

- $n_b - n_x$ for the heuristic algorithm
- $(n_b - n_x)(n_b + n_x + 1)/2$ for the backward elimination algorithm
- $n_x(2n_b - n_x + 1)/2$ for the forward selection algorithm

		Number of solved QP problems		
n_b	n_x	Heuristic	Backward	Forward
50	10	40	1 220	455
	40	10	455	1 220
500	50	450	123 975	23 775
	450	50	23 775	123 975
1 500	100	1 400	1 120 700	145 050
	1 000	500	625 250	1 000 500

Index sampling (Eurostoxx 50, June 2012)

Table 14: Sampling the SX5E index with the heuristic algorithm

k	Stock	b_i	$\sigma\left(x_{(k)} \mid b\right)$
1	Nokia	0.45	0.18
2	Carrefour	0.60	0.23
3	Repsol	0.71	0.28
4	Unibail-Rodamco	0.99	0.30
5	Muenchener Rueckver	1.34	0.32
6	RWE	1.18	0.36
7	Koninklijke Philips	1.07	0.41
8	Generali	1.06	0.45
9	CRH	0.82	0.51
10	Volkswagen	1.34	0.55
42	LVMH	2.39	3.67
43	Telefonica	3.08	3.81
44	Bayer	3.51	4.33
45	Vinci	1.46	5.02
46	BBVA	2.13	6.53
47	Sanofi	5.38	7.26
48	Allianz	2.67	10.76
49	Total	5.89	12.83
50	Siemens	4.36	30.33

Index sampling (Eurostoxx 50, June 2012)

Table 15: Sampling the SX5E index with the backward elimination algorithm

k	Stock	b_i	$\sigma\left(x_{(k)} \mid b\right)$
1	Iberdrola	1.05	0.11
2	France Telecom	1.48	0.18
3	Carrefour	0.60	0.22
4	Muenchener Rueckver	1.34	0.26
5	Repsol	0.71	0.30
6	BMW	1.37	0.34
7	Generali	1.06	0.37
8	RWE	1.18	0.41
9	Koninklijke Philips	1.07	0.44
10	Air Liquide	2.10	0.48
42	GDF Suez	1.92	3.49
43	Bayer	3.51	3.88
44	BNP Paribas	2.26	4.42
45	Total	5.89	4.99
46	LVMH	2.39	5.74
47	Allianz	2.67	7.15
48	Sanofi	5.38	8.90
49	BBVA	2.13	12.83
50	Siemens	4.36	30.33

Index sampling (Eurostoxx 50, June 2012)

Table 16: Sampling the SX5E index with the forward selection algorithm

k	Stock	b_i	$\sigma\left(x_{(k)} \mid b\right)$
1	Siemens	4.36	12.83
2	Banco Santander	3.65	8.86
3	Bayer	3.51	6.92
4	Eni	3.32	5.98
5	Allianz	2.67	5.11
6	LVMH	2.39	4.55
7	France Telecom	1.48	3.93
8	Carrefour	0.60	3.62
9	BMW	1.37	3.35
41	Société Générale	1.07	0.50
42	CRH	0.82	0.45
43	Air Liquide	2.10	0.41
44	RWE	1.18	0.37
45	Nokia	0.45	0.33
46	Unibail-Rodamco	0.99	0.28
47	Repsol	0.71	0.24
48	Essilor	1.17	0.18
49	Muenchener Rueckver	1.34	0.11
50	Iberdrola	1.05	0.00

Index sampling

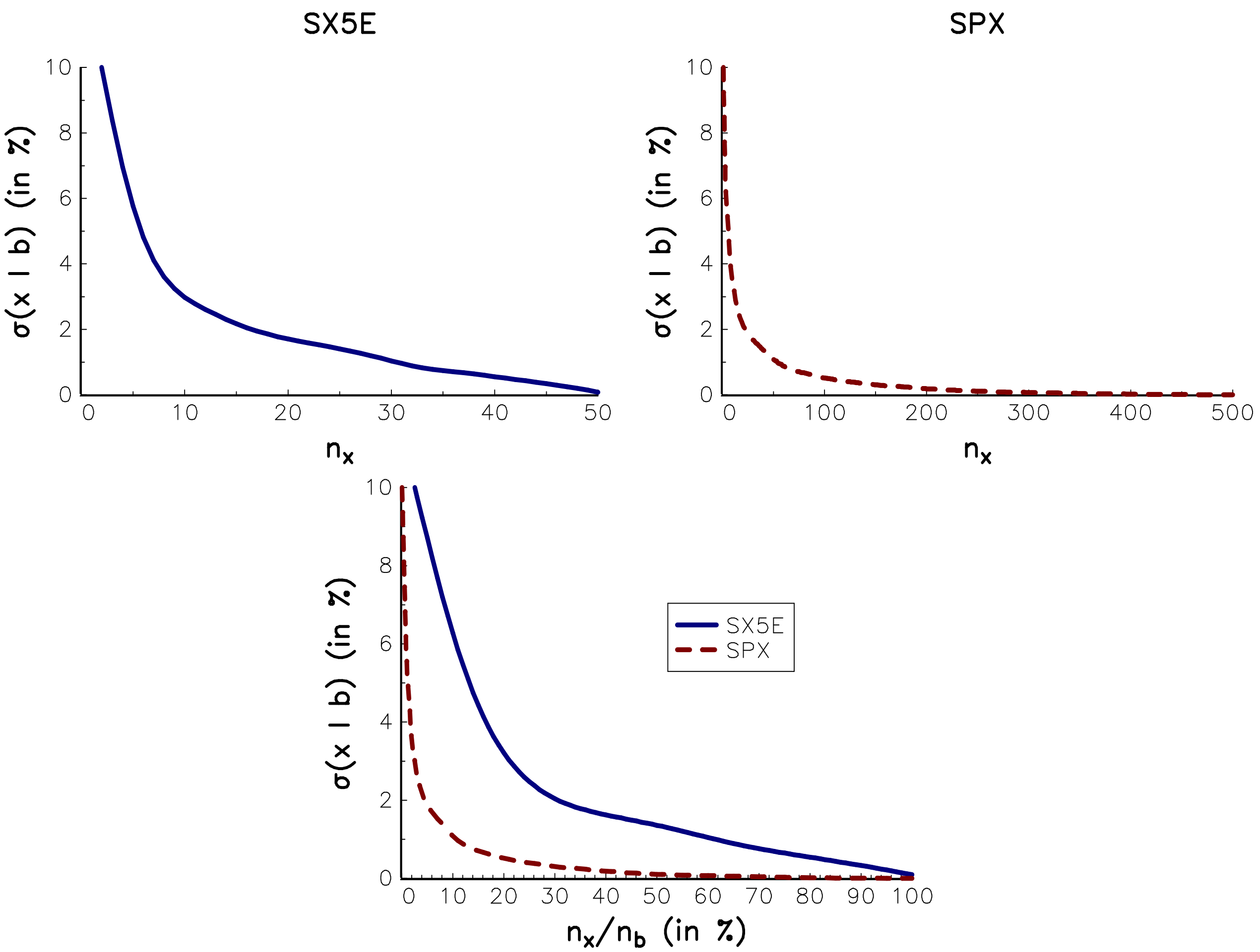


Figure 15: Sampling the SX5E and SPX indices (June 2012)

The impact of weight constraints

We specify the optimization problem as follows:

$$\min \frac{1}{2} x^\top \Sigma x$$
$$\text{u.c.} \left\{ \begin{array}{l} \mathbf{1}_n^\top x = 1 \\ \mu^\top x \geq \mu^\star \\ x \in \mathcal{C} \end{array} \right.$$

where $\mathcal{C}$ is the set of weights constraints. We define:

- the unconstrained portfolio $x^\star$ or $x^\star(\mu, \Sigma)$:

$$\mathcal{C} = \mathbb{R}^n$$

- the constrained portfolio $\tilde{x}$:

$$\mathcal{C}\left(x^-, x^+\right) = \left\{x \in \mathbb{R}^n : x_i^- \leq x_i \leq x_i^+\right\}$$

The impact of weight constraints

Theorem

Jagannathan and Ma (2003) show that the constrained portfolio is the solution of the unconstrained problem:

$$\tilde{x} = x^{\star}\left(\tilde{\mu}, \tilde{\Sigma}\right)$$

with:

$$\left\{ \begin{array}{l} \tilde{\mu} = \mu \\ \tilde{\Sigma} = \Sigma + \left(\lambda^{+} - \lambda^{-}\right) \mathbf{1}_n^{\top} + \mathbf{1}_n \left(\lambda^{+} - \lambda^{-}\right)^{\top} \end{array} \right.$$

where λ^{-} and λ^{+} are the Lagrange coefficients vectors associated to the lower and upper bounds.

$\Rightarrow$ Introducing weights constraints is equivalent to introduce a shrinkage method or to introduce some relative views (similar to the Black-Litterman approach).

The impact of weight constraints

Proof (step 1)

Without weight constraints, the expression of the Lagrangian is:

$$\mathcal{L}\left(x;\lambda_0,\lambda_1\right)=\frac{1}{2}x^\top\Sigma x-\lambda_0\left(\mathbf{1}_n^\top x-1\right)-\lambda_1\left(\mu^\top x-\mu^\star\right)$$

with $\lambda_0\geq 0$ and $\lambda_1\geq 0$. The first-order conditions are:

$$\left\{\begin{array}{l}\Sigma x-\lambda_0\mathbf{1}_n-\lambda_1\mu=\mathbf{0}_n\\ \mathbf{1}_n^\top x-1=0\\ \mu^\top x-\mu^\star=0\end{array}\right.$$

We deduce that the solution $x^\star$ depends on the vector of expected return μ and the covariance matrix Σ and we note $x^\star=x^\star\left(\mu,\Sigma\right)$

The impact of weight constraints

Proof (step 2)

If we impose now the weight constraints $\mathcal{C}\left(x^-, x^+\right)$, we have:

$$\begin{aligned} \mathcal{L}\left(x;\lambda_0,\lambda_1,\lambda^-,\lambda^+\right) &= \frac{1}{2}x^\top \Sigma x - \lambda_0\left(\mathbf{1}_n^\top x - 1\right) - \lambda_1\left(\mu^\top x - \mu^\star\right) - \\ & \lambda^{-\top}\left(x - x^-\right) - \lambda^{+\top}\left(x^+ - x\right) \end{aligned}$$

with $\lambda_0 \geq 0$, $\lambda_1 \geq 0$, $\lambda_i^- \geq 0$ and $\lambda_i^+ \geq 0$. In this case, the Kuhn-Tucker conditions are:

$$\left\{\begin{array}{l} \Sigma x - \lambda_0 \mathbf{1}_n - \lambda_1 \mu - \lambda^- + \lambda^+ = \mathbf{0}_n \\ \mathbf{1}_n^\top x - 1 = 0 \\ \mu^\top x - \mu^\star = 0 \\ \min\left(\lambda_i^-, x_i - x_i^-\right) = 0 \\ \min\left(\lambda_i^+, x_i^+ - x_i\right) = 0 \end{array}\right.$$

The impact of weight constraints

Proof (step 3)

Given a constrained portfolio $\tilde{x}$, it is possible to find a covariance matrix $\tilde{\Sigma}$ such that $\tilde{x}$ is the solution of unconstrained mean-variance portfolio. Let $\mathcal{E} = \left\{\tilde{\Sigma} > 0 : \tilde{x} = x^{\star}\left(\mu, \tilde{\Sigma}\right)\right\}$ denote the corresponding set:

$$\mathcal{E} = \left\{\tilde{\Sigma} > 0 : \tilde{\Sigma}\tilde{x} - \lambda_0 \mathbf{1}_n - \lambda_1 \mu = \mathbf{0}_n\right\}$$

Of course, the set $\mathcal{E}$ contains several solutions. From a financial point of view, we are interested in covariance matrices $\tilde{\Sigma}$ that are close to Σ. Jagannathan and Ma note that the matrix $\tilde{\Sigma}$ defined by:

$$\tilde{\Sigma} = \Sigma + \left(\lambda^+ - \lambda^-\right)\mathbf{1}_n^\top + \mathbf{1}_n\left(\lambda^+ - \lambda^-\right)^\top$$

is a solution of $\mathcal{E}$

The impact of weight constraints

Proof (step 4)

Indeed, we have:

$$\begin{aligned}
\tilde{\Sigma}\tilde{x} &= \Sigma\tilde{x} + \left(\lambda^{+} - \lambda^{-}\right) \mathbf{1}_n^\top \tilde{x} + \mathbf{1}_n \left(\lambda^{+} - \lambda^{-}\right)^\top \tilde{x} \\
&= \Sigma\tilde{x} + \left(\lambda^{+} - \lambda^{-}\right) + \mathbf{1}_n \left(\lambda^{+} - \lambda^{-}\right)^\top \tilde{x} \\
&= \lambda_0 \mathbf{1}_n + \lambda_1 \mu + \mathbf{1}_n \left(\lambda_0 \mathbf{1}_n + \lambda_1 \mu - \Sigma\tilde{x}\right)^\top \tilde{x} \\
&= \lambda_0 \mathbf{1}_n + \lambda_1 \mu + \mathbf{1}_n \left(\lambda_0 + \lambda_1 \mu^\star - \tilde{x}^\top \Sigma \tilde{x}\right) \\
&= \left(2\lambda_0 - \tilde{x}^\top \Sigma \tilde{x} + \lambda_1 \mu^\star\right) \mathbf{1}_n + \lambda_1 \mu
\end{aligned}$$

It proves that $\tilde{x}$ is the solution of the unconstrained optimization problem. The Lagrange coefficients $\lambda_0^\star$ and $\lambda_1^\star$ for the unconstrained problem are respectively equal to $2\tilde{\lambda}_0 - \tilde{x}^\top \Sigma \tilde{x} + \tilde{\lambda}_1 \mu^\star$ and $\tilde{\lambda}_1$ where $\tilde{\lambda}_0$ and $\tilde{\lambda}_1$ are the Lagrange coefficient for the constrained problem. Moreover, $\tilde{\Sigma}$ is generally a positive definite matrix

The impact of weight constraints

Example 13

We consider four assets. Their expected returns are equal to 5%, 6%, 8% and 6% while their volatilities are equal to 15%, 20%, 25% and 30%. The correlation matrix of asset returns is given by the following matrix:

$$C = \begin{pmatrix} 1.00 & & & \\ 0.10 & 1.00 & & \\ 0.40 & 0.70 & 1.00 & \\ 0.50 & 0.40 & 0.80 & 1.00 \end{pmatrix}$$

Given these parameters, the global minimum variance portfolio is equal to:

$$x^{\star} = \begin{pmatrix} 72.742\% \\ 49.464\% \\ -20.454\% \\ -1.753\% \end{pmatrix}$$

The impact of weight constraints

Table 17: Minimum variance portfolio when $x_i \geq 10\%$

$x_i^\star$	$\tilde{x}_i$	λ_i^-	λ_i^+	$\tilde{\sigma}_i$	$\tilde{\rho}_{i,j}$			
72.742	56.195	0.000	0.000	15.00	100.00			
49.464	23.805	0.000	0.000	20.00	10.00	100.00		
−20.454	10.000	1.190	0.000	19.67	10.50	58.71	100.00	
−1.753	10.000	1.625	0.000	23.98	17.38	16.16	67.52	100.00

Table 18: Minimum variance portfolio when $10\% \leq x_i \leq 40\%$

$x_i^\star$	$\tilde{x}_i$	λ_i^-	λ_i^+	$\tilde{\sigma}_i$	$\tilde{\rho}_{i,j}$			
72.742	40.000	0.000	0.915	20.20	100.00			
49.464	40.000	0.000	0.000	20.00	30.08	100.00		
−20.454	10.000	0.915	0.000	21.02	35.32	61.48	100.00	
−1.753	10.000	1.050	0.000	26.27	39.86	25.70	73.06	100.00

The impact of weight constraints

Table 19: Mean-variance portfolio when $10\% \leq x_i \leq 40\%$ and $\mu^\star = 6\%$

$x_i^\star$	$\tilde{x}_i$	λ_i^-	λ_i^+	$\tilde{\sigma}_i$	$\tilde{\rho}_{i,j}$			
65.866	40.000	0.000	0.125	15.81	100.00			
26.670	30.000	0.000	0.000	20.00	13.44	100.00		
32.933	20.000	0.000	0.000	25.00	41.11	70.00	100.00	
−25.470	10.000	1.460	0.000	24.66	23.47	19.06	73.65	100.00

Table 20: MSR portfolio when $10\% \leq x_i \leq 40\%$

$x_i^\star$	$\tilde{x}_i$	λ_i^-	λ_i^+	$\tilde{\sigma}_i$	$\tilde{\rho}_{i,j}$			
51.197	40.000	0.000	0.342	17.13	100.00			
50.784	39.377	0.000	0.000	20.00	18.75	100.00		
−21.800	10.000	0.390	0.000	23.39	36.25	66.49	100.00	
19.818	10.623	0.000	0.000	30.00	50.44	40.00	79.96	100.00

Variations on the efficient frontier

Exercise

We consider an investment universe of four assets. We assume that their expected returns are equal to 5%, 6%, 8% and 6%, and their volatilities are equal to 15%, 20%, 25% and 30%. The correlation matrix is:

$$\rho = \begin{pmatrix} 100\% & & & \\ 10\% & 100\% & & \\ 40\% & 70\% & 100\% & \\ 50\% & 40\% & 80\% & 100\% \end{pmatrix}$$

We note x_i the weight of the i^{th} asset in the portfolio. We only impose that the sum of the weights is equal to 100%.

Variations on the efficient frontier

Question 1

Represent the efficient frontier by considering the following values of γ: -1, -0.5, -0.25, 0, 0.25, 0.5, 1 and 2.

Variations on the efficient frontier

We deduce that the covariance matrix is:

$$\Sigma = \begin{pmatrix} 2.250 & 0.300 & 1.500 & 2.250 \\ 0.300 & 4.000 & 3.500 & 2.400 \\ 1.500 & 3.500 & 6.250 & 6.000 \\ 2.250 & 2.400 & 6.000 & 9.000 \end{pmatrix} \times 10^{-2}$$

We then have to solve the γ-formulation of the Markowitz problem:

$$\begin{aligned} x^{\star}(\gamma) & = & \arg\min \frac{1}{2} x^{\top} \Sigma x - \gamma x^{\top} \mu \\ & \text{u.c.} & \mathbf{1}_n^{\top} x = 1 \end{aligned}$$

We obtain the results[1] given in Table 21. We represent the efficient frontier in Figure 16.

[1] The weights, expected returns and volatilities are expressed in %.

Variations on the efficient frontier

Table 21: Solution of Question 1

γ	−1.00	−0.50	−0.25	0.00	0.25	0.50	1.00	2.00
$x_1^\star$	94.04	83.39	78.07	72.74	67.42	62.09	51.44	30.15
$x_2^\star$	120.05	84.76	67.11	49.46	31.82	14.17	−21.13	−91.72
$x_3^\star$	−185.79	−103.12	−61.79	−20.45	20.88	62.21	144.88	310.22
$x_4^\star$	71.69	34.97	16.61	−1.75	−20.12	−38.48	−75.20	−148.65
$\mu(x^\star)$	1.34	3.10	3.98	4.86	5.74	6.62	8.38	11.90
$\sigma(x^\star)$	22.27	15.23	12.88	12.00	12.88	15.23	22.27	39.39

Variations on the efficient frontier

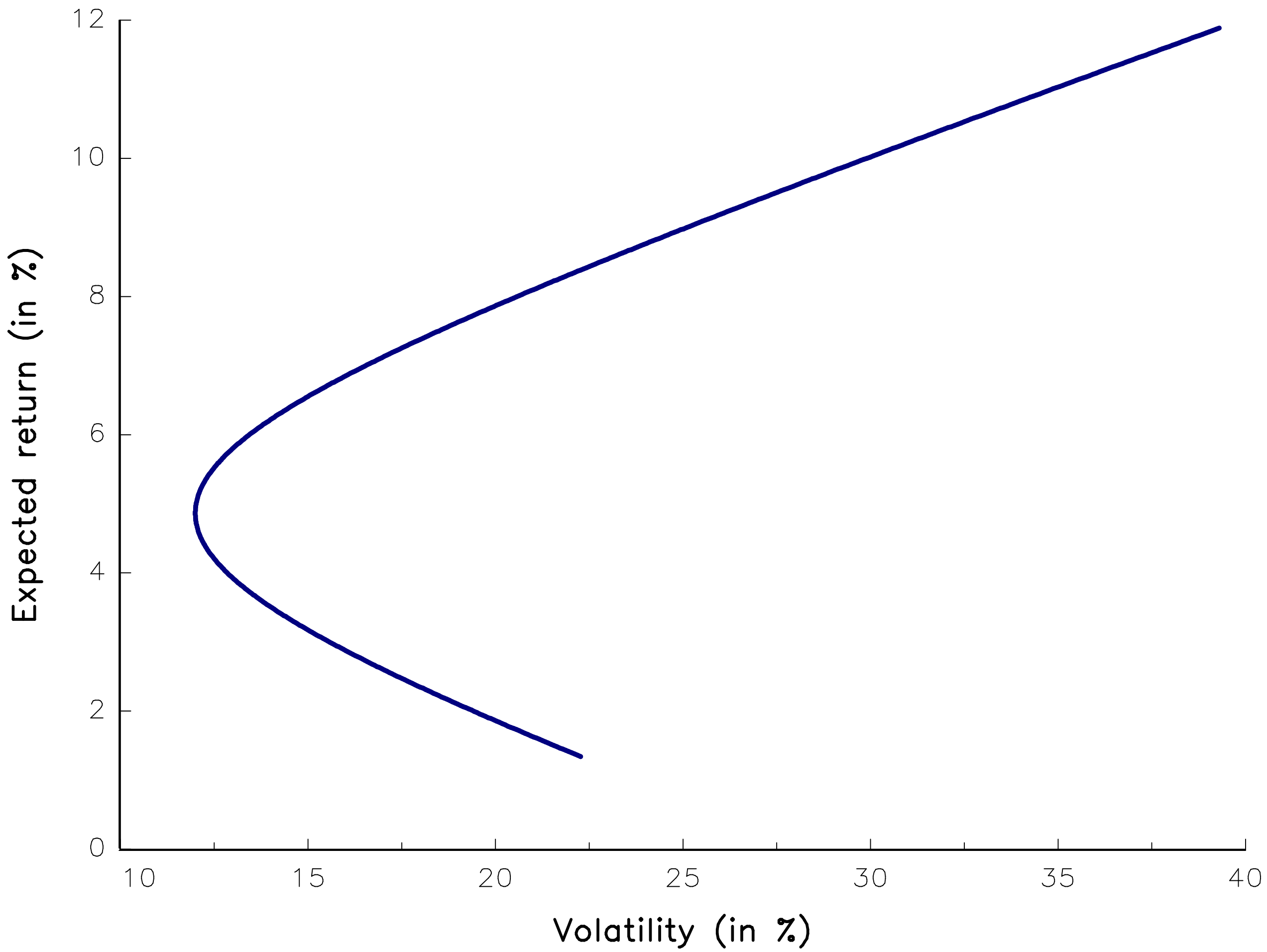


Figure 16: Markowitz efficient frontier

Variations on the efficient frontier

Question 2

Calculate the minimum variance portfolio. What are its expected return and its volatility?

Variations on the efficient frontier

We solve the γ-problem with $\gamma = 0$. The minimum variance portfolio is then $x_1^\star = 72.74\%$, $x_2^\star = 49.46\%$, $x_3^\star = -20.45\%$ and $x_4^\star = -1.75\%$. We deduce that $\mu(x^\star) = 4.86\%$ and $\sigma(x^\star) = 12.00\%$.

Variations on the efficient frontier

Question 3

Calculate the optimal portfolio which has an ex-ante volatility $\sigma^{\star}$ equal to 10%. Same question if $\sigma^{\star} = 15\%$ and $\sigma^{\star} = 20\%$.

Variations on the efficient frontier

There is no solution when the target volatility $\sigma^\star$ is equal to 10% because the minimum variance portfolio has a volatility larger than 10%. Finding the optimized portfolio for $\sigma^\star = 15\%$ or $\sigma^\star = 20\%$ is equivalent to solving a σ-problem. If $\sigma^\star = 15\%$ (resp. $\sigma^\star = 20\%$), we obtain an implied value of γ equal to 0.48 (resp. 0.85). Results are given in the following Table:

$\sigma^\star$	15.00	20.00
$x_1^\star$	62.52	54.57
$x_2^\star$	15.58	−10.75
$x_3^\star$	58.92	120.58
$x_4^\star$	−37.01	−64.41
$\mu(x^\star)$	6.55	7.87
γ	0.48	0.85

Variations on the efficient frontier

Question 4

We note $x^{(1)}$ the minimum variance portfolio and $x^{(2)}$ the optimal portfolio with $\sigma^{\star} = 20\%$. We consider the set of portfolios $x^{(\alpha)}$ defined by the relationship:

$$x^{(\alpha)} = (1 - \alpha)\, x^{(1)} + \alpha x^{(2)}$$

In the previous efficient frontier, place the portfolios $x^{(\alpha)}$ when α is equal to -0.5, -0.25, 0, 0.1, 0.2, 0.5, 0.7 and 1. What do you observe? Comment on this result.

Variations on the efficient frontier

Let $x^{(\alpha)}$ be the portfolio defined by the relationship $x^{(\alpha)} = (1-\alpha)\, x^{(1)} + \alpha x^{(2)}$ where $x^{(1)}$ is the minium variance portfolio and $x^{(2)}$ is the optimized portfolio with a 20% ex-ante volatility. We obtain the following results:

α	$\sigma\left(x^{(\alpha)}\right)$	$\mu\left(x^{(\alpha)}\right)$
−0.50	14.42	3.36
−0.25	12.64	4.11
0.00	12.00	4.86
0.10	12.10	5.16
0.20	12.41	5.46
0.50	14.42	6.36
0.70	16.41	6.97
1.00	20.00	7.87

We have reported these portfolios in Figure 17. We notice that they are located on the efficient frontier. This is perfectly normal because we know that a combination of two optimal portfolios corresponds to another optimal portfolio.

Variations on the efficient frontier

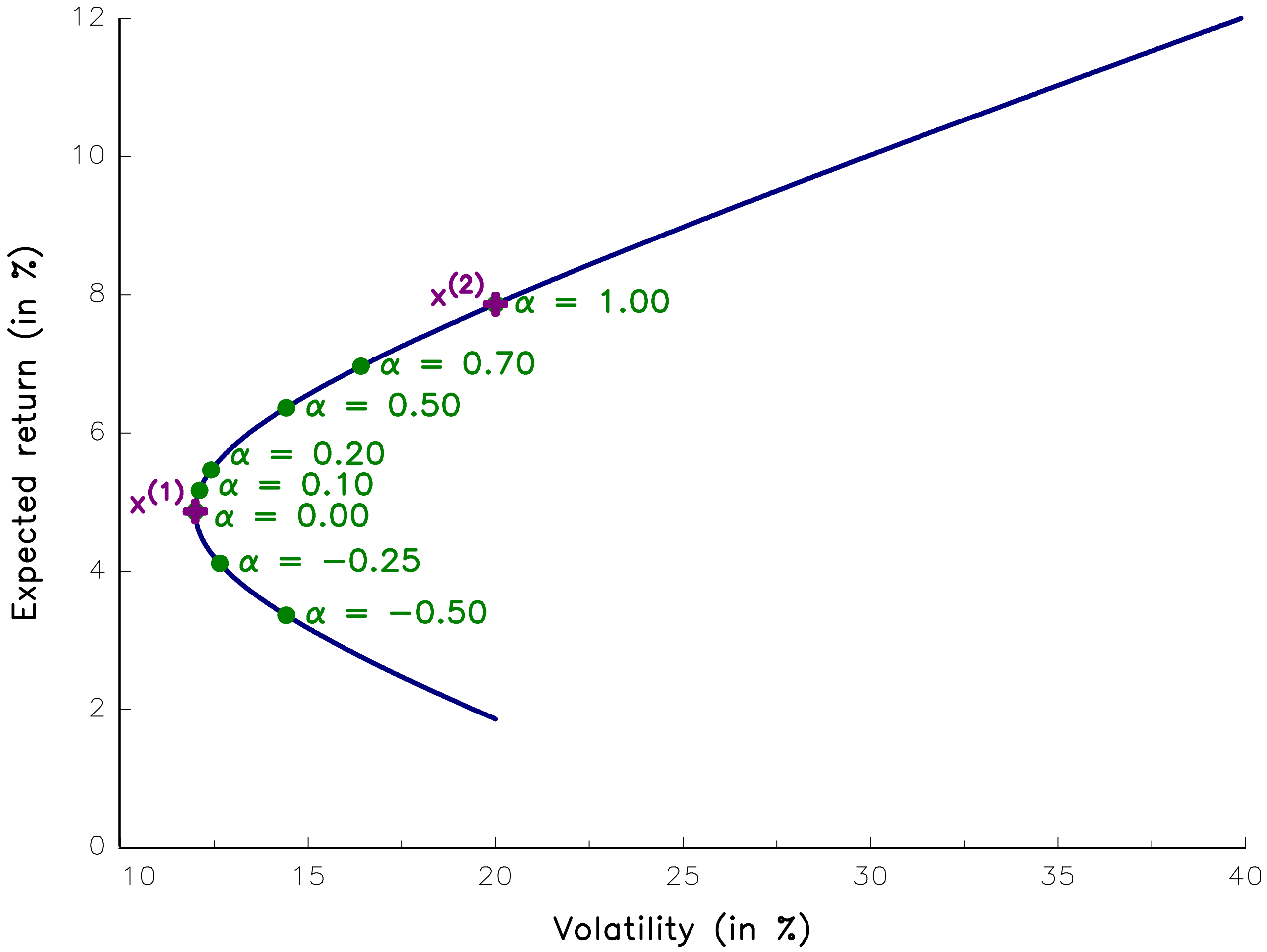


Figure 17: Mean-variance diagram of portfolios $x^{(\alpha)}$

Variations on the efficient frontier

Question 5

Repeat Questions 3 and 4 by considering the constraint $0 \leq x_i \leq 1$. Explain why we do not retrieve the same observation.

Variations on the efficient frontier

If we consider the constraint $0 \leq x_i \leq 1$, the γ-formulation of the Markowitz problem becomes:

$$\begin{aligned} x^{\star}(\gamma) &= \arg\min \frac{1}{2} x^{\top} \Sigma x - \gamma x^{\top} \mu \\ \text{u.c.} & \left\{ \begin{array}{l} \mathbf{1}_n^{\top} x = 1 \\ \mathbf{0}_n \leq x \leq \mathbf{1}_n \end{array} \right. \end{aligned}$$

Variations on the efficient frontier

We obtain the following results:

$\sigma^{\star}$	MV	12.00	15.00	20.00
$x_1^{\star}$	65.49	✓	45.59	24.88
$x_2^{\star}$	34.51	✓	24.74	4.96
$x_3^{\star}$	0.00	✓	29.67	70.15
$x_4^{\star}$	0.00	✓	0.00	0.00
$\mu(x^{\star})$	5.35	✓	6.14	7.15
$\sigma(x^{\star})$	12.56	✓	15.00	20.00
γ	0.00	✓	0.62	1.10

We observe that we cannot target a volatility $\sigma^{\star} = 10\%$. Moreover, the expected return $\mu(x^{\star})$ of the optimal portfolios are reduced due to the additional constraints.

Variations on the efficient frontier

Question 6

We now include in the investment universe a fifth asset corresponding to the risk-free asset. Its return is equal to 3%.

Variations on the efficient frontier

Question 6.a

Define the expected return vector and the covariance matrix of asset returns.

Variations on the efficient frontier

We have:

$$\mu = \begin{pmatrix} 5.0 \\ 6.0 \\ 8.0 \\ 6.0 \\ 3.0 \end{pmatrix} \times 10^{-2}$$

and:

$$\Sigma = \begin{pmatrix} 2.250 & 0.300 & 1.500 & 2.250 & 0.000 \\ 0.300 & 4.000 & 3.500 & 2.400 & 0.000 \\ 1.500 & 3.500 & 6.250 & 6.000 & 0.000 \\ 2.250 & 2.400 & 6.000 & 9.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \end{pmatrix} \times 10^{-2}$$

Variations on the efficient frontier

Question 6.b

Deduce the efficient frontier by solving directly the quadratic problem.

Variations on the efficient frontier

We solve the γ-problem and obtain the efficient frontier given in Figure 18.

Variations on the efficient frontier

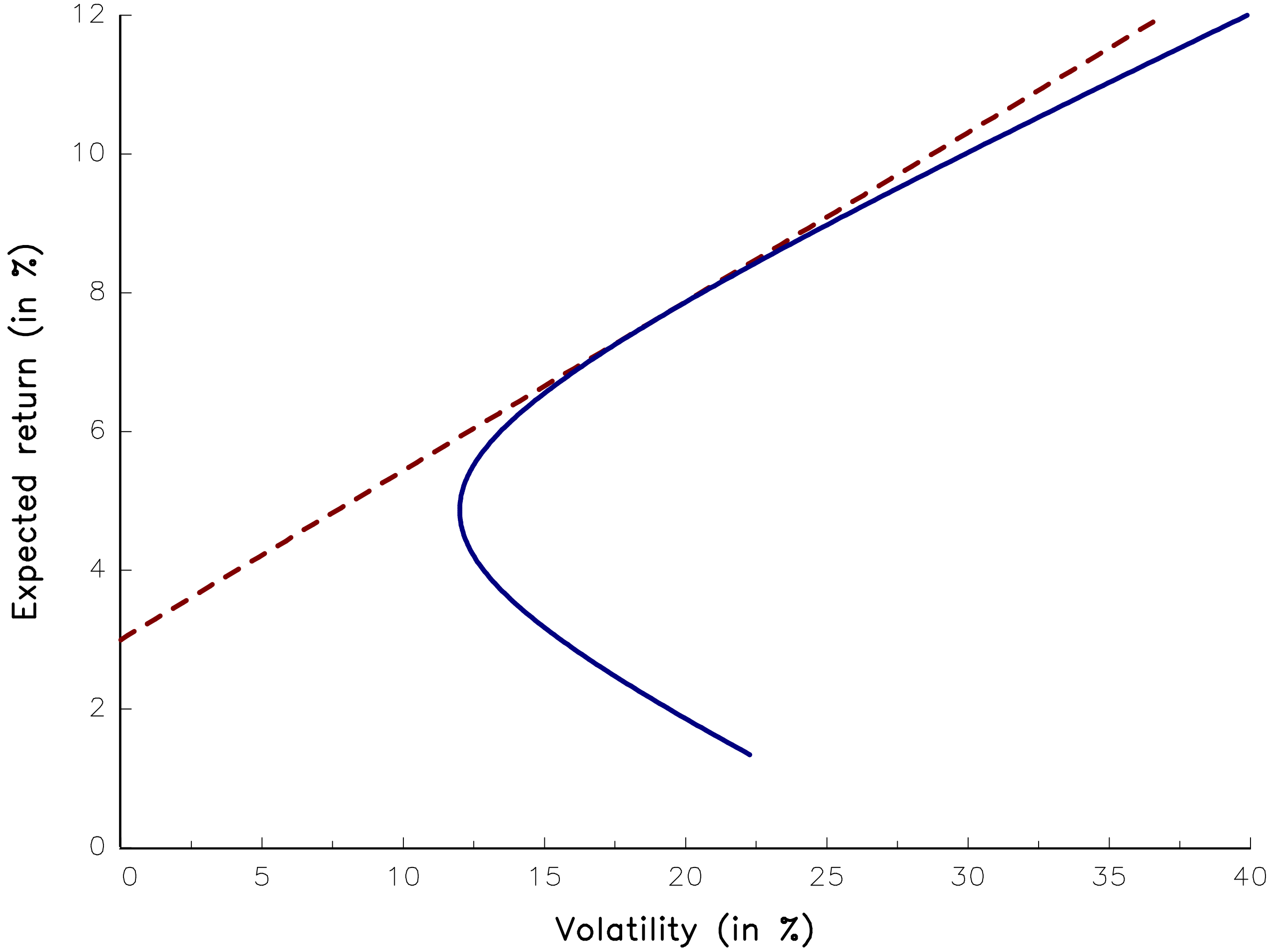


Figure 18: Efficient frontier when the risk-free asset is introduced

Variations on the efficient frontier

Question 6.c

What is the shape of the efficient frontier? Comment on this result.

Variations on the efficient frontier

This efficient frontier is a straight line. This line passes through the risk-free asset and is tangent to the efficient frontier of Figure 16. This question is a direct application of the *Separation Theorem* of Tobin.

Variations on the efficient frontier

Question 6.d

Choose two arbitrary portfolios $x^{(1)}$ and $x^{(2)}$ of this efficient frontier. Deduce the Sharpe ratio of the tangency portfolio.

Variations on the efficient frontier

We consider two optimized portfolios of this efficient frontier. They corresponds to $\gamma = 0.25$ and $\gamma = 0.50$. We obtain the following results:

γ	0.25	0.50
$x_1^\star$	18.23	36.46
$x_2^\star$	-1.63	-3.26
$x_3^\star$	34.71	69.42
$x_4^\star$	-18.93	-37.86
$x_5^\star$	67.62	35.24
$\mu(x^\star)$	4.48	5.97
$\sigma(x^\star)$	6.09	12.18

Variations on the efficient frontier

The first portfolio has an expected return equal to 4.48% and a volatility equal to 6.09%. The weight of the risk-free asset is 67.62%. This explains the low volatility of this portfolio. For the second portfolio, the weight of the risk-free asset is lower and equal to 35.24%. The expected return and the volatility are then equal to 5.97% and 12.18%. We note $x^{(1)}$ and $x^{(2)}$ these two portfolios. By definition, the Sharpe ratio of the market portfolio $x^{\star}$ is the tangency of the line. We deduce that:

$$\begin{aligned} \mathrm{SR}\left(x^{\star} \mid r\right) &= \frac{\mu\left(x^{(2)}\right) - \mu\left(x^{(1)}\right)}{\sigma\left(x^{(2)}\right) - \sigma\left(x^{(1)}\right)} \\ &= \frac{5.97 - 4.48}{12.18 - 6.09} \\ &= 0.2436 \end{aligned}$$

The Sharpe ratio of the market portfolio $x^{\star}$ is then equal to 0.2436.

Variations on the efficient frontier

Question 6.e

Calculate then the composition of the tangency portfolio from $x^{(1)}$ and $x^{(2)}$.

Variations on the efficient frontier

By construction, every portfolio $x^{(\alpha)}$ which belongs to the tangency line is a linear combination of two portfolios $x^{(1)}$ and $x^{(2)}$ of this efficient frontier:

$$x^{(\alpha)} = (1-\alpha)\, x^{(1)} + \alpha x^{(2)}$$

The market portfolio $x^{\star}$ is the portfolio $x^{(\alpha)}$ which has a zero weight in the risk-free asset. We deduce that the value $\alpha^{\star}$ which corresponds to the market portfolio satisfies the following relationship:

$$(1-\alpha^{\star})\, x_5^{(1)} + \alpha^{\star} x_5^{(2)} = 0$$

because the risk-free asset is the fifth asset of the portfolio.

Variations on the efficient frontier

It follows that:

$$\begin{aligned}\alpha^{\star} &= \frac{x_5^{(1)}}{x_5^{(1)} - x_5^{(2)}} \\ &= \frac{67.62}{67.62 - 35.24} \\ &= 2.09\end{aligned}$$

We deduce that the market portfolio is:

$$x^{\star} = (1 - 2.09) \cdot \begin{pmatrix} 18.23 \\ -1.63 \\ 34.71 \\ -18.93 \\ 67.62 \end{pmatrix} + 2.09 \cdot \begin{pmatrix} 36.46 \\ -3.26 \\ 69.42 \\ -37.86 \\ 35.24 \end{pmatrix} = \begin{pmatrix} 56.30 \\ -5.04 \\ 107.21 \\ -58.46 \\ 0.00 \end{pmatrix}$$

We check that the Sharpe ratio of this portfolio is 0.2436.

Variations on the efficient frontier

Question 7

We consider the general framework with n risky assets whose vector of expected returns is μ and the covariance matrix of asset returns is Σ while the return of the risk-free asset is r. We note $\tilde{x}$ the portfolio invested in the $n+1$ assets. We have:

$$\tilde{x} = \begin{pmatrix} x \\ x_r \end{pmatrix}$$

with x the weight vector of risky assets and x_r the weight of the risk-free asset. We impose the following constraint:

$$\sum_{i=1}^{n} \tilde{x}_i = \sum_{i=1}^{n} x_i = 1$$

Variations on the efficient frontier

Question 7.a

Define $\tilde{\mu}$ and $\tilde{\Sigma}$ the vector of expected returns and the covariance matrix of asset returns associated with the $n+1$ assets.

Variations on the efficient frontier

We have:

$$\tilde{\mu} = \begin{pmatrix} \mu \\ r \end{pmatrix}$$

and:

$$\tilde{\Sigma} = \begin{pmatrix} \Sigma & \mathbf{0}_n \\ \mathbf{0}_n^\top & 0 \end{pmatrix}$$

Variations on the efficient frontier

Question 7.b

By using the Markowitz ϕ-problem, retrieve the *Separation Theorem* of Tobin.

Variations on the efficient frontier

If we include the risk-free asset, the Markowitz ϕ-problem becomes:

$$\begin{aligned}\tilde{x}^{\star}(\phi) &= \arg\max \tilde{x}^{\top}\tilde{\mu} - \frac{\phi}{2}\tilde{x}^{\top}\tilde{\Sigma}\tilde{x} \\ \text{u.c.} &\quad \mathbf{1}_n^{\top}\tilde{x} = 1\end{aligned}$$

We note that the objective function can be written as follows:

$$\begin{aligned}f(\tilde{x}) &= \tilde{x}^{\top}\tilde{\mu} - \frac{\phi}{2}\tilde{x}^{\top}\tilde{\Sigma}\tilde{x} \\ &= x^{\top}\mu + x_r r - \frac{\phi}{2}x^{\top}\Sigma x \\ &= g(x, x_r)\end{aligned}$$

The constraint becomes $\mathbf{1}_n^{\top}x + x_r = 1$. We deduce that the Lagrange function is:

$$\mathcal{L}(x, x_r; \lambda_0) = x^{\top}\mu + x_r r - \frac{\phi}{2}x^{\top}\Sigma x - \lambda_0\left(\mathbf{1}_n^{\top}x + x_r - 1\right)$$

Variations on the efficient frontier

The first-order conditions are:

$$\left\{\begin{array}{l}\partial_x \mathcal{L}(x, x_r; \lambda_0) = \mu - \phi\Sigma x - \lambda_0 \mathbf{1}_n = \mathbf{0}_n \\ \partial_{x_r} \mathcal{L}(x, x_r; \lambda_0) = r - \lambda_0 = 0 \\ \partial_{\lambda_0} \mathcal{L}(x, x_r; \lambda_0) = \mathbf{1}_n^\top x + x_r - 1 = 0\end{array}\right.$$

The solution of the optimization problem is then:

$$\left\{\begin{array}{l}x^\star = \phi^{-1}\Sigma^{-1}(\mu - r\mathbf{1}_n) \\ \lambda_0^\star = r \\ x_r^\star = 1 - \phi^{-1}\mathbf{1}_n^\top \Sigma^{-1}(\mu - r\mathbf{1}_n)\end{array}\right.$$

Let $x_0^\star$ be the following portfolio:

$$x_0^\star = \frac{\Sigma^{-1}(\mu - r\mathbf{1}_n)}{\mathbf{1}_n^\top \Sigma^{-1}(\mu - r\mathbf{1}_n)}$$

Variations on the efficient frontier

We can then write the solution of the optimization problem in the following way:

$$\left\{\begin{array}{l} x^\star = \alpha x_0^\star \\ \lambda_0^\star = r \\ x_r^\star = 1 - \alpha \\ \alpha = \phi^{-1} \mathbf{1}_n^\top \Sigma^{-1} \left(\mu - r \mathbf{1}_n\right) \end{array}\right.$$

The first equation indicates that the relative proportions of risky assets in the optimized portfolio remain constant. If $\phi = \phi_0 = \mathbf{1}_n^\top \Sigma^{-1} \left(\mu - r \mathbf{1}_n\right)$, then $x^\star = x_0^\star$ and $x_r^\star = 0$. We deduce that $x_0^\star$ is the tangency portfolio. If $\phi \neq \phi_0$, $x^\star$ is proportional to $x_0^\star$ and the wealth invested in the risk-free asset is the complement $(1 - \alpha)$ to obtain a total exposure equal to 100%. We retrieve then the separation theorem:

$$\tilde{x}^\star = \underbrace{\alpha \cdot \left(\begin{array}{c} x_0^\star \\ 0 \end{array}\right)}_{\text{risky assets}} + \underbrace{(1 - \alpha) \cdot \left(\begin{array}{c} \mathbf{0}_n \\ 1 \end{array}\right)}_{\text{risk-free asset}}$$

Beta coefficient

Question 1

We consider an investment universe of n assets with:

$$R = \begin{pmatrix} R_1 \\ \vdots \\ R_n \end{pmatrix} \sim \mathcal{N}(\mu, \Sigma)$$

The weights of the market portfolio (or the benchmark) are $b = (b_1, \ldots, b_n)$.

Beta coefficient

Question 1.a

Define the beta β_i of asset i with respect to the market portfolio.

Beta coefficient

The beta of an asset is the ratio between its covariance with the market portfolio return and the variance of the market portfolio return. In the CAPM theory, we have:

$$\mathbb{E}[R_i] = r + \beta_i \left(\mathbb{E}[R(b)] - r\right)$$

where R_i is the return of asset i, $R(b)$ is the return of the market portfolio and r is the risk-free rate. The beta β_i of asset i is:

$$\beta_i = \frac{\operatorname{cov}(R_i, R(b))}{\operatorname{var}(R(b))}$$

Let Σ be the covariance matrix of asset returns. We have $\operatorname{cov}(R, R(b)) = \Sigma b$ and $\operatorname{var}(R(b)) = b^\top \Sigma b$. We deduce that:

$$\beta_i = \frac{(\Sigma b)_i}{b^\top \Sigma b}$$

Beta coefficient

Question 1.b

Let X_1, X_2 and X_3 be three random variables. Show that:

$$\operatorname{cov}\left(c_1 X_1 + c_2 X_2, X_3\right) = c_1 \operatorname{cov}\left(X_1, X_3\right) + c_2 \operatorname{cov}\left(X_2, X_3\right)$$

Beta coefficient

We recall that the mathematical operator $\mathbb{E}$ is bilinear. Let c be the covariance $\operatorname{cov}(c_1 X_1 + c_2 X_2, X_3)$. We then have:

$$\begin{aligned} c &= \mathbb{E}\left[\left(c_1 X_1 + c_2 X_2 - \mathbb{E}\left[c_1 X_1 + c_2 X_2\right]\right)\left(X_3 - \mathbb{E}\left[X_3\right]\right)\right] \\ &= \mathbb{E}\left[\left(c_1\left(X_1 - \mathbb{E}\left[X_1\right]\right) + c_2\left(X_2 - \mathbb{E}\left[X_2\right]\right)\right)\left(X_3 - \mathbb{E}\left[X_3\right]\right)\right] \\ &= c_1 \mathbb{E}\left[\left(X_1 - \mathbb{E}\left[X_1\right]\right)\left(X_3 - \mathbb{E}\left[X_3\right]\right)\right] + c_2 \mathbb{E}\left[\left(X_2 - \mathbb{E}\left[X_2\right]\right)\left(X_3 - \mathbb{E}\left[X_3\right]\right)\right] \\ &= c_1 \operatorname{cov}\left(X_1, X_3\right) + c_2 \operatorname{cov}\left(X_2, X_3\right) \end{aligned}$$

Beta coefficient

Question 1.c

We consider the asset portfolio $x = (x_1, \ldots, x_n)$ such that $\sum_{i=1}^{n} x_i = 1$. What is the relationship between the beta $\beta(x \mid b)$ of the portfolio and the betas β_i of the assets?

Beta coefficient

We have:

$$\begin{aligned}
\beta(x \mid b) &= \frac{\operatorname{cov}(R(x), R(b))}{\operatorname{var}(R(b))} = \frac{\operatorname{cov}(x^\top R, b^\top R)}{\operatorname{var}(b^\top R)} \\
&= \frac{x^\top \mathbb{E}\left[(R-\mu)(R-\mu)^\top\right] b}{b^\top \mathbb{E}\left[(R-\mu)(R-\mu)^\top\right] b} \\
&= \frac{x^\top \Sigma b}{b^\top \Sigma b} = x^\top \frac{\Sigma b}{b^\top \Sigma b} = x^\top \beta = \sum_{i=1}^{n} x_i \beta_i
\end{aligned}$$

with $\beta = (\beta_1, \ldots, \beta_n)$. The beta of portfolio x is then the weighted mean of asset betas. Another way to show this result is to exploit the result of Question 1.b. We have:

$$\beta(x \mid b) = \frac{\operatorname{cov}\left(\sum_{i=1}^{n} x_i R_i, R(b)\right)}{\operatorname{var}(R(b))} = \sum_{i=1}^{n} x_i \frac{\operatorname{cov}(R_i, R(b))}{\operatorname{var}(R(b))} = \sum_{i=1}^{n} x_i \beta_i$$

Beta coefficient

Question 1.d

Calculate the beta of the portfolios $x^{(1)}$ and $x^{(2)}$ with the following data:

i	1	2	3	4	5
β_i	0.7	0.9	1.1	1.3	1.5
$x_i^{(1)}$	0.5	0.5			
$x_i^{(2)}$	0.25	0.25	0.5	0.5	-0.5

Beta coefficient

We obtain $\beta\left(x^{(1)} \mid b\right) = 0.80$ and $\beta\left(x^{(2)} \mid b\right) = 0.85$.

Beta coefficient

Question 2

We assume that the market portfolio is the equally weighted portfolio[a].

[a] We have $b_i = n^{-1}$.

Beta coefficient

Question 2.a

Show that $\sum_{i=1}^{n} \beta_i = n$.

Beta coefficient

The weights of the market portfolio are then $b = n^{-1}\mathbf{1}_n$. We have:

$$\beta = \frac{\operatorname{cov}\left(R, R\left(b\right)\right)}{\operatorname{var}\left(R\left(b\right)\right)} = \frac{\Sigma b}{b^\top \Sigma b} = \frac{n^{-1}\Sigma \mathbf{1}_n}{n^{-2}\left(\mathbf{1}_n^\top \Sigma \mathbf{1}_n\right)} = n\frac{\Sigma \mathbf{1}_n}{\left(\mathbf{1}_n^\top \Sigma \mathbf{1}_n\right)}$$

We deduce that:

$$\sum_{i=1}^n \beta_i = \mathbf{1}_n^\top \beta = \mathbf{1}_n^\top n\frac{\Sigma \mathbf{1}_n}{\left(\mathbf{1}_n^\top \Sigma \mathbf{1}_n\right)} = n\frac{\mathbf{1}_n^\top \Sigma \mathbf{1}_n}{\left(\mathbf{1}_n^\top \Sigma \mathbf{1}_n\right)} = n$$

Beta coefficient

Question 2.b

We consider the case $n = 3$. Show that $\beta_1 \geq \beta_2 \geq \beta_3$ implies $\sigma_1 \geq \sigma_2 \geq \sigma_3$ if $\rho_{i,j} = 0$.

Beta coefficient

If $\rho_{i,j} = 0$, we have:

$$\beta_i = n \frac{\sigma_i^2}{\sum_{j=1}^n \sigma_j^2}$$

We deduce that:

$$\begin{aligned} \beta_1 \geq \beta_2 \geq \beta_3 \quad & \Rightarrow \quad n \frac{\sigma_1^2}{\sum_{j=1}^3 \sigma_j^2} \geq n \frac{\sigma_2^2}{\sum_{j=1}^3 \sigma_j^2} \geq n \frac{\sigma_3^2}{\sum_{j=1}^3 \sigma_j^2} \\ & \Rightarrow \quad \sigma_1^2 \geq \sigma_2^2 \geq \sigma_3^2 \\ & \Rightarrow \quad \sigma_1 \geq \sigma_2 \geq \sigma_3 \end{aligned}$$

Beta coefficient

Question 2.c

What is the result if the correlation is uniform $\rho_{i,j} = \rho$?

Beta coefficient

If $\rho_{i,j} = \rho$, it follows that:

$$\begin{aligned}
\beta_i &\propto \sigma_i^2 + \sum_{j \neq i} \rho \sigma_i \sigma_j \\
&= \sigma_i^2 + \rho \sigma_i \sum_{j \neq i} \sigma_j + \rho \sigma_i^2 - \rho \sigma_i^2 \\
&= (1 - \rho) \sigma_i^2 + \rho \sigma_i \sum_{j=1}^{n} \sigma_j \\
&= f(\sigma_i)
\end{aligned}$$

with:

$$f(z) = (1 - \rho) z^2 + \rho z \sum_{j=1}^{n} \sigma_j$$

Beta coefficient

The first derivative of $f(z)$ is:

$$f'(z) = 2(1-\rho)z + \rho \sum_{j=1}^{n} \sigma_j$$

If $\rho \geq 0$, then $f(z)$ is an increasing function for $z \geq 0$ because $(1-\rho) \geq 0$ and $\rho \sum_{j=1}^{n} \sigma_j \geq 0$. This explains why the previous result remains valid:

$$\beta_1 \geq \beta_2 \geq \beta_3 \Rightarrow \sigma_1 \geq \sigma_2 \geq \sigma_3 \quad \text{if} \quad \rho_{i,j} = \rho \geq 0$$

If $-(n-1)^{-1} \leq \rho < 0$, then f' is decreasing if $z < -2^{-1}\rho(1-\rho)^{-1} \sum_{j=1}^{n} \sigma_j$ and increasing otherwise. We then have:

$$\beta_1 \geq \beta_2 \geq \beta_3 \nRightarrow \sigma_1 \geq \sigma_2 \geq \sigma_3 \quad \text{if} \quad \rho_{i,j} = \rho < 0$$

In fact, the result remains valid in most cases. To obtain a counter-example, we must have large differences between the volatilities and a correlation close to $-(n-1)^{-1}$. For example, if $\sigma_1 = 5\%$, $\sigma_2 = 6\%$, $\sigma_3 = 80\%$ and $\rho = -49\%$, we have $\beta_1 = -0.100$, $\beta_2 = -0.115$ and $\beta_3 = 3.215$.

Beta coefficient

Question 2.d

Find a general example such that $\beta_1 > \beta_2 > \beta_3$ and $\sigma_1 < \sigma_2 < \sigma_3$.

Beta coefficient

We assume that $\sigma_1 = 15\%$, $\sigma_2 = 20\%$, $\sigma_3 = 22\%$, $\rho_{1,2} = 70\%$, $\rho_{1,3} = 20\%$ and $\rho_{2,3} = -50\%$. It follows that $\beta_1 = 1.231$, $\beta_2 = 0.958$ and $\beta_3 = 0.811$. We thus have found an example such that $\beta_1 > \beta_2 > \beta_3$ and $\sigma_1 < \sigma_2 < \sigma_3$.

Beta coefficient

Question 2.e

Do we have $\sum_{i=1}^{n} \beta_i < n$ or $\sum_{i=1}^{n} \beta_i > n$ if the market portfolio is not equally weighted?

Beta coefficient

There is no reason that we have either $\sum_{i=1}^{n} \beta_i < n$ or $\sum_{i=1}^{n} \beta_i > n$. Let us consider the previous numerical example. If $b = (5\%, 25\%, 70\%)$, we obtain $\sum_{i=1}^{3} \beta_i = 1.808$ whereas if $b = (20\%, 40\%, 40\%)$, we have $\sum_{i=1}^{3} \beta_i = 3.126$.

Beta coefficient

Question 3

We search a market portfolio $b \in \mathbb{R}^n$ such that the betas are the same for all the assets: $\beta_i = \beta_j = \beta$.

Beta coefficient

Question 3.a

Show that there is an obvious solution which satisfies $\beta = 1$.

Beta coefficient

We have:

$$\begin{aligned}\sum_{i=1}^{n} b_i \beta_i &= \sum_{i=1}^{n} b_i \frac{(\Sigma b)_i}{b^\top \Sigma b} \\ &= b^\top \frac{\Sigma b}{b^\top \Sigma b} \\ &= 1\end{aligned}$$

If $\beta_i = \beta_j = \beta$, then $\beta = 1$ is an obvious solution because the previous relationship is satisfied:

$$\sum_{i=1}^{n} b_i \beta_i = \sum_{i=1}^{n} b_i = 1$$

Beta coefficient

Question 3.b

Show that this solution is unique and corresponds to the minimum variance portfolio.

Beta coefficient

If $\beta_i = \beta_j = \beta$, then we have:

$$\sum_{i=1}^{n} b_i \beta = 1 \Leftrightarrow \beta = \frac{1}{\sum_{i=1}^{n} b_i} = 1$$

β can only take one value, the solution is then unique. We know that the marginal volatilities are the same in the case of the minimum variance portfolio x (TR-RPB, page 173):

$$\frac{\partial\, \sigma\left(x\right)}{\partial\, x_i} = \frac{\partial\, \sigma\left(x\right)}{\partial\, x_j}$$

with $\sigma\left(x\right) = \sqrt{x^\top \Sigma x}$ the volatility of the portfolio x.

Beta coefficient

It follows that:

$$\frac{(\Sigma x)_i}{\sqrt{x^\top \Sigma x}} = \frac{(\Sigma x)_j}{\sqrt{x^\top \Sigma x}}$$

By dividing the two terms by $\sqrt{x^\top \Sigma x}$, we obtain:

$$\frac{(\Sigma x)_i}{x^\top \Sigma x} = \frac{(\Sigma x)_j}{x^\top \Sigma x}$$

The asset betas are then the same in the minimum variance portfolio. Because we have:

$$\left\{ \begin{array}{l} \beta_i = \beta_j \\ \sum_{i=1}^n x_i \beta_i = 1 \end{array} \right.$$

we deduce that:

$$\beta_i = 1$$

Beta coefficient

Question 4

We assume that $b \in [0,1]^n$.

Beta coefficient

Question 4.a

Show that if one asset has a beta greater than one, there exists another asset which has a beta smaller than one.

Beta coefficient

We have:

$$\begin{aligned} & \sum_{i=1}^{n} b_i \beta_i = 1 \\ \Leftrightarrow \quad & \sum_{i=1}^{n} b_i \beta_i = \sum_{i=1}^{n} b_i \\ \Leftrightarrow \quad & \sum_{i=1}^{n} b_i \beta_i - \sum_{i=1}^{n} b_i = 0 \\ \Leftrightarrow \quad & \sum_{i=1}^{n} b_i \left(\beta_i - 1 \right) = 0 \end{aligned}$$

Beta coefficient

We obtain the following system of equations:

$$\left\{ \begin{array}{l} \sum_{i=1}^{n} b_i \left(\beta_i - 1\right) = 0 \\ b_i \geq 0 \end{array} \right.$$

Let us assume that the asset j has a beta greater than 1. We then have:

$$\left\{ \begin{array}{l} b_j \left(\beta_j - 1\right) + \sum_{i \neq j} b_i \left(\beta_i - 1\right) = 0 \\ b_i \geq 0 \end{array} \right.$$

It follows that $b_j \left(\beta_j - 1\right) > 0$ because $b_j > 0$ (otherwise the beta is zero). We must therefore have $\sum_{i \neq j} x_i \left(\beta_i - 1\right) < 0$. Because $b_i \geq 0$, it is necessary that at least one asset has a beta smaller than 1.

Beta coefficient

Question 4.b

We consider the case $n = 3$. Find a covariance matrix Σ and a market portfolio b such that one asset has a negative beta.

Beta coefficient

We use standard notations to represent Σ. We seek a portfolio such that $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_3 < 0$. To simplify this problem, we assume that the three assets have the same volatility. We also obtain the following system of inequalities:

$$\left\{ \begin{array}{l} b_1 + b_2 \rho_{1,2} + b_3 \rho_{1,3} > 0 \\ b_1 \rho_{1,2} + b_2 + b_3 \rho_{2,3} > 0 \\ b_1 \rho_{1,3} + b_2 \rho_{2,3} + b_3 < 0 \end{array} \right.$$

It is sufficient that $b_1 \rho_{1,3} + b_2 \rho_{2,3}$ is negative and b_3 is small. For example, we may consider $b_1 = 50\%$, $b_2 = 45\%$, $b_3 = 5\%$, $\rho_{1,2} = 50\%$, $\rho_{1,3} = 0\%$ and $\rho_{2,3} = -50\%$. We obtain $\beta_1 = 1.10$, $\beta_2 = 1.03$ and $\beta_3 = -0.27$.

Beta coefficient

Question 5

We report the return $R_{i,t}$ and $R_t(b)$ of asset i and market portfolio b at different dates:

t	1	2	3	4	5	6
$R_{i,t}$	−22	−11	−10	−8	13	11
$R_t(b)$	−26	−9	−10	−10	16	14
t	7	8	9	10	11	12
$R_{i,t}$	21	13	−30	−6	−5	−5
$R_t(b)$	14	15	−22	−7	−11	2
t	13	14	15	16	17	18
$R_{i,t}$	19	−17	2	−24	25	−7
$R_t(b)$	15	−15	−1	−23	15	−6

Beta coefficient

Question 5.a

Estimate the beta of the asset.

Beta coefficient

We perform the linear regression $R_{i,t} = \alpha_i + \beta_i R_t(b) + \varepsilon_{i,t}$ and we obtain $\hat{\beta}_i = 1.06$.

Beta coefficient

Question 5.b

What is the proportion of the asset volatility explained by the market?

Beta coefficient

We deduce that the contribution c_i of the market factor is (TR-RPB, page 16):

$$c_i = \frac{\beta_i^2 \operatorname{var}(R(b))}{\operatorname{var}(R_i)} = 90.62\%$$

Black-Litterman model

Exercise

We consider a universe of three assets. Their volatilities are 20%, 20% and 15%. The correlation matrix of asset returns is:

$$\rho = \begin{pmatrix} 1.00 & & \\ 0.50 & 1.00 & \\ 0.20 & 0.60 & 1.00 \end{pmatrix}$$

We would like to implement a trend-following strategy. For that, we estimate the trend of each asset and the volatility of the trend. We obtain the following results:

Asset	1	2	3
$\hat{\mu}$	10%	-5%	15%
$\sigma(\hat{\mu})$	4%	2%	10%

We assume that the neutral portfolio is the equally weighted portfolio.

Black-Litterman model

Question 1

Find the optimal portfolio if the constraint of the tracking error volatility is set to 1%, 2%, 3%, 4% and 5%.

Black-Litterman model

We consider the portfolio optimization problem in the presence of a benchmark (TR-RPB, page 17). We obtain the following results (expressed in %):

$\sigma(x^\star \mid b)$	1.00	2.00	3.00	4.00	5.00
$x_1^\star$	35.15	36.97	38.78	40.60	42.42
$x_2^\star$	26.32	19.30	12.28	5.26	-1.76
$x_3^\star$	38.53	43.74	48.94	54.14	59.34
$\mu(x^\star \mid b)$	1.31	2.63	3.94	5.25	6.56

Black-Litterman model

Question 2

In order to tilt the neutral portfolio, we now consider the Black-Litterman model. The risk-free rate is set to 0.

Black-Litterman model

Question 2.a

Find the implied risk premium of the assets if we target a Sharpe ratio equal to 0.50. What is the value of ϕ?

Black-Litterman model

Let b be the benchmark (that is the equally weighted portfolio). We recall that the implied risk aversion parameter is:

$$\phi = \frac{\mathrm{SR}\left(b \mid r\right)}{\sqrt{b^\top \Sigma b}}$$

and the implied risk premium is:

$$\tilde{\mu} = r + \mathrm{SR}\left(b \mid r\right) \frac{\Sigma b}{\sqrt{b^\top \Sigma b}}$$

We obtain $\phi = 3.4367$ and:

$$\tilde{\mu} = \begin{pmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \\ \tilde{\mu}_3 \end{pmatrix} = \begin{pmatrix} 7.56\% \\ 8.94\% \\ 5.33\% \end{pmatrix}$$

Black-Litterman model

Question 2.b

How does one incorporate a trend-following strategy in the Black-Litterman model? Give the P, Q and Ω matrices.

Black-Litterman model

In this case, the views of the portfolio manager corresponds to the trends observed in the market. We then have[2]:

$$\begin{aligned} P &= I_3 \\ Q &= \hat{\mu} \\ \Omega &= \operatorname{diag}\left(\sigma^2\left(\hat{\mu}_1\right), \ldots, \sigma^2\left(\hat{\mu}_n\right)\right) \end{aligned}$$

The views $P\mu = Q + \varepsilon$ become:

$$\mu = \hat{\mu} + \varepsilon$$

with $\varepsilon \sim \mathcal{N}\left(\mathbf{0}_3, \Omega\right)$.

[2] If we suppose that the trends are not correlated.

Black-Litterman model

Question 2.c

Calculate the conditional expectation $\bar{\mu} = \mathbb{E}\left[\mu \mid P\mu = Q + \varepsilon\right]$ if we assume that $\Gamma = \tau\Sigma$ and $\tau = 0.01$.

Black-Litterman model

We have (TR-RPB, page 25):

$$\begin{aligned}\bar{\mu} &= E\left[\mu \mid P\mu = Q + \varepsilon\right] \\ &= \tilde{\mu} + \Gamma P^{\top}\left(P\Gamma P^{\top} + \Omega\right)^{-1}\left(Q - P\tilde{\mu}\right) \\ &= \tilde{\mu} + \tau\Sigma\left(\tau\Sigma + \Omega\right)^{-1}\left(\hat{\mu} - \tilde{\mu}\right)\end{aligned}$$

We obtain:

$$\bar{\mu} = \begin{pmatrix} \bar{\mu}_1 \\ \bar{\mu}_2 \\ \bar{\mu}_3 \end{pmatrix} = \begin{pmatrix} 5.16\% \\ 2.38\% \\ 2.47\% \end{pmatrix}$$

Black-Litterman model

Question 2.d

Find the Black-Litterman optimized portfolio.

Black-Litterman model

We optimize the quadratic utility function with $\phi = 3.4367$. The Black-Litterman portfolio is then:

$$x^{\star} = \begin{pmatrix} x_1^{\star} \\ x_2^{\star} \\ x_3^{\star} \end{pmatrix} = \begin{pmatrix} 56.81\% \\ -23.61\% \\ 66.80\% \end{pmatrix}$$

Its volatility tracking error is $\sigma\left(x^{\star} \mid b\right) = 8.02\%$ and its alpha is $\mu\left(x^{\star} \mid b\right) = 10.21\%$.

Black-Litterman model

Question 3

We would like to compute the Black-Litterman optimized portfolio, corresponding to a 3% tracking error volatility.

Black-Litterman model

Question 3.a

What is the Black-Litterman portfolio when $\tau = 0$ and $\tau = +\infty$?

Black-Litterman model

- If $\tau = 0$, $\bar{\mu} = \tilde{\mu}$. The BL portfolio x is then equal to the neutral portfolio b.
- We also have:

$$\begin{aligned} \lim_{\tau \to \infty} \bar{\mu} &= \tilde{\mu} + \lim_{\tau \to \infty} \tau \Sigma^{\top} \left(\tau \Sigma + \Omega \right)^{-1} \left(\hat{\mu} - \tilde{\mu} \right) \\ &= \tilde{\mu} + \left(\hat{\mu} - \tilde{\mu} \right) \\ &= \hat{\mu} \end{aligned}$$

In this case, $\bar{\mu}$ is independent from the implied risk premium $\hat{\mu}$ and is exactly equal to the estimated trends $\hat{\mu}$. The BL portfolio x is then the Markowitz optimized portfolio with the given value of ϕ.

Black-Litterman model

Question 3.b

Using the previous results, apply the bisection algorithm and find the Black-Litterman optimized portfolio, which corresponds to a 3% tracking error volatility.

Black-Litterman model

We would like to find the BL portfolio such that $\sigma(x \mid b) = 3\%$. We know that $\sigma(x \mid b) = 0$ if $\tau = 0$. Thanks to Question 2.d, we also know that $\sigma(x \mid b) = 8.02\%$ if $\tau = 1\%$. It implies that the optimal portfolio corresponds to a specific value of τ which is between 0 and 1%. If we apply the bi-section algorithm, we find that:

$$\tau^\star = 0.242\%$$

. The composition of the optimal portfolio is then

$$x^\star = \begin{pmatrix} x_1^\star \\ x_2^\star \\ x_3^\star \end{pmatrix} = \begin{pmatrix} 41.18\% \\ 11.96\% \\ 46.85\% \end{pmatrix}$$

We obtain an alpha equal to 3.88%, which is a little bit smaller than the alpha of 3.94% obtained for the TE portfolio.

Black-Litterman model

Question 3.c

Compare the relationship between $\sigma(x \mid b)$ and $\mu(x \mid b)$ of the Black-Litterman model with the one of the tracking error model. Comment on these results.

Black-Litterman model

We have reported the relationship between $\sigma(x \mid b)$ and $\mu(x \mid b)$ in Figure 19. We notice that the information ratio of BL portfolios is very close to the information ratio of TE portfolios. We may explain that because of the homogeneity of the estimated trends $\hat{\mu}_i$ and the volatilities $\sigma(\hat{\mu}_i)$. If we suppose that $\sigma(\hat{\mu}_1) = 1\%$, $\sigma(\hat{\mu}_2) = 5\%$ and $\sigma(\hat{\mu}_3) = 15\%$, we obtain the relationship $\#2$. In this case, the BL model produces a smaller information ratio than the TE model. We explain this because $\bar{\mu}$ is the right measure of expected return for the BL model whereas it is $\hat{\mu}$ for the TE model. We deduce that the ratios $\bar{\mu}_i / \hat{\mu}_i$ are more volatile for the parameter set $\#2$, in particular when τ is small.

Black-Litterman model

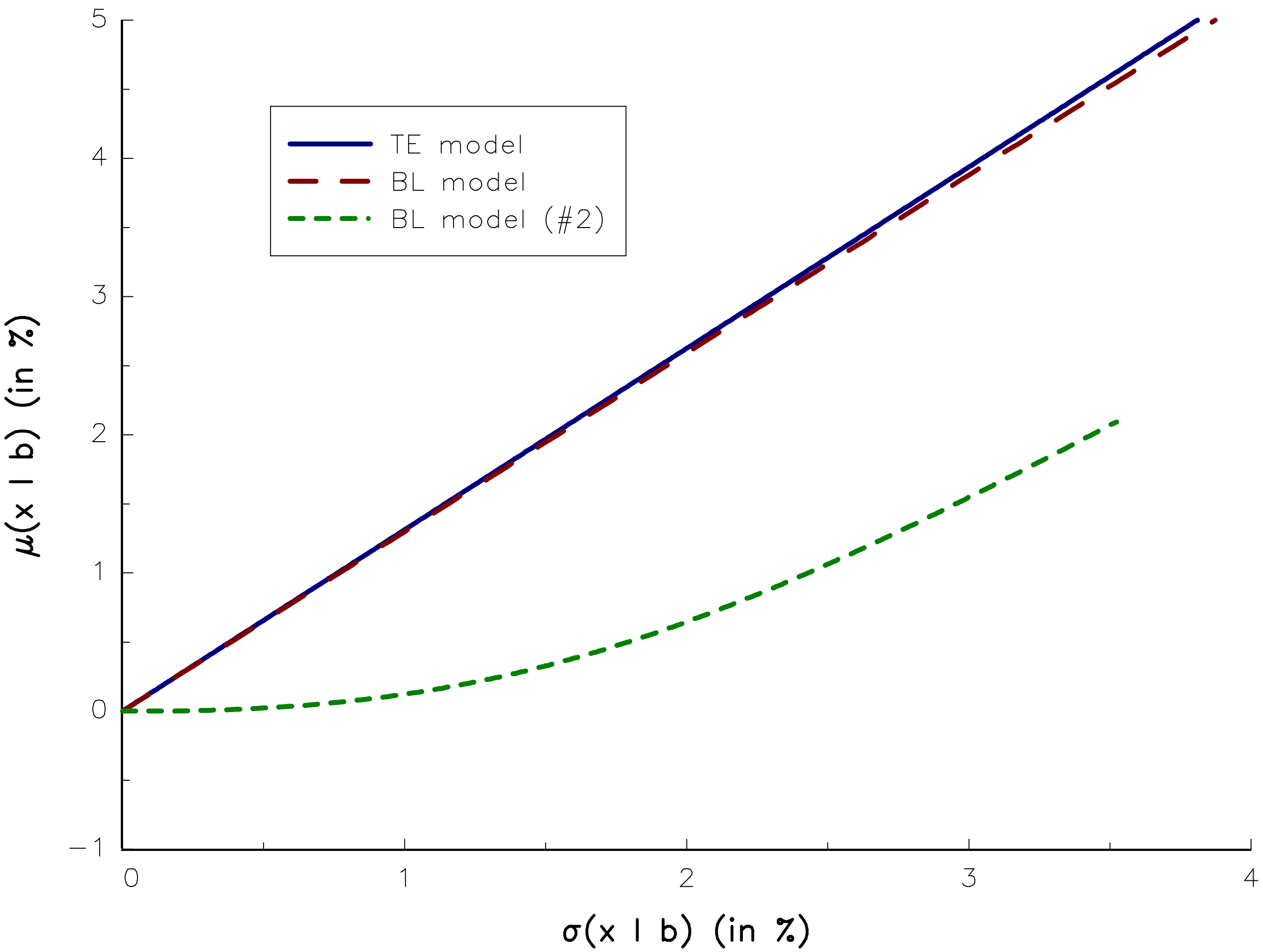


Figure 19: Efficient frontier of TE and BL portfolios

Main references

- RONCALLI, T. (2013)
  Introduction to Risk Parity and Budgeting, Chapman and Hall/CRC Financial Mathematics Series, Chapter 1.
- RONCALLI, T. (2013)
  Introduction to Risk Parity and Budgeting — Companion Book, Chapman and Hall/CRC Financial Mathematics Series, Chapter 1.

References I

- BLACK, F. and LITTERMAN, R.B. (1992)
  Global Portfolio Optimization, *Financial Analysts Journal*, 48(5), pp. 28-43.
- BOURGERON, T., LEZMI, E., and RONCALLI, T. (2018)
  Robust Asset Allocation for Robo-Advisors, *arXiv*, 1902.05710, `https://arxiv.org/abs/1902.07449`.
- BRUDER, B., GAUSSEL, N., RICHARD, J.C., and RONCALLI, T. (2013)
  Regularization of Portfolio Allocation, *SSRN*, `www.ssrn.com/abstract=2767358`.
- JAGANNATHAN, R. and MA, T. (2003)
  Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps, *Journal of Finance*, 58(4), pp. 1651-1684.

References II

Laloux, L., Cizeau, P., Bouchaud, J-P. and Potters, M. (1999)
Noise Dressing of Financial Correlation Matrices, *Physical Review Letters*, 83(7), pp. 1467-1470.*

Ledoit, O. and Wolf, M. (2003)
Improved Estimation of the Covariance Matrix of Stock Returns With an Application to Portfolio Selection, *Journal of Empirical Finance*, 10(5), pp. 603-621.

Ledoit, O. and Wolf, M. (2004)
Honey, I Shrunk the Sample Covariance Matrix, *Journal of Portfolio Management*, 30(4), pp. 110-119.

Markowitz, H. (1952)
Portfolio Selection, *Journal of Finance*, 7(1), pp. 77-91.

References III

- Michaud, R.O. (1989)
  The Markowitz Optimization Enigma: Is Optimized Optimal?, *Financial Analysts Journal*, 45(1), pp. 31-42.

- Scherer, B. (2007)
  Portfolio Construction & Risk Budgeting, Third edition, Risk Books.

- Sharpe, W.F. (1964)
  Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance*, 19(3), pp. 425-442.

- Stevens, G.V.G. (1998)
  On the Inverse of the Covariance Matrix in Portfolio analysis, *Journal of Finance*, 53(5), pp. 1821-1827.

- Tobin, J. (1958)
  Liquidity Preference as Behavior Towards Risk, *Review of Economic Studies*, 25(2), pp. 65-86.

Asset Management
Lecture 2. Risk Budgeting

Thierry Roncalli*

*University of Paris-Saclay

January 2021

Agenda

- Lecture 1: Portfolio Optimization
- **Lecture 2: Risk Budgeting**
- Lecture 3: Smart Beta, Factor Investing and Alternative Risk Premia
- Lecture 4: Green and Sustainable Finance, ESG Investing and Climate Risk
- Lecture 5: Machine Learning in Asset Management

Portfolio optimization & portfolio diversification

Example 1

- We consider an investment universe of 5 assets
- (μ_i, σ_i) are respectively equal to $(8\%, 12\%)$, $(7\%, 10\%)$, $(7.5\%, 11\%)$, $(8.5\%, 13\%)$ and $(8\%, 12\%)$
- The correlation matrix is $\mathcal{C}_5(\rho)$ with $\rho = 60\%$

The optimal portfolio $x^\star$ such that $\sigma(x^\star) = 10\%$ is equal to:

$$x^\star = \begin{pmatrix} 23.97\% \\ 6.42\% \\ 16.91\% \\ 28.73\% \\ 23.97\% \end{pmatrix}$$

Portfolio optimization & portfolio diversification

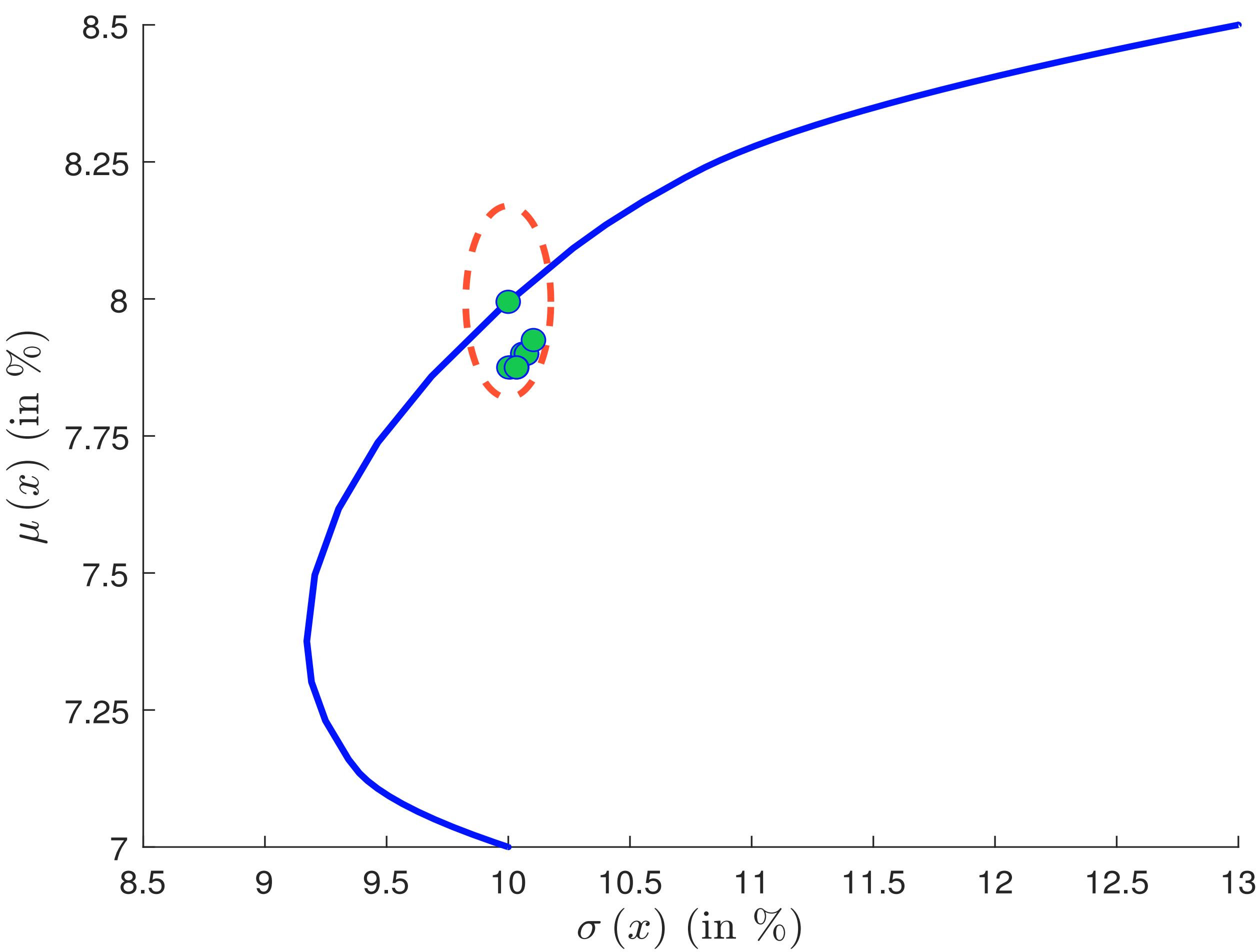


Figure 20: Optimized portfolios versus optimal diversified portfolios

Portfolio optimization & portfolio diversification

Table 22: Some equivalent mean-variance portfolios

x_1	23.97		5	5	35	35	50	5	5	10
x_2	6.42	25		25	10	25	10	30		25
x_3	16.91	5	40		10	5	15		45	10
x_4	28.73	35	20	30	5	35	10	35	20	45
x_5	23.97	35	35	40	40		15	30	30	10
$\mu(x)$	7.99	7.90	7.90	7.90	7.88	7.90	7.88	7.88	7.88	7.93
$\sigma(x)$	10.00	10.07	10.06	10.07	10.01	10.07	10.03	10.00	10.03	10.10

$\Rightarrow$ These portfolios have very different compositions, but lead to very close mean-variance features

Some of these portfolios appear more balanced and more diversified than the optimized portfolio

Other methods to build a portfolio

1. Weight budgeting (WB)
2. Risk budgeting (RB)
3. Performance budgeting (PB)

Ex-ante analysis
$\neq$
Ex-post analysis

Important result

RB = PB

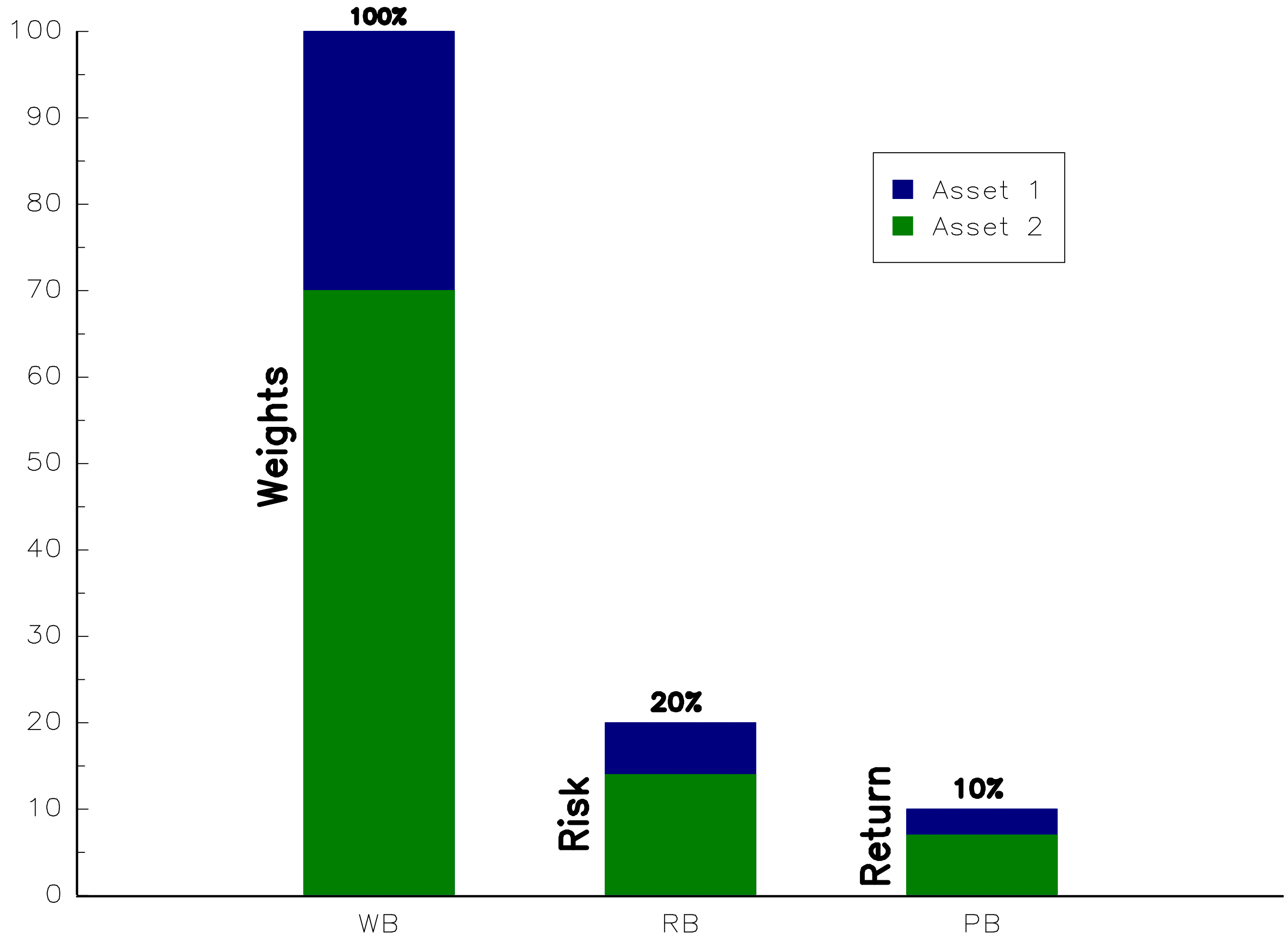


Figure 21: The 30/70 rule

Weight budgeting versus risk budgeting

Let $x = (x_1, \ldots, x_n)$ be the weights of n assets in the portfolio. Let $\mathcal{R}(x_1, \ldots, x_n)$ be a coherent and convex risk measure. We have:

$$
\begin{aligned}
\mathcal{R}(x_1, \ldots, x_n) &= \sum_{i=1}^{n} x_i \cdot \frac{\partial\, \mathcal{R}(x_1, \ldots, x_n)}{\partial\, x_i} \\
&= \sum_{i=1}^{n} \mathcal{RC}_i(x_1, \ldots, x_n)
\end{aligned}
$$

Let $b = (b_1, \ldots, b_n)$ be a vector of budgets such that $b_i \geq 0$ and $\sum_{i=1}^{n} b_i = 1$. We consider two allocation schemes:

1. Weight budgeting (WB)

$$x_i = b_i$$

2. Risk budgeting (RB)

$$\mathcal{RC}_i = b_i \cdot \mathcal{R}(x_1, \ldots, x_n)$$

Importance of the coherency and convexity properties

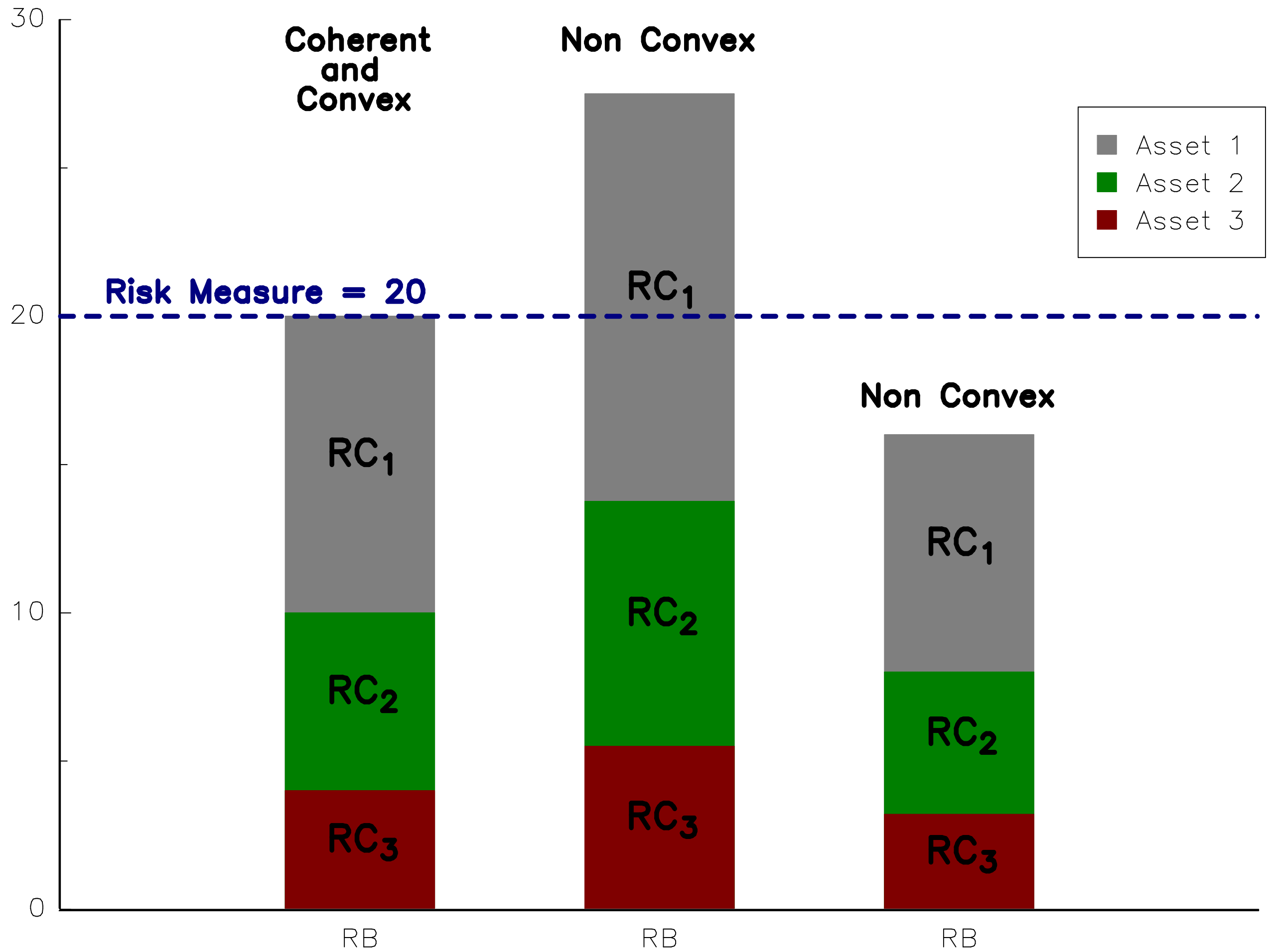


Figure 22: Risk Measure = 20 with a 50/30/20 budget rule

Application to the volatility risk measure

Let Σ be the covariance matrix of the assets returns. We note x the vector of the portfolio's weights:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

It follows that the portfolio volatility is equal to:

$$\sigma(x) = \sqrt{x^\top \Sigma x}$$

Computation of the marginal volatilities

The vector of marginal volatilities is equal to:

$$\begin{aligned}
\frac{\partial\,\sigma(x)}{\partial\,x} &= \begin{pmatrix} \dfrac{\partial\,\sigma(x)}{\partial\,x_1} \\ \vdots \\ \dfrac{\partial\,\sigma(x)}{\partial\,x_n} \end{pmatrix} \\
&= \frac{\partial}{\partial\,x}\left(x^\top \Sigma x\right)^{1/2} \\
&= \frac{1}{2}\left(x^\top \Sigma x\right)^{1/2-1}(2\Sigma x) \\
&= \frac{\Sigma x}{\sqrt{x^\top \Sigma x}}
\end{aligned}$$

It follows that the marginal volatility of Asset i is given by:

$$\frac{\partial\,\sigma(x)}{\partial\,x_i} = \frac{(\Sigma x)_i}{\sqrt{x^\top \Sigma x}} = \sum_{j=1}^n \frac{\rho_{i,j}\sigma_i\sigma_j x_j}{\sigma(x)} = \sigma_i \sum_{j=1}^n x_j \frac{\rho_{i,j}\sigma_j}{\sigma(x)}$$

Computation of the risk contributions

We deduce that the risk contribution of the i^{th} asset is then:

$$\begin{aligned}\mathcal{RC}_i &= x_i \cdot \frac{\partial\, \sigma(x)}{\partial\, x_i} \\ &= \frac{x_i \cdot (\Sigma x)_i}{\sqrt{x^\top \Sigma x}} \\ &= \sigma_i x_i \sum_{j=1}^{n} x_j \frac{\rho_{i,j}\sigma_j}{\sigma(x)}\end{aligned}$$

The Euler allocation principle

We verify that the volatility satisfies the full allocation property:

$$\begin{aligned}\sum_{i=1}^{n}\mathcal{RC}_i &= \sum_{i=1}^{n}\sigma_i x_i \sum_{j=1}^{n} x_j \frac{\rho_{i,j}\sigma_j}{\sigma(x)} = \frac{1}{\sigma(x)}\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \rho_{i,j}\sigma_i\sigma_j \\ &= \frac{\sigma^2(x)}{\sigma(x)} = \sigma(x)\end{aligned}$$

An alternative proof uses the definition of the dot product:

$$a \cdot b = \sum_{i=1}^{n} a_i b_i = a^\top b$$

Indeed, we have:

$$\sum_{i=1}^{n}\mathcal{RC}_i = \sum_{i=1}^{n}\frac{x_i \cdot (\Sigma x)_i}{\sqrt{x^\top \Sigma x}} = \frac{1}{\sqrt{x^\top \Sigma x}}\sum_{i=1}^{n} x_i \cdot (\Sigma x)_i = \frac{1}{\sqrt{x^\top \Sigma x}} x^\top \Sigma x = \sigma(x)$$

Definition of the risk contribution

Definition

The marginal risk contribution of Asset i is:

$$\mathcal{MR}_i = \frac{\partial\,\sigma\left(x\right)}{\partial\,x_i} = \frac{\left(\Sigma x\right)_i}{\sqrt{x^\top \Sigma x}}$$

The absolute risk contribution of Asset i is:

$$\mathcal{RC}_i = x_i \frac{\partial\,\sigma\left(x\right)}{\partial\,x_i} = \frac{x_i \cdot \left(\Sigma x\right)_i}{\sqrt{x^\top \Sigma x}}$$

The relative risk contribution of Asset i is:

$$\mathcal{RC}_i^\star = \frac{\mathcal{RC}_i}{\sigma\left(x\right)} = \frac{x_i \cdot \left(\Sigma x\right)_i}{x^\top \Sigma x}$$

The Euler allocation principle

Remark

We have $\sum_{i=1}^{n} \mathcal{RC}_i = \sigma(x)$ *and* $\sum_{i=1}^{n} \mathcal{RC}_i^{\star} = 100\%$.

Application

Example 2

We consider three assets. We assume that their expected returns are equal to zero whereas their volatilities are equal to 30%, 20% and 15%. The correlation of asset returns is given by the following matrix:

$$\rho = \begin{pmatrix} 1.00 & & \\ 0.80 & 1.00 & \\ 0.50 & 0.30 & 1.00 \end{pmatrix}$$

We consider the portfolio x, which is given by:

$$x = \begin{pmatrix} 50\% \\ 20\% \\ 30\% \end{pmatrix}$$

Application

Using the relationship $\Sigma_{i,j} = \rho_{i,j}\sigma_i\sigma_j$, we deduce that the covariance matrix is[3]:

$$\Sigma = \begin{pmatrix} 9.00 & 4.80 & 2.25 \\ 4.80 & 4.00 & 0.90 \\ 2.25 & 0.90 & 2.25 \end{pmatrix} \times 10^{-2}$$

It follows that the variance of the portfolio is:

$$\begin{aligned} \sigma^2(x) &= 0.50^2 \times 0.09 + 0.20^2 \times 0.04 + 0.30^2 \times 0.0225 + \\ &\quad 2 \times 0.50 \times 0.20 \times 0.0480 + 2 \times 0.50 \times 0.30 \times 0.0225 + \\ &\quad 2 \times 0.20 \times 0.30 \times 0.0090 \\ &= 4.3555\% \end{aligned}$$

The volatility is then $\sigma(x) = \sqrt{4.3555\%} = 20.8698\%$.

[3] The covariance term between assets 1 and 2 is equal to $\Sigma_{1,2} = 80\% \times 30\% \times 20\%$ or $\Sigma_{1,2} = 4.80\%$

Application

The computation of the marginal volatilities gives:

$$\frac{\Sigma x}{\sqrt{x^\top \Sigma x}} = \frac{1}{20.8698\%} \begin{pmatrix} 6.1350\% \\ 3.4700\% \\ 1.9800\% \end{pmatrix} = \begin{pmatrix} 29.3965\% \\ 16.6269\% \\ 9.4874\% \end{pmatrix}$$

Application

Finally, we obtain the risk contributions by multiplying the weights by the marginal volatilities:

$$x \circ \frac{\Sigma x}{\sqrt{x^\top \Sigma x}} = \begin{pmatrix} 50\% \\ 20\% \\ 30\% \end{pmatrix} \circ \begin{pmatrix} 29.3965\% \\ 16.6269\% \\ 9.4874\% \end{pmatrix} = \begin{pmatrix} 14.6982\% \\ 3.3254\% \\ 2.8462\% \end{pmatrix}$$

We verify that the sum of risk contributions is equal to the volatility:

$$\sum_{i=1}^{3} \mathcal{RC}_i = 14.6982\% + 3.3254\% + 2.8462\% = 20.8698\%$$

Application

Table 23: Risk decomposition of the portfolio's volatility (Example 2)

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^{\star}$
1	50.00	29.40	14.70	70.43
2	20.00	16.63	3.33	15.93
3	30.00	9.49	2.85	13.64
$\sigma(x)$			20.87	

The ERC portfolio

Definition

- Let Σ be the covariance matrix of asset returns
- The risk measure corresponds to the volatility risk measure
- The ERC portfolio is the **unique** portfolio x such that the risk contributions are equal:

$$\mathcal{RC}_i = \mathcal{RC}_j \Leftrightarrow \frac{x_i \cdot (\Sigma x)_i}{\sqrt{x^\top \Sigma x}} = \frac{x_j \cdot (\Sigma x)_j}{\sqrt{x^\top \Sigma x}}$$

ERC = Equal Risk Contribution

The concept of risk budgeting

Example 3

- 3 assets
- Volatilities are respectively equal to 20%, 30% and 15%
- Correlations are set to 60% between the 1st asset and the 2nd asset and 10% between the first two assets and the 3rd asset
- Budgets are set to 50%, 25% and 25%
- For the ERC (Equal Risk Contribution) portfolio, all the assets have the same risk budget

Weight budgeting (or traditional approach)

Asset	Weight	Marginal Risk	Risk Contribution Absolute	Risk Contribution Relative
1	**50.00%**	17.99%	9.00%	54.40%
2	**25.00%**	25.17%	6.29%	38.06%
3	**25.00%**	4.99%	1.25%	7.54%
Volatility			16.54%	

Risk budgeting approach

Asset	Weight	Marginal Risk	Risk Contribution Absolute	Risk Contribution Relative
1	41.62%	16.84%	7.01%	**50.00%**
2	15.79%	22.19%	3.51%	**25.00%**
3	42.58%	8.23%	3.51%	**25.00%**
Volatility			14.02%	

ERC approach

Asset	Weight	Marginal Risk	Risk Contribution Absolute	Risk Contribution Relative
1	30.41%	15.15%	4.61%	**33.33%**
2	20.28%	22.73%	4.61%	**33.33%**
3	49.31%	9.35%	4.61%	**33.33%**
Volatility			13.82%	

The concept of risk budgeting

We have:

$$\sigma\left(50\%, 25\%, 25\%\right) = 16.54\%$$

The marginal risk for the first asset is:

$$\frac{\partial\,\sigma\left(x\right)}{\partial\,x_1} = \lim_{\varepsilon \to 0} \frac{\sigma\left(x_1 + \varepsilon, x_2, x_3\right) - \sigma\left(x_1, x_2, x_3\right)}{\left(x_1 + \varepsilon\right) - x_1}$$

If $\varepsilon = 1\%$, we have:

$$\sigma\left(0.51, 0.25, 0.25\right) = 16.72\%$$

We deduce that:

$$\frac{\partial\,\sigma\left(x\right)}{\partial\,x_1} \simeq \frac{16.72\% - 16.54\%}{1\%} = 18.01\%$$

whereas the true value is equal to:

$$\frac{\partial\,\sigma\left(x\right)}{\partial\,x_1} = 17.99\%$$

The concept of risk budgeting

Example 4

- 3 assets
- Volatilities are respectively 30%, 20% and 15%
- Correlations are set to 80% between the 1st asset and the 2nd asset, 50% between the 1st asset and the 3rd asset and 30% between the 2nd asset and the 3rd asset

Weight budgeting (or traditional) approach

Asset	Weight	Marginal Risk	Risk Contribution Absolute	Risk Contribution Relative
1	**50.00%**	29.40%	14.70%	70.43%
2	**20.00%**	16.63%	3.33%	15.93%
3	**30.00%**	9.49%	2.85%	13.64%
Volatility			20.87%	

Risk budgeting approach

Asset	Weight	Marginal Risk	Risk Contribution Absolute	Risk Contribution Relative
1	31.15%	28.08%	8.74%	**50.00%**
2	21.90%	15.97%	3.50%	**20.00%**
3	46.96%	11.17%	5.25%	**30.00%**
Volatility			17.49%	

ERC approach

Asset	Weight	Marginal Risk	Risk Contribution Absolute	Risk Contribution Relative
1	19.69%	27.31%	5.38%	**33.33%**
2	32.44%	16.57%	5.38%	**33.33%**
3	47.87%	11.23%	5.38%	**33.33%**
Volatility			16.13%	

The concept of risk budgeting

Question

We assume that the portfolio's wealth is set to \$1 000. Calculate the nominal volatility of the previous WB, RB and ERC portfolios.

The concept of risk budgeting

We have:

$$\begin{aligned}\sigma\left(x_{\mathrm{wb}}\right) &= 10^{3} \times 20.87\% = \$208.7 \\ \sigma\left(x_{\mathrm{rb}}\right) &= 10^{3} \times 17.49\% = \$174.9 \\ \sigma\left(x_{\mathrm{erc}}\right) &= 10^{3} \times 16.13\% = \$161.3\end{aligned}$$

The concept of risk budgeting

Question

We increase the exposure of the 3 portfolios by \$10 as follows:

$$\Delta x = \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{pmatrix} = \begin{pmatrix} \$1 \\ \$5 \\ \$4 \end{pmatrix}$$

Calculate the nominal volatility of these new portfolios.

The concept of risk budgeting

By assuming that $\Delta x \simeq 0$, we have:

$$\begin{array}{rcl}\sigma\left(x_{\mathrm{wb}}+\Delta x\right) & \approx & (\$500+\$1)\times 0.2940 + \\ & & (\$200+\$5)\times 0.1663 + \\ & & (\$300+\$4)\times 0.0949 \\ & \approx & \$210.2\end{array}$$

$\sigma\left(x_{\mathrm{rb}}+\Delta x\right) \approx \$176.4$ and $\sigma\left(x_{\mathrm{erc}}+\Delta x\right) \approx \$162.9$.

Uniform correlation

- We assume a constant correlation matrix $\mathcal{C}_n(\rho)$, meaning that $\rho_{i,j} = \rho$ for all $i \neq j$
- We have:

$$\begin{aligned}
(\Sigma x)_i &= \sum_{k=1}^{n} \rho_{i,k}\sigma_i\sigma_k x_k \\
&= \sigma_i^2 x_i + \rho\sigma_i \sum_{k\neq i} \sigma_k x_k \\
&= \sigma_i^2 x_i + \rho\sigma_i \sum_{k=1}^{n} \sigma_k x_k - \rho\sigma_i^2 x_i \\
&= (1-\rho) x_i\sigma_i^2 + \rho\sigma_i \sum_{k=1}^{n} x_k\sigma_k \\
&= \sigma_i \left((1-\rho) x_i\sigma_i + \rho \sum_{k=1}^{n} x_k\sigma_k \right)
\end{aligned}$$

Uniform correlation

- Since we have:

$$\mathcal{RC}_i = \frac{x_i \left(\Sigma x\right)_i}{\sigma\left(x\right)}$$

we deduce that $\mathcal{RC}_i = \mathcal{RC}_j$ is equivalent to:

$$x_i \sigma_i \left(\left(1 - \rho\right) x_i \sigma_i + \rho \sum_{k=1}^{n} x_k \sigma_k \right) = x_j \sigma_j \left(\left(1 - \rho\right) x_j \sigma_j + \rho \sum_{k=1}^{n} x_k \sigma_k \right)$$

It follows that $x_i \sigma_i = x_j \sigma_j$. Because $\sum_{i=1}^n x_i = 1$, we deduce that:

$$x_i = \frac{\sigma_i^{-1}}{\sum_{j=1}^n \sigma_j^{-1}}$$

Result

The weight allocated to Asset i is inversely proportional to its volatility and does not depend on the value of the correlation

Minimum uniform correlation

- The global minimum variance portfolio is equal to:

$$x_{\mathrm{gmv}} = \frac{\Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^\top \Sigma^{-1}\mathbf{1}_n}$$

- Let $\Sigma = \sigma\sigma^\top \circ \mathcal{C}_n(\rho)$ be the covariance matrix with $\mathcal{C}_n(\rho)$ the constant correlation matrix
- We have:

$$\Sigma^{-1} = \Gamma \circ \mathcal{C}_n^{-1}(\rho)$$

with $\Gamma_{i,j} = \sigma_i^{-1}\sigma_j^{-1}$ and:

$$\mathcal{C}_n^{-1}(\rho) = \frac{\rho\mathbf{1}_n\mathbf{1}_n^\top - ((n-1)\rho + 1) I_n}{(n-1)\rho^2 - (n-2)\rho - 1}$$

Minimum uniform correlation

- We deduce that the expression of the GMV weights are:

$$x_{\mathrm{gmv},i} = \frac{-\left(\left(n-1\right)\rho+1\right)\sigma_i^{-2}+\rho\sum_{j=1}^{n}\left(\sigma_i\sigma_j\right)^{-1}}{\sum_{k=1}^{n}\left(-\left(\left(n-1\right)\rho+1\right)\sigma_k^{-2}+\rho\sum_{j=1}^{n}\left(\sigma_k\sigma_j\right)^{-1}\right)}$$

- The lower bound of $\mathcal{C}_n\left(\rho\right)$ is achieved for $\rho=-\left(n-1\right)^{-1}$
- In this case, the solution becomes:

$$x_{\mathrm{gmv},i} = \frac{\sum_{j=1}^{n}\left(\sigma_i\sigma_j\right)^{-1}}{\sum_{k=1}^{n}\sum_{j=1}^{n}\left(\sigma_k\sigma_j\right)^{-1}} = \frac{\sigma_i^{-1}}{\sum_{k=1}^{n}\sigma_k^{-1}}$$

Result

The ERC portfolio is equal to the GMV portfolio when the correlation is at its lowest possible value:

$$\lim_{\rho\rightarrow-\left(n-1\right)^{-1}} x_{\mathrm{gmv}} = x_{\mathrm{erc}}$$

Uniform volatility

- If all volatilities are equal, i.e. $\sigma_i = \sigma$ for all i, the risk contribution becomes:

$$\mathcal{RC}_i = \frac{\left(\sum_{k=1}^n x_i x_k \rho_{i,k}\right) \sigma^2}{\sigma(x)}$$

- The ERC portfolio verifies then:

$$x_i \left(\sum_{k=1}^n x_k \rho_{i,k}\right) = x_j \left(\sum_{k=1}^n x_k \rho_{j,k}\right)$$

- We deduce that:

$$x_i = \frac{\left(\sum_{k=1}^n x_k \rho_{i,k}\right)^{-1}}{\sum_{j=1}^n \left(\sum_{k=1}^n x_k \rho_{j,k}\right)^{-1}}$$

Uniform volatility

Result

The weight of asset i is inversely proportional to the weighted average of correlations of Asset i

Remark

Contrary to the previous case, this solution is endogenous since x_i is a function of itself directly

General case

- In the general case, we have:

$$\beta_i = \beta\left(\mathbf{e}_i \mid x\right) = \frac{\mathbf{e}_i^\top \Sigma x}{x^\top \Sigma x} = \frac{(\Sigma x)_i}{\sigma^2(x)}$$

and:

$$\mathcal{RC}_i = \frac{x_i (\Sigma x)_i}{\sigma(x)} = \sigma(x)\, x_i \beta_i$$

- We deduce that $\mathcal{RC}_i = \mathcal{RC}_j$ is equivalent to:

$$x_i \beta_i = x_j \beta_j$$

- It follows that:

$$x_i = \frac{\beta_i^{-1}}{\sum_{j=1}^n \beta_j^{-1}}$$

General case

- We notice that:

$$\sum_{i=1}^{n} x_i \beta_i = \sum_{i=1}^{n} \frac{\mathcal{RC}_i}{\sigma(x)} = \frac{1}{\sigma(x)} \sum_{i=1}^{n} \mathcal{RC}_i = 1$$

  and:

$$\sum_{i=1}^{n} x_i \beta_i = \sum_{i=1}^{n} \left(\frac{1}{\sum_{j=1}^{n} \beta_j^{-1}} \right) = 1$$

  It follows that:

$$\frac{1}{\sum_{j=1}^{n} \beta_j^{-1}} = \frac{1}{n}$$

- We finally obtain:

$$x_i = \frac{1}{n \beta_i}$$

General case

Result

The weight of Asset i is proportional to the inverse of its beta:

$$x_i \propto \beta_i^{-1}$$

Remark

This solution is endogenous since x_i is a function of itself because $\beta_i = \beta(\mathbf{e}_i \mid x)$.

General case

Example 5

We consider an investment universe of four assets with $\sigma_1 = 15\%$, $\sigma_2 = 20\%$, $\sigma_3 = 30\%$ and $\sigma_4 = 10\%$. The correlation of asset returns is given by the following matrix:

$$\rho = \begin{pmatrix} 1.00 & & & \\ 0.50 & 1.00 & & \\ 0.00 & 0.20 & 1.00 & \\ -0.10 & 0.40 & 0.70 & 1.00 \end{pmatrix}$$

General case

Table 24: Composition of the ERC portfolio (Example 5)

Asset	x_i	$\mathcal{MR}_i$	β_i	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	31.34%	8.52%	0.80	2.67%	25.00%
2	17.49%	15.27%	1.43	2.67%	25.00%
3	13.05%	20.46%	1.92	2.67%	25.00%
4	38.12%	7.00%	0.66	2.67%	25.00%
Volatility				10.68%	

We verify that:

$$x_1 = \frac{1}{(4 \times 0.7978)} = 31.34\%$$

Existence and uniqueness

We consider the following optimization problem:

$$\begin{aligned} y^{\star}(c) &= \arg\min \frac{1}{2} y^{\top} \Sigma y \\ &\text{u.c.} \quad \sum_{i=1}^{n} \ln y_i \geq c \end{aligned}$$

The Lagrange function is equal to:

$$\mathcal{L}(y; \lambda_c) = \frac{1}{2} y^{\top} \Sigma y - \lambda_c \left(\sum_{i=1}^{n} \ln y_i - c \right)$$

At the optimum, we have:

$$\frac{\partial \mathcal{L}(y; \lambda_c, \lambda)}{\partial y} = \mathbf{0}_n \Leftrightarrow (\Sigma y)_i - \frac{\lambda_c}{y_i} = 0$$

Existence and uniqueness

It follows that:

$$y_i \left(\Sigma y\right)_i = \lambda_c$$

or equivalently:

$$\mathcal{RC}_i = \mathcal{RC}_j$$

Since we minimize a convex function subject to a lower convex bound, the solution $y^{\star}\left(c\right)$ exists and is unique

Existence and uniqueness

Question

What is the difference between $y^\star(c)$ and $y^\star(c')$?

Let $y' = \alpha y^\star(c)$. The first-order conditions are:

$$y_i^\star(c)\left(\Sigma y^\star(c)\right)_i = \lambda_c$$

and:

$$y_i'\left(\Sigma y'\right)_i = \alpha^2 \lambda_c = \lambda_{c'}$$

Since $\lambda_c \neq 0$, the Kuhn-Tucker condition becomes:

$$\min\left(\lambda_c, \sum_{i=1}^n \ln y_i^\star(c) - c\right) = 0 \Leftrightarrow \sum_{i=1}^n \ln y_i^\star(c) - c = 0$$

Existence and uniqueness

It follows that:

$$\sum_{i=1}^{n} \ln \frac{y_i'(c)}{\alpha} = c$$

or:

$$\sum_{i=1}^{n} \ln y_i'(c) = c + n \ln \alpha = c'$$

We deduce that:

$$\alpha = \exp\left(\frac{c' - c}{n}\right)$$

$y^\star(c')$ is a scaled solution of $y^\star(c)$:

$$y^\star(c') = \exp\left(\frac{c' - c}{n}\right) y^\star(c)$$

Existence and uniqueness

The ERC portfolio is the solution $y^{\star}(c)$ such that $\sum_{i=1}^{n} y_i^{\star}(c) = 1$:

$$x_{\mathrm{erc}} = \frac{y^{\star}(c)}{\sum_{i=1}^{n} y_i^{\star}(c)}$$

and corresponds to the following value of the logarithmic barrier:

$$c_{\mathrm{erc}} = c - n \ln \sum_{i=1}^{n} y_i^{\star}(c)$$

Existence and uniqueness

Theorem

Because of the previous results, x_{erc} exists and is unique. This is the solution of the following optimization problem[a]:

$$\begin{aligned} x_{\mathrm{erc}} &= \arg\min \frac{1}{2} x^\top \Sigma x \\ \text{u.c.} &\quad \left\{ \begin{array}{l} \sum_{i=1}^{n} \ln x_i \geq c_{\mathrm{erc}} \\ \mathbf{1}_n^\top x = 1 \\ \mathbf{0}_n \leq x \leq \mathbf{1}_n \end{array} \right. \end{aligned}$$

[a] We can add the last two constraints because they do not change the solution

Location of the ERC portfolio

The global global minimum variance portfolio is defined by:

$$\begin{array}{rcl} x_{\mathrm{gmv}} & = & \arg\min \sigma(x) \\ & \text{u.c.} & \mathbf{1}_n^\top x = 1 \end{array}$$

We have:

$$\mathcal{L}(x;\lambda_0) = \sigma(x) - \lambda_0\left(\mathbf{1}_n^\top x - 1\right)$$

The first-order condition is:

$$\frac{\partial\,\mathcal{L}(x;\lambda_0)}{\partial\,x} = \mathbf{0}_n \Leftrightarrow \frac{\partial\,\sigma(x)}{\partial\,x} - \lambda_0\mathbf{1}_n = \mathbf{0}_n$$

Location of the ERC portfolio

Theorem

The global minimum variance portfolio satisfies:

$$\frac{\partial \sigma(x)}{\partial x_i} = \frac{\partial \sigma(x)}{\partial x_j}$$

The marginal volatilities are then the same.

Location of the ERC portfolio

The equally-weighted portfolio is defined by:

$$x_i = \frac{1}{n}$$

We deduce that:

$$x_i = x_j$$

Location of the ERC portfolio

We have:

$$x_i = x_j \qquad \text{(EW)}$$

$$\frac{\partial\, \sigma\left(x\right)}{\partial\, x_i} = \frac{\partial\, \sigma\left(x\right)}{\partial\, x_j} \qquad \text{(GMV)}$$

$$x_i \frac{\partial\, \sigma\left(x\right)}{\partial\, x_i} = x_j \frac{\partial\, \sigma\left(x\right)}{\partial\, x_j} \qquad \text{(ERC)}$$

The ERC portfolio is a combination of GMV and EW portfolios

Volatility of the ERC portfolio

We consider the following optimization problem:

$$
\begin{aligned}
x^{\star}(c) &= \arg\min \frac{1}{2} x^{\top} \Sigma x \\
\text{u.c.} & \left\{
\begin{array}{l}
\sum_{i=1}^{n} \ln x_i \geq c \\
\mathbf{1}_n^{\top} x = 1 \\
\mathbf{0}_n \leq x \leq \mathbf{1}_n
\end{array}
\right.
\end{aligned}
$$

- We know that there exists a scalar c_{erc} such that:

$$
x^{\star}(c_{\mathrm{erc}}) = x_{\mathrm{erc}}
$$

- If $c = -\infty$, the logarithmic barrier constraint vanishes and we have:

$$
x^{\star}(-\infty) = x_{\mathrm{mv}}
$$

where x_{mv} is the long-only minimum variance portfolio

Volatility of the ERC portfolio

- We notice that the function $f(x) = \sum_{i=1}^{n} \ln x_i$ such that $\mathbf{1}_n^\top x = 1$ reaches its maximum when:

$$\frac{1}{x_i} = \lambda_0$$

implying that $x_i = x_j = n^{-1}$. In this case, we have:

$$c_{\max} = \sum_{i=1}^{n} \ln \frac{1}{n} = -n \ln n$$

- If $c = -n \ln n$, we have:

$$x^\star(-n \ln n) = x_{\mathrm{ew}}$$

- Because we have a convex minimization problem and a lower convex bound, we deduce that:

$$c_2 \geq c_1 \Leftrightarrow \sigma(x^\star(c_2)) \geq \sigma(x^\star(c_1))$$

Volatility of the ERC portfolio

Theorem

We obtain the following inequality:

$$\sigma\left(x_{\mathrm{mv}}\right) \leq \sigma\left(x_{\mathrm{erc}}\right) \leq \sigma\left(x_{\mathrm{ew}}\right)$$

The ERC portfolio may be viewed as a portfolio "between" the MV portfolio and the EW portfolio.

Remark

The ERC portfolio is a form of variance-minimizing portfolio subject to a constraint of sufficient diversification in terms of weights

Relationship with naive diversification $(1/n)$

Optimality of the ERC portfolio

Let us consider the tangency (or maximum Sharpe ratio) portfolio defined by:

$$x_{\mathrm{msr}} = \arg\max \frac{\mu(x) - r}{\sigma(x)}$$

where $\mu(x) = x^\top \mu$ and $\sigma(x) = \sqrt{x^\top \Sigma x}$. We recall that the portfolio is MSR if and only if:

$$\frac{\partial_{x_i}\, \mu(x) - r}{\partial_{x_i}\, \sigma(x)} = \frac{\mu(x) - r}{\sigma(x)}$$

Therefore, the MSR portfolio x_{msr} verifies the following relationship:

$$\begin{aligned} \mu - r\mathbf{1}_n &= \left(\frac{\mu(x_{\mathrm{msr}}) - r}{\sigma^2(x_{\mathrm{msr}})}\right) \Sigma x_{\mathrm{msr}} \\ &= \mathrm{SR}(x_{\mathrm{msr}} \mid r) \frac{\Sigma x_{\mathrm{msr}}}{\sigma(x_{\mathrm{msr}})} \end{aligned}$$

Optimality of the ERC portfolio

- If we assume a constant correlation matrix, the ERC portfolio is defined by:

$$x_i = \frac{c}{\sigma_i}$$

where $c = \left(\sum_{j=1}^{n} \sigma_j^{-1}\right)^{-1}$

- We have:

$$(\Sigma x)_i = \sum_{j=1}^{n} \rho_{i,j}\sigma_i\sigma_j x_j = c\sigma_i \sum_{j=1}^{n} \rho_{i,j} = c\sigma_i \left(1 + \rho\left(n-1\right)\right)$$

- We deduce that:

$$\frac{\partial\,\sigma\left(x\right)}{\partial\,x_i} = c\frac{\sigma_i\left(\left(1-\rho\right)+\rho n\right)}{\sigma\left(x\right)}$$

Optimality of the ERC portfolio

- The portfolio volatility is equal to:

$$\begin{aligned}\sigma^2(x) &= \sigma(x)\sum\nolimits_{i=1}^{n} x_i \frac{\partial\,\sigma(x)}{\partial\,x_i} \\ &= \sigma(x)\sum\nolimits_{i=1}^{n} \frac{c}{\sigma_i}\cdot c\frac{\sigma_i\left((1-\rho)+\rho n\right)}{\sigma(x)} \\ &= nc^2\left((1-\rho)+\rho n\right)\end{aligned}$$

- The ERC portfolio is the MSR portfolio if and only if:

$$\begin{aligned}\mu_i - r &= \left(\frac{\sum_{j=1}^{n}(\mu_j - r)\,x_j}{\sigma^2(x)}\right)(\Sigma x)_i \\ &= \left(\frac{\sum_{j=1}^{n}(\mu_j - r)\,c\sigma_j^{-1}}{nc^2\left((1-\rho)+\rho n\right)}\right)c\sigma_i\left(1+\rho(n-1)\right) \\ &= \left(\frac{1}{n}\sum_{j=1}^{n}\frac{\mu_j - r}{\sigma_j}\right)\sigma_i\end{aligned}$$

Optimality of the ERC portfolio

- We can write this condition as follows:

$$\mu_i = r + \mathrm{SR} \cdot \sigma_i$$

where:

$$\mathrm{SR} = \frac{1}{n} \sum_{j=1}^{n} \frac{\mu_j - r}{\sigma_j}$$

Theorem

The ERC portfolio is the tangency or MSR portfolio if and only if the correlation is uniform and the Sharpe ratio is the same for all the assets

Optimality of the ERC portfolio

Example 6

We consider an investment universe of five assets. The volatilities are respectively equal to 5%, 7%, 9%, 10% and 15%. The risk-free rate is equal to 2%. The correlation is uniform.

Optimality of the ERC portfolio

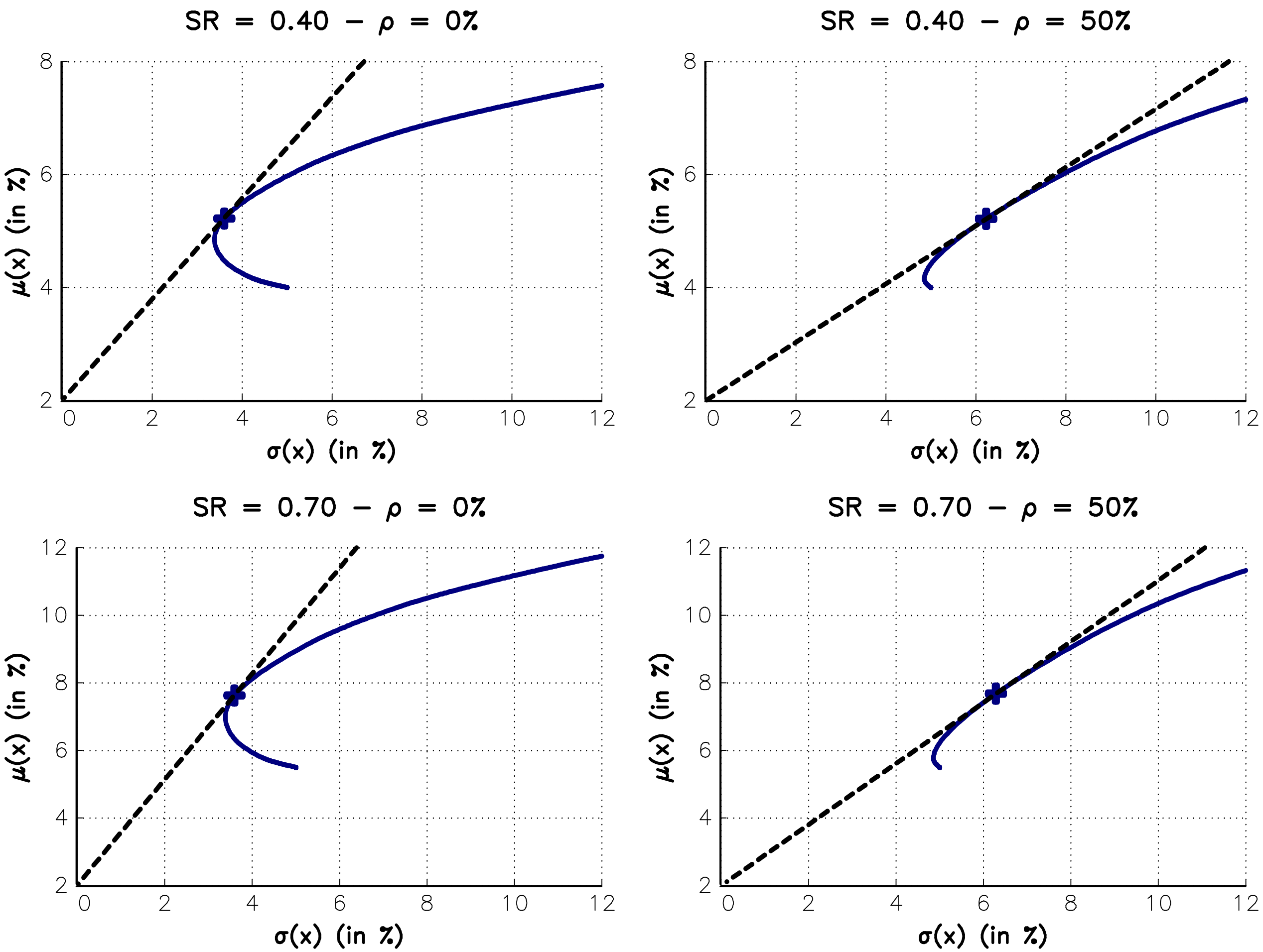


Figure 23: Location of the ERC portfolio in the mean-variance diagram when the Sharpe ratios are the same (Example 6)

Optimality of the ERC portfolio

Example 7

We consider an investment universe of five assets. The volatilities are respectively equal to 5%, 7%, 9%, 10% and 15%. The correlation matrix is equal to:

$$\rho = \begin{pmatrix} 1.00 & & & & \\ 0.50 & 1.00 & & & \\ 0.25 & 0.25 & 1.00 & & \\ 0.00 & 0.00 & 0.00 & 1.00 & \\ -0.25 & -0.25 & -0.25 & 0.00 & 1.00 \end{pmatrix}$$

Optimality of the ERC portfolio

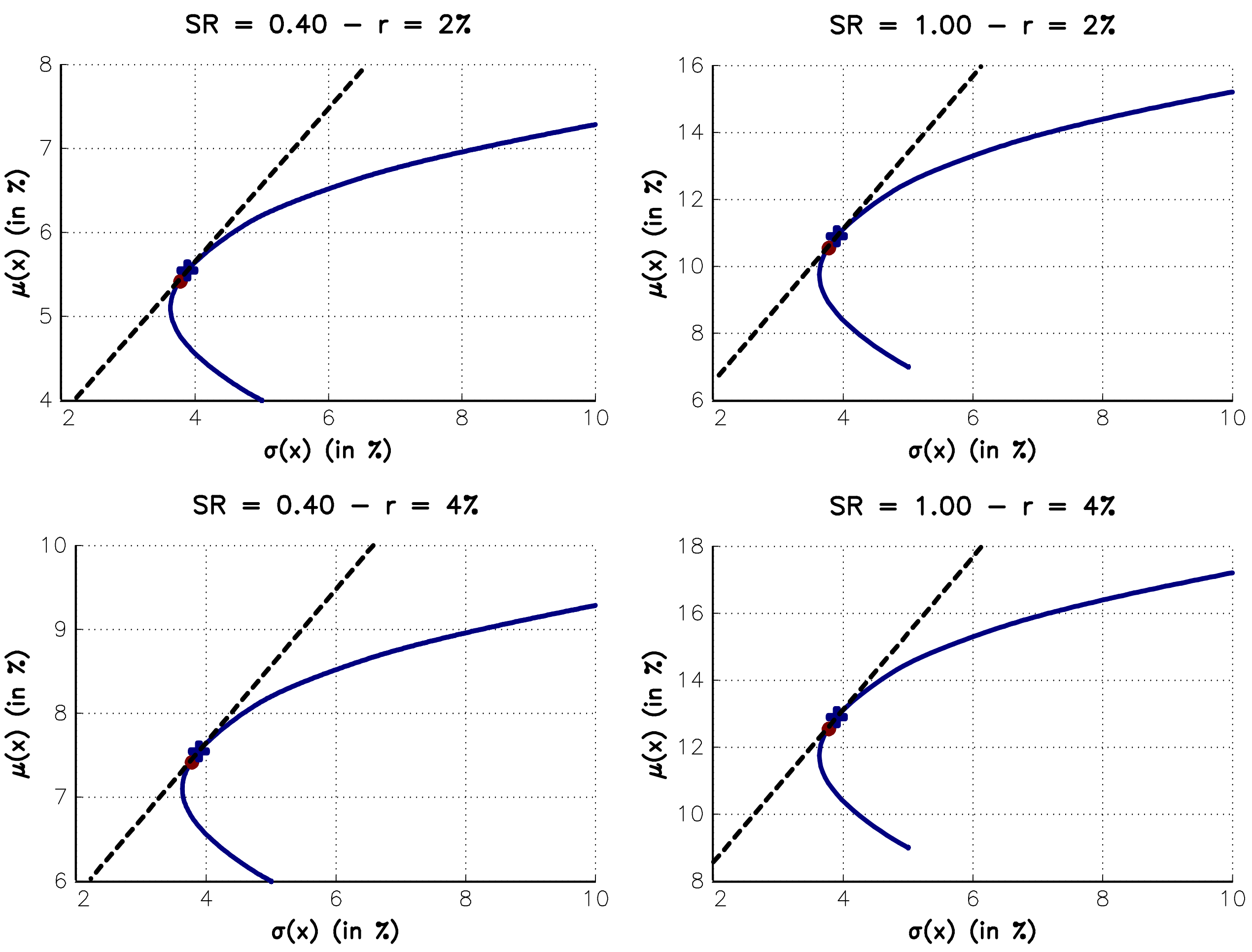


Figure 24: Location of the ERC portfolio in the mean-variance diagram when the Sharpe ratios are the same (Example 7)

The SQP approach

- The ERC portfolio satisfies:

$$x_i \cdot (\Sigma x)_i = x_j \cdot (\Sigma x)_j$$

or:

$$x_i \cdot (\Sigma x)_i = \frac{x^\top \Sigma x}{n}$$

- We deduce that:

$$\begin{array}{rcl} x_{\mathrm{erc}} & = & \arg\min f(x) \\ & \text{u.c.} & \left\{ \begin{array}{l} \mathbf{1}_n^\top x = 1 \\ \mathbf{0}_n \leq x \leq \mathbf{1}_n \end{array} \right. \end{array}$$

and $f(x_{\mathrm{erc}}) = 0$

Remark

The optimization problem is solved using the sequential quadratic programming (or SQP) algorithm

The SQP approach

- We can choose:

$$f(x) = \sum_{i=1}^{n} \left(x_i \cdot (\Sigma x)_i - \frac{1}{n} x^\top \Sigma x \right)^2$$

or:

$$f(x; b) = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(x_i \cdot (\Sigma x)_i - x_j \cdot (\Sigma x)_j \right)^2$$

The Jacobi approach

- We have:

$$\beta_i(x) = \frac{(\Sigma x)_i}{x^\top \Sigma x}$$

- The ERC portfolio satisfies:

$$x_i = \frac{\beta_i^{-1}(x)}{\sum_{j=1}^n \beta_j^{-1}(x)}$$

or:

$$x_i \propto \frac{1}{(\Sigma x)_i}$$

The Jacobi approach

The Jacobi algorithm consists in finding the fixed point by considering the following iterations:

1. We set $k \leftarrow 0$ and we note $x^{(0)}$ the vector of starting values[4]
2. At iteration $k+1$, we compute:

$$y_i^{(k+1)} \propto \frac{1}{\beta_i\left(x^{(k)}\right)} = \frac{1}{\left(\Sigma x^{(k)}\right)_i}$$

   and:

$$x_i^{(k+1)} = \frac{y_i^{(k+1)}}{\sum_{j=1}^n y_j^{(k+1)}}$$

3. We iterate Step 2 until convergence

[4] For instance, we can use the following rule:

$$x_i^{(0)} = \frac{\sigma_i^{-1}}{\sum_{j=1}^n \sigma_j^{-1}}$$

The Newton-Raphson approach

We consider the following optimization problem:

$$x^{\star} = \arg\min f(x)$$

The Newton-Raphson iteration is defined by:

$$x^{(k+1)} = x^{(k)} - \Delta x^{(k)}$$

where $\Delta x^{(k)}$ is the inverse of the Hessian matrix of $f\left(x^{(k)}\right)$ times the gradient vector of $f\left(x^{(k)}\right)$:

$$\Delta x^{(k)} = \left[\partial_x^2 f\left(x^{(k)}\right)\right]^{-1} \partial_x f\left(x^{(k)}\right)$$

The Newton-Raphson approach

- We consider the Lagrange function:

$$f(y) = \frac{1}{2} y^\top \Sigma y - \lambda_c \sum_{i=1}^{n} \ln y_i$$

- We choose a value of λ_c (e.g. $\lambda_c = 1$)
- We note y^{-m} the vector $n \times 1$ matrix with elements $\left(y_1^{-m}, \ldots, y_n^{-m}\right)$ and $\operatorname{diag}\left(y^{-m}\right)$ the $n \times n$ diagonal matrix with elements $\left(y_1^{-m}, \ldots, y_n^{-m}\right)$:

$$\operatorname{diag}\left(y^{-m}\right) = \begin{pmatrix} y_1^{-m} & 0 & & 0 \\ 0 & y_2^{-m} & & \\ & & \ddots & 0 \\ 0 & & 0 & y_n^{-m} \end{pmatrix}$$

The Newton-Raphson approach

- We apply the Newton-Raphson algorithm with:

$$\partial_y f(y) = \Sigma y - \lambda_c y^{-1}$$

and:

$$\partial_y^2 f(y) = \Sigma + \lambda_c \operatorname{diag}\left(y^{-2}\right)$$

- The solution is given by:

$$x_{\text{erc}} = \frac{y^\star}{\sum_{i=1}^n y_i^\star}$$

The Newton-Raphson approach

- For the starting value $y_i^{(0)}$, we can assume that the correlations are uniform:

$$y_i^{(0)} = \frac{\sigma_i^{-1}}{\sum_{j=1}^n \sigma_j^{-1}}$$

- At the optimum, we recall that $\lambda_c = y_i^\star \cdot (\Sigma y^\star)_i$. We deduce that:

$$\lambda_c = \frac{1}{n} \sum_{i=1}^n y_i^\star \cdot (\Sigma y^\star)_i = \frac{\sigma^2 (y^\star)}{n}$$

Therefore, we can choose:

$$\lambda_c = \frac{\sigma^2 \left(y^{(0)} \right)}{n}$$

The Newton-Raphson approach

- From a numerical point of view, it may be important to control the magnitude order α of $y^\star$ (e.g. $\alpha = 10\%$, $\alpha = 1$ or $\alpha = 10$). For instance, we don't want that the magnitude order is 10^{-5} or 10^5. In this case, we can use the following rule:

$$\lambda_c = n\alpha^2\sigma^2\left(x_{\mathrm{erc}}\right)$$

- For example, if $n = 10$ and $\alpha = 5$, and we guess that the volatility of the ERC portfolio is around 10%, we set:

$$\lambda_c = 10 \times 5^2 \times 0.10^2 = 2.5$$

The CCD approach

Table 25: Cyclical coordinate descent algorithm

The goal is to find the solution $x^{\star} = \arg\min f(x)$
We initialize the vector $x^{(0)}$
Set $k \leftarrow 0$
repeat
  for $i = 1 : n$ **do**
    $x_i^{(k+1)} = \arg\min_{\varkappa} f\left(x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}, \varkappa, x_{i+1}^{(k)}, \ldots, x_n^{(k)}\right)$
  end for
  $k \leftarrow k + 1$
until convergence
return $x^{\star} \leftarrow x^{(k)}$

The CCD approach

We have:

$$\mathcal{L}(y;\lambda_c) = \arg\min \frac{1}{2} y^\top \Sigma y - \lambda_c \sum_{i=1}^{n} \ln y_i$$

The first-order condition is equal to:

$$\frac{\partial\,\mathcal{L}(y;\lambda)}{\partial\, y_i} = (\Sigma y)_i - \frac{\lambda_c}{y_i} = 0$$

or:

$$y_i \cdot (\Sigma y)_i - \lambda_c = 0$$

It follows that:

$$\sigma_i^2 y_i^2 + \left(\sigma_i \sum_{j \neq i} \rho_{i,j} \sigma_j y_j\right) y_i - \lambda_c = 0$$

The CCD approach

We recognize a second-degree equation:

$$\alpha_i y_i^2 + \beta_i y_i + \gamma_i = 0$$

1. The polynomial function is convex because we have $\alpha_i = \sigma_i^2 > 0$
2. The product of the roots is negative:

$$y_i' y_i'' = \frac{\gamma_i}{\alpha_i} = -\frac{\lambda_c}{\sigma_i^2} < 0$$

3. The discriminant is positive:

$$\Delta = \beta_i^2 - 4\alpha_i\gamma_i = \left(\sigma_i \sum_{j \neq i} \rho_{i,j}\sigma_j y_j\right)^2 + 4\sigma_i^2\lambda_c > 0$$

We always have two solutions with opposite signs. We deduce that the solution is the positive root of the second-degree equation:

$$y_i^\star = y_i'' = \frac{-\beta_i + \sqrt{\beta_i^2 - 4\alpha_i\gamma_i}}{2\alpha_i}$$

The CCD approach

The CCD algorithm consists in iterating the following formula:

$$y_i^{(k+1)} = \frac{-\beta_i^{(k+1)} + \sqrt{\left(\beta_i^{(k+1)}\right)^2 - 4\alpha_i^{(k+1)}\gamma_i^{(k+1)}}}{2\alpha_i^{(k+1)}}$$

where:

$$\begin{array}{rcl} \alpha_i^{(k+1)} & = & \sigma_i^2 \\ \beta_i^{(k+1)} & = & \sigma_i \left(\sum_{j<i} \rho_{i,j}\sigma_j y_j^{(k+1)} + \sum_{j>i} \rho_{i,j}\sigma_j y_j^{(k)}\right) \\ \gamma_i^{(k+1)} & = & -\lambda_c \end{array}$$

The ERC portfolio is the scaled solution $y^\star$:

$$x_{\mathrm{erc}} = \frac{y^\star}{\sum_{i=1}^n y_i^\star}$$

Efficiency of the algorithms

$$\text{CCD} \succ \text{NR} \succ \text{SQP} \succ \text{Jacobi}$$

Definition of RB portfolios

Definition

A risk budgeting (RB) portfolio x satisfies the following conditions:

$$\left\{\begin{array}{l} \mathcal{RC}_1 = b_1 \mathcal{R}(x) \\ \vdots \\ \mathcal{RC}_i = b_i \mathcal{R}(x) \\ \vdots \\ \mathcal{RC}_n = b_n \mathcal{R}(x) \end{array}\right.$$

where $\mathcal{R}(x)$ is a coherent and convex risk measure and $b = (b_1, \ldots, b_n)$ is a vector of risk budgets such that $b_i \geq 0$ and $\sum_{i=1}^n b_i = 1$

Definition of RB portfolios

Remark

The ERC portfolio is a particular case of RB portfolios when $\mathcal{R}(x) = \sigma(x)$ and $b_i = \frac{1}{n}$

Coherent risk measure

1. Subadditivity
$$\mathcal{R}(x_1 + x_2) \leq \mathcal{R}(x_1) + \mathcal{R}(x_2)$$
2. Homogeneity
$$\mathcal{R}(\lambda x) = \lambda \mathcal{R}(x) \quad \text{if } \lambda \geq 0$$
3. Monotonicity
$$\text{if } x_1 \prec x_2, \text{ then } \mathcal{R}(x_1) \geq \mathcal{R}(x_2)$$
4. Translation invariance
$$\text{if } m \in \mathbb{R}, \text{ then } \mathcal{R}(x + m) = \mathcal{R}(x) - m$$

Convex risk measure

The convexity property is defined as follows:

$$\mathcal{R}\left(\lambda x_1 + (1-\lambda) x_2\right) \leq \lambda \mathcal{R}\left(x_1\right) + (1-\lambda) \mathcal{R}\left(x_2\right)$$

This condition means that diversification should not increase the risk

Euler allocation principle

This property is necessary for the Euler allocation principle:

$$\mathcal{R}\left(x\right) = \sum_{i=1}^{n} x_i \frac{\partial\,\mathcal{R}\left(x\right)}{\partial\,x_i}$$

Some risk measures

The portfolio loss is $L(x) = -R(x)$ where $R(x)$ is the portfolio return. We consider then different risk measures:

- Volatility of the loss

$$\mathcal{R}(x) = \sigma(L(x)) = \sigma(x)$$

- Standard deviation-based risk measure

$$\mathcal{R}(x) = \mathrm{SD}_c(x) = \mathbb{E}[L(x)] + c \cdot \sigma(L(x)) = -\mu(x) + c \cdot \sigma(x)$$

- Value-at-risk

$$\mathcal{R}(x) = \mathrm{VaR}_\alpha(x) = \inf\{\ell : \Pr\{L(x) \leq \ell\} \geq \alpha\}$$

- Expected shortfall

$$\mathcal{R}(x) = \mathrm{ES}_\alpha(x) = \mathbb{E}[L(x) \mid L(x) \geq \mathrm{VaR}_\alpha(x)] = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(x)\,\mathrm{d}u$$

Gaussian risk measures

We assume that the asset returns are normally distributed: $R \sim \mathcal{N}(\mu, \Sigma)$

We have:

$$\begin{aligned}
\sigma(x) &= \sqrt{x^\top \Sigma x} \\
\mathrm{SD}_c(x) &= -x^\top \mu + c \cdot \sqrt{x^\top \Sigma x} \\
\mathrm{VaR}_\alpha(x) &= -x^\top \mu + \Phi^{-1}(\alpha) \sqrt{x^\top \Sigma x} \\
\mathrm{ES}_\alpha(x) &= -x^\top \mu + \frac{\sqrt{x^\top \Sigma x}}{(1-\alpha)} \phi\left(\Phi^{-1}(\alpha)\right)
\end{aligned}$$

Gaussian risk contributions

- Volatility $\sigma(x)$

$$\mathcal{RC}_i = x_i \cdot \frac{(\Sigma x)_i}{\sqrt{x^\top \Sigma x}}$$

- Standard deviation-based risk measure $\mathrm{SD}_c(x)$

$$\mathcal{RC}_i = x_i \cdot \left(-\mu_i + c \frac{(\Sigma x)_i}{\sqrt{x^\top \Sigma x}} \right)$$

- Value-at-risk $\mathrm{VaR}_\alpha(x)$

$$\mathcal{RC}_i = x_i \cdot \left(-\mu_i + \Phi^{-1}(\alpha) \frac{(\Sigma x)_i}{\sqrt{x^\top \Sigma x}} \right)$$

- Expected shortfall $\mathrm{ES}_\alpha(x)$

$$\mathcal{RC}_i = x_i \cdot \left(-\mu_i + \frac{(\Sigma x)_i}{(1-\alpha)\sqrt{x^\top \Sigma x}} \phi\left(\Phi^{-1}(\alpha)\right) \right)$$

Gaussian risk contributions

Example 8

We consider three assets. We assume that their expected returns are equal to zero whereas their volatilities are equal to 30%, 20% and 15%. The correlation of asset returns is given by the following matrix:

$$\rho = \begin{pmatrix} 1.00 & & \\ 0.80 & 1.00 & \\ 0.50 & 0.30 & 1.00 \end{pmatrix}$$

The portfolio is equal to $(50\%, 20\%, 30\%)$.

Gaussian risk contributions

Table 26: Risk decomposition of the portfolio (Example 8)

$\mathcal{R}(x)$	Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
Volatility	1	50.00	29.40	14.70	70.43
	2	20.00	16.63	3.33	15.93
	3	30.00	9.49	2.85	13.64
	$\sigma(x)$			20.87	
Value-at-risk	1	50.00	68.39	34.19	70.43
	2	20.00	38.68	7.74	15.93
	3	30.00	22.07	6.62	13.64
	$\mathrm{VaR}_{99\%}(x)$			48.55	
Expected shortfall	1	50.00	78.35	39.17	70.43
	2	20.00	44.31	8.86	15.93
	3	30.00	25.29	7.59	13.64
	$\mathrm{ES}_{99\%}(x)$			55.62	

Gaussian risk contributions

Example 9

We consider three assets. We assume that their expected returns are equal to 10%, 5% and 8% whereas their volatilities are equal to 30%, 20% and 15%. The correlation of asset returns is given by the following matrix:

$$\rho = \begin{pmatrix} 1.00 & & \\ 0.80 & 1.00 & \\ 0.50 & 0.30 & 1.00 \end{pmatrix}$$

The portfolio is equal to $(50\%, 20\%, 30\%)$.

Gaussian risk contributions

Table 27: Risk decomposition of the portfolio (Example 9)

$\mathcal{R}(x)$	Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
Volatility	1	50.00	29.40	14.70	70.43
	2	20.00	16.63	3.33	15.93
	3	30.00	9.49	2.85	13.64
	$\sigma(x)$			20.87	
Value-at-risk	1	50.00	58.39	29.19	72.71
	2	20.00	33.68	6.74	16.78
	3	30.00	14.07	4.22	10.51
	$\mathrm{VaR}_{99\%}(x)$			40.15	
Expected shortfall	1	50.00	68.35	34.17	72.37
	2	20.00	39.31	7.86	16.65
	3	30.00	17.29	5.19	10.98
	$\mathrm{ES}_{99\%}(x)$			47.22	

Non-Gaussian risk contributions

They are not frequently used in asset management and portfolio allocation, except in the case of skewed assets (Bruder *et al.*, 2016; Lezmi *et al.*, 2018)

Non-parametric risk contributions are given in Chapter 2 in Roncalli (2013)

Gaussian RB portfolios

Example 10

We consider three assets. We assume that their expected returns are equal to 10%, 5% and 8% whereas their volatilities are equal to 30%, 20% and 15%. The correlation of asset returns is given by the following matrix:

$$\rho = \left(\begin{array}{ccc} 1.00 & & \\ 0.80 & 1.00 & \\ 0.50 & 0.30 & 1.00 \end{array} \right)$$

The risk budgets are equal to $(50\%, 20\%, 30\%)$.

Gaussian RB portfolios

Table 28: Risk budgeting portfolios (Example 10)

$\mathcal{R}(x)$	Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^{\star}$
Volatility	1	31.14	28.08	8.74	50.00
	2	21.90	15.97	3.50	20.00
	3	46.96	11.17	5.25	30.00
	$\sigma(x)$			17.49	
Value-at-risk	1	29.18	54.47	15.90	50.00
	2	20.31	31.30	6.36	20.00
	3	50.50	18.89	9.54	30.00
	$\mathrm{VaR}_{99\%}(x)$			31.79	
Expected shortfall	1	29.48	64.02	18.87	50.00
	2	20.54	36.74	7.55	20.00
	3	49.98	22.65	11.32	30.00
	$\mathrm{ES}_{99\%}(x)$			37.74	

Special cases

- The case of uniform correlation[5] $\rho_{i,j} = \rho$
  1. Minimum correlation

$$x_i\left(-\frac{1}{n-1}\right) = \frac{\sigma_i^{-1}}{\sum_{j=1}^n \sigma_j^{-1}}$$

  2. Zero correlation

$$x_i\left(0\right) = \frac{\sqrt{b_i}\sigma_i^{-1}}{\sum_{j=1}^n \sqrt{b_j}\sigma_j^{-1}}$$

  3. Maximum correlation

$$x_i\left(1\right) = \frac{b_i\sigma_i^{-1}}{\sum_{j=1}^n b_j\sigma_j^{-1}}$$

- The general case

$$x_i = \frac{b_i\beta_i^{-1}}{\sum_{j=1}^n b_j\beta_j^{-1}}$$

where β_i is the beta of Asset i with respect to the RB portfolio

[5] The solution is noted $x_i\left(\rho\right)$.

Existence and uniqueness

We have:

$$\frac{\partial\,\sigma(x)}{\partial\,x_i} = \frac{x_i\sigma_i^2 + \sigma_i \sum_{j\neq i} x_j\rho_{i,j}\sigma_j}{\sigma(x)}$$

Suppose that the risk budget b_k is equal to zero. This means that:

$$x_k\left(x_k\sigma_k^2 + \sigma_k \sum_{j\neq k} x_j\rho_{k,j}\sigma_j\right) = 0$$

We obtain two solutions:

1. The first one is:
$$x_k' = 0$$
2. The second one verifies:
$$x_k'' = -\frac{\sum_{j\neq k} x_j\rho_{k,j}\sigma_j}{\sigma_k}$$

Existence and uniqueness

- If $\rho_{k,j} \geq 0$ for all j, we have $\sum_{j \neq k} x_j \rho_{k,j} \sigma_j \geq 0$ because $x_j \geq 0$ and $\sigma_j > 0$. This implies that $x''_k \leq 0$ meaning that $x'_k = 0$ is the unique positive solution
- The only way to have $x''_k > 0$ is to have some negative correlations $\rho_{k,j}$. In this case, this implies that:

$$\sum_{j \neq k} x_j \rho_{k,j} \sigma_j < 0$$

- If we consider a universe of three assets, this constraint is verified for $k = 3$ and a covariance matrix such that $\rho_{1,3} < 0$ and $\rho_{2,3} < 0$

Existence and uniqueness

Example 11

We have $\sigma_1 = 20\%$, $\sigma_2 = 10\%$, $\sigma_3 = 5\%$, $\rho_{1,2} = 50\%$, $\rho_{1,3} = -25\%$ and $\rho_{2,3} = -25\%$

If the risk budgets are equal to $(50\%, 50\%, 0\%)$, the two solutions are:

$$(33.33\%, 66.67\%, 0\%)$$

and:

$$(20\%, 40\%, 40\%)$$

Two questions

1. How many solutions do we have in the general case?
2. Which solution is the best?

Existence and uniqueness

Table 29: First solution (Example 11)

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	33.33	17.32	5.77	50.00
2	66.67	8.66	5.77	50.00
3	0.00	-1.44	0.00	0.00
Volatility			11.55	

Table 30: Second solution (Example 11)

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	20.00	16.58	3.32	50.00
2	40.00	8.29	3.32	50.00
3	40.00	0.00	0.00	0.00
Volatility			6.63	

Existence and uniqueness

The case with strictly positive risk budgets

- We consider the following optimization problem:

$$\begin{aligned} y^{\star} &= \arg\min \mathcal{R}(y) \\ \text{u.c.} & \left\{ \begin{array}{l} \sum_{i=1}^{n} b_i \ln y_i \geq c \\ y \geq \mathbf{0}_n \end{array} \right. \end{aligned}$$

  where c is an arbitrary constant
- The associated Lagrange function is:

$$\mathcal{L}(y;\lambda,\lambda_c) = \mathcal{R}(y) - \lambda^{\top} y - \lambda_c \left(\sum_{i=1}^{n} b_i \ln y_i - c \right)$$

  where $\lambda \in \mathbb{R}^n$ and $\lambda_c \in \mathbb{R}$

Existence and uniqueness

The case with strictly positive risk budgets

- The solution $y^{\star}$ verifies the following first-order condition:

$$\frac{\partial\,\mathcal{L}\left(y;\lambda,\lambda_c\right)}{\partial\,y_i}=\frac{\partial\,\mathcal{R}\left(y\right)}{\partial\,y_i}-\lambda_i-\lambda_c\frac{b_i}{y_i}=0$$

- The Kuhn-Tucker conditions are:

$$\left\{\begin{array}{l}\min\left(\lambda_i,y_i\right)=0\\ \min\left(\lambda_c,\sum_{i=1}^{n}b_i\ln y_i-c\right)=0\end{array}\right.$$

Existence and uniqueness

The case with strictly positive risk budgets

- Because $\ln y_i$ is not defined for $y_i = 0$, it follows that $y_i > 0$ and $\lambda_i = 0$
- We note that the constraint $\sum_{i=1}^{n} b_i \ln y_i = c$ is necessarily reached (because the solution cannot be $y^{\star} = \mathbf{0}_n$), then $\lambda_c > 0$ and we have:

$$y_i \frac{\partial \mathcal{R}(y)}{\partial y_i} = \lambda_c b_i$$

- We verify that the risk contributions are proportional to the risk budgets:

$$\mathcal{RC}_i = \lambda_c b_i$$

Existence and uniqueness

The case with strictly positive risk budgets

Theorem

The optimization program has a unique solution and the RB portfolio is equal to:

$$x_{\mathrm{rb}} = \frac{y^\star}{\sum_{i=1}^n y_i^\star}$$

Remark

We note that the convexity property of the risk measure is essential to the existence and uniqueness of the RB portfolio. If $\mathcal{R}(x)$ is not convex, the preceding analysis becomes invalid.

Existence and uniqueness

Effect on the solution of setting risk budgets to zero

- Let $\mathcal{N}$ be the set of assets such that $b_i = 0$
- The Lagrange function becomes:

$$\mathcal{L}(y; \lambda, \lambda_c) = \mathcal{R}(y) - \lambda^\top y - \lambda_c \left(\sum_{i \notin \mathcal{N}} b_i \ln y_i - c \right)$$

Existence and uniqueness

Effect on the solution of setting risk budgets to zero

- The solution $y^\star$ verifies the following first-order conditions:

$$\frac{\partial\,\mathcal{L}(y;\lambda,\lambda_c)}{\partial\,y_i}=\begin{cases}\partial_{y_i}\,\mathcal{R}(y)-\lambda_i-\lambda_c b_i y_i^{-1}=0 & \text{if} \quad i\notin\mathcal{N}\\ \partial_{y_i}\,\mathcal{R}(y)-\lambda_i=0 & \text{if} \quad i\in\mathcal{N}\end{cases}$$

  - If $i\notin\mathcal{N}$, the previous analysis is valid and we verify that risk contributions are proportional to the risk budgets:

$$y_i\frac{\partial\,\mathcal{R}(y)}{\partial\,y_i}=\lambda_c b_i$$

  - If $i\in\mathcal{N}$, we must distinguish two cases:
    1. If $y_i=0$, it implies that $\lambda_i>0$ and $\partial_{y_i}\,\mathcal{R}(y)>0$
    2. In the other case, if $y_i>0$, it implies that $\lambda_i=0$ and $\partial_{y_i}\,\mathcal{R}(y)=0$
  - The solution $y_i=0$ or $y_i>0$ if $i\in\mathcal{N}$ will then depend on the structure of the covariance matrix Σ (in the case of a Gaussian risk measure)

Existence and uniqueness

Effect on the solution of setting risk budgets to zero

Theorem

We conclude that the solution $y^{\star}$ of the optimization problem exists and is unique even if some risk budgets are set to zero. As previously, we deduce the normalized RB portfolio x_{rb} by scaling $y^{\star}$. This solution, noted $\mathcal{S}_1$, satisfies the following relationships:

$$\left\{ \begin{array}{ll} \mathcal{RC}_i = x_i \cdot \partial_{x_i} \mathcal{R}(x) = b_i & \text{if} \quad i \notin \mathcal{N} \\ \left\{ \begin{array}{c} x_i = 0 \text{ and } \partial_{x_i} \mathcal{R}(x) > 0 \quad (i) \\ \text{or} \\ x_i > 0 \text{ and } \partial_{x_i} \mathcal{R}(x) = 0 \quad (ii) \end{array} \right. & \text{if} \quad i \in \mathcal{N} \end{array} \right.$$

The conditions (i) and (ii) are mutually exclusive for one asset $i \in \mathcal{N}$, but not necessarily for all the assets $i \in \mathcal{N}$.

Existence and uniqueness

Effect on the solution of setting risk budgets to zero

The previous analysis implies that there may be several solutions to the following non-linear system when $b_i = 0$ for $i \in \mathcal{N}$:

$$\left\{ \begin{array}{c} \mathcal{RC}_1 = b_1 \mathcal{R}(x) \\ \vdots \\ \mathcal{RC}_i = b_i \mathcal{R}(x) \\ \vdots \\ \mathcal{RC}_n = b_n \mathcal{R}(x) \end{array} \right.$$

- Let $\mathcal{N} = \mathcal{N}_1 \bigsqcup \mathcal{N}_2$ where $\mathcal{N}_1$ is the set of assets verifying the condition (i) and $\mathcal{N}_2$ is the set of assets verifying the condition (ii)
- The number of solutions is equal to 2^m where $m = |\mathcal{N}_2|$ is the cardinality of $\mathcal{N}_2$

Existence and uniqueness

Effect on the solution of setting risk budgets to zero

We note $\mathcal{S}_2$ the solution with $x_i = 0$ for all assets such that $b_i = 0$. Even if $\mathcal{S}_2$ is the solution expected by the investor, the only acceptable solution is $\mathcal{S}_1$. Indeed, if we impose $b_i = \varepsilon_i$ where $\varepsilon_i > 0$ is a small number for $i \in \mathcal{N}$, we obtain:

$$\lim_{\varepsilon_i \to 0} \mathcal{S} = \mathcal{S}_1$$

The solution converges to $\mathcal{S}_1$, and not to $\mathcal{S}_2$ or the other solutions

Existence and uniqueness

Effect on the solution of setting risk budgets to zero

Remark

The non-linear system is not well-defined, whereas the optimization problem is the right approach to define a RB portfolio

Definition

A RB portfolio is a minimum risk portfolio subject to a diversification constraint, which is defined by the logarithmic barrier function

Existence and uniqueness

Example 12

We consider a universe of three assets with $\sigma_1 = 20\%$, $\sigma_2 = 10\%$ and $\sigma_3 = 5\%$. The correlation of asset returns is given by the following matrix:

$$\rho = \begin{pmatrix} 1.00 & & \\ 0.50 & 1.00 & \\ \rho_{1,3} & \rho_{2,3} & 1.00 \end{pmatrix}$$

We would like to build a RB portfolio such that the risk budgets with respect to the volatility risk measure are $(50\%, 50\%, 0\%)$. Moreover, we assume that $\rho_{1,3} = \rho_{2,3}$.

Existence and uniqueness

Table 31: RB solutions when the risk budget b_3 is equal to 0 (Example 12)

$\rho_{1,3} = \rho_{2,3}$	Solution		1	2	3	$\sigma(x)$
-25%	$\mathcal{S}_1$	x_i	20.00%	40.00%	40.00%	6.63%
		$\mathcal{MR}_i$	16.58%	8.29%	0.00%	
		$\mathcal{RC}_i$	50.00%	50.00%	0.00%	
	$\mathcal{S}_2$	x_i	33.33%	66.67%	0.00%	11.55%
		$\mathcal{MR}_i$	17.32%	8.66%	-1.44%	
		$\mathcal{RC}_i$	50.00%	50.00%	0.00%	
	$\mathcal{S}_1'$	x_i	19.23%	38.46%	42.31%	6.38%
		$\mathcal{MR}_i$	16.42%	8.21%	0.15%	
		$\mathcal{RC}_i$	49.50%	49.50%	1.00%	
25%	$\mathcal{S}_1$	x_i	33.33%	66.67%	0.00%	11.55%
		$\mathcal{MR}_i$	17.32%	8.66%	1.44%	
		$\mathcal{RC}_i$	50.00%	50.00%	0.00%	

Existence and uniqueness

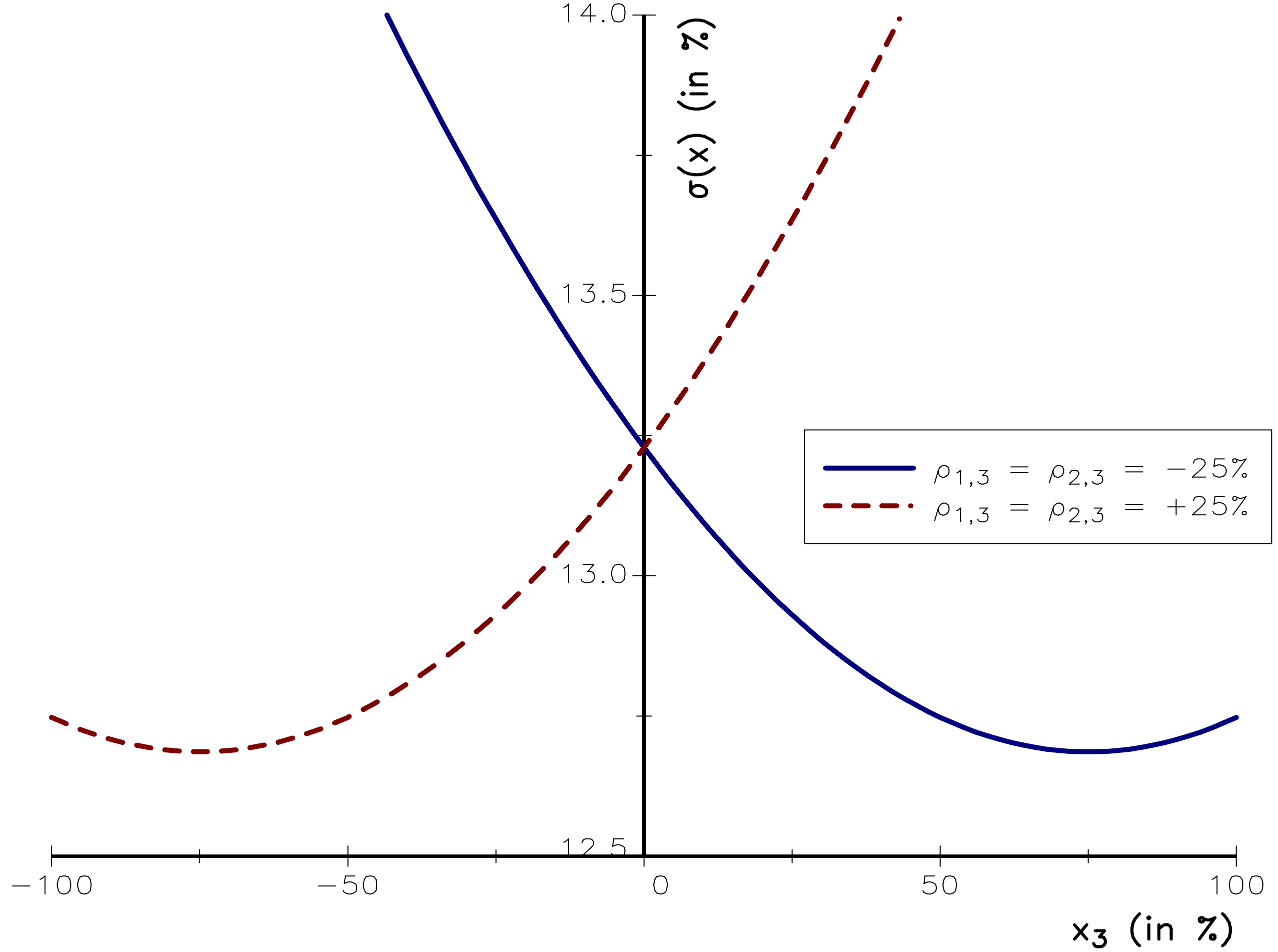


Figure 25: Evolution of the portfolio's volatility with respect to x_3

Location of the RB portfolio

We have:

$$\frac{x_i}{b_i} = \frac{x_j}{b_j} \tag{WB}$$

$$\frac{\partial\,\mathcal{R}(x)}{\partial\,x_i} = \frac{\partial\,\mathcal{R}(x)}{\partial\,x_j} \tag{MR}$$

$$\frac{1}{b_i}\left(x_i \frac{\partial\,\mathcal{R}(x)}{\partial\,x_i}\right) = \frac{1}{b_j}\left(x_j \frac{\partial\,\mathcal{R}(x)}{\partial\,x_j}\right) \tag{ERC}$$

The RB portfolio is a combination of MR (long-only minimum risk) and WB (weight budgeting) portfolios

Risk of the RB portfolio

Theorem

We obtain the following inequality:

$$\mathcal{R}(x_{\mathrm{mr}}) \leq \mathcal{R}(x_{\mathrm{rb}}) \leq \mathcal{R}(x_{\mathrm{wb}})$$

The RB portfolio may be viewed as a portfolio "between" the MR portfolio and the WB portfolio

Diversification index

Definition

The diversification index is equal to:

$$\begin{aligned}\mathcal{D}(x) &= \frac{\mathcal{R}\left(\sum_{i=1}^{n} L_i\right)}{\sum_{i=1}^{n} \mathcal{R}(L_i)} \\ &= \frac{\mathcal{R}(x)}{\sum_{i=1}^{n} x_i \mathcal{R}(\mathbf{e}_i)}\end{aligned}$$

Diversification index

- The diversification index is the ratio between the risk measure of portfolio x and the weighted risk measure of the assets
- If $\mathcal{R}$ is a coherent risk measure, we have $\mathcal{D}(x) \leq 1$
- If $\mathcal{D}(x) = 1$, it implies that the losses are comonotonic
- If $\mathcal{R}$ is the volatility risk measure, we obtain:

$$\mathcal{D}(x) = \frac{\sqrt{x^\top \Sigma x}}{\sum_{i=1}^n x_i \sigma_i}$$

It takes the value one if the asset returns are perfectly correlated meaning that the correlation matrix is $\mathcal{C}_n(1)$

Concentration index

- Let $\pi \in \mathbb{R}_+^n$ such that $\mathbf{1}_n^\top \pi = 1 \Rightarrow \pi$ is a probability distribution
- The probability distribution π^+ is perfectly concentrated if there exists one observation i_0 such that $\pi_{i_0}^+ = 1$ and $\pi_i^+ = 0$ if $i \neq i_0$
- When n tends to $+\infty$, the limit distribution is noted π_∞^+
- On the opposite, the probability distribution π^- such that $\pi_i^- = 1/n$ for all $i = 1, \ldots, n$ has no concentration

Concentration index

Definition

A concentration index is a mapping function $\mathcal{C}(\pi)$ such that $\mathcal{C}(\pi)$ increases with concentration and verifies:

$$\mathcal{C}\left(\pi^{-}\right) \leq \mathcal{C}(\pi) \leq \mathcal{C}\left(\pi^{+}\right)$$

- For instance, if π represents the weights of the portfolio, $\mathcal{C}(\pi)$ measures then the weight concentration
- By construction, $\mathcal{C}(\pi)$ reaches the minimum value if the portfolio is equally weighted
- To measure the risk concentration of the portfolio, we define π as the distribution of the risk contributions. In this case, the portfolio corresponding to the lower bound $\mathcal{C}\left(\pi^{-}\right) = 0$ is the ERC portfolio

Herfindahl index

Definition

The Herfindahl index associated with π is defined as:

$$\mathcal{H}(\pi) = \sum_{i=1}^{n} \pi_i^2$$

- This index takes the value 1 for the probability distribution π^+ and $1/n$ for the distribution with uniform probabilities π^-
- To scale the statistics onto $[0,1]$, we consider the normalized index $\mathcal{H}^\star(\pi)$ defined as follows:

$$\mathcal{H}^\star(\pi) = \frac{n\mathcal{H}(\pi) - 1}{n - 1}$$

Gini index

- The Gini index is based on the Lorenz curve of inequality
- Let X and Y be two random variables. The Lorenz curve $y = \mathbb{L}(x)$ is defined by the following parameterization:

$$\begin{cases} x = \Pr\{X \leq x\} \\ y = \Pr\{Y \leq y \mid X \leq x\} \end{cases}$$

- The Lorenz curve admits two limit cases
  1. If the portfolio is perfectly concentrated, the distribution of the weights corresponds to π_{∞}^{+}
  2. On the opposite, the least concentrated portfolio is the equally weighted portfolio and the Lorenz curve is the bisecting line $y = x$

Gini index

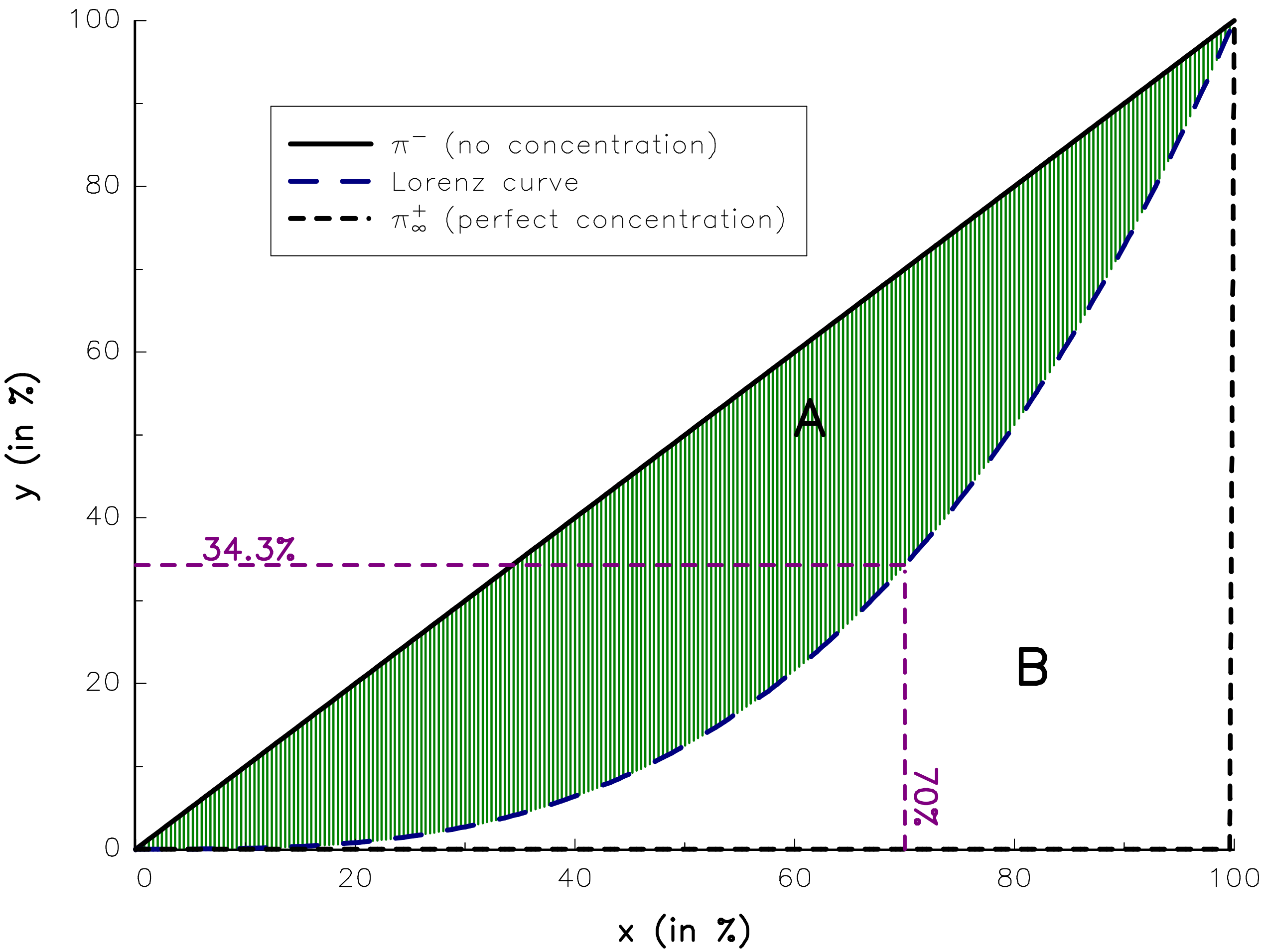


Figure 26: Geometry of the Lorenz curve

Gini index

Definition

The Gini index is then defined as:

$$\mathcal{G}(\pi) = \frac{A}{A+B}$$

with A the area between $\mathbb{L}(\pi^-)$ and $\mathbb{L}(\pi)$, and B the area between $\mathbb{L}(\pi)$ and $\mathbb{L}(\pi^+_\infty)$

Gini index

- By construction, we have $\mathcal{G}(\pi^-) = 0$, $\mathcal{G}(\pi^+_\infty) = 1$ and:

$$\begin{aligned}
\mathcal{G}(\pi) &= \frac{(A+B)-B}{A+B} \\
&= 1 - \frac{1}{A+B} B \\
&= 1 - 2\int_0^1 \mathbb{L}(x)\,\mathrm{d}x
\end{aligned}$$

In the case when π is a discrete probability distribution, we obtain:

$$\mathcal{G}(\pi) = \frac{2\sum_{i=1}^n i\pi_{i:n}}{n\sum_{i=1}^n \pi_{i:n}} - \frac{n+1}{n}$$

where $\{\pi_{1:n}, \ldots, \pi_{n:n}\}$ are the ordered statistics of $\{\pi_1, \ldots, \pi_n\}$.

Shannon entropy

Definition

The Shannon entropy is equal to:

$$\mathcal{I}(\pi) = -\sum_{i=1}^{n} \pi_i \ln \pi_i$$

- The diversity index corresponds to the statistic:

$$\mathcal{I}^{\star}(\pi) = \exp(\mathcal{I}(\pi))$$

- We have $\mathcal{I}^{\star}(\pi^{-}) = n$ and $\mathcal{I}^{\star}(\pi^{+}) = 1$

Impact of the reparametrization on the asset universe

- We consider a set of m primary assets $(\mathcal{A}'_1, \ldots, \mathcal{A}'_m)$ with a covariance matrix Ω
- We define n synthetic assets $(\mathcal{A}_1, \ldots, \mathcal{A}_n)$ which are composed of the primary assets
- We denote $W = (w_{i,j})$ the weight matrix such that $w_{i,j}$ is the weight of the primary asset $\mathcal{A}'_j$ in the synthetic asset $\mathcal{A}_i$ (we have $\sum_{j=1}^{m} w_{i,j} = 1$)
- The covariance matrix of the synthetic assets Σ is equal to $W \Omega W^{\top}$
- The synthetic assets can be interpreted as portfolios of the primary assets
- For example, $\mathcal{A}'_j$ may represent a stock whereas $\mathcal{A}_i$ may be an index

Impact of the reparametrization on the asset universe

1. We consider a portfolio $x = (x_1, \ldots, x_n)$ defined with respect to the synthetic assets. We have:

$$\mathcal{RC}_i = x_i \cdot \frac{(\Sigma x)_i}{\sqrt{x^\top \Sigma x}}$$

2. We also define the portfolio with respect to the primary assets. In this case, the composition is $y = (y_1, \ldots, y_m)$ where $y_j = \sum_{i=1}^n x_i w_{i,j}$ (or $y = W^\top x$). We have:

$$\mathcal{RC}_j = y_j \cdot \frac{(\Omega y)_j}{\sqrt{y^\top \Omega y}}$$

Impact of the reparametrization on the asset universe

Example 13

We have six primary assets. The volatility of these assets is respectively 20%, 30%, 25%, 15%, 10% and 30%. We assume that the assets are not correlated. We consider two equally weighted synthetic assets with:

$$W = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 & & \\ & & 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

Impact of the reparametrization on the asset universe

Table 32: Risk decomposition of Portfolio #1 with respect to the synthetic assets (Example 13)

Asset i	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
$\mathcal{A}_1$	36.00	9.44	3.40	33.33
$\mathcal{A}_2$	38.00	8.90	3.38	33.17
$\mathcal{A}_3$	26.00	13.13	3.41	33.50

Table 33: Risk decomposition of Portfolio #1 with respect to the primary assets (Example 13)

Asset j	y_j	$\mathcal{MR}_j$	$\mathcal{RC}_j$	$\mathcal{RC}_j^\star$
$\mathcal{A}'_1$	9.00	3.53	0.32	3.12
$\mathcal{A}'_2$	9.00	7.95	0.72	7.02
$\mathcal{A}'_3$	31.50	19.31	6.08	59.69
$\mathcal{A}'_4$	31.50	6.95	2.19	21.49
$\mathcal{A}'_5$	9.50	0.93	0.09	0.87
$\mathcal{A}'_6$	9.50	8.39	0.80	7.82

Impact of the reparametrization on the asset universe

Table 34: Risk decomposition of Portfolio #2 with respect to the synthetic assets (Example 13)

Asset i	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
$\mathcal{A}_1$	48.00	9.84	4.73	49.91
$\mathcal{A}_2$	50.00	9.03	4.51	47.67
$\mathcal{A}_3$	2.00	11.45	0.23	2.42

Table 35: Risk decomposition of Portfolio #2 with respect to the primary assets (Example 13)

Asset j	y_j	$\mathcal{MR}_j$	$\mathcal{RC}_j$	$\mathcal{RC}_j^\star$
$\mathcal{A}'_1$	12.00	5.07	0.61	6.43
$\mathcal{A}'_2$	12.00	11.41	1.37	14.46
$\mathcal{A}'_3$	25.50	16.84	4.29	45.35
$\mathcal{A}'_4$	25.50	6.06	1.55	16.33
$\mathcal{A}'_5$	12.50	1.32	0.17	1.74
$\mathcal{A}'_6$	12.50	11.88	1.49	15.69

Impact of the reparametrization on the asset universe

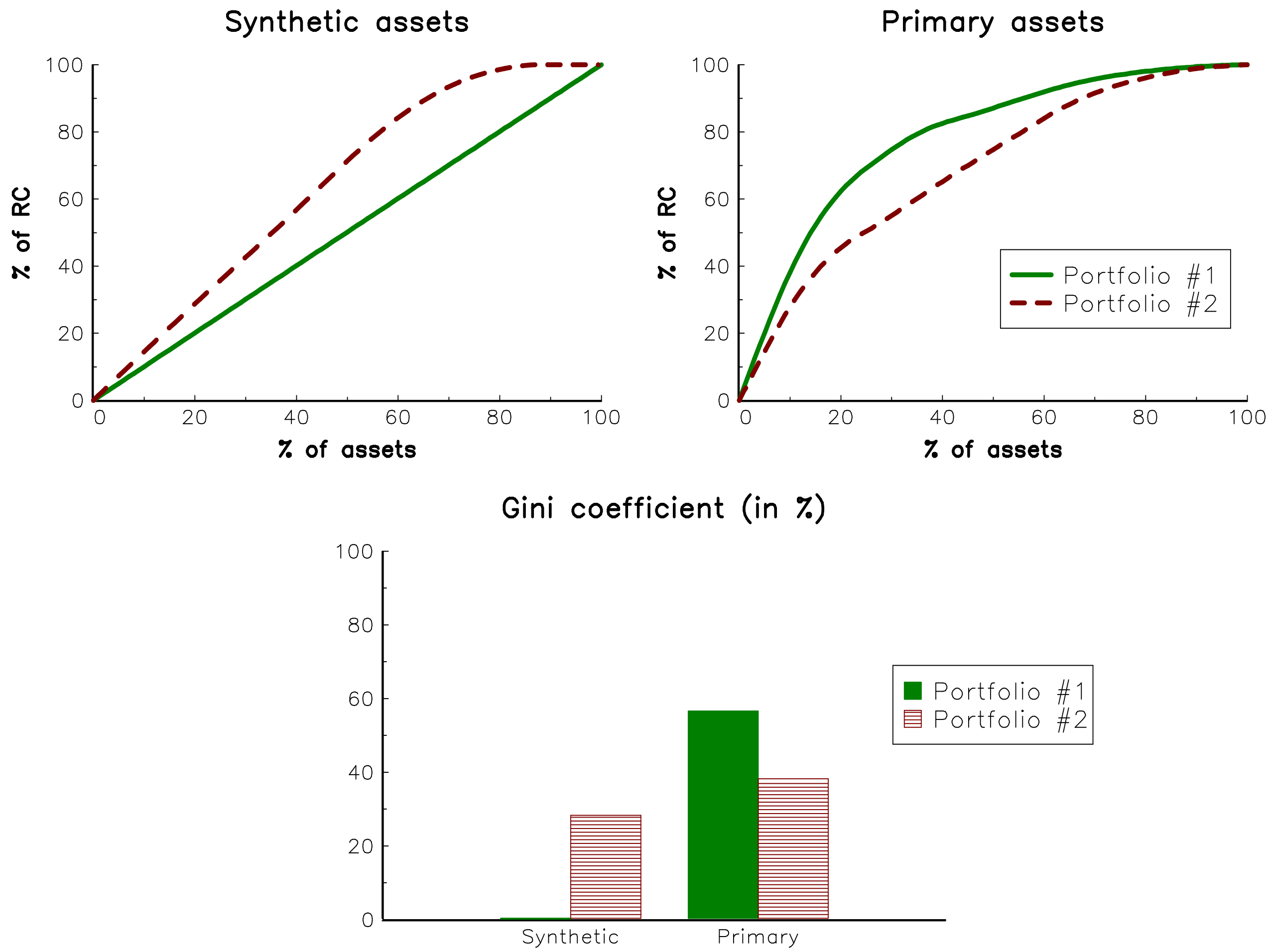


Figure 27: Lorenz curve of risk contributions (Example 13)

Risk decomposition with respect to the risk factors

- We consider a set of n assets $\{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ and a set of m risk factors $\{\mathcal{F}_1, \ldots, \mathcal{F}_m\}$
- R_t is the $(n \times 1)$ vector of asset returns at time t
- Σ is the covariance matrix of asset returns
- $\mathcal{F}_t$ is the $(m \times 1)$ vector of factor returns at time t
- Ω is the covariance matrix of factor returns

Risk decomposition with respect to the risk factors

Linear factor model

We consider the linear factor model:

$$R_t = A\mathcal{F}_t + \varepsilon_t$$

where $\mathcal{F}_t$ and ε_t are two uncorrelated random vectors, ε_t is a centered random vector $(n \times 1)$ of covariance D and A is the $(n \times m)$ loadings matrix

We have the following relationship:

$$\Sigma = A\Omega A^\top + D$$

Risk decomposition with respect to the risk factors

We decompose the portfolio's asset exposures x by the portfolio's risk factors exposures y in the following way:

$$x = B^{+}y + \tilde{B}^{+}\tilde{y}$$

where:

- B^{+} is the Moore-Penrose inverse of $A^{\top}$
- $\tilde{B}^{+}$ is any $n \times (n - m)$ matrix that spans the left nullspace of B^{+}
- $\tilde{y}$ corresponds to $n - m$ residual (or additional) factors that have no economic interpretation

It follows that:

$$\left\{ \begin{array}{l} y = A^{\top}x \\ \tilde{y} = \tilde{B}x \end{array} \right.$$

where $\tilde{B} = \ker\left(A^{\top}\right)^{\top}$

Risk decomposition with respect to the risk factors

Risk decomposition I

- We can show that the marginal risk of the j^{th} factor exposure is given by:

$$\mathcal{MR}\left(\mathcal{F}_j\right) = \frac{\partial\,\sigma\left(x\right)}{\partial\,y_j} = \frac{\left(A^{+}\Sigma x\right)_j}{\sigma\left(x\right)}$$

whereas its risk contribution is equal to:

$$\mathcal{RC}\left(\mathcal{F}_j\right) = y_j \frac{\partial\,\sigma\left(x\right)}{\partial\,y_j} = \frac{\left(A^{\top} x\right)_j \cdot \left(A^{+}\Sigma x\right)_j}{\sigma\left(x\right)}$$

Risk decomposition with respect to the risk factors

Risk decomposition II

- For the residual factors, we have:

$$\mathcal{MR}\left(\tilde{\mathcal{F}}_j\right) = \frac{\partial\, \sigma\left(x\right)}{\partial\, \tilde{y}_j} = \frac{\left(\tilde{B}\Sigma x\right)_j}{\sigma\left(x\right)}$$

and:

$$\mathcal{RC}\left(\tilde{\mathcal{F}}_j\right) = \tilde{y}_j \frac{\partial\, \sigma\left(x\right)}{\partial\, \tilde{y}_j} = \frac{\left(\tilde{B}x\right)_j \cdot \left(\tilde{B}\Sigma x\right)_j}{\sigma\left(x\right)}$$

Risk decomposition with respect to the risk factors

Remark

We can show that these risk contributions satisfy the allocation principle:

$$\sigma(x) = \sum_{j=1}^{m} \mathcal{RC}(\mathcal{F}_j) + \sum_{j=1}^{n-m} \mathcal{RC}(\tilde{\mathcal{F}}_j)$$

Risk decomposition with respect to the risk factors

Let $\operatorname{pinv}(C)$ and $\operatorname{null}(C)$ be the Moore-Penrose pseudo-inverse and the orthonormal basis for the right null space of C

1. Computation of A^{+}

$$A^{+} = \operatorname{pinv}(A) = \left(A^{\top} A\right)^{-1} A^{\top}$$

2. Computation of B

$$B = A^{\top}$$

3. Computation of B^{+}

$$B^{+} = \operatorname{pinv}(B) = B^{\top} \left(B B^{\top}\right)^{-1}$$

4. Computation of $\tilde{B}$

$$\tilde{B} = \operatorname{pinv}\left(\operatorname{null}\left({B^{+}}^{\top}\right)\right) \cdot \left(I_n - B^{+} A^{\top}\right)$$

Risk decomposition with respect to the risk factors

Remark

The previous results can be extended to other coherent and convex risk measures (Roncalli and Weisang, 2016)

Risk decomposition with respect to the risk factors

Example 14

We consider an investment universe with four assets and three factors. The loadings matrix A is:

$$A = \begin{pmatrix} 0.9 & 0.0 & 0.5 \\ 1.1 & 0.5 & 0.0 \\ 1.2 & 0.3 & 0.2 \\ 0.8 & 0.1 & 0.7 \end{pmatrix}$$

The three factors are uncorrelated and their volatilities are 20%, 10% and 10%. We assume a diagonal matrix D with specific volatilities 10%, 15%, 10% and 15%.

Risk decomposition with respect to the risk factors

The correlation matrix of asset returns is (in %):

$$\rho = \begin{pmatrix} 100.0 & & & \\ 69.0 & 100.0 & & \\ 79.5 & 76.4 & 100.0 & \\ 66.2 & 57.2 & 66.3 & 100.0 \end{pmatrix}$$

and their volatilities are respectively equal to 21.19%, 27.09%, 26.25% and 23.04%.

Risk decomposition with respect to the risk factors

We obtain that:

$$A^{+} = \begin{pmatrix} 1.260 & -0.383 & 1.037 & -1.196 \\ -3.253 & 2.435 & -1.657 & 2.797 \\ -0.835 & 0.208 & -1.130 & 2.348 \end{pmatrix}$$

and:

$$\tilde{B} = \begin{pmatrix} 0.533 & 0.452 & -0.692 & -0.183 \end{pmatrix}$$

Risk decomposition with respect to the risk factors

Table 36: Risk decomposition of the EW portfolio with respect to the assets (Example 14)

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	25.00	18.81	4.70	21.97
2	25.00	23.72	5.93	27.71
3	25.00	24.24	6.06	28.32
4	25.00	18.83	4.71	22.00
Volatility			21.40	

Table 37: Risk decomposition of the EW portfolio with respect to the risk factors (Example 14)

Factor	y_j	$\mathcal{MR}_j$	$\mathcal{RC}_j$	$\mathcal{RC}_j^\star$
$\mathcal{F}_1$	100.00	17.22	17.22	80.49
$\mathcal{F}_2$	22.50	9.07	2.04	9.53
$\mathcal{F}_3$	35.00	6.06	2.12	9.91
$\tilde{\mathcal{F}}_1$	2.75	0.52	0.01	0.07
Volatility			21.40	

Risk factor parity (or RFP) portfolios

RFP portfolios are defined by:

$$\mathcal{RC}\left(\mathcal{F}_j\right) = b_j \mathcal{R}\left(x\right)$$

They are computed using the following optimization problem:

$$\begin{aligned}\left(y^\star, \breve{y}^\star\right) &= \arg\min \sum_{j=1}^{m} \left(\mathcal{RC}\left(\mathcal{F}_j\right) - b_j \mathcal{R}\left(x\right)\right)^2 \\ \text{u.c.} \quad & \mathbf{1}_n^\top \left(B^+ y + \breve{B}^+ \breve{y}\right) = 1\end{aligned}$$

Risk factor parity (or RFP) portfolios

Example 15

We consider an investment universe with four assets and three factors. The loadings matrix A is:

$$A = \begin{pmatrix} 0.9 & 0.0 & 0.5 \\ 1.1 & 0.5 & 0.0 \\ 1.2 & 0.3 & 0.2 \\ 0.8 & 0.1 & 0.7 \end{pmatrix}$$

The three factors are uncorrelated and their volatilities are 20%, 10% and 10%. We assume a diagonal matrix D with specific volatilities 10%, 15%, 10% and 15%. We consider the following factor risk budgets:

$$b = (49\%, 25\%, 25\%)$$

Risk factor parity (or RFP) portfolios

Table 38: Risk decomposition of the RFP portfolio with respect to the risk factors (Example 15)

Factor	y_j	$\mathcal{MR}_j$	$\mathcal{RC}_j$	$\mathcal{RC}_j^\star$
$\mathcal{F}_1$	93.38	11.16	10.42	49.00
$\mathcal{F}_2$	24.02	22.14	5.32	25.00
$\mathcal{F}_3$	39.67	13.41	5.32	25.00
$\tilde{\mathcal{F}}_1$	16.39	1.30	0.21	1.00
Volatility			21.27	

Table 39: Risk decomposition of the RFP portfolio with respect to the assets (Example 15)

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	15.08	17.44	2.63	12.36
2	38.38	23.94	9.19	43.18
3	0.89	21.82	0.20	0.92
4	45.65	20.29	9.26	43.54
Volatility			21.27	

Minimizing the risk concentration between the risk factors

We now consider the following problem:

$$\mathcal{RC}\left(\mathcal{F}_j\right) \simeq \mathcal{RC}\left(\mathcal{F}_k\right)$$

$\Rightarrow$ The portfolios are computed by minimizing the risk concentration between the risk factors

Remark

We can use the Herfindahl index, the Gini index or the Shanon entropy

Minimizing the risk concentration between the risk factors

Example 16

We consider an investment universe with four assets and three factors. The loadings matrix A is:

$$A = \begin{pmatrix} 0.9 & 0.0 & 0.5 \\ 1.1 & 0.5 & 0.0 \\ 1.2 & 0.3 & 0.2 \\ 0.8 & 0.1 & 0.7 \end{pmatrix}$$

The three factors are uncorrelated and their volatilities are 20%, 10% and 10%. We assume a diagonal matrix D with specific volatilities 10%, 15%, 10% and 15%.

Minimizing the risk concentration between the risk factors

Table 40: Risk decomposition of the balanced RFP portfolio with respect to the risk factors (Example 16)

Factor	y_j	$\mathcal{MR}_j$	$\mathcal{RC}_j$	$\mathcal{RC}_j^\star$
$\mathcal{F}_1$	91.97	7.91	7.28	33.26
$\mathcal{F}_2$	25.78	28.23	7.28	33.26
$\mathcal{F}_3$	42.22	17.24	7.28	33.26
$\tilde{\mathcal{F}}_1$	6.74	0.70	0.05	0.21
Volatility			21.88	

Table 41: Risk decomposition of the balanced RFP portfolio with respect to the assets (Example 16)

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	0.30	16.11	0.05	0.22
2	39.37	23.13	9.11	41.63
3	0.31	20.93	0.07	0.30
4	60.01	21.09	12.66	57.85
Volatility			21.88	

Minimizing the risk concentration between the risk factors

We have $\mathcal{H}^{\star} = 0$, $\mathcal{G} = 0$ and $\mathcal{I}^{\star} = 3$

Minimizing the risk concentration between the risk factors

Table 42: Balanced RFP portfolios with $x_i \geq 10\%$ (Example 16)

Criterion	$\mathcal{H}(x)$	$\mathcal{G}(x)$	$\mathcal{I}(x)$
x_1	10.00	10.00	10.00
x_2	22.08	18.24	24.91
x_3	10.00	10.00	10.00
x_4	57.92	61.76	55.09
$\mathcal{H}^{\star}$	0.0436	0.0490	0.0453
$\mathcal{G}$	0.1570	0.1476	0.1639
$\mathcal{I}^{\star}$	2.8636	2.8416	2.8643

Justification of diversified funds

Investor Profiles

1. Conservative (low risk)
2. Moderate (medium risk)
3. Aggressive (high risk)

Fund Profiles

1. Defensive (20% equities and 80% bonds)
2. Balanced (50% equities and 50% bonds)
3. Dynamic (80% equities and 20% bonds)

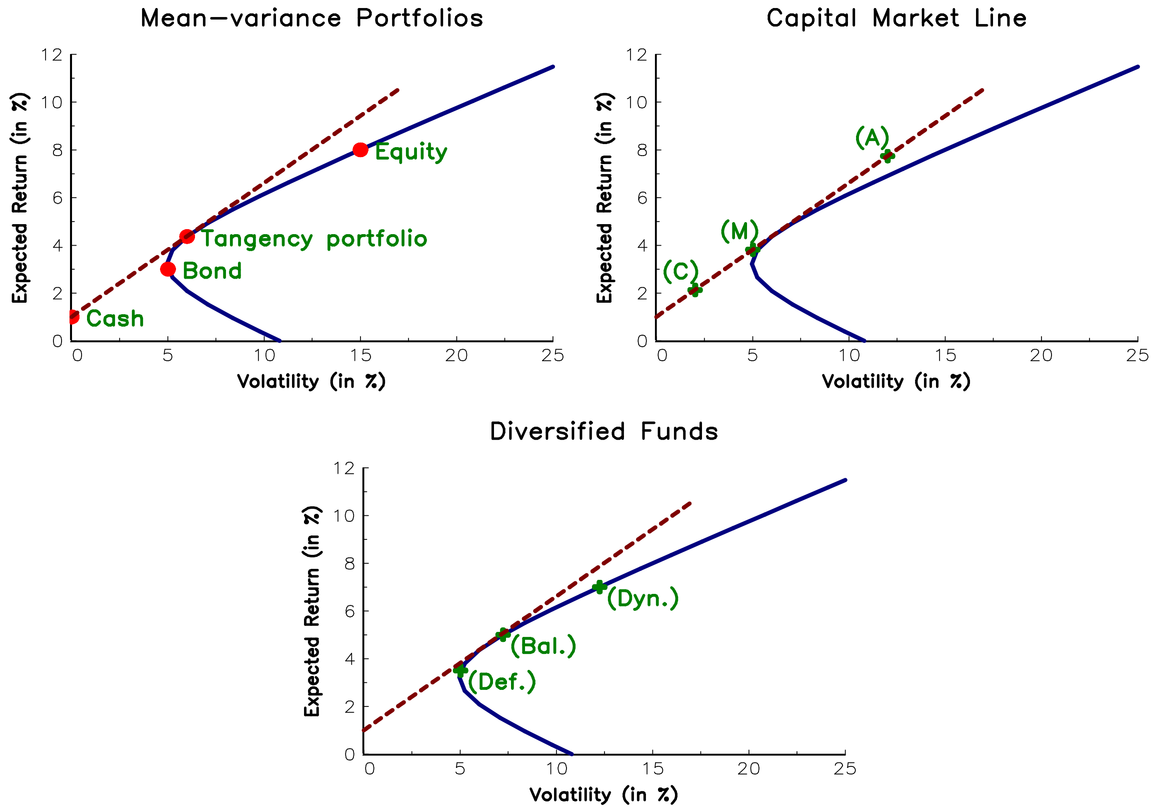


Figure 28: The asset allocation puzzle

What type of diversification is offered by diversified funds?

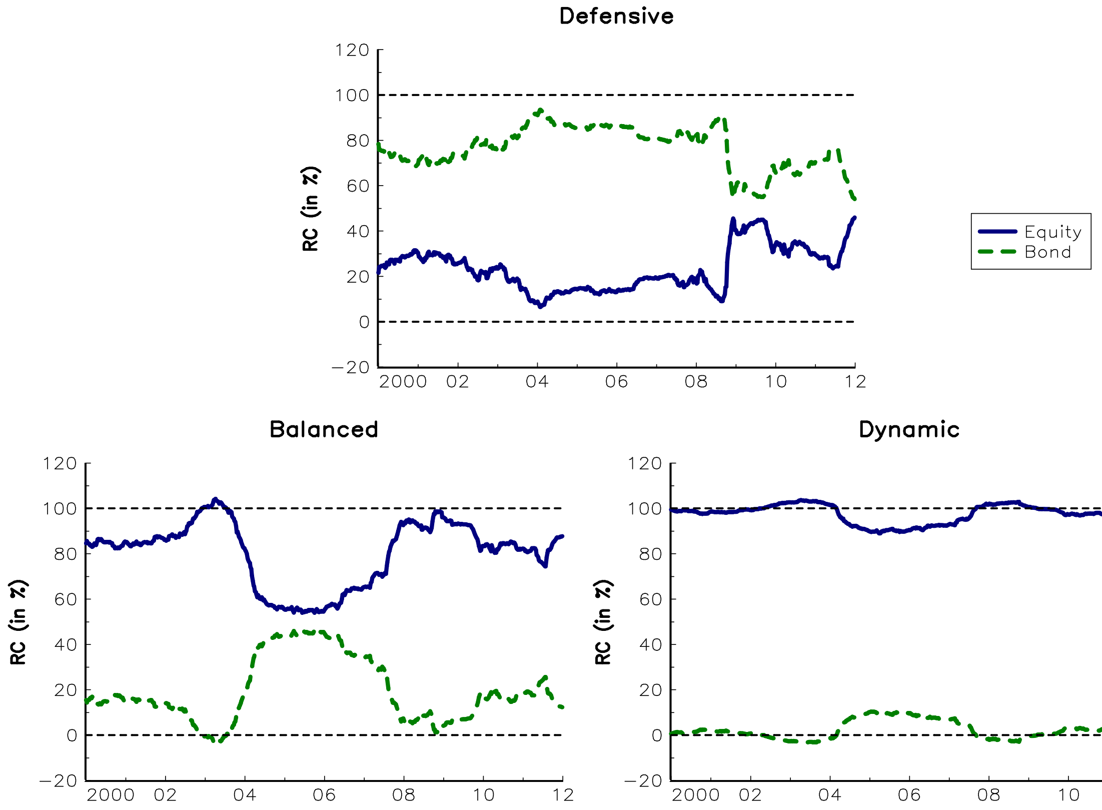


Figure 29: Equity (MSCI World) and bond (WGBI) risk contributions

Diversified funds
=
Marketing idea?

- Contrarian constant-mix strategy
- Deleverage of an equity exposure
- Low risk diversification
- No mapping between fund profiles and investor profiles
- Static weights
- Dynamic risk contributions

Risk-balanced allocation

- Multi-dimensional target volatility strategy
- Trend-following portfolio (if negative correlation between return and risk)
- Dynamic weights
- Static risk contributions (risk budgeting)
- High diversification

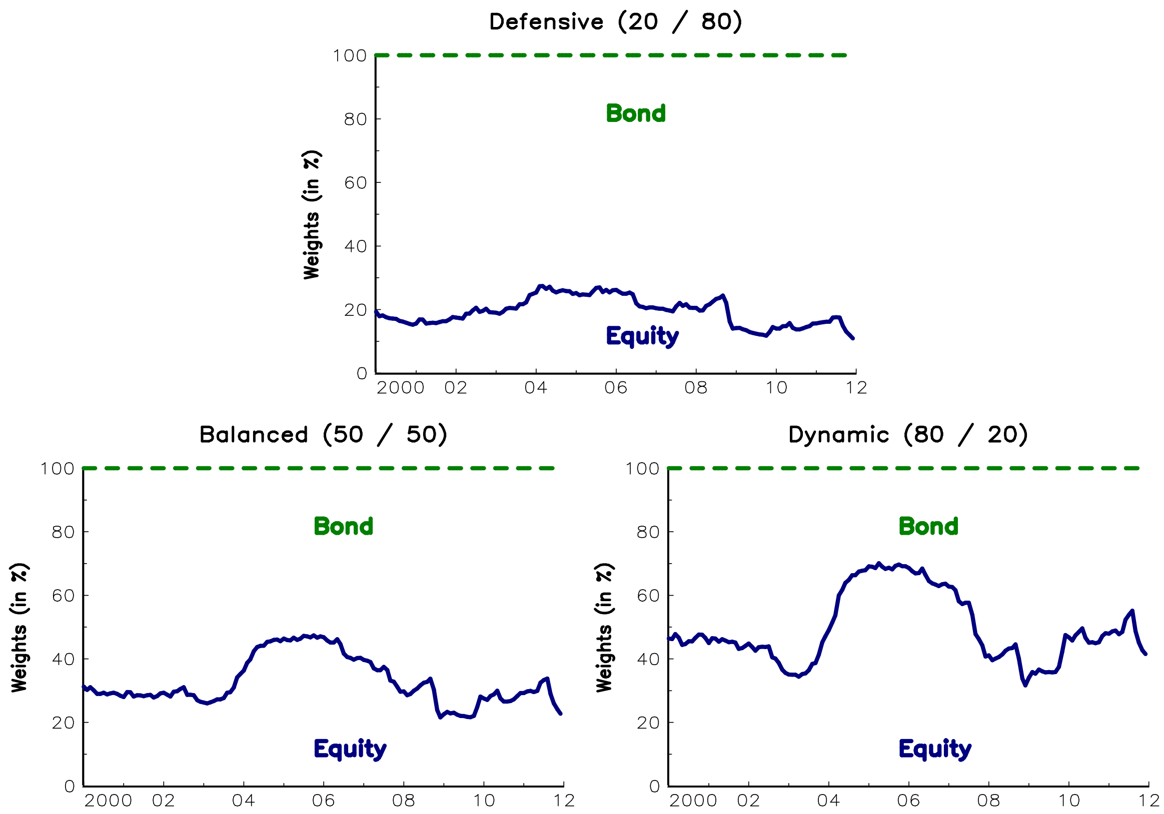


Figure 30: Equity and bond allocation

Characterization of the stock/bond market portfolio

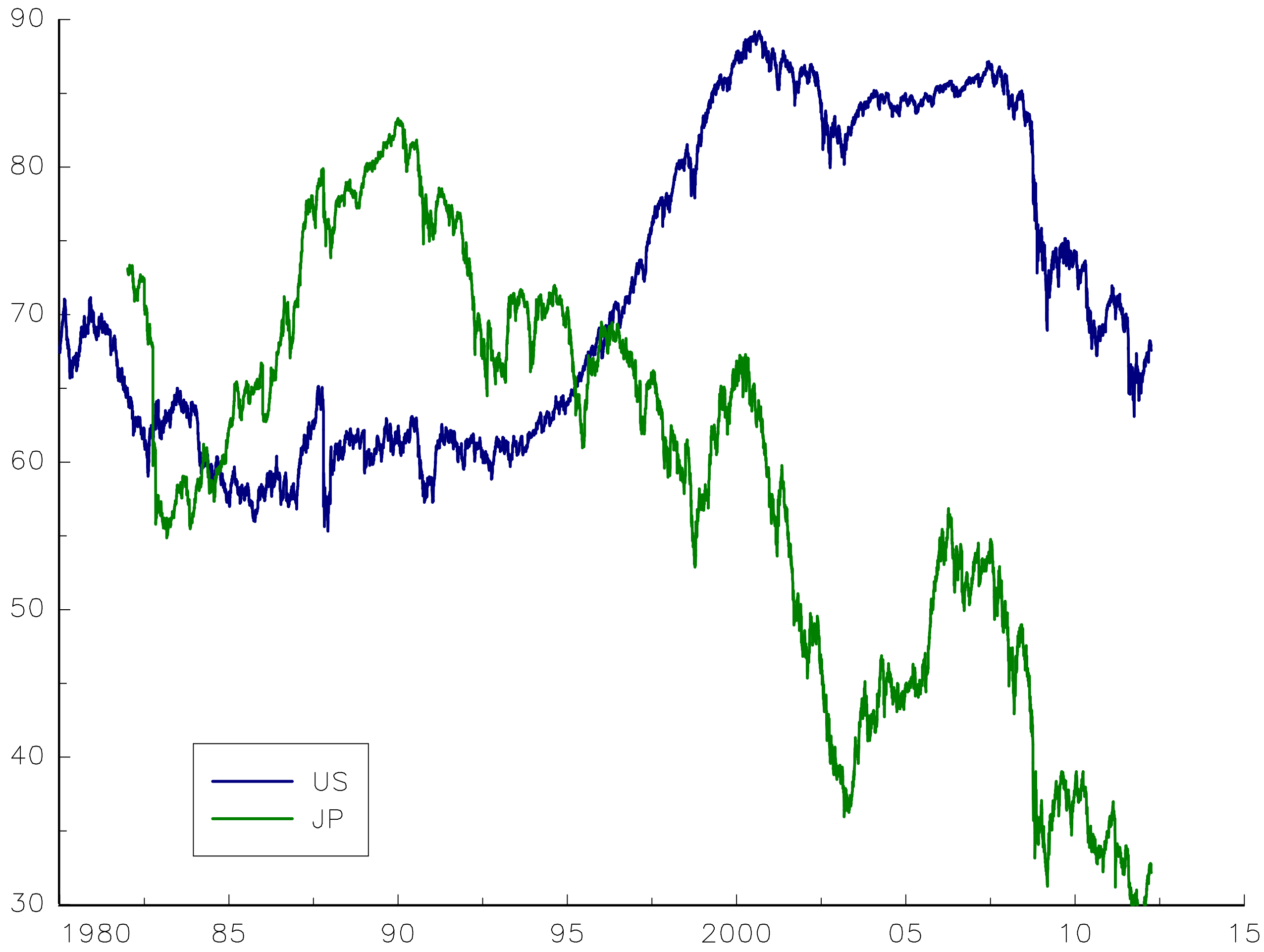


Figure 31: Evolution of the equity weight for United States and Japan

Characterization of the stock/bond market portfolio

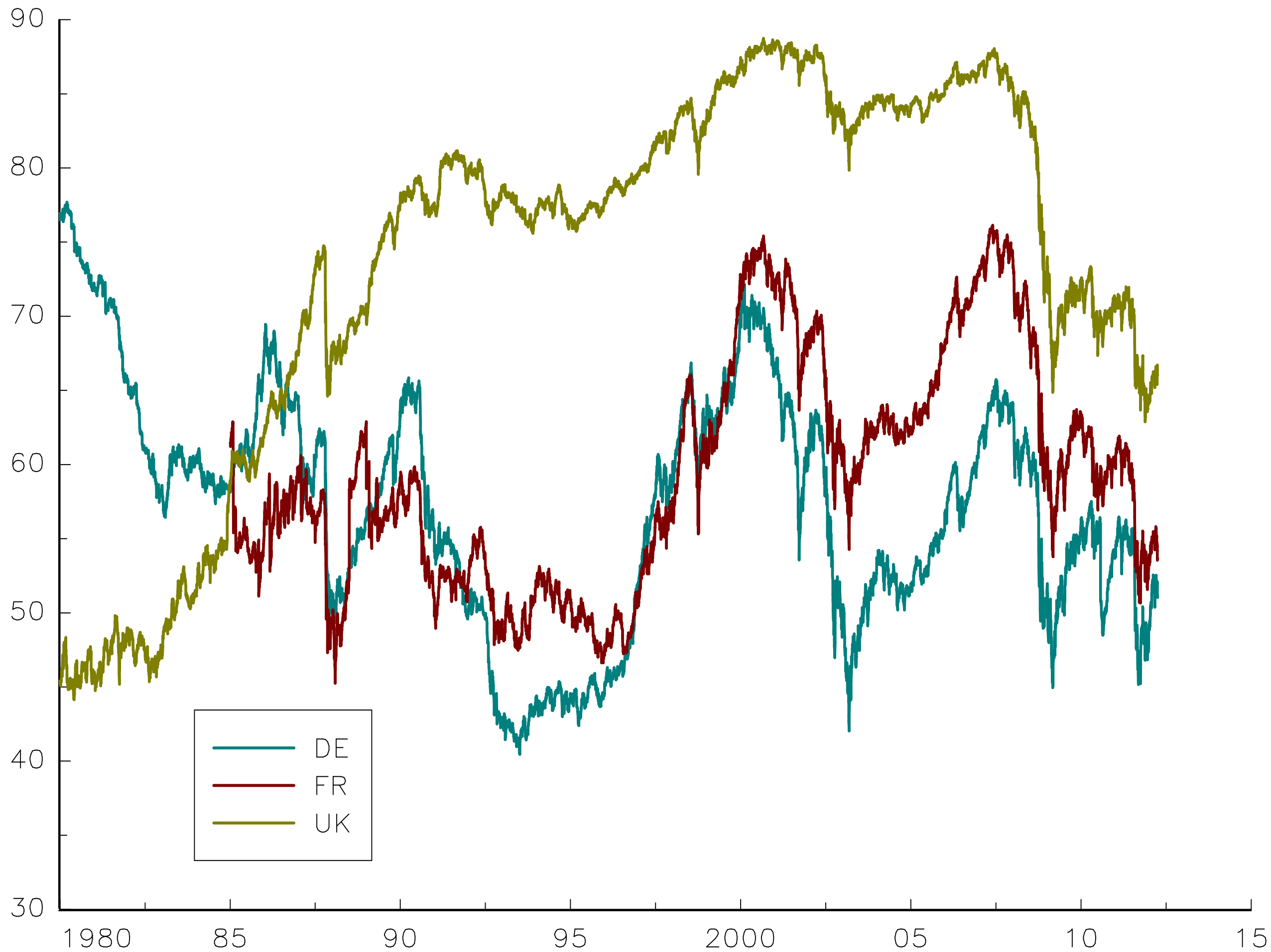


Figure 32: Evolution of the equity weight for Germany, France and UK

Link between risk premium and risk contribution

Let π_i and π_M be the risk premium of Asset i and the market risk premium. We have:

$$\begin{aligned}
\pi_i &= \beta_i \cdot \pi_M \\
&= \frac{\operatorname{cov}(R_i, R_M)}{\sigma(R_M)} \cdot \frac{\pi_M}{\sigma(R_M)} \\
&= \frac{\partial\, \sigma(x_M)}{\partial\, x_i} \cdot \mathrm{SR}(x_M)
\end{aligned}$$

The risk premium of Asset i is then proportional to the marginal volatility of Asset i with respect to the market portfolio

Foundation of the risk budgeting approach

For the tangency portfolio, we have:

performance contribution = risk contribution

Link between risk premium and risk contribution

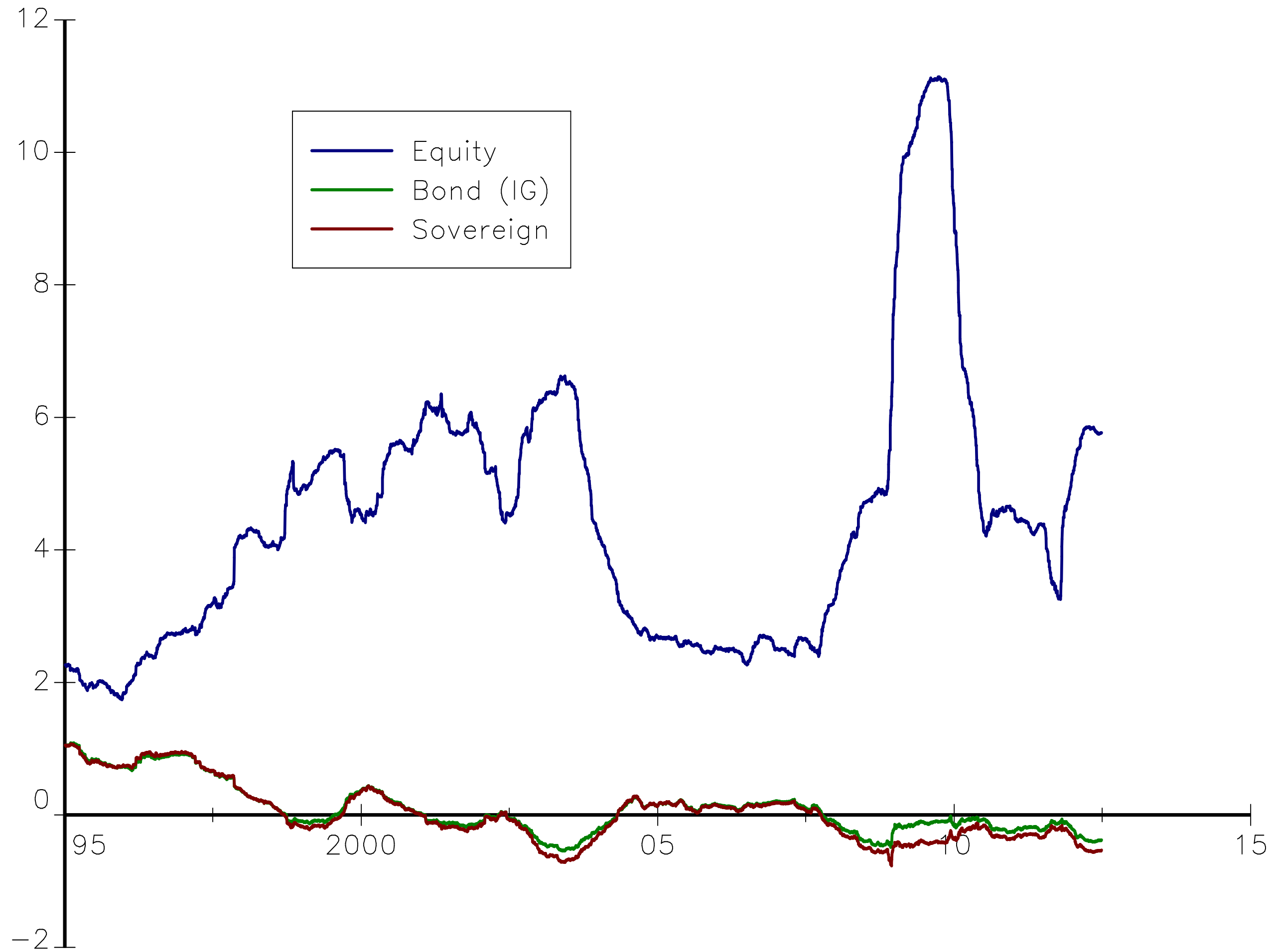


Figure 33: Risk premia (in %) for the US market portfolio ($\mathrm{SR}(x_M) = 25\%$)

Link between risk premium and risk contribution

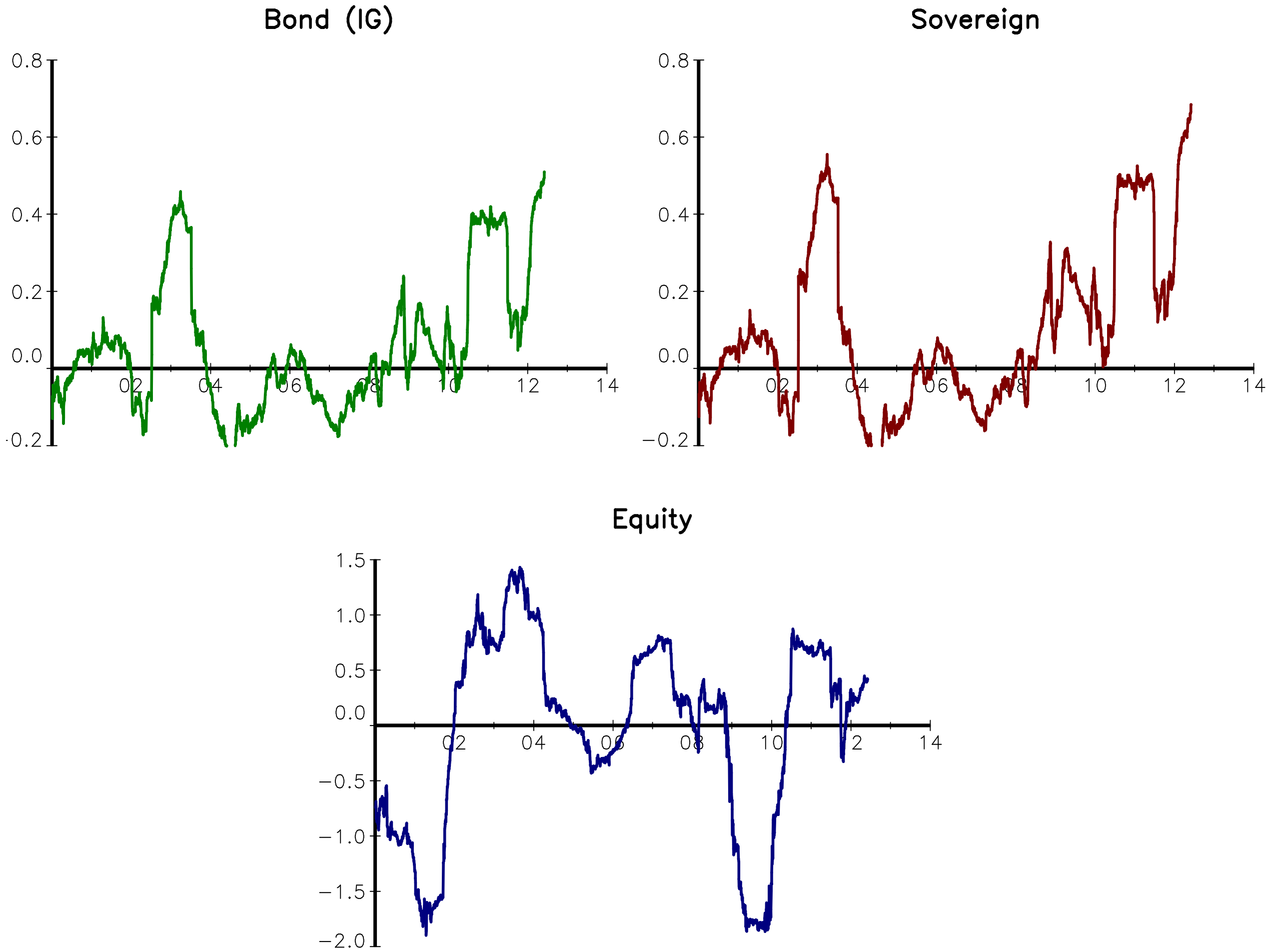


Figure 34: Difference (in %) between EURO and US risk premia ($\mathrm{SR}(x_M) = 25\%$)

Sharpe theory of risk premia

The one-factor risk model

We deduce that:

$$R_i - R_f = \alpha_i + \underbrace{\beta_i \cdot (R_M - R_f)}_{\text{Systematic Risk}} + \underbrace{\varepsilon_i}_{\text{Specific Risk}}$$

We necessarily have:

1. $\alpha_i = 0$
2. $\mathbb{E}[\varepsilon_i] = 0$

$\Rightarrow$ On average, only the systematic risk is rewarded, not the idiosyncratic risk

Sharpe theory of risk premia

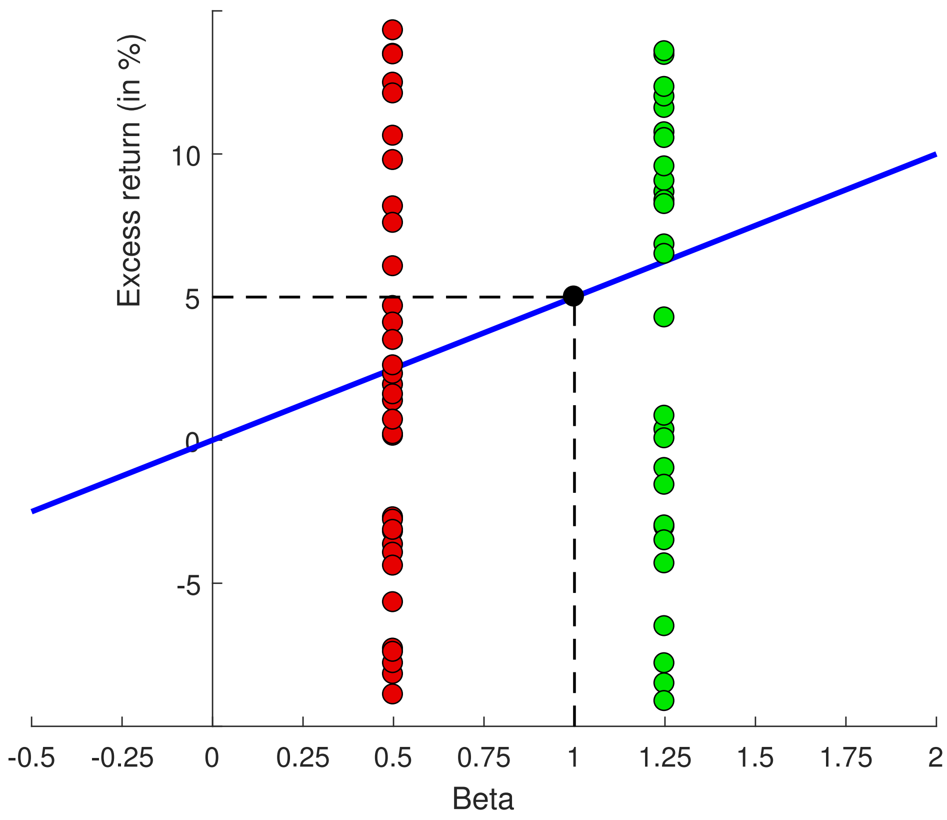


Figure 35: The security market line (SML)

- Risk premium is an increasing function of the systematic risk
- Risk premium may be negative (meaning that some assets can have a return lower than the risk-free asset!)
- More risk $\neq$ more return

Black-Litterman theory of risk premia

In the Black-Litterman model, the expected (or ex-ante/implied) risk premia are equal to:

$$\tilde{\pi} = \tilde{\mu} - r = \mathrm{SR}\left(x \mid r\right) \frac{\Sigma x}{\sqrt{x^\top \Sigma x}}$$

where $\mathrm{SR}\left(x \mid r\right)$ is the expected Sharpe ratio of the portfolio.

Black-Litterman theory of risk premia

Example 17

We consider four assets. Their expected returns are equal to 5%, 6%, 8% and 6% while their volatilities are equal to 15%, 20%, 25% and 30%. The correlation matrix of asset returns is given by the following matrix:

$$C = \begin{pmatrix} 1.00 & & & \\ 0.10 & 1.00 & & \\ 0.40 & 0.70 & 1.00 & \\ 0.50 & 0.40 & 0.80 & 1.00 \end{pmatrix}$$

We also assume that the return of the risk-free asset is equal to 1.5%.

Black-Litterman theory of risk premia

Table 43: Black-Litterman risk premia (Example 17)

Asset	CAPM		Black-Litterman			
	π_i	$x_i^\star$	x_i	$\tilde{\pi}_i$	x_i	$\tilde{\pi}_i$
#1	3.50%	63.63%	25.00%	2.91%	40.00%	3.33%
#2	4.50%	19.27%	25.00%	4.71%	30.00%	4.97%
#3	6.50%	50.28%	25.00%	7.96%	20.00%	7.69%
#4	4.50%	−33.17%	25.00%	9.07%	10.00%	8.18%
$\mu(x)$	6.37%		6.25%		6.00%	
$\sigma(x)$	14.43%		18.27%		15.35%	
$\tilde{\mu}(x)$	6.37%		7.66%		6.68%	

Black-Litterman theory of risk premia

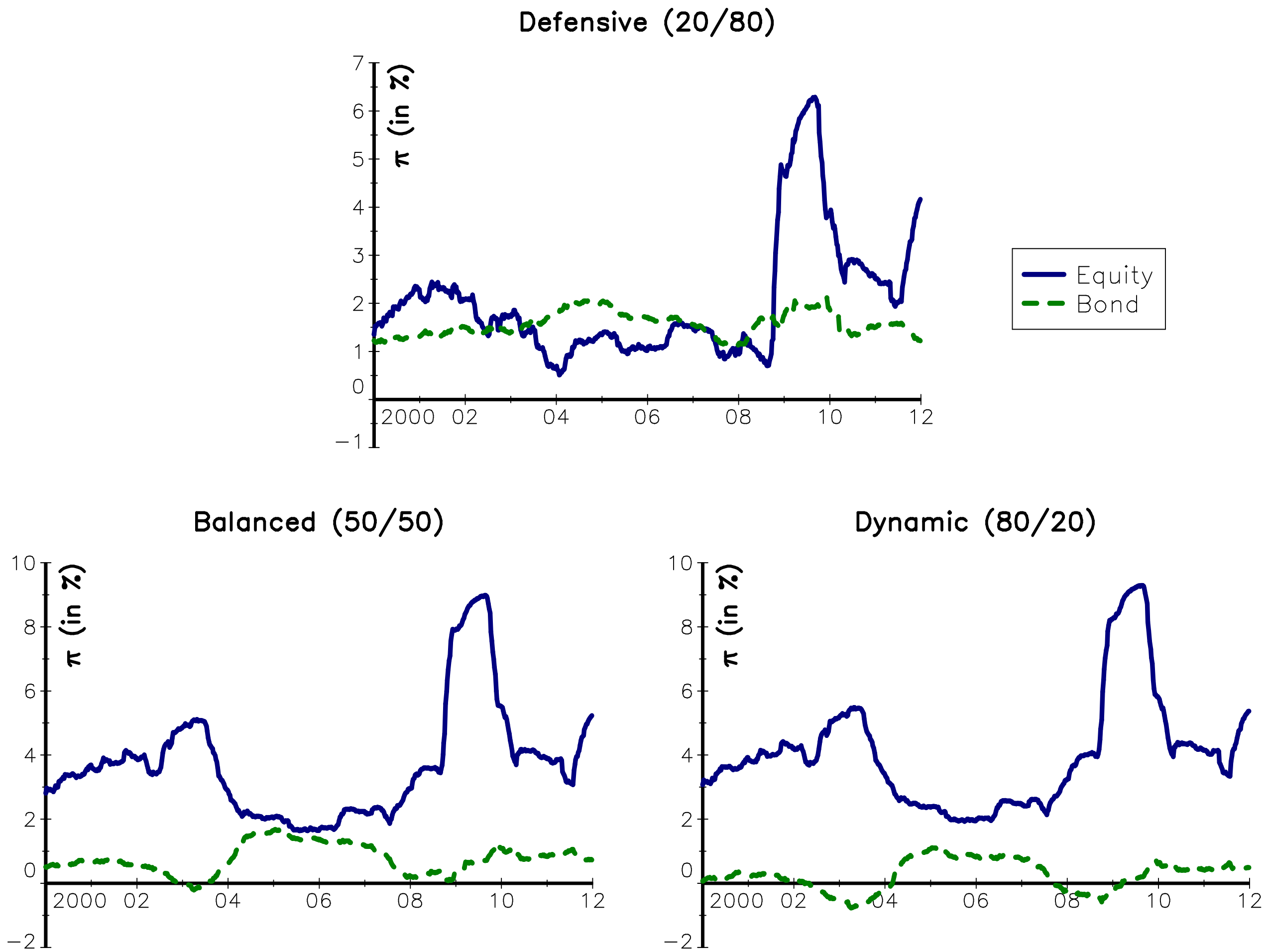


Figure 36: Equity and bond implied risk premia for diversified funds

Performance assets versus hedging assets

- We recall that:

$$\tilde{\pi} = \mathrm{SR}\left(x \mid r\right) \frac{\partial\, \sigma\left(x\right)}{\partial\, x}$$

where $\sigma\left(x\right)$ is the volatility of portfolio x

- We have:

$$\begin{aligned} \frac{\partial\, \sigma\left(x\right)}{\partial\, x_i} &= \frac{\left(\Sigma x\right)_i}{\sigma\left(x\right)} \\ &= \frac{\left(x_i \sigma_i^2 + \sigma_i \sum_{j \neq i} x_j \rho_{i,j} \sigma_j\right)}{\sigma\left(x\right)} \end{aligned}$$

- We deduce that

$$\tilde{\pi}_i = \mathrm{SR}\left(x \mid r\right) \frac{\left(x_i \sigma_i^2 + \sigma_i \sum_{j \neq i} x_j \rho_{i,j} \sigma_j\right)}{\sigma\left(x\right)}$$

Performance assets versus hedging assets

In the two-asset case, we obtain:

$$\tilde{\pi}_1 = c(x)\left(\underbrace{x_1\sigma_1^2}_{\text{variance}} + \underbrace{\rho\sigma_1\sigma_2(1-x_1)}_{\text{covariance}}\right)$$

and:

$$\tilde{\pi}_2 = c(x)\left(\underbrace{x_2\sigma_2^2}_{\text{variance}} + \underbrace{\rho\sigma_1\sigma_2(1-x_2)}_{\text{covariance}}\right)$$

where $c(x)$ is equal to $\mathrm{SR}(x \mid r)/\sigma(x)$ and ρ is the cross-correlation between the two asset returns

Performance assets versus hedging assets

In the two-asset case, the implied risk premium becomes:

$$\tilde{\pi}_i = \frac{\mathrm{SR}\left(x \mid r\right)}{\sigma\left(x\right)} \left(\underbrace{x_i \cdot \sigma_i^2}_{\text{variance}} + \underbrace{\left(1 - x_i\right) \cdot \rho \sigma_i \sigma_j}_{\text{covariance}} \right)$$

There are two components in the risk premium:

- a variance risk component, which is an increasing function of the volatility and the weight of the asset
- a (positive or negative) covariance risk component, which depends on the correlation between asset returns

Performance asset versus hedging asset

- When $\tilde{\pi}_i > 0$, the asset i is a performance asset for Portfolio x
- When $\tilde{\pi}_i < 0$, the asset i is a hedging asset for Portfolio x

Performance assets versus hedging assets

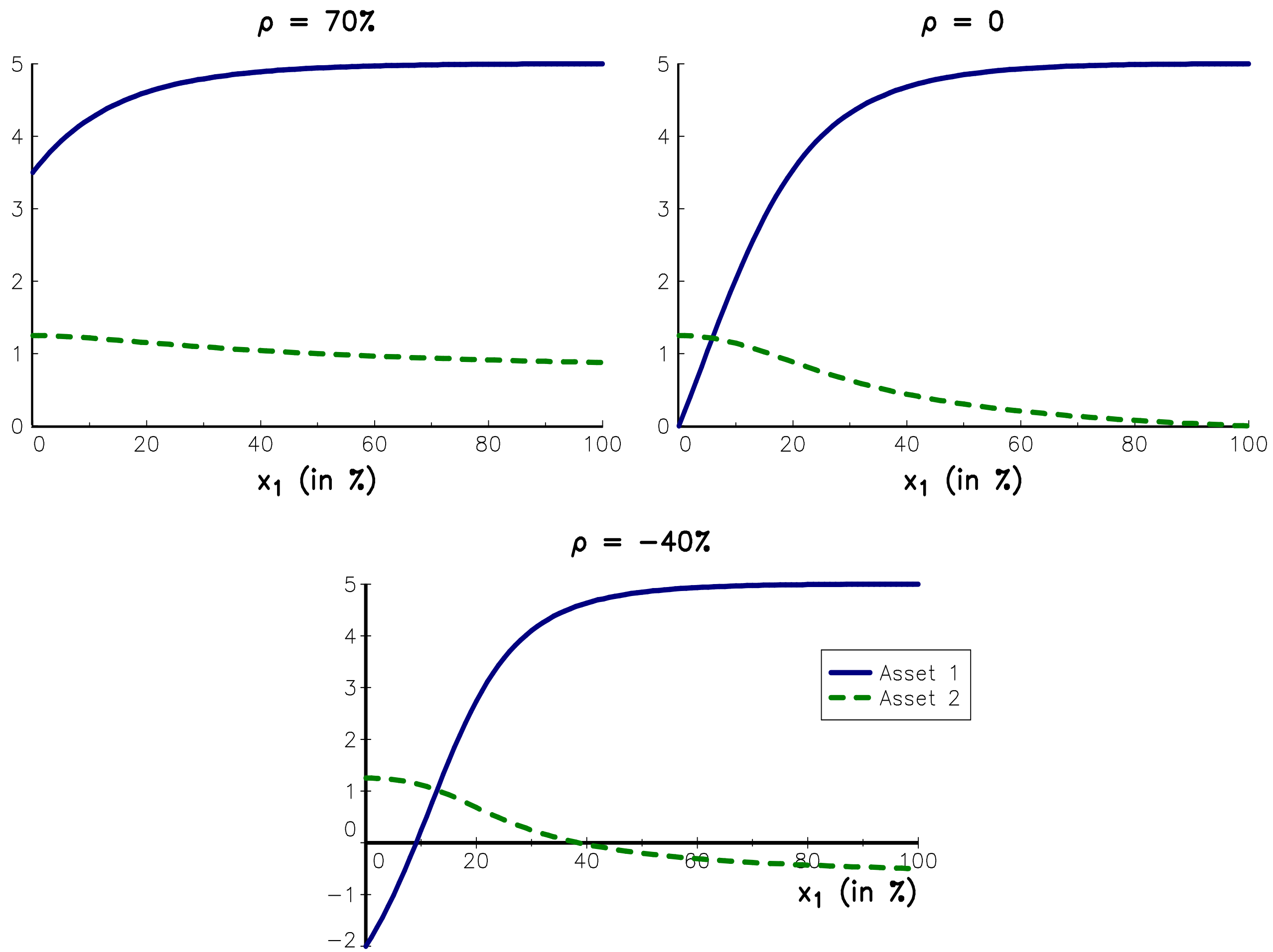


Figure 37: Impact of the correlation on the expected risk premium ($\sigma_1 = 20\%$, $\sigma_2 = 5\%$ and $\mathrm{SR}(x) = 0.25$)

Are bonds performance or hedging assets?

- Stocks are always considered as performance assets, while bonds may be performance or hedging assets, depending on the region and the period[6]
- 1990-2008: (Sovereign) bonds were perceived as performance assets
- The 2008 GFC has strengthened the fly-to-quality characteristic of bonds
- 2013-2017: Bonds are now more and more perceived as hedging assets[7]

Diversified stock-bond portfolios $\Rightarrow$ Deleveraged equity portfolios

[6] For instance bonds were hedging assets in 2008 and performance assets in 2011

[7] This is particular true in the US and Europe, where the implied risk premium is negative. In Japan, the implied risk premium continue to be positive

Diversified versus risk parity funds

Table 44: Statistics of diversified and risk parity portfolios (2000-2012)

Portfolio	$\hat{\mu}_{1Y}$	$\hat{\sigma}_{1Y}$	SR	$\mathcal{MDD}$	γ_1	γ_2
Defensive	5.41	6.89	0.42	−17.23	0.19	2.67
Balanced	3.68	9.64	0.12	−33.18	−0.13	3.87
Dynamic	1.70	14.48	−0.06	−48.90	−0.18	5.96
Risk parity	5.12	7.29	0.36	−21.22	0.08	2.65
Static	4.71	7.64	0.29	−23.96	0.03	2.59
Leveraged RP	6.67	9.26	0.45	−23.74	0.01	0.78

- The 60/40 constant mix strategy is not the right benchmark
- Results depend on the investment universe (number/granularity of asset classes)
- What is the impact of rising interest rates?

Optimality of the RB portfolio

We consider the utility function:

$$\mathcal{U}(x) = (\mu(x) - r) - \phi \mathcal{R}(x)$$

Portfolio x is optimal if the vector of expected risk premia satisfies this relationship:

$$\tilde{\pi} = \phi \frac{\partial \mathcal{R}(x)}{\partial x}$$

If the RB portfolio is optimal, we deduce that the (excess) performance contribution $\mathcal{PC}_i$ of asset i is proportional to its risk budget:

$$\begin{array}{rcl} \mathcal{PC}_i & = & x_i \tilde{\pi}_i \\ & = & \phi \cdot \mathcal{RC}_i \\ & \propto & b_i \end{array}$$

Optimality of the RB portfolio

In the Black-Litterman approach of risk premia, we have:

$$\tilde{\pi}_i = \tilde{\mu}_i - r = \mathrm{SR}\left(x \mid r\right) \frac{\left(\Sigma x\right)_i}{\sqrt{x^\top \Sigma x}}$$

This implies that the (excess) performance contribution is equal to:

$$\begin{aligned}\mathcal{PC}_i &= \mathrm{SR}\left(x \mid r\right) \frac{x_i \cdot \left(\Sigma x\right)_i}{\sqrt{x^\top \Sigma x}} \\ &= \mathrm{SR}\left(x \mid r\right) \cdot \mathcal{RC}_i\end{aligned}$$

where $\mathrm{SR}\left(x \mid r\right)$ is the expected Sharpe ratio of the RB portfolio

Optimality of the RB portfolio

Remark

From an ex-ante point of view, performance budgeting and risk budgeting are equivalent

Optimality of the RB portfolio

Example 18

We consider a universe of four assets. The volatilities are respectively 10%, 20%, 30% and 40%. The correlation of asset returns is given by the following matrix:

$$\rho = \begin{pmatrix} 1.00 & & & \\ 0.80 & 1.00 & & \\ 0.20 & 0.20 & 1.00 & \\ 0.20 & 0.20 & 0.50 & 1.00 \end{pmatrix}$$

The risk-free rate is equal to zero

Optimality of the RB portfolio

Table 45: Implied risk premia when $b = (20\%, 25\%, 40\%, 15\%)$ (Example 18)

Asset	x_i	$\mathcal{MR}_i$	$\tilde{\mu}_i$	$\mathcal{PC}_i$	$\mathcal{PC}_i^{\star}$
1	40.91	7.10	3.55	1.45	20.00
2	25.12	14.46	7.23	1.82	25.00
3	25.26	23.01	11.50	2.91	40.00
4	8.71	25.04	12.52	1.09	15.00
Expected return				7.27	

Table 46: Implied risk premia when $b = (10\%, 10\%, 10\%, 70\%)$ (Example 18)

Asset	x_i	$\mathcal{MR}_i$	$\tilde{\mu}_i$	$\mathcal{PC}_i$	$\mathcal{PC}_i^{\star}$
1	35.88	5.27	2.63	0.94	10.00
2	17.94	10.53	5.27	0.94	10.00
3	10.18	18.56	9.28	0.94	10.00
4	35.99	36.75	18.37	6.61	70.00
Expected return				9.45	

Main result

There is no neutral allocation. Every portfolio corresponds to an active bet.

Variation on the ERC portfolio

Question 1

We note Σ the covariance matrix of asset returns.

Variation on the ERC portfolio

Question 1.a

What is the risk contribution $\mathcal{RC}_i$ of asset i with respect to portfolio x?

Variation on the ERC portfolio

Let $\mathcal{R}(x)$ be a risk measure of the portfolio x. If this risk measure satisfies the Euler principle, we have (TR-RPB, page 78):

$$\mathcal{R}(x) = \sum_{i=1}^{n} x_i \frac{\partial \, \mathcal{R}(x)}{\partial \, x_i}$$

We can then decompose the risk measure as a sum of asset contributions. This is why we define the risk contribution $\mathcal{RC}_i$ of asset i as the product of the weight by the marginal risk:

$$\mathcal{RC}_i = x_i \frac{\partial \, \mathcal{R}(x)}{\partial \, x_i}$$

When the risk measure is the volatility $\sigma(x)$, it follows that:

$$\begin{aligned} \mathcal{RC}_i &= x_i \frac{(\Sigma x)_i}{\sqrt{x^\top \Sigma x}} \\ &= \frac{x_i \left(\sum_{k=1}^{n} \rho_{i,k} \sigma_i \sigma_k x_k \right)}{\sigma(x)} \end{aligned}$$

Variation on the ERC portfolio

Question 1.b

Define the ERC portfolio.

Variation on the ERC portfolio

The ERC portfolio corresponds to the risk budgeting portfolio when the risk measure is the return volatility $\sigma(x)$ and when the risk budgets are the same for all the assets (TR-RPB, page 119). It means that $\mathcal{RC}_i = \mathcal{RC}_j$, that is:

$$x_i \frac{\partial\, \sigma(x)}{\partial\, x_i} = x_j \frac{\partial\, \sigma(x)}{\partial\, x_j}$$

Variation on the ERC portfolio

Question 1.c

Calculate the variance of the risk contributions. Define an optimization program to compute the ERC portfolio. Find an equivalent maximization program based on the $\mathcal{L}^2$ norm.

Variation on the ERC portfolio

We have:

$$\begin{aligned}\overline{\mathcal{RC}} &= \frac{1}{n}\sum_{i=1}^{n}\mathcal{RC}_i \\ &= \frac{1}{n}\sigma(x)\end{aligned}$$

It follows that:

$$\begin{aligned}\operatorname{var}(\mathcal{RC}) &= \frac{1}{n}\sum_{i=1}^{n}\left(\mathcal{RC}_i - \overline{\mathcal{RC}}\right)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\mathcal{RC}_i - \frac{1}{n}\sigma(x)\right)^2 \\ &= \frac{1}{n^2\sigma(x)}\sum_{i=1}^{n}\left(n x_i (\Sigma x)_i - \sigma^2(x)\right)^2\end{aligned}$$

Variation on the ERC portfolio

To compute the ERC portfolio, we may consider the following optimization program:

$$x^{\star} = \arg\min \sum_{i=1}^{n} \left(n x_i \left(\Sigma x\right)_i - \sigma^2 \left(x\right)\right)^2$$

Because we know that the ERC portfolio always exists (TR-RPB, page 108), the objective function at the optimum $x^{\star}$ is necessarily equal to 0. Another equivalent optimization program is to consider the L^2 norm. In this case, we have (TR-RPB, page 102):

$$x^{\star} = \arg\min \sum_{i=1}^{n} \sum_{j=1}^{n} \left(x_i \cdot \left(\Sigma x\right)_i - x_j \cdot \left(\Sigma x\right)_j\right)^2$$

Variation on the ERC portfolio

Question 1.d

Let $\beta_i(x)$ be the beta of asset i with respect to portfolio x. Show that we have the following relationship in the ERC portfolio:

$$x_i \beta_i(x) = x_j \beta_j(x)$$

Propose a numerical algorithm to find the ERC portfolio.

Variation on the ERC portfolio

We have:

$$\begin{aligned}\beta_i(x) &= \frac{(\Sigma x)_i}{x^\top \Sigma x} \\ &= \frac{\mathcal{MR}_i}{\sigma(x)}\end{aligned}$$

We deduce that:

$$\begin{aligned}\mathcal{RC}_i &= x_i \cdot \mathcal{MR}_i \\ &= x_i \beta_i(x) \sigma(x)\end{aligned}$$

The relationship $\mathcal{RC}_i = \mathcal{RC}_j$ becomes:

$$x_i \beta_i(x) = x_j \beta_j(x)$$

It means that the weight is inversely proportional to the beta:

$$x_i \propto \frac{1}{\beta_i(x)}$$

Variation on the ERC portfolio

We can use the Jacobi power algorithm (TR-RPB, page 308). Let $x^{(k)}$ be the portfolio at iteration k. We define the portfolio $x^{(k+1)}$ as follows:

$$x^{(k+1)} = \frac{\beta_i^{-1}\left(x^{(k)}\right)}{\sum_{j=1}^{n} \beta_j^{-1}\left(x^{(k)}\right)}$$

Starting from an initial portfolio $x^{(0)}$, the limit portfolio is the ERC portfolio if the algorithm converges:

$$\lim_{k \rightarrow \infty} x^{(k)} = x_{\mathrm{erc}}$$

Variation on the ERC portfolio

Question 1.e

We suppose that the volatilities are 15%, 20% and 25% and that the correlation matrix is:

$$\rho = \begin{pmatrix} 100\% & & \\ 50\% & 100\% & \\ 40\% & 30\% & 100\% \end{pmatrix}$$

Compute the ERC portfolio using the beta algorithm.

Variation on the ERC portfolio

Starting from the EW portfolio, we obtain for the first five iterations:

k	0	1	2	3	4	5
$x_1^{(k)}$ (in %)	33.3333	43.1487	40.4122	41.2314	40.9771	41.0617
$x_2^{(k)}$ (in %)	33.3333	32.3615	31.9164	32.3529	32.1104	32.2274
$x_3^{(k)}$ (in %)	33.3333	24.4898	27.6714	26.4157	26.9125	26.7109
$\beta_1\left(x^{(k)}\right)$	0.7326	0.8341	0.8046	0.8147	0.8113	0.8126
$\beta_2\left(x^{(k)}\right)$	0.9767	1.0561	1.0255	1.0397	1.0337	1.0363
$\beta_3\left(x^{(k)}\right)$	1.2907	1.2181	1.2559	1.2405	1.2472	1.2444

Variation on the ERC portfolio

The next iterations give the following results:

k	6	7	8	9	10	11
$x_1^{(k)}$ (in %)	41.0321	41.0430	41.0388	41.0405	41.0398	41.0401
$x_2^{(k)}$ (in %)	32.1746	32.1977	32.1878	32.1920	32.1902	32.1909
$x_3^{(k)}$ (in %)	26.7933	26.7593	26.7734	26.7676	26.7700	26.7690
$\beta_1\left(x^{(k)}\right)$	0.8121	0.8123	0.8122	0.8122	0.8122	0.8122
$\beta_2\left(x^{(k)}\right)$	1.0352	1.0356	1.0354	1.0355	1.0355	1.0355
$\beta_3\left(x^{(k)}\right)$	1.2456	1.2451	1.2453	1.2452	1.2452	1.2452

Variation on the ERC portfolio

Finally, the algorithm converges after 14 iterations with the following stopping criteria:

$$\sup_i \left| x_i^{(k+1)} - x_i^{(k)} \right| \leq 10^{-6}$$

and we obtain the following results:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	41.04%	12.12%	4.97%	33.33%
2	32.19%	15.45%	4.97%	33.33%
3	26.77%	18.58%	4.97%	33.33%

Variation on the ERC portfolio

Question 2

We now suppose that the return of asset i satisfies the CAPM model:

$$R_i = \beta_i R_m + \varepsilon_i$$

where R_m is the return of the market portfolio and ε_i is the idiosyncratic risk. We note $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$. We assume that $R_m \perp \varepsilon$, $\operatorname{var}(R_m) = \sigma_m^2$ and $\operatorname{cov}(\varepsilon) = D = \operatorname{diag}(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2)$.

Variation on the ERC portfolio

Question 2.a

Calculate the risk contribution $\mathcal{RC}_i$.

Variation on the ERC portfolio

We have:

$$\Sigma = \beta\beta^\top \sigma_m^2 + \operatorname{diag}\left(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2\right)$$

We deduce that:

$$\begin{aligned}\mathcal{RC}_i &= \frac{x_i \left(\sum_{k=1}^n \beta_i \beta_k \sigma_m^2 x_k + \tilde{\sigma}_i^2 x_i\right)}{\tilde{\sigma}\left(x\right)} \\ &= \frac{x_i \beta_i B + x_i^2 \tilde{\sigma}_i^2}{\sigma\left(x\right)}\end{aligned}$$

with:

$$B = \sum_{k=1}^n x_k \beta_k \sigma_m^2$$

Variation on the ERC portfolio

Question 2.b

We assume that $\beta_i = \beta_j$. Show that the ERC weight x_i is a decreasing function of the idiosyncratic volatility $\tilde{\sigma}_i$.

Variation on the ERC portfolio

Using Equation 2.a, we deduce that the ERC portfolio satisfies:

$$x_i \beta_i B + x_i^2 \tilde{\sigma}_i^2 = x_j \beta_j B + x_j^2 \tilde{\sigma}_j^2$$

or:

$$\left(x_i \beta_i - x_j \beta_j\right) B = \left(x_j^2 \tilde{\sigma}_j^2 - x_i^2 \tilde{\sigma}_i^2\right)$$

Variation on the ERC portfolio

If $\beta_i = \beta_j = \beta$, we have:

$$(x_i - x_j)\,\beta B = \left(x_j^2 \tilde{\sigma}_j^2 - x_i^2 \tilde{\sigma}_i^2\right)$$

Because $\beta > 0$, we deduce that:

$$\begin{array}{rcl} x_i > x_j & \Leftrightarrow & x_j^2 \tilde{\sigma}_j^2 - x_i^2 \tilde{\sigma}_i^2 > 0 \\ & \Leftrightarrow & x_j \tilde{\sigma}_j > x_i \tilde{\sigma}_i \\ & \Leftrightarrow & \tilde{\sigma}_i < \tilde{\sigma}_j \end{array}$$

We conclude that the weight x_i is a decreasing function of the specific volatility $\tilde{\sigma}_i$.

Variation on the ERC portfolio

Question 2.c

We assume that $\tilde{\sigma}_i = \tilde{\sigma}_j$. Show that the ERC weight x_i is a decreasing function of the sensitivity β_i to the common factor.

Variation on the ERC portfolio

If $\tilde{\sigma}_i = \tilde{\sigma}_j = \tilde{\sigma}$, we have:

$$\left(x_i \beta_i - x_j \beta_j\right) B = \left(x_j^2 - x_i^2\right) \tilde{\sigma}^2$$

We deduce that:

$$\begin{aligned} x_i > x_j & \Leftrightarrow \left(x_i \beta_i - x_j \beta_j\right) B < 0 \\ & \Leftrightarrow x_i \beta_i < x_j \beta_j \\ & \Leftrightarrow \beta_i < \beta_j \end{aligned}$$

We conclude that the weight x_i is a decreasing function of the sensitivity β_i.

Variation on the ERC portfolio

Question 2.d

We consider the numerical application: $\beta_1 = 1$, $\beta_2 = 0.9$, $\beta_3 = 0.8$, $\beta_4 = 0.7$, $\tilde{\sigma}_1 = 5\%$, $\tilde{\sigma}_2 = 5\%$, $\tilde{\sigma}_3 = 10\%$, $\tilde{\sigma}_4 = 10\%$, and $\sigma_m = 20\%$. Find the ERC portfolio.

Variation on the ERC portfolio

We obtain the following results:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^{\star}$
1	21.92%	19.73%	4.32%	25.00%
2	24.26%	17.83%	4.32%	25.00%
3	25.43%	17.00%	4.32%	25.00%
4	28.39%	15.23%	4.32%	25.00%

Weight concentration of a portfolio

Question 1

We consider the Lorenz curve defined by:

$$\begin{array}{rcl} [0,1] & \longrightarrow & [0,1] \\ x & \longmapsto & \mathbb{L}(x) \end{array}$$

We assume that $\mathbb{L}$ is an increasing function and $\mathbb{L}(x) > x$.

Weight concentration of a portfolio

Question 1.a

Represent graphically the function $\mathbb{L}$ and define the Gini coefficient $\mathcal{G}$ associated with $\mathbb{L}$.

Weight concentration of a portfolio

We have represented the function $y = \mathcal{L}(x)$ in Figure 38. It verifies $\mathcal{L}(x) \geq x$ and $\mathcal{L}(x) \leq 1$. The Gini coefficient is defined as follows (TR-RPB, page 127):

$$
\begin{aligned}
G &= \frac{A}{A+B} \\
&= \left(\int_0^1 \mathcal{L}(x) \,\mathrm{d}x - \frac{1}{2} \right) \Big/ \frac{1}{2} \\
&= 2 \int_0^1 \mathcal{L}(x) \,\mathrm{d}x - 1
\end{aligned}
$$

Weight concentration of a portfolio

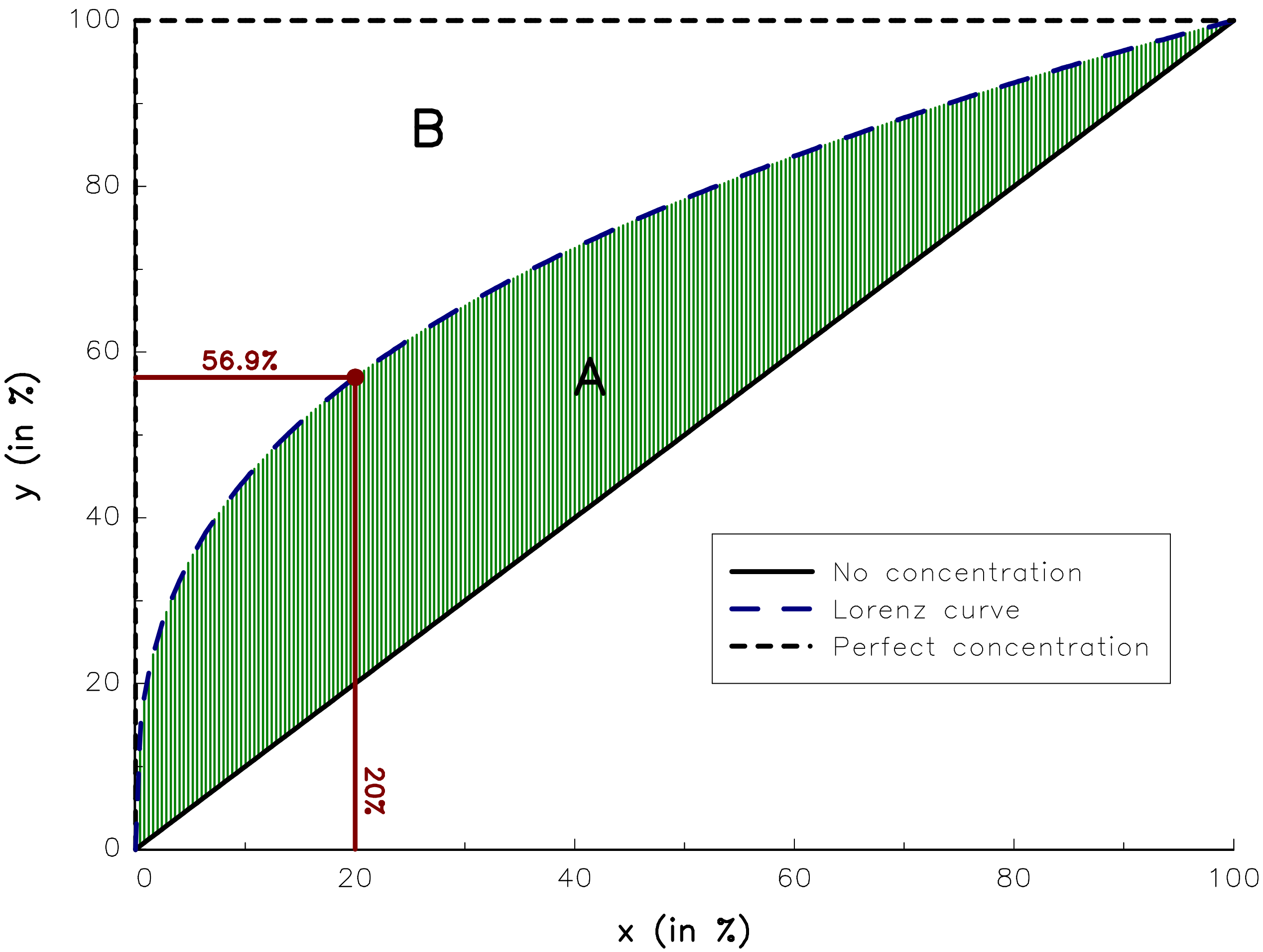


Figure 38: Lorenz curve

Weight concentration of a portfolio

Question 1.b

We set $\mathbb{L}_{\alpha}(x) = x^{\alpha}$ with $\alpha \geq 0$. Is the function $\mathbb{L}_{\alpha}$ a Lorenz curve? Calculate the Gini coefficient $\mathcal{G}(\alpha)$ with respect to α. Deduce $\mathcal{G}(0)$, $\mathcal{G}\left(\frac{1}{2}\right)$ and $\mathcal{G}(1)$.

Weight concentration of a portfolio

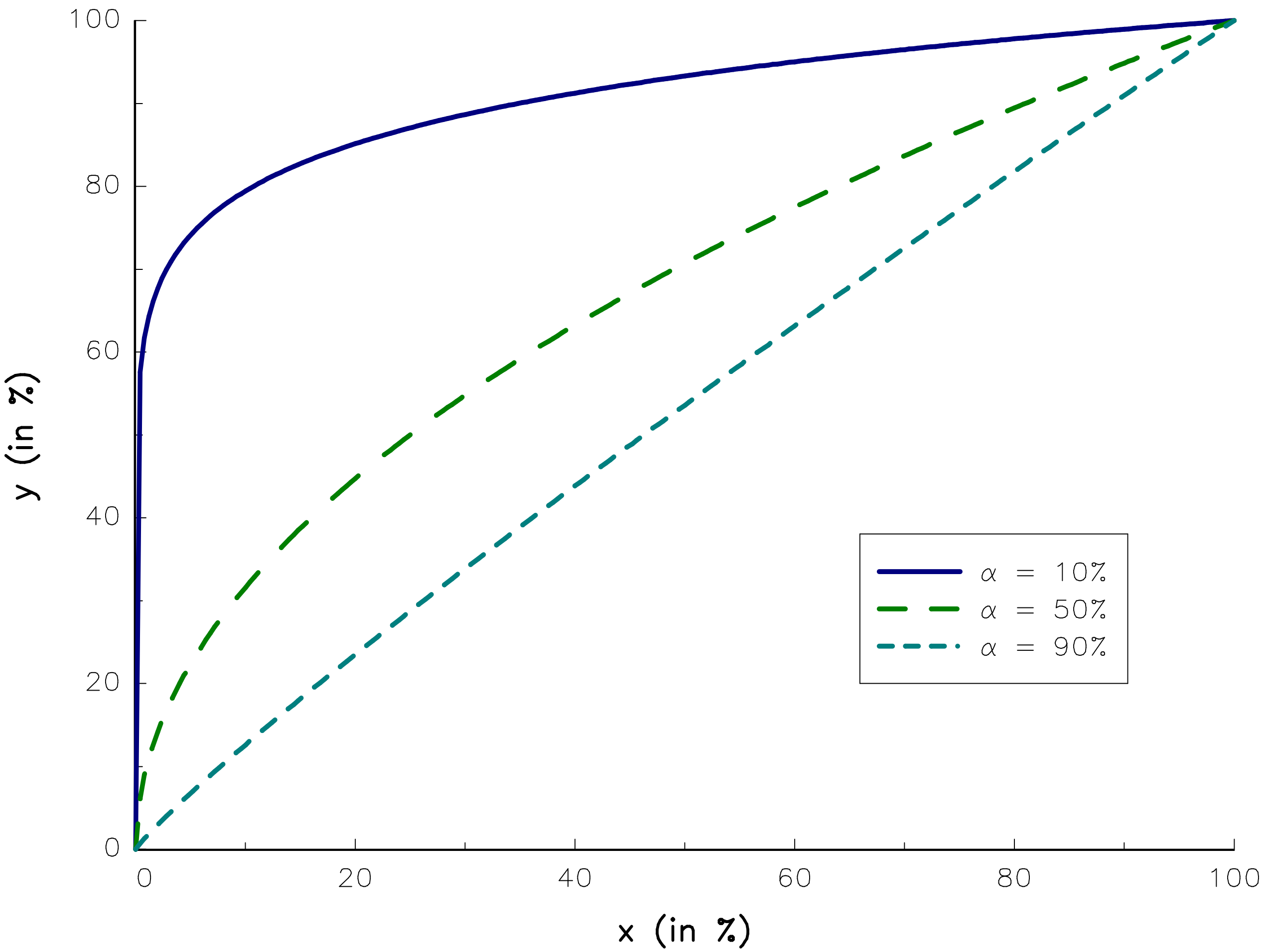


Figure 39: Function $y = x^{\alpha}$

Weight concentration of a portfolio

If $\alpha \geq 0$, the function $\mathcal{L}_\alpha(x) = x^\alpha$ is increasing. We have $\mathcal{L}_\alpha(1) = 1$, $\mathcal{L}_\alpha(x) \leq 1$ and $\mathcal{L}_\alpha(x) \geq x$. We deduce that $\mathcal{L}_\alpha$ is a Lorenz curve. For the Gini index, we have:

$$
\begin{aligned}
\mathcal{G}(\alpha) &= 2\int_0^1 x^\alpha \,\mathrm{d}x - 1 \\
&= 2\left[\frac{x^{\alpha+1}}{\alpha+1}\right]_0^1 - 1 \\
&= \frac{1-\alpha}{1+\alpha}
\end{aligned}
$$

We deduce that $\mathcal{G}(0) = 1$, $\mathcal{G}\left(\frac{1}{2}\right) = 1/3$ et $\mathcal{G}(1) = 0$. $\alpha = 0$ corresponds to the perfect concentration whereas $\alpha = 1$ corresponds to the perfect equality.

Weight concentration of a portfolio

Question 2

Let w be a portfolio of n assets. We suppose that the weights are sorted in a descending order: $w_1 \geq w_2 \geq \ldots \geq w_n$.

Weight concentration of a portfolio

Question 2.a

We define $\mathbb{L}_w(x)$ as follows:

$$\mathbb{L}_w(x) = \sum_{j=1}^{i} w_j \quad \text{if} \quad \frac{i}{n} \leq x < \frac{i+1}{n}$$

with $\mathbb{L}_w(0) = 0$. Is the function $\mathbb{L}_w$ a Lorenz curve ? Calculate the Gini coefficient with respect to the weights w_i. In which cases does $\mathcal{G}$ take the values 0 and 1?

Weight concentration of a portfolio

We have $\mathcal{L}_w(0) = 0$ and $\mathcal{L}_w(1) = \sum_{j=1}^n w_j = 1$. If $x_2 \geq x_1$, we have $\mathcal{L}_w(x_2) \geq \mathcal{L}_w(x_2)$. $\mathcal{L}_w$ is then a Lorenz curve. The Gini coefficient is equal to:

$$\begin{aligned}\mathcal{G} &= 2\int_0^1 \mathcal{L}(x)\,\mathrm{d}x - 1 \\ &= \frac{2}{n}\sum_{i=1}^n \sum_{j=1}^i w_j - 1\end{aligned}$$

If $w_j = n^{-1}$, we have:

$$\begin{aligned}\lim_{n\to\infty}\mathcal{G} &= \lim_{n\to\infty}\frac{2}{n}\sum_{i=1}^n \frac{i}{n} - 1 \\ &= \lim_{n\to\infty}\frac{2}{n}\cdot\frac{n(n+1)}{2n} - 1 \\ &= \lim_{n\to\infty}\frac{1}{n} = 0\end{aligned}$$

Weight concentration of a portfolio

If $w_1 = 1$, we have:

$$\begin{aligned} \lim_{n\to\infty} \mathcal{G} &= \lim_{n\to\infty} 1 - \frac{1}{n} \\ &= 1 \end{aligned}$$

We note that the perfect equality does not correspond to the case $\mathcal{G} = 0$ except in the asymptotic case. This is why we may slightly modify the definition of $\mathcal{L}_w(x)$:

$$\mathcal{L}_w(x) = \begin{cases} \sum_{j=1}^{i} w_j & \text{if} \quad x = n^{-1} i \\ \sum_{j=1}^{i} w_j + w_{i+1}(nx - i) & \text{if} \quad n^{-1} i < x < n^{-1}(i+1) \end{cases}$$

While the previous definition corresponds to a constant piecewise function, this one defines an affine piecewise function. In this case, the computation of the Gini index is done using a trapezoidal integration:

$$\mathcal{G} = \frac{2}{n}\left(\sum_{i=1}^{n-1}\sum_{j=1}^{i} w_j + \frac{1}{2}\right) - 1$$

Weight concentration of a portfolio

Question 2.b

The definition of the Herfindahl index is:

$$\mathcal{H} = \sum_{i=1}^{n} w_i^2$$

In which cases does $\mathcal{H}$ take the value 1? Show that $\mathcal{H}$ reaches its maximum when $w_i = n^{-1}$. What is the interpretation of this result?

Weight concentration of a portfolio

The Herfindahl index is equal to 1 if the portfolio is concentrated in only one asset. We seek to minimize $\mathcal{H} = \sum_{i=1}^n w_i^2$ under the constraint $\sum_{i=1}^n w_i = 1$. The Lagrange function is then:

$$f\left(w_1, \ldots, w_n; \lambda\right) = \sum_{i=1}^n w_i^2 - \lambda \left(\sum_{i=1}^n w_i - 1 \right)$$

The first-order conditions are $2w_i - \lambda = 0$. We deduce that $w_i = w_j$. $\mathcal{H}$ reaches its minimum when $w_i = n^{-1}$. It corresponds to the equally weighted portfolio. In this case, we have:

$$\mathcal{H} = \frac{1}{n}$$

Weight concentration of a portfolio

Question 2.c

We set $\mathcal{N} = \mathcal{H}^{-1}$. What does the statistic $\mathcal{N}$ mean?

Weight concentration of a portfolio

The statistic $\mathcal{N}$ is the degree of freedom or the equivalent number of equally weighted assets. For instance, if $\mathcal{H} = 0.5$, then $\mathcal{N} = 2$. It is a portfolio equivalent to two equally weighted assets.

Weight concentration of a portfolio

Question 3

We consider an investment universe of five assets. We assume that their asset returns are not correlated. The volatilities are given in the table below:

σ_i	2%	5%	10%	20%	30%
$w_i^{(1)}$		10%	20%	30%	40%
$w_i^{(2)}$	40%	20%		30%	10%
$w_i^{(3)}$	20%	15%	25%	35%	5%

Weight concentration of a portfolio

Question 3.a

Find the minimum variance portfolio $w^{(4)}$.

Weight concentration of a portfolio

The minimum variance portfolio is equal to:

$$w^{(4)} = \begin{pmatrix} 82.342\% \\ 13.175\% \\ 3.294\% \\ 0.823\% \\ 0.366\% \end{pmatrix}$$

Weight concentration of a portfolio

Question 3.b

Calculate the Gini and Herfindahl indices and the statistic $\mathcal{N}$ for the four portfolios $w^{(1)}$, $w^{(2)}$, $w^{(3)}$ and $w^{(4)}$.

Weight concentration of a portfolio

For each portfolio, we sort the weights in descending order. For the portfolio $w^{(1)}$, we have $w_1^{(1)} = 40\%$, $w_2^{(1)} = 30\%$, $w_3^{(1)} = 20\%$, $w_4^{(1)} = 10\%$ and $w_5^{(1)} = 0\%$. It follows that:

$$\begin{aligned}
\mathcal{H}\left(w^{(1)}\right) &= \sum_{i=1}^{5}\left(w_i^{(1)}\right)^2 \\
&= 0.10^2 + 0.20^2 + 0.30^2 + 0.40^2 \\
&= 0.30
\end{aligned}$$

We also have:

$$\begin{aligned}
\mathcal{G}\left(w^{(1)}\right) &= \frac{2}{5}\left(\sum_{i=1}^{4}\sum_{j=1}^{i}\tilde{w}_j^{(1)} + \frac{1}{2}\right) - 1 \\
&= \frac{2}{5}\left(0.40 + 0.70 + 0.90 + 1.00 + \frac{1}{2}\right) - 1 \\
&= 0.40
\end{aligned}$$

Weight concentration of a portfolio

For the portfolios $w^{(2)}$, $w^{(3)}$ and $w^{(4)}$, we obtain $\mathcal{H}\left(w^{(2)}\right) = 0.30$, $\mathcal{H}\left(w^{(3)}\right) = 0.25$, $\mathcal{H}\left(w^{(4)}\right) = 0.70$, $\mathcal{G}\left(w^{(2)}\right) = 0.40$, $\mathcal{G}\left(w^{(3)}\right) = 0.28$ and $\mathcal{G}\left(w^{(4)}\right) = 0.71$. We have $\mathcal{N}\left(w^{(2)}\right) = \mathcal{N}\left(w^{(1)}\right) = 3.33$, $\mathcal{N}\left(w^{(3)}\right) = 4.00$ and $\mathcal{N}\left(w^{(4)}\right) = 1.44$.

Weight concentration of a portfolio

Question 3.c

Comment on these results. What differences do you make between portfolio concentration and portfolio diversification?

Weight concentration of a portfolio

All the statistics show that the least concentrated portfolio is $w^{(3)}$. The most concentrated portfolio is paradoxically the minimum variance portfolio $w^{(4)}$. We generally assimilate variance optimization to diversification optimization. We show in this example that diversifying in the Markowitz sense does not permit to minimize the concentration.

The optimization problem of the ERC portfolio

Question 1

We consider four assets. Their volatilities are equal to 10%, 15%, 20% and 25% whereas the correlation matrix of asset returns is:

$$\rho = \begin{pmatrix} 100\% & & & \\ 60\% & 100\% & & \\ 40\% & 40\% & 100\% & \\ 30\% & 30\% & 20\% & 100\% \end{pmatrix}$$

The optimization problem of the ERC portfolio

Question 1.a

Find the long-only minimum variance, ERC and equally weighted portfolios.

The optimization problem of the ERC portfolio

The weights of the three portfolios are:

Asset	MV	ERC	EW
1	87.51%	37.01%	25.00%
2	4.05%	24.68%	25.00%
3	4.81%	20.65%	25.00%
4	3.64%	17.66%	25.00%

The optimization problem of the ERC portfolio

Question 1.b

We consider the following portfolio optimization problem:

$$x^{\star}(c) = \arg\min \sqrt{x^{\top}\Sigma x} \quad \text{u.c.} \quad \begin{cases} \sum_{i=1}^{n} \ln x_i \geq c \\ \mathbf{1}_n^{\top} x = 1 \\ x \geq \mathbf{0}_n \end{cases} \tag{1}$$

with Σ the covariance matrix of asset returns. We note λ_c and λ_0 the Lagrange coefficients associated with the constraints $\sum_{i=1}^{n} \ln x_i \geq c$ and $\mathbf{1}_n^{\top} x = 1$. Write the Lagrange function of the optimization problem. Deduce then an equivalent optimization problem that is easier to solve than Problem (1).

The optimization problem of the ERC portfolio

The Lagrange function is:

$$\begin{aligned}\mathcal{L}(x;\lambda,\lambda_0,\lambda_c) &= \sqrt{x^\top \Sigma x} - \lambda^\top x - \lambda_0 \left(\mathbf{1}_n^\top x - 1\right) - \lambda_c \left(\sum_{i=1}^n \ln x_i - c\right) \\ &= \left(\sqrt{x^\top \Sigma x} - \lambda_c \sum_{i=1}^n \ln x_i\right) - \lambda^\top x - \lambda_0 \left(\mathbf{1}_n^\top x - 1\right) + \lambda_c c\end{aligned}$$

We deduce that an equivalent optimization problem is:

$$\begin{aligned}\tilde{x}^\star(\lambda_c) &= \arg\min \sqrt{\tilde{x}^\top \Sigma \tilde{x}} - \lambda_c \sum_{i=1}^n \ln \tilde{x}_i \\ &\text{u.c.} \quad \left\{\begin{array}{l}\mathbf{1}_n^\top \tilde{x} = 1 \\ \tilde{x} \geq \mathbf{0}_n\end{array}\right.\end{aligned}$$

The optimization problem of the ERC portfolio

We notice a strong difference between the two problems because they don't use the same control variable. However, the control variable c of the first problem may be deduced from the solution of the second problem:

$$c = \sum_{i=1}^{n} \ln \tilde{x}_i^\star \left(\lambda_c \right)$$

We also know that (TR-RPB, page 131):

$$c_- \leq \sum_{i=1}^{n} \ln x_i \leq c_+$$

where $c_- = \sum_{i=1}^{n} \ln \left(x_{\mathrm{mv}} \right)_i$ and $c_+ = -n \ln n$. It follows that:

$$\left\{ \begin{array}{ll} x^\star \left(c \right) = \tilde{x}^\star \left(0 \right) & \text{if } c \leq c_- \\ x^\star \left(c \right) = \tilde{x}^\star \left(\infty \right) & \text{if } c \geq c_+ \end{array} \right.$$

If $c \in \left] c_-, c_+ \right[$, there exists a scalar $\lambda_c > 0$ such that:

$$x^\star \left(c \right) = \tilde{x}^\star \left(\lambda_c \right)$$

The optimization problem of the ERC portfolio

Question 1.c

Represent the relationship between λ_c and $\sigma\left(x^{\star}\left(c\right)\right)$, c and $\sigma\left(x^{\star}\left(c\right)\right)$ and $\mathcal{I}^{\star}\left(x^{\star}\left(c\right)\right)$ and $\sigma\left(x^{\star}\left(c\right)\right)$ where $\mathcal{I}^{\star}\left(x\right)$ is the diversity index of the weights.

The optimization problem of the ERC portfolio

For a given value $\lambda_c \in [0, +\infty[$, we solve numerically the second problem and find the optimized portfolio $\tilde{x}^\star(\lambda_c)$. Then, we calculate $c = \sum_{i=1}^n \ln \tilde{x}_i^\star(\lambda_c)$ and deduce that $x^\star(c) = \tilde{x}^\star(\lambda_c)$. We finally obtain $\sigma(x^\star(c)) = \sigma(\tilde{x}^\star(\lambda_c))$ and $\mathcal{I}^\star(x^\star(c)) = \mathcal{I}^\star(\tilde{x}^\star(\lambda_c))$. The relationships between λ_c, c, $\mathcal{I}^\star(x^\star(c))$ and $\sigma(x^\star(c))$ are reported in Figure 40.

The optimization problem of the ERC portfolio

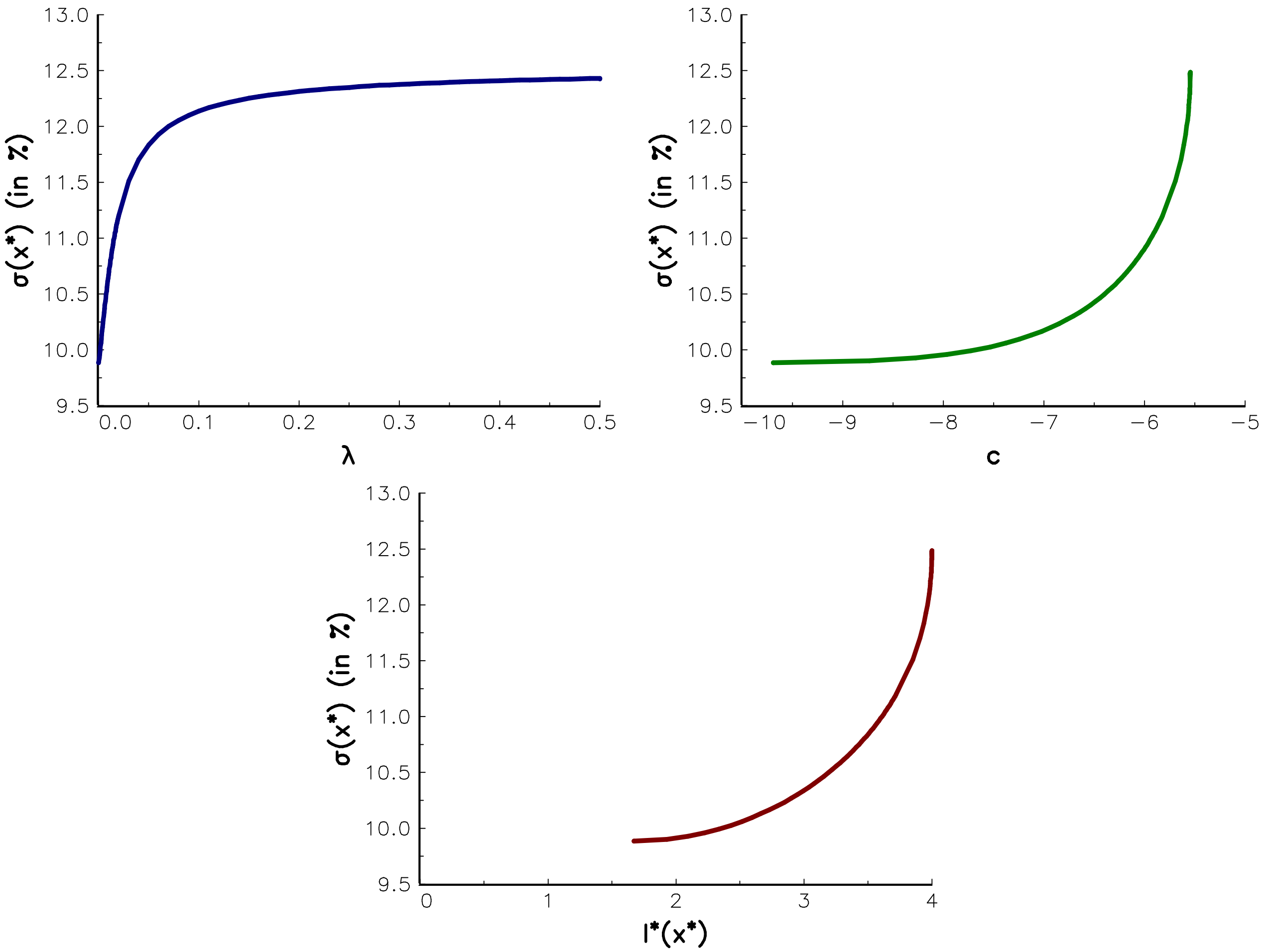


Figure 40: Relationship between λ_c, c, $\mathcal{I}^{\star}\left(x^{\star}\left(c\right)\right)$ and $\sigma\left(x^{\star}\left(c\right)\right)$

The optimization problem of the ERC portfolio

Question 1.d

Represent the relationship between λ_c and $\mathcal{I}^{\star}(\mathcal{RC})$, c and $\mathcal{I}^{\star}(\mathcal{RC})$ and $\mathcal{I}^{\star}(x^{\star}(c))$ and $\mathcal{I}^{\star}(\mathcal{RC})$ where $\mathcal{I}^{\star}(\mathcal{RC})$ is the diversity index of the risk contributions.

The optimization problem of the ERC portfolio

If we consider $\mathcal{I}^{\star}\left(\mathcal{RC}\right)$ in place of $\sigma\left(x^{\star}\left(c\right)\right)$, we obtain Figure 41.

The optimization problem of the ERC portfolio

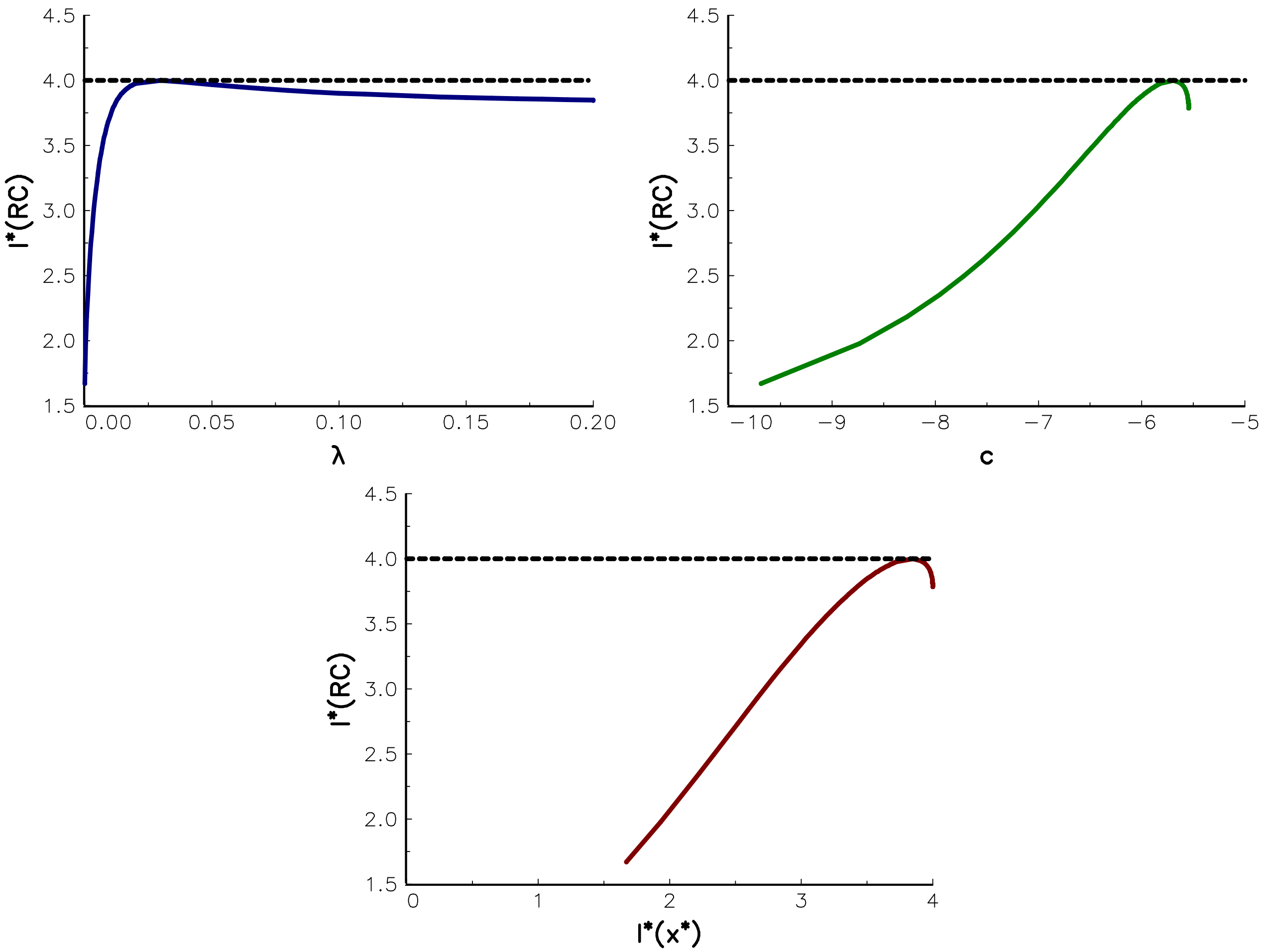


Figure 41: Relationship between λ_c, c, $\mathcal{I}^{\star}\left(x^{\star}\left(c\right)\right)$ and $\mathcal{I}^{\star}\left(\mathcal{RC}\right)$

The optimization problem of the ERC portfolio

Question 1.e

Draw the relationship between $\sigma\left(x^{\star}\left(c\right)\right)$ and $\mathcal{I}^{\star}\left(\mathcal{RC}\right)$. Identify the ERC portfolio.

The optimization problem of the ERC portfolio

In Figure 42, we have reported the relationship between $\sigma\left(x^{\star}\left(c\right)\right)$ and $\mathcal{I}^{\star}\left(\mathcal{RC}\right)$. The ERC portfolio satisfies the equation $\mathcal{I}^{\star}\left(\mathcal{RC}\right)=n$.

The optimization problem of the ERC portfolio

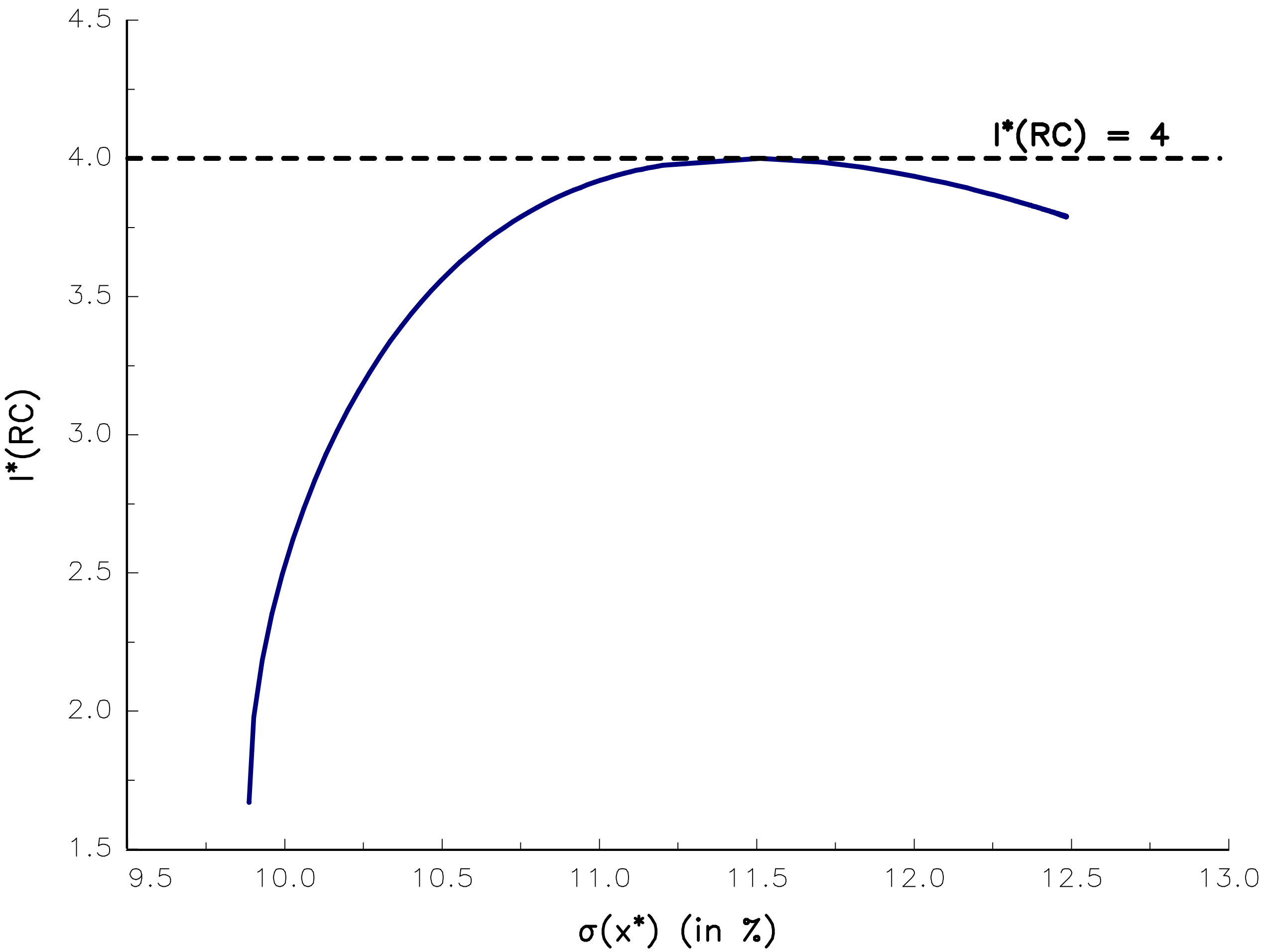


Figure 42: Relationship between $\sigma\left(x^{\star}\left(c\right)\right)$ and $\mathcal{I}^{\star}\left(\mathcal{RC}\right)$

The optimization problem of the ERC portfolio

Question 2

We now consider a slight modification of the previous optimization problem:

$$\begin{aligned} x^{\star}(c) &= \arg\min \sqrt{x^\top \Sigma x} \\ \text{u.c.} & \left\{ \begin{array}{l} \sum_{i=1}^{n} \ln x_i \geq c \\ x \geq \mathbf{0}_n \end{array} \right. \end{aligned} \tag{2}$$

The optimization problem of the ERC portfolio

Question 2.a

Why does the optimization problem (1) not define the ERC portfolio?

The optimization problem of the ERC portfolio

Let us consider the optimization problem when we impose the constraint $\mathbf{1}_n^\top x = 1$. The first-order condition is:

$$\frac{\partial\,\sigma(x)}{\partial\,x_i} - \lambda_i - \lambda_0 - \frac{\lambda_c}{x_i} = 0$$

Because $x_i > 0$, we deduce that $\lambda_i = 0$ and:

$$x_i \frac{\partial\,\sigma(x)}{\partial\,x_i} = \lambda_0 x_i + \lambda_c$$

If this solution corresponds to the ERC portfolio, we obtain:

$$\mathcal{RC}_i = \mathcal{RC}_j \Leftrightarrow \lambda_0 x_i + \lambda_c = \lambda_0 x_j + \lambda_c$$

If $\lambda_0 \neq 0$, we deduce that:

$$x_i = x_j$$

It corresponds to the EW portfolio meaning that the assumption $\mathcal{RC}_i = \mathcal{RC}_j$ is false.

The optimization problem of the ERC portfolio

Question 2.b

Find the optimized portfolio of the optimization problem (2) when c is equal to -10. Calculate the corresponding risk allocation.

The optimization problem of the ERC portfolio

If c is equal to -10, we obtain the following results:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	12.65%	7.75%	0.98%	25.00%
2	8.43%	11.63%	0.98%	25.00%
3	7.06%	13.89%	0.98%	25.00%
4	6.03%	16.25%	0.98%	25.00%
$\sigma(x)$			3.92%	

The optimization problem of the ERC portfolio

Question 2.c

Same question if $c = 0$.

The optimization problem of the ERC portfolio

If c is equal to 0, we obtain the following results:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^{\star}$
1	154.07%	7.75%	11.94%	25.00%
2	102.72%	11.63%	11.94%	25.00%
3	85.97%	13.89%	11.94%	25.00%
4	73.50%	16.25%	11.94%	25.00%
$\sigma(x)$			47.78%	

The optimization problem of the ERC portfolio

Question 2.d

Demonstrate then that the solution to the second optimization problem is:

$$x^{\star}(c) = \exp\left(\frac{c - c_{\mathrm{erc}}}{n}\right) x_{\mathrm{erc}}$$

where $c_{\mathrm{erc}} = \sum_{i=1}^{n} \ln x_{\mathrm{erc},i}$. Comment on this result.

The optimization problem of the ERC portfolio

In this case, the first-order condition is:

$$\frac{\partial\,\sigma\left(x\right)}{\partial\,x_i} - \lambda_i - \frac{\lambda_c}{x_i} = 0$$

As previously, $\lambda_i = 0$ because $x_i > 0$ and we obtain:

$$x_i \frac{\partial\,\sigma\left(x\right)}{\partial\,x_i} = \lambda_c$$

The solution of the second optimization problem is then a non-normalized ERC portfolio because $\sum_{i=1}^n x_i$ is not necessarily equal to 1. If we note $c_{\mathrm{erc}} = \sum_{i=1}^n \ln\left(x_{\mathrm{erc}}\right)_i$, we deduce that:

$$\begin{aligned} x_{\mathrm{erc}} &= \arg\min \sqrt{x^\top \Sigma x} \\ &\text{u.c.} \quad \left\{ \begin{array}{l} \sum_{i=1}^n \ln x_i \geq c_{\mathrm{erc}} \\ x \geq \mathbf{0}_n \end{array} \right. \end{aligned}$$

The optimization problem of the ERC portfolio

Let $x^\star(c)$ be the portfolio defined by:

$$x^\star(c) = \exp\left(\frac{c - c_{\mathrm{erc}}}{n}\right) x_{\mathrm{erc}}$$

We have $x^\star(c) > \mathbf{0}_n$,

$$\sqrt{x^\star(c)^\top \Sigma x^\star(c)} = \exp\left(\frac{c - c_{\mathrm{erc}}}{n}\right) \sqrt{x_{\mathrm{erc}}^\top \Sigma x_{\mathrm{erc}}}$$

and:

$$\begin{aligned}
\sum_{i=1}^n \ln x_i^\star(c) &= \sum_{i=1}^n \ln\left(\exp\left(\frac{c - c_{\mathrm{erc}}}{n}\right) x_{\mathrm{erc}}\right)_i \\
&= c - c_{\mathrm{erc}} + \sum_{i=1}^n \ln\left(x_{\mathrm{erc}}\right)_i \\
&= c
\end{aligned}$$

We conclude that $x^\star(c)$ is the solution of the optimization problem.

The optimization problem of the ERC portfolio

$x^{\star}(c)$ is then a leveraged ERC portfolio if $c > c_{\mathrm{erc}}$ and a deleveraged ERC portfolio if $c < c_{\mathrm{erc}}$.

In our example, c_{erc} is equal to -5.7046. If $c = -10$, we have:

$$\exp\left(\frac{c - c_{\mathrm{erc}}}{n}\right) = 34.17\%$$

We verify that the solution of Question 2.b is such that $\sum_{i=1}^{n} x_i = 34.17\%$ and $\mathcal{RC}_i^{\star} = \mathcal{RC}_j^{\star}$.

If $c = 0$, we obtain:

$$\exp\left(\frac{c - c_{\mathrm{erc}}}{n}\right) = 416.26\%$$

In this case, the solution is a leveraged ERC portfolio.

The optimization problem of the ERC portfolio

Question 2.e

Show that there exists a scalar c such that the Lagrange coefficient λ_0 of the optimization problem (1) is equal to zero. Deduce then that the volatility of the ERC portfolio is between the volatility of the long-only minimum variance portfolio and the volatility of the equally weighted portfolio:

$$\sigma\left(x_{\mathrm{mv}}\right) \leq \sigma\left(x_{\mathrm{erc}}\right) \leq \sigma\left(x_{\mathrm{ew}}\right)$$

The optimization problem of the ERC portfolio

From the previous question, we know that the ERC optimization portfolio is the solution of the second optimization problem if we use c_{erc} for the control variable. In this case, we have $\sum_{i=1}^{n} x_i^\star \left(c_{\mathrm{erc}}\right) = 1$ meaning that x_{erc} is also the solution of the first optimization problem. We deduce that $\lambda_0 = 0$ if $c = c_{\mathrm{erc}}$. The first optimization problem is a convex problem with a convex inequality constraint. The objective function is then an increasing function of the control variable c:

$$c_1 \leq c_2 \Rightarrow \sigma\left(x^\star\left(c_1\right)\right) \geq \sigma\left(x^\star\left(c_2\right)\right)$$

The optimization problem of the ERC portfolio

We have seen that the minimum variance portfolio corresponds to $c = -\infty$, that the EW portfolio is obtained with $c = -n \ln n$ and that the ERC portfolio is the solution of the optimization problem when c is equal to c_{erc}. Moreover, we have $-\infty \leq c_{\mathrm{erc}} \leq -n \ln n$. We deduce that the volatility of the ERC portfolio is between the volatility of the long-only minimum variance portfolio and the volatility of the equally weighted portfolio:

$$\sigma\left(x_{\mathrm{mv}}\right) \leq \sigma\left(x_{\mathrm{erc}}\right) \leq \sigma\left(x_{\mathrm{ew}}\right)$$

Risk parity funds

Question 1

We consider a universe of three asset classes[a] which are stocks (S), bonds (B) and commodities (C). We have computed the one-year historical covariance matrix of asset returns for different dates and we obtain the following results (all the numbers are expressed in %):

	31/12/1999			31/12/2002			30/12/2005		
σ_i	12.40	5.61	12.72	20.69	7.36	13.59	7.97	7.01	16.93
$\rho_{i,j}$	100.00			100.00			100.00		
	−5.89	100.00		−36.98	100.00		29.25	100.00	
	−4.09	−7.13	100.00	22.74	−13.12	100.00	15.75	15.05	100.00
	31/12/2007			31/12/2008			31/12/2010		
σ_i	12.94	5.50	14.54	33.03	9.73	29.00	16.73	6.88	16.93
$\rho_{i,j}$	100.00	−25.76		100.00			100.00		
	−25.76	100.00		−16.26	100.00		15.31	100.00	
	31.91	6.87	100.00	47.31	9.13	100.00	64.13	15.46	100.00

[a] In fact, we use the MSCI World index, the Citigroup WGBI index and the DJ UBS Commodity index to represent these asset classes.

Risk parity funds

Question 1.a

Compute the weights and the volatility of the risk parity[a] (RP portfolio) portfolios for the different dates.

[a]Here, risk parity refers to the ERC portfolio when we do not take into account the correlations.

Risk parity funds

The RP portfolio is defined as follows:

$$x_i = \frac{\sigma_i^{-1}}{\sum_{j=1}^{n} \sigma_j^{-1}}$$

We obtain the following results:

Date	1999	2002	2005	2007	2008	2010
S	23.89%	18.75%	38.35%	23.57%	18.07%	22.63%
B	52.81%	52.71%	43.60%	55.45%	61.35%	55.02%
C	23.29%	28.54%	18.05%	20.98%	20.58%	22.36%
$\sigma(x)$	4.83%	6.08%	6.26%	5.51%	11.64%	8.38%

Risk parity funds

Question 1.b

Same question by considering the ERC portfolio.

Risk parity funds

In the ERC portfolio, the risk contributions are equal for all the assets:

$$\mathcal{RC}_i = \mathcal{RC}_j$$

with:

$$\mathcal{RC}_i = \frac{x_i \cdot (\Sigma x)_i}{\sqrt{x^\top \Sigma x}} \tag{3}$$

We obtain the following results:

Date	1999	2002	2005	2007	2008	2010
S	23.66%	18.18%	37.85%	23.28%	17.06%	20.33%
B	53.12%	58.64%	43.18%	59.93%	66.39%	59.61%
C	23.22%	23.18%	18.97%	16.79%	16.54%	20.07%
$\sigma(x)$	4.82%	5.70%	6.32%	5.16%	10.77%	7.96%

Risk parity funds

Question 1.c

What do you notice about the volatility of RP and ERC portfolios? Explain these results.

Risk parity funds

We notice that $\sigma(x_{\mathrm{erc}}) \leq \sigma(x_{\mathrm{rp}})$ except for the year 2005. This date corresponds to positive correlations between assets. Moreover, the correlation between stocks and bonds is the highest. Starting from the RP portfolio, it is then possible to approach the ERC portfolio by reducing the weights of stocks and bonds and increasing the weight of commodities. At the end, we find an ERC portfolio that has a slightly higher volatility.

Risk parity funds

Question 1.d

Find the analytical expression of the volatility $\sigma(x)$, the marginal risk $\mathcal{MR}_i$, the risk contribution $\mathcal{RC}_i$ and the normalized risk contribution $\mathcal{RC}_i^\star$ in the case of RP portfolios.

Risk parity funds

The volatility of the RP portfolio is:

$$\begin{aligned}
\sigma(x) &= \frac{1}{\sum_{j=1}^{n} \sigma_j^{-1}} \sqrt{\left(\sigma^{-1}\right)^\top \Sigma \sigma^{-1}} \\
&= \frac{1}{\sum_{j=1}^{n} \sigma_j^{-1}} \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{\sigma_i \sigma_j} \rho_{i,j} \sigma_i \sigma_j} \\
&= \frac{1}{\sum_{j=1}^{n} \sigma_j^{-1}} \sqrt{n + 2 \sum_{i>j} \rho_{i,j}} \\
&= \frac{1}{\sum_{j=1}^{n} \sigma_j^{-1}} \sqrt{n \left(1 + (n-1)\bar{\rho}\right)}
\end{aligned}$$

where $\bar{\rho}$ is the average correlation between asset returns.

Risk parity funds

For the marginal risk, we obtain:

$$
\begin{aligned}
\mathcal{MR}_i &= \frac{\left(\Sigma \sigma^{-1}\right)_i}{\sigma(x) \sum_{j=1}^{n} \sigma_j^{-1}} \\
&= \frac{1}{\sqrt{n\left(1+(n-1)\bar{\rho}\right)}} \sum_{j=1}^{n} \rho_{i,j} \sigma_i \sigma_j \frac{1}{\sigma_j} \\
&= \frac{\sigma_i}{\sqrt{n\left(1+(n-1)\bar{\rho}\right)}} \sum_{j=1}^{n} \rho_{i,j} \\
&= \frac{\sigma_i \bar{\rho}_i \sqrt{n}}{\sqrt{1+(n-1)\bar{\rho}}}
\end{aligned}
$$

where $\bar{\rho}_i$ is the average correlation of asset i with the other assets (including itself).

Risk parity funds

The expression of the risk contribution is then:

$$\begin{aligned}\mathcal{RC}_i &= \frac{\sigma_i^{-1}}{\sum_{j=1}^n \sigma_j^{-1}} \frac{\sigma_i \bar{\rho}_i \sqrt{n}}{\sqrt{1+(n-1)\bar{\rho}}} \\ &= \frac{\bar{\rho}_i \sqrt{n}}{\sqrt{1+(n-1)\bar{\rho}} \sum_{j=1}^n \sigma_j^{-1}}\end{aligned}$$

We deduce that the normalized risk contribution is:

$$\begin{aligned}\mathcal{RC}_i^\star &= \frac{\bar{\rho}_i \sqrt{n}}{\sigma(x) \sqrt{1+(n-1)\bar{\rho}} \sum_{j=1}^n \sigma_j^{-1}} \\ &= \frac{\bar{\rho}_i}{1+(n-1)\bar{\rho}}\end{aligned}$$

Risk parity funds

Question 1.e

Compute the normalized risk contributions of the previous RP portfolios. Comment on these results.

Risk parity funds

We obtain the following normalized risk contributions:

Date	1999	2002	2005	2007	2008	2010
S	33.87%	34.96%	34.52%	32.56%	34.45%	36.64%
B	32.73%	20.34%	34.35%	24.88%	24.42%	26.70%
C	33.40%	44.69%	31.14%	42.57%	41.13%	36.67%

We notice that the risk contributions are not exactly equal for all the assets. Generally, the risk contribution of bonds is lower than the risk contribution of equities, which is itself lower than the risk contribution of commodities.

Risk parity funds

Question 2

We consider four parameter sets of risk budgets:

Set	b_1	b_2	b_3
#1	45%	45%	10%
#2	70%	10%	20%
#3	20%	70%	10%
#4	25%	25%	50%

Risk parity funds

Question 2.a

Compute the RB portfolios for the different dates.

Risk parity funds

We obtain the following RB portfolios:

Date	b_i	1999	2002	2005	2007	2008	2010
S	45%	26.83%	22.14%	42.83%	27.20%	20.63%	25.92%
B	45%	59.78%	66.10%	48.77%	66.15%	73.35%	67.03%
C	10%	13.39%	11.76%	8.40%	6.65%	6.02%	7.05%
S	70%	40.39%	29.32%	65.53%	39.37%	33.47%	46.26%
B	10%	37.63%	51.48%	19.55%	47.18%	52.89%	37.76%
C	20%	21.98%	19.20%	14.93%	13.45%	13.64%	15.98%
S	20%	17.55%	16.02%	25.20%	18.78%	12.94%	13.87%
B	70%	69.67%	71.70%	66.18%	74.33%	80.81%	78.58%
C	10%	12.78%	12.28%	8.62%	6.89%	6.24%	7.55%
S	25%	21.69%	15.76%	34.47%	20.55%	14.59%	16.65%
B	25%	48.99%	54.03%	39.38%	55.44%	61.18%	53.98%
C	50%	29.33%	30.21%	26.15%	24.01%	24.22%	29.37%

Risk parity funds

Question 2.b

Compute the implied risk premium $\tilde{\pi}_i$ of the assets for these portfolios if we assume a Sharpe ratio equal to 0.40.

Risk parity funds

To compute the implied risk premium $\tilde{\pi}_i$, we use the following formula (TR-RPB, page 274):

$$\begin{aligned} \tilde{\pi}_i &= \mathrm{SR}\left(x \mid r\right) \cdot \mathcal{MR}_i \\ &= \mathrm{SR}\left(x \mid r\right) \cdot \frac{\left(\Sigma x\right)_i}{\sigma\left(x\right)} \end{aligned}$$

where $\mathrm{SR}\left(x \mid r\right)$ is the Sharpe ratio of the portfolio.

Risk parity funds

We obtain the following results:

Date	b_i	1999	2002	2005	2007	2008	2010
S	45%	3.19%	4.60%	2.49%	3.15%	8.64%	5.20%
B	45%	1.43%	1.54%	2.19%	1.29%	2.43%	2.01%
C	10%	1.42%	1.92%	2.82%	2.86%	6.58%	4.24%
S	70%	4.05%	6.45%	2.86%	4.31%	11.56%	6.32%
B	10%	0.62%	0.52%	1.37%	0.51%	1.04%	1.11%
C	20%	2.13%	2.81%	3.59%	3.61%	8.11%	5.23%
S	20%	2.06%	2.68%	1.91%	1.93%	5.61%	3.91%
B	70%	1.82%	2.10%	2.54%	1.71%	3.14%	2.42%
C	10%	1.42%	1.75%	2.79%	2.64%	5.82%	3.60%
S	25%	2.33%	3.78%	1.98%	2.74%	8.06%	5.13%
B	25%	1.03%	1.10%	1.74%	1.02%	1.92%	1.58%
C	50%	3.45%	3.95%	5.23%	4.69%	9.71%	5.82%

Risk parity funds

Question 2.c

Comment on these results.

Risk parity funds

We have:

$$x_i \tilde{\pi}_i = \mathrm{SR}\left(x \mid r\right) \cdot \mathcal{RC}_i$$

We deduce that:

$$\tilde{\pi}_i \propto \frac{b_i}{x_i}$$

x_i is generally an increasing function of b_i. As a consequence, the relationship between the risk budgets b_i and the risk premiums $\tilde{\pi}_i$ is not necessarily increasing. However, we notice that the bigger the risk budget, the higher the risk premium. This is easily explained. If an investor allocates more risk budget to one asset class than another investor, he thinks that the risk premium of this asset class is higher than the other investor.

Risk parity funds

However, we must be careful. This interpretation is valid if we compare two sets of risk budgets. It is false if we compare the risk budgets among themselves. For instance, if we consider the third parameter set, the risk budget of bonds is 70% whereas the risk budget of stocks is 20%. It does not mean that the risk premium of bonds is higher than the risk premium of equities. In fact, we observe the contrary. If we would like to compare risk budgets among themselves, the right measure is the implied Sharpe ratio, which is equal to:

$$\begin{aligned}\mathrm{SR}_i &= \frac{\tilde{\pi}_i}{\sigma_i} \\ &= \mathrm{SR}\left(x \mid r\right) \cdot \frac{\mathcal{MR}_i}{\sigma_i}\end{aligned}$$

For instance, if we consider the most diversified portfolio, the marginal risk is proportional to the volatility and we retrieve the result that Sharpe ratios are equal if the MDP is optimal.

Main references

Roncalli, T. (2013)
Introduction to Risk Parity and Budgeting, Chapman and Hall/CRC Financial Mathematics Series, Chapter 2.

Roncalli, T. (2013)
Introduction to Risk Parity and Budgeting — Companion Book, Chapman and Hall/CRC Financial Mathematics Series, Chapter 2.

References I

- Bruder, B., Kostyuchyk, N., and Roncalli, T. (2016)
  Risk Parity Portfolios with Skewness Risk: An Application to Factor Investing and Alternative Risk Premia, *SSRN*, www.ssrn.com/abstract=2813384.

- Lezmi, E., Malongo, H., Roncalli, T., and Sobotka, R. (2018)
  Portfolio Allocation with Skewness Risk: A Practical Guide, *SSRN*, www.ssrn.com/abstract=3201319.

- Maillard, S., Roncalli, T. and Teïletche, J. (2010)
  The Properties of Equally Weighted Risk Contribution Portfolios, *Journal of Portfolio Management*, 36(4), pp. 60-70.

- Roncalli, T. (2015)
  Introducing Expected Returns into Risk Parity Portfolios: A New Framework for Asset Allocation, *Bankers, Markets & Investors*, 138, pp. 18-28.

References II

Roncalli, T., and Weisang, G. (2016)
Risk Parity Portfolios with Risk Factors, *Quantitative Finance*, 16(3), pp. 377-388

Scherer, B. (2007)
Portfolio Construction & Risk Budgeting, Third edition, Risk Books.

Asset Management
Lecture 3. Smart Beta, Factor Investing and Alternative Risk Premia

Thierry Roncalli*

*University of Paris-Saclay

January 2021

Agenda

- Lecture 1: Portfolio Optimization
- Lecture 2: Risk Budgeting
- **Lecture 3: Smart Beta, Factor Investing and Alternative Risk Premia**
- Lecture 4: Green and Sustainable Finance, ESG Investing and Climate Risk
- Lecture 5: Machine Learning in Asset Management

Cap-weighted indexation and modern portfolio theory

Rationale of market-cap indexation

- **Separation Theorem**: there is one unique risky portfolio owned by investors called the tangency portfolio (Tobin, 1958)
- **CAPM**: the tangency portfolio is the Market portfolio, best represented by the capitalization-weighted index (Sharpe, 1964)
- **Performance of active management**: negative alpha in equity mutual funds on average (Jensen, 1968)
- **EMH**: markets are efficient (Fama, 1970)
- **Passive management**: launch of the first index fund (John McQuown, Wells Fargo Investment Advisors, Samsonite Luggage Corporation, 1971)
- **First S&P 500 index fund** by Wells Fargo and American National Bank in Chicago (1973)
- The **first listed ETF** was the SPDRs (Ticker: SPY) in 1993

Index funds

Mutual Fund (MF)

A mutual fund is a **collective investment fund** that are regulated and sold to the general public

Exchange Traded Fund (ETF)

It is a mutual fund which trades **intra-day** on a securities exchange (thanks to market makers)

Exchange Traded Product (ETP)

It is a security that is **derivatively-priced** and that trades intra-day on an exchange. ETPs includes exchange traded funds (ETFs), exchange traded vehicles (ETVs), exchange traded notes (ETNs) and certificates.

Pros of market-cap indexation

- A convenient and **recognized approach** to participate to broad equity markets
- **Management simplicity**: low turnover & transaction costs

Construction of an equity index

- We consider an index universe composed of n stocks
- Let $P_{i,t}$ be the price of the i^{th} stock and $R_{i,t}$ be the corresponding return between times $t-1$ and t:

$$R_{i,t} = \frac{P_{i,t}}{P_{i,t-1}} - 1$$

- The value of the index B_t at time t is defined by:

$$B_t = \varphi \sum_{i=1}^{n} N_i P_{i,t}$$

where φ is a scaling factor and N_i is the total number of shares issued by the company i

Construction of an equity index

- Another expression of B_t is[8]:

$$\begin{aligned} B_t &= \varphi \sum_{i=1}^{n} N_i P_{i,t-1} \left(1 + R_{i,t}\right) \\ &= B_{t-1} \frac{\sum_{i=1}^{n} N_i P_{i,t-1} \left(1 + R_{i,t}\right)}{\sum_{i=1}^{n} N_i P_{i,t-1}} \\ &= B_{t-1} \sum_{i=1}^{n} w_{i,t-1} \left(1 + R_{i,t}\right) \end{aligned}$$

where $w_{i,t-1}$ is the weight of the i^{th} stock in the index:

$$w_{i,t-1} = \frac{N_i P_{i,t-1}}{\sum_{i=1}^{n} N_i P_{i,t-1}}$$

- The computation of the index value B_t can be done at the closing time t and also in an intra-day basis

[8] B_0 can be set to an arbitrary value (e.g. 100, 500, 1000 or 5000)

Construction of an equity index

Remark

The previous computation is purely theoretical because the portfolio corresponds to all the shares outstanding of the n stocks $\Rightarrow$ impossible to hold this portfolio

Remark

Most of equity indices use floating shares[a] instead of shares outstanding

[a] They indicate the number of shares available for trading

Replication of an equity index

- In order to replicate this index, we must build a hedging strategy that consists in investing in stocks
- Let S_t be the value of the strategy (or the index fund):

$$S_t = \sum_{i=1}^{n} n_{i,t} P_{i,t}$$

where $n_{i,t}$ is the number of stock i held between $t-1$ and t

- The tracking error is the difference between the return of the strategy and the return of the index:

$$e_t \left(S \mid B \right) = R_{S,t} - R_{B,t}$$

Replication of an equity index

The quality of the replication process is measured by the volatility $\sigma(e_t(S \mid B))$ of the tracking error. We may distinguish several cases:

1. Index funds with low tracking error volatility (less than 10 bps) $\Rightarrow$ physical replication or synthetic replication
2. Index funds with moderate tracking error volatility (between 10 bps and 50 bps) $\Rightarrow$ sampling replication
3. Index funds with higher tracking error volatility (larger than 50 bps) $\Rightarrow$ equity universes with liquidity problems and enhanced/tilted index funds

Replication of an equity index

- In a capitalization-weighted index, the weights are given by:

$$w_{i,t} = \frac{C_{i,t}}{\sum_{j=1}^{n} C_{j,t}} = \frac{N_{i,t} P_{i,t}}{\sum_{j=1}^{n} N_{j,t} P_{j,t}}$$

where $N_{i,t}$ and $C_{i,t} = N_{i,t} P_{i,t}$ are the number of shares outstanding and the market capitalization of the i^{th} stock

- If we have a perfect match at time $t-1$:

$$\frac{n_{i,t-1} P_{i,t-1}}{\sum_{i=1}^{n} n_{i,t-1} P_{i,t-1}} = w_{i,t-1}$$

we have a perfect match at time t:

$$n_{i,t} = n_{i,t-1}$$

Replication of an equity index

- We do not need to rebalance the hedging portfolio because of the relationship:

$$n_{i,t}P_{i,t} \propto w_{i,t}P_{i,t}$$

- Therefore, it is not necessarily to adjust the portfolio of the strategy (except if there are subscriptions or redemptions)

A CW index fund remains the most efficient investment in terms of management simplicity, turnover and transaction costs

Cons of market-cap indexation

- Trend-following strategy: momentum bias leads to bubble risk exposure as weight of best performers ever increases
  $\Rightarrow$ Mid 2007, financial stocks represent 40% of the Eurostoxx 50 index
- Growth bias as high valuation multiples stocks weight more than low-multiple stocks with equivalent realized earnings.
  $\Rightarrow$ Mid 2000, the 8 stocks of the technology/telecom sectors represent 35% of the Eurostoxx 50 index
  $\Rightarrow$ $2^1/_2$ years later after the dot.com bubble, these two sectors represent 12%
- Concentrated portfolios
  $\Rightarrow$ The top 100 market caps of the S&P 500 account for around 70%
- Lack of risk diversification and high drawdown risk: no portfolio construction rules leads to concentration issues (e.g. sectors, stocks).

Cons of market-cap indexation

Some illustrations

- Mid 2000: 8 Technology/Telecom stocks represent 35% of the Eurostoxx 50 index
- In 2002: 7.5% of the Eurostoxx 50 index is invested into Nokia with a volatility of 70%
- Dec. 2006: 26.5% of the MSCI World index is invested in financial stocks
- June 2007: 40% of the Eurostoxx 50 is invested into Financials
- January 2013: 20% of the S&P 500 stocks represent 68% of the S&P 500 risk
- Between 2002 and 2012, two stocks contribute on average to more than 20% of the monthly performance of the Eurostoxx 50 index

Cons of market-cap indexation

Table 47: Weight and risk concentration of several equity indices (June 29, 2012)

Ticker	Weights				Risk contributions			
	$\mathcal{G}(x)$	$\mathbb{L}(x)$ 10%	$\mathbb{L}(x)$ 25%	$\mathbb{L}(x)$ 50%	$\mathcal{G}(x)$	$\mathbb{L}(x)$ 10%	$\mathbb{L}(x)$ 25%	$\mathbb{L}(x)$ 50%
SX5P	30.8	24.1	48.1	71.3	26.3	19.0	40.4	68.6
SX5E	31.2	23.0	46.5	72.1	31.2	20.5	44.7	73.3
INDU	33.2	23.0	45.0	73.5	35.8	25.0	49.6	75.9
BEL20	39.1	25.8	49.4	79.1	45.1	25.6	56.8	82.5
DAX	44.0	27.5	56.0	81.8	47.3	27.2	59.8	84.8
CAC	47.4	34.3	58.3	82.4	44.1	31.9	57.3	79.7
AEX	52.2	37.2	61.3	86.0	51.4	35.3	62.0	84.7
HSCEI	54.8	39.7	69.3	85.9	53.8	36.5	67.2	85.9
NKY	60.2	47.9	70.4	87.7	61.4	49.6	70.9	88.1
UKX	60.8	47.5	73.1	88.6	60.4	46.1	72.8	88.7
SXXE	61.7	49.2	73.5	88.7	63.9	51.6	75.3	90.1
SPX	61.8	52.1	72.0	87.8	59.3	48.7	69.9	86.7
MEXBOL	64.6	48.2	75.1	91.8	65.9	45.7	78.6	92.9
IBEX	64.9	51.7	77.3	90.2	68.3	58.2	80.3	91.4
SXXP	65.6	55.0	76.4	90.1	64.2	52.0	75.5	90.0
NDX	66.3	58.6	77.0	89.2	64.6	56.9	74.9	88.6
TWSE	79.7	73.4	86.8	95.2	79.7	72.6	87.3	95.7
TPX	80.8	72.8	88.8	96.3	83.9	77.1	91.0	97.3
KOSPI	86.5	80.6	93.9	98.0	89.3	85.1	95.8	98.8

$\mathcal{G}(x)$ = Gini coefficient, $\mathbb{L}(x)$ = Lorenz curve

Cons of market-cap indexation

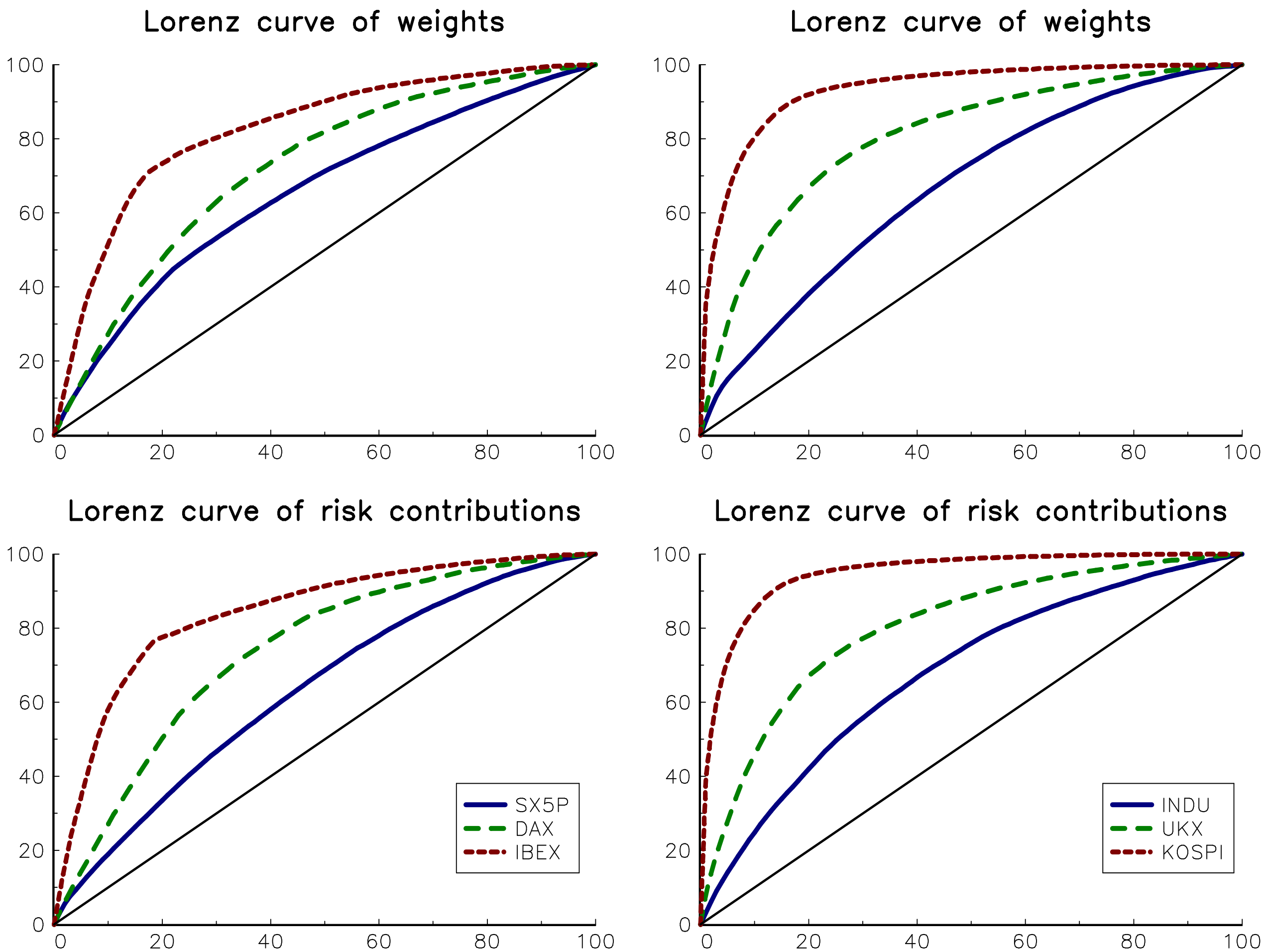


Figure 43: Lorenz curve of several equity indices (June 29, 2012)

Capturing the equity risk premium

	APPLE	EXXON	MSFT	J&J	IBM	PFIZER	CITI	McDO
	Cap-weighted allocation (in %)							
Dec. 1999	1.05	12.40	38.10	7.94	12.20	12.97	11.89	3.46
Dec. 2004	1.74	22.16	19.47	12.61	11.00	13.57	16.76	2.70
Dec. 2008	6.54	35.03	14.92	14.32	9.75	10.30	3.15	5.98
Dec. 2010	18.33	22.84	14.79	10.52	11.29	8.69	8.51	5.02
Dec. 2012	26.07	20.55	11.71	10.12	11.27	9.62	6.04	4.61
Jun. 2013	20.78	19.80	14.35	11.64	11.36	9.51	7.79	4.77
	Implied risk premium (in %)							
Dec. 1999	5.96	2.14	8.51	3.61	5.81	5.91	6.19	2.66
Dec. 2004	3.88	2.66	2.79	2.03	2.32	3.90	3.02	1.86
Dec. 2008	9.83	11.97	10.48	6.24	7.28	8.06	17.15	6.28
Dec. 2010	5.38	3.85	4.42	2.29	3.66	3.76	6.52	2.54
Dec. 2012	5.87	2.85	3.58	1.44	2.80	1.77	5.91	1.88
Jun. 2013	5.59	2.79	3.60	1.55	2.92	1.91	5.24	1.82
	Expected performance contribution (in %)							
Dec. 1999	1.01	4.31	52.63	4.66	11.52	12.43	11.94	1.49
Dec. 2004	2.41	21.04	19.44	9.15	9.12	18.93	18.11	1.79
Dec. 2008	6.60	43.00	16.04	9.17	7.28	8.52	5.55	3.85
Dec. 2010	23.58	21.01	15.62	5.77	9.89	7.81	13.27	3.05
Dec. 2012	42.41	16.23	11.61	4.04	8.73	4.71	9.88	2.40
Jun. 2013	33.96	16.18	15.10	5.28	9.69	5.32	11.93	2.53

Alternative-weighted indexation

Definition

Alternative-weighted indexation aims at building passive indexes where the weights are not based on market capitalization

Alternative-weighted indexation

Three kinds of responses:

1. Fundamental indexation (capturing *alpha*?)
   1. Dividend yield indexation
   2. RAFI indexation
2. Risk-based indexation (capturing *diversification*?)
   1. Equally weighted portfolio
   2. Minimum variance portfolio
   3. Equal risk contribution portfolio
   4. Most diversified portfolio
3. Factor investing (capturing *normal returns or beta*? *abnormal returns or alpha*?)
   1. The market risk factor is not the only systematic risk factor
   2. Other factors: size, value, momentum, low beta, quality, etc.

Alternative-weighted indexation

2008

Smart Beta

=

Fundamental Indexation

+

Risk-Based Indexation

Today

Smart Beta

=

Risk-Based Indexation

+

Factor Investing

Equally-weighted portfolio

- The underlying idea of the equally weighted or '$1/n$' portfolio is to define a portfolio independently from the estimated statistics and properties of stocks
- If we assume that it is impossible to predict return and risk, then attributing an equal weight to all of the portfolio components constitutes a natural choice
- We have:

$$x_i = x_j = \frac{1}{n}$$

Equally-weighted portfolio

The portfolio volatility is equal to:

$$\begin{aligned}\sigma^2(x) &= \sum_{i=1}^{n} x_i^2\sigma_i^2 + 2\sum_{i>j} x_i x_j \rho_{i,j}\sigma_i\sigma_j \\ &= \frac{1}{n^2}\left(\sum_{i=1}^{n}\sigma_i^2 + 2\sum_{i>j}\rho_{i,j}\sigma_i\sigma_j\right)\end{aligned}$$

If we assume that $\sigma_i \leq \sigma_{\max}$ and $0 \leq \rho_{i,j} \leq \rho_{\max}$, we obtain:

$$\begin{aligned}\sigma^2(x) &\leq \frac{1}{n^2}\left(\sum_{i=1}^{n}\sigma_{\max}^2 + 2\sum_{i>j}\rho_{\max}\sigma_{\max}^2\right) \\ &\leq \frac{1}{n^2}\left(n\sigma_{\max}^2 + 2\frac{n(n-1)}{2}\rho_{\max}\sigma_{\max}^2\right) \\ &\leq \left(\frac{1+(n-1)\rho_{\max}}{n}\right)\sigma_{\max}^2\end{aligned}$$

Equally-weighted portfolio

We deduce that:

$$\lim_{n \to \infty} \sigma(x) \leq \sigma_{\max}(x) = \sigma_{\max}\sqrt{\rho_{\max}}$$

Table 48: Value of $\sigma_{\max}(x)$ (in %)

		$\sigma_{\max}$ (in %)					
		5.00	10.00	15.00	20.00	25.00	30.00
$\rho_{\max}$ (in %)	10.00	1.58	3.16	4.74	6.32	7.91	9.49
	20.00	2.24	4.47	6.71	8.94	11.18	13.42
	30.00	2.74	5.48	8.22	10.95	13.69	16.43
	40.00	3.16	6.32	9.49	12.65	15.81	18.97
	50.00	3.54	7.07	10.61	14.14	17.68	21.21
	75.00	4.33	8.66	12.99	17.32	21.65	25.98
	90.00	4.74	9.49	14.23	18.97	23.72	28.46
	99.00	4.97	9.95	14.92	19.90	24.87	29.85

Equally-weighted portfolio

If the volatilities are the same ($\sigma_i = \sigma$) and the correlation matrix is constant ($\rho_{i,j} = \rho$), we deduce that:

$$\sigma(x) = \sigma \sqrt{\frac{1 + (n-1)\rho}{n}}$$

Correlations are more important than volatilities to benefit from diversification (= risk reduction)

Equally-weighted portfolio

Result

The main interest of the EW portfolio is the volatility reduction

It is called "naive diversification"

Equally-weighted portfolio

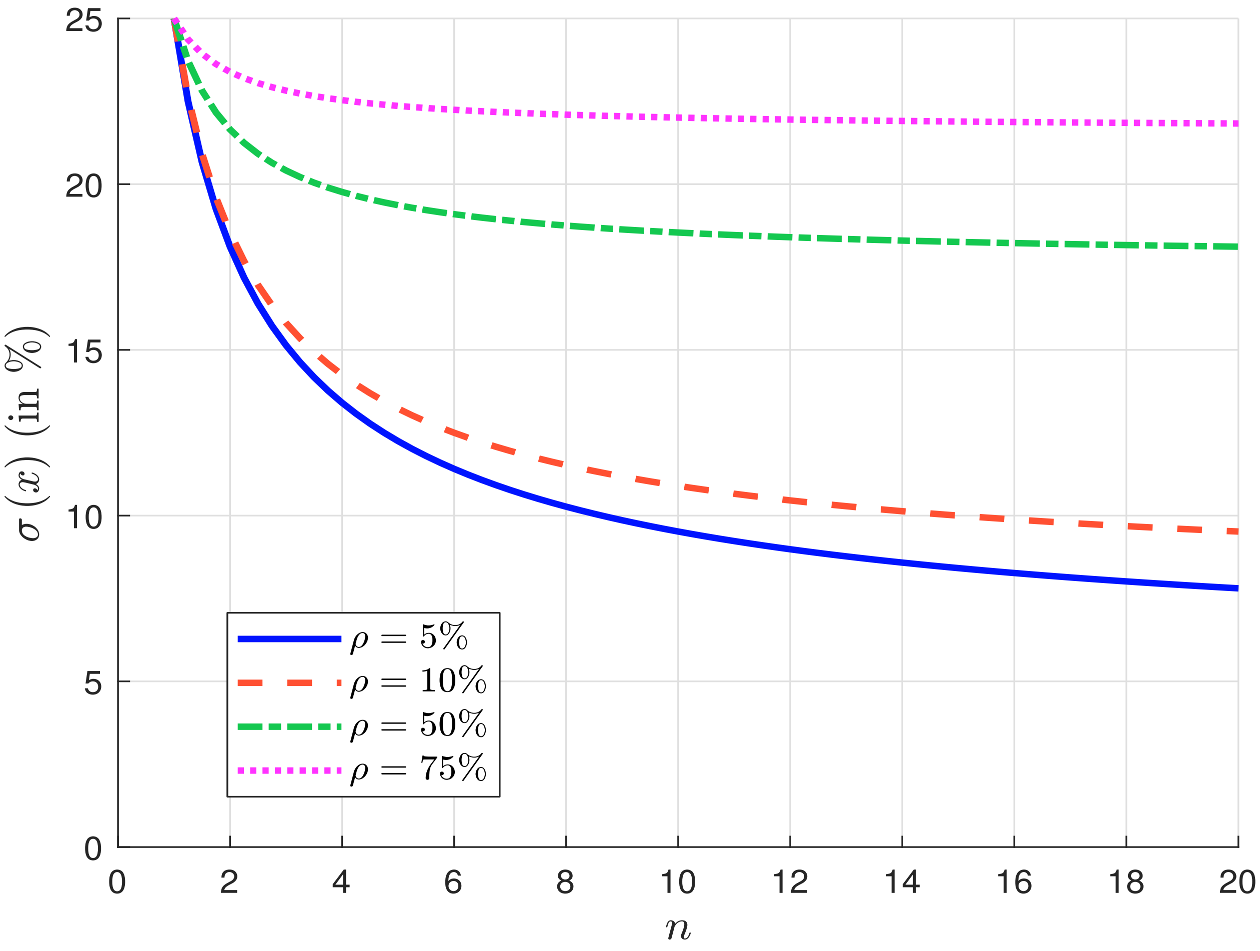


Figure 44: Illustration of the diversification effect ($\sigma = 25\%$)

Equally-weighted portfolio

Another interest of the EW portfolio is its good out-of-sample performance:

> *"We evaluate the out-of-sample performance of the sample-based mean-variance model, and its extensions designed to reduce estimation error, relative to the naive $1/n$ portfolio. Of the 14 models we evaluate across seven empirical datasets, none is consistently better than the $1/n$ rule in terms of Sharpe ratio, certainty-equivalent return, or turnover, which indicates that, out of sample, the gain from optimal diversification is more than offset by estimation error" (DeMiguel et al., 2009)*

Minimum variance portfolio

The global minimum variance (GMV) portfolio corresponds to the following optimization program:

$$\begin{aligned} x_{\mathrm{gmv}} &= \arg\min \frac{1}{2} x^\top \Sigma x \\ \text{u.c.} & \quad \mathbf{1}_n^\top x = 1 \end{aligned}$$

Minimum variance portfolio

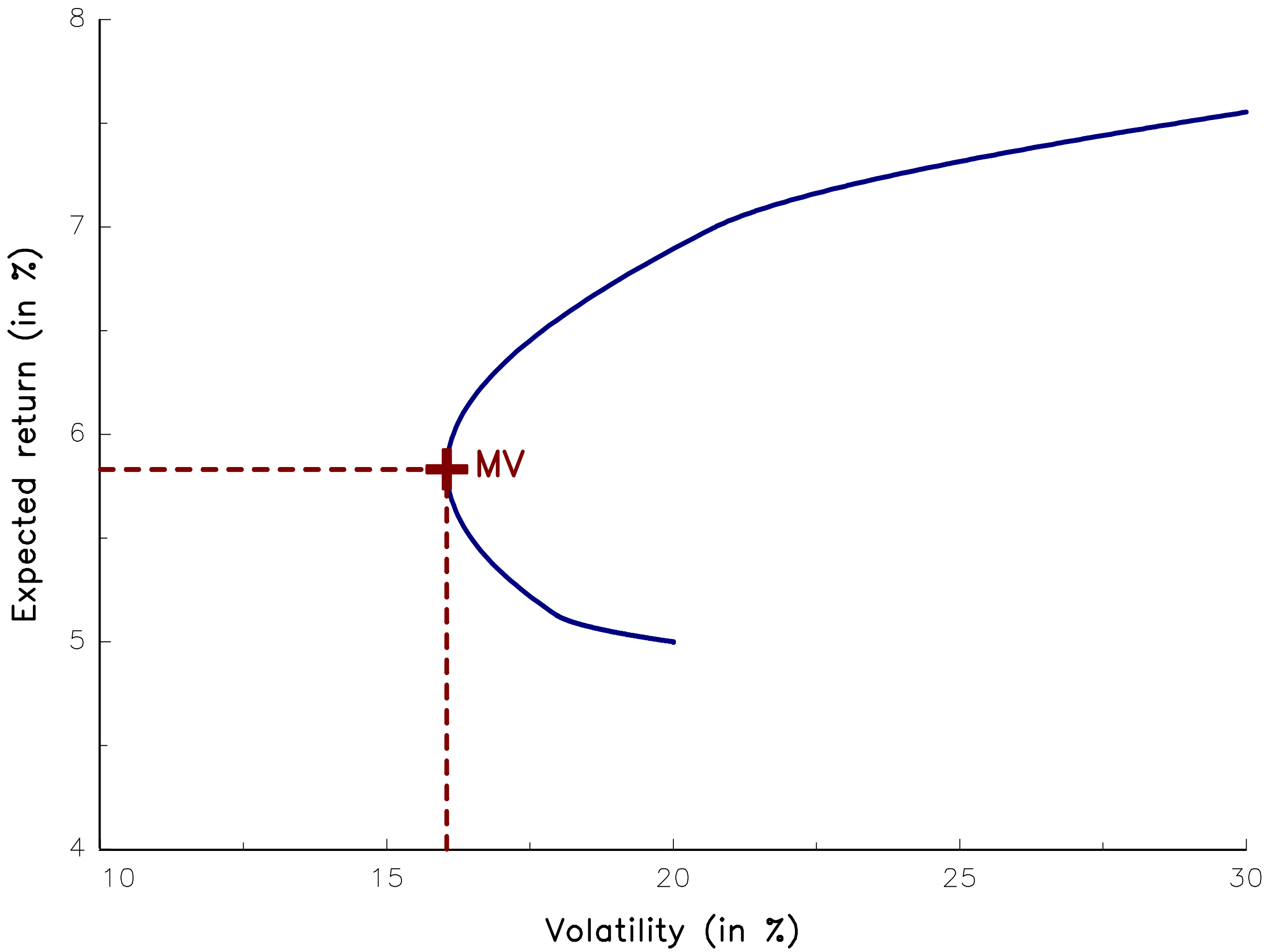


Figure 45: Location of the minimum variance portfolio in the efficient frontier

Minimum variance portfolio

The Lagrange function is equal to:

$$\mathcal{L}(x;\lambda_0) = \frac{1}{2}x^\top \Sigma x - \lambda_0 \left(\mathbf{1}_n^\top x - 1\right)$$

The first-order condition is:

$$\frac{\partial\,\mathcal{L}(x;\lambda_0)}{\partial\,x} = \Sigma x - \lambda_0 \mathbf{1}_n = \mathbf{0}_n$$

We deduce that:

$$x = \lambda_0 \Sigma^{-1} \mathbf{1}_n$$

Since we have $\mathbf{1}_n^\top x = 1$, the Lagrange multiplier satisfies:

$$\mathbf{1}_n^\top \left(\lambda_0 \Sigma^{-1} \mathbf{1}_n\right) = 1$$

or:

$$\lambda_0 = \frac{1}{\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n}$$

Minimum variance portfolio

Theorem

The GMV portfolio is given by the following formula:

$$x_{\mathrm{gmv}} = \frac{\Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^\top \Sigma^{-1}\mathbf{1}_n}$$

Minimum variance portfolio

The volatility of the GMV portfolio is equal to:

$$\begin{aligned}
\sigma^2(x_{\text{gmv}}) &= x_{\text{gmv}}^\top \Sigma x_{\text{gmv}} \\
&= \frac{\mathbf{1}_n^\top \Sigma^{-1}}{\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n} \Sigma \frac{\Sigma^{-1} \mathbf{1}_n}{\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n} \\
&= \frac{\mathbf{1}_n^\top \Sigma^{-1} \Sigma \Sigma^{-1} \mathbf{1}_n}{\left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right)^2} \\
&= \frac{\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n}{\left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right)^2} \\
&= \frac{1}{\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n}
\end{aligned}$$

Another expression of the GMV portfolio is:

$$x_{\text{gmv}} = \sigma^2(x_{\text{gmv}}) \Sigma^{-1} \mathbf{1}_n$$

Minimum variance portfolio

Example 1

The investment universe is made up of 4 assets. The volatility of these assets is respectively equal to 20%, 18%, 16% and 14%, whereas the correlation matrix is given by:

$$\rho = \begin{pmatrix} 1.00 & & & \\ 0.50 & 1.00 & & \\ 0.40 & 0.20 & 1.00 & \\ 0.10 & 0.40 & 0.70 & 1.00 \end{pmatrix}$$

Minimum variance portfolio

We have:

$$\Sigma = \begin{pmatrix} 400.00 & 180.00 & 128.00 & 28.00 \\ 180.00 & 324.00 & 57.60 & 100.80 \\ 128.00 & 57.60 & 256.00 & 156.80 \\ 28.00 & 100.80 & 156.80 & 196.00 \end{pmatrix} \times 10^4$$

It follows that:

$$\Sigma^{-1} = \begin{pmatrix} 54.35 & -37.35 & -50.55 & 51.89 \\ -37.35 & 62.97 & 41.32 & -60.11 \\ -50.55 & 41.32 & 124.77 & -113.85 \\ 51.89 & -60.11 & -113.85 & 165.60 \end{pmatrix}$$

Minimum variance portfolio

We deduce that:

$$\Sigma^{-1}\mathbf{1}_4 = \begin{pmatrix} 18.34 \\ 6.83 \\ 1.69 \\ 43.53 \end{pmatrix}$$

We also have $\mathbf{1}_4^\top \Sigma^{-1}\mathbf{1}_4 = 70.39$, $\sigma^2(x_{\mathrm{gmv}}) = 1/70.39 = 1.4206\%$ and $\sigma(x_{\mathrm{gmv}}) = \sqrt{1.4206\%} = 11.92\%$. Finally, we obtain:

$$x_{\mathrm{gmv}} = \frac{\Sigma^{-1}\mathbf{1}_4}{\mathbf{1}_4^\top \Sigma^{-1}\mathbf{1}_4} = \begin{pmatrix} 26.05\% \\ 9.71\% \\ 2.41\% \\ 61.84\% \end{pmatrix}$$

We verify that $\sum_{i=1}^4 x_{\mathrm{gmv},i} = 100\%$ and $\sqrt{x_{\mathrm{gmv}}^\top \Sigma x_{\mathrm{gmv}}} = 11.92\%$

Minimum variance portfolio

- If we assume that the correlation matrix is constant – $C = \mathcal{C}_n(\rho)$, the covariance matrix is $\Sigma = \sigma\sigma^\top \circ \mathcal{C}_n(\rho)$ with $\mathcal{C}_n(\rho)$ the constant correlation matrix. We deduce that:

$$\Sigma^{-1} = \Gamma \circ \mathcal{C}_n^{-1}(\rho)$$

with $\Gamma_{i,j} = \sigma_i^{-1}\sigma_j^{-1}$ and:

$$\mathcal{C}_n^{-1}(\rho) = \frac{\rho \mathbf{1}_n \mathbf{1}_n^\top - ((n-1)\rho + 1) I_n}{(n-1)\rho^2 - (n-2)\rho - 1}$$

- By using the trace property $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, we can show that:

$$x_{\mathrm{gmv},i} = \frac{-((n-1)\rho + 1)\sigma_i^{-2} + \rho\sum_{j=1}^n (\sigma_i\sigma_j)^{-1}}{\sum_{k=1}^n \left(-((n-1)\rho + 1)\sigma_k^{-2} + \rho\sum_{j=1}^n (\sigma_k\sigma_j)^{-1}\right)}$$

Minimum variance portfolio

- The denominator is the scaling factor such that $\mathbf{1}_n^\top x_{\mathrm{gmv}} = 1$. We deduce that the optimal weights are given by the following relationship:

$$x_{\mathrm{gmv},i} \propto \frac{((n-1)\rho + 1)}{\sigma_i^2} - \frac{\rho}{\sigma_i} \sum_{j=1}^{n} \frac{1}{\sigma_j}$$

Minimum variance portfolio

Here are some special cases:

1. The lower bound of $C_n(\rho)$ is achieved for $\rho = -(n-1)^{-1}$ and we have:

$$\begin{aligned} x_{\mathrm{gmv},i} &\propto -\frac{\rho}{\sigma_i}\sum_{j=1}^{n}\frac{1}{\sigma_j} \\ &\propto \frac{1}{\sigma_i} \end{aligned}$$

   The weights are proportional to the inverse volatilities (GMV = ERC)

2. If the assets are uncorrelated ($\rho = 0$), we obtain:

$$x_i \propto \frac{1}{\sigma_i^2}$$

   The weights are proportional to the inverse variances

Minimum variance portfolio

3. If the assets are perfectly correlated ($\rho = 1$), we have:

$$x_{\mathrm{gmv},i} \propto \frac{1}{\sigma_i}\left(\frac{n}{\sigma_i} - \sum_{j=1}^{n}\frac{1}{\sigma_j}\right)$$

We deduce that:

$$\begin{aligned} x_{\mathrm{gmv},i} \geq 0 \quad & \Leftrightarrow \quad \frac{n}{\sigma_i} - \sum_{j=1}^{n}\frac{1}{\sigma_j} \geq 0 \\ & \Leftrightarrow \quad \sigma_i \leq \left(\frac{1}{n}\sum_{j=1}^{n}\sigma_j^{-1}\right)^{-1} \\ & \Leftrightarrow \quad \sigma_i \leq \bar{H}\left(\sigma_1, \ldots, \sigma_n\right) \end{aligned}$$

where $\bar{H}\left(\sigma_1, \ldots, \sigma_n\right)$ is the harmonic mean of volatilities

Minimum variance portfolio

Example 2

We consider a universe of four assets. Their volatilities are respectively equal to 4%, 6%, 8% and 10%. We assume also that the correlation matrix C is uniform and is equal to $\mathcal{C}_n(\rho)$.

Minimum variance portfolio

Table 49: Global minimum variance portfolios

Asset	ρ = −20%	0%	20%	50%	70%	90%	99%
1	44.35	53.92	65.88	90.65	114.60	149.07	170.07
2	25.25	23.97	22.36	19.04	15.83	11.20	8.38
3	17.32	13.48	8.69	−1.24	−10.84	−24.67	−33.09
4	13.08	8.63	3.07	−8.44	−19.58	−35.61	−45.37
$\sigma(x^\star)$	1.93	2.94	3.52	3.86	3.62	2.52	0.87

Table 50: Long-only minimum variance portfolios

Asset	ρ = −20%	0%	20%	50%	70%	90%	99%
1	44.35	53.92	65.88	85.71	100.00	100.00	100.00
2	25.25	23.97	22.36	14.29	0.00	0.00	0.00
3	17.32	13.48	8.69	0.00	0.00	0.00	0.00
4	13.08	8.63	3.07	0.00	0.00	0.00	0.00
$\sigma(x^\star)$	1.93	2.94	3.52	3.93	4.00	4.00	4.00

Minimum variance portfolio

In practice, we impose no short selling constraints

$\Downarrow$

Smart beta products (funds and indices) corresponds
to long-only minimum variance portfolios

Minimum variance portfolio

Remark

The minimum variance strategy is related to the low beta effect (Black, 1972; Frazzini and Pedersen, 2014) or the low volatility anomaly (Haugen and Baker, 1991).

Minimum variance portfolio

We consider the single-factor model of the CAPM:

$$R_i = \alpha_i + \beta_i R_m + \varepsilon_i$$

We have:

$$\Sigma = \beta\beta^\top \sigma_m^2 + D$$

where:

- $\beta = (\beta_1, \ldots, \beta_n)$ is the vector of betas
- σ_m^2 is the variance of the market portfolio
- $D = \operatorname{diag}\left(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2\right)$ is the diagonal matrix of specific variances

Minimum variance portfolio

Sherman-Morrison-Woodbury formula

Suppose u and v are two $n \times 1$ vectors and A is an invertible $n \times n$ matrix. We can show that:

$$\left(A + uv^\top\right)^{-1} = A^{-1} - \frac{1}{1 + v^\top A^{-1} u} A^{-1} u v^\top A^{-1}$$

Minimum variance portfolio

We have:

$$\Sigma = D + (\sigma_m \beta)(\sigma_m \beta)^\top$$

We apply the Sherman-Morrison-Woodbury with $A = D$ and $u = v = \sigma_m \beta$:

$$\begin{aligned}\Sigma^{-1} &= \left(D + (\sigma_m \beta)(\sigma_m \beta)^\top\right)^{-1} \\ &= D^{-1} - \frac{1}{1 + (\sigma_m \beta)^\top D^{-1} (\sigma_m \beta)} D^{-1} (\sigma_m \beta)(\sigma_m \beta)^\top D^{-1} \\ &= D^{-1} - \frac{\sigma_m^2}{1 + \sigma_m^2 (\beta^\top D^{-1} \beta)} \left(D^{-1}\beta\right)\left(D^{-1}\beta\right)^\top\end{aligned}$$

Minimum variance portfolio

We have:

$$D^{-1}\beta = \tilde{\beta}$$

with $\tilde{\beta}_i = \beta_i / \tilde{\sigma}_i^2$ and:

$$\begin{aligned} \varphi &= \beta^\top D^{-1} \beta \\ &= \tilde{\beta}^\top \beta \\ &= \sum_{i=1}^n \frac{\beta_i^2}{\tilde{\sigma}_i^2} \end{aligned}$$

We obtain:

$$\Sigma^{-1} = D^{-1} - \frac{\sigma_m^2}{1 + \varphi \sigma_m^2} \tilde{\beta} \tilde{\beta}^\top$$

The GMV portfolio is equal to:

$$\begin{aligned} x_{\text{gmv}} &= \sigma^2 \left(x_{\text{gmv}} \right) \Sigma^{-1} \mathbf{1}_n \\ &= \sigma^2 \left(x_{\text{gmv}} \right) \left(D^{-1} \mathbf{1}_n - \frac{\sigma_m^2}{1 + \varphi \sigma_m^2} \tilde{\beta} \tilde{\beta}^\top \mathbf{1}_n \right) \end{aligned}$$

Minimum variance portfolio

It follows that:

$$\begin{aligned} x_{\mathrm{gmv},i} &= \sigma^2\left(x_{\mathrm{gmv}}\right)\left(\frac{1}{\tilde{\sigma}_i^2} - \frac{\sigma_m^2\left(\tilde{\beta}^\top \mathbf{1}_n\right)}{1+\varphi\sigma_m^2}\frac{\beta_i}{\tilde{\sigma}_i^2}\right) \\ &= \frac{\sigma^2\left(x_{\mathrm{gmv}}\right)}{\tilde{\sigma}_i^2}\left(1 - \frac{\beta_i}{\beta^\star}\right) \end{aligned}$$

where:

$$\beta^\star = \frac{1+\varphi\sigma_m^2}{\sigma_m^2\left(\tilde{\beta}^\top \mathbf{1}_n\right)}$$

The minimum variance portfolio is positively exposed to stocks with low beta:

$$\left\{\begin{array}{l} \beta_i < \beta^\star \Rightarrow x_{\mathrm{gmv},i} > 0 \\ \beta_i > \beta^\star \Rightarrow x_{\mathrm{gmv},i} < 0 \end{array}\right.$$

Moreover, the absolute weight is a decreasing function of the idiosyncratic volatility: $\tilde{\sigma}_i \searrow \Rightarrow \left|x_{\mathrm{gmv},i}\right| \nearrow$

Minimum variance portfolio

The previous formula has been found by Scherer (2011). Clarke et al. (2011) have extended it to the long-only minimum variance:

$$x_{\mathrm{mv},i} = \frac{\sigma^2 \left(x_{\mathrm{gmv}}\right)}{\tilde{\sigma}_i^2} \left(1 - \frac{\beta_i}{\beta^\star}\right)$$

where the threshold $\beta^\star$ is defined as follows:

$$\beta^\star = \frac{1 + \sigma_m^2 \sum_{\beta_i < \beta^\star} \tilde{\beta}_i \beta_i}{\sigma_m^2 \sum_{\beta_i < \beta^\star} \tilde{\beta}_i}$$

In this case, if $\beta_i > \beta^\star$, $x_i^\star = 0$

Minimum variance portfolio

Example 3

We consider an investment universe of five assets. Their beta is respectively equal to 0.9, 0.8, 1.2, 0.7 and 1.3 whereas their specific volatility is 4%, 12%, 5%, 8% and 5%. We also assume that the market portfolio volatility is equal to 25%.

Minimum variance portfolio

- In the case of the GMV portfolio, we have $\varphi = 1879.26$ and $\beta^{\star} = 1.0972$
- In the case of the long-only MV portfolio, we have $\varphi = 121.01$ and $\beta^{\star} = 0.8307$

Table 51: Composition of the MV portfolio

Asset	β_i	$\tilde{\beta}_i$	x_i Unconstrained	x_i Long-only
1	0.90	562.50	147.33	0.00
2	0.80	55.56	24.67	9.45
3	1.20	480.00	−49.19	0.00
4	0.70	109.37	74.20	90.55
5	1.30	520.00	−97.01	0.00
Volatility			11.45	19.19

Minimum variance portfolio

In practice, we use a constrained long-only optimization program:

$$
\begin{aligned}
x^{\star} &= \arg\min \frac{1}{2} x^{\top} \Sigma x \\
\text{u.c.} &\quad \left\{
\begin{array}{l}
\mathbf{1}_n^{\top} x = 1 \\
\mathbf{0}_n \leq x \leq \mathbf{1}_n \\
x \in \mathcal{DC}
\end{array}
\right.
\end{aligned}
$$

$\Rightarrow$ we need to impose some diversification constraints ($x \in \mathcal{DC}$) because Markowitz optimization leads to corner solutions that are not diversified

Minimum variance portfolio

Three main approaches:

1. In order to reduce the concentration of a few number of assets, we can use upper bound on the weights:

$$x_i \leq x_i^+$$

For instance, we can set $x_i \leq 5\%$, meaning that the weight of an asset cannot be larger than 5%. We can also impose lower and upper bounds by sector:

$$s_j^- \leq \sum_{i \in \mathcal{S}_j} x_i \leq s_j^+$$

For instance, if we impose that $3\% \leq \sum_{i \in \mathcal{S}_j} x_i \leq 20\%$, this implied that the weight of each sector must be between 3% and 20%.

Minimum variance portfolio

2. We can impose some constraints with respect to the benchmark composition:
$$\frac{b_i}{m} \leq x_i \leq m \cdot b_i$$
where b_i is the weight of asset i in the benchmark (or index) b. For instance, if $m = 2$, the weight of asset i cannot be lower than 50% of its weight in the benchmark. It cannot also be greater than twice of its weight in the benchmark.

3. The third approach consists of imposing a weight diversification based on the Herfindahl index:
$$\mathcal{H}(x) = \sum_{i=1}^{n} x_i$$

Minimum variance portfolio

- The inverse of the Herfindahl index is called the effective number of bets (ENB):
$$\mathcal{N}(x) = \mathcal{H}^{-1}(x)$$
- $\mathcal{N}(x)$ represents the equivalent number of equally-weighted assets. We can impose a sufficient number of effective bets:
$$\mathcal{N}(x) \geq \mathcal{N}_{\min}$$
- During the period 2000-2020, the ENB of the S&P 500 index is between 90 and 130:
$$90 \leq \mathcal{N}(b) \leq 130$$
- During the same period, the ENB of the S&P 500 minimum variance portfolio is between 15 and 30:
$$15 \leq \mathcal{N}(x) \leq 30$$
- We conclude that the S&P 500 minimum variance portfolio is less diversified than the S&P 500 index

Minimum variance portfolio

We can impose:

$$\mathcal{N}(x) \geq m \cdot \mathcal{N}(b)$$

For instance, if $m = 1.5$, the ENB of the S&P 500 minimum variance portfolio will be 50% larger than the ENB of the S&P 500 index
We notice that:

$$\begin{aligned} \mathcal{N}(x) \geq \mathcal{N}_{\min} \quad & \Leftrightarrow \quad \mathcal{H}(x) \leq \mathcal{N}_{\min}^{-1} \\ & \Leftrightarrow \quad x^\top x \leq \mathcal{N}_{\min}^{-1} \end{aligned}$$

The optimization problem becomes:

$$\begin{aligned} x^\star(\lambda) &= \arg\min \frac{1}{2} x^\top \Sigma x + \lambda \left(x^\top x - \mathcal{N}_{\min}^{-1} \right) \\ & \text{u.c.} \quad \left\{ \begin{array}{l} \mathbf{1}_n^\top x = 1 \\ \mathbf{0}_n \leq x \leq \mathbf{1}_n \end{array} \right. \end{aligned}$$

Minimum variance portfolio

We can rewrite the objective function as follows:

$$\mathcal{L}(x;\lambda) = \frac{1}{2}x^\top \Sigma x + \lambda x^\top I_n x = \frac{1}{2}x^\top (\Sigma + 2\lambda I_n) x$$

We obtain a standard minimum variance optimization problem where the covariance matrix is shrunk

Remark

The optimal solution is found by applying the bisection algorithm to the QP problem in order to match the constraint:

$$\mathcal{N}(x^\star(\lambda)) = \mathcal{N}_{\min}$$

An alternative approach is to consider the ADMM algorithm (these numerical problems are studied in Lecture 5)

Most diversified portfolio

Definition

Choueifaty and Coignard (2008) introduce the concept of diversification ratio:

$$\mathcal{DR}(x) = \frac{\sum_{i=1}^n x_i \sigma_i}{\sigma(x)} = \frac{x^\top \sigma}{\sqrt{x^\top \Sigma x}}$$

$\mathcal{DR}(x)$ is the ratio between the weighted average volatility and the portfolio volatility

- The diversification ratio of a portfolio fully invested in one asset is equal to one:

$$\mathcal{DR}(e_i) = 1$$

- In the general case, it is larger than one:

$$\mathcal{DR}(x) \geq 1$$

Most diversified portfolio

The most diversified portfolio (or MDP) is defined as the portfolio which maximizes the diversification ratio:

$$\begin{aligned} x^{\star} &= \arg\max \ln \mathcal{DR}(x) \\ \text{u.c.} & \left\{ \begin{array}{l} \mathbf{1}_n^\top x = 1 \\ \mathbf{0}_n \leq x \leq \mathbf{1}_n \end{array} \right. \end{aligned}$$

Most diversified portfolio

The associated Lagrange function is equal to:

$$\begin{aligned}\mathcal{L}(x;\lambda_0,\lambda) &= \ln\left(\frac{x^\top\sigma}{\sqrt{x^\top\Sigma x}}\right)+\lambda_0\left(\mathbf{1}_n^\top x-1\right)+\lambda^\top\left(x-\mathbf{0}_n\right)\\ &= \ln\left(x^\top\sigma\right)-\frac{1}{2}\ln\left(x^\top\Sigma x\right)+\lambda_0\left(\mathbf{1}_n^\top x-1\right)+\lambda^\top x\end{aligned}$$

The first-order condition is:

$$\frac{\partial\,\mathcal{L}(x;\lambda_0,\lambda)}{\partial\,x}=\frac{\sigma}{x^\top\sigma}-\frac{\Sigma x}{x^\top\Sigma x}+\lambda_0\mathbf{1}_n+\lambda=\mathbf{0}_n$$

whereas the Kuhn-Tucker conditions are:

$$\min\left(\lambda_i,x_i\right)=0\qquad\text{for } i=1,\ldots,n$$

Most diversified portfolio

The constraint $\mathbf{1}_n^\top x = 1$ can always be matched because:

$$\mathcal{DR}\left(\varphi \cdot x\right) = \mathcal{DR}\left(x\right)$$

We deduce that the MDP $x^\star$ satisfies:

$$\frac{\Sigma x^\star}{x^{\star\top} \Sigma x^\star} = \frac{\sigma}{x^{\star\top}\sigma} + \lambda$$

or:

$$\begin{aligned}
\Sigma x^\star &= \frac{\sigma^2\left(x^\star\right)}{x^{\star\top}\sigma}\sigma + \lambda\sigma^2\left(x^\star\right) \\
&= \frac{\sigma\left(x^\star\right)}{\mathcal{DR}\left(x^\star\right)}\sigma + \lambda\sigma^2\left(x^\star\right)
\end{aligned}$$

If the long-only constraint is not imposed, we have $\lambda = \mathbf{0}_n$

Most diversified portfolio

The correlation between a portfolio x and the MDP $x^{\star}$ is given by:

$$\begin{aligned}\rho\left(x, x^{\star}\right) & = \frac{x^{\top}\Sigma x^{\star}}{\sigma\left(x\right)\sigma\left(x^{\star}\right)} \\ & = \frac{1}{\sigma\left(x\right)\mathcal{DR}\left(x^{\star}\right)}x^{\top}\sigma + \frac{\sigma\left(x^{\star}\right)}{\sigma\left(x\right)}x^{\top}\lambda \\ & = \frac{\mathcal{DR}\left(x\right)}{\mathcal{DR}\left(x^{\star}\right)} + \frac{\sigma\left(x^{\star}\right)}{\sigma\left(x\right)}x^{\top}\lambda\end{aligned}$$

Most diversified portfolio

If $x^{\star}$ is the long-only MDP, we obtain (because $\lambda \geq \mathbf{0}_n$ and $x^{\top}\lambda \geq 0$):

$$\rho\left(x, x^{\star}\right) \geq \frac{\mathcal{DR}\left(x\right)}{\mathcal{DR}\left(x^{\star}\right)}$$

whereas we have for the unconstrained MDP:

$$\rho\left(x, x^{\star}\right) = \frac{\mathcal{DR}\left(x\right)}{\mathcal{DR}\left(x^{\star}\right)}$$

The 'core property' of the MDP

> *"The long-only MDP is the long-only portfolio such that the correlation between any other long-only portfolio and itself is greater than or equal to the ratio of their diversification ratios" (Choueifaty et al., 2013)*

Most diversified portfolio

The correlation between Asset i and the MDP is equal to:

$$\begin{aligned} \rho\left(e_i, x^\star\right) &= \frac{\mathcal{DR}\left(e_i\right)}{\mathcal{DR}\left(x^\star\right)} + \frac{\sigma\left(x^\star\right)}{\sigma\left(e_i\right)} e_i^\top \lambda \\ &= \frac{1}{\mathcal{DR}\left(x^\star\right)} + \frac{\sigma\left(x^\star\right)}{\sigma_i} \lambda_i \end{aligned}$$

Most diversified portfolio

Because $\lambda_i = 0$ if $x_i^\star > 0$ and $\lambda_i > 0$ if $x_i^\star = 0$, we deduce that:

$$\rho\left(e_i, x^\star\right) = \frac{1}{\mathcal{DR}\left(x^\star\right)} \quad \text{if} \quad x_i^\star > 0$$

and:

$$\rho\left(e_i, x^\star\right) \geq \frac{1}{\mathcal{DR}\left(x^\star\right)} \quad \text{if} \quad x_i^\star = 0$$

Most diversified portfolio

Another diversification concept

"Any stock not held by the MDP is more correlated to the MDP than any of the stocks that belong to it. Furthermore, all stocks belonging to the MDP have the same correlation to it. [...] This property illustrates that all assets in the universe are effectively represented in the MDP, even if the portfolio does not physically hold them. [...] This is consistent with the notion that the most diversified portfolio is the un-diversifiable portfolio" (Choueifaty et al., 2013)

Most diversified portfolio

Remark

In the case when the long-only constraint is omitted, we have $\rho(e_i, x^\star) = \rho(e_j, x^\star)$ meaning that the correlation with the MDP is the same for all the assets

Most diversified portfolio

Example 4

We consider an investment universe of four assets. Their volatilities are equal to 20%, 10%, 20% and 25%. The correlation of asset returns is given by the following matrix:

$$\rho = \left(\begin{array}{cccc} 1.00 & & & \\ 0.80 & 1.00 & & \\ 0.40 & 0.30 & 1.00 & \\ 0.50 & 0.10 & -0.10 & 1.00 \end{array} \right)$$

Most diversified portfolio

Table 52: Composition of the MDP

Asset	Unconstrained		Long-only	
	$x_i^\star$	$\rho(e_i, x^\star)$	$x_i^\star$	$\rho(e_i, x^\star)$
1	−18.15	61.10	0.00	73.20
2	61.21	61.10	41.70	62.40
3	29.89	61.10	30.71	62.40
4	27.05	61.10	27.60	62.40
$\sigma(x^\star)$	9.31		10.74	
$\mathcal{DR}(x^\star)$	1.64		1.60	

Most diversified portfolio

Assumption $\mathcal{H}_0$: all the assets have the same Sharpe ratio

$$\frac{\mu_i - r}{\sigma_i} = s$$

Under $\mathcal{H}_0$, the diversification ratio of portfolio x is proportional to its Sharpe ratio:

$$\begin{aligned}
\mathcal{DR}(x) &= \frac{1}{s} \frac{\sum_{i=1}^n x_i (\mu_i - r)}{\sigma(x)} \\
&= \frac{1}{s} \frac{x^\top \mu - r}{\sigma(x)} \\
&= \frac{1}{s} \cdot \mathrm{SR}(x \mid r)
\end{aligned}$$

Most diversified portfolio

Optimality of the MDP

Under $\mathcal{H}_0$, maximizing the diversification ratio is then equivalent to maximizing the Sharpe ratio:

$$\text{MDP} = \text{MSR}$$

Most diversified portfolio

In the CAPM framework, Clarke *et al.* (2013) showed that:

$$x_i^\star = \mathcal{DR}\left(x^\star\right) \frac{\sigma_i \sigma\left(x^\star\right)}{\tilde{\sigma}_i^2}\left(1 - \frac{\rho_{i,m}}{\rho^\star}\right)$$

where $\sigma_i = \sqrt{\beta_i^2 \sigma_m^2 + \tilde{\sigma}_i^2}$ is the volatility of asset i, $\rho_{i,m} = \beta_i \sigma_m / \sigma_i$ is the correlation between asset i and the market portfolio and $\rho^\star$ is the threshold correlation given by this formula:

$$\rho^\star = \left(1 + \sum_{i=1}^n \frac{\rho_{i,m}^2}{1 - \rho_{i,m}^2}\right) \Bigg/ \left(\sum_{i=1}^n \frac{\rho_{i,m}}{1 - \rho_{i,m}^2}\right)$$

The weights are then strictly positive if $\rho_{i,m} < \rho^\star$

Most diversified portfolio

The MDP tends to be less concentrated than the MV portfolio because:

$$
\begin{aligned}
x_{\mathrm{mv},i} &= \frac{1}{\tilde{\sigma}_i^2} \times \cdots \\
x_{\mathrm{mdp},i} &= \frac{\sigma_i}{\tilde{\sigma}_i^2} \times \cdots \approx \frac{1}{\tilde{\sigma}_i} \times \cdots + \cdots
\end{aligned}
$$

ERC portfolio

In Lecture 2, we have seen that the ERC portfolio corresponds to the portfolio such that the risk contribution from each stock is made equal

The main advantages of the ERC allocation are the following:

1. It defines a portfolio that is well diversified in terms of risk and weights
2. Like the three previous risk-based methods, it does not depend on any expected returns hypothesis
3. It is less sensitive to small changes in the covariance matrix than MV or MDP portfolios (Demey *et al.*, 2010)

ERC portfolio

In the CAPM framework, Clarke *et al.* (2013) showed:

$$x_i^\star = \frac{\sigma^2(x^\star)}{\tilde{\sigma}_i^2}\left(\sqrt{\frac{\beta_i^2}{\beta^{\star 2}} + \frac{\tilde{\sigma}_i^2}{n\sigma^2(x^\star)}} - \frac{\beta_i}{\beta^\star}\right)$$

where:

$$\beta^\star = \frac{2\sigma^2(x^\star)}{\beta(x^\star)\,\sigma_m^2}$$

It follows that:

$$\lim_{n\rightarrow\infty} x_{\mathrm{erc}} = x_{\mathrm{ew}}$$

Comparison of the 4 Methods

Equally-weighted (EW)

- Weights are equal
- Easy to understand
- Contrarian strategy with a take-profit scheme
- The least concentrated in terms of weights
- Do not depend on risks

Most Diversified Portfolio (MDP)

- Also known as the Max Sharpe Ratio (MSR) portfolio of EDHEC
- Based on the assumption that sharpe ratio is equal for all stocks
- It is the tangency portfolio if the previous assumption is verified
- Sensitive to the covariance matrix

Minimum variance (MV)

- Low volatility portfolio
- The only optimal portfolio not depending on expected returns assumptions
- Good out of sample performance
- Concentrated portfolios
- Sensitive to the covariance matrix

Equal Risk Contribution (ERC)

- Risk contributions are equal
- Highly diversified portfolios
- Less sensitive to the covariance matrix (than the MV and MDP portfolios)
- Not efficient for universe with a large number of stocks (equivalent to the EW portfolio)

Some properties

In terms of bets

$$\begin{array}{ll} \exists i : w_i = 0 & \text{(MV - MDP)} \\ \forall i : w_i \neq 0 & \text{(EW - ERC)} \end{array}$$

In terms of risk factors

$$\begin{array}{rcll} x_i & = & x_j & \text{(EW)} \\ \frac{\partial\, \sigma(x)}{\partial\, x_i} & = & \frac{\partial\, \sigma(x)}{\partial\, x_j} & \text{(MV)} \\ x_i \cdot \frac{\partial\, \sigma(x)}{\partial\, x_i} & = & x_j \cdot \frac{\partial\, \sigma(x)}{\partial\, x_j} & \text{(ERC)} \\ \frac{1}{\sigma_i} \cdot \frac{\partial\, \sigma(x)}{\partial\, x_i} & = & \frac{1}{\sigma_j} \cdot \frac{\partial\, \sigma(x)}{\partial\, x_j} & \text{(MDP)} \end{array}$$

Some properties

Proof for the MDP portfolio

For the unconstrained MDP portfolio, we recall that the first-order condition is given by:

$$\frac{\partial \mathcal{L}(x; \lambda_0, \lambda)}{\partial x_i} = \frac{\sigma_i}{x^\top \sigma} - \frac{(\Sigma x)_i}{x^\top \Sigma x} = 0$$

The scaled marginal volatility is then equal to the inverse of the diversification ratio of the MDP:

$$\begin{aligned} \frac{1}{\sigma_i} \cdot \frac{\partial \sigma(x)}{\partial x_i} &= \frac{1}{\sigma_i} \cdot \frac{(\Sigma x)_i}{\sqrt{x^\top \Sigma x}} \\ &= \frac{\sigma(x)}{\sigma_i} \cdot \frac{(\Sigma x)_i}{x^\top \Sigma x} \\ &= \frac{\sigma(x)}{x^\top \sigma} = \frac{1}{\mathcal{DR}(x)} \end{aligned}$$

Application to the Eurostoxx 50 index

Table 53: Composition in % (January 2010)

	CW	MV	ERC	MDP	1/n	MV 10%	MDP 10%	MV 5%	MDP 5%
TOTAL	6.1		2.1		2			5.0	
BANCO SANTANDER	5.8		1.3		2				
TELEFONICA SA	5.0	31.2	3.5		2	10.0		5.0	5.0
SANOFI-AVENTIS	3.6	12.1	4.5	15.5	2	10.0	10.0	5.0	5.0
E.ON AG	3.6		2.1		2				1.4
BNP PARIBAS	3.4		1.1		2				
SIEMENS AG	3.2		1.5		2				
BBVA(BILB-VIZ-ARG)	2.9		1.4		2				
BAYER AG	2.9		2.6	3.7	2	2.2	5.0	5.0	5.0
ENI	2.7		2.1		2				
GDF SUEZ	2.5		2.6	4.5	2		5.4	5.0	5.0
BASF SE	2.5		1.5		2				
ALLIANZ SE	2.4		1.4		2				
UNICREDIT SPA	2.3		1.1		2				
SOC GENERALE	2.2		1.2	3.9	2		3.7		5.0
UNILEVER NV	2.2	11.4	3.7	10.8	2	10.0	10.0	5.0	5.0
FRANCE TELECOM	2.1	14.9	4.1	10.2	2	10.0	10.0	5.0	5.0
NOKIA OYJ	2.1		1.8	4.5	2		4.8		5.0
DAIMLER AG	2.1		1.3		2				
DEUTSCHE BANK AG	1.9		1.0		2				
DEUTSCHE TELEKOM	1.9		3.2	2.6	2	5.7	3.7	5.0	5.0
INTESA SANPAOLO	1.9		1.3		2				
AXA	1.8		1.0		2				
ARCELORMITTAL	1.8		1.0		2				
SAP AG	1.8	21.0	3.4	11.2	2	10.0	10.0	5.0	5.0

	CW	MV	ERC	MDP	1/n	MV 10%	MDP 10%	MV 5%	MDP 5%
RWE AG (NEU)	1.7	2.7	2.7		2	7.0		5.0	
ING GROEP NV	1.6		0.8	0.4	2				
DANONE	1.6	1.9	3.4	1.8	2	8.7	3.3	5.0	5.0
IBERDROLA SA	1.6		2.5		2	5.1		5.0	1.2
ENEL	1.6		2.1		2			5.0	2.9
VIVENDI SA	1.6	2.8	3.1	4.5	2	10.0	5.9	5.0	5.0
ANHEUSER-BUSCH INB	1.6	0.2	2.7	10.9	2	2.1	10.0	5.0	5.0
ASSIC GENERALI SPA	1.6		1.8		2				
AIR LIQUIDE(L')	1.4		2.1		2			5.0	
MUENCHENER RUECKVE	1.3		2.1	2.1	2		3.1	5.0	5.0
SCHNEIDER ELECTRIC	1.3		1.5		2				
CARREFOUR	1.3	1.0	2.7	1.3	2	3.7	2.5	5.0	5.0
VINCI	1.3		1.6		2				
LVMH MOET HENNESSY	1.2		1.8		2				
PHILIPS ELEC(KON)	1.2		1.4		2				
L'OREAL	1.1	0.8	2.8		2	5.5		5.0	5.0
CIE DE ST-GOBAIN	1.0		1.1		2				
REPSOL YPF SA	0.9		2.0		2			5.0	
CRH	0.8		1.7	5.1	2		5.2		5.0
CREDIT AGRICOLE SA	0.8		1.1		2				
DEUTSCHE BOERSE AG	0.7		1.5		2				1.9
TELECOM ITALIA SPA	0.7		2.0		2				2.5
ALSTOM	0.6		1.5		2				
AEGON NV	0.4		0.7		2				
VOLKSWAGEN AG	0.2		1.8	7.1	2		7.4		5.0
Total of components	50	11	50	17	50	14	16	20	23

Some examples

To compare the risk-based methods, we report:

- The weights x_i in %
- The relative risk contributions $\mathcal{RC}_i$ in %
- The weight concentration $\mathcal{H}^{\star}(x)$ in % and the risk concentration $\mathcal{H}^{\star}(\mathcal{RC})$ in % where $\mathcal{H}^{\star}$ is the modified Herfindahl index[9]
- The portfolio volatility $\sigma(x)$ in %
- The diversification ratio $\mathcal{DR}(x)$

[9] We have:

$$\mathcal{H}^{\star}(\pi) = \frac{n\mathcal{H}(\pi) - 1}{n - 1} \in [0, 1]$$

Some examples

Example 5

We consider an investment universe with four assets. We assume that the volatility σ_i is the same and equal to 20% for all four assets. The correlation matrix C is equal to:

$$C = \begin{pmatrix} 100\% & & & \\ 80\% & 100\% & & \\ 0\% & 0\% & 100\% & \\ 0\% & 0\% & -50\% & 100\% \end{pmatrix}$$

Some examples

Table 54: Weights and risk contributions (Example 5)

Asset	EW		MV		MDP		ERC	
	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$
1	25.00	4.20	10.87	0.96	10.87	0.96	17.26	2.32
2	25.00	4.20	10.87	0.96	10.87	0.96	17.26	2.32
3	25.00	1.17	39.13	3.46	39.13	3.46	32.74	2.32
4	25.00	1.17	39.13	3.46	39.13	3.46	32.74	2.32
$\mathcal{H}^{\star}(x)$	0.00		10.65		10.65		3.20	
$\sigma(x)$	10.72		8.85		8.85		9.26	
$\mathcal{DR}(x)$	1.87		2.26		2.26		2.16	
$\mathcal{H}^{\star}(\mathcal{RC})$	10.65		10.65		10.65		0.00	

Some examples

Example 6

We modify the previous example by introducing differences in volatilities. They are 10%, 20%, 30% and 40% respectively. The correlation matrix remains the same as in Example 5.

Some examples

Table 55: Weights and risk contributions (Example 6)

Asset	EW		MV		MDP		ERC	
	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$
1	25.00	1.41	74.48	6.43	27.78	1.23	38.36	2.57
2	25.00	3.04	0.00	0.00	13.89	1.23	19.18	2.57
3	25.00	1.63	15.17	1.31	33.33	4.42	24.26	2.57
4	25.00	5.43	10.34	0.89	25.00	4.42	18.20	2.57
$\mathcal{H}^\star(x)$	0.00		45.13		2.68		3.46	
$\sigma(x)$	11.51		8.63		11.30		10.29	
$\mathcal{DR}(x)$	2.17		1.87		2.26		2.16	
$\mathcal{H}^\star(\mathcal{RC})$	10.31		45.13		10.65		0.00	

Some examples

Example 7

We now reverse the volatilities of Example 6. They are now equal to 40%, 30%, 20% and 10%.

Some examples

Table 56: Weights and risk contributions (Example 7)

Asset	EW		MV		MDP		ERC	
	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$
1	25.00	9.32	0.00	0.00	4.18	0.74	7.29	1.96
2	25.00	6.77	4.55	0.29	5.57	0.74	9.72	1.96
3	25.00	1.09	27.27	1.74	30.08	2.66	27.66	1.96
4	25.00	0.00	68.18	4.36	60.17	2.66	55.33	1.96
$\mathcal{H}^{\star}(x)$	0.00		38.84		27.65		19.65	
$\sigma(x)$	17.18		6.40		6.80		7.82	
$\mathcal{DR}(x)$	1.46		2.13		2.26		2.16	
$\mathcal{H}^{\star}(\mathcal{RC})$	27.13		38.84		10.65		0.00	

Some examples

Example 8

We consider an investment universe of four assets. The volatility is respectively equal to 15%, 30%, 45% and 60% whereas the correlation matrix C is equal to:

$$C = \begin{pmatrix} 100\% & & & \\ 10\% & 100\% & & \\ 30\% & 30\% & 100\% & \\ 40\% & 20\% & -50\% & 100\% \end{pmatrix}$$

Some examples

Table 57: Weights and risk contributions (Example 8)

Asset	EW		MV		MDP		ERC	
	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$
1	25.00	2.52	82.61	11.50	0.00	0.00	40.53	4.52
2	25.00	5.19	17.39	2.42	0.00	0.00	22.46	4.52
3	25.00	3.89	0.00	0.00	57.14	12.86	21.12	4.52
4	25.00	9.01	0.00	0.00	42.86	12.86	15.88	4.52
$\mathcal{H}^{\star}(x)$	0.00		61.69		34.69		4.61	
$\sigma(x)$	20.61		13.92		25.71		18.06	
$\mathcal{DR}(x)$	1.82		1.27		2.00		1.76	
$\mathcal{H}^{\star}(\mathcal{RC})$	7.33		61.69		33.33		0.00	

Some examples

Example 9

Now we consider an example with six assets. The volatilities are 25%, 20%, 15%, 18%, 30% and 20% respectively. We use the following correlation matrix:

$$C = \begin{pmatrix} 100\% & & & & & \\ 20\% & 100\% & & & & \\ 60\% & 60\% & 100\% & & & \\ 60\% & 60\% & 60\% & 100\% & & \\ 60\% & 60\% & 60\% & 60\% & 100\% & \\ 60\% & 60\% & 60\% & 60\% & 60\% & 100\% \end{pmatrix}$$

Some examples

Table 58: Weights and risk contributions (Example 9)

Asset	EW		MV		MDP		ERC	
	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$	x_i	$\mathcal{RC}_i$
1	16.67	3.19	0.00	0.00	44.44	8.61	14.51	2.72
2	16.67	2.42	6.11	0.88	55.56	8.61	18.14	2.72
3	16.67	2.01	65.16	9.33	0.00	0.00	21.84	2.72
4	16.67	2.45	22.62	3.24	0.00	0.00	18.20	2.72
5	16.67	4.32	0.00	0.00	0.00	0.00	10.92	2.72
6	16.67	2.75	6.11	0.88	0.00	0.00	16.38	2.72
$\mathcal{H}^{\star}(x)$	0.00		37.99		40.74		0.83	
$\sigma(x)$	17.14		14.33		17.21		16.31	
$\mathcal{DR}(x)$	1.24		1.14		1.29		1.25	
$\mathcal{H}^{\star}(\mathcal{RC})$	1.36		37.99		40.00		0.00	

Some examples

Example 10

To illustrate how the MV and MDP portfolios are sensitive to specific risks, we consider a universe of n assets with volatility equal to 20%. The structure of the correlation matrix is the following:

$$C = \begin{pmatrix} 100\% & & & & \\ \rho_{1,2} & 100\% & & & \\ 0 & \rho & 100\% & & \\ \vdots & \vdots & \ddots & 100\% & \\ 0 & \rho & \cdots & \rho & 100\% \end{pmatrix}$$

Some examples

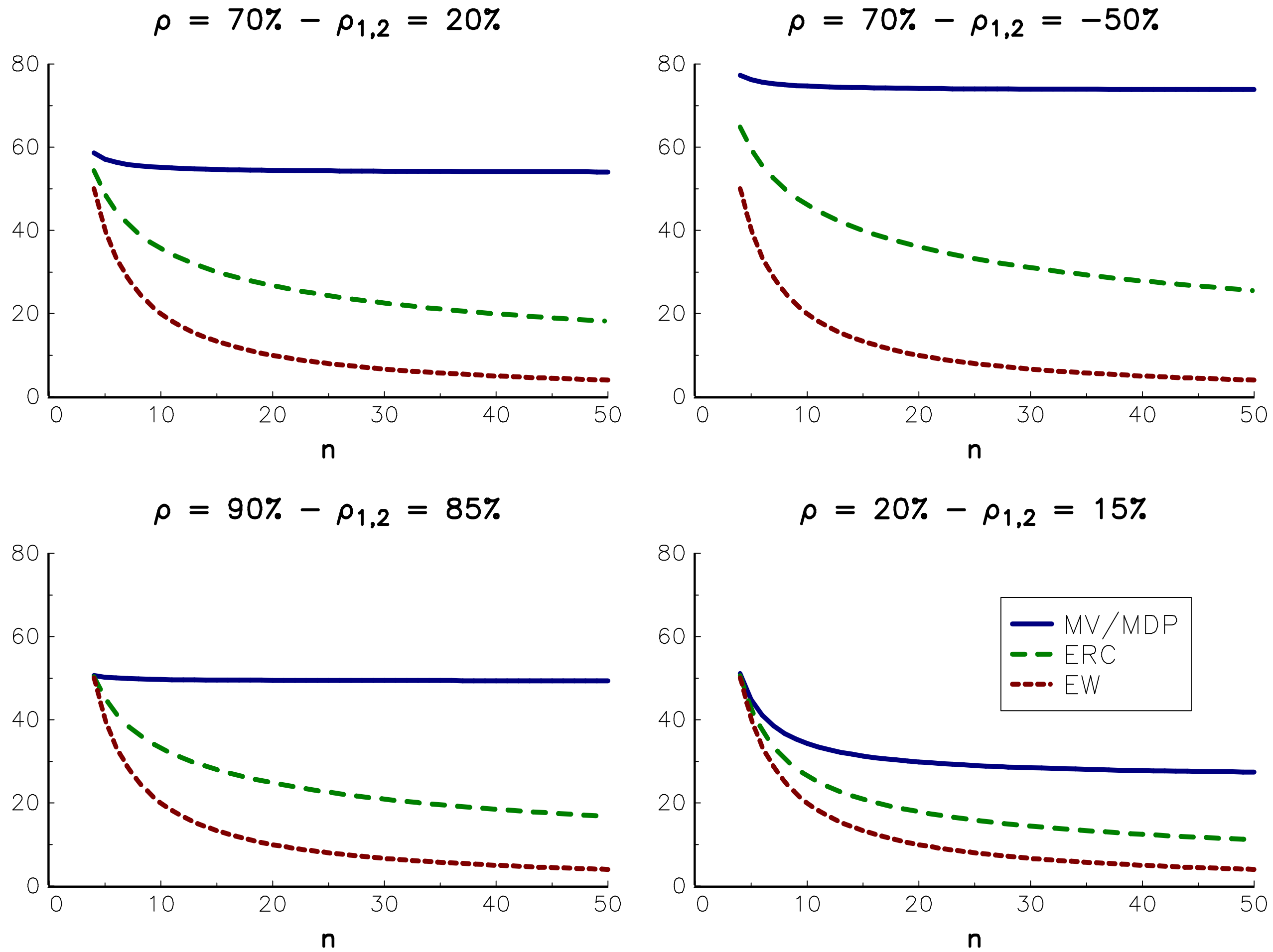


Figure 46: Weight of the first two assets in AW portfolios (Example 10)

Some examples

Example 11

We assume that asset returns follow the one-factor CAPM model. The idiosyncratic volatility $\tilde{\sigma}_i$ is set to 5% for all the assets whereas the volatility of the market portfolio σ_m is equal to 25%.

Some examples

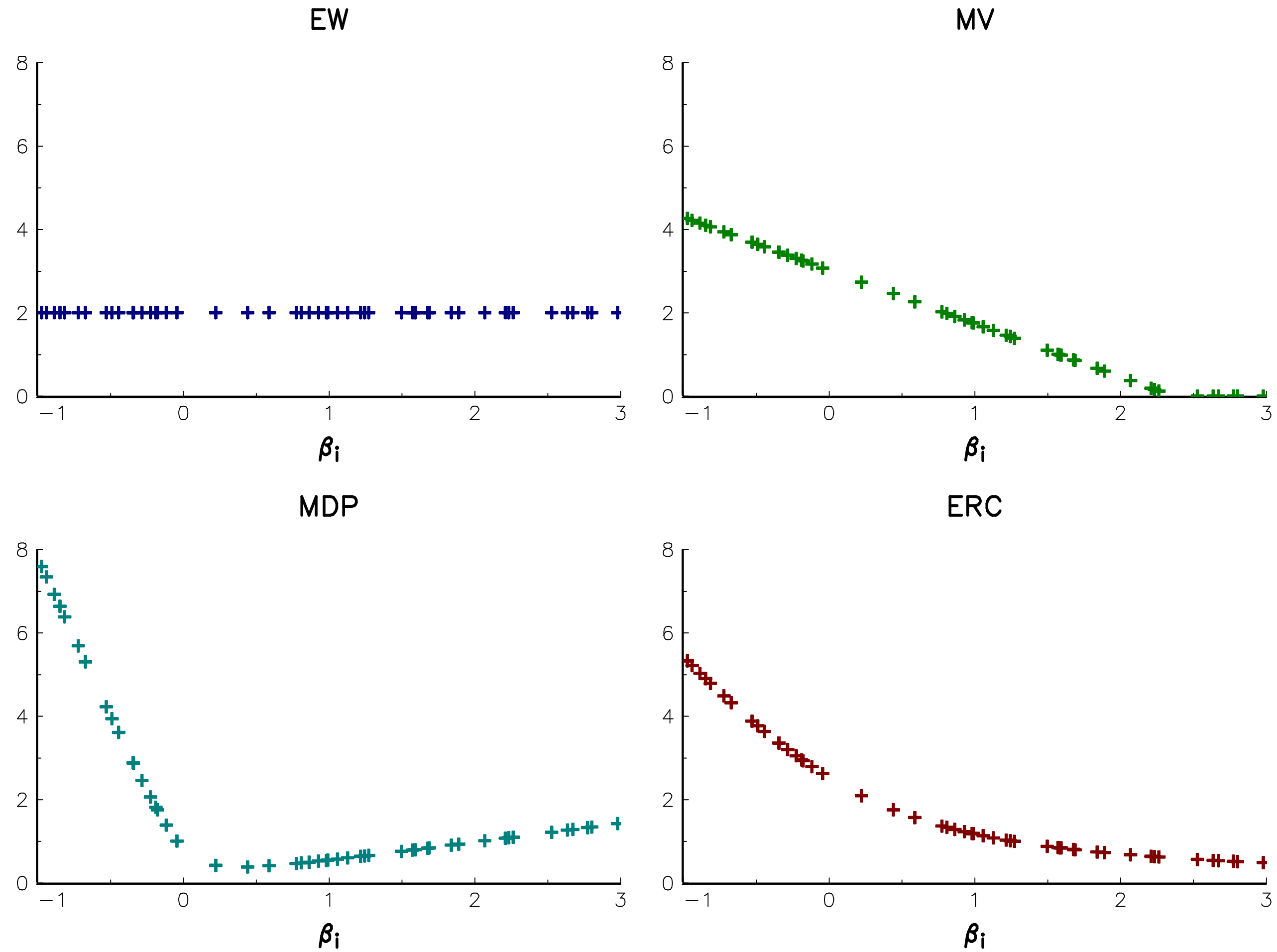


Figure 47: Weight with respect to the asset beta β_i (Example 11)

Smart beta products

- MSCI Equal Weighted Indexes (EW)
  www.msci.com/msci-equal-weighted-indexes
- S&P 500 Equal Weight Index (EW)
  www.spglobal.com/spdji/en/indices/equity/sp-500-equal-weight-index
- FTSE UK Equally Weighted Index Series (EW)
  www.ftserussell.com/products/indices/equally-weighted
- FTSE Global Minimum Variance Index Series (MV)
  www.ftserussell.com/products/indices/min-variance
- MSCI Minimum Volatility Indexes (MV)
  www.msci.com/msci-minimum-volatility-indexes
- S&P 500 Minimum Volatility Index (MV)
  www.spglobal.com/spdji/en/indices/strategy/sp-500-minimum-volatility-index
- FTSE Global Equal Risk Contribution Index Series (ERC)
  www.ftserussell.com/products/indices/erc
- TOBAM MaxDiv Index Series (MDP)
  www.tobam.fr/maximum-diversification-indexes

Smart beta products

Largest ETF issuers in Europe

1. iShares (BlackRock)
2. Xtrackers (DWS)
3. Lyxor ETF
4. UBS ETF
5. Amundi ETF

Largest ETF issuers in US

1. iShares (BlackRock)
2. SPDR (State Street)
3. Vanguard
4. Invesco PowerShares
5. First Trust

- Specialized smart beta ETF issuers: Wisdom Tree (US), Ossiam (Europe), Research affiliates (US), etc.
- Smart beta fund managers in Europe: Amundi, Ossiam, Quoniam, Robeco, Seeyond, Tobam, Unigestion, etc.
- ETFs, mutual funds, mandates

The case of bonds

Two main problems:

1. Benchmarks = debt-weighted indexation (the weights are based on the notional amount of the debt)
2. Fund management driven by the search of yield with little consideration for **credit risk** (carry position $\neq$ arbitrage position)

$\Rightarrow$ Time to rethink bond indexes? (Toloui, 2010)

The case of bonds

Two main problems:

1. Benchmarks = debt-weighted indexation (the weights are based on the notional amount of the debt)
2. Fund management driven by the search of yield with little consideration for **credit risk** (carry position $\neq$ arbitrage position)

$\Rightarrow$ Time to rethink bond indexes? (Toloui, 2010)

Bond indexation

Debt weighting

It is defined by:

$$x_i = \frac{\mathrm{DEBT}_i}{\sum_{i=1}^n \mathrm{DEBT}_i}$$

GDP weighting

It is defined by:

$$x_i = \frac{\mathrm{GDP}_i}{\sum_{i=1}^n \mathrm{GDP}_i}$$

Risk budgeting

It is defined by:

$$b_i = \frac{\mathrm{DEBT}_i}{\sum_{i=1}^n \mathrm{DEBT}_i}$$

or:

$$b_i = \frac{\mathrm{GDP}_i}{\sum_{i=1}^n \mathrm{GDP}_i}$$

$\Rightarrow$ The offering is very small compared to equity indices because of the liquidity issues (see Roncalli (2013), Chapter 4 for more details)

From CAPM to factor investing

How to define risk factors?

Risk factors are common factors that explain the cross-section variance of expected returns

- 1964: Market or MKT (or BETA) factor
- 1972: Low beta or BAB factor
- 1981: Size or SMB factor
- 1985: Value or HML factor
- 1991: Low volatility or VOL factor
- 1993: Momentum or WML factor
- 2000: Quality or QMJ factor

Systematic risk factors $\neq$ Idiosyncratic risk factors

Beta(s) $\neq$ Alpha(s)

Alpha or beta?

At the security level, there is a lot of idiosyncratic risk or alpha[10]:

	Common Risk	Idiosyncratic Risk
GOOGLE	47%	53%
NETFLIX	24%	76%
MASTERCARD	50%	50%
NOKIA	32%	68%
TOTAL	89%	11%
AIRBUS	56%	44%

Carhart's model with 4 factors, 2010-2014
Source: Roncalli (2017)

[10]The linear regression is:

$$R_i = \alpha_i + \sum_{j=1}^{n_{\mathcal{F}}} \beta_i^j \mathcal{F}_j + \varepsilon_i$$

In our case, we measure the alpha as $1 - \mathfrak{R}_i^2$ where:

$$\mathfrak{R}_i^2 = 1 - \frac{\sigma^2(\varepsilon_i)}{\sigma^2(R_i)}$$

The concept of alpha

- Jensen (1968) – **How to measure the performance of active fund managers?**

$$R_t^F = \alpha + \beta R_t^{MKT} + \varepsilon_t$$

Fund	Return	Rank	Beta	Alpha	Rank
A	12%	Best	1.0	−2%	Worst
B	11%	Worst	0.5	4%	Best

Market return = 14%

$\Rightarrow \bar{\alpha} = -\text{fees}$

- It is the beginning of passive management:
  - John McQuown (Wells Fargo Bank, 1971)
  - Rex Sinquefield (American National Bank, 1973)

Active management and performance persistence

- Hendricks *et al.* (1993) – **Hot Hands in Mutual Funds**

$$\operatorname{cov}\left(\alpha_t^{\text{Jensen}}, \alpha_{t-1}^{\text{Jensen}}\right) > 0$$

where:

$$\alpha_t^{\text{Jensen}} = R_t^F - \beta^{\mathrm{MKT}} R_t^{MKT}$$

$\Rightarrow$ The persistence of the performance of active management is due to the **persistence of the alpha**

Risk factors and active management

- Grinblatt *et al.* (1995) – **Momentum investors versus Value investors**

  "*77% of mutual funds are momentum investors*"

- Carhart (1997):

$$\left\{ \begin{array}{l} \operatorname{cov}\left(\alpha_t^{\text{Jensen}}, \alpha_{t-1}^{\text{Jensen}}\right) > 0 \\ \operatorname{cov}\left(\alpha_t^{\text{Carhart}}, \alpha_{t-1}^{\text{Carhart}}\right) = 0 \end{array} \right.$$

  where:

$$\alpha_t^{\text{Carhart}} = R_t^F - \beta^{\text{MKT}} R_t^{\text{MKT}} - \beta^{\text{SMB}} R_t^{\text{SMB}} - \beta^{\text{HML}} R_t^{\text{HML}} - \beta^{\text{WML}} R_t^{\text{WML}}$$

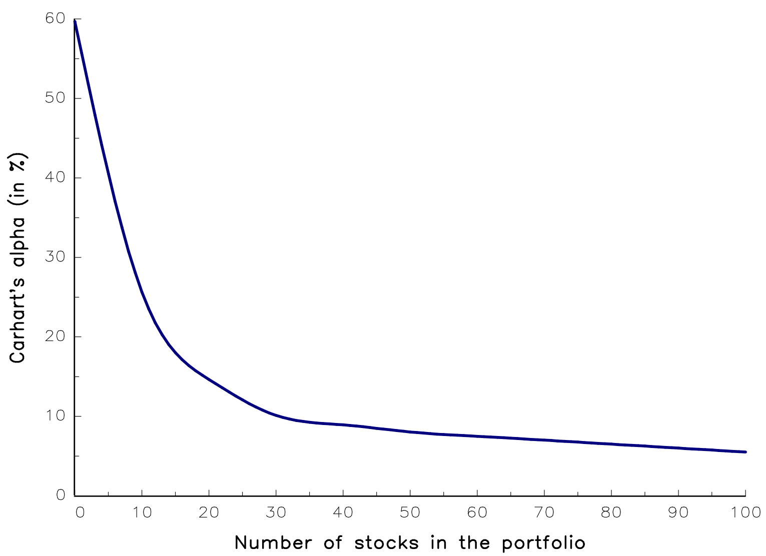
$\Rightarrow$ The (short-term) persistence of the performance of active management is due to the (short-term) **persistence of the performance of risk factors**

Diversification and alpha

David Swensen's rule for effective stock picking

Concentrated portfolio ⇒ No more than 20 bets?

Figure 48: Carhart's alpha decreases with the number of holding assets

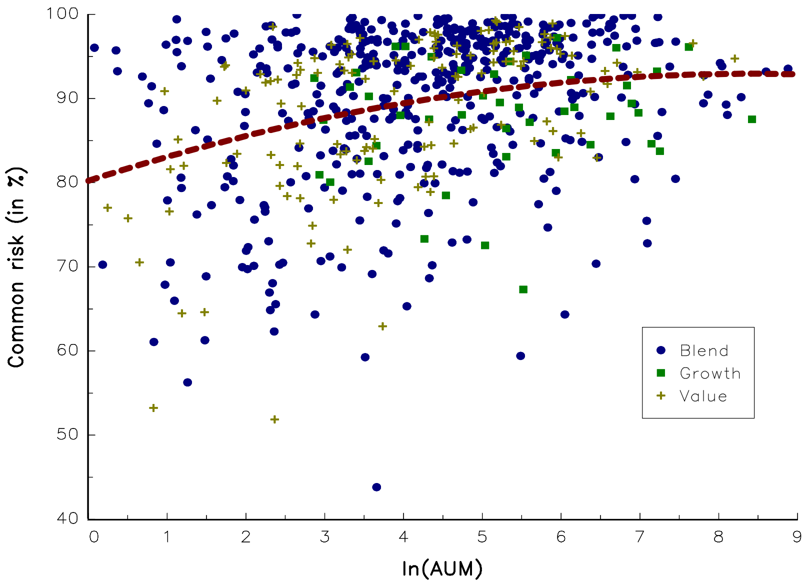


US equity markets, 2000-2014
Source: Roncalli (2017)

"*If you can identify six wonderful businesses, that is all the diversification you need. And you will make a lot of money. And I can guarantee that going into the seventh one instead of putting more money into your first one is going to be a terrible mistake. Very few people have gotten rich on their seventh best idea.*" (Warren Buffett, University of Florida, 1998).

Diversification and alpha

Figure 49: What proportion of return variance is explained by the 4-factor model?



Morningstar database, 880 mutual funds, European equities
Carhart's model with 4 factors, 2010-2014
Source: Roncalli (2017)

How many bets are there in large portfolios of institutional investors?

1986 Less than 10% of institutional portfolio return is explained by security picking and market timing (Brinson *et al.*, 1986)

2009 Professors' Report on the Norwegian GPFG: Risk factors represent 99.1% of the fund return variation (Ang *et al.*, 2009)

Risk factors versus alpha

What lessons can we draw from this?

Idiosyncratic risks and specific bets disappear in (large) diversified portfolios. Performance of institutional investors is then exposed to (common) risk factors.

Alpha is not scalable, but risk factors are scalable

$\Rightarrow$ Risk factors are the only bets that are compatible with diversification

Alpha

- Concentration
- Scarce?

$\neq$

Beta(s)

- Diversification
- Easy access?

Factor investing and active management

Misconception about active management

- Active management $= \boldsymbol{\alpha}$
- Passive management $= \boldsymbol{\beta}$

In this framework, passive management encompasses cap-weighted indexation, risk-based indexation and factor investing because these management styles do not pretend to create alpha

Factor investing and active management

"The question is when is active management good? The answer is never"

Eugene Fama, Morningstar ETF conference, September 2014

"So people say, 'I'm not going to try to beat the market. The market is all-knowing.' But how in the world can the market be all-knowing, if nobody is trying – well, not as many people – are trying to beat it?"

Robert Shiller, CNBC, November 2017

Factor investing and active management

- Discretionary active management $\Rightarrow$ specific/idiosyncratic risks & rule-based management $\Rightarrow$ factor investing and systematic risks?

- Are common risk factors exogenous or endogenous?
- Do risk factors exist without active management?

Risk factors first, active management second

or

Active management first, risk factors second

- Factor investing needs active investing
- Imagine a world without active managers, stock pickers, hedge funds, etc.

$\Rightarrow$ **Should active management be reduced to alpha management?**

Factor investing and active management

- Market risk factor = average of active management
- Low beta/low volatility strategies begin to be implemented in 2003-2004 (after the dot.com crisis)
- Quality strategies begin to be implemented in 2009-2010 (after the GFC crisis)

Alpha strategy $\Rightarrow$ **Risk Factor** (or a beta strategy)

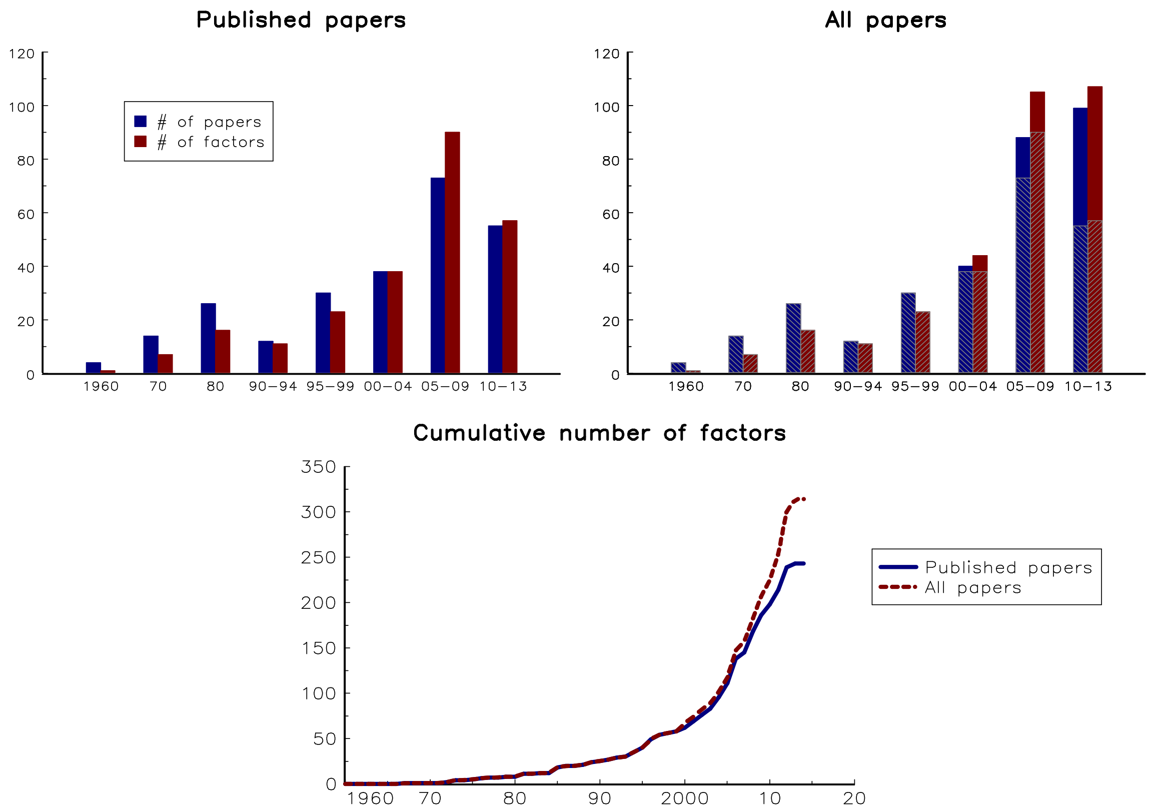
Factor investing and active management

α or β?

"[...] When an alpha strategy is massively invested, it has an enough impact on the structure of asset prices to become a risk factor.
[...] Indeed, an alpha strategy becomes a common market risk factor once it represents a significant part of investment portfolios and explains the cross-section dispersion of asset returns" (Roncalli, 2020)

The factor zoo

Figure 50: Harvey *et al.* (2016)



"*Now we have a zoo of new factors*" (Cochrane, 2011).

Factors, factors everywhere

"Standard predictive regressions fail to reject the hypothesis that the party of the U.S. President, the weather in Manhattan, global warming, El Niño, sunspots, or the conjunctions of the planets, are significantly related to anomaly performance. These results are striking, and quite surprising. In fact, some readers may be inclined to reject some of this paper's conclusions solely on the grounds of plausibility. I urge readers to consider this option carefully, however, as doing do so entails rejecting the standard methodology on which the return predictability literature is built."(Novy-Marx, 2014).

⇒ MKT, SMB, HML, WML, STR, LTR, VOL, IVOL, BAB, QMJ, LIQ, TERM, CARRY, DIV, JAN, CDS, GDP, INF, etc.

The alpha puzzle (Cochrane, 2011)

- Chaos

$$\mathbb{E}[R_i] - R_f = \boxed{\alpha_i}$$

- Sharpe (1964)

$$\mathbb{E}[R_i] - R_f = \beta_i^m (\mathbb{E}[R_m] - R_f)$$

- Chaos again

$$\mathbb{E}[R_i] - R_f = \boxed{\alpha_i} + \beta_i^m (\mathbb{E}[R_m] - R_f)$$

- Fama and French (1992)

$$\mathbb{E}[R_i] - R_f = \beta_i^m (\mathbb{E}[R_m] - R_f) + \beta_i^{smb}\mathbb{E}[R_{smb}] + \beta_i^{hml}\mathbb{E}[R_{hml}]$$

This is not the end of the story...

The alpha puzzle (Cochrane, 2011)

It's just the beginning!

- Chaos again

$$\mathbb{E}[R_i] - R_f = \boxed{\alpha_i} + \beta_i^m (\mathbb{E}[R_m] - R_f) + \beta_i^{smb}\mathbb{E}[R_{smb}] + \beta_i^{hml}\mathbb{E}[R_{hml}]$$

- Carhart (1997)

$$\mathbb{E}[R_i] - R_f = \beta_i^m (\mathbb{E}[R_m] - R_f) + \beta_i^{smb}\mathbb{E}[R_{smb}] + \beta_i^{hml}\mathbb{E}[R_{hml}] + \beta_i^{wml}\mathbb{E}[R_{wml}]$$

- Chaos again

$$\begin{aligned}\mathbb{E}[R_i] - R_f &= \boxed{\alpha_i} + \beta_i^m (\mathbb{E}[R_m] - R_f) + \beta_i^{smb}\mathbb{E}[R_{smb}] + \\ &\quad \beta_i^{hml}\mathbb{E}[R_{hml}] + \beta_i^{wml}\mathbb{E}[R_{wml}]\end{aligned}$$

- Etc.

How can alpha always come back?

The alpha puzzle (Cochrane, 2011)

1. **Because academic backtesting is not the real life**

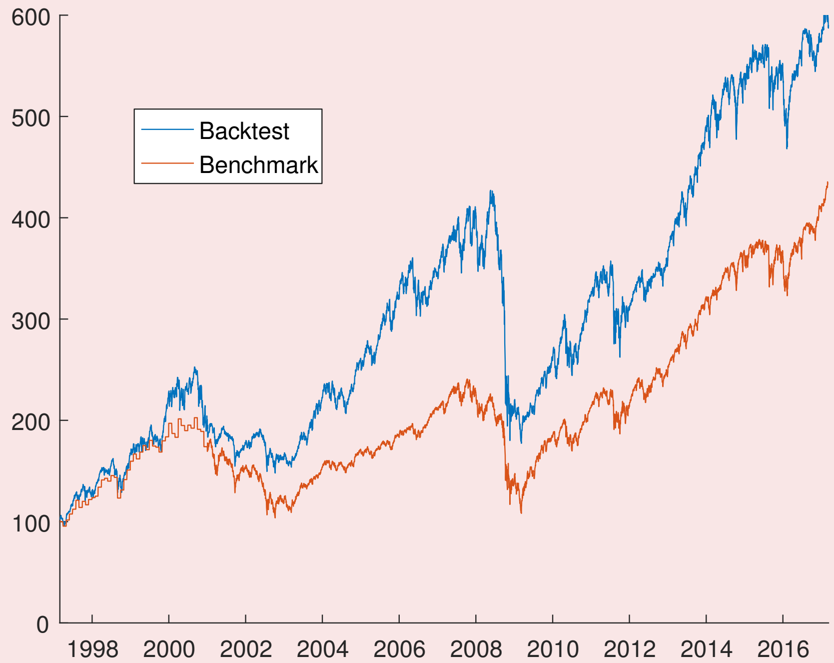
2. **Because risk factors are not independent in practice**

3. **Because the explanatory power of risk factors is time-varying**

4. **Because alpha and beta are highly related (beta strategy = successful alpha strategy)**

The issue of backtesting

Backtesting syndrome



The blue line is above the red line $\Rightarrow$ it's OK!

$\Rightarrow$ Analytical models are important to understand a risk factor

The professional consensus

There is now a consensus among professionals that five factors are sufficient for the equity markets:

1. **Size**

   Small cap stocks $\neq$ Large cap stocks

2. **Value**

   Value stocks $\neq$ Non-value stocks (including growth stocks)

3. **Momentum**

   Past winners $\neq$ Past loosers

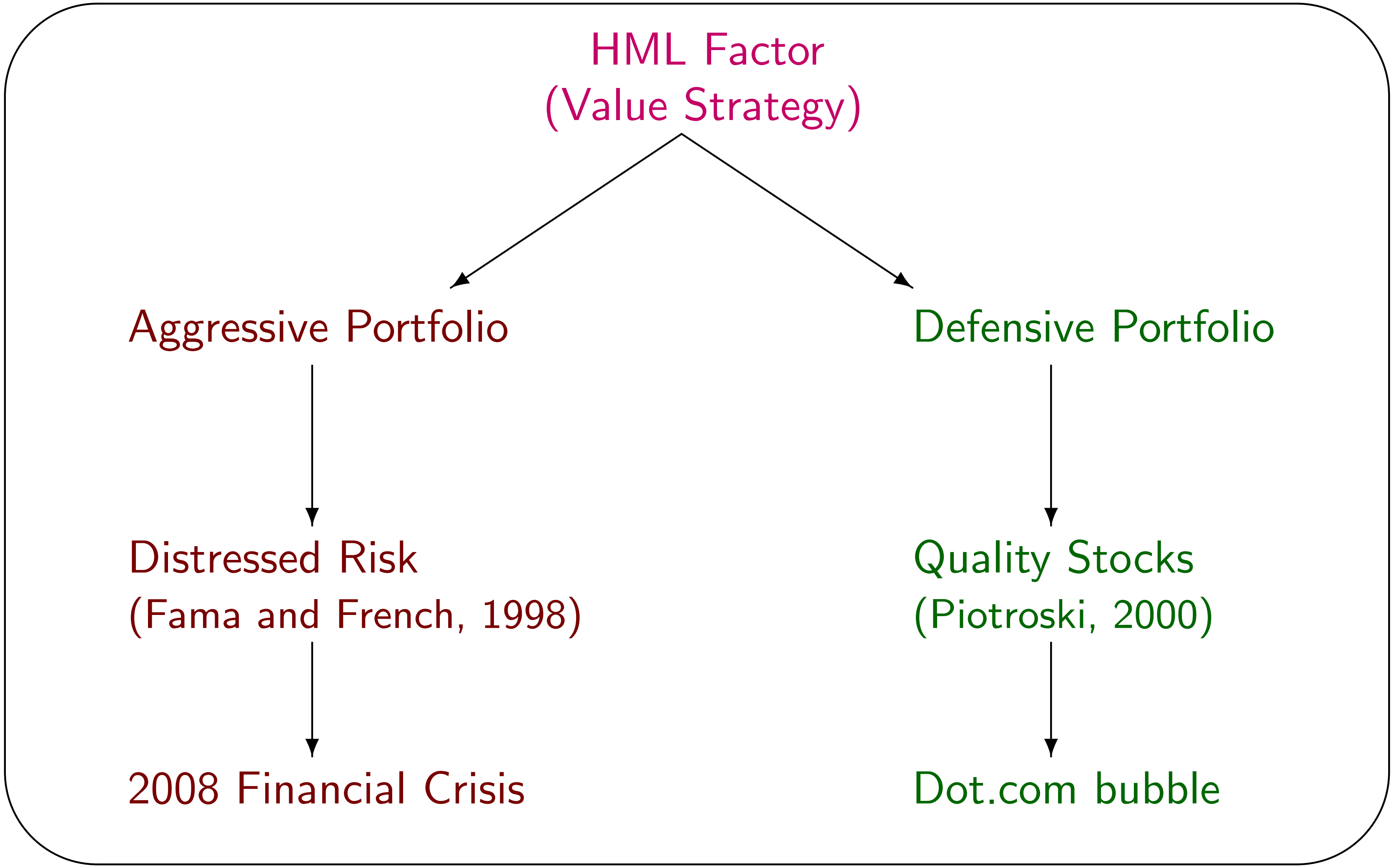
4. **Low-volatility**

   Low-vol (or low-beta) stocks $\neq$ High-vol (or high-beta stocks)

5. **Quality**

   Quality stocks $\neq$ Non-quality stocks (including junk stocks)

The example of the value risk factor



The example of the dividend yield risk factor

- Book-to-price (value risk factor):

$$\mathrm{B2P} = \frac{B}{P}$$

- Dividend yield (carry risk factor):

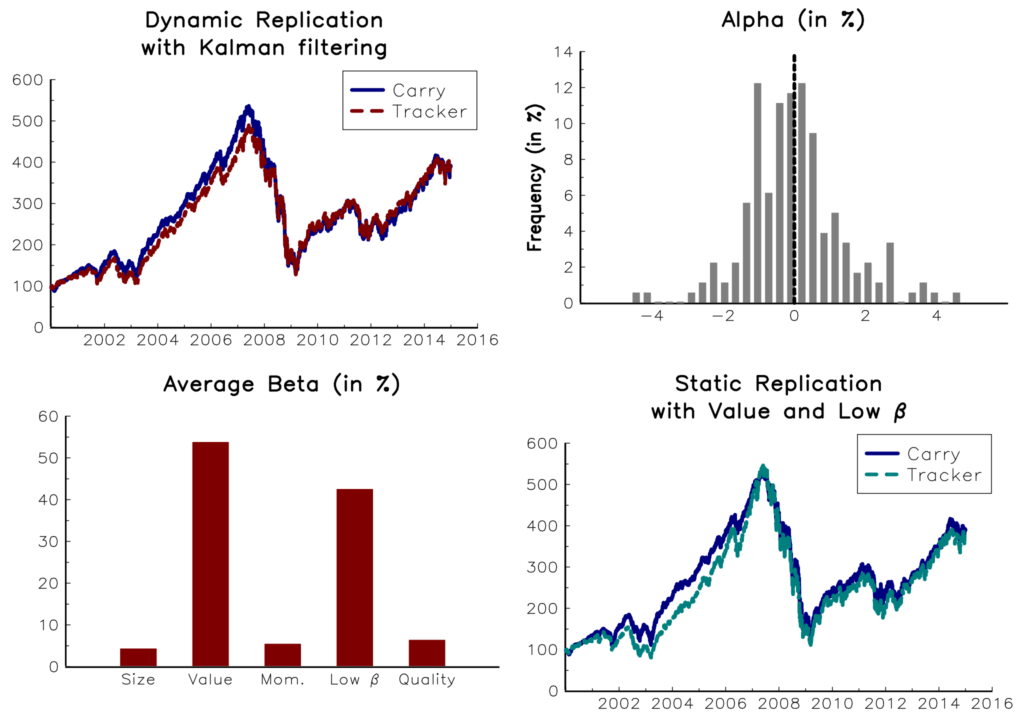
$$\begin{aligned}\mathrm{DY} &= \frac{D}{P} \\ &= \frac{D}{B} \times \frac{B}{P} \\ &= \mathrm{D2B} \times \mathrm{B2P}\end{aligned}$$

- Value component (book and dividend = low-frequency, price = high-frequency)
- Low-volatility component (bond-like stocks)

Risk factors are not orthogonal, they are correlated

The example of the dividend yield risk factor

Figure 51: Value, low beta and carry are not orthogonal risk factors



Source: Richard and Roncalli (2015)

$$\text{Carry} \simeq 60\% \text{ Value} + 40\% \text{ Low-volatility}$$

The example of the dividend yield risk factor

- Why Size + Value + Momentum + Low-volatility + Quality?
- Why not Size + Carry + Momentum + Low-volatility + Quality or Size + Carry + Momentum + Value + Quality?
- Because:

$$\begin{aligned} \text{Carry} &\simeq 60\%\ \text{Value} + 40\%\ \text{Low-volatility} \\ \text{Value} &\simeq 167\%\ \text{Carry} - 67\%\ \text{Low-volatility} \\ \text{Low-volatility} &\simeq 250\%\ \text{Carry} - 150\%\ \text{Value} \end{aligned}$$

Question

Why Value + Momentum + Low-volatility + Quality
and not
Size + Value + Momentum + Low-volatility + Quality?

General approach

- We consider a universe $\mathcal{U}$ of stocks (e.g. the MSCI World Index)
- We define a rebalancing period (e.g. every month, every quarter or every year)
- At each rebalancing date t_τ:
  - We define a score $\mathbb{S}_i(t_\tau)$ for each stock i
  - Stocks with high scores are selected to form the long exposure $\mathcal{L}(t_\tau)$ of the risk factor
  - Stocks with low scores are selected to form the short exposure $\mathcal{S}(t_\tau)$ of the risk factor
- We specify a weighting scheme $w_i(t_\tau)$ (e.g. value weighted or equally weighted)

General approach

- The performance of the risk factor between two rebalancing dates corresponds to the performance of the long/short portfolio:

$$\mathcal{F}(t) = \mathcal{F}(t_\tau) \cdot \left(\sum_{i \in \mathcal{L}(t_\tau)} w_i(t_\tau)(1 + R_i(t)) - \sum_{i \in \mathcal{S}(t_\tau)} w_i(t_\tau)(1 + R_i(t)) \right)$$

  where $t \in]t_\tau, t_{\tau+1}]$ and $\mathcal{F}(t_0) = 100$.

- In the case of a long-only risk factor, we only consider the long portfolio:

$$\mathcal{F}(t) = \mathcal{F}(t_\tau) \cdot \left(\sum_{i \in \mathcal{L}(t_\tau)} w_i(t_\tau)(1 + R_i(t)) \right)$$

The Fama-French approach

The SMB and HML factors are defined as follows:

$$\mathrm{SMB}_t = \frac{1}{3}\left(R_t\left(\mathrm{SV}\right) + R_t\left(\mathrm{SN}\right) + R_t\left(\mathrm{SG}\right)\right) - \frac{1}{3}\left(R_t\left(\mathrm{BV}\right) + R_t\left(\mathrm{BN}\right) + R_t\left(\mathrm{BG}\right)\right)$$

and:

$$\mathrm{HML}_t = \frac{1}{2}\left(R_t\left(\mathrm{SV}\right) + R_t\left(\mathrm{BV}\right)\right) - \frac{1}{2}\left(R_t\left(\mathrm{SG}\right) + R_t\left(\mathrm{BG}\right)\right)$$

with the following 6 portfolios[11]:

	Value	Neutral	Growth
Small	SV	SN	SG
Big	BV	BN	BG

[11]We have:

- The scores are the market equity (ME) and the book equity to market equity (BE/ME)
- The size breakpoint is the median market equity (Small = 50% and Big = 50%)
- The value breakpoints are the 30th and 70th percentiles of BE/ME (Value = 30%, Neutral = 40% and Growth = 30%)

The Fama-French approach

Homepage of Kenneth R. French

You can download data at the following webpage:

https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

- Asia Pacific ex Japan
- Developed
- Developed ex US
- Europe
- Japan
- North American
- US

Quintile portfolios

In this approach, we form five quintile portfolios:

- Q_1 corresponds to the stocks with the highest scores (top 20%)
- Q_2, Q_3 and Q_4 are the second, third and fourth quintile portfolios
- Q_5 corresponds to the stocks with the lowest scores (bottom 20%)

$\Rightarrow$ The long/short risk factor is the performance of $Q_1 - Q_5$, whereas the long-only risk factor is the performance of Q_1

The construction of risk factors

Table 59: An illustrative example

Asset	Score	Rank	Quintile	Selected	L/S	Weight
A_1	1.1	3	Q_2			
A_2	0.5	4	Q_2			
A_3	−1.3	9	Q_5	✓	**Short**	−50%
A_4	1.5	2	Q_1	✓	**Long**	+50%
A_5	−2.8	10	Q_5	✓	**Short**	−50%
A_6	0.3	5	Q_3			
A_7	0.1	6	Q_3			
A_8	2.3	1	Q_1	✓	**Long**	+50%
A_9	−0.7	8	Q_4			
A_{10}	−0.3	7	Q_4			

The scoring system

Variable selection

- Size: market capitalization
- Value: Price to book, price to earnings, price to cash flow, dividend yield, etc.
- Momentum = one-year price return ex 1 month, 13-month price return minus one-month price return, etc.
- Low volatility = one-year rolling volatility, one-year rolling beta, etc.
- Quality: Profitability, leverage, ROE, Debt to Assets, etc.

The scoring system

Variable combination

- Z-score averaging
- Ranking system
- Bottom exclusion
- Etc.

⇒ Finally, we obtain one score for each stock (e.g. the value score, the quality score, etc.)

Single-factor exposure versus multi-factor portfolio

Single-factor

- Trading bet
- Tactical asset allocation (TAA)
- If the investor believe that value stocks will outperform growth stocks in the next six months, he will overweight value stocks or add an exposure on the value risk factor
- Active management

Multi-factor

- Long-term bet
- Strategic asset allocation (SAA)
- The investor believe that a factor investing portfolio allows to better capture the equity risk premium than a CW index
- Factor investing portfolio = diversified portfolio (across risk factors)

Multi-factor portfolio

- Long/short: Market + Size + Value + Momentum + Low-volatility + Quality
- Long-only: Size + Value + Momentum + Low-volatility + Quality (because the market risk factor is replicated by the other risk factors)

Risk factors in sovereign bonds

> *"Market participants have long recognized the importance of identifying the common factors that affect the returns on U.S. government bonds and related securities. To explain the variation in these returns, it is critical to distinguish the systematic risks that have a general impact on the returns of most securities from the specific risks that influence securities individually and hence a negligible effect on a diversified portfolio" (Litterman and Scheinkman, 1991, page 54).*

$\Rightarrow$ The 3-factor model of Litterman and Scheinkman (1991) is based on the PCA analysis:

- the level of the yield curve
- the steepness of the yield curve
- the curvature of the yield curve

Conventional bond model

- Let $B_i(t, D_i)$ be the zero-coupon bond price with maturity D_i:

$$B_i(t, D_i) = e^{-(R(t)+S_i(t))\,D_i}$$

where $R(t)$ is the risk-free interest rate and $S_i(t)$ is the credit spread

- L-CAPM of Acharya and Pedersen (2005):

$$R_i(t) = \underbrace{\underbrace{(R(t) + S_i(t))\,D_i}_{\text{Gross return}} - L_i(t)}_{\text{Net return}}$$

where $L_i(t)$ is the illiquidity cost of Bond i

Conventional bond model

We deduce that:

$$B_i(t, D_i) = e^{-((R(t)+S_i(t))\,D_i - L_i(t))}$$

and:

$$\begin{aligned}
\mathrm{d}\ln B_i(t, D_i) &= -D_i\,\mathrm{d}R(t) - D_i\,\mathrm{d}S_i(t) + \mathrm{d}L_i(t) \\
&= -D_i\,\mathrm{d}R(t) - \mathrm{DTS}_i(t)\,\frac{\mathrm{d}S_i(t)}{S_i(t)} + \mathrm{d}L_i(t)
\end{aligned}$$

where $\mathrm{DTS}_i(t) = D_i\,S_{i,t}$ is the duration-time-spread factor

Conventional bond model

Liquidity premia (Acharya and Pedersen, 2005)

The illiquidity premium $\mathrm{d}L_{i,t}$ can be decomposed into an illiquidity level component $\mathbb{E}\left[L_{i,t}\right]$ and three illiquidity covariance risks:

1. $\beta\left(L_i, L_M\right)$
   An asset that becomes illiquid when the market becomes illiquid should have a higher risk premium.
2. $\beta\left(R_i, L_M\right)$
   An asset that perform well in times of market illiquidity should have a lower risk premium.
3. $\beta\left(L_i, R_M\right)$
   Investors accept a lower risk premium on assets that are liquid in a bear market.

Conventional bond model

By assuming that:

$$\mathrm{d}L_{i,t} = \alpha_i(t) + \beta(L_i, L_M)\,\mathrm{d}L_M(t)$$

where α_i is the liquidity return that is not explained by the liquidity commonality, we obtain:

$$R_i(t) = \alpha_i(t) - D_i\,\mathrm{d}R(t) - \mathrm{DTS}_i(t)\,\frac{\mathrm{d}S_i(t)}{S_i(t)} + \beta(L_i, L_M)\,\mathrm{d}L_M(t)$$

or:

$$R_i(t) = a(t) - D_i\,\mathrm{d}R(t) - \mathrm{DTS}_i(t)\,\frac{\mathrm{d}S_i(t)}{S_i(t)} + \beta(L_i, L_M)\,\mathrm{d}L_M(t) + u_i(t)$$

Risk factors in corporate bonds

Conventional bond model (or the 'equivalent' CAPM for bonds)

The total return $R_i(t)$ of Bond i at time t is equal to:

$$R_i(t) = a(t) - \mathrm{MD}_i(t)\, R^I(t) - \mathrm{DTS}_i(t)\, R^S(t) + \mathrm{LTP}_i(t)\, R^L(t) + u_i(t)$$

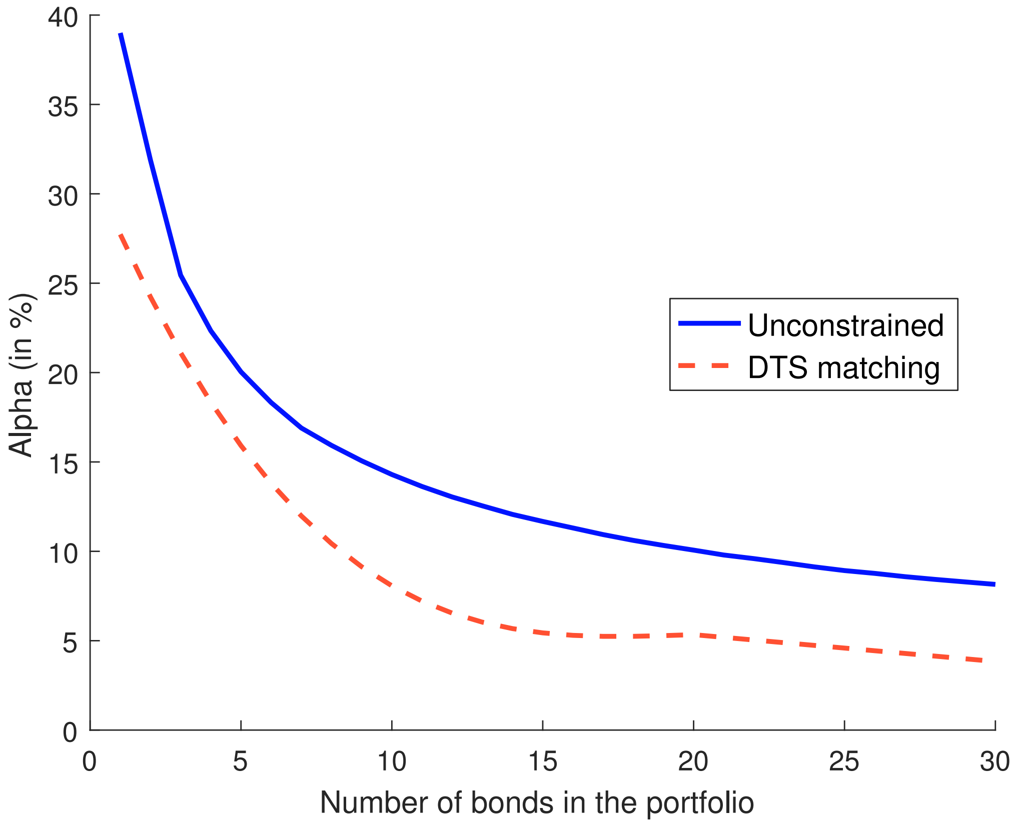
where:

- $a(t)$ is the constant/carry/zero intercept
- $\mathrm{MD}_i(t)$ is the modified duration
- $\mathrm{DTS}_i(t)$ is the duration-times-spread
- $\mathrm{LTP}_i(t)$ is the liquidity-time-price
- $u_i(t)$ is the residual

$\Rightarrow$ $R^I(t)$, $R^S(t)$ and $R^L(t)$ are the return components due to interest rate movements, credit spread variation and liquidity dynamics.

Risk factors in corporate bonds

Figure 52: Conventional alpha decreases with the number of holding assets



EURO IG corporate bonds, 2000-2015
Source: Amundi Research (2017)

- There is less traditional alpha in the bond market than in the stock market

Risk factors in corporate bonds

Since 2015

- Houweling and van Zundert (2017) — HZ
- Bektic, Neugebauer, Wegener and Wenzler (2017) — BNWW
- Israel, Palhares and Richardson (2017) — IPR
- Bektic, Wenzler, Wegener, Schiereck and Spielmann (2019) — BWWSS
- Ben Slimane, De Jong, Dumas, Fredj, Sekine and Srb (2019) — BDDFSS
- Etc.

Risk factors in corporate bonds

Study	HZ	BWWSS	IPR	BNWW
Period	1994-2015	1996-2016 (US) 2000-2016 (EU)	1997-2015	1999-2016
Universe	Bloomberg Barclays US IG & HY	BAML US IG & HY, EU IG	BAML US IG & HY	BAML US IG & HY
Investment		1Y variation in total assets		
Low risk	Short maturity + High rating		Leverage × Duration × Profitability	1Y equity beta
Momentum	6M bond return		6M bond return + 6M stock return	1Y stock return
Profitability		Earnings-to-book		
Size	Market value of issuer	Market capitalization		Market capitalization
Value	Comparing OAS to Maturity × Rating × 3M OAS variation	Price-to-book	Comparing OAS to Duration × Rating × Bond return volatility + Implied default probability	Price-to-book

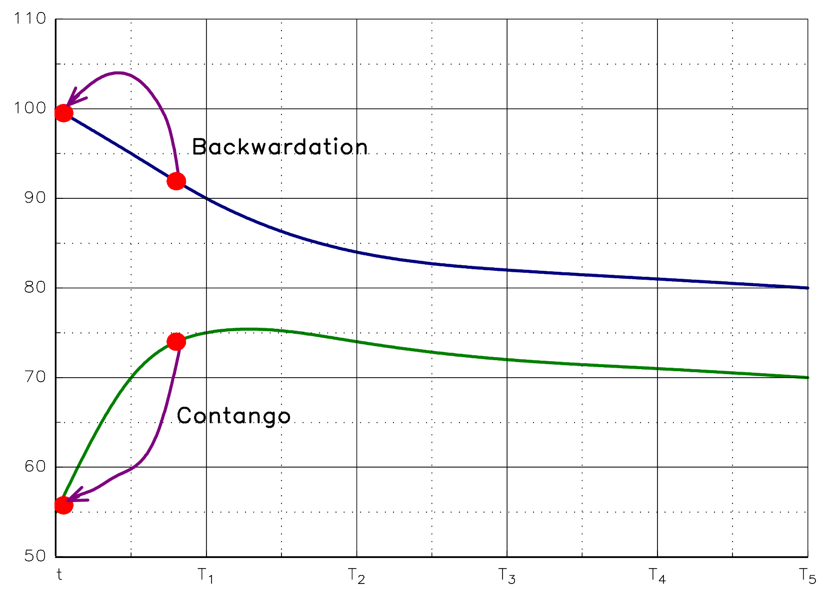
Risk factors in currency markets

- What are the main risk factors for explaining the cross-section of currency returns?
  1. Momentum (cross-section or time-series)
  2. Carry
  3. Value (short-term, medium-term or long-term)
- The dynamics of some currencies are mainly explained by:
  - Common risk factors (e.g. NZD or CAD)
  - Idiosyncratic risk factors (e.g. IDR or PEN)
- Carry-oriented currency? (e.g. JPY $\neq$ CHF)

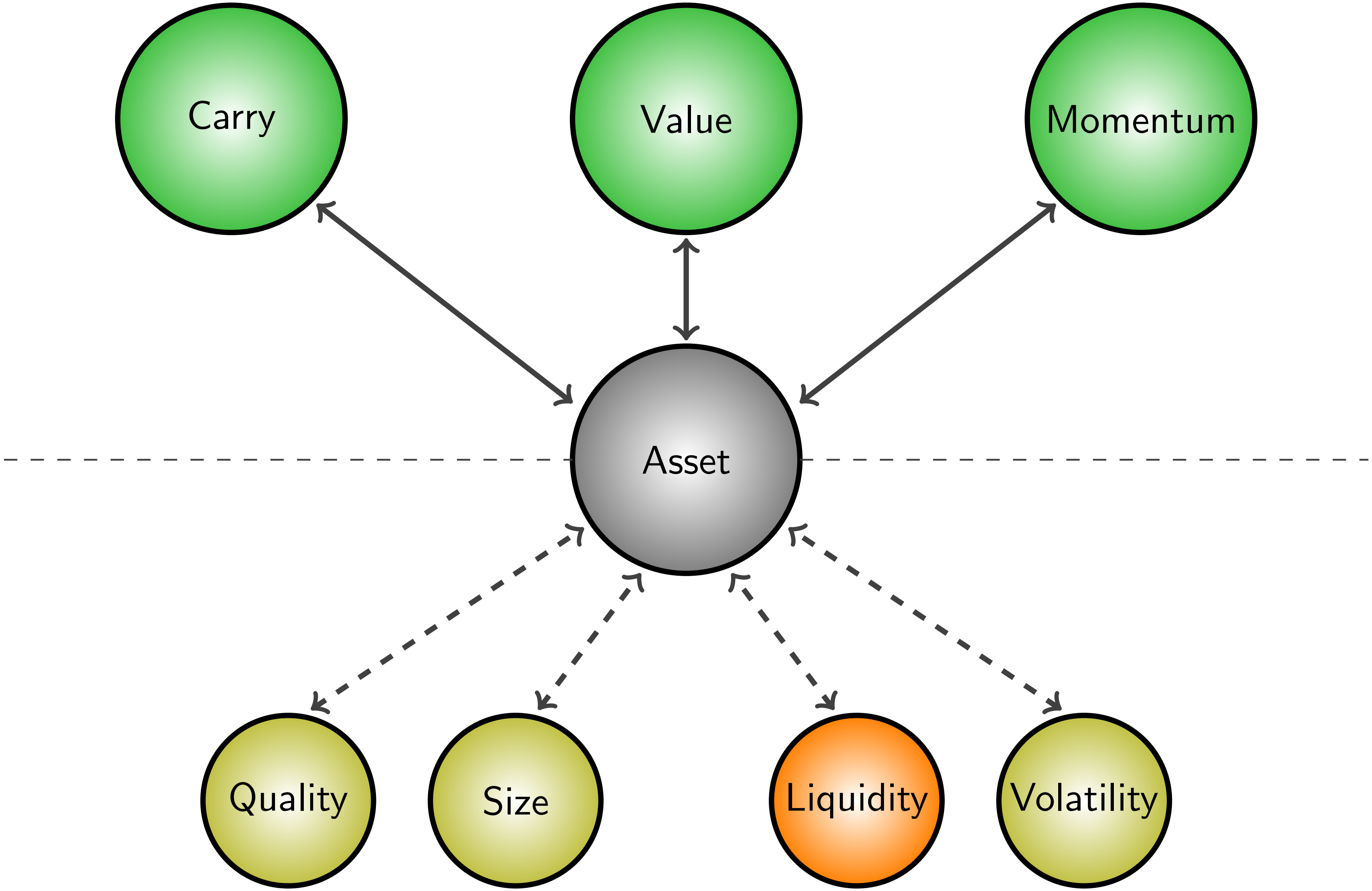
Risk factors in commodities

- Two universal strategies:
  - Contango/backwardation strategy
  - Trend-following strategy
- **CTA = Commodity Trading Advisor**
- Only two risk factors?
  - Carry
  - Momentum

Figure 53: Contango and backwardation movements in commodity futures contracts



Factor analysis of an asset



Factor analysis of an asset

Carry

- Yield
- Income generation
- Risk arbitrage

Value

- Fair price
- Overvalued / undervalued
- Fundamental

Momentum

- Price dynamics
- Trend-following
- Mean-reverting / Reversal

Liquidity

- Tradability property (transaction cost, execution time, scarcity)
- Supply/demande imbalance
- Bad times $\neq$ good times

The concept of alternative risk premia

There are many definitions of ARP:

- ARP $\approx$ factor investing (FI)
  (ARP = long/short portfolios, FI = long portfolios)
- ARP $\approx$ all the other risk premia (RP) than the equity and bond risk premia
- ARP $\approx$ quantitative investment strategies (QIS)

Sell-side

- CIBs & brokers
- ARP = QIS

Buy-side

- Asset managers & asset owners
- ARP = FI (for asset managers)
- ARP = RP (for asset owners)

The concept of alternative risk premia

Alternative Risk Premia

Alternative (or real) assets

- Private equity
- Private debt
- Real estate
- Infrastructure

Traditional financial assets

- Long/short risk factors in equities, rates, credit, currencies & commodities
- Risk premium strategy (e.g. carry, momentum, value, etc.)

The concept of alternative risk premia

- A risk premium is the expected excess return by the investor in order to accept the risk ⇒ any (risky) investment strategy has a risk premium!
- Generally, the term "*risk premium*" is associated to asset classes:
  - The equity risk premium
  - The risk premium of high yield bonds
- This means that a risk premium is the expected excess return by the investor in order to accept a future economic risk that cannot be diversifiable
  - For instance, the risk premium of a security does not integrate its specific risk

The concept of alternative risk premia

- What is the relationship between a risk factor and a risk premium?
  - A rewarded risk factor may correspond a to risk premium, while a non-rewarded risk factor is not a risk premium
  - A risk premium can be a risk factor if it helps to explain the cross-section of expected returns
  - The case of cat bonds:

Risk premium	✔
Risk factor	✘

Risk premia & non-diversifiable risk

Consumption-based model (Lucas, 1978; Cochrane, 2001)

A risk premium is a compensation for accepting (systematic) risk in **bad times**.

We have:

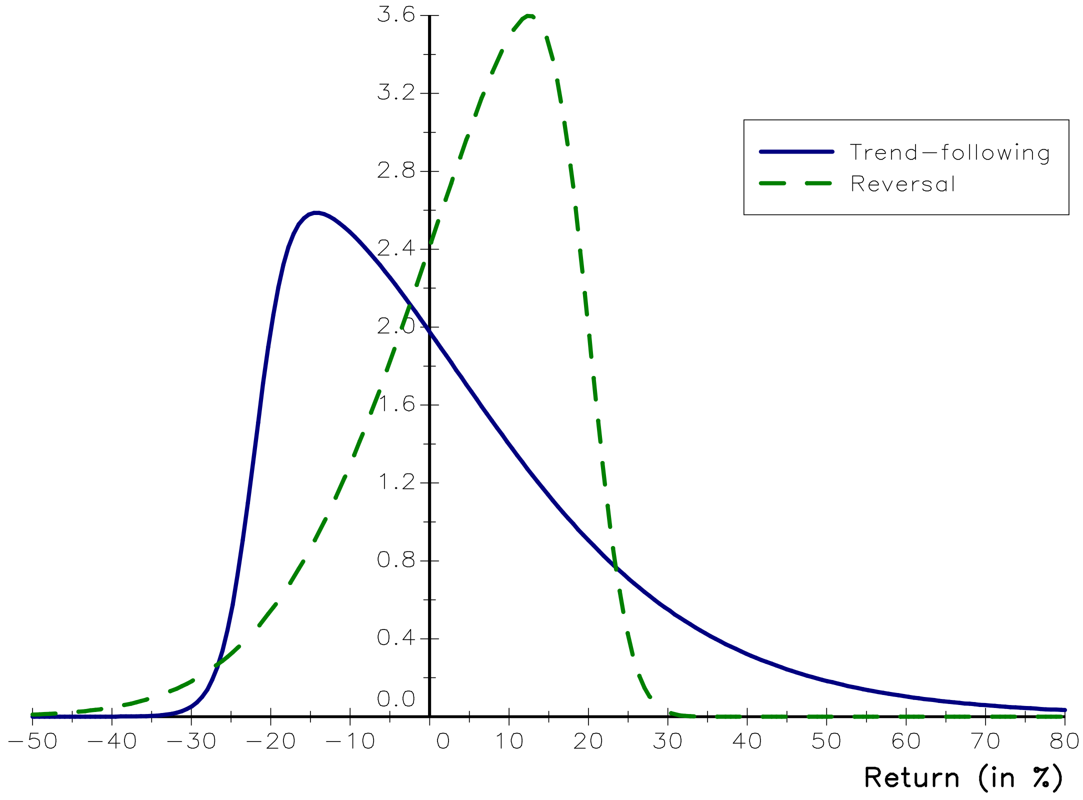
$$\underbrace{\mathbb{E}_t\left[R_{t+1} - R_{f,t}\right]}_{\text{Risk premium}} \propto -\underbrace{\rho\left(u'\left(C_{t+1}\right), R_{t+1}\right)}_{\text{Correlation term}} \times \underbrace{\sigma\left(u'\left(C_{t+1}\right)\right)}_{\text{Smoothing term}} \times \underbrace{\sigma\left(R_{t+1}\right)}_{\text{Volatility term}}$$

where R_{t+1} is the one-period return of the asset, $R_{f,t}$ is the risk-free rate, C_{t+1} is the future consumption and $u(C)$ is the utility function.

Main results

- Hedging assets help to smooth the consumption $\Rightarrow$ low or negative risk premium
- In bad times, risk premium strategies are correlated and have a negative performance ($\neq$ all-weather strategies)

Risk premia & bad times



The market must reward contrarian and value investors, not momentum investors

Behavioral finance and limits to arbitrage

Bounded rationality

Barberis and Thaler (2003), A Survey of Behavioral Finance.

Decisions of the other economic agents

$\Downarrow$

Feedback effects on our decisions!

Killing Homo Economicus

[...] "conventional economics assumes that people are highly rational, super rational and unemotional. They can calculate like a computer and have no self-control problems" (Richard Thaler, 2009).

"The people I study are humans that are closer to Homer Simpson" (Richard Thaler, 2017).

Behavioral finance and social preferences

- For example, momentum may be a rational behavior if the investor is not informed and his objective is to minimize the loss with respect to the 'average' investor.
- Absolute loss $\neq$ relative loss
- Loss aversion and performance asymmetry
- Imitations between institutional investors $\Rightarrow$ benchmarking
- Home bias

What does the theory become if utility maximization includes the performance of other economic agents?

$\Rightarrow$ **The crowning glory of tracking error and relative performance!**

Behavioral finance and market anomalies

Previously

Positive expected excess returns are explained by:

- risk premia

Today

Positive expected excess returns are explained by:

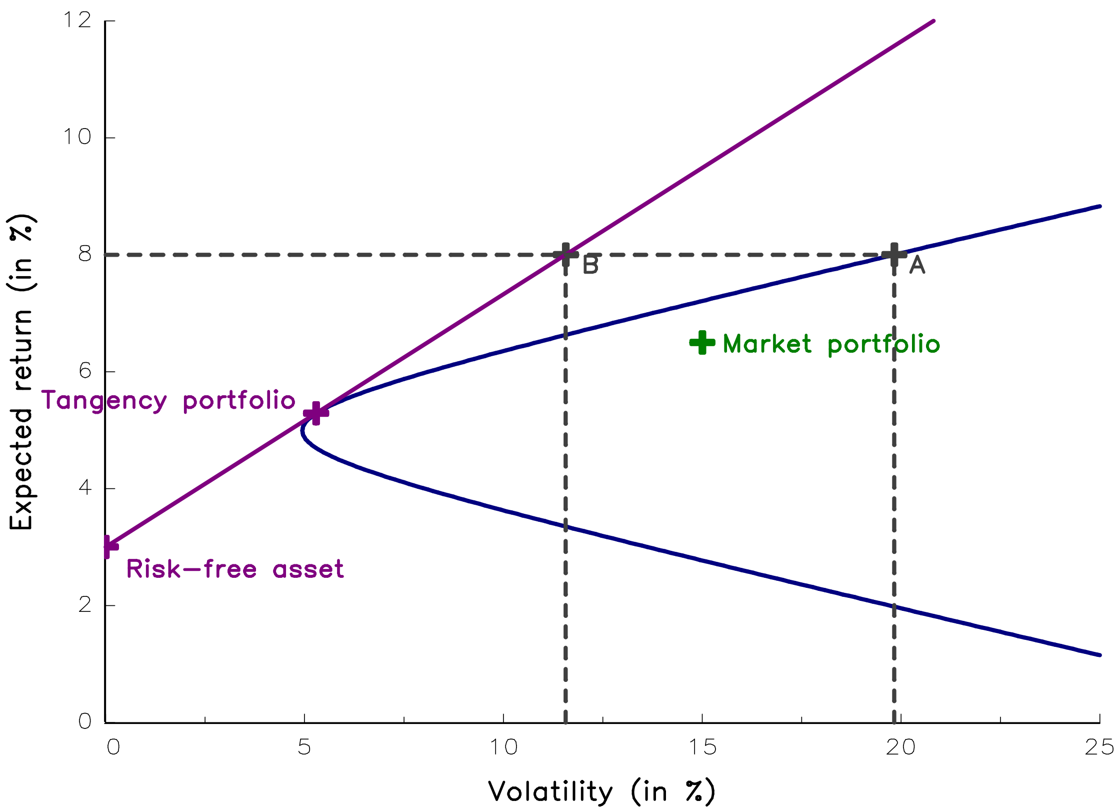
- risk premia
- or market anomalies

Market anomalies correspond to trading strategies that have delivered good performance in the past, but their performance cannot be explained by the existence of a systematic risk (in bad times). Their performance can only be explained by behavioral theories.

$\Rightarrow$ Momentum, low risk and quality risk factors are three market anomalies

The case of low risk assets

Figure 54: What is the impact of borrowing constraints on the market portfolio?



- The investor that targets a 8% expected return must choose Portfolio B
- The demand for high beta assets is higher than this predicted by CAPM
- This effect is called the low beta anomaly

Low risk assets have a higher Sharpe ratio than high risk assets

Skewness risk premia & market anomalies

Characterization of alternative risk premia

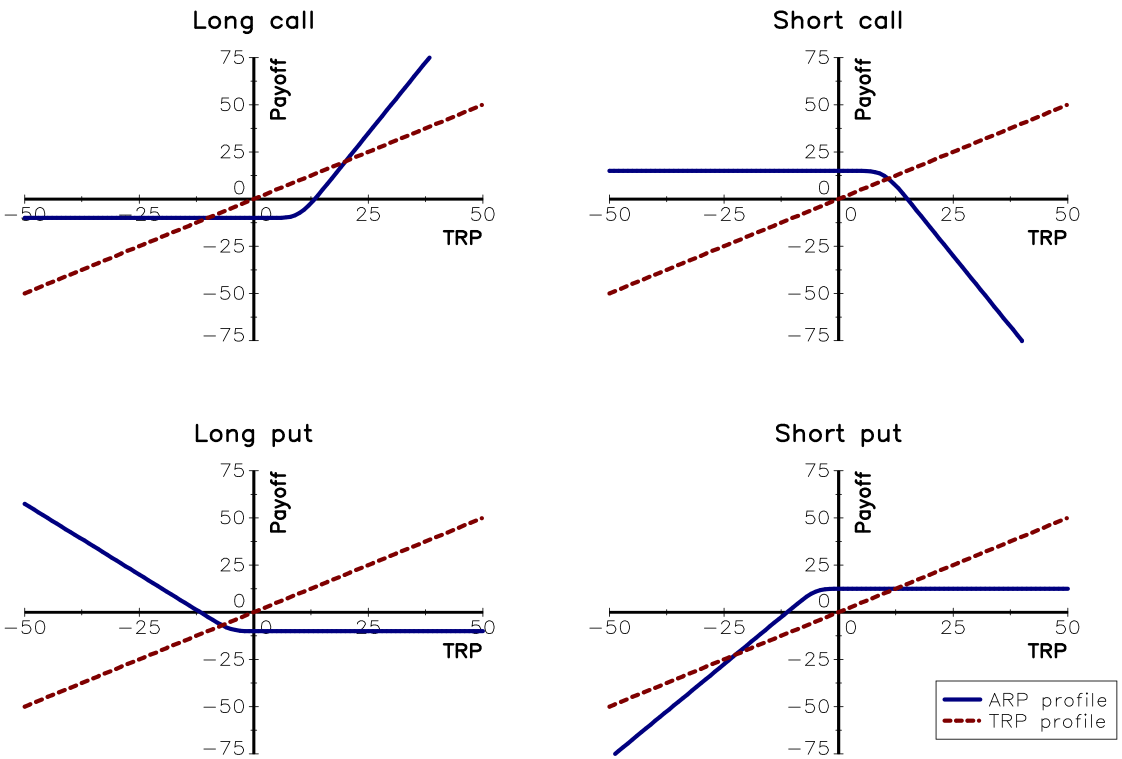
- An alternative risk premium (ARP) is a risk premium, which is not traditional
  - Traditional risk premia (TRP): equities, sovereign/corporate bonds
  - Currencies and some commodities are not TRP
- The drawdown of an ARP must be positively correlated to bad times
  - Risk premia $\neq$ insurance against bad times
  - (SMB, HML) $\neq$ WML
- Risk premia are an increasing function of the volatility and a decreasing function of the skewness

In the market practice, alternative risk premia recover:

1. Skewness risk premia (or pure risk premia), which present high negative skewness and potential large drawdown
2. Markets anomalies

Payoff function of alternative risk premia

Figure 55: Which option profile may be considered as a skewness risk premium?



- ~~Long call~~ (risk adverse)
- ~~Short call~~ (market anomaly)
- ~~Long put~~ (insurance)
- Short put

$\Rightarrow$ SMB, HML, ~~WML~~, ~~BAB~~, ~~QMJ~~

A myriad of alternative risk premia?

Figure 56: Mapping of risk premia strategies (based on existing products)

Strategy	Equities	Rates	Credit	Currencies	Commodities
Carry	Dividend futures High dividend yield	Forward rate bias Term structure slope Cross-term-structure	Forward rate bias	Forward rate bias	Forward rate bias Term structure slope Cross-term-structure
Event	Buyback Merger arbitrage				
Growth	Growth				
Liquidity	Amihud liquidity	Turn-of-the-month	Turn-of-the-month		Turn-of-the-month
Low beta	Low beta Low volatility				
Momentum	Cross-section Time-series	Cross-section Time-series	Time-series	Cross-section Time-series	Cross-section Time-series
Quality	Quality				
Reversal	Time-series Variance	Time-series		Time-series	Time-series
Size	Size				
Value	Value	Value	Value	PPP REER, BEER, FEER NATREX	Value
Volatility	Carry Term structure	Carry		Carry	Carry

Source: Roncalli (2017)

The carry risk premium

Underlying idea

Definition

- The investor takes an investment risk
- This investment risk is rewarded by a high and known yield
- Financial theory predicts a negative mark-to-market return that may reduce or write off the performance
- The investor hopes that the impact of the mark-to-market will be lower than the predicted value

$\Rightarrow$ Carry strategies are highly related to the concept of risk arbitrage[12]

- The carry risk premium is extensively studied by Koijen *et al.* (2018)
- The carry risk premium has a short put option profile

[12] An example is the carry strategy between pure money market instruments and commercial papers = not the same credit risk, not the same maturity risk, but the investor believes that the default will never occur!

The carry risk premium

Not one but several carry strategies

- Equity
  - Carry on dividend futures
  - Carry on stocks with high dividend yields (HDY)
- Rates (sovereign bonds)
  - Carry on the yield curve (term structure & roll-down)
- Credit (corporate bons)
  - Carry on bonds with high spreads
  - High yield strategy
- Currencies
  - Carry on interest rate differentials (uncovered interest rate parity)
- Commodities
  - Carry on contango & backwardation movements
- Volatility
  - Carry on option implied volatilities
  - Short volatility strategy

$\Rightarrow$ Many implementation methods: security-slope, cross-asset, long/short, long-only, basis arbitrage, etc.

The carry risk premium
Analytical model

- Let X_t be the capital allocated at time t to finance a futures position (or an unfunded forward exposure) on asset S_t
- By assuming that the futures price expires at the future spot price ($F_{t+1} = S_{t+1}$), Koijen *et al.* (2018) showed that:

$$
\begin{aligned}
R_{t+1}(X_t) - R_f &= \frac{F_{t+1} - F_t}{X_t} \\
&= \frac{S_{t+1} - F_t}{X_t} \\
&= \frac{S_t - F_t}{X_t} + \frac{\mathbb{E}_t[S_{t+1}] - S_t}{X_t} + \frac{S_{t+1} - \mathbb{E}_t[S_{t+1}]}{X_t}
\end{aligned}
$$

The carry risk premium

Analytical model

- At time $t+1$, the excess return of this investment is then equal to:

$$R_{t+1}(X_t) - R_f = \mathcal{C}_t + \frac{\mathbb{E}_t[\Delta S_{t+1}]}{X_t} + \varepsilon_{t+1}$$

  where $\varepsilon_{t+1} = (S_{t+1} - \mathbb{E}_t[S_{t+1}])/X_t$ is the unexpected price change and $\mathcal{C}_t$ is the carry:

$$\mathcal{C}_t = \frac{S_t - F_t}{X_t}$$

- It follows that the expected excess return is the sum of the carry and the expected price change:

$$\mathbb{E}_t[R_{t+1}(X_t)] - R_f = \mathcal{C}_t + \frac{\mathbb{E}_t[\Delta S_{t+1}]}{X_t}$$

- The nature of these two components is different:
  1. The carry is an ex-ante observable quantity (known value)
  2. The price change depends on the dynamic model of S_t (unknown value)

The carry risk premium

Analytical model

- If we assume that the spot price does not change (no-arbitrage assumption $\mathcal{H}$), the expected excess return is equal to the carry:

$$\frac{\mathbb{E}_t\left[\Delta S_{t+1}\right]}{X_t} = -\mathcal{C}_t$$

- The carry investor will prefer Asset i to Asset j if the carry of Asset i is higher:

$$\mathcal{C}_{i,t} \geq \mathcal{C}_{j,t} \Longrightarrow A_i \succ A_j$$

- The carry strategy would then be long on high carry assets and short on low carry assets.

Remark

In the case of a fully-collateralized position $X_t = F_t$, the value of the carry becomes:

$$\mathcal{C}_t = \frac{S_t}{F_t} - 1$$

The carry risk premium

Currency carry (or the carry trade strategy)

- Let S_t, i_t and r_t be the spot exchange rate, the domestic interest rate and the foreign interest rate for the period $[t, t+1]$
- The forward exchange rate F_t is equal to:

$$F_t = \frac{1 + i_t}{1 + r_t} S_t$$

- The carry is approximately equal to the interest rate differential:

$$\mathcal{C}_t = \frac{r_t - i_t}{1 + i_t} \simeq r_t - i_t$$

The carry risk premium

Currency carry (or the carry trade strategy)

- The carry strategy is long on currencies with high interest rates and short on currencies with low interest rates
- We can consider the following carry scoring (or ranking) system:

$$\mathcal{C}_t = r_t$$

Uncovered interest rate parity (UIP)

- An interest rate differential of 10% $\Rightarrow$ currency depreciation of 10% per year
- In 10 years, we must observe a depreciation of 65%!

The carry risk premium

Currency carry (or the carry trade strategy)

ARS	Argentine peso	KRW	Korean won
AUD	Australian dollar	LTL	Lithuanian litas
BGN	Bulgarian lev	LVL	Latvian lats
BHD	Bahraini dinar	MXN	Mexican peso
BRL	Brazilian real	MYR	Malaysian ringgit
CAD	Canadian dollar	NOK	Norwegian krone
CHF	Swiss franc	NZD	New Zealand dollar
CLP	Chilean peso	PEN	Peruvian new sol
CNY/RMB	Chinese yuan (Renminbi)	PHP	Philippine peso
COP	Colombian peso	PLN	Polish zloty
CZK	Czech koruna	RON	new Romanian leu
DKK	Danish krone	RUB	Russian rouble
EUR	Euro	SAR	Saudi riyal
GBP	Pound sterling	SEK	Swedish krona
HKD	Hong Kong dollar	SGD	Singapore dollar
HUF	Hungarian forint	THB	Thai baht
IDR	Indonesian rupiah	TRY	Turkish lira
ILS	Israeli new shekel	TWD	new Taiwan dollar
INR	Indian rupee	USD	US dollar
JPY	Japanese yen	ZAR	South African rand

The carry risk premium

Currency carry (or the carry trade strategy)

Baku *et al.* (2019, 2020) consider the most liquid currencies:

- **G10** AUD, CAD, CHF, EUR, GBP, JPY, NOK, NZD, SEK and USD
- **EM** BRL, CLP, CZK, HUF, IDR, ILS, INR, KRW, MXN, PLN, RUB, SGD, TRY, TWD and ZAR
- **G25** G10 + EM

They build currency risk factors using the following characteristics:

- The portfolio is equally-weighted and rebalanced every month
- The portfolio is notional-neutral (number of long exposures = number of short exposures)
- 3/3 for G10, 4/4 for EM and 7/7 for G25
- The long (resp. short) exposures correspond to the highest (resp. lowest) scores

The carry risk premium

Currency carry (or the carry trade strategy)

- Scoring system: $\mathbb{S}_{i,t} = \mathcal{C}_{i,t} = r_{i,t}$
- The carry strategy is long on currencies with high interest rates and short on currencies with low interest rates

Table 60: Risk/return statistics of the carry risk factor (2000-2018)

	G10	EM	G25
Excess return (in %)	3.75	11.21	7.22
Volatility (in %)	9.35	9.12	8.18
Sharpe ratio	0.40	1.23	0.88
Maximum drawdown (in %)	−31.60	−25.27	−17.89

Source: Baku *et al.* (2019, 2020)

The carry risk premium

Currency carry (or the carry trade strategy)

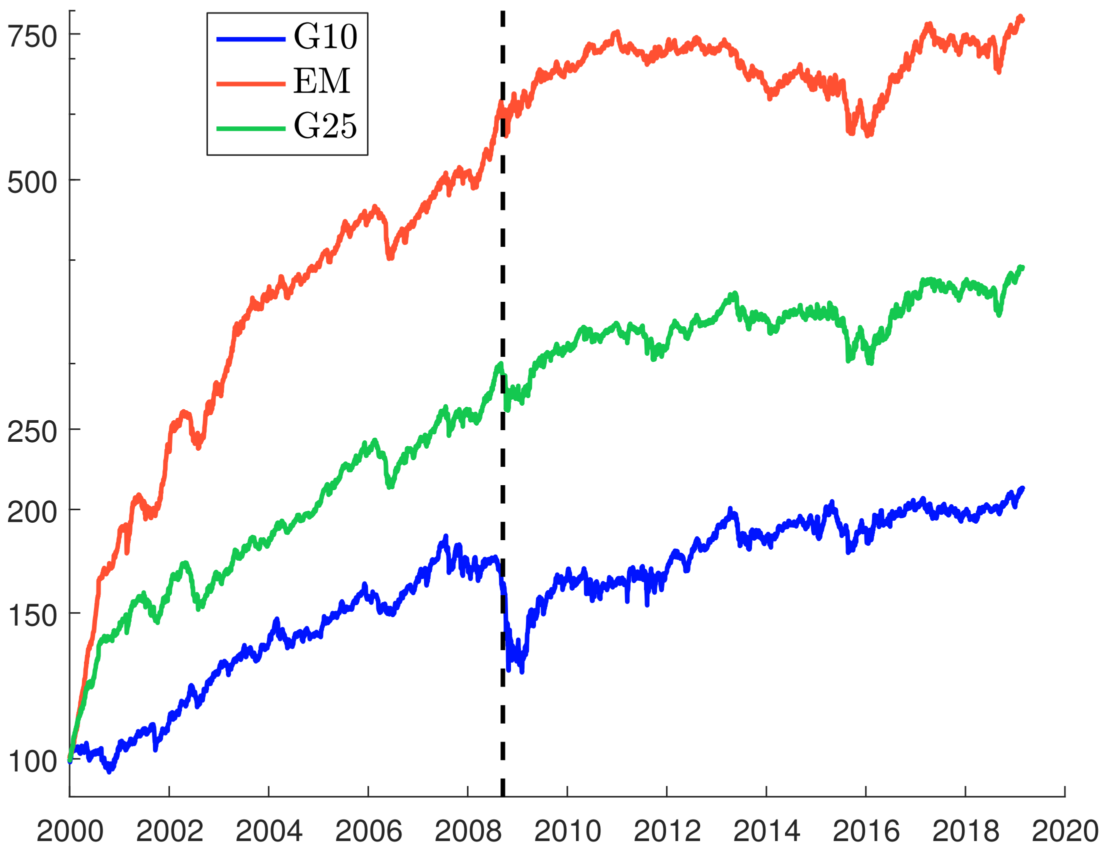


Figure 57: Cumulative performance of the carry risk factor

Source: Baku *et al.* (2019, 2020)

The carry risk premium

Equity carry

- We have:

$$\mathcal{C}_t \simeq \frac{\mathbb{E}_t\left[D_{t+1}\right]}{S_t} - r_t$$

where $\mathbb{E}_t\left[D_{t+1}\right]$ is the risk-neutral expected dividend for time $t+1$

- If we assume that dividends are constant, the carry is the difference between the dividend yield $\mathcal{y}_t$ and the risk-free rate r_t:

$$\mathcal{C}_t = \mathcal{y}_t - r_t$$

- The carry strategy is long on stocks with high dividend yields and short on stocks with low dividend yields
- This strategy may be improved by considering forecasts of dividends. In this case, we have:

$$\mathcal{C}_t \simeq \frac{\mathbb{E}_t\left[D_{t+1}\right]}{S_t} - r_t = \frac{D_t + \mathbb{E}_t\left[\Delta D_{t+1}\right]}{S_t} - r_t = \mathcal{y}_t + g_t - r_t$$

where g_t is the expected dividend growth

The carry risk premium

Equity carry

Carry strategy with dividend futures

Another carry strategy concerns dividend futures. The underlying idea is to take a long position on dividend futures where the dividend premium is the highest and a short position on dividend futures where the dividend premium is the lowest. Because dividend futures are on equity indices, the market beta exposure is generally hedged.

Why do we observe a premium on dividend futures?

$\Rightarrow$ Because of the business of structured products and options

The carry risk premium

Bond carry

- The price of a zero-coupon bond with maturity date T is equal to:

$$B_t(T) = e^{-(T-t)R_t(T)}$$

where $R_t(T)$ is the corresponding zero-coupon rate

- Let $F_t(T, m)$ denote the forward interest rate for the period $[T, T+m]$, which is defined as follows:

$$B_t(T+m) = e^{-mF_t(T,m)} B_t(T)$$

We deduce that:

$$F_t(T, m) = -\frac{1}{m} \ln \frac{B_t(T+m)}{B_t(T)}$$

It follows that the instantaneous forward rate is given by this equation:

$$F_t(T) = F_t(T, 0) = \frac{-\partial \ln B_t(T)}{\partial T}$$

The carry risk premium
Bond carry

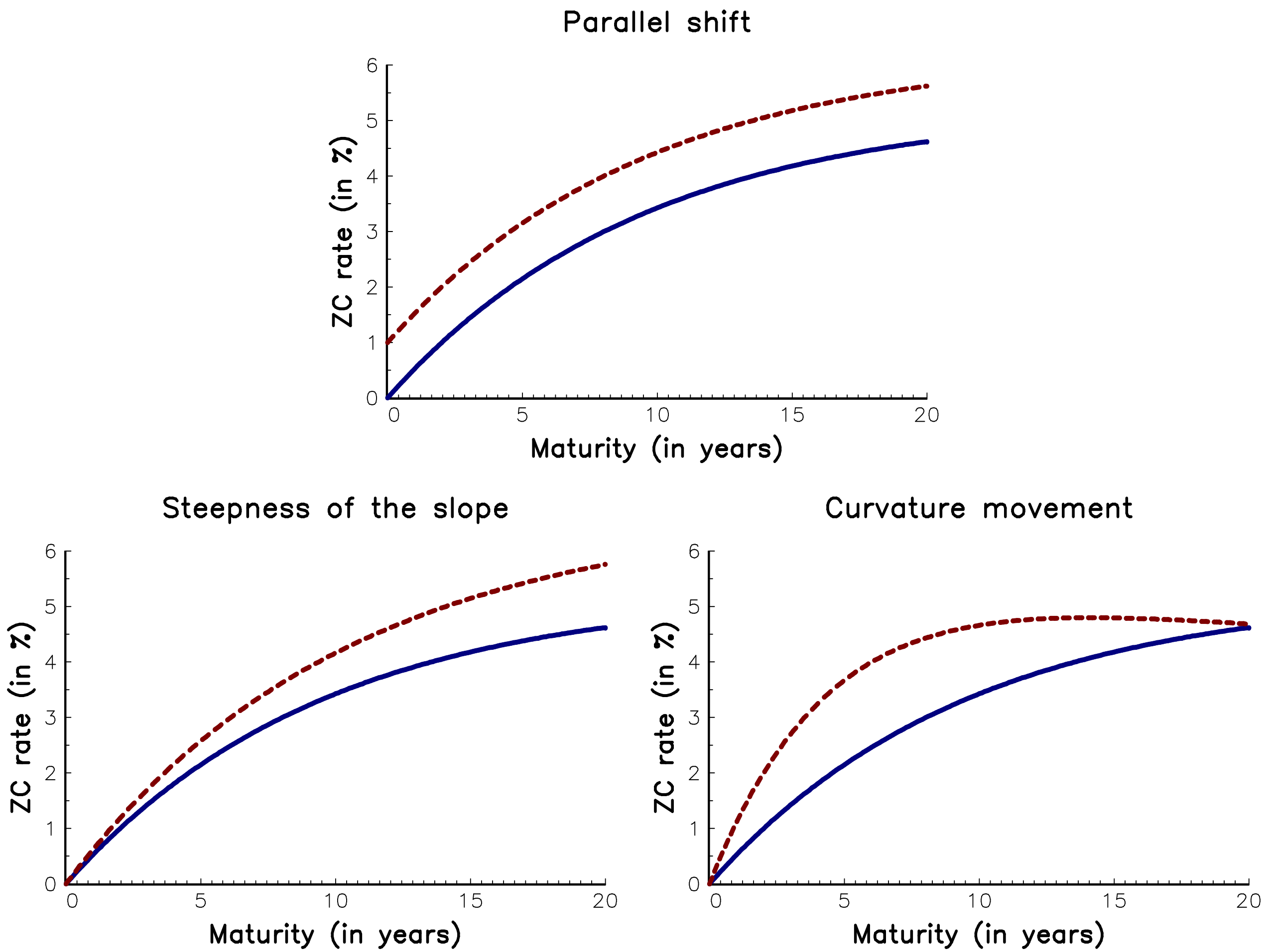


Figure 58: Movements of the yield curve

The carry risk premium

Bond carry

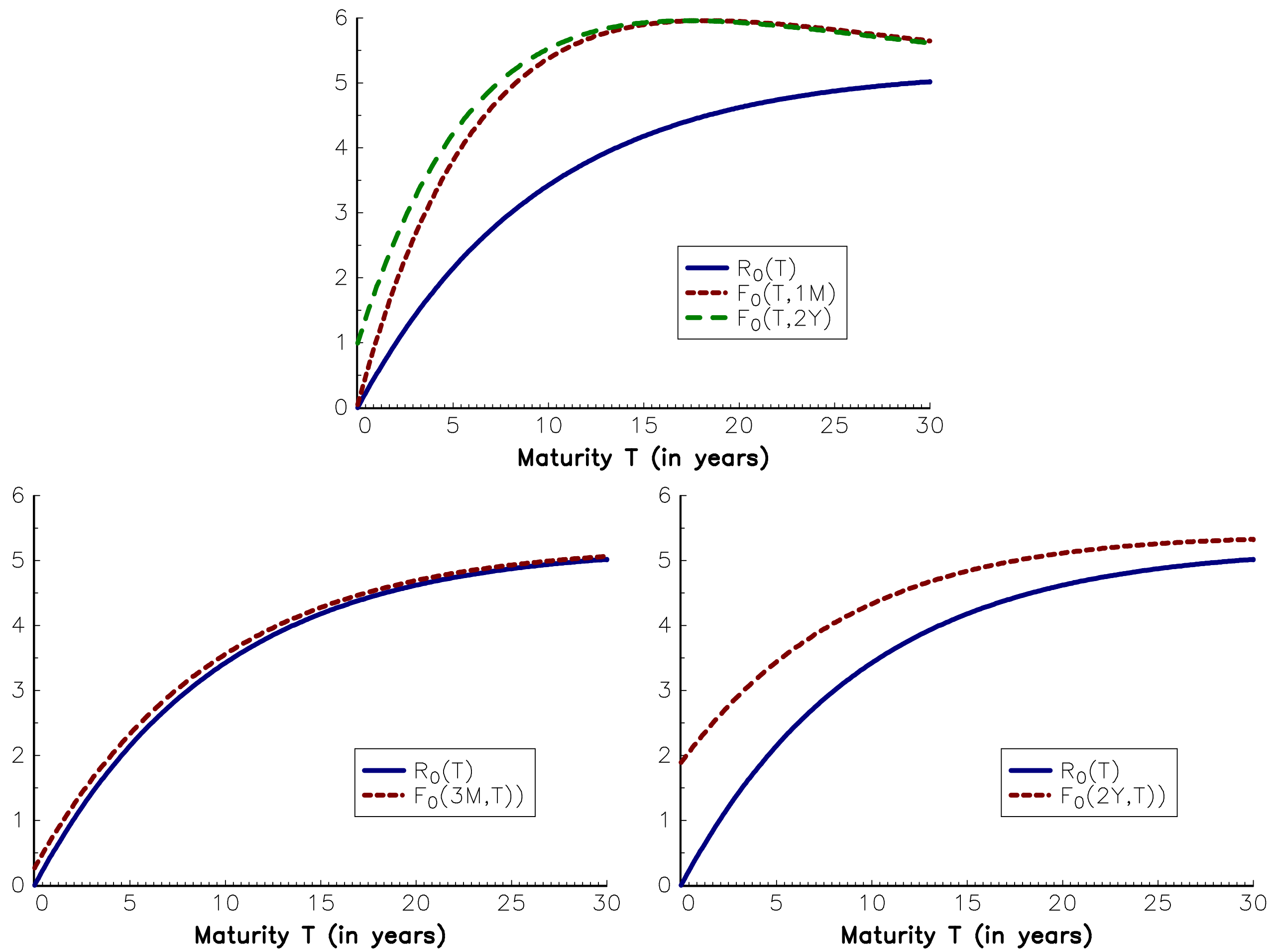


Figure 59: Sport and forward interest rates

The carry risk premium

Bond carry

1. The first carry strategy ("*forward rate bias*") consists in being long the forward contract on the forward rate $F_t(T, m)$ and selling it at time $t + \mathrm{d}t$ with $t + \mathrm{d}t \leq T$
   - Forward rates are generally higher than spot rates
   - Under the hypothesis $(\mathcal{H})$ that the yield curve does not change, rolling forward rate agreements can then capture the term premium and the roll down
   - The carry of this strategy is equal to:

$$\mathcal{C}_t = \underbrace{R_t(T) - r_t}_{\text{term premium}} + \underbrace{\partial_{\bar{T}} \bar{R}_t(\bar{T})}_{\text{roll down}}$$

where $\bar{R}_t(\bar{T})$ is the zero-coupon rate with a constant time to maturity $\bar{T} = T - t$

The carry risk premium

Bond carry

Implementation

We notice that the difference is higher for long maturities. However, the risk associated with such a strategy is that of a rise in interest rates. This is why this carry strategy is generally implemented by using short-term maturities (less than two years)

The carry risk premium
Bond carry

2. The second carry strategy ("*carry slope*") corresponds to a long position in the bond with maturity T_2 and a short position in the bond with maturity T_1
   - The exposure of the two legs are adjusted in order to obtain a duration-neutral portfolio
   - This strategy is known as the slope carry trade
   - We have:

$$\mathcal{C}_t = \underbrace{\left(R_t\left(T_2\right) - r_t\right) - \frac{D_2\left(T_1\right)}{D_t\left(T_1\right)}\left(R_t\left(T_1\right) - r_t\right)}_{\text{duration neutral slope}} + \underbrace{\partial_{\bar{T}}\,\bar{R}_t\left(\bar{T}_2\right) - \frac{D_2\left(T_1\right)}{D_t\left(T_1\right)}\partial_{\bar{T}}\,\bar{R}_t\left(\bar{T}_1\right)}_{\text{duration neutral roll down}}$$

The carry risk premium

Bond carry

Implementation

The classical carry strategy is long 10Y/short 2Y

The carry risk premium

Bond carry

3. The third carry strategy ("*cross-carry slope*") is a variant of the second carry strategy when we consider the yield curves of several countries

Implementation

The portfolio is long on positive or higher slope carry and short on negative or lower slope carry

The carry risk premium

Credit carry

We consider a long position on a corporate bond and a short position on the government bond with the same duration

The carry is equal to:

$$\mathcal{C}_t = \underbrace{\mathcal{s}_t(T)}_{\text{spread}} + \underbrace{\partial_{\bar{T}} \bar{R}_t^{\star}(\bar{T}) - \partial_{\bar{T}} \bar{R}_t(\bar{T})}_{\text{roll down difference}}$$

where $\mathcal{s}_t(T) = R_t^{\star}(T) - R_t(T)$ is the credit spread, $R_t^{\star}(T)$ is the yield-to-maturity of the credit bond and $R_t^{\star}(T)$ is the yield-to-maturity of the government bond

The carry risk premium
Credit carry

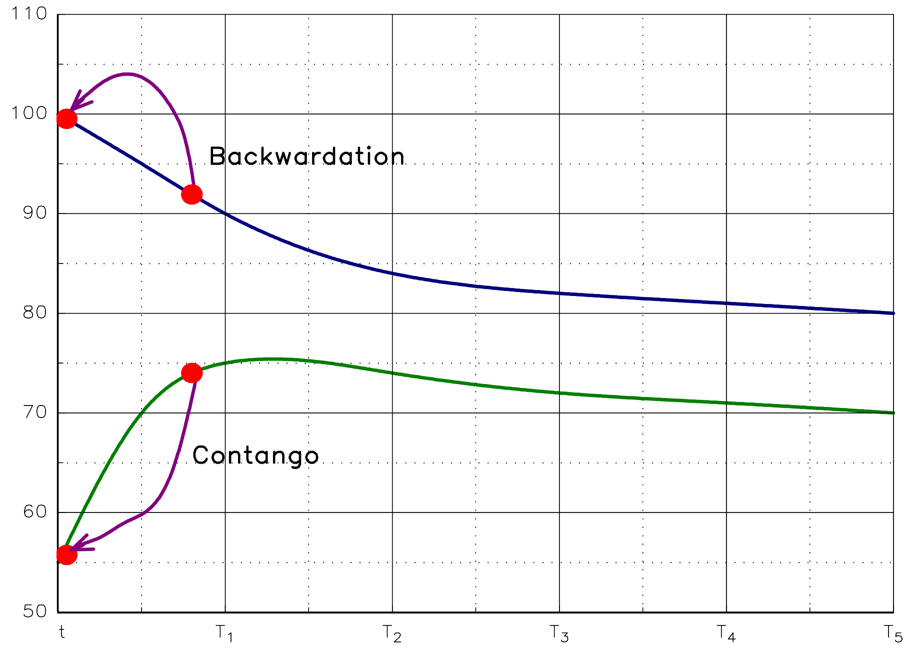
Two implementations

1. The first one is to build a long/short portfolio with corporate bond indices or baskets. The bond universe can be investment grade or high yield. In the case of HY bonds, the short exposure can be an IG bond index
2. The second approach consists in using credit default swaps (CDS). Typically, we sell credit protection on HY credit default indices (e.g. CDX.NA.HY) and buy protection on IG credit default indices (e.g. CDX.NA.IG)

The carry risk premium

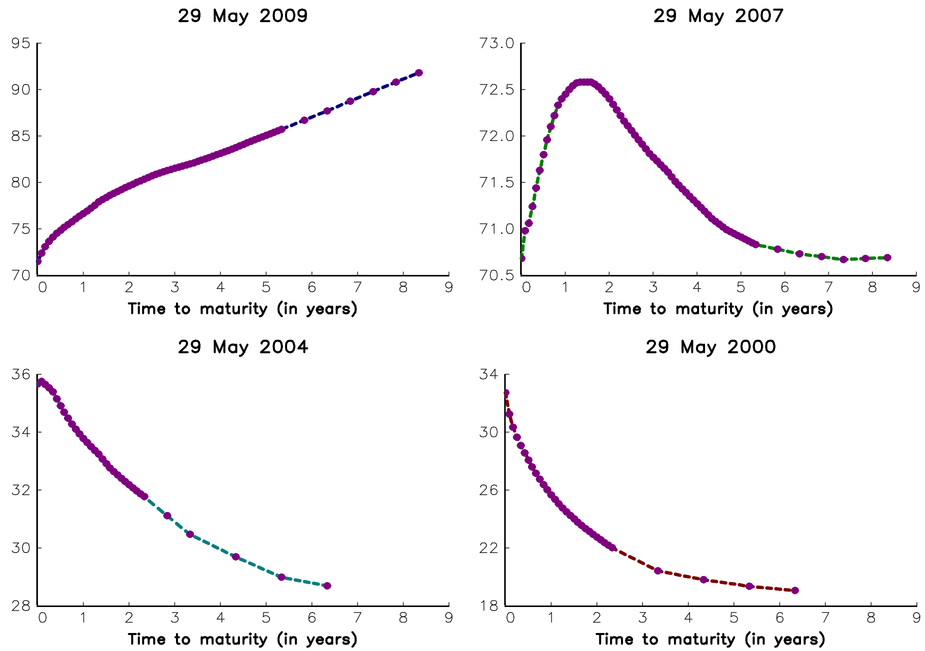
Commodity carry

Figure 60: Contango and backwardation movements in commodity futures contracts



Source: Roncalli (2013)

Figure 61: Term structure of crude oil futures contracts



Source: Roncalli (2013)

The carry risk premium

Volatility carry (or the short volatility strategy)

Volatility carry risk premium

- Long volatility $\Rightarrow$ negative carry ($\neq$ structural exposure)
- Short volatility $\Rightarrow$ positive carry, but the highest skewness risk
- The P&L of selling and delta-hedging an option is equal to:

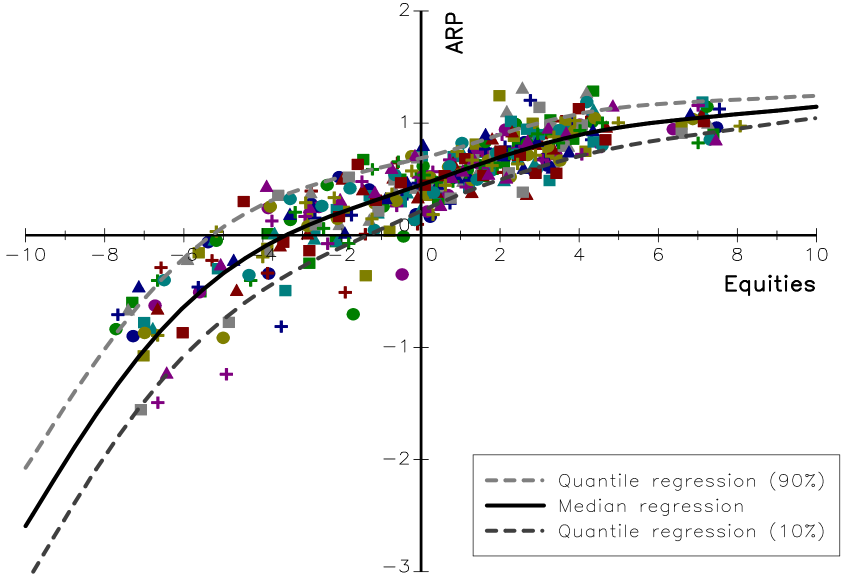
$$\Pi = \frac{1}{2} \int_0^T e^{r(T-t)} S_t^2 \Gamma_t \left(\Sigma_t^2 - \sigma_t^2 \right) \mathrm{d}t$$

  where S_t is the price of the underlying asset, Γ_t is the gamma coefficient, Σ_t is the implied volatility and σ_t is the realized volatility
- $\Sigma_t \geq \sigma_t \Longrightarrow \Pi > 0$
- 3 main reasons:
  1. Asymmetric risk profile between the seller and the buyer
  2. Hedging demand imbalances
  3. Liquidity preferences

The carry risk premium

Volatility carry (or the short volatility strategy)

Figure 62: Non-parametric payoff of the US short volatility strategy



- Income generation
- Short put option profile
- Strategic asset allocation ($\neq$ tactical asset allocation)
- Time horizon is crucial!

It is a skewness risk premium!

Carry strategies exhibit concave payoffs

The value risk premium

Definition

- Let $S_{i,t}$ be the market price of Asset i
- Let $S_i^\star$ be the fundamental price (or the fair value) of Asset i
- The value of Asset i is the relative difference between the two prices:

$$\mathcal{V}_{i,t} = \frac{S_i^\star - S_{i,t}}{S_{i,t}}$$

- The value investor will prefer Asset i to Asset j if the value of Asset i is higher:

$$\mathcal{V}_{i,t} \geq \mathcal{V}_{j,t} \Longrightarrow A_i \succ A_j$$

The value risk premium

The value strategy is an active management bet

- The price of Asset i is undervalued if and only if its value is negative:

$$\mathcal{V}_{i,t} \leq 0 \Leftrightarrow S_i^\star \leq S_{i,t}$$

  The value investor should sell securities with negative values

- The price of Asset i is overvalued if and only if its value is positive:

$$\mathcal{V}_{i,t} \geq 0 \Leftrightarrow S_i^\star \geq S_{i,t}$$

  The value investor should buy securities with positive values

Remark

While carry is an **objective** measure, value is a **subjective** measure, because the fair value is different from one investor to another (e.g. stock picking = value strategy)

The value risk premium

Computing the fair value

We need a model to estimate the fundamental price $S_i^{\star}$:

- Stocks: discounted cash flow (DCF) method, fundamental measures (B2P, PE, DY, EBITDA/EV, etc.), machine learning model, etc.
- Sovereign bonds: macroeconomic model, flows model, etc.
- Corporate bonds: Merton model, structural model, econometric model, etc.
- Foreign exchange rates: purchasing power parity (PPP), real effective exchange rate (REER), BEER, FEER, NATREX, etc.
- Commodities: statistical model (5-year average price), etc.

The value risk premium

Equity value

The equity strategy

If we assume that the weight of asset i is proportional to its book-to-price:

$$w_{i,t} \propto \frac{B_{i,t}}{P_{i,t}}$$

We obtain:

$$w_{i,t} = \underbrace{B_{i,t} \Big/ \sum_{j=1}^{n} B_{j,t}}_{\text{Fundamental component}} \times \underbrace{\sum_{j=1}^{n} P_{j,t} \Big/ P_{i,t}}_{\text{Reversal component}} \times \underbrace{\text{a cross-effect term}}_{\simeq \text{ constant}}$$

The value risk factor can be decomposed into two main components:

- a fundamental indexation pattern
- a reversal-based pattern

$\Rightarrow$ Reversal strategies $\approx$ value strategies

The value risk premium

Equity value

- In equities, the frequency of the reversal pattern is ≤ 1 month or $\geq$ 18 months
- In currencies and commodities, the frequency of the reversal pattern is very short (one or two weeks) or very long (≥ 3 years)

$\Rightarrow$ Value strategy in currencies and commodities?

The value risk premium

The payoff of the equity value risk premium

- We consider two Eurozone Value indices calculated by the same index sponsor
- The index sponsor uses the same stock selection process
- The index sponsor uses two different weighting schemes:
  - The first index considers a capitalization-weighted portfolio
  - The second index considers a minimum variance portfolio

$\Rightarrow$ We recall that the payoff of the low-volatility strategy is long put + short call

The value risk premium

The payoff of the equity value risk premium

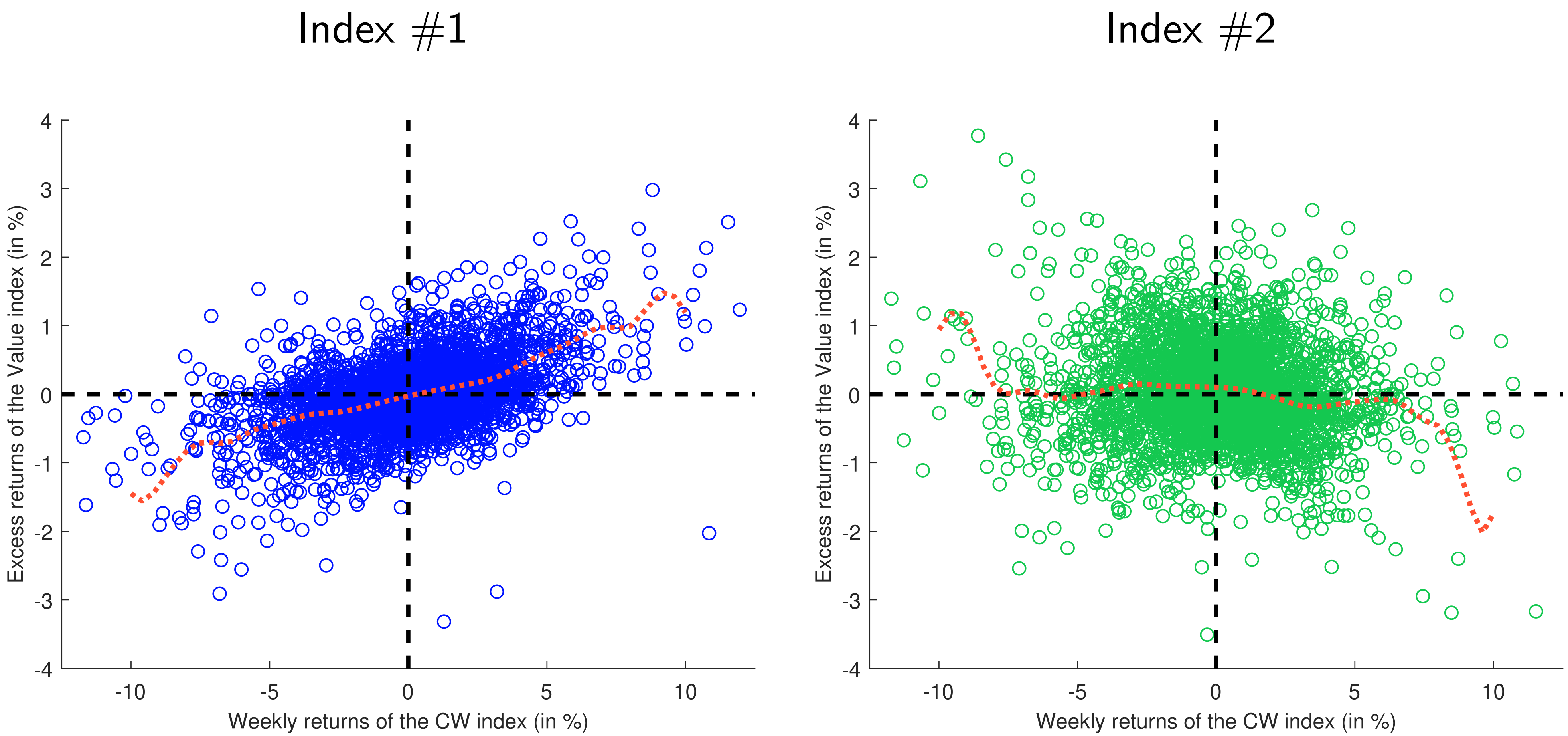


Figure 63: Which Eurozone value index has the right payoff?

The value risk premium

The payoff of the equity value risk premium

Answer

The payoff of the equity value risk premium is:

Short Put + **Long Call**

⇒ It is a skewness risk premium too!

- The design of the strategy is crucial (some weighting schemes may change or destroy the desired payoff!)
- Are the previous results valid for other asset classes, e.g. rates or currencies?

The value risk premium

Misunderstanding of the equity value risk premium

The dot-com crisis (2000-2003)

If we consider the S&P 500 index, we obtain:

- 55% of stocks post a negative performance

$$\approx 75\% \text{ of MC}$$

- 45% of stocks post a positive performance

$$\text{Maximum drawdown} = 49\ \%$$

Small caps stocks $\nearrow$
Value stocks $\nearrow$

The GFC crisis (2008)

If we consider the S&P 500 index, we obtain:

- 95% of stocks post a negative performance

$$\approx 97\% \text{ of MC}$$

- 5% of stocks post a positive performance

$$\text{Maximum drawdown} = 56\ \%$$

Small caps stocks $\searrow$
Value stocks $\searrow$

What is the impact of the liquidity risk premium?

The value risk premium

Extension to other asset classes

- Corporate bonds
  - Houweling and van Zundert (2017)
  - Ben Slimane *et al.* (2019)
  - Roncalli (2020)
- Currencies
  - MacDonald (1995)
  - Menkhoff *et al.* (2016)
  - Baku *et al.* (2019, 2020)

The momentum risk premium

Definition

- Let $S_{i,t}$ be the market price of Asset i
- We assume that:

$$\frac{\mathrm{d}S_{i,t}}{S_{i,t}} = \mu_{i,t}\,\mathrm{d}t + \sigma_{i,t}\,\mathrm{d}W_{i,t}$$

- The momentum of Asset i corresponds to its past trend:

$$\mathcal{M}_{i,t} = \hat{\mu}_{i,t}$$

- The momentum investor will prefer Asset i to Asset j if the momentum of Asset i is higher:

$$\mathcal{M}_{i,t} \geq \mathcal{M}_{j,t} \Longrightarrow A_i \succ A_j$$

The momentum risk premium

Computing the momentum measure

- Past return (e.g. one-month, three-month, one-year, etc.)

$$\mathcal{M}_{i,t} = \frac{S_{i,t} - S_{i,t-h}}{S_{i,t-h}}$$

- Lagged past return[13]
- Econometric and statistical trend estimators (see Bruder *et al.* (2011) for a survey)

[13] For instance, the WML risk factor is generally implemented using the one-month lag of the twelve-month return:

$$\mathcal{M}_{i,t} = \frac{S_{i,t-1\mathrm{M}} - S_{i,t-13\mathrm{M}}}{S_{i,t-13\mathrm{M}}}$$

because the stock market is reversal within a one-month time horizon

The momentum risk premium

Three momentum strategies

1. Cross-section momentum (CSM)

$$\mathcal{M}_{i,t} \geq \mathcal{M}_{j,t} \Longrightarrow A_i \succ A_j$$

2. Time-series momentum (TSM)

$$\mathcal{M}_{i,t} > 0 \Longrightarrow A_i \succ 0 \text{ and } \mathcal{M}_{i,t} < 0 \Longrightarrow A_i \prec 0$$

3. Reversal strategy:

$$\mathcal{M}_{i,t} \geq \mathcal{M}_{j,t} \Longrightarrow A_i \prec A_j$$

Remark

Generally, the momentum risk premium corresponds to the CSM or TSM strategies. When we speak about momentum strategies, we can also include reversal strategies, which are more considered as trading strategies with high turnover ratios and very short holding periods (generally intra-day or daily frequency, less than one week most of the time)

The momentum risk premium

Cross-section versus time-series

Time-series momentum (TSM)

- The portfolio is long (resp. short) on the asset if it has a positive (resp. negative) momentum
- This strategy is also called "trend-following" or "trend-continuation"
- HF: CTA and managed futures
- Between asset classes

Cross-section momentum (CSM)

- The portfolio is long (resp. short) on assets that present a momentum higher (resp. lower) than the others
- This strategy is also called "winners minus losers" (or WML) by Carhart (1997)
- Within an asset class (equities, currencies)

$\Rightarrow$ These two momentum risk premia are very different and not well understood!

The momentum risk premium

Understanding the TSM strategy

Some results (Jusselin *et al.*, 2017)

- EWMA is the optimal trend estimator (Kalman-Bucy filtering)
- Two components
  - a short-term component given by the payoff (dynamics)
  - a long-term component given by the trading impact (performance)
- Main important parameters
  - The Sharpe ratio
  - The duration of the moving average
  - The correlation matrix
  - The term structure of the volatility
- Too much leverage kills momentum (high ruin probability)

The momentum risk premium

Understanding the TSM strategy

Some results (Jusselin *et al.*, 2017)

- The issue of diversification
  - Time-series momentum likes zero-correlated assets (e.g. multi-asset momentum premium)
  - Cross-section momentum likes highly correlated assets (e.g. equity momentum factor)
  - The number of assets decreases the P&L dispersion
  - The symmetry puzzle
  - The n/ρ trade-off
- Short-term versus long-term momentum
  - Short-term momentum is more risky than long-term momentum
  - The Sharpe ratio of long-term momentum is higher
  - The choice of the EWMA duration is more crucial for long-term momentum

The momentum risk premium

Understanding the TSM strategy

Some results (Jusselin *et al.*, 2017)

- The momentum strategy outperforms the buy-and-hold strategy when the Sharpe ratio is lower than 35%
- The specific nature of equities and bonds
  - Performance of equity momentum is explained by leverage patterns
  - Performance of bond momentum is explained by frequency patterns
- A lot of myths about the performance of CTAs (equity contribution, option profile, hedging properties)
- Momentum strategies are not alpha or absolute return strategies, but diversification strategies

The momentum risk premium

Trend-following strategies (or TSM) exhibit a convex payoff

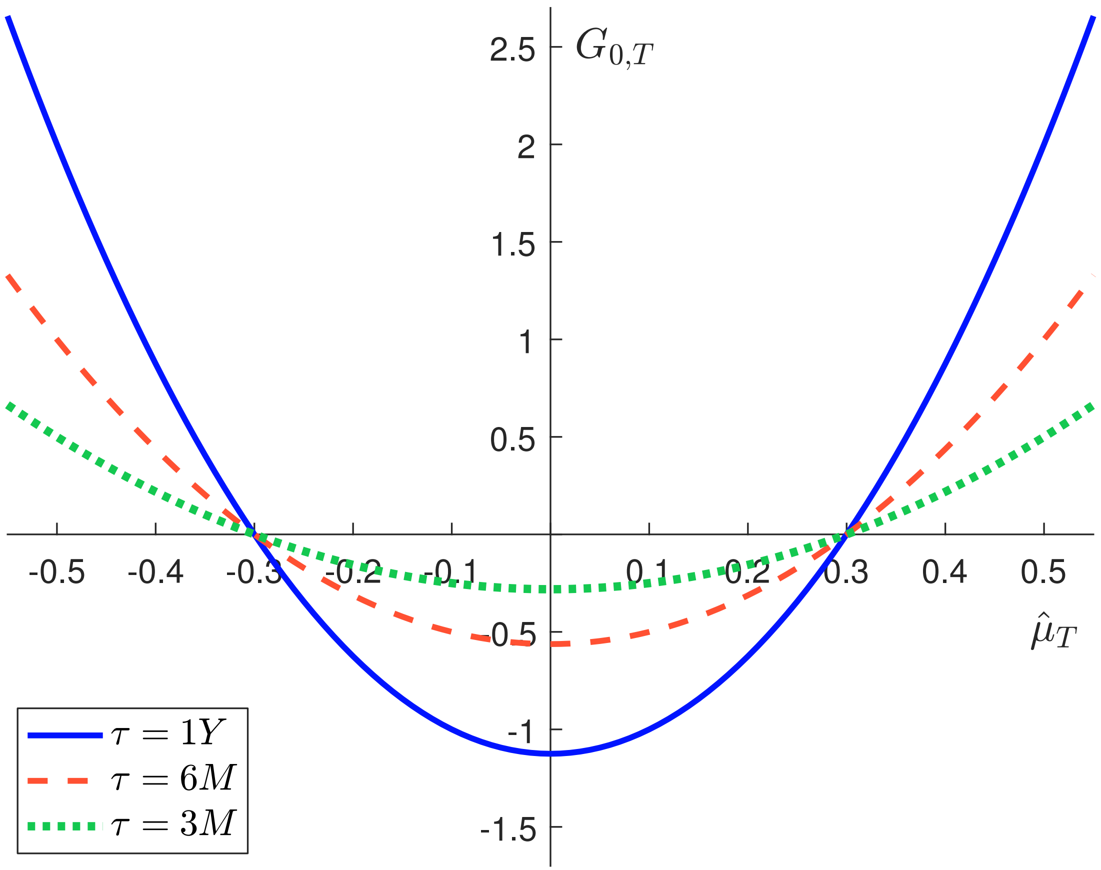


Figure 64: Option profile of the trend-following strategy

- λ is the parameter of the EWMA estimator
- $\tau = 1/\lambda$ is the duration of the EWMA estimator
- Market anomaly: the systematic risk is limited in bad times
- **Trend-following strategies exhibit a convex payoff**

The momentum risk premium

The loss of a trend-following strategy is bounded

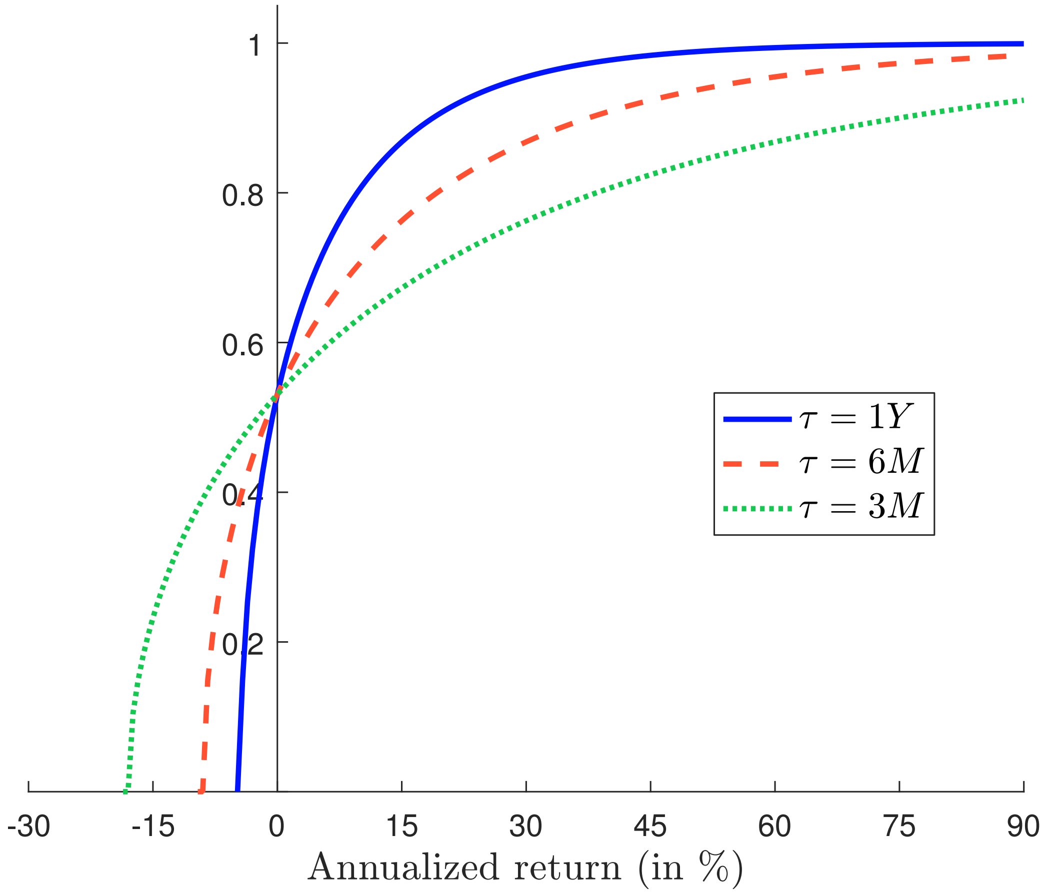


Figure 65: Cumulative distribution function of g_t ($s_t = 0$)

- s_t is the Sharpe ratio
- g_t is the trading impact
- **The loss is bounded**
- The gain may be infinite
- The return variance of short-term momentum strategies is larger than the return variance of long-term momentum strategies
- The skewness is positive

The momentum risk premium

Trend-following strategies exhibit positive skewness

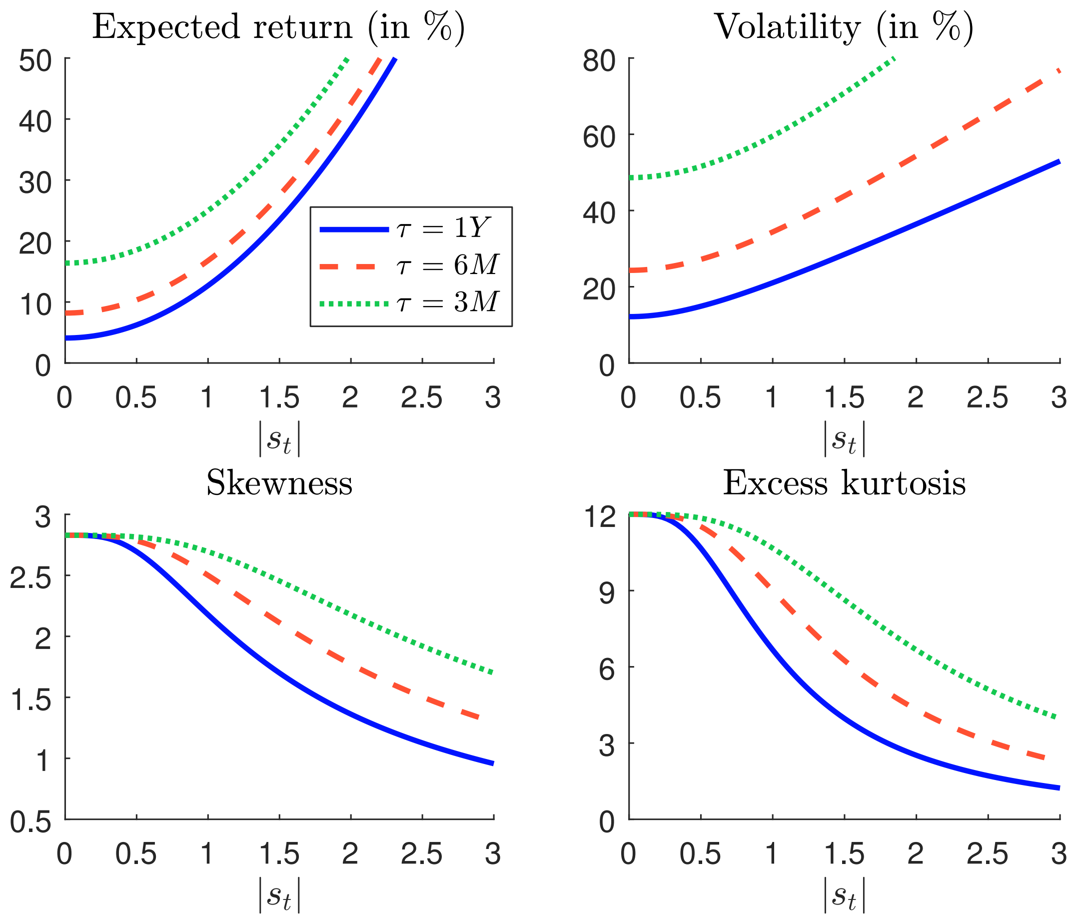


Figure 66: Statistical moments of the momentum strategy

- Short-term trend-following strategies are more risky than long-term trend-following strategies
- The skewness is positive
- It is a market anomaly, not a skewness risk premium

The momentum risk premium

Short-term versus long-term trend-following strategies

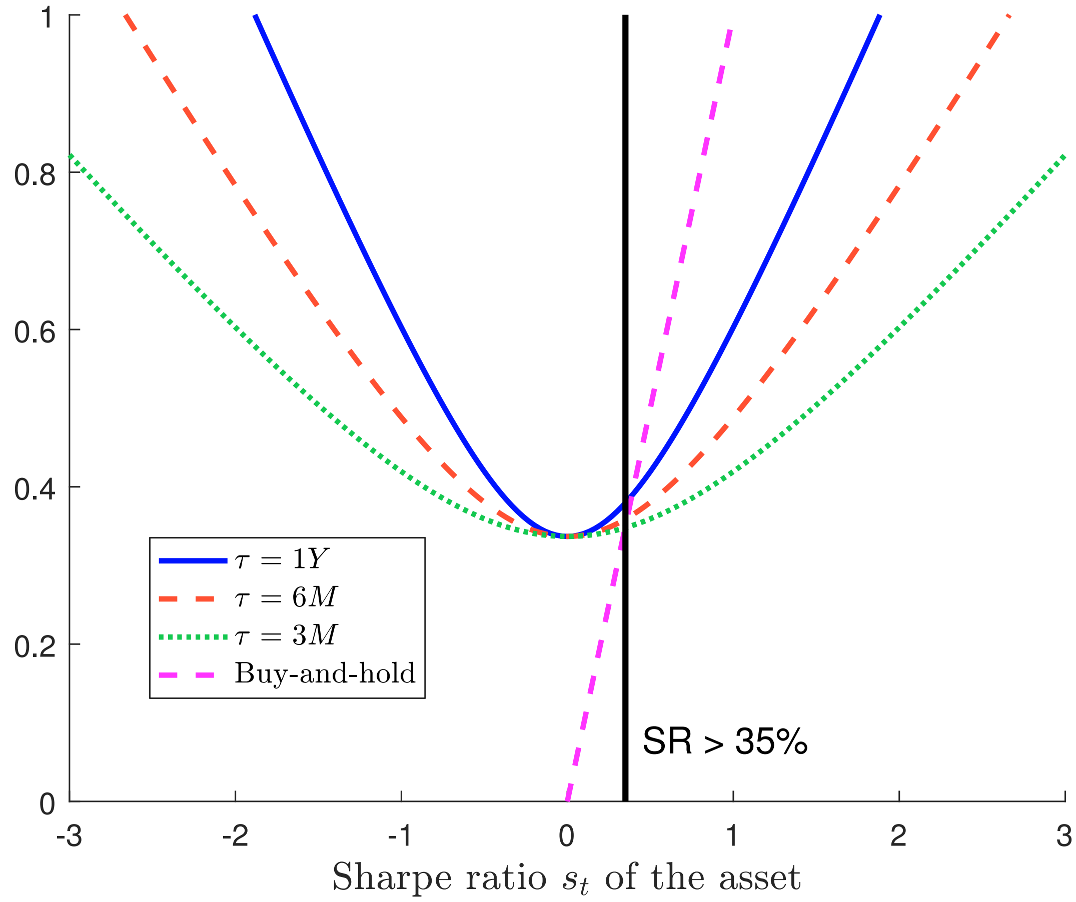


Figure 67: Sharpe ratio of the momentum strategy

- When the Sharpe ratio of the underlying is lower than 35%, the momentum strategy dominates the buy-and-hold strategy
- The Sharpe ratio of long-term momentum strategies is higher than the Sharpe ratio of short-term momentum strategies

The momentum risk premium

Relationship with the Black-Scholes robustness

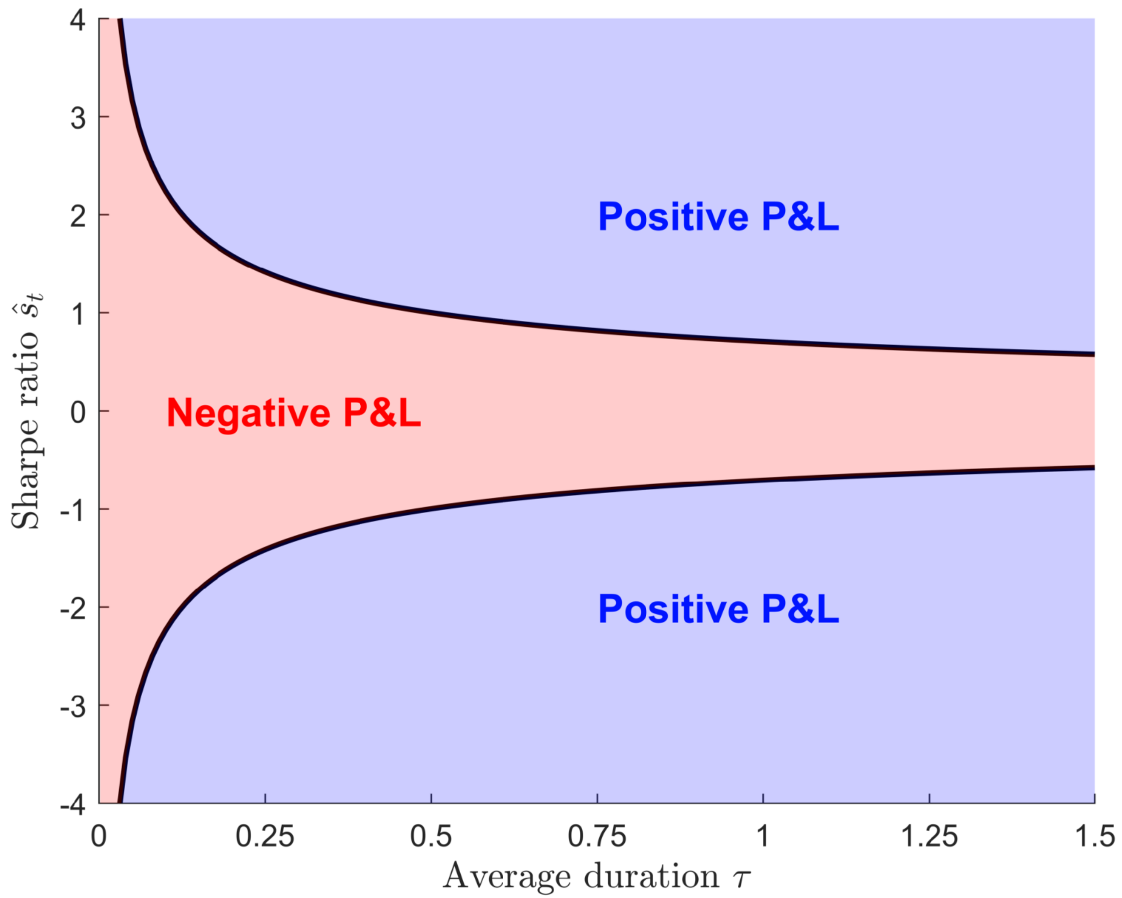


Figure 68: Admissible region for positive P&L

- Delta-hedging: implied volatility vs realized volatility
- Trend-following: duration vs realized Sharpe ratio
- The critical value for the Sharpe ratio is 1.41 for 3M and 0.71 for 1Y

The momentum risk premium

Impact of the correlation on trend-following strategies

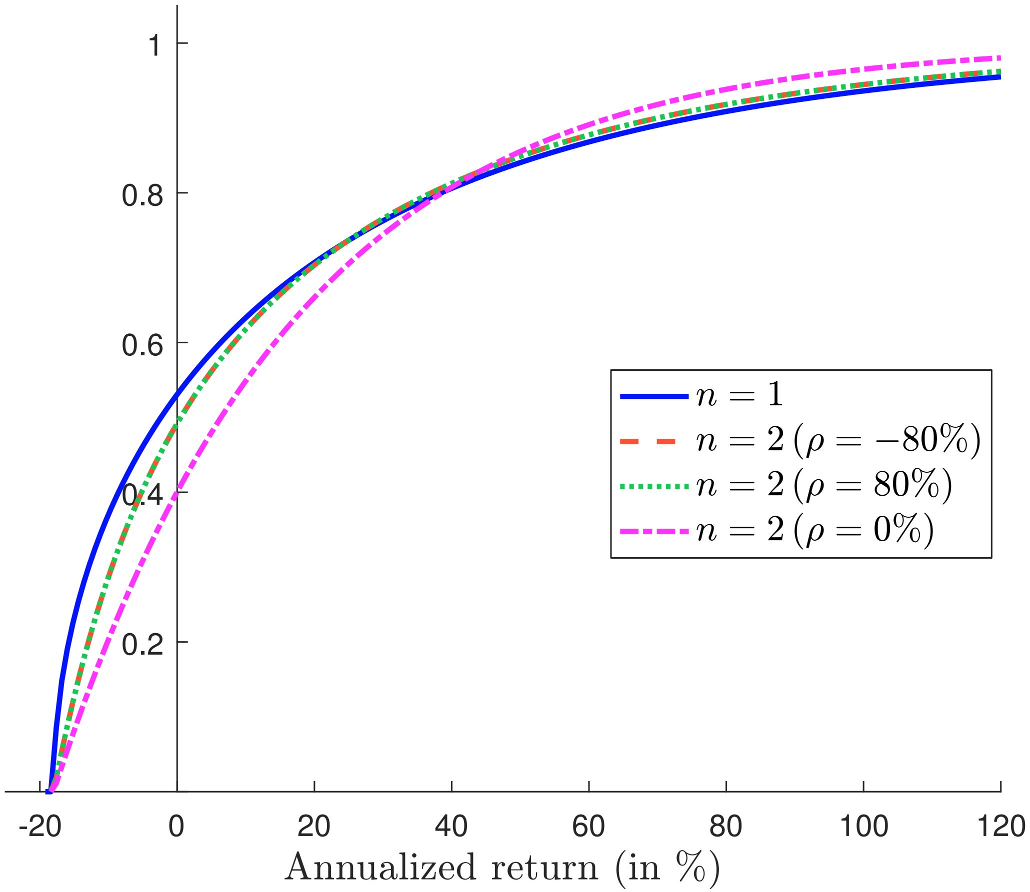


Figure 69: Cumulative distribution function of g_t ($s_t = 0$)

- Sign of correlation does not matter when the Sharpe ratio of assets is zero
- Symmetry puzzle

positive correlation
=
negative correlation

The momentum risk premium

Correlation and diversification

Long-only versus long/short diversification

We consider a portfolio (α_1, α_2) composed of two assets. We have:

$$\sigma(\rho) = \sqrt{\alpha_1^2\sigma_1^2 + 2\rho\alpha_1\alpha_2\sigma_1\sigma_2 + \alpha_2^2\sigma_2^2}$$

- In the case of a long-only portfolio, the best case for diversification is reached when the correlation is equal to -1:

$$|\alpha_1\sigma_1 - \alpha_2\sigma_2| = \sigma(-1) \leq \sigma(\rho) \leq \sigma(1) = \alpha_1\sigma_1 + \alpha_2\sigma_2$$

- In the case of a long/short portfolio, we generally have $\operatorname{sgn}(\alpha_1\alpha_2) = \operatorname{sgn}(\rho)$. Therefore, the best case for diversification is reached when the correlation is equal to zero: $\sigma(0) \leq \sigma(\rho)$. Indeed, when the correlation is -1, the investor is long on one asset and short on the other asset, implying that this is the same bet.

The momentum risk premium

The number of assets/correlation trade-off

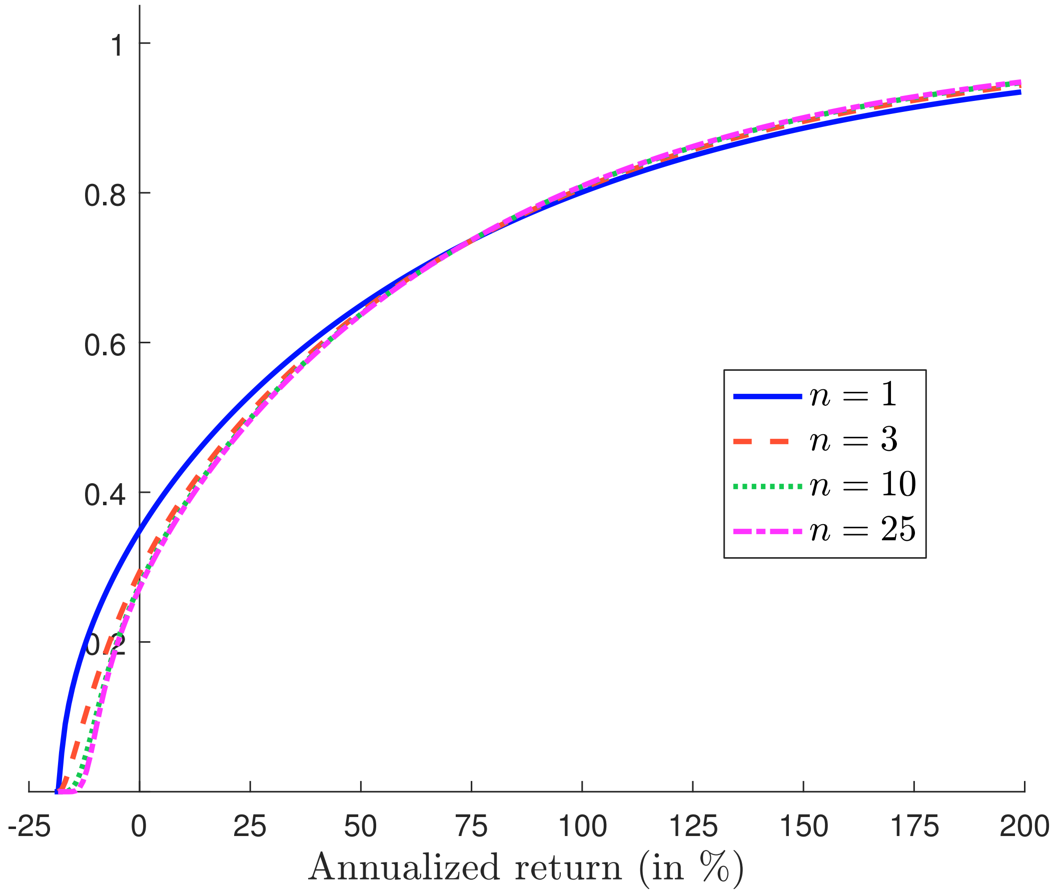


Figure 70: Impact of the number of assets on $\Pr\{g_t \leq g\}$ ($s_t = 2$, $\rho = 80\%$)

- **Correlation is not the friend of time-series momentum**
- A momentum strategy prefers a few number of assets with high Sharpe ratio absolute values than a large number of assets with low Sharpe ratio absolute values

The momentum risk premium

TSM versus CSM

Time-series momentum

- Absolute trends

$$\begin{cases} \hat{\mu}_{i,t} \geq 0 \Rightarrow e_{i,t} \geq 0 \\ \hat{\mu}_{i,t} < 0 \Rightarrow e_{i,t} < 0 \end{cases}$$

- CTA hedge funds
- Alternative risk premia in multi-asset portfolios

Cross-section momentum

- Relative trends

$$\begin{cases} \hat{\mu}_{i,t} \geq \bar{\mu}_t \Rightarrow e_{i,t} \geq 0 \\ \hat{\mu}_{i,t} < \bar{\mu}_t \Rightarrow e_{i,t} < 0 \end{cases}$$

where:

$$\bar{\mu}_t = \frac{1}{n} \sum_{j=1}^{n} \hat{\mu}_{j,t}$$

- Statistical arbitrage / relative value
- Factor investing in equity portfolios

Beta strategy or Alpha strategy?

The momentum risk premium

Performance of cross-section momentum risk premium

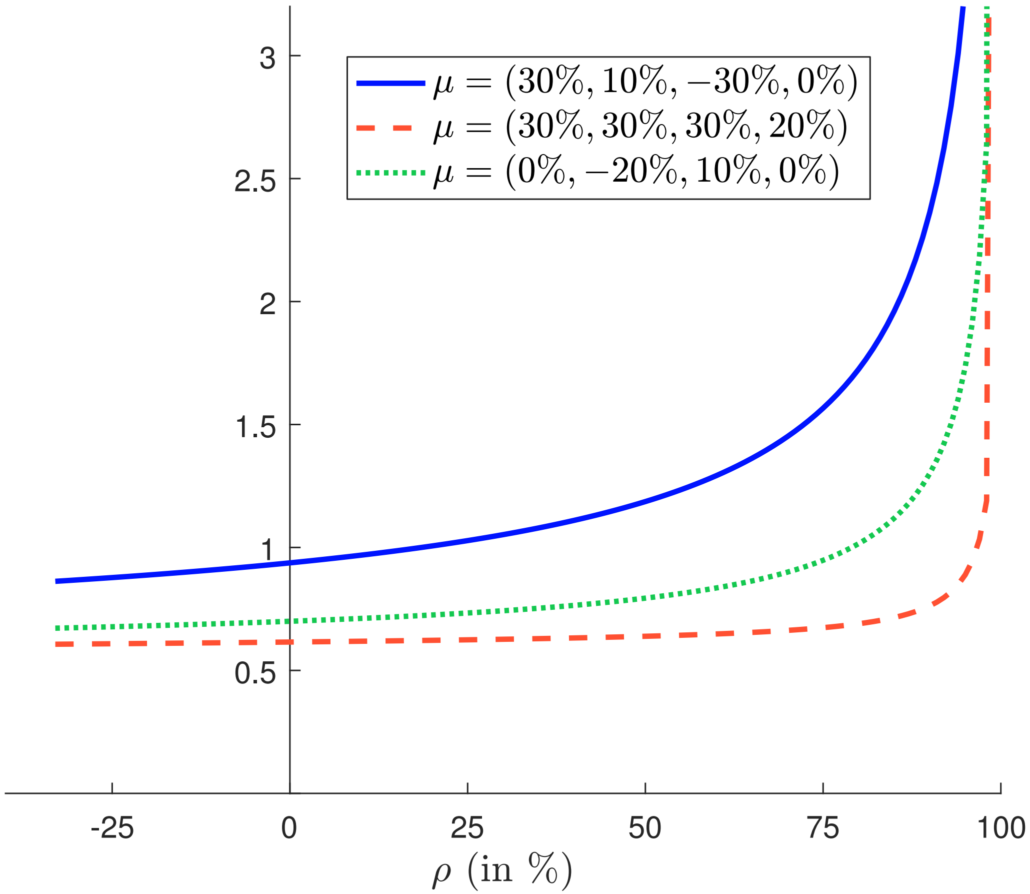


Figure 71: Sharpe ratio of the CSM strategy

- **Correlation is the friend of cross-section momentum!**
- Statistical arbitrage / relative value

The momentum risk premium

Naive replication of the SG CTA Index

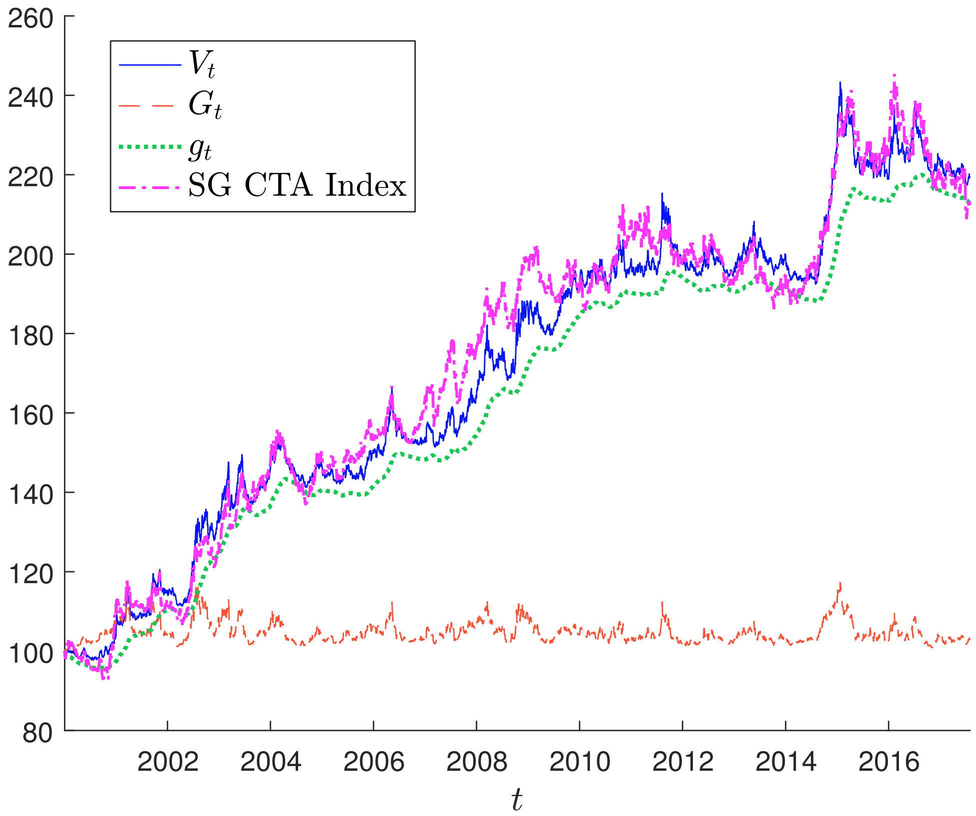


Figure 72: Comparison between the cumulative performance of the naive replication strategy and the SG CTA Index

- The performance of trend-followers comes from the trading impact
- Currencies and commodities are the main contributors!
- Mixing asset classes is the key point in order to capture the diversification premium

The momentum risk premium

Trend-following strategies benefit from traditional risk premia

Table 61: Exposure average of the trend-following strategy (in %)

Asset Class	Average Exposure	Short Exposure	Long Exposure	Short Frequency	Long Frequency
Bond	58%	−100%	122%	29%	**71%**
Equity	52%	−88%	**160%**	44%	56%
Currency	**18%**	−103%	115%	45%	55%
Commodity	**23%**	−108%	113%	41%	59%

- The specific nature of bonds: long exposure frequency $>$ short exposure frequency; long leverage $\approx$ short leverage
- The specific nature of equities: short exposure frequency $\approx$ long exposure frequency; long leverage $>$ short leverage

The momentum risk premium

The myth of short selling

- Equity and bond momentum strategies benefit from the existence of a risk premium
- Currency and commodity momentum strategies benefit from (positive / negative) trend patterns
- Leverage management $\succ$ short management
- The case of equities in the 2008 GFC, the stock-bond correlation and the symmetry puzzle

The good performance of CTAs in 2008 is not explained by their short exposure in equities, but by their long exposure in bonds

The momentum risk premium

The reversal strategy

- The reversal strategy may be defined as the opposite of the momentum strategy (CSM or TSM)
- It is also known as the mean-reverting strategy

How to reconciliate reversal and trend-following strategies?

Because they don't use the same trend windows and holding periods[14]

[14] Generally, reversal strategies use short-term or very long-term trends while trend-following strategies use medium-term trends

The momentum risk premium

The reversal strategy

The mean-reverting (or autocorrelation) strategy

- Let $R_{i,t} = \ln S_{i,t} - \ln S_{i,t-1}$ be the one-period return
- We note $\rho_i(h) = \rho(R_{i,t}, R_{i,t-h})$ the autocorrelation function
- Asset i exhibits a mean-reverting pattern if the short-term autocorrelation $\rho_i(1)$ is negative
- In this case, the short-term reversal is defined by the product of the autocorrelation and the current return:

$$\mathcal{R}_{i,t} = \rho_i(1) \cdot R_{i,t}$$

- The short-term reversal strategy is then defined by the following rule:

$$\mathcal{R}_{i,t} \geq \mathcal{R}_{j,t} \Longrightarrow i \succ j$$

The momentum risk premium

The reversal strategy

First implementation of the autocorrelation strategy

- If $\mathcal{R}_{i,t}$ is positive, meaning that the current return $R_{i,t}$ is negative, we should buy the asset, because a negative return is followed by a positive return on average
- If $\mathcal{R}_{i,t}$ is negative, meaning that the current return $R_{i,t}$ is positive, we should sell the asset, because a positive return is followed by a negative return on average

The momentum risk premium

The reversal strategy

The variance swap strategy

- We assume that the one-period asset return follows an AR(1) process:

$$R_{i,t} = \rho R_{i,t-1} + \varepsilon_t$$

  where $|\rho| < 1$, $\varepsilon_t \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$ and $\operatorname{cov}\left(\varepsilon_t, \varepsilon_{t-j}\right) = 0$ for $j \geq 1$
- Let $\mathrm{RV}\left(h\right)$ be the annualized realized variance of the h-period asset return $R_{i,t}\left(h\right) = \ln S_{i,t} - \ln S_{i,t-h}$
- Hamdan *et al.* (2016) showed that:

$$\mathbb{E}\left[\mathrm{RV}\left(h\right)\right] = \phi\left(h\right) \mathbb{E}\left[\mathrm{RV}\left(1\right)\right]$$

  where:

$$\phi\left(h\right) = 1 + 2\rho \frac{1 - \rho^{h-1}}{1 - \rho} - 2 \sum_{j=1}^{h-1} \frac{j}{h} \rho^j$$

The momentum risk premium

The reversal strategy

The variance swap strategy

- We notice that:

$$\lim_{h\to\infty} \mathbb{E}\left[\mathrm{RV}\left(h\right)\right] = \left(1 + \frac{2\rho}{1-\rho}\right) \cdot \mathbb{E}\left[\mathrm{RV}\left(1\right)\right]$$

- When the autocorrelation is negative, this implies that the long-term frequency variance is lower than the short-term frequency variance
- More generally, we have:

$$\left\{ \begin{array}{ll} \mathbb{E}\left[\mathrm{RV}\left(h\right)\right] < \mathbb{E}\left[\mathrm{RV}\left(1\right)\right] & \text{if } \rho < 0 \\ \mathbb{E}\left[\mathrm{RV}\left(h\right)\right] \geq \mathbb{E}\left[\mathrm{RV}\left(1\right)\right] & \text{otherwise} \end{array} \right.$$

The momentum risk premium

The reversal strategy

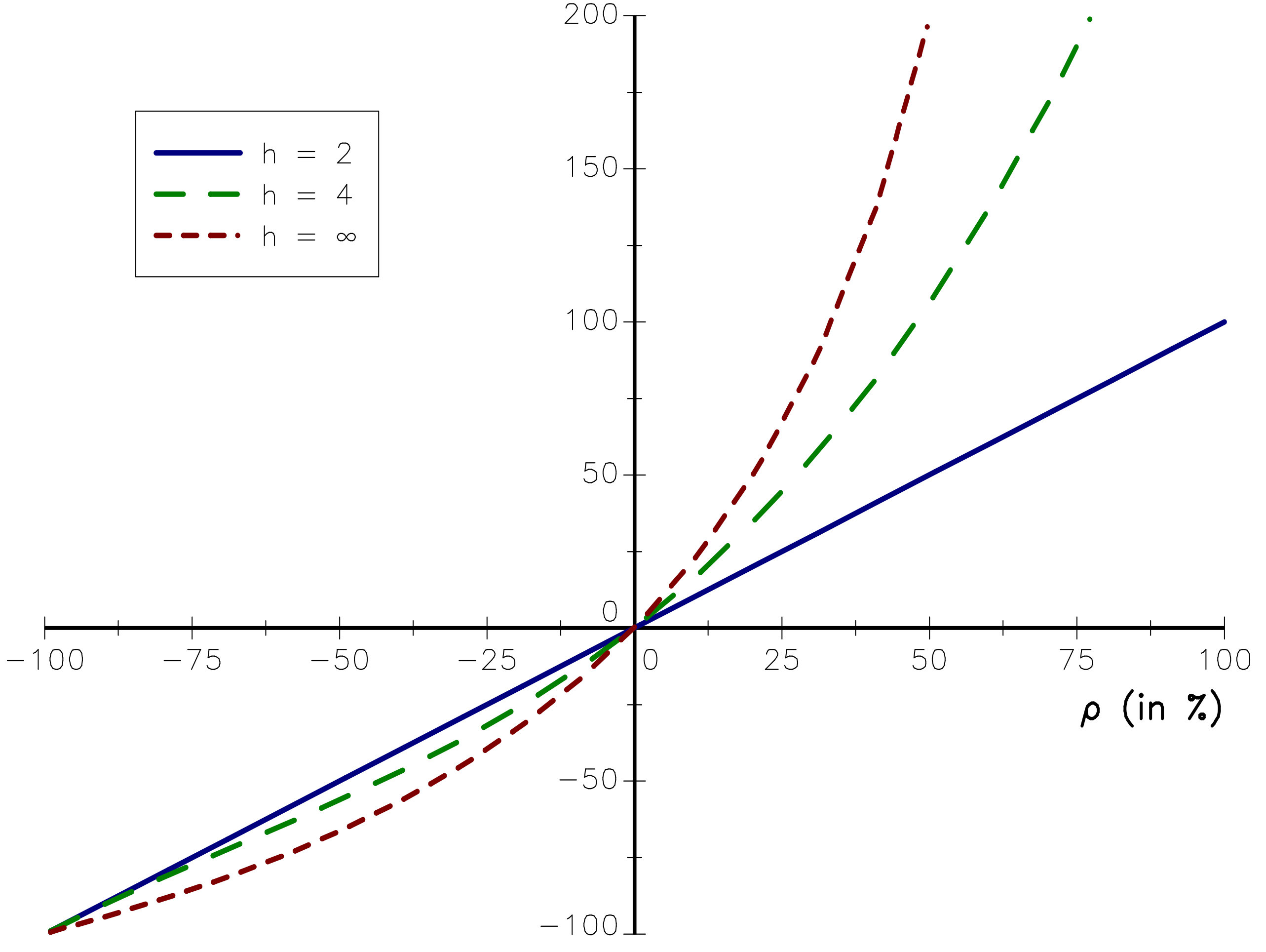


Figure 73: Variance ratio $(\mathrm{RV}(h) - \mathrm{RV}(1)) / \mathrm{RV}(1)$ (in %)

The momentum risk premium

The reversal strategy

Second implementation of the autocorrelation strategy

- The spread between daily/weekly and weekly/monthly variance swaps depends on the autocorrelation of daily returns
- The reversal strategy consists in being long on the daily/weekly variance swaps and short on the weekly/monthly variance swaps

The momentum risk premium

The reversal strategy

The long-term reversal strategy

- The long-term return reversal is defined by the difference between long-run and short-period average prices:

$$\mathcal{R}_{i,t} = \bar{S}_{i,t}^{LT} - \bar{S}_{i,t}^{ST}$$

- Typically, $\bar{S}_{i,t}^{ST}$ is the average price over the last year and $\bar{S}_{i,t}^{LT}$ is the average price over the last five years
- The long-term return reversal strategy follows the same rule as the short-term reversal strategy
- This reversal strategy is equivalent to a value strategy because the long-run average price can be viewed as an estimate of the fundamental price in some asset classes

The momentum risk premium

The reversal strategy

Implementation of the long-term reversal strategy

- If $\mathcal{R}_{i,t}$ is positive, the long-term mean of the asset price is above its short-term mean $\Rightarrow$ we should buy the asset
- If $\mathcal{R}_{i,t}$ is negative, the long-term mean of the asset price is below its short-term mean $\Rightarrow$ we should sell the asset

The liquidity risk premium

What means "*liquidity risk*"?

"[...] there is also broad belief among users of financial liquidity — traders, investors and central bankers — that the principal challenge is not the **average level** *of financial liquidity ... but its* **variability** *and uncertainty" (Persaud, 2003).*

The liquidity risk premium

The liquidity-adjusted CAPM

L-CAPM (Acharya and Pedersen, 2005)

We note L_i the relative (stochastic) illiquidity cost of Asset i. At the equilibrium, we have:

$$\mathbb{E}\left[R_i - L_i\right] - R_f = \tilde{\beta}_i \left(\mathbb{E}\left[R_M - L_M\right] - R_f\right)$$

where:

$$\tilde{\beta}_i = \frac{\operatorname{cov}\left(R_i - L_i, R_M - L_M\right)}{\operatorname{var}\left(R_M - L_M\right)}$$

CAPM in the frictionless economy

$\Downarrow$

CAPM in net returns (including illiquidity costs)

The liquidity risk premium

The liquidity-adjusted CAPM

- The liquidity-adjusted beta can be decomposed into four beta(s):

$$\tilde{\beta}_i = \beta_i + \beta\left(L_{i,}, L_M\right) - \beta\left(R_{i,}, L_M\right) - \beta\left(L_{i,}, R_M\right)$$

  where:

  - $\beta_i = \beta\left(R_{i,} R_M\right)$ is the standard market beta;
  - $\beta\left(L_{i,}, L_M\right)$ is the beta associated to the commonality in liquidity with the market liquidity;
  - $\beta\left(R_{i,}, L_M\right)$ is the beta associated to the return sensitivity to market liquidity;
  - $\beta\left(L_{i,}, R_M\right)$ is the beta associated to the liquidity sensitivity to market returns.

- The risk premium is equal to:

$$\begin{aligned}\pi_i &= \mathbb{E}\left[L_i\right] + \left(\beta_i + \beta\left(L_{i,}, L_M\right)\right)\pi_M - \\ &\quad \left(\tilde{\beta}_i \mathbb{E}\left[L_M\right] + \left(\beta\left(R_{i,}, L_M\right) + \beta\left(L_{i,}, R_M\right)\right)\pi_M\right)\end{aligned}$$

The liquidity risk premium

The liquidity-adjusted CAPM

Acharya and Pedersen (2005)

If assets face some liquidity costs, the relationship between the risk premium and the beta of asset i becomes:

$$\mathbb{E}\left[R_i\right] - R_f = \alpha_i + \beta_i \left(\mathbb{E}\left[R_M\right] - R_f\right)$$

where α_i is a function of the relative liquidity of Asset i with respect to the market portfolio and the liquidity beta(s):

$$\begin{aligned}\alpha_i &= \left(\mathbb{E}\left[L_i\right] - \tilde{\beta}_i \mathbb{E}\left[L_M\right]\right) + \\ & \quad \beta\left(L_{i,}, L_M\right)\pi_M - \beta\left(R_{i,}, L_M\right)\pi_M - \beta\left(L_{i,}, R_M\right)\pi_M\end{aligned}$$

The liquidity risk premium

Disentangling the liquidity alpha

- We deduce that:

$$\alpha_i \neq \mathbb{E}[L_i]$$

meaning that the risk premium of an illiquid asset is not the systematic risk premium plus a premium due the illiquidity level:

$$\mathbb{E}[R_i] - R_f \neq \mathbb{E}[L_i] + \beta_i (\mathbb{E}[R_M] - R_f)$$

- The 4 liquidity premia are highly correlated[15] ($\mathbb{E}[L_i]$, $\beta(L_{i,}, L_M)$, $\beta(R_{i,}, L_M)$ and $\beta(L_{i,}, R_M)$).
- Acharaya and Pedersen (2005) found that $\mathbb{E}[L_i]$ represents 75% of α_i on average. The 25% remaining are mainly explained by the liquidity sensitivity to market returns – $\beta(L_{i,}, R_M)$.

[15] For instance, we have $\rho(\beta(L_{i,}, L_M), \beta(R_{i,}, L_M)) = -57\%$, $\rho(\beta(L_{i,}, L_M), \beta(L_{i,}, R_M)) = -94\%$ and $\rho(\beta(R_{i,}, L_M), \beta(L_{i,}, R_M)) = 73\%$.

The liquidity risk premium

Three liquidity risks

In fact, we have:

$$\alpha_i = \text{illiquidity level} + \text{illiquidity covariance risks}$$

1. $\beta\left(L_{i,}, L_M\right)$
   - An asset that becomes illiquid when the market becomes illiquid should have a higher risk premium
   - Substitution effects when the market becomes illiquid
2. $\beta\left(R_{i,}, L_M\right)$
   - Assets that perform well in times of market illiquidity should have a lower risk premium
   - Relationship with solvency constraints
3. $\beta\left(L_{i,}, R_M\right)$
   - Investors accept a lower risk premium on assets that are liquid in a bear market
   - Selling markets $\neq$ buying markets

The liquidity risk premium

How does market liquidity impact risk premia?

Three main impacts

- Effect on the risk premium
- Effect on the price dynamics
  If liquidity is persistent, negative shock to liquidity implies low current returns and high predicted future returns:
  $$\operatorname{cov}\left(L_{i,t}, R_{i,t}\right) < 0 \text{ and } \partial_{L_{i,t}} \mathbb{E}_t\left[R_{i,t+1}\right] > 0$$
- Effect on portfolio management

- Sovereign bonds
- Corporate bonds
- Stocks
- Small caps
- Private equities

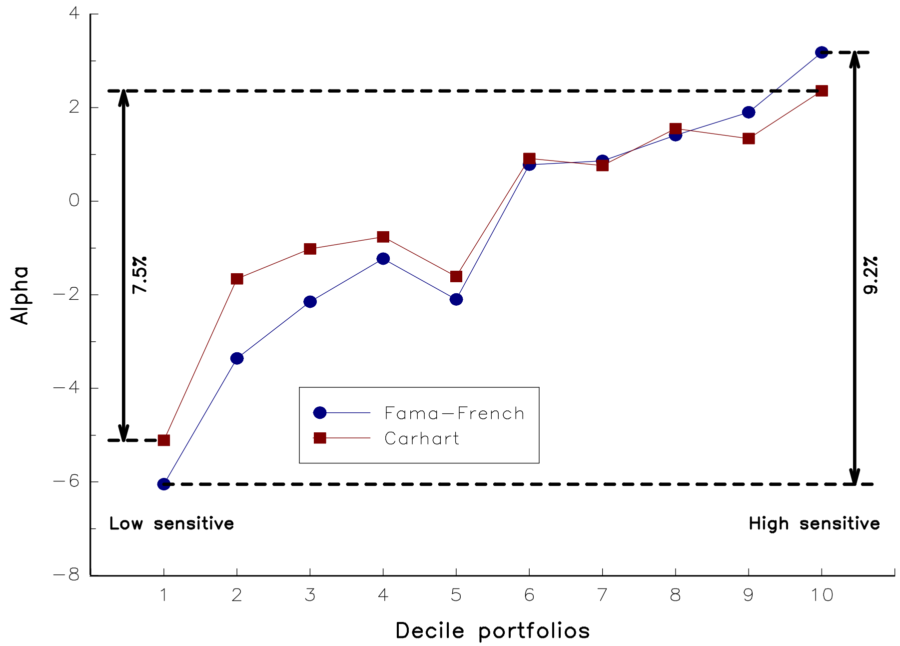
The liquidity risk premium

Application to stocks

Pastor and Stambaugh (2003) include a liquidity premium in the Fama-French-Carhart model:

$$\begin{aligned}\mathbb{E}\left[R_i\right] - R_f &= \beta_i^M \left(\mathbb{E}\left[R_M\right] - R_f\right) + \beta_i^{SMB}\mathbb{E}\left[R_{SMB}\right] + \beta_i^{HML}\mathbb{E}\left[R_{HML}\right] + \\ & \quad \beta_i^{WML}\mathbb{E}\left[R_{WML}\right] + \beta_i^{LIQ}\mathbb{E}\left[R_{LIQ}\right]\end{aligned}$$

where LIQ measures the shock or innovation of the aggregate liquidity.



Alphas of decile portfolios sorted on predicted liquidity beta(s)

Long Q10 / Short Q1:

- 9.2% wrt 3F Fama-French model
- 7.5% wrt 4F Carhart model

The liquidity risk premium

Impact of the liquidity on the stock market

The dot-com crisis (2000-2003)

If we consider the S&P 500 index, we obtain:

- 55% of stocks post a negative performance

$$\approx 75\% \text{ of MC}$$

- 45% of stocks post a positive performance

$$\text{Maximum drawdown} = 49\ \%$$

Small caps stocks $\nearrow$
Value stocks $\nearrow$

The GFC crisis (2008)

If we consider the S&P 500 index, we obtain:

- 95% of stocks post a negative performance

$$\approx 97\% \text{ of MC}$$

- 5% of stocks post a positive performance

$$\text{Maximum drawdown} = 56\ \%$$

Small caps stocks $\searrow$
Value stocks $\searrow$

The liquidity risk premium

The specific status of the stock market

The interconnectedness nature of illiquid assets and liquid assets: the example of the Global Financial Crisis

- Subprime crisis ⇔ banks (credit risk)
- Banks ⇔ asset management, e.g. hedge funds (funding & leverage risk)
- Asset management ⇔ equity market (liquidity risk)
- Equity market ⇔ banks (asset-price & collateral risk)

The equity market is the ultimate liquidity provider:
GFC $\gg$ internet bubble

Remark

$1/3$ *of the losses in the stock market is explained by the liquidity supply*

The liquidity risk premium

Relationship between diversification & liquidity

During good times

- Medium correlation between liquid assets
- Illiquid assets have low impact on liquid assets
- Low substitution effects

Main effect:

$$\mathbb{E}[L_i]$$

During bad times

- High correlation between liquid assets
- Illiquid assets have a high impact on liquid assets
- High substitution effects

Main effects:

$$\beta(L_i, R_M) \text{ and } \beta(R_i, L_M)$$

The skewness puzzle

Skewness aggregation $\neq$ volatility aggregation

When we accumulate long/short skewness risk premia in a portfolio, the volatility of this portfolio decreases dramatically, but its skewness risk generally increases!

- Skewness diversification $\neq$ volatility diversification

$$\begin{aligned}\sigma(X_1 + X_2) &\leq \sigma(X_1) + \sigma(X_2) \\ |\gamma_1(X_1 + X_2)| &\not\leq |\gamma_1(X_1) + \gamma_1(X_2)|\end{aligned}$$

Skewness is not a convex risk measure

The skewness puzzle

Example 12

We assume that (X_1, X_2) follows a bivariate log-normal distribution $\mathcal{LN}\left(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho\right)$. This implies that $\ln X_1 \sim \mathcal{N}\left(\mu_1, \sigma_1^2\right)$, $\ln X_2 \sim \mathcal{N}\left(\mu_2, \sigma_2^2\right)$ and ρ is the correlation between $\ln X_1$ and $\ln X_2$.

The skewness puzzle

We recall that the skewness of X_1 is equal to:

$$\gamma_1 (X_1) = \frac{\mu_3 (X_1)}{\mu_2^{3/2} (X_1)} = \frac{e^{3\sigma_1^2} - 3e^{\sigma_1^2} + 2}{\left(e^{\sigma_1^2} - 1\right)^{3/2}}$$

whereas the skewness of $X_1 + X_2$ is equal to:

$$\gamma_1 (X_1 + X_2) = \frac{\mu_3 (X_1 + X_2)}{\mu_2^{3/2} (X_1 + X_2)}$$

where $\mu_n (X)$ is the n^{th} central moment of X

The skewness puzzle

In order to find the skewness of the sum $X_1 + X_2$, we need a preliminary result. By denoting $X = \alpha_1 \ln X_1 + \alpha_2 \ln X_2$, we have[16]:

$$\mathbb{E}\left[e^X\right] = e^{\mu_X + \frac{1}{2}\sigma_X^2}$$

where:

$$\mu_X = \alpha_1 \mu_1 + \alpha_2 \mu_2$$

and:

$$\sigma_X^2 = \alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2 + 2\alpha_1 \alpha_2 \rho \sigma_1 \sigma_2$$

It follows that:

$$\mathbb{E}\left[X_1^{\alpha_1} X_2^{\alpha_2}\right] = e^{\alpha_1 \mu_1 + \alpha_2 \mu_2 + \frac{1}{2}\left(\alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2 + 2\alpha_1 \alpha_2 \rho \sigma_1 \sigma_2\right)}$$

[16] Because X is a Gaussian random variable

The skewness puzzle

We have:

$$\mu_2(X_1 + X_2) = \mu_2(X_1) + \mu_2(X_2) + 2\operatorname{cov}(X_1, X_2)$$

where:

$$\mu_2(X_1) = e^{2\mu_1 + \sigma_1^2}\left(e^{\sigma_1^2} - 1\right)$$

and:

$$\operatorname{cov}(X_1, X_2) = \left(e^{\rho\sigma_1\sigma_2} - 1\right) e^{\mu_1 + \frac{1}{2}\sigma_1^2} e^{\mu_2 + \frac{1}{2}\sigma_2^2}$$

The skewness puzzle

For the third moment of $X_1 + X_2$, we use the following formula:

$$\mu_3 (X_1 + X_2) = \mu_3 (X_1) + \mu_3 (X_2) + 3 (\mathrm{cov}(X_1, X_1, X_2) + \mathrm{cov}(X_1, X_2, X_2))$$

where:

$$\mu_3 (X_1) = e^{2\mu_1 + \frac{3}{2}\sigma_1^2} \left(e^{3\sigma_1^2} - 3e^{\sigma_1^2} + 2 \right)$$

and:

$$\mathrm{cov}(X_1, X_1, X_2) = \left(e^{\rho\sigma_1\sigma_2} - 1 \right) e^{2\mu_1 + \sigma_1^2 + \mu_2 + \frac{\sigma_2^2}{2}} \left(e^{\sigma_1^2 + \rho\sigma_1\sigma_2} + e^{\sigma_2^2} - 2 \right)$$

The skewness puzzle

We deduce that:

$$\gamma_1 (X_1 + X_2) = \frac{\mu_3 (X_1 + X_2)}{\mu_2^{3/2} (X_1 + X_2)}$$

where:

$$\begin{aligned}\mu_2 (X_1 + X_2) &= e^{2\mu_1 + \sigma_1^2} \left(e^{\sigma_1^2} - 1 \right) + e^{2\mu_2 + \sigma_2^2} \left(e^{\sigma_2^2} - 1 \right) + \\ & \quad 2 \left(e^{\rho \sigma_1 \sigma_2} - 1 \right) e^{\mu_1 + \frac{1}{2}\sigma_1^2} e^{\mu_2 + \frac{1}{2}\sigma_2^2}\end{aligned}$$

and:

$$\begin{aligned}\mu_3 (X_1 + X_2) &= e^{2\mu_1 + \frac{3}{2}\sigma_1^2} \left(e^{3\sigma_1^2} - 3e^{\sigma_1^2} + 2 \right) + e^{2\mu_2 + \frac{3}{2}\sigma_2^2} \left(e^{3\sigma_2^2} - 3e^{\sigma_2^2} + 2 \right) + \\ & \quad 3 \left(e^{\rho \sigma_1 \sigma_2} - 1 \right) e^{2\mu_1 + \sigma_1^2 + \mu_2 + \frac{\sigma_2^2}{2}} \left(e^{\sigma_1^2 + \rho \sigma_1 \sigma_2} + e^{\sigma_2^2} - 2 \right) + \\ & \quad 3 \left(e^{\rho \sigma_1 \sigma_2} - 1 \right) e^{\mu_1 + \frac{1}{2}\sigma_1^2 + 2\mu_2 + \sigma_2^2} \left(e^{\sigma_2^2 + \rho \sigma_1 \sigma_2} + e^{\sigma_1^2} - 2 \right)\end{aligned}$$

The skewness puzzle

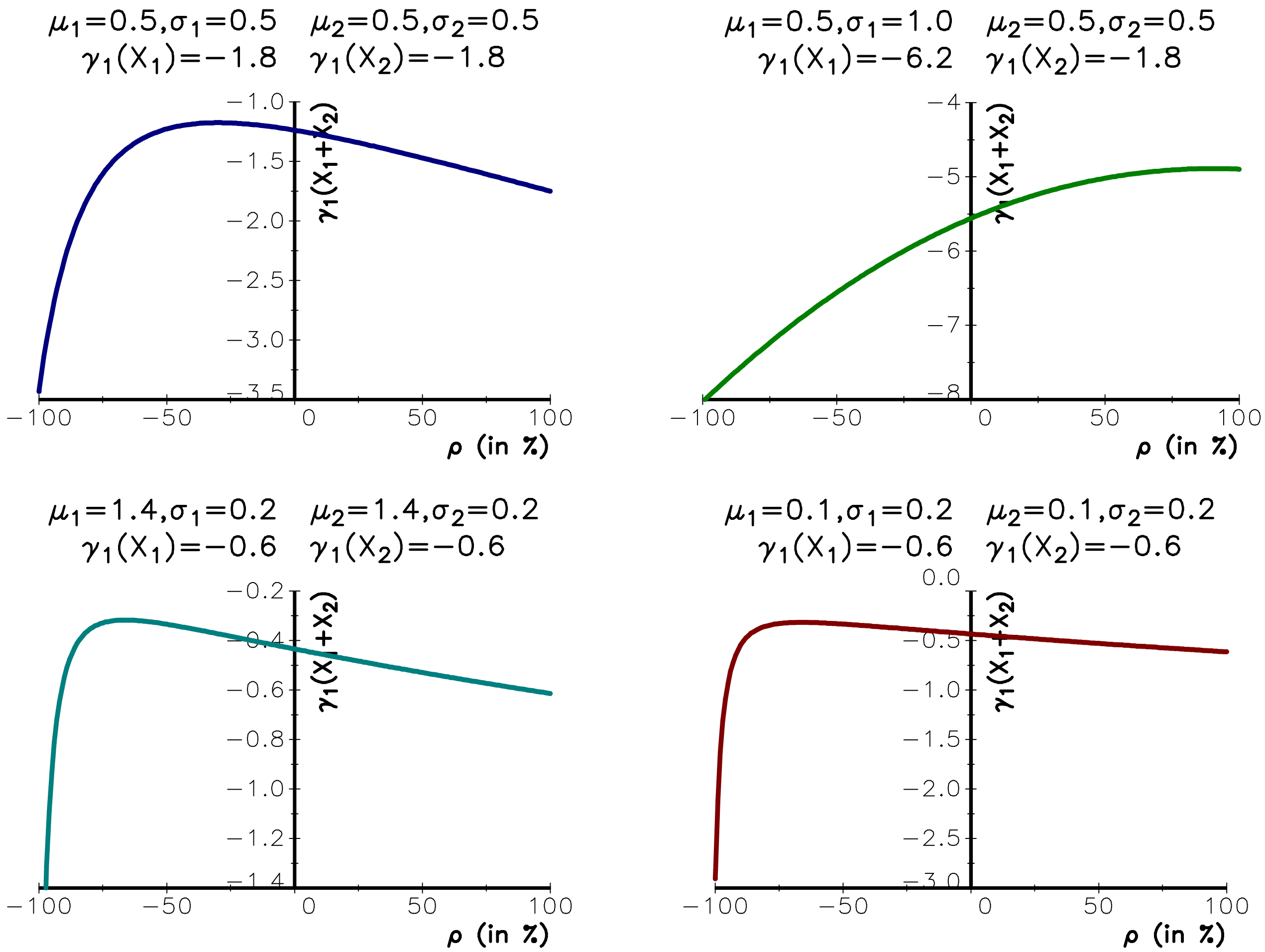


Figure 74: Skewness aggregation of the random vector $(-X_1, -X_3)$

The skewness puzzle

Why?

- Volatility diversification works very well with L/S risk premia:

$$\sigma(R(x)) \approx \frac{\bar{\sigma}}{\sqrt{n}}$$

- Drawdown diversification don't work very well because bad times are correlated and are difficult to hedge:

$$\mathrm{DD}(x) \approx \overline{\mathrm{DD}}$$

The skewness puzzle

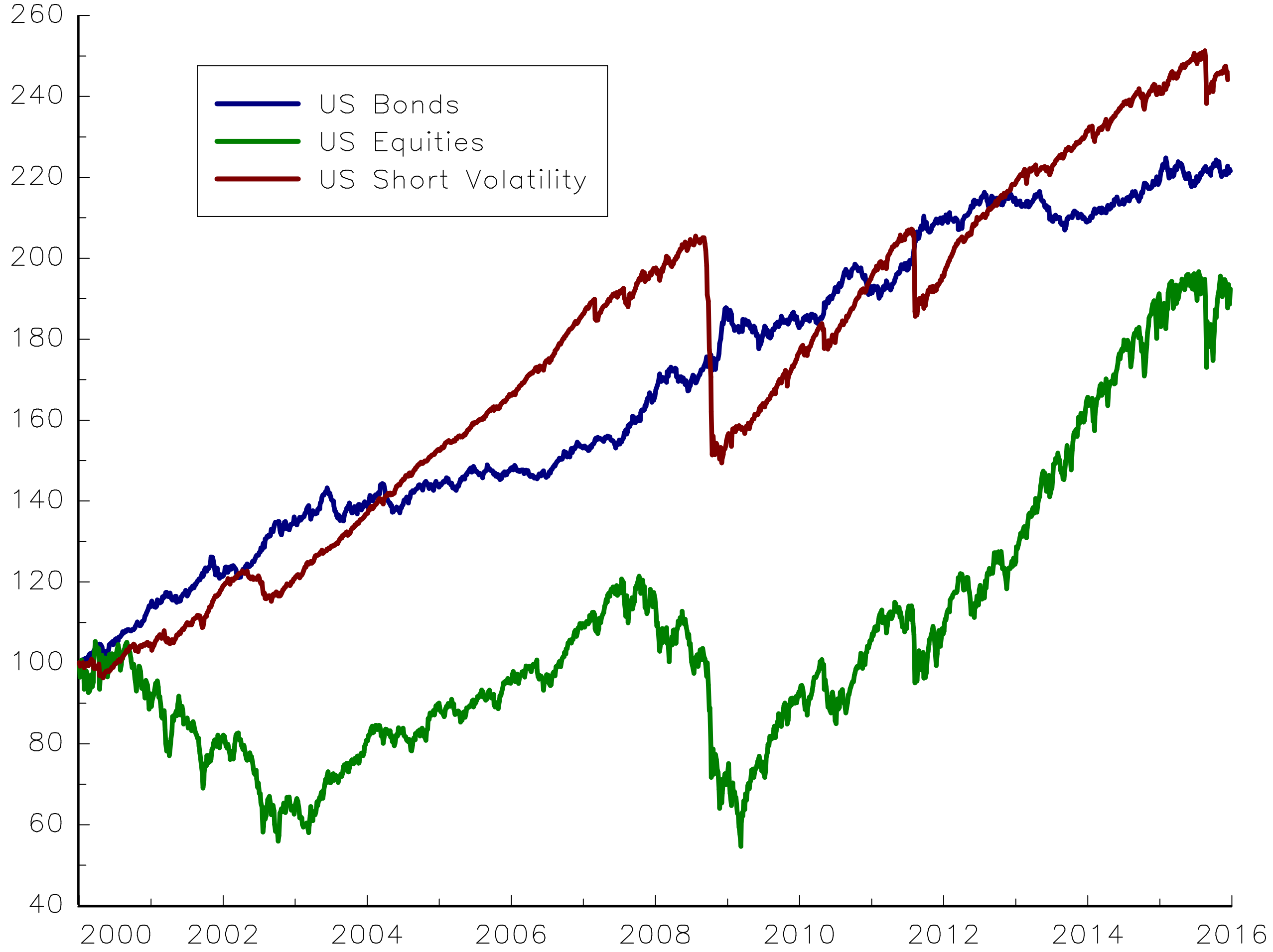


Figure 75: Cumulative performance of US 10Y bonds, US equities and US short volatility

Source: Bruder *et al.* (2016)

The correlation puzzle

We consider the Gaussian random vector (R_1, R_2, R_3), whose volatilities are equal to 25%, 12% and 9.76%. The correlation matrix is given by:

$$C = \begin{pmatrix} 100\% & & \\ -25.00\% & 100\% & \\ 55.31\% & 66.84\% & 100\% \end{pmatrix}$$

Good diversification? (correlation approach)

If R_i represents an asset return (or an excess return), we conclude that (R_1, R_2, R_3) is a well-diversified investment universe

Bad diversification? (payoff approach)

However, we have:

$$R_3 = 0.30 R_1 + 0.70 R_2$$

The correlation puzzle

Fantasies about correlations

- Negative correlations are good for diversification
- Positive correlations are bad for diversification

- If $\rho(R_1, R_2)$ is close to -1, can we hedge Asset 1 with Asset 2?
- If $\rho(R_1, R_2)$ is close to -1, can we diversify Asset 1 with Asset 2?
- If $\rho(R_1, R_2)$ is close to $+1$, can we hedge Asset 1 with a short position on Asset 2?
- If $\rho(R_1, R_2)$ is close to $+1$, can we diversify Asset 1 with a short position on Asset 2?
- Does $\rho(R_1, R_2) = -70\%$ correspond to a better diversification pattern than $\rho(R_1, R_2) = +70\%$?

There is a confusion between diversification and hedging!

The payoff approach

Table 62: Correlation matrix between asset classes (2000-2016)

		Equity				Bond			
		US	Euro	UK	Japan	US	Euro	UK	Japan
Equity	US	100%							
	Euro	78%	100%						
	UK	79%	87%	100%					
	Japan	53%	57%	55%	100%				
Bond	US	−35%	−39%	−32%	−29%	100%			
	Euro	−17%	−16%	−16%	−16%	58%	100%		
	UK	−31%	−37%	−30%	−31%	72%	63%	100%	
	Japan	−17%	−18%	−16%	−33%	37%	31%	36%	100%

Correlation = Pearson correlation = Linear correlation

The payoff approach

Let us consider a Gaussian random vector defined as follows:

$$\begin{pmatrix} Y \\ X \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix}, \begin{pmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{pmatrix}\right)$$

The conditional distribution of Y given $X = x$ is a MN distribution:

$$\mu_{y|x} = \mathbb{E}\left[Y \mid X = x\right] = \mu_y + \Sigma_{yx}\Sigma_{xx}^{-1}\left(x - \mu_x\right)$$

and:

$$\Sigma_{yy|x} = \sigma^2\left[Y \mid X = x\right] = \Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}$$

We deduce that:

$$\begin{array}{rcl} Y & = & \mu_y + \Sigma_{yx}\Sigma_{xx}^{-1}\left(x - \mu_x\right) + u \\ & = & \underbrace{\left(\mu_y - \Sigma_{yx}\Sigma_{xx}^{-1}\mu_x\right)}_{\beta_0} + \underbrace{\Sigma_{yx}\Sigma_{xx}^{-1}}_{\beta^\top} x + u \end{array}$$

where u is a centered Gaussian random variable with variance $s^2 = \Sigma_{yy|x}$.

The payoff approach

Correlation = linear payoff

It follows that the payoff function is defined by the curve:

$$y = f(x)$$

where:

$$\begin{aligned} f(x) &= \mathbb{E}[R_2 \mid R_1 = x] \\ &= \left(\mu_2 - \beta_{2|1}\mu_1\right) + \beta_{2|1}x \end{aligned}$$

The payoff approach

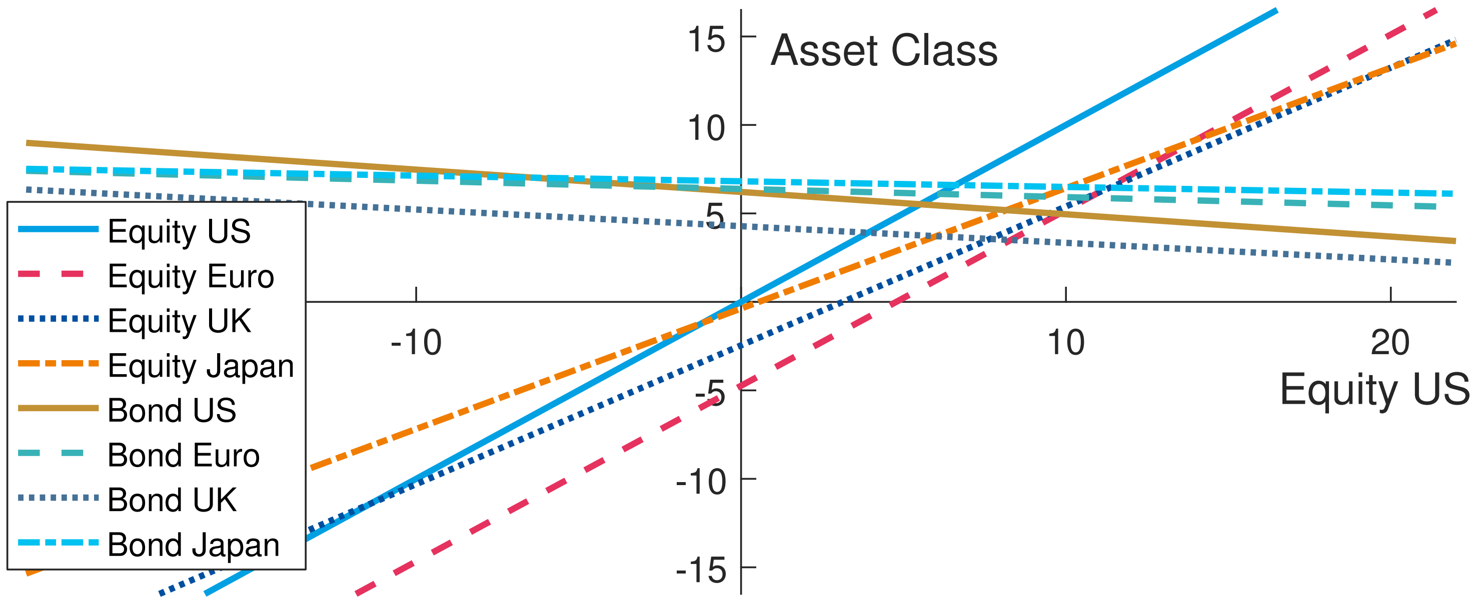


Figure 76: Linear payoff function with respect to the S&P 500 Index

A long-only diversified stock-bond portfolio makes sense!

The payoff approach

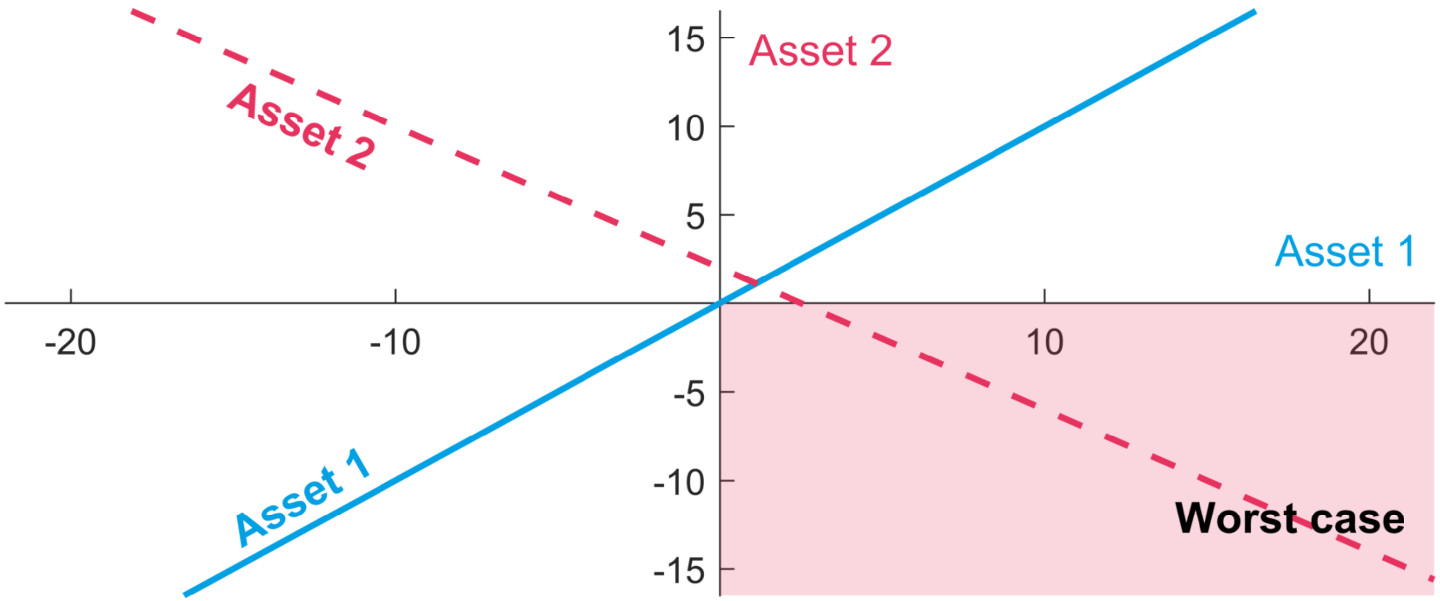


Figure 77: Worst diversification case

What is good diversification? What is bad diversification?

Negative correlation does not necessarily imply good diversification!

The payoff approach

Concave payoff

- Negative skewness
- Positive vega
- Hit ratio $\geq 50\%$
- Gain frequency $>$ loss frequency
- Average gain $<$ average loss
- Positively correlated with bad times

Convex payoff

- Positive skewness
- Negative vega
- Hit ratio $\leq 50\%$
- Gain frequency $<$ loss frequency
- Average gain $>$ average loss
- Negatively correlated with bad times?

Volatility Carry $\neq$ **Time-series Momentum**

The payoff approach

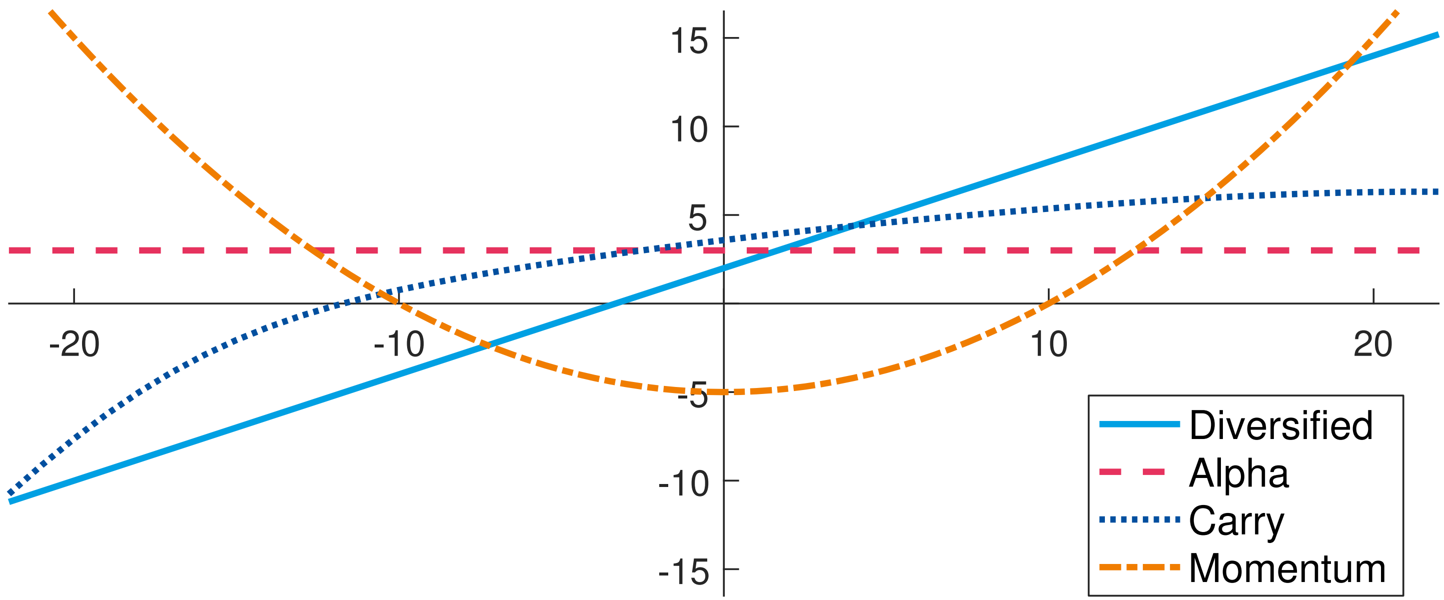


Figure 78: What does portfolio optimization produce with convex and concave strategies?

- Momentum = low allocation during good times and high allocation after bad times
- Carry = high allocation during good times and low allocation after bad times

The payoff approach

The magic formula

Long-run positive correlations, but...

...negative correlations is bad times ☺

The payoff approach

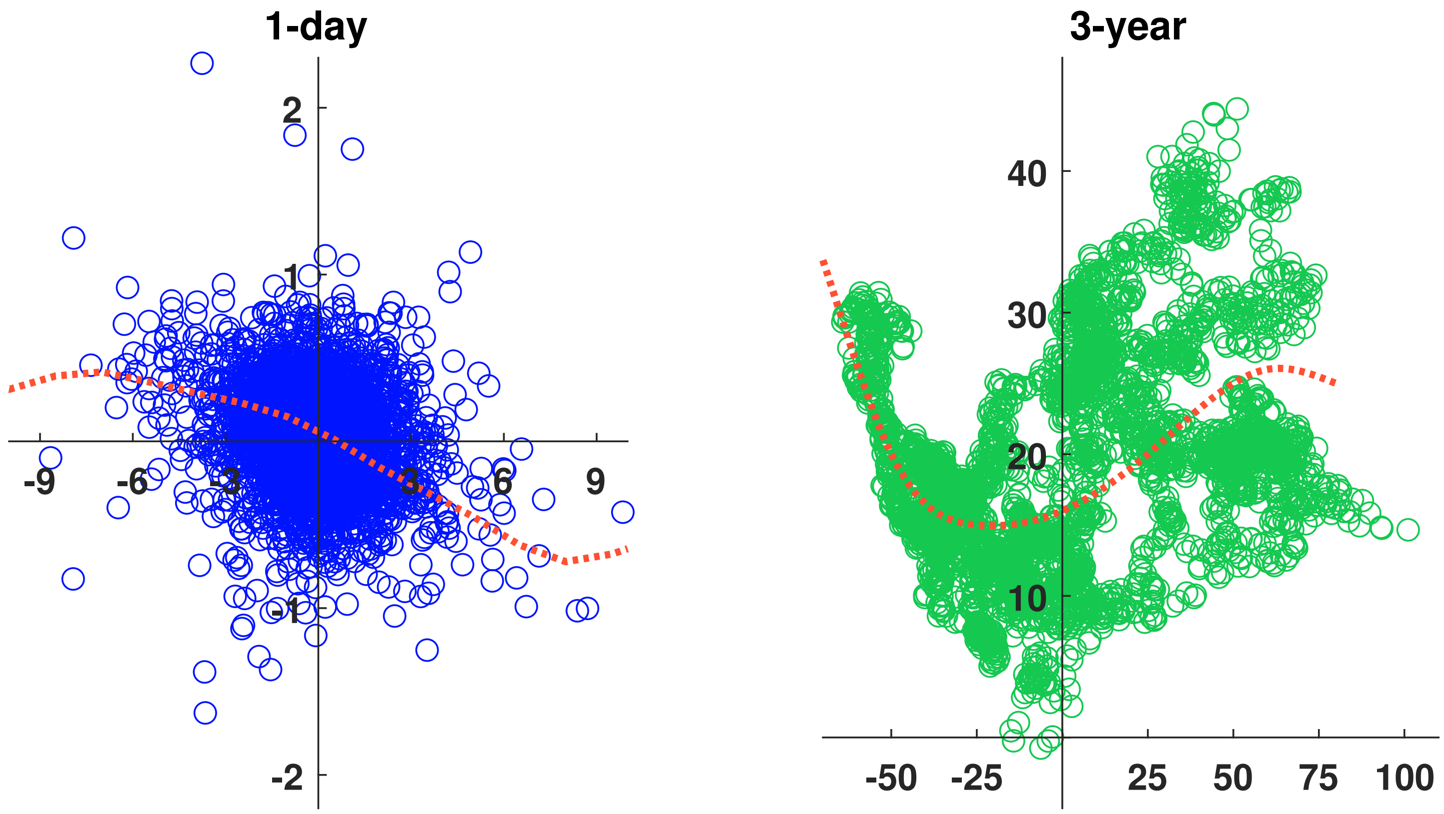


Figure 79: Stock/bond payoff (EUR)

Daily diversification is different than 3-year diversification

Equally-weighted portfolio

Exercise

We note Σ the covariance matrix of n asset returns. In what follows, we consider the equally weighted portfolio based on the universe of these n assets.

Equally-weighted portfolio

Question 1

Let $\Sigma_{i,j} = \rho_{i,j}\sigma_i\sigma_j$ be the elements of the covariance matrix Σ.

Equally-weighted portfolio

Question 1.a

Compute the volatility $\sigma(x)$ of the EW portfolio.

Equally-weighted portfolio

The elements of the covariance matrix are $\Sigma_{i,j} = \rho_{i,j}\sigma_i\sigma_j$. If we consider a portfolio $x = (x_1, \ldots, x_n)$, its volatility is:

$$\begin{aligned}\sigma(x) &= \sqrt{x^\top \Sigma x} \\ &= \sqrt{\sum_{i=1}^{n} x_i^2\sigma_i^2 + 2\sum_{i>j} x_i x_j \rho_{i,j}\sigma_i\sigma_j}\end{aligned}$$

For the equally weighted portfolio, we have $x_i = n^{-1}$ and:

$$\sigma(x) = \frac{1}{n}\sqrt{\sum_{i=1}^{n} \sigma_i^2 + 2\sum_{i>j} \rho_{i,j}\sigma_i\sigma_j}$$

Equally-weighted portfolio

Question 1.b

Let $\sigma_0(x)$ and $\sigma_1(x)$ be the volatility of the EW portfolio when the asset returns are respectively independent and perfectly correlated. Calculate $\sigma_0(x)$ and $\sigma_1(x)$.

Equally-weighted portfolio

We have:

$$\sigma_0(x) = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\sigma_i^2}$$

and:

$$\begin{aligned}
\sigma_1(x) &= \frac{1}{n}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i\sigma_j} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\sigma_i\sum_{j=1}^{n}\sigma_j} \\
&= \frac{1}{n}\sqrt{\left(\sum_{i=1}^{n}\sigma_i\right)^2} = \frac{\sum_{i=1}^{n}\sigma_i}{n} \\
&= \bar{\sigma}
\end{aligned}$$

Equally-weighted portfolio

Question 1.c

We assume that the volatilities are the same. Find the expression of the portfolio volatility with respect to the mean correlation $\bar{\rho}$. What is the value of $\sigma(x)$ when $\bar{\rho}$ is equal to zero? What is the value of $\sigma(x)$ when n tends to $+\infty$?

Equally-weighted portfolio

If $\sigma_i = \sigma_j = \sigma$, we obtain:

$$\sigma(x) = \frac{\sigma}{n}\sqrt{n + 2\sum_{i>j}\rho_{i,j}}$$

Let $\bar{\rho}$ be the mean correlation. We have:

$$\bar{\rho} = \frac{2}{n^2 - n}\sum_{i>j}\rho_{i,j}$$

We deduce that:

$$\sum_{i>j}\rho_{i,j} = \frac{n(n-1)}{2}\bar{\rho}$$

Equally-weighted portfolio

We finally obtain:

$$\begin{aligned}\sigma(x) &= \frac{\sigma}{n}\sqrt{n + n(n-1)\bar{\rho}} \\ &= \sigma\sqrt{\frac{1+(n-1)\bar{\rho}}{n}}\end{aligned}$$

When $\bar{\rho}$ is equal to zero, the volatility $\sigma(x)$ is equal to $\sigma/\sqrt{n}$. When the number of assets tends to $+\infty$, it follows that:

$$\lim_{n\to\infty} \sigma(x) = \sigma\sqrt{\bar{\rho}}$$

Equally-weighted portfolio

Question 1.d

We assume that the correlations are uniform ($\rho_{i,j} = \rho$). Find the expression of the portfolio volatility as a function of $\sigma_0(x)$ and $\sigma_1(x)$. Comment on this result.

Equally-weighted portfolio

If $\rho_{i,j} = \rho$, we obtain:

$$\begin{aligned}
\sigma(x) &= \frac{1}{n}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{i,j}\sigma_i\sigma_j} \\
&= \frac{1}{n}\sqrt{\sum_{i=1}^{n}\sigma_i^2 + \rho\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i\sigma_j - \rho\sum_{i=1}^{n}\sigma_i^2} \\
&= \frac{1}{n}\sqrt{(1-\rho)\sum_{i=1}^{n}\sigma_i^2 + \rho\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i\sigma_j}
\end{aligned}$$

Equally-weighted portfolio

We have:

$$\sum_{i=1}^{n} \sigma_i^2 = n^2 \sigma_0^2 (x)$$

and:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_i \sigma_j = n^2 \sigma_1^2 (x)$$

It follows that:

$$\sigma (x) = \sqrt{(1 - \rho) \sigma_0^2 (x) + \rho \sigma_1^2 (x)}$$

When the correlation is uniform, the variance $\sigma^2 (x)$ is the weighted average between $\sigma_0^2 (x)$ and $\sigma_1^2 (x)$.

Equally-weighted portfolio

Question 2.a

Compute the normalized risk contributions $\mathcal{RC}_i^\star$ of the EW portfolio.

Equally-weighted portfolio

The risk contributions are equal to:

$$\mathcal{RC}_i^\star = \frac{x_i \cdot (\Sigma x)_i}{\sigma^2(x)}$$

In the case of the EW portfolio, we obtain:

$$\begin{aligned}\mathcal{RC}_i^\star &= \frac{\sum_{j=1}^n \rho_{i,j}\sigma_i\sigma_j}{n^2\sigma^2(x)} \\ &= \frac{\sigma_i^2 + \sigma_i \sum_{j\neq i} \rho_{i,j}\sigma_j}{n^2\sigma^2(x)}\end{aligned}$$

Equally-weighted portfolio

Question 2.b

Deduce the risk contributions $\mathcal{RC}_i^\star$ when the asset returns are respectively independent and perfectly correlated[a].

[a] We note them $\mathcal{RC}_{0,i}^\star$ and $\mathcal{RC}_{1,i}^\star$.

Equally-weighted portfolio

If asset returns are independent, we have:

$$\mathcal{RC}_{0,i}^{\star} = \frac{\sigma_i^2}{\sum_{i=1}^{n} \sigma_i^2}$$

In the case of perfect correlation, we obtain:

$$\begin{aligned}
\mathcal{RC}_{1,i}^{\star} &= \frac{\sigma_i^2 + \sigma_i \sum_{j \neq i} \sigma_j}{n^2 \bar{\sigma}^2} \\
&= \frac{\sigma_i \sum_j \sigma_j}{n^2 \bar{\sigma}^2} \\
&= \frac{\sigma_i}{n \bar{\sigma}}
\end{aligned}$$

Equally-weighted portfolio

Question 2.c

Show that the risk contribution $\mathcal{RC}_i$ is proportional to the ratio between the mean correlation of asset i and the mean correlation of the asset universe when the volatilities are the same.

Equally-weighted portfolio

If $\sigma_i = \sigma_j = \sigma$, we obtain:

$$\begin{aligned}\mathcal{RC}_i^\star &= \frac{\sigma^2 + \sigma^2 \sum_{j \neq i} \rho_{i,j}}{n^2 \sigma^2 (x)} \\ &= \frac{\sigma^2 + (n-1)\, \sigma^2 \bar{\rho}_i}{n^2 \sigma^2 (x)} \\ &= \frac{1 + (n-1)\, \bar{\rho}_i}{n\left(1 + (n-1)\, \bar{\rho}\right)}\end{aligned}$$

It follows that:

$$\lim_{n \to \infty} \frac{1 + (n-1)\, \bar{\rho}_i}{1 + (n-1)\, \bar{\rho}} = \frac{\bar{\rho}_i}{\bar{\rho}}$$

We deduce that the risk contributions are proportional to the ratio between the mean correlation of asset i and the mean correlation of the asset universe.

Equally-weighted portfolio

Question 2.d

We assume that the correlations are uniform ($\rho_{i,j} = \rho$). Show that the risk contribution $\mathcal{RC}_i$ is a weighted average of $\mathcal{RC}^{\star}_{0,i}$ and $\mathcal{RC}^{\star}_{1,i}$.

Equally-weighted portfolio

We recall that we have:

$$\sigma(x) = \sqrt{(1-\rho)\,\sigma_0^2(x) + \rho\sigma_1^2(x)}$$

It follows that:

$$\begin{aligned}\mathcal{RC}_i &= x_i \cdot \frac{(1-\rho)\,\sigma_0(x)\,\partial_{x_i}\sigma_0(x) + \rho\sigma_1(x)\,\partial_{x_i}\sigma_1(x)}{\sqrt{(1-\rho)\,\sigma_0^2(x) + \rho\sigma_1^2(x)}} \\ &= \frac{(1-\rho)\,\sigma_0(x)\,\mathcal{RC}_{0,i} + \rho\sigma_1(x)\,\mathcal{RC}_{1,i}}{\sqrt{(1-\rho)\,\sigma_0^2(x) + \rho\sigma_1^2(x)}}\end{aligned}$$

We then obtain:

$$\mathcal{RC}_i^\star = \frac{(1-\rho)\,\sigma_0^2(x)}{\sigma^2(x)}\mathcal{RC}_{0,i}^\star + \frac{\rho\sigma_1(x)}{\sigma^2(x)}\mathcal{RC}_{1,i}^\star$$

We verify that the risk contribution $\mathcal{RC}_i$ is a weighted average of $\mathcal{RC}_{0,i}^\star$ and $\mathcal{RC}_{1,i}^\star$.

Equally-weighted portfolio

Question 3

We suppose that the return of asset i satisfies the CAPM model:

$$R_i = \beta_i R_m + \varepsilon_i$$

where R_m is the return of the market portfolio and ε_i is the specific risk. We note $\beta = (\beta_1, \ldots, \beta_n)$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$. We assume that $R_m \perp \varepsilon$, $\operatorname{var}(R_m) = \sigma_m^2$ and $\operatorname{cov}(\varepsilon) = D = \operatorname{diag}(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2)$.

Equally-weighted portfolio

Question 3.a

Calculate the volatility of the EW portfolio.

Equally-weighted portfolio

We have:

$$\Sigma = \beta\beta^\top \sigma_m^2 + D$$

We deduce that:

$$\sigma(x) = \frac{1}{n}\sqrt{\sigma_m^2 \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_i \beta_j + \sum_{i=1}^{n} \tilde{\sigma}_i^2}$$

Equally-weighted portfolio

Question 3.b

Calculate the risk contribution $\mathcal{RC}_i$.

Equally-weighted portfolio

The risk contributions are equal to:

$$\mathcal{RC}_i = \frac{x_i \cdot (\Sigma x)_i}{\sigma(x)}$$

In the case of the EW portfolio, we obtain:

$$\begin{aligned} \mathcal{RC}_i &= \frac{x_i \cdot \left(\sigma_m^2 \beta_i \sum_{j=1}^n x_j \beta_j + x_i \tilde{\sigma}_i^2\right)}{n^2 \sigma(x)} \\ &= \frac{\sigma_m^2 \beta_i \sum_{j=1}^n \beta_j + \tilde{\sigma}_i^2}{n^2 \sigma(x)} \\ &= \frac{n \sigma_m^2 \beta_i \bar{\beta} + \tilde{\sigma}_i^2}{n^2 \sigma(x)} \end{aligned}$$

Equally-weighted portfolio

Question 3.c

Show that $\mathcal{RC}_i$ is approximately proportional to β_i if the number of assets is large. Illustrate this property using a numerical example.

Equally-weighted portfolio

When the number of assets is large and $\beta_i > 0$, we obtain:

$$\mathcal{RC}_i \simeq \frac{\sigma_m^2 \beta_i \bar{\beta}}{n\sigma(x)}$$

because $\bar{\beta} > 0$. We deduce that the risk contributions are approximately proportional to the beta coefficients:

$$\mathcal{RC}_i^\star \simeq \frac{\beta_i}{\sum_{j=1}^n \beta_j}$$

In Figure 80, we compare the exact and approximated values of $\mathcal{RC}_i^\star$. For that, we simulate β_i and $\tilde{\sigma}_i$ with $\beta_i \sim \mathcal{U}_{[0.5,1.5]}$ and $\tilde{\sigma}_i \sim \mathcal{U}_{[0,20\%]}$ whereas σ_m is set to 25%. We notice that the approximated value is very close to the exact value when n increases.

Equally-weighted portfolio

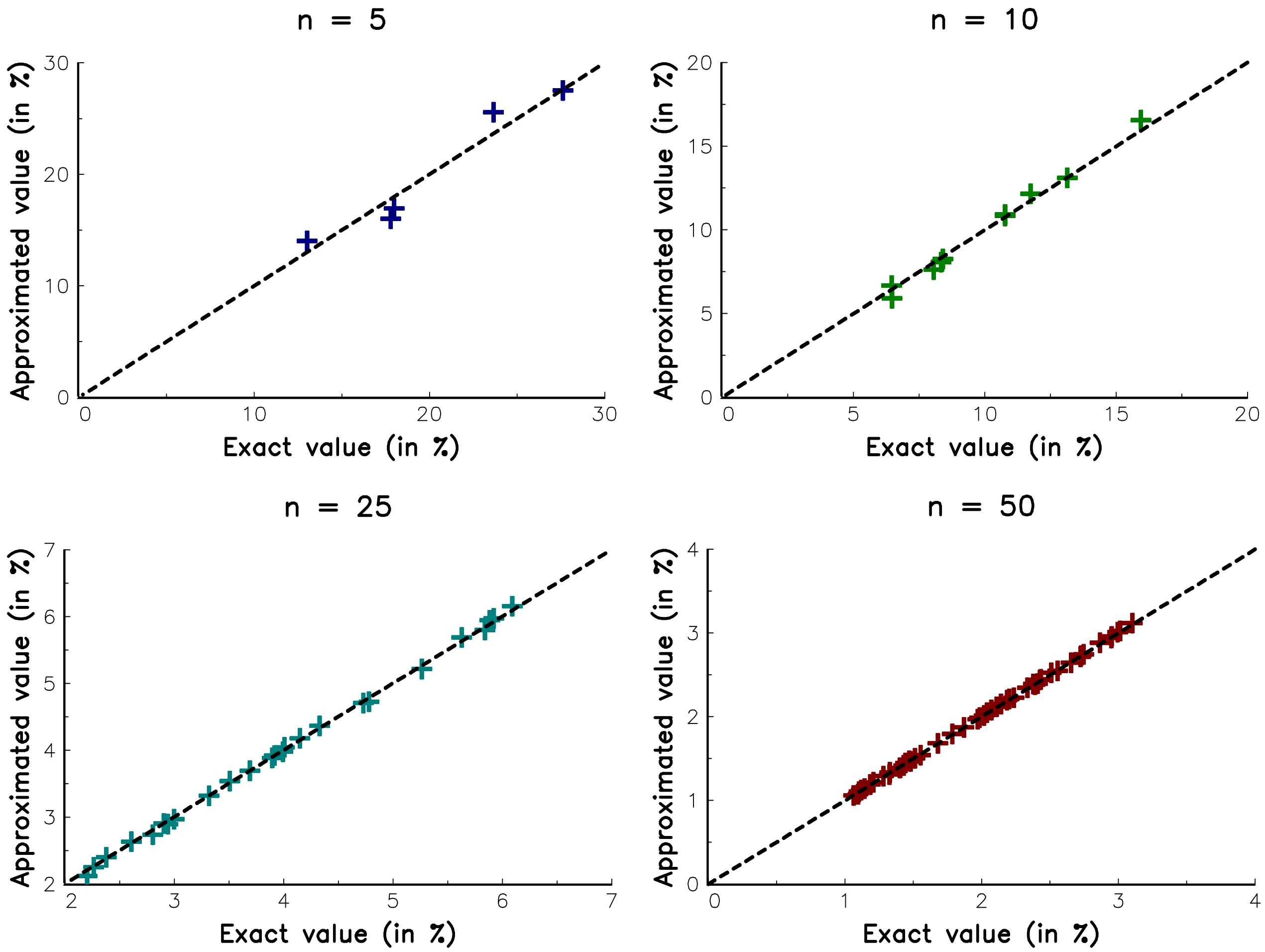


Figure 80: Comparing the exact and approximated values of $\mathcal{RC}_i^\star$

Most diversified portfolio

Exercise

We consider a universe of n assets. We note $\sigma = (\sigma_1, \ldots, \sigma_n)$ the vector of volatilities and Σ the covariance matrix.

Most diversified portfolio

Question 1

In what follows, we consider non-constrained optimized portfolios.

Most diversified portfolio

Question 1.a

Define the diversification ratioDiversification ratio $\mathcal{DR}(x)$ by considering a general risk measure $\mathcal{R}(x)$. How can one interpret this measure from a risk allocation perspective?

Most diversified portfolio

Let $\mathcal{R}(x)$ be the risk measure of the portfolio x. We note $\mathcal{R}_i = \mathcal{R}(\mathbf{e}_i)$ the risk associated to the i^{th} asset. The diversification ratio is the ratio between the weighted mean of the individual risks and the portfolio risk (TR-RPB, page 168):

$$\mathcal{DR}(x) = \frac{\sum_{i=1}^{n} x_i \mathcal{R}_i}{\mathcal{R}(x)}$$

If we assume that the risk measure satisfies the Euler allocation principle, we have:

$$\mathcal{DR}(x) = \frac{\sum_{i=1}^{n} x_i \mathcal{R}_i}{\sum_{i=1}^{n} \mathcal{RC}_i}$$

Most diversified portfolio

Question 1.b

We assume that the weights of the portfolio are positive. Show that $\mathcal{DR}(x) \geq 1$ for all risk measures satisfying the Euler allocation principle. Find an upper bound of $\mathcal{DR}(x)$.

Most diversified portfolio

If $\mathcal{R}(x)$ satisfies the Euler allocation principle, we know that $\mathcal{R}_i \geq \mathcal{MR}_i$ (TR-RPB, page 78). We deduce that:

$$\begin{aligned} \mathcal{DR}(x) & \geq \frac{\sum_{i=1}^{n} x_i \mathcal{R}_i}{\sum_{i=1}^{n} x_i \mathcal{R}_i} \\ & \geq 1 \end{aligned}$$

Let x_{mr} be the portfolio that minimizes the risk measure. We have:

$$\mathcal{DR}(x) \leq \frac{\sup \mathcal{R}_i}{\mathcal{R}(x_{\mathrm{mr}})}$$

Most diversified portfolio

Question 1.c

We now consider the volatility risk measure. Calculate the upper bound of $\mathcal{DR}(x)$.

Most diversified portfolio

If we consider the volatility risk measure, the minimum risk portfolio is the minimum variance portfolio. We have (TR-RPB, page 164):

$$\sigma(x_{\mathrm{mv}}) = \frac{1}{\sqrt{\mathbf{1}_n^\top \Sigma \mathbf{1}_n}}$$

We deduce that:

$$\mathcal{DR}(x) \leq \sqrt{\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n} \cdot \sup \sigma_i$$

Most diversified portfolio

Question 1.d

What is the most diversified portfolio (or MDP)? In which case does it correspond to the tangency portfolio? Deduce the analytical expression of the MDP and calculate its volatility.

Most diversified portfolio

The MDP is the portfolio which maximizes the diversification ratio when the risk measure is the volatility (TR-RPB, page 168). We have:

$$\begin{aligned} x^{\star} &= \arg\max \mathcal{DR}(x) \\ \text{u.c.} &\quad \mathbf{1}_n^{\top} x = 1 \end{aligned}$$

Most diversified portfolio

If we consider that the risk premium $\pi_i = \mu_i - r$ of the asset i is proportional to its volatility σ_i, we obtain:

$$
\begin{aligned}
\mathrm{SR}(x \mid r) &= \frac{\mu(x) - r}{\sigma(x)} \\
&= \frac{\sum_{i=1}^{n} x_i (\mu_i - r)}{\sigma(x)} \\
&= s \frac{\sum_{i=1}^{n} x_i \sigma_i}{\sigma(x)} \\
&= s \cdot \mathcal{DR}(x)
\end{aligned}
$$

where s is the coefficient of proportionality. Maximizing the diversification ratio is equivalent to maximizing the Sharpe ratio.

Most diversified portfolio

We recall that the expression of the tangency portfolio is:

$$x^{\star} = \frac{\Sigma^{-1} \left(\mu - r\mathbf{1}_n\right)}{\mathbf{1}_n^{\top} \Sigma^{-1} \left(\mu - r\mathbf{1}_n\right)}$$

We deduce that the weights of the MDP are:

$$x^{\star} = \frac{\Sigma^{-1} \sigma}{\mathbf{1}_n^{\top} \Sigma^{-1} \sigma}$$

The volatility of the MDP is then:

$$\begin{aligned} \sigma\left(x^{\star}\right) &= \sqrt{\frac{\sigma^{\top} \Sigma^{-1}}{\mathbf{1}_n^{\top} \Sigma^{-1} \sigma} \Sigma \frac{\Sigma^{-1} \sigma}{\mathbf{1}_n^{\top} \Sigma^{-1} \sigma}} \\ &= \frac{\sqrt{\sigma^{\top} \Sigma^{-1} \sigma}}{\mathbf{1}_n^{\top} \Sigma^{-1} \sigma} \end{aligned}$$

Most diversified portfolio

Question 1.e

Demonstrate then that the weights of the MDP are in some sense proportional to $\Sigma^{-1}\sigma$.

Most diversified portfolio

We recall that another expression of the unconstrained tangency portfolio is:

$$x^{\star} = \frac{\sigma^{2}\left(x^{\star}\right)}{\left(\mu\left(x^{\star}\right) - r\right)} \Sigma^{-1}\left(\mu - r\mathbf{1}_{n}\right)$$

We deduce that the MDP is also:

$$x^{\star} = \frac{\sigma^{2}\left(x^{\star}\right)}{\bar{\sigma}\left(x^{\star}\right)} \Sigma^{-1}\sigma$$

where $\bar{\sigma}\left(x^{\star}\right) = x^{\star\top}\sigma$. Nevertheless, this solution is endogenous.

Most diversified portfolio

Question 2

We suppose that the return of asset i satisfies the CAPM:

$$R_i = \beta_i R_m + \varepsilon_i$$

where R_m is the return of the market portfolio and ε_i is the specific risk. We note $\beta = (\beta_1, \ldots, \beta_n)$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$. We assume that $R_m \perp \varepsilon$, $\operatorname{var}(R_m) = \sigma_m^2$ and $\operatorname{cov}(\varepsilon) = D = \operatorname{diag}(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2)$.

Most diversified portfolio

Question 2.a

Compute the correlation $\rho_{i,m}$ between the asset return and the market return. Deduce the relationship between the specific risk $\tilde{\sigma}_i$ and the total risk σ_i of asset i.

Most diversified portfolio

We have:

$$\operatorname{cov}\left(R_i, R_m\right) = \beta_i \sigma_m^2$$

We deduce that:

$$\begin{aligned} \rho_{i,m} &= \frac{\operatorname{cov}\left(R_i, R_m\right)}{\sigma_i \sigma_m} \\ &= \beta_i \frac{\sigma_m}{\sigma_i} \end{aligned} \tag{4}$$

and:

$$\begin{aligned} \tilde{\sigma}_i &= \sqrt{\sigma_i^2 - \beta_i^2 \sigma_m^2} \\ &= \sigma_i \sqrt{1 - \rho_{i,m}^2} \end{aligned} \tag{5}$$

Most diversified portfolio

Question 2.b

Show that the solution of the MDP may be written as:

$$x_i^\star = \mathcal{DR}\left(x^\star\right) \frac{\sigma_i \sigma\left(x^\star\right)}{\tilde{\sigma}_i^2} \left(1 - \frac{\rho_{i,m}}{\rho^\star}\right) \tag{6}$$

with $\rho^\star$ a scalar to be determined.

Most diversified portfolio

We know that (TR-RPB, page 167):

$$\Sigma^{-1} = D^{-1} - \frac{1}{\sigma_m^{-2} + \tilde{\beta}^\top \beta} \tilde{\beta} \tilde{\beta}^\top$$

where $\tilde{\beta}_i = \beta_i / \tilde{\sigma}_i^2$.

Most diversified portfolio

We deduce that:

$$x^{\star} = \frac{\sigma^2(x^{\star})}{\bar{\sigma}(x^{\star})}\left(D^{-1}\sigma - \frac{1}{\sigma_m^{-2} + \tilde{\beta}^{\top}\beta}\tilde{\beta}\tilde{\beta}^{\top}\sigma\right)$$

and:

$$\begin{aligned} x_i^{\star} &= \frac{\sigma^2(x^{\star})}{\bar{\sigma}(x^{\star})}\left(\frac{\sigma_i}{\tilde{\sigma}_i^2} - \frac{\tilde{\beta}^{\top}\sigma}{\sigma_m^{-2} + \tilde{\beta}^{\top}\beta}\tilde{\beta}_i\right) \\ &= \frac{\sigma_i\sigma^2(x^{\star})}{\bar{\sigma}(x^{\star})\tilde{\sigma}_i^2}\left(1 - \frac{\tilde{\beta}^{\top}\sigma}{\sigma_m^{-1} + \sigma_m\tilde{\beta}^{\top}\beta}\frac{\sigma_m\tilde{\sigma}_i^2\tilde{\beta}_i}{\sigma_i}\right) \\ &= \frac{\sigma_i\sigma^2(x^{\star})}{\bar{\sigma}(x^{\star})\tilde{\sigma}_i^2}\left(1 - \frac{\tilde{\beta}^{\top}\sigma}{\sigma_m^{-1} + \sigma_m\tilde{\beta}^{\top}\beta}\rho_{i,m}\right) \\ &= \mathcal{DR}(x^{\star})\frac{\sigma_i\sigma(x^{\star})}{\tilde{\sigma}_i^2}\left(1 - \frac{\rho_{i,m}}{\rho^{\star}}\right) \end{aligned}$$

Most diversified portfolio

Using Equations (4) and (5), $\rho^\star$ is defined as follows:

$$\begin{aligned}
\rho^\star &= \frac{\sigma_m^{-1} + \sigma_m \tilde{\beta}^\top \beta}{\tilde{\beta}^\top \sigma} \\
&= \left(1 + \sum_{j=1}^{n} \frac{\sigma_m^2 \beta_j^2}{\tilde{\sigma}_j^2}\right) \Bigg/ \left(\sum_{j=1}^{n} \frac{\sigma_m \beta_j \sigma_j}{\tilde{\sigma}_j^2}\right) \\
&= \left(1 + \sum_{j=1}^{n} \frac{\rho_{j,m}^2}{1 - \rho_{j,m}^2}\right) \Bigg/ \left(\sum_{j=1}^{n} \frac{\rho_{j,m}}{1 - \rho_{j,m}^2}\right)
\end{aligned}$$

Most diversified portfolio

Question 2.c

In which case is the optimal weight $x_i^\star$ positive?

Most diversified portfolio

The optimal weight $x_i^\star$ is positive if:

$$1 - \frac{\rho_{i,m}}{\rho^\star} \geq 0$$

or equivalently:

$$\rho_{i,m} \leq \rho^\star$$

Most diversified portfolio

Question 2.d

Are the weights of the MDP a decreasing or an increasing function of the specific risk $\tilde{\sigma}_i$?

Most diversified portfolio

We recall that:

$$\begin{aligned}\rho_{i,m} &= \beta_i \frac{\sigma_m}{\sigma_i} \\ &= \frac{\beta_i \sigma_m}{\sqrt{\beta_i^2 \sigma_m^2 + \tilde{\sigma}_i^2}}\end{aligned}$$

If $\beta_i < 0$, an increase of the idiosyncratic volatility $\tilde{\sigma}_i$ increases $\rho_{i,m}$ and decreases the ratio $\sigma_i / \tilde{\sigma}_i^2$. We deduce that the weight is a decreasing function of the specific volatility $\tilde{\sigma}_i$. If $\beta_i > 0$, an increase of the idiosyncratic volatility $\tilde{\sigma}_i$ decreases $\rho_{i,m}$ and decreases the ratio $\sigma_i / \tilde{\sigma}_i^2$. We cannot conclude in this case.

Most diversified portfolio

Question 3

In this question, we illustrate that the MDP may be very different than the minimum variance portfolio.

Most diversified portfolio

Question 3.a

In which case does the MDP coincide with the minimum variance portfolio?

Most diversified portfolio

The MDP coincide with the MV portfolio when the volatility is the same for all the assets.

Most diversified portfolio

Question 3.b

We consider the following parameter values:

i	1	2	3	4
β_i	0.80	0.90	1.10	1.20
$\tilde{\sigma}_i$	0.02	0.05	0.15	0.15

with $\sigma_m = 20\%$. Calculate the unconstrained MDP with Formula (6). Compare it with the unconstrained MV portfolio. What is the result if we consider a long-only portfolio?

Most diversified portfolio

The formula cannot be used directly, because it depends on $\sigma(x^\star)$ and $\mathcal{DR}(x^\star)$. However, we notice that:

$$x_i^\star \propto \frac{\sigma_i}{\tilde{\sigma}_i^2}\left(1 - \frac{\rho_{i,m}}{\rho^\star}\right)$$

It suffices then to rescale these weights to obtain the solution. Using the numerical values of the parameters, $\rho^\star = 98.92\%$ and we obtain the following results:

	β_i	$\rho_{i,m}$	$x_i \in \mathbb{R}$ MDP	$x_i \in \mathbb{R}$ MV	$x_i \geq 0$ MDP	$x_i \geq 0$ MV
$x_1^\star$	0.80	99.23%	−27.94%	211.18%	0.00%	100.00%
$x_2^\star$	0.90	96.35%	43.69%	−51.98%	25.00%	0.00%
$x_3^\star$	1.10	82.62%	43.86%	−24.84%	39.24%	0.00%
$x_4^\star$	1.20	84.80%	40.39%	−34.37%	35.76%	0.00%
$\sigma(x^\star)$			24.54%	13.42%	23.16%	16.12%

Most diversified portfolio

Question 3.c

We assume that the volatility of the assets is 10%, 10%, 50% and 50% whereas the correlation matrix of asset returns is:

$$\rho = \begin{pmatrix} 1.00 & & & \\ 0.90 & 1.00 & & \\ 0.80 & 0.80 & 1.00 & \\ 0.00 & 0.00 & -0.25 & 1.00 \end{pmatrix}$$

Calculate the (unconstrained and long-only) MDP and MV portfolios.

Most diversified portfolio

The results are:

	$x_i \in \mathbb{R}$		$x_i \geq 0$	
	MDP	MV	MDP	MV
$x_1^\star$	−36.98%	60.76%	0.00%	48.17%
$x_2^\star$	−36.98%	60.76%	0.00%	48.17%
$x_3^\star$	91.72%	−18.54%	50.00%	0.00%
$x_4^\star$	82.25%	−2.98%	50.00%	3.66%
$\sigma(x^\star)$	48.59%	6.43%	30.62%	9.57%

Most diversified portfolio

Question 3.d

Comment on these results.

Most diversified portfolio

These two examples show that the MDP may have a different behavior than the minimum variance portfolio. Contrary to the latter, the most diversified portfolio is not necessarily a low-beta or a low-volatility portfolio.

Computation of risk-based portfolios

Exercise

We consider a universe of five assets. Their expected returns are 6%, 10%, 6%, 8% and 12% whereas their volatilities are equal to 10%, 20%, 15%, 25% and 30%. The correlation matrix of asset returns is defined as follows:

$$\rho = \left(\begin{array}{rrrrr} 100\% & & & & \\ 60\% & 100\% & & & \\ 40\% & 50\% & 100\% & & \\ 30\% & 30\% & 20\% & 100\% & \\ 20\% & 10\% & 10\% & -50\% & 100\% \end{array} \right)$$

We assume that the risk-free rate is equal to 2%.

Computation of risk-based portfolios

Question 1

We consider unconstrained portfolios. For each portfolio, compute the risk decomposition.

Computation of risk-based portfolios

Question 1.a

Find the tangency portfolio.

Computation of risk-based portfolios

To compute the unconstrained tangency portfolio, we use the analytical formula (TR-RPB, page 14):

$$x^{\star} = \frac{\Sigma^{-1} \left(\mu - r \mathbf{1}_n\right)}{\mathbf{1}_n^{\top} \Sigma^{-1} \left(\mu - r \mathbf{1}_n\right)}$$

We obtain the following results:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^{\star}$
1	11.11%	6.56%	0.73%	5.96%
2	17.98%	13.12%	2.36%	19.27%
3	2.55%	6.56%	0.17%	1.37%
4	33.96%	9.84%	3.34%	27.31%
5	34.40%	16.40%	5.64%	46.09%

Computation of risk-based portfolios

Question 1.b

Determine the equally weighted portfolio.

Computation of risk-based portfolios

We obtain the following results for the equally weighted portfolio:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	20.00%	7.47%	1.49%	13.43%
2	20.00%	15.83%	3.17%	28.48%
3	20.00%	9.98%	2.00%	17.96%
4	20.00%	9.89%	1.98%	17.80%
5	20.00%	12.41%	2.48%	22.33%

Computation of risk-based portfolios

Question 1.c

Compute the minimum variance portfolio.

Computation of risk-based portfolios

For the minimum variance portfolio, we have:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^{\star}$
1	74.80%	9.08%	6.79%	74.80%
2	−15.04%	9.08%	−1.37%	−15.04%
3	21.63%	9.08%	1.96%	21.63%
4	10.24%	9.08%	0.93%	10.24%
5	8.36%	9.08%	0.76%	8.36%

Computation of risk-based portfolios

Question 1.d

Calculate the most diversified portfolio.

Computation of risk-based portfolios

For the most diversified portfolio, we have:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^{\star}$
1	-14.47%	4.88%	-0.71%	-5.34%
2	4.83%	9.75%	0.47%	3.56%
3	18.94%	7.31%	1.38%	10.47%
4	49.07%	12.19%	5.98%	45.24%
5	41.63%	14.63%	6.09%	46.06%

Computation of risk-based portfolios

Question 1.e

Find the ERC portfolio.

Computation of risk-based portfolios

For the ERC portfolio, we have:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	27.20%	7.78%	2.12%	20.00
2	13.95%	15.16%	2.12%	20.00
3	20.86%	10.14%	2.12%	20.00
4	19.83%	10.67%	2.12%	20.00
5	18.16%	11.65%	2.12%	20.00

Computation of risk-based portfolios

Question 1.f

Compare the expected return $\mu(x)$, the volatility $\sigma(x)$ and the Sharpe ratio $\mathrm{SR}(x \mid r)$ of the different portfolios. Calculate then the tracking error volatility $\sigma(x \mid b)$, the beta $\beta(x \mid b)$ and the correlation $\rho(x \mid b)$ if we assume that the benchmark b is the tangency portfolio.

Computation of risk-based portfolios

We recall the definition of the statistics:

$$
\begin{aligned}
\mu(x) &= \mu^\top x \\
\sigma(x) &= \sqrt{x^\top \Sigma x} \\
\mathrm{SR}(x \mid r) &= \frac{\mu(x) - r}{\sigma(x)} \\
\sigma(x \mid b) &= \sqrt{(x-b)^\top \Sigma (x-b)} \\
\beta(x \mid b) &= \frac{x^\top \Sigma b}{b^\top \Sigma b} \\
\rho(x \mid b) &= \frac{x^\top \Sigma b}{\sqrt{x^\top \Sigma x}\sqrt{b^\top \Sigma b}}
\end{aligned}
$$

Computation of risk-based portfolios

We obtain the following results:

Statistic	$x^\star$	x_{ew}	x_{mv}	x_{mdp}	x_{erc}
$\mu(x)$	9.46%	8.40%	6.11%	9.67%	8.04%
$\sigma(x)$	12.24%	11.12%	9.08%	13.22%	10.58%
$\mathrm{SR}(x \mid r)$	60.96%	57.57%	45.21%	58.03%	57.15%
$\sigma(x \mid b)$	0.00%	4.05%	8.21%	4.06%	4.35%
$\beta(x \mid b)$	100.00%	85.77%	55.01%	102.82%	81.00%
$\rho(x \mid b)$	100.00%	94.44%	74.17%	95.19%	93.76%

We notice that all the portfolios present similar performance in terms of Sharpe Ratio. The minimum variance portfolio shows the smallest Sharpe ratio, but it also shows the lowest correlation with the tangency portfolio.

Computation of risk-based portfolios

Question 2

Same questions if we impose the long-only portfolio constraint.

Computation of risk-based portfolios

The tangency portfolio, the equally weighted portfolio and the ERC portfolio are already long-only. For the minimum variance portfolio, we obtain:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	65.85%	9.37%	6.17%	65.85%
2	0.00%	13.11%	0.00%	0.00%
3	16.72%	9.37%	1.57%	16.72%
4	9.12%	9.37%	0.85%	9.12%
5	8.32%	9.37%	0.78%	8.32%

Computation of risk-based portfolios

For the most diversified portfolio, we have:

Asset	x_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
1	0.00%	5.50%	0.00%	0.00%
2	1.58%	9.78%	0.15%	1.26%
3	16.81%	7.34%	1.23%	10.04%
4	44.13%	12.23%	5.40%	43.93%
5	37.48%	14.68%	5.50%	44.77%

Computation of risk-based portfolios

The results become:

Statistic	$x^{\star}$	x_{ew}	x_{mv}	x_{mdp}	x_{erc}
$\mu(x)$	9.46%	8.40%	6.68%	9.19%	8.04%
$\sigma(x)$	12.24%	11.12%	9.37%	12.29%	10.58%
$\mathrm{SR}(x \mid r)$	60.96%	57.57%	49.99%	58.56%	57.15%
$\sigma(x \mid b)$	0.00%	4.05%	7.04%	3.44%	4.35%
$\beta(x \mid b)$	100.00%	85.77%	62.74%	96.41%	81.00%
$\rho(x \mid b)$	100.00%	94.44%	82.00%	96.06%	93.76%

Building a carry trade exposure

Question 1

We would like to build a carry trade strategy using a *cash neutral* portfolio with equal weights and a notional amount of $100 mn. We use the data given in Table 63. The holding period is equal to three months.

Table 63: Three-month interest rates (March, 15$^{\text{th}}$ 2000)

Currency	AUD	CAD	CHF	EUR	GBP
Interest rate (in %)	5.74	5.37	2.55	3.79	6.21
Currency	JPY	NOK	NZD	SEK	USD
Interest rate (in %)	0.14	5.97	6.24	4.18	6.17

Building a carry trade exposure

Question 1.a

Build the carry trade exposure with two funding currencies and two asset currencies.

Building a carry trade exposure

We rank the currencies according to their interest rate from the lowest to the largest value:

1. JPY	2. CHF	3. EUR	4. SEK	5. CAD
6. AUD	7. NOK	8. USD	9. GBP	10. NZD

We deduce that the carry trade portfolio is:

1. long \$50 mn on NZD
2. long \$50 mn on GBP
3. short \$50 mn on JPY
4. short \$50 mn on CHF

Building a carry trade exposure

Question 1.b

Same question with five funding currencies and two asset currencies.

Building a carry trade exposure

The portfolio becomes:

1. long \$50 mn on NZD and GBP
2. short \$20 mn on JPY, CHF, EUR, SEK and CAD

Building a carry trade exposure

Question 1.c

What is the specificity of the portfolio if we use five funding currencies and five asset currencies.

Building a carry trade exposure

The portfolio is:

1. long \$20 mn on NZD, GBP, USD, NOK and AUD
2. short \$20 mn on JPY, CHF, EUR, SEK and CAD

The asset notional is not equal to the funding notional, because the funding notional is equal to \$100 mn and the asset notional is equal to \$80 mn. Indeed, we don't need to invest the \$20 mn USD exposure since the portfolio currency is the US dollar.

Building a carry trade exposure

Question 1.d

Calculate an approximation of the carry trade P&L if we assume that the spot foreign exchange rates remain constant during the next three months.

Building a carry trade exposure

If we consider the last portfolio, we have:

$$\begin{aligned}
\mathrm{PnL} &\approx 20 \times \frac{1}{4}(6.24\% + 6.21\% + 6.17\% + 5.97\% + 5.74\%) - \\
&\quad 20 \times \frac{1}{4}(0.14\% + 2.55\% + 3.79\% + 4.18\% + 5.37\%) \\
&= \$0.71 \text{ mn}
\end{aligned}$$

If the spot foreign exchange rates remain constant during the next three months, the quarterly P&L is approximated equal to \$710 000.

Building a carry trade exposure

Question 2

We consider the data given in Tables 64 and 65.

Table 64: Three-month interest rates (March, 21$^{\mathrm{th}}$ 2005)

Currency	BRL	CZK	HUF	KRW	MXN
Interest rate (in %)	18.23	2.45	8.95	3.48	8.98
Currency	PLN	SGD	THB	TRY	TWD
Interest rate (in %)	6.63	1.44	2.00	19.80	1.30

Table 65: Annualized volatility of foreign exchange rates (March, 21$^{\mathrm{th}}$ 2005)

Currency	BRL	CZK	HUF	KRW	MXN
Volatility (in %)	11.19	12.57	12.65	6.48	6.80
Currency	PLN	SGD	THB	TRY	TWD
Volatility (in %)	11.27	4.97	4.26	11.61	4.12

Building a carry trade exposure

Question 2.a

Let Σ be the covariance matrix of the currency returns. Which expected returns are used by the carry investor? Write the mean-variance optimization problem if we assume a cash neutral portfolio.

Building a carry trade exposure

Let $\mathcal{C}_i$ and $\mathcal{C} = (\mathcal{C}_1, \ldots, \mathcal{C}_n)$ be the carry of Currency i and the vector of carry values. The carry investor assumes that $\mu_i = \mathcal{C}_i$. We deduce that the mean-variance optimization problem is:

$$\begin{aligned} x^\star(\gamma) &= \arg\min \frac{1}{2} x^\top \Sigma x - \gamma x^\top \mathcal{C} \\ &\text{u.c.} \quad \mathbf{1}_n^\top x = 0 \end{aligned}$$

The constraint $\mathbf{1}_n^\top x = 0$ indicates that the portfolio is cash neutral. If we target a portfolio volatility $\sigma^\star$, we use the bisection algorithm in order to find the optimal value of γ such that:

$$\sigma(x^\star(\gamma)) = \sigma^\star$$

Building a carry trade exposure

Question 2.b

By assuming a zero correlation between the currencies, calibrate the cash neutral portfolio when the objective function is to target a 3% portfolio volatility.

Building a carry trade exposure

We obtain the following solution:

Currency	BRL	CZK	HUF	KRW	MXN
Weight	15.05%	−1.28%	4.11%	−1.57%	14.30%
Currency	PLN	SGD	THB	TRY	TWD
Weight	2.76%	−13.59%	−14.42%	15.52%	−20.87%

Building a carry trade exposure

Question 2.c

Same question if we use the following correlation matrix:

$$\rho = \begin{pmatrix}
1.00 & & & & & & & & & \\
0.30 & 1.00 & & & & & & & & \\
0.38 & 0.80 & 1.00 & & & & & & & \\
0.00 & 0.04 & 0.08 & 1.00 & & & & & & \\
0.50 & 0.30 & 0.34 & 0.12 & 1.00 & & & & & \\
0.35 & 0.70 & 0.78 & 0.06 & 0.30 & 1.00 & & & & \\
0.33 & 0.49 & 0.56 & 0.29 & 0.27 & 0.53 & 1.00 & & & \\
0.30 & 0.34 & 0.34 & 0.38 & 0.29 & 0.35 & 0.53 & 1.00 & & \\
0.43 & 0.39 & 0.48 & 0.10 & 0.38 & 0.41 & 0.35 & 0.43 & 1.00 & \\
0.03 & 0.07 & 0.06 & 0.63 & 0.09 & 0.07 & 0.30 & 0.40 & 0.20 & 1.00
\end{pmatrix}$$

Building a carry trade exposure

The solution becomes:

Currency	BRL	CZK	HUF	KRW	MXN
Weight	13.69%	−9.45%	4.58%	17.31%	6.56%
Currency	PLN	SGD	THB	TRY	TWD
Weight	2.07%	−17.79%	−20.86%	17.98%	−14.10%

Building a carry trade exposure

Question 2.d

Calculate the carry of this optimized portfolio. For each currency, deduce the maximum value of the devaluation (or revaluation) rate that is compatible with a positive P&L.

Building a carry trade exposure

The carry of the portfolio is equal to:

$$\mathcal{C}(x) = \sum_{i=1}^{n} x_i \cdot \mathcal{C}_i$$

We find $\mathcal{C}(x) = 6.7062\%$ per year. We deduce that the maximum value of the devaluation or revaluation rate D_i that is compatible with a positive P&L is equal to:

$$D_i = \frac{6.7062\%}{4} = 1.6765\%$$

This figure is valid for an exposure of 100%.

Building a carry trade exposure

By considering the weights, we deduce that:

$$D_i = -\frac{\mathcal{C}(x)}{4x_i}$$

Finally, we obtain the following compatible devaluation (negative sign $-$) and revaluation (positive sign $+$) rates:

Currency	BRL	CZK	HUF	KRW	MXN
D_i	-12.25%	$+17.75\%$	-36.64%	-9.69%	-25.55%
Currency	PLN	SGD	THB	TRY	TWD
D_i	-81.08%	$+9.43\%$	$+8.04\%$	-9.32%	$+11.89\%$

Building a carry trade exposure

Question 2.e

Repeat Question 2.b assuming that the volatility target is equal 5%. Calculate the leverage ratio. Comment on these results.

Building a carry trade exposure

We obtain the following results:

Currency	BRL	CZK	HUF	KRW	MXN
Weight	25.08%	−2.13%	6.84%	−2.62%	23.83%
Currency	PLN	SGD	THB	TRY	TWD
Weight	4.60%	−22.65%	−24.03%	25.86%	−34.78%

The leverage ratio of this portfolio is equal to $\sum_{i=1}^{n} |x_i| = 172.43\%$, whereas it is equal to 103.47% and 124.37% for the portfolios of Questions 2.b and 2.c. This is perfectly normal because the leverage is proportional to the volatility.

Building a carry trade exposure

Question 2.f

Find the analytical solution of the optimal portfolio $x^{\star}$ when we target a volatility $\sigma^{\star}$.

Building a carry trade exposure

The Lagrange function is equal to:

$$\mathcal{L}(x;\lambda_0) = \frac{1}{2}x^\top \Sigma x - \gamma x^\top \mathcal{C} + \lambda_0 \left(\mathbf{1}_n^\top x - 0\right)$$

The first-order condition is equal to:

$$\frac{\partial\, \mathcal{L}(x;\lambda_0)}{\partial\, x} = \Sigma x - \gamma \mathcal{C} + \lambda_0 \mathbf{1}_n = \mathbf{0}_n$$

It follows that:

$$x = \Sigma^{-1}\left(\gamma \mathcal{C} - \lambda_0 \mathbf{1}_n\right)$$

Building a carry trade exposure

The cash neutral constraint implies that:

$$\mathbf{1}_n^\top \Sigma^{-1} \left(\gamma \mathcal{C} - \lambda_0 \mathbf{1}_n\right) = 0$$

We deduce that:

$$\lambda_0 = \gamma \frac{\mathbf{1}_n^\top \Sigma^{-1} \mathcal{C}}{\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n}$$

Therefore, the optimal solution is equal to:

$$x^\star = \frac{\gamma \Sigma^{-1}}{\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n} \left(\left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right) \mathcal{C} - \left(\mathbf{1}_n^\top \Sigma^{-1} \mathcal{C}\right) \mathbf{1}_n\right)$$

Building a carry trade exposure

The volatility of the optimal portfolio is equal:

$$\begin{aligned}
\sigma^2(x^\star) &= x^{\star\top}\Sigma x^\star \\
&= \left(\gamma\mathcal{C}^\top - \lambda_0 \mathbf{1}_n^\top\right)\Sigma^{-1}\Sigma\Sigma^{-1}\left(\gamma\mathcal{C} - \lambda_0\mathbf{1}_n\right) \\
&= \left(\gamma\mathcal{C}^\top - \lambda_0 \mathbf{1}_n^\top\right)\Sigma^{-1}\left(\gamma\mathcal{C} - \lambda_0\mathbf{1}_n\right) \\
&= \gamma^2\mathcal{C}^\top\Sigma^{-1}\mathcal{C} + \lambda_0^2\mathbf{1}_n^\top\Sigma^{-1}\mathbf{1}_n - 2\gamma\lambda_0\mathcal{C}^\top\Sigma^{-1}\mathbf{1}_n \\
&= \gamma^2\left(\mathcal{C}^\top\Sigma^{-1}\mathcal{C} - \frac{\left(\mathbf{1}_n^\top\Sigma^{-1}\mathcal{C}\right)^2}{\mathbf{1}_n^\top\Sigma^{-1}\mathbf{1}_n}\right) \\
&= \frac{\gamma^2}{\mathbf{1}_n^\top\Sigma^{-1}\mathbf{1}_n}\left(\left(\mathbf{1}_n^\top\Sigma^{-1}\mathbf{1}_n\right)\left(\mathcal{C}^\top\Sigma^{-1}\mathcal{C}\right) - \left(\mathbf{1}_n^\top\Sigma^{-1}\mathcal{C}\right)^2\right)
\end{aligned}$$

Building a carry trade exposure

We deduce that:

$$\gamma = \frac{\sqrt{\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n}}{\sqrt{\left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right)\left(\mathcal{C}^\top \Sigma^{-1} \mathcal{C}\right) - \left(\mathbf{1}_n^\top \Sigma^{-1} \mathcal{C}\right)^2}} \sigma\left(x^\star\right)$$

Finally, we obtain:

$$\begin{aligned} x^\star &= \sigma\left(x^\star\right) \frac{\Sigma^{-1}\left(\left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right) \mathcal{C} - \left(\mathbf{1}_n^\top \Sigma^{-1} \mathcal{C}\right) \mathbf{1}_n\right)}{\sqrt{\left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right)^2 \left(\mathcal{C}^\top \Sigma^{-1} \mathcal{C}\right) - \left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right)\left(\mathbf{1}_n^\top \Sigma^{-1} \mathcal{C}\right)^2}} \\ &= \sigma^\star \frac{\Sigma^{-1}\left(\left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right) \mathcal{C} - \left(\mathbf{1}_n^\top \Sigma^{-1} \mathcal{C}\right) \mathbf{1}_n\right)}{\sqrt{\left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right)^2 \left(\mathcal{C}^\top \Sigma^{-1} \mathcal{C}\right) - \left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right)\left(\mathbf{1}_n^\top \Sigma^{-1} \mathcal{C}\right)^2}} \end{aligned}$$

Building a carry trade exposure

Question 2.g

We assume that the correlation matrix is the identity matrix I_n. Find the expression of the threshold value $\mathcal{C}^{\star}$ such that all currencies with a carry $\mathcal{C}_i$ larger than $\mathcal{C}^{\star}$ form the long leg of the portfolio.

Building a carry trade exposure

We recall that:

$$x^{\star} \propto \Sigma^{-1}\left(\left(\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n\right) \mathcal{C} - \left(\mathbf{1}_n^\top \Sigma^{-1} \mathcal{C}\right) \mathbf{1}_n\right)$$

If $\rho = I_n$, we have:

$$\mathbf{1}_n^\top \Sigma^{-1} \mathbf{1}_n = \sum_{j=1}^{n} \frac{1}{\sigma_j^2}$$

and:

$$\mathbf{1}_n^\top \Sigma^{-1} \mathcal{C} = \sum_{j=1}^{n} \frac{\mathcal{C}_j}{\sigma_j^2}$$

We deduce that:

$$x_i^{\star} \propto \frac{1}{\sigma_i^2}\left(\left(\sum_{j=1}^{n} \frac{1}{\sigma_j^2}\right) \mathcal{C}_i - \left(\sum_{j=1}^{n} \frac{\mathcal{C}_j}{\sigma_j^2}\right)\right)$$

Building a carry trade exposure

The portfolio is long on the currency i if:

$$\mathcal{C}_i \geq \mathcal{C}^\star$$

where:

$$\mathcal{C}^\star = \left(\sum_{j=1}^{n} \frac{1}{\sigma_j^2} \right)^{-1} \left(\sum_{j=1}^{n} \frac{\mathcal{C}_j}{\sigma_j^2} \right) = \sum_{j=1}^{n} \omega_j \mathcal{C}_j$$

and:

$$\omega_j = \frac{\sigma_j^{-2}}{\sum_{k=1}^{n} \sigma_k^{-2}}$$

$\mathcal{C}^\star$ is the weighted mean of the carry values and the weights are inversely proportional to the variance of the currency returns.

Main references

- RONCALLI, T. (2013)
  Introduction to Risk Parity and Budgeting, Chapman and Hall/CRC Financial Mathematics Series, Chapter 3.
- RONCALLI, T. (2013)
  Introduction to Risk Parity and Budgeting — Companion Book, Chapman and Hall/CRC Financial Mathematics Series, Chapter 3.
- RONCALLI, T. (2017)
  Alternative Risk Premia: What Do We Know?, in Jurczenko, E. (Ed.), *Factor Investing and Alternative Risk Premia*, ISTE Press – Elsevier.

References I

Acharya, V.V., and Pedersen, L.H. (2005)
Asset Pricing with Liquidity Risk, *Journal of Financial Economics*, 77(2), pp. 375-410.

Baku, E., Fortes, R., Hervé, K., Lezmi, E., Malongo, H., Roncalli, T., and Xu, J. (2019)
Factor Investing in Currency Markets: Does it Make Sense?, *Amundi Working Paper*, 89, www.research-center.amundi.com.

Baku, E., Fortes, R., Hervé, K., Lezmi, E., Malongo, H., Roncalli, T., and Xu, J. (2020)
Factor Investing in Currency Markets: Does it Make Sense?, *Journal of Portfolio Management*, 46(2), pp. 141-155.

References II

BARBERIS, N., and THALER, R.H. (2003)
A Survey of Behavioral Finance, in Constantinides, G.M., Harris, M. and Stulz, R.M. (Eds), *Handbook of the Economics of Finance*, 1(B), Elsevier, pp. 1053-1128.

BEKTIC, D., WENZLER, J.S., WEGENER, M., SCHIERECK, D., and SPIELMANN, T. (2019)
Extending Fama-French Factors to Corporate Bond Markets, *Journal of Portfolio Management*, 45(3), pp. 141-158.

BEKTIC, D., NEUGEBAUER, U., WEGENER, M. and WENZLER, J.S. (2017)
Common Equity Factors in Corporate Bond Markets, in Jurczenko, E. (Ed.), *Factor Investing and Alternative Risk Premia*, ISTE Press-Elsevier.

References III

- Ben Slimane, M., De Jong, M., Dumas, J-M., Fredj, H., Sekine, T., and Srb, M. (2019)
  Traditional and Alternative Factors in Investment Grade Corporate Bond Investing, *Amundi Working Paper*, 78, www.research-center.amundi.com.
- Black, F. (1972)
  Capital Market Equilibrium with Restricted Borrowing, *Journal of Business*, 45(3), pp. 444-455.
- Bruder, B., Dao, T.L., Richard, J.C., and Roncalli, T. (2011)
  Trend Filtering Methods for Momentum Strategies, *SSRN*, www.ssrn.com/abstract=2289097.
- Carhart, M.M. (1997)
  On Persistence in Mutual Fund Performance, *Journal of Finance*, 52(1), pp. 57-82.

References IV

- Cazalet, Z., and Roncalli, T. (2014)
  Facts and Fantasies About Factor Investing, *SSRN*,
  www.ssrn.com/abstract=2524547.
- Choueifaty, Y., and Coignard, Y. (2008)
  Toward Maximum Diversification, *Journal of Portfolio Management*,
  35(1), pp. 40-51.
- Choueifaty, Y., Froidure, T., and Reynier, J. (2013)
  Properties of the Most Diversified Portflio, *Journal of Investment Strategies*, 2(2), pp. 49-70.
- Clarke, R.G., de Silva, H., and Thorley, S. (2011)
  Minimum Variance Portfolio Composition, *Journal of Portfolio Management*, 37(2), pp. 31-45.

References V

CLARKE, R.G., DE SILVA, H., and THORLEY, S. (2013)
Risk Parity, Maximum Diversification, and Minimum Variance: An Analytic Perspective, *Journal of Portfolio Management*, 39(3), pp. 39-53.

COCHRANE, J.H. (2001)
Asset Pricing, Princeton University Press.

COCHRANE, J.H. (2011)
Presidential Address: Discount Rates, *Journal of Finance*, 66(4), pp. 1047-1108.

DEMIGUEL, V., GARLAPPI, L., and UPPAL, R. (2009)
Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?, *Review of Financial Studies*, 22(5), pp. 1915-1953.

References VI

Fama, E.F., and French, K.R. (1992)
The Cross-Section of Expected Stock Returns, *Journal of Finance*, 47(2), pp. 427-465.

Fama, E.F., and French, K.R. (1993)
Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, 33(1), pp. 3-56.

Frazzini, A, and Pedersen, L.H. (2014)
Betting Against Beta, *Journal of Financial Economics*, 111(1), pp. 1-25.

Grinblatt, M., Titman, S., and Wermers, R. (1995)
Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior, *American Economic Review*, 85(5), pp. 1088-1105.

References VII

- Hamdan, R., Pavlowsky, F., Roncalli, T., and Zheng, B. (2016)
  A Primer on Alternative Risk Premia, *SSRN*,
  www.ssrn.com/abstract=2766850.

- Haugen, R.A. and Baker, N.L. (1991)
  The Efficient Market Inefficiency of Capitalization-weighted Stock Portfolios, *Journal of Portfolio Management*, 17(3), pp. 35-40.

- Harvey, C.R., Liu, Y., and Zhu, H. (2016)
  ...and the Cross-Section of Expected Returns, *Review of Financial Studies*, 29(1), pp. 5-68.

- Houweling, P., and Van Zundert, J. (2017)
  Factor Investing in the Corporate Bond Market, *Financial Analysts Journal*, 73(2), pp. 100-115.

References VIII

Ilmanen, A. (2011)
Expected Returns: An Investor's Guide to Harvesting Market Rewards, Wiley.

Israel, R., Palhares, D., and Richardson, S.A. (2018)
Common Factors in Corporate Bond and Bond Fund Returns, *Journal of Investment Management*, 16(2), pp. 17-46

Jensen, M.C. (1968)
The Performance of Mutual Funds in the Period 1945-1964, *Journal of Finance*, 23(2), pp. 389-416.

Jurczenko, E., Michel, T. and Teïletche, J. (2015)
A Unified Framework for Risk-based Investing, *Journal of Investment Strategies*, 4(4), pp. 1-29.

References IX

Jusselin, P., Lezmi, E., Malongo, H., Masselin, C., Roncalli, T., and Dao, T-L. (2017)
Understanding the Momentum Risk Premium: An In-Depth Journey Through Trend-Following Strategies, *SSRN*, www.ssrn.com/abstract=3042173.

Koijen, R.S.J., Moskowitz, T.J., Pedersen, L.H., and Vrugt, E.B. (2018)
Carry, *Journal of Financial Economics*, 127(2), pp. 197-225.

Lucas, R.E. (1978)
Asset Prices in an Exchange Economy, *Econometrica*, 46(6), pp. 1429-1445.

MacDonald, R. (1995)
Long-Run Exchange Rate Modeling: a Survey of the Recent Evidence, *IMF Staff Papers*, 42(3), pp. 437-489.

References X

Menkhoff, L., Sarno, L., Schmeling, M., and Schrimpf, A. (2016)
Currency Value, *Review of Financial Studies*, 30(2), pp. 416-441.

Roncalli, T. (2017)
ESG & Factor Investing: A New Stage Has Been Reached, *Amundi ViewPoint*, www.research-center.amundi.com.

Roncalli, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series.

Scherer, B. (2011)
A Note on the Returns from Minimum Variance Investing, *Journal of Empirical Finance*, 18(4), pp. 652-660.

Asset Management
Lecture 4. Green and Sustainable Finance, ESG Investing and Climate Risk

Thierry Roncalli*

*University of Paris-Saclay

January 2021

Agenda

- Lecture 1: Portfolio Optimization
- Lecture 2: Risk Budgeting
- Lecture 3: Smart Beta, Factor Investing and Alternative Risk Premia
- **Lecture 4: Green and Sustainable Finance, ESG Investing and Climate Risk**
- Lecture 5: Machine Learning in Asset Management

Introduction to sustainable finance

Sustainable investing

Sustainable investing is an investment approach that considers environmental, social and governance (ESG) factors in portfolio selection and management

Socially responsible investing (SRI)

Socially responsible investing (SRI) is an investment strategy that is considered socially responsible, because it invests in companies that have ethical practices

Environmental, Social and Governance (ESG)

Environmental, Social, and Corporate Governance (ESG) refers to the factors that measure the sustainability of an investment

Introduction to sustainable finance

Sustainable Investing

$\approx$

Socially Responsible Investing (SRI)

$\approx$

Environmental, Social, and Governance (ESG)

Introduction to sustainable finance

Figure 81: The raison d'être of ESG investing

Introduction to sustainable finance

ESG financial ecosystem

- Asset owners (pension funds, sovereign wealth funds (SWF), insurance and institutional investors, retail investors, etc.)
- Asset managers
- ESG rating agencies
- ESG index sponsors
- Banks
- ESG associations (GSIA, UNPRI, etc.)
- Regulators and international bodies (governments, financial and industry regulators, central banks, etc.)
- **Issuers** (equities, bonds, loans, etc.)

ESG investing $\Leftrightarrow$ ESG financing

ESG regulations

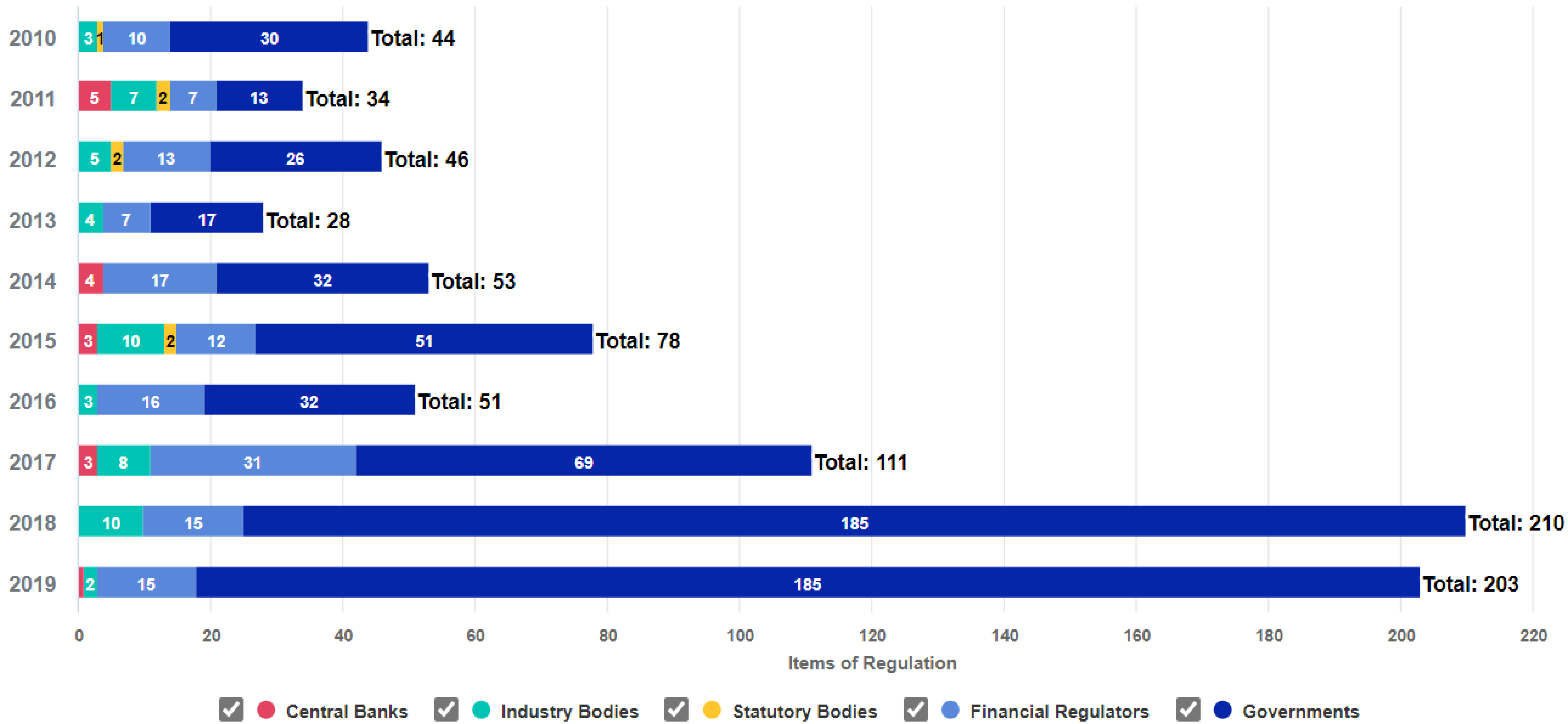


Figure 82: List of ESG regulations (MSCI, Who will regulate ESG?)

ESG regulations

Visit the MSCI website

`https://www.msci.com/who-will-regulate-esg`

and obtain the detailed list of regulations

by year, country, regulator, regulated investors, etc.

ESG regulators

The example of ESMA

ESMA strategy on sustainable finance

1. Completing the regulatory framework on transparency obligations via the Disclosures Regulation (joint technical standards with EBA and EIOPA)
2. TRV (trends, risks and vulnerabilities) reporting of sustainable finance
3. Analyse financial risks from climate change, including potentially climate-related stress testing
4. Convergence of national supervisory practices on ESG factors
5. Participating in the EU taxonomy on sustainable finance
6. Ensuring ESG guidelines are implemented by regulated entities (e.g. asset managers)

ESG regulators
The example of central banks

Figure 83: Network of Central Banks and Supervisors for Greening the Financial System (NGFS)

Go the NGFS website (`https://www.ngfs.net`) and download the NGFS climate scenarios

ESG associations

GLOBAL SUSTAINABLE INVESTMENT ALLIANCE

Figure 84: Global Sustainable Investment Alliance (GSIA)

http://www.gsi-alliance.org

ESG associations

GSIA members

- The European Sustainable Investment Forum (Eurosif), `http://www.eurosif.org`
- Responsible Investment Association Australasia (RIAA), `https://responsibleinvestment.org`
- Responsible Investment Association Canada (RIA Canada), `https://www.riacanada.ca`
- UK Sustainable Investment & Finance Association (UKSIF), `https://www.uksif.org`
- The Forum for Sustainable & Responsible Investment (US SIF), `https://www.ussif.org`
- Dutch Association of Investors for Sustainable Development (VBDO), `https://www.vbdo.nl/en/`
- Japan Sustainable Investment Forum (JSIF), `https://japansif.com/english`

ESG associations

Figure 85: Principles for Responsible Investment (PRI)

https://www.unpri.org

ESG associations

PRI (or UNPRI)

- Early 2005: UN Secretary-General Kofi Annan invited a group of the world's largest institutional investors to join a process to develop the Principles for Responsible Investment
- April 2006: The Principles were launched at the New York Stock Exchange
- 6 ESG principles

ESG associations

Signatories' commitment

"As institutional investors, we have a duty to act in the best long-term interests of our beneficiaries. In this fiduciary role, we believe that environmental, social, and corporate governance (ESG) issues can affect the performance of investment portfolios (to varying degrees across companies, sectors, regions, asset classes and through time).
We also recognise that applying these Principles may better align investors with broader objectives of society. Therefore, where consistent with our fiduciary responsibilities, we commit to the following:

- Principle 1: We will incorporate ESG issues into investment analysis and decision-making processes.
- Principle 2: We will be active owners and incorporate ESG issues into our ownership policies and practices.
- Principle 3: We will seek appropriate disclosure on ESG issues by the entities in which we invest.
- Principle 4: We will promote acceptance and implementation of the Principles within the investment industry.
- Principle 5: We will work together to enhance our effectiveness in implementing the Principles.
- Principle 6: We will each report on our activities and progress towards implementing the Principles.

The Principles for Responsible Investment were developed by an international group of institutional investors reflecting the increasing relevance of environmental, social and corporate governance issues to investment practices. The process was convened by the United Nations Secretary-General.
In signing the Principles, we as investors publicly commit to adopt and implement them, where consistent with our fiduciary responsibilities. We also commit to evaluate the effectiveness and improve the content of the Principles over time. We believe this will improve our ability to meet commitments to beneficiaries as well as better align our investment activities with the broader interests of society.
We encourage other investors to adopt the Principles."

Source: `https://www.unpri.org`

ESG associations

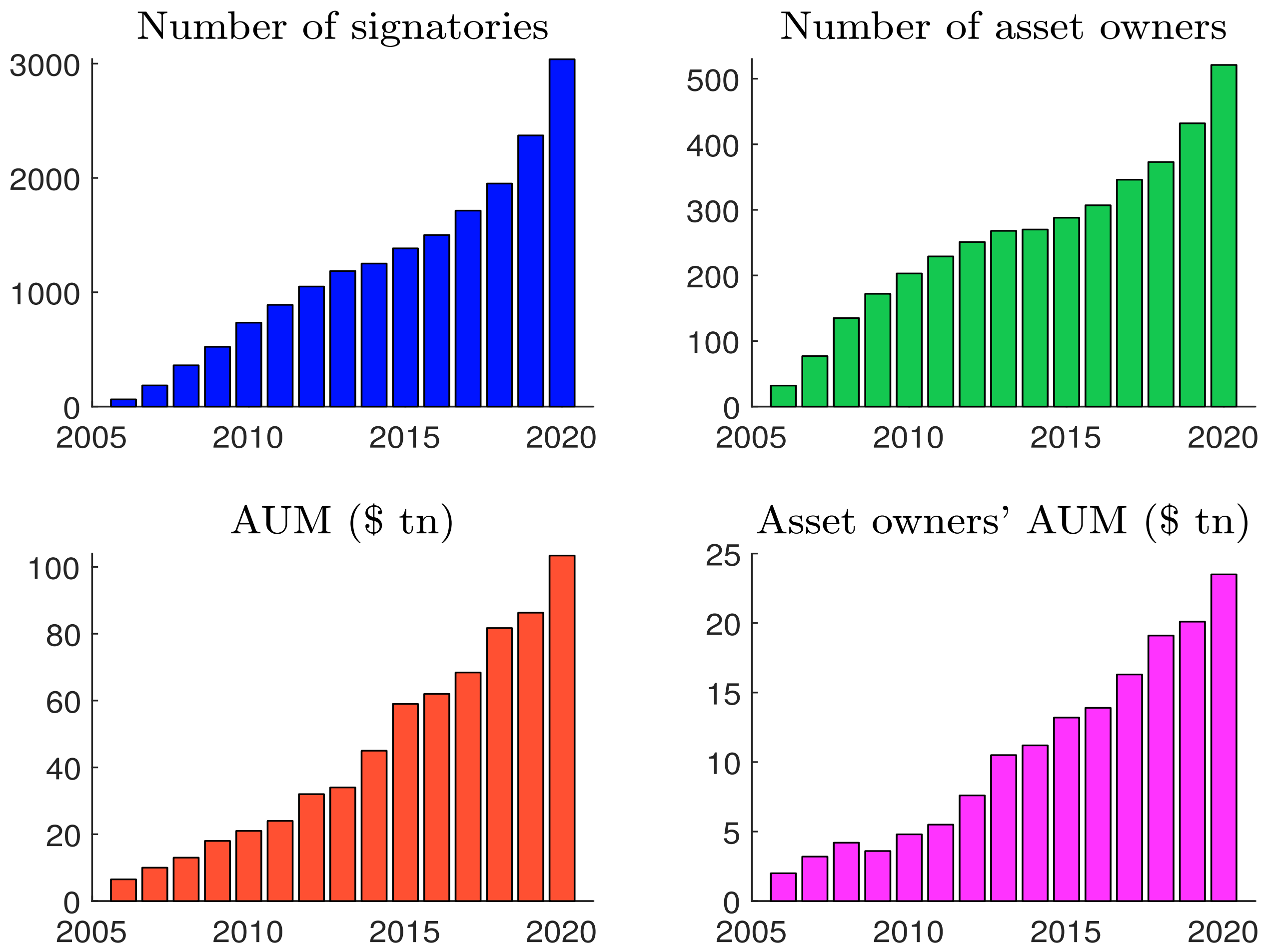


Figure 86: PRI Signatory growth

Source: `https://www.unpri.org`

The issuer point of view of ESG

Corporate financial performance (CFP)

- Friedman (2007)
- Shareholder theory
- Corporations have no social responsibility to the public or society
- Their only responsibility is to its shareholders (profit maximization)

Corporate social responsibility (CSR)

- Freeman (2010)
- Stakeholder theory
- Corporations create negative externalities
- They must have social and moral responsibilities
- Impact on the cost-of-capital and business risk

ESG strategies

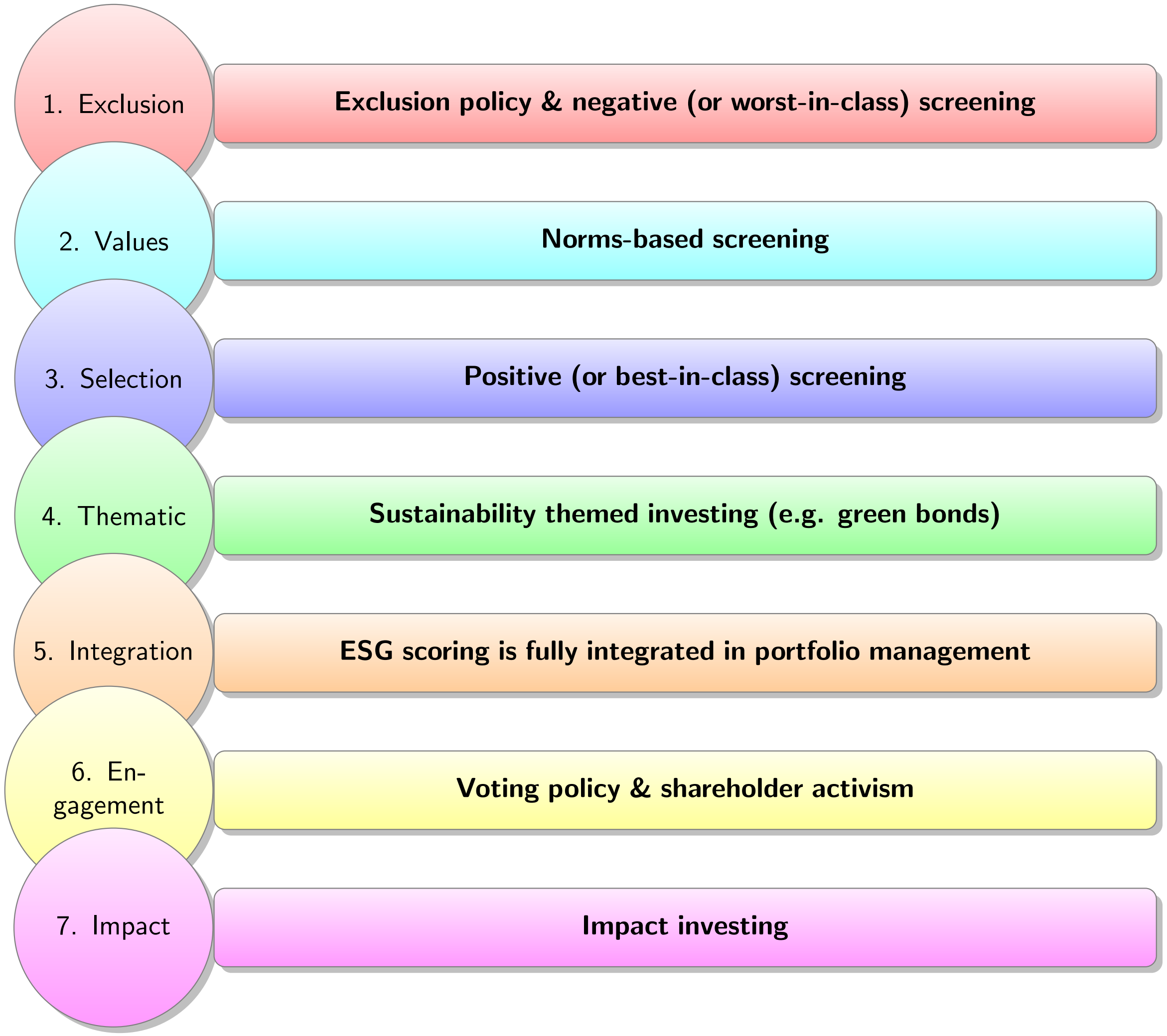


Figure 87: Categorisation of ESG strategies (Eurosif, 2019)

ESG strategies

Exclusion/Negative Screening

The exclusion from a fund or portfolio of certain sectors, companies or practices based on specific ESG criteria (worst-in-class)

Values/Norms-based Screening (or Red Flags)

Screening of investments against minimum standards of business practice based on international norms, such as those issued by the OECD, ILO, UN and UNICEF[a]

[a] In Europe, the top exclusion criteria are (1) controversial weapons (Ottawa and Oslo treaties), (2), tobacco, (3) all weapons, (4) gambling, (5) pornography, (6) nuclear energy, (7) alcohol, (8) GMO and (9) animal testing (Eurosif, 2019)

Source: Global Sustainable Investment Alliance (2018)

ESG strategies

Selection/Positive Screening

Investment in sectors, companies or projects selected for positive ESG performance relative to industry peers (best-in-class)

Thematic/Sustainability Themed Investing

Investment in themes or assets specifically related to sustainability (for example clean energy, green technology or sustainable agriculture)

ESG Integration

The systematic and explicit inclusion by investment managers of environmental, social and governance factors into financial analysis

Source: Global Sustainable Investment Alliance (2018)

ESG strategies

Engagement/Shareholder Action

The use of shareholder power to influence corporate behavior, including through direct corporate engagement (i.e., communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is guided by comprehensive ESG guidelines.

Impact Investing

Targeted investments aimed at solving social or environmental problems, and including community investing, where capital is specifically directed to traditionally underserved individuals or communities, as well as financing that is provided to businesses with a clear social or environmental purpose

Source: Global Sustainable Investment Alliance (2018)

The market of ESG investing

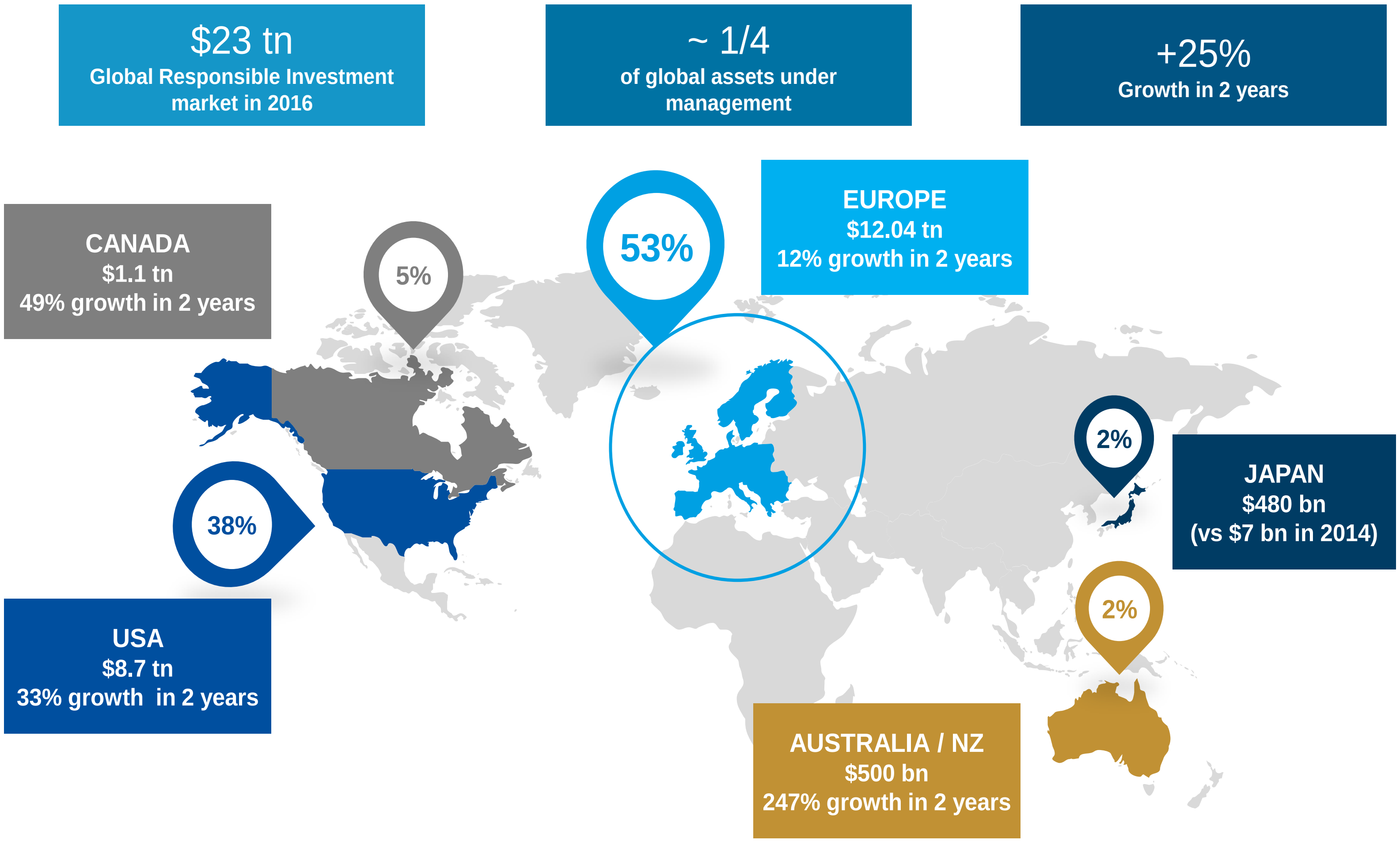


Figure 88: ESG at the start of 2016

Source: Global Sustainable Investment Alliance (2017)

The market of ESG investing

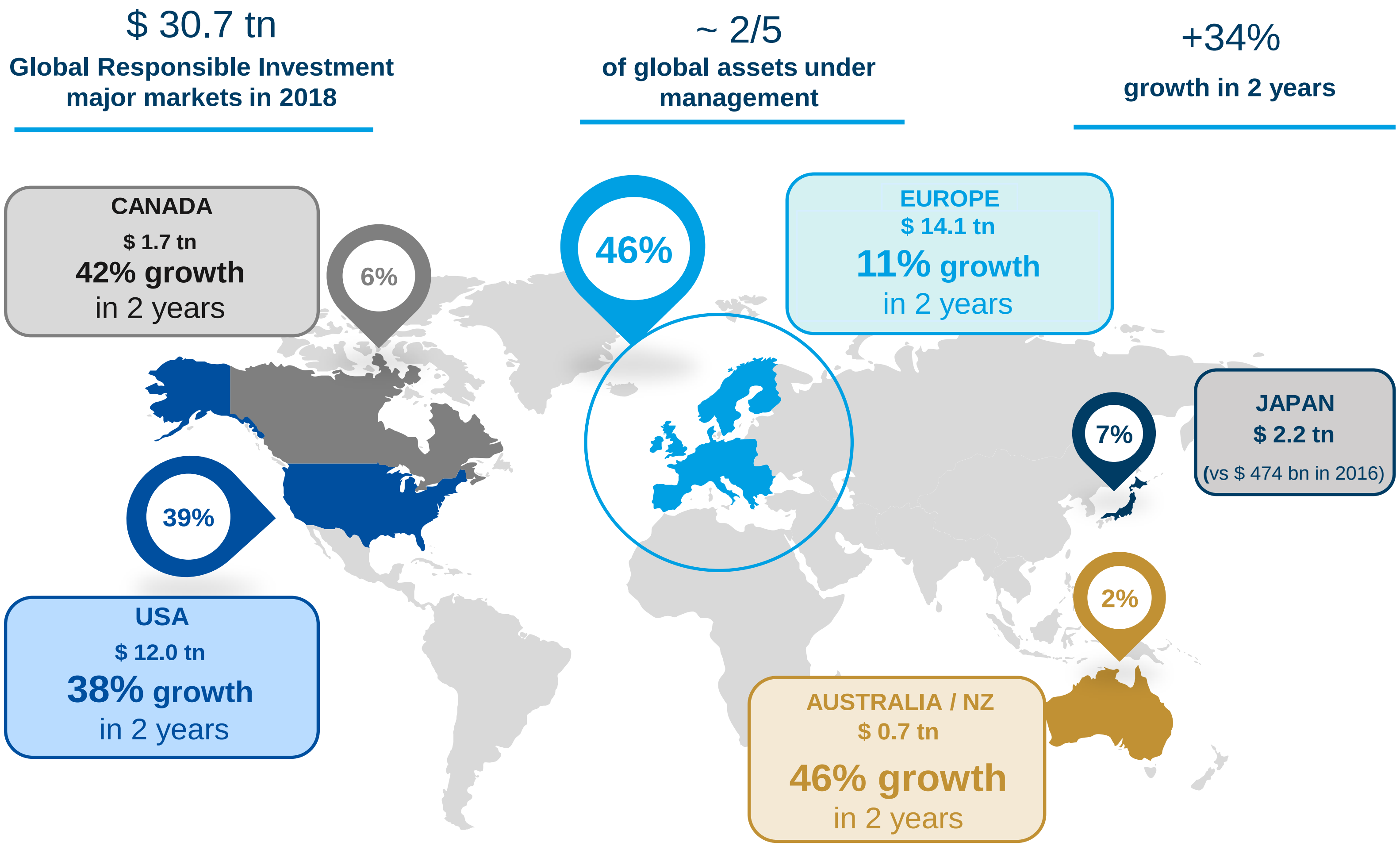


Figure 89: ESG at the start of 2018

Source: Global Sustainable Investment Alliance (2019)

The market of ESG investing

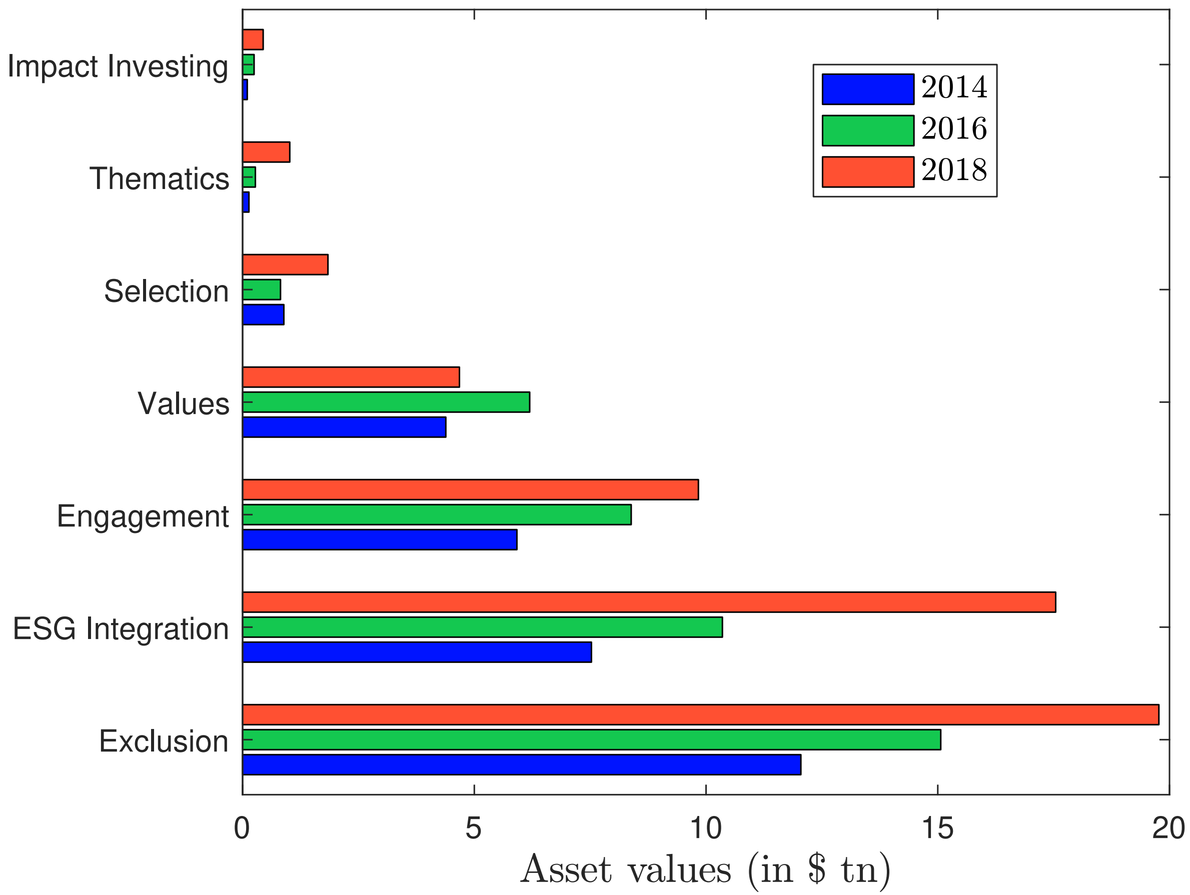


Figure 90: Asset values of ESG strategies between 2014 and 2018

Source: Global Sustainable Investment Alliance (2015, 2017, 2019)

The market of ESG investing

Table 66: Annual growth of ESG strategies

	2014-2016	2016-2018
Exclusion	11.7%	14.6%
ESG Integration	17.4%	30.2%
Engagement	18.9%	8.3%
Values	19.0%	−13.1%
Selection	7.6%	50.1%
Thematics	55.1%	92.0%
Impact Investing	56.8%	33.7%

Source: Global Sustainable Investment Alliance (2015, 2017, 2019)

The concept of ESG investing

Environmental, **S**ocial and **G**overnance (ESG)

- ESG **analysis**: extra-financial analysis $\neq$ financial analysis
- ESG **scoring**: quantitative measures of ESG dimensions
- ESG **ratings**: provide a grade (e.g. AAA, AA, A, etc.) to an issuer ($\approx$ credit ratings)
- ESG **screening**: process of scanning and filtering issuers based on ESG analysis and scoring ($\approx$ stock screening, bond screening, stock picking)
- ESG **investment process**: define how the investment process integrates ESG
- ESG **reporting**: provide ESG information and measures on the investment portfolio (e.g. ESG risk of the portfolio vs ESG risk of the benchmark, repartition of ESG ratings, top/bottom ESG issuers, etc.)

ESG rating agencies

Major players

- ISS ESG (Deutsche Börse)
- MSCI ESG
- Sustainalytics (Morningstar)
- Thomson Reuters
- Vigeo-Eiris (Moody's)

Other players

- Beyond Ratings (LSE)
- Bloomberg ESG
- RobecoSAM (S&P)
- Refinitiv (LSE)
- TrueValue Labs (Factset)

Specialized climate data providers

- CDP
- Trucost (S&P)

ESG data

- ESG requires a lot of data and **alternative** data
- For example, Sustainalytics ESG Data includes 220 ESG indicators and 450 fields, and covers over 12 000 companies
- Where to find the data?
  - Public data
    - Standardized data (regulatory reporting)
    - Non-standardized data (self reporting)
  - Private data
  - Proprietary data
  - Questionnaire/survey
  - Analyst scores

ESG data

Examples of data

- Corporate annual reports
- Corporate environmental and social reports
- Carbon Disclosure Project (CDP) responses
- US Bureau of Labor Statistics
- Thomson Financial
- World Bank (WB)

ESG data

Examples of alternative data

- Energy Data Analytics Lab research (Duke university)
  `https://energy.duke.edu/research/energy-data/resources`
- Food and Agriculture Organization (FAO)
  `http://www.fao.org`
- UK Reporting of Injuries, Diseases and Dangerous Occurrences Regulations (RIDDOR)
  `https://www.hse.gov.uk/riddor`
- World Health Organization (WHO)
  `https://www.who.int`
- World Bank Governance Indicators (WGI)
  `https://info.worldbank.org/governance/wgi`
- World Resources Institute (WRI)
  `https://www.wri.org`

ESG (alternative) data

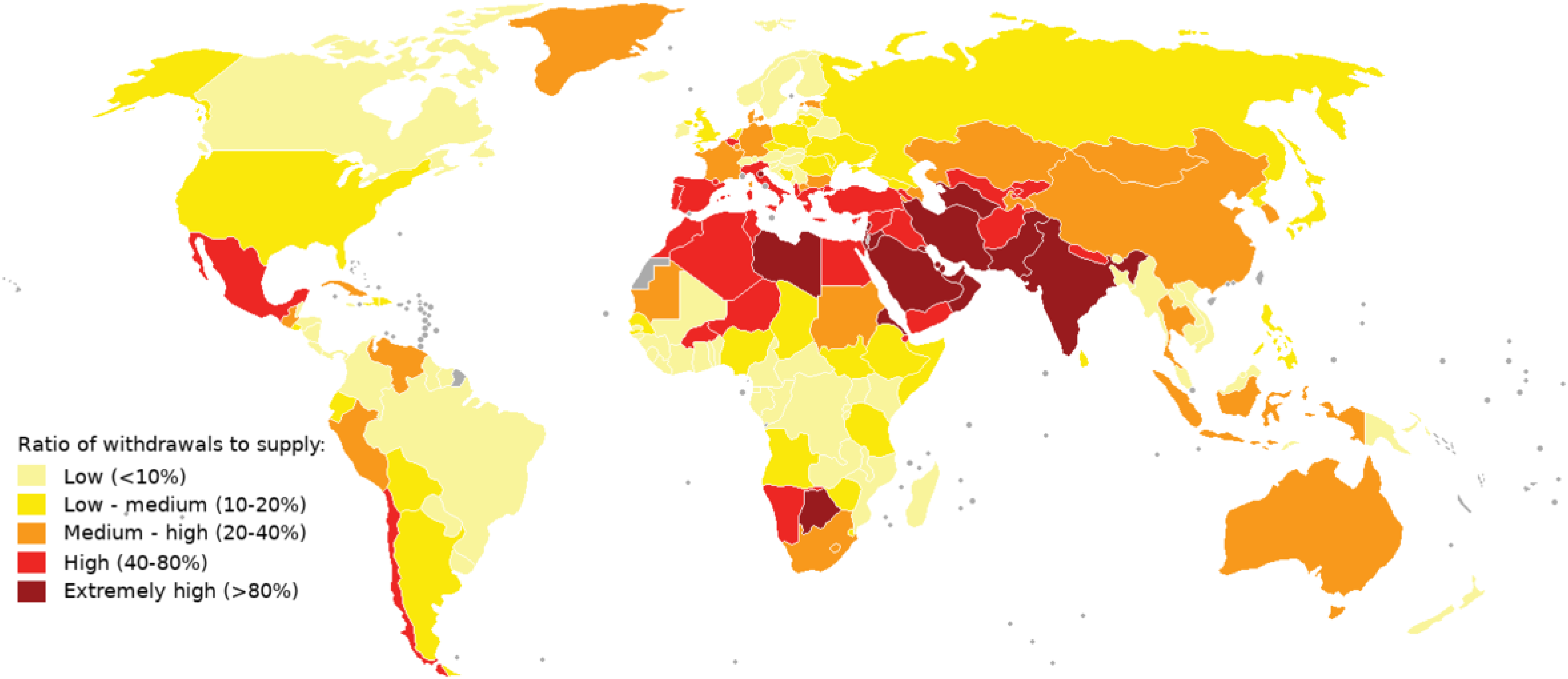


Figure 91: WRI Water Stress 2019

Source: World Resources Institute (WRI), www.wri.org

ESG (alternative) data

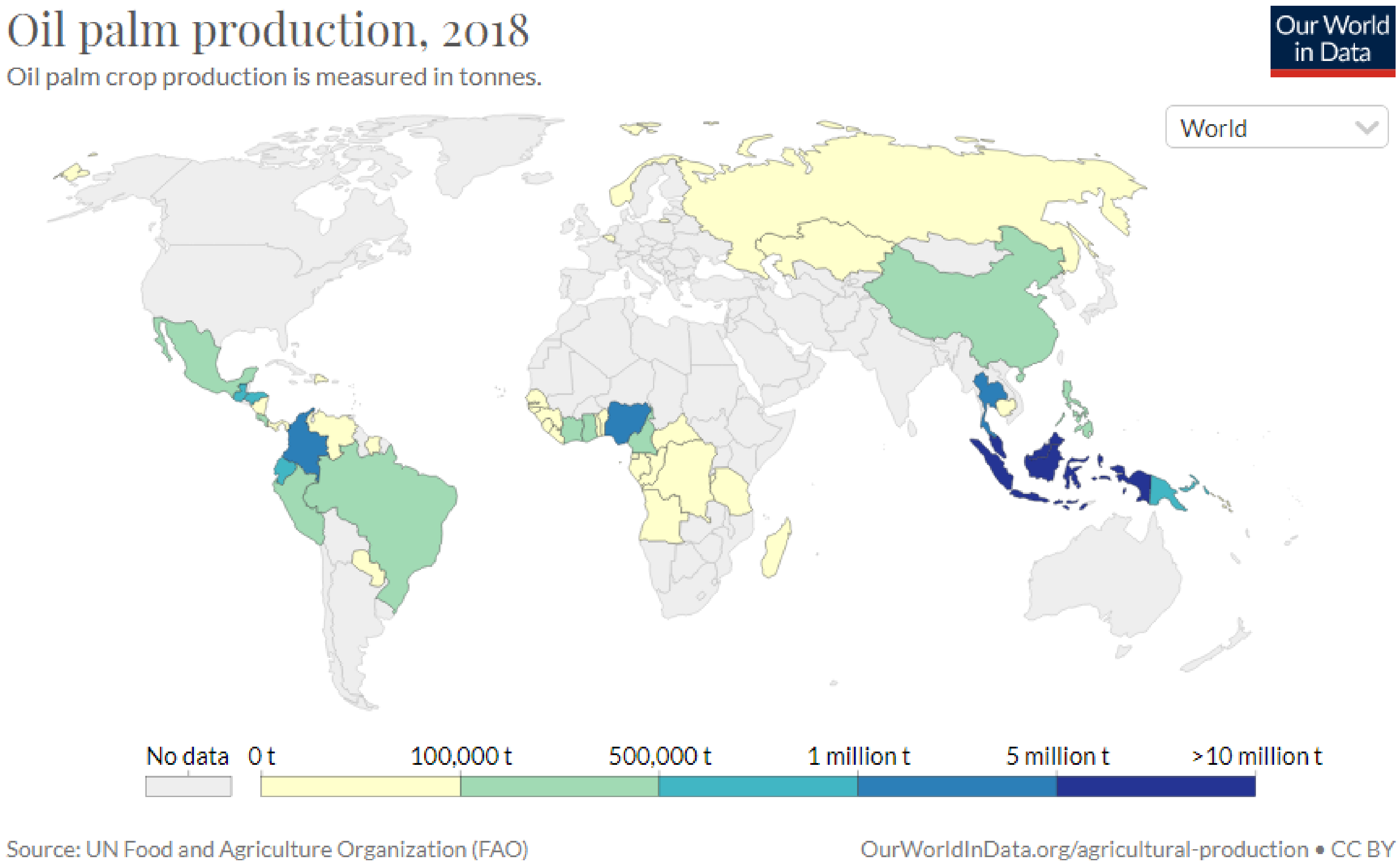


Figure 92: Oil palm production in 2018

Source: Our World in Data, `https://ourworldindata.org/grapher/palm-oil-production`

ESG (alternative) data

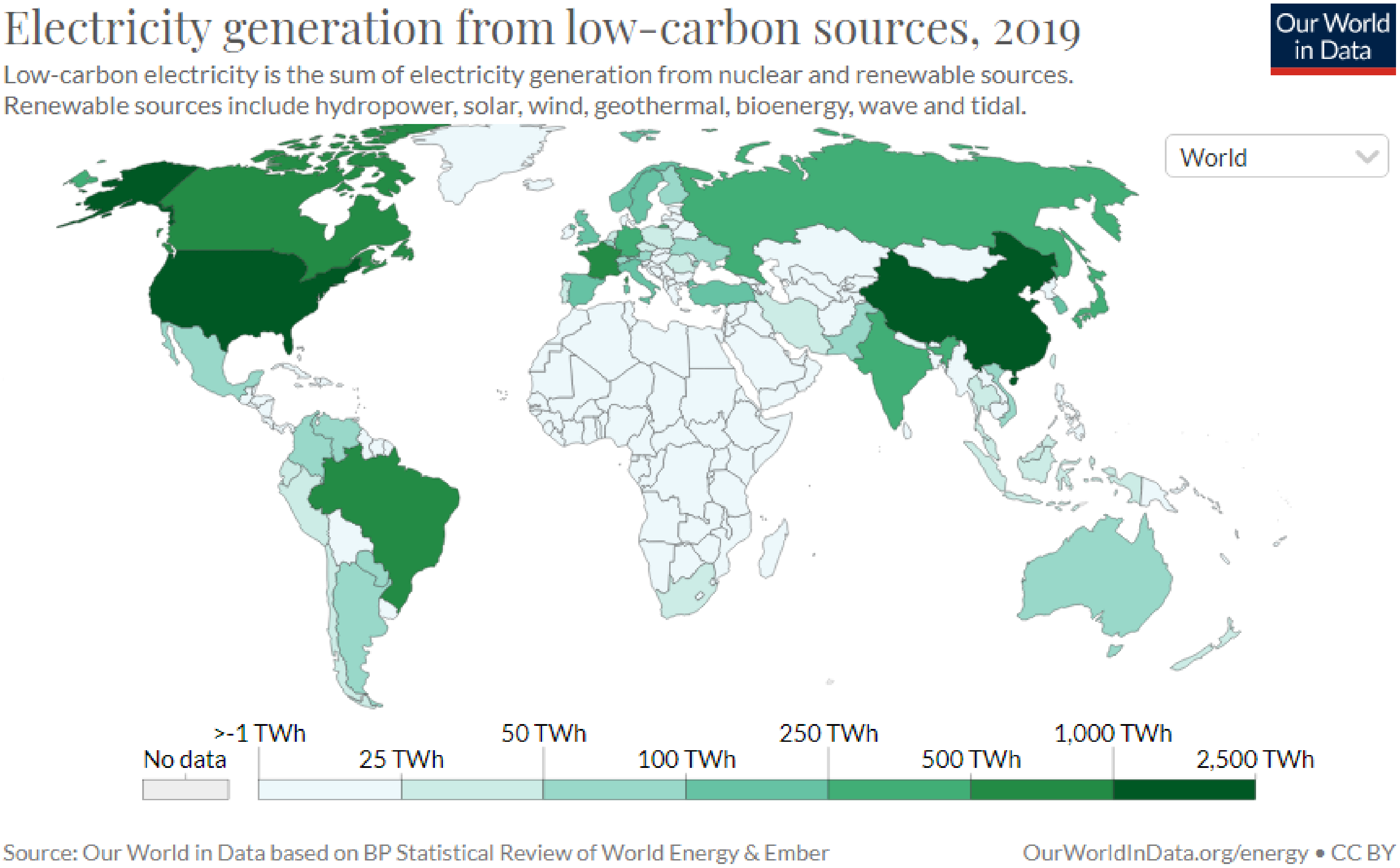


Figure 93: Electricity generation from low-carbon sources in 2019

Source: Our World in Data, https://ourworldindata.org/grapher/low-carbon-electricity

ESG scoring system

Table 67: An example of ESG criteria (corporate issuers)

Environmental	Social	Governance
Carbon emissions	Employment conditions	Board independence
Energy use	Community involvement	Corporate behaviour
Pollution	Gender equality	Audit and control
Waste disposal	Diversity	Executive compensation
Water use	Stakeholder opposition	Shareholder' rights
Renewable energy	Access to medicine	CSR strategy
Green cars*		
Green financing*		

$^{(\star)}$means a specific criterion related to one or several sectors (Green cars $\Rightarrow$ Automobiles, Green financing $\Rightarrow$ Financials)

ESG scoring system

Table 68: An example of ESG criteria (sovereign issuers)

Environmental

- Carbon emissions
- Energy transition risk
- Fossil fuel exposure
- Emissions reduction target
- Physical risk exposure
- Green economy

Social

- Income inequality
- Living standards
- Non-discrimination
- Health & security
- Local communities and human rights
- Social cohesion
- Access to education

Governance

- Political stability
- Institutional strength
- Levels of corruption
- Rule of law
- Government and regulatory effectiveness
- Rights of shareholders

ESG scoring system

Sovereign ESG Data Framework

- World Bank
- Data may be download at the following webpage:
  `https://datatopics.worldbank.org/esg/framework.html`
- **E**: 27 variables
- **S**: 22 variables
- **G**: 18 variables

ESG scoring system

Table 69: Sovereign ESG Data Framework (World Bank)

Environmental

- Emissions & pollution (5)
- Natural capital endowment and management (6)
- Energy use & security (7)
- Environment/ climate risk & resilience (6)
- Food security (3)

Social

- Education & skills (3)
- Employment (3)
- Demography (3)
- Poverty & inequality (4)
- Health & nutrition (5)
- Access to services (4)

Governance

- Human rights (2)
- Government effectiveness (2)
- Stability & rule of law (4)
- Economic environment (3)
- Gender (4)
- Innovation (3)

ESG scoring system

- Most of ESG scoring systems are based on scoring trees
- Raw data are normalized in order to obtain features $X_1, \ldots, X_m$
- Features $X_1, \ldots, X_m$ are aggregated to obtain sub-scores $\mathcal{S}_1, \ldots, \mathcal{S}_n$:

$$\mathcal{S}_i = \sum_{j=1}^{m} \omega_{i,j}^{(1)} X_j$$

- Sub-scores $\mathcal{S}_1, \ldots, \mathcal{S}_n$ are aggregated to obtain the final score $\mathcal{S}$:

$$\mathcal{S} = \sum_{i=1}^{n} \omega_{i}^{(2)} \mathcal{S}_i$$

The two-level tree structure can be extended to multi-level tree structures
For example, in the case of a three-level tree structure, we have:

$$\text{Features} \Rightarrow \text{sub-sub-scores} \Rightarrow \text{sub-scores} \Rightarrow \text{final score}$$

ESG scoring system

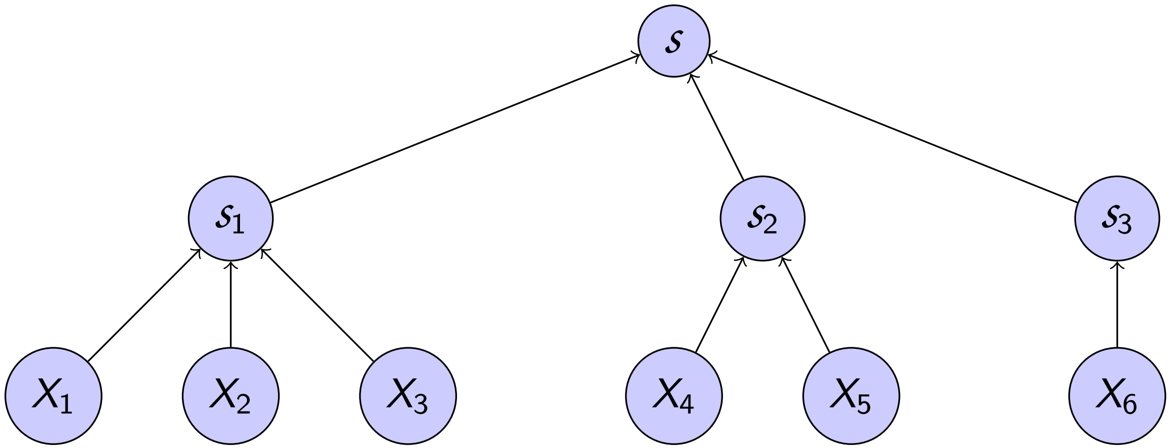


Figure 94: A two-level tree structure

We assume that:

- Level 1
  1. $\omega_{1,1}^{(1)} = 50\%$
     $\omega_{2,1}^{(1)} = 25\%$
     $\omega_{3,1}^{(1)} = 25\%$
  2. $\omega_{4,2}^{(1)} = 50\%$
     $\omega_{5,2}^{(1)} = 50\%$
  3. $\omega_{6,3}^{(1)} = 100\%$
- Level 2: $\omega_1^{(2)} = \omega_2^{(2)} = \omega_3^{(2)} = 33.33\%$

ESG scoring system

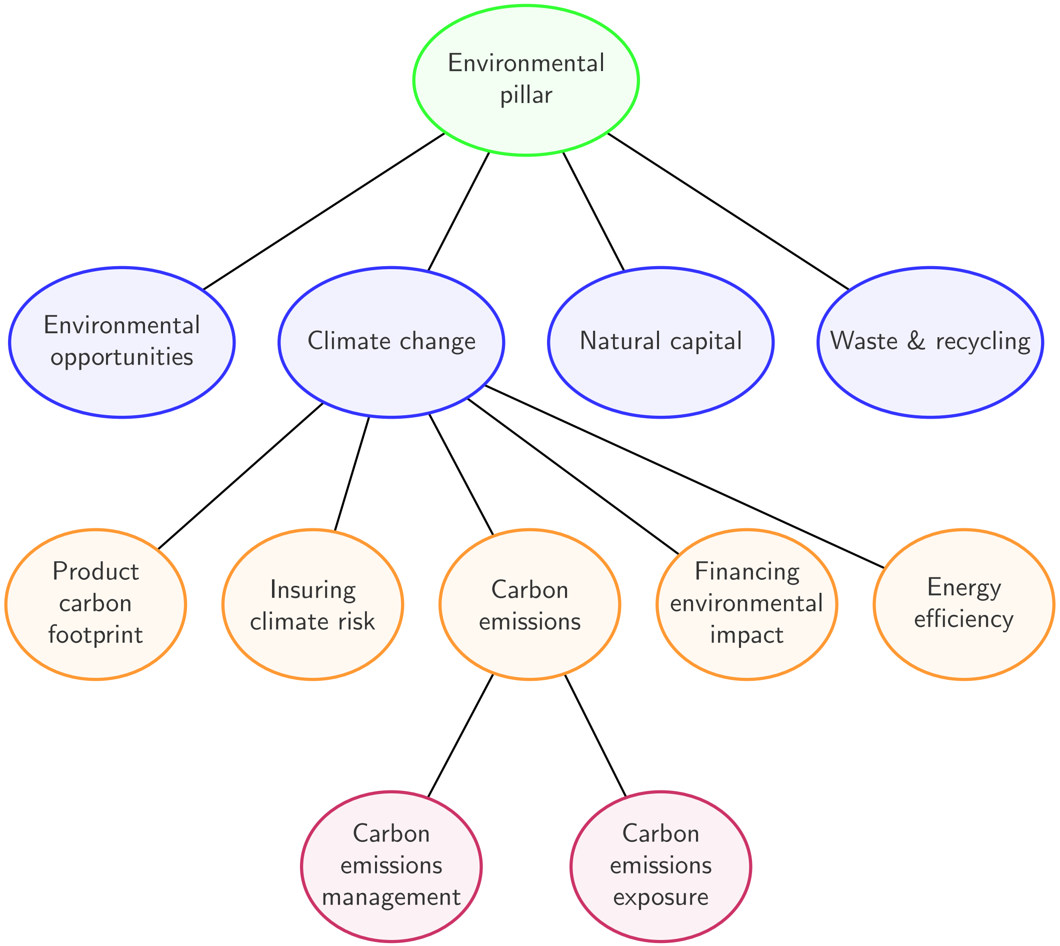


Figure 95: An example of ESG scoring tree (MSCI methodology)

Source: MSCI (2020)

ESG scoring system

Raw data and scores have to be normalized

Why? Because to facilitate the aggregation process

Several normalization approaches:

- $0-1$ normalization: $X_j \in [0,1] \Rightarrow \mathcal{s}_i \in [0,1]$
- $0-100$ normalization: $X_j \in [0,100] \Rightarrow \mathcal{s}_i \in [0,100]$
- z-score normalisation:

$$z_{i,j} = \frac{X_{i,j} - \hat{\mu}(X_j)}{\hat{\sigma}(X_j)}$$

- Empirical normalization using the empirical probability distribution ($0-1$ normalization)

ESG scoring system

Table 70: Computation of z-score

Observation	X_1	z_1	X_2	z_2
1	70.4000	−0.0015	0.0340	0.6911
2	31.3000	−1.0089	0.1000	1.3918
3	66.0000	−0.1149	−0.1660	−1.4321
4	84.2000	0.3540	−0.0590	−0.2962
5	91.7000	0.5472	−0.0280	0.0329
6	53.4000	−0.4395	0.0420	0.7760
7	49.6000	−0.5375	−0.1670	−1.4427
8	133.4000	1.6216	0.0470	0.8291
9	5.1000	−1.6840	−0.1210	−0.9544
10	119.5000	1.2635	0.0070	0.4045
Mean	70.4600	0.0000	−0.0311	0.0000
Standard deviation	38.8127	1.0000	0.0942	1.0000

We have $z_{1,8} = \frac{133.4-70.46}{38.8127} = 1.6216$ and $z_{2,1} = \frac{0.0340-(0.0311)}{0.0942} = 0.6911$

ESG scoring system

Sector neutrality

- Most of ESG scoring systems are sector neutral
- The normalization is done at the sector level, not at the universe level
- ESG scores are then relative (with respect to a sector), not absolute
- Best-in-class/worst-in-class issuers $\neq$ best/worst issuers

ESG rating system

We need a mapping function $\mathcal{M}_{\text{apping}}$ to transform the ESG score $\mathcal{S}$ into an ESG rating $\mathcal{R}$

MSCI methodology

$$\begin{array}{rccl} \mathcal{M}_{\text{apping}}: & [0,10] & \longrightarrow & \{\text{AAA},\text{AA},\text{A},\text{BBB},\text{BB},\text{B},\text{CCC}\} \\ & \mathcal{S} & \longmapsto & \mathcal{R} = \mathcal{M}_{\text{apping}}(\mathcal{S}) \end{array}$$

- If $\mathcal{S} \in \left[0, \frac{10}{7}\right]$, $\mathcal{M}_{\text{apping}}(\mathcal{S}) = \text{CCC}$
- If $\mathcal{S} \in \left[\frac{10}{7}, \frac{2\times 10}{7}\right]$, $\mathcal{M}_{\text{apping}}(\mathcal{S}) = \text{B}$
- If $\mathcal{S} \in \left[\frac{2\times 10}{7}, \frac{3\times 10}{7}\right]$, $\mathcal{M}_{\text{apping}}(\mathcal{S}) = \text{BB}$
- If $\mathcal{S} \in \left[\frac{3\times 10}{7}, \frac{4\times 10}{7}\right]$, $\mathcal{M}_{\text{apping}}(\mathcal{S}) = \text{BBB}$
- If $\mathcal{S} \in \left[\frac{4\times 10}{7}, \frac{5\times 10}{7}\right]$, $\mathcal{M}_{\text{apping}}(\mathcal{S}) = \text{A}$
- If $\mathcal{S} \in \left[\frac{5\times 10}{7}, \frac{6\times 10}{7}\right]$, $\mathcal{M}_{\text{apping}}(\mathcal{S}) = \text{AA}$
- If $\mathcal{S} \in \left[\frac{6\times 10}{7}, 10\right]$, $\mathcal{M}_{\text{apping}}(\mathcal{S}) = \text{AAA}$

ESG rating system

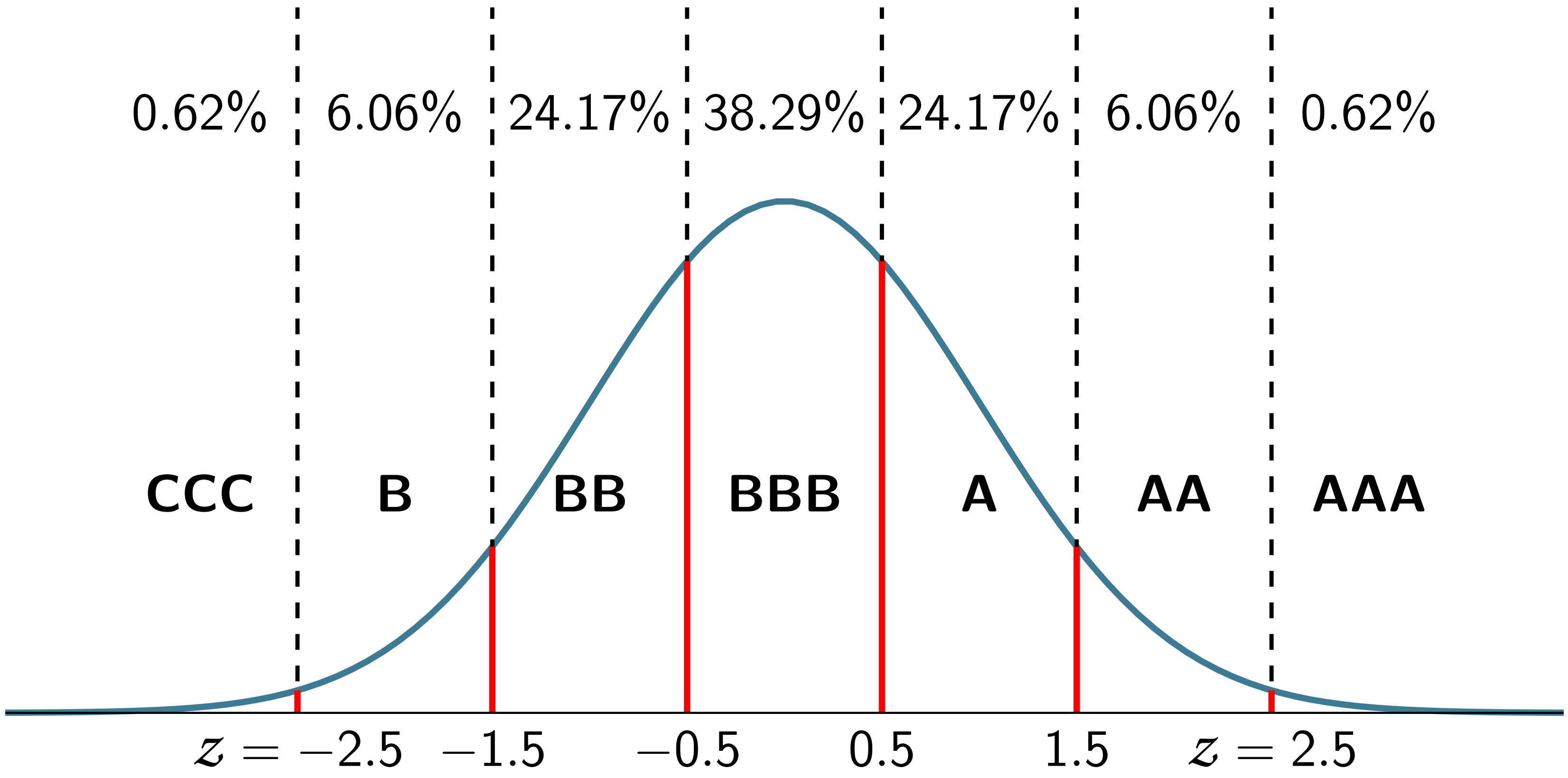


Figure 96: From ESG scores to ESG ratings (Gaussian mapping* of the z-score)

*We have $\Phi(-2.5) = 0.62\%$, $\Phi(-1.5) - \Phi(-2.5) = 6.06\%$, $\Phi(-0.5) - \Phi(-1.5) = 24.17\%$, $\Phi(0.5) - \Phi(-0.5) = 38.29\%$, $\Phi(1.5) - \Phi(0.5) = 24.17\%$, $\Phi(2.5) - \Phi(1.5) = 6.06\%$ and $1 - \Phi(2.5) = 0.62\%$

ESG ratings versus credit ratings

Credit rating

- What is the question?
  Measuring the 1Y PD
- Rating correlation $\geq 90\%$
  Convergence in the 1990s
- **Absolute** rating
  $\Rightarrow$ Facilitates comparison
- More stable
- Accounting standards

ESG rating

- What is the question?
  ???
- Rating correlation $\leq 40\%$
  European issuers $>$ American issuers $>$ Japanese issuers (≈ 0)
- **Relative** rating
  $\Rightarrow$ Complicates comparison
- Less stable
- ESG standardization and the issue of self-reporting

What can we anticipate? $\Rightarrow$ Strong convergence for subcomponents, (more or less) convergence for **E**, **S**, and **G** ratings, but not for **ESG** ratings

The example of Tesla!

What is the performance of ESG investing?

Impact on stock returns

- Stock financial performance $\neq$ corporate financial performance
- Heterogenous results
- Return-oriented or risk-oriented investment style?
- Mixed results

What is the performance of ESG investing?

Academic findings

- Relationship between shareholder rights and "*higher firm value, higher profits, higher sales growth, lower capital expenditures, and [...] fewer corporate acquisitions*" (Gompers *et al.*, 2003)
- Positive relation between high corporate social responsibility and low cost of equity capital (El Ghoul *et al.*, 2011): "*Employee Relations, Environmental Policies, Product Strategies lower the firms' cost of equity*"
- Corporate financial performance is a U-shape function of corporate social performance (Barnett and Salomon, 2012)
- Cultural differences explain the diversity and differences in intentions ('Value' or 'Values' oriented) of the currently available ESG data (Eccles and Stroehle, 2018)
- Negative/neutral impact: Schröder (2007), Hong and Kacperczyk (2009)

Mixed results

What is the performance of ESG investing?

We consider the two studies conducted by Amundi Quantitative Research:

- 2010-2017
  Bennani, L., Le Guenedal, T., Lepetit, F., Ly, L., Mortier, V., Roncalli, T., and Sekine T. (2018), How ESG Investing Has Impacted the Asset Pricing in the Equity Market, Amundi Discussion Paper, DP-39-2018, `https://research-center.amundi.com`
- 2018-2019
  Drei, A., Le Guenedal, T., Lepetit, F., Mortier, V., Roncalli, T., and Sekine T. (2020), ESG Investing in Recent Years: New Insights from Old Challenges, Amundi Discussion Paper, DP-42-2019, `https://research-center.amundi.com`

2010 – 2017: From hell to heaven

- ESG investing tended to penalize both passive and active ESG investors between 2010 and 2013
- Contrastingly, ESG investing was a source of outperformance from 2014 to 2017 in Europe and North America
- Two success stories between 2014 and 2017: **E**nvironmental in North America and **G**overnance in the Eurozone
- ESG was a risk factor (or a beta strategy) in the Eurozone, whereas it was an alpha strategy in North America

Active management

Sorted portfolio methodology

Sorted-portfolio approach

- Sorted-based approach of Fama-French (1992)
- At each rebalancing date t, we rank the stocks according to their Amundi **ESG** z-score $s_{i,t}$
- We form the five quintile portfolios Q_i for $i = 1, \ldots, 5$
- The portfolio Q_i is invested during the period $]t, t+1]$:
  - Q_1 corresponds to the best-in-class portfolio (best scores)
  - Q_5 corresponds to the worst-in-class portfolio (worst scores)
- Quarterly rebalancing
- Universe: MSCI World Index
- Equally-weighted and sector-neutral portfolio (and region-neutral for the world universe)

Performance of ESG active management (2010 – 2017)

North America

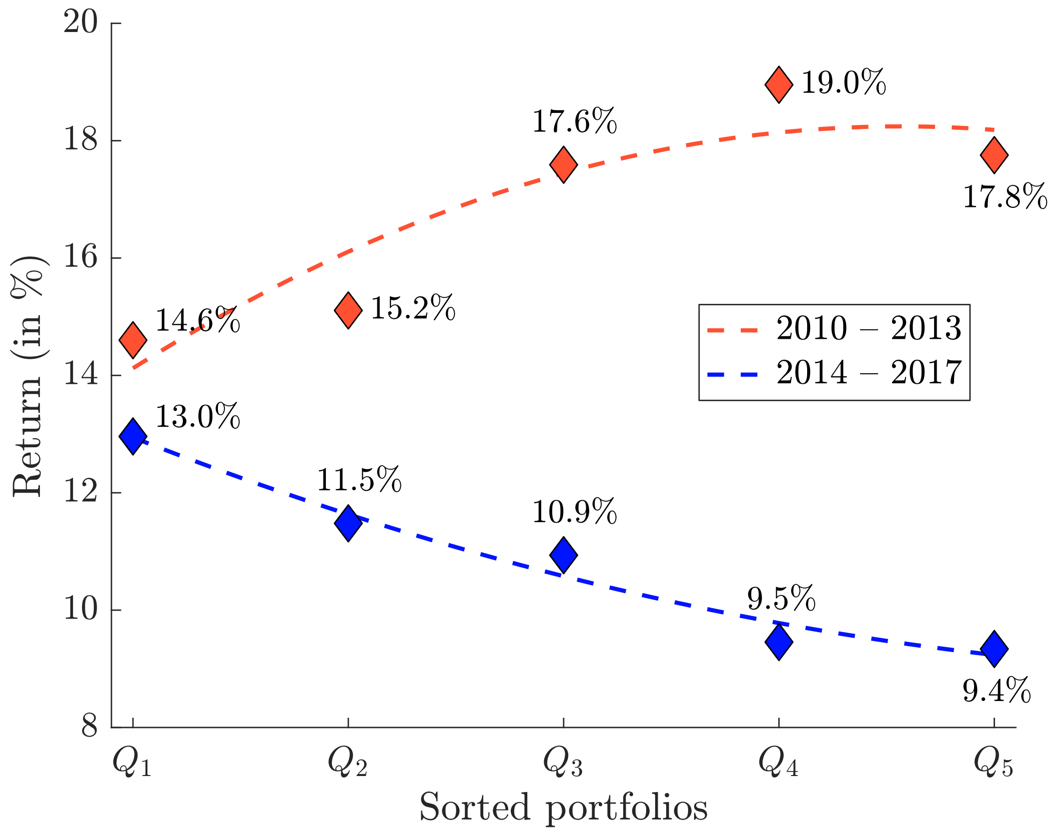


Figure 97: Annualized return of **ESG** sorted portfolios (North America)

Source: Amundi Quantitative Research (2018)

Performance of ESG active management (2010 – 2017)

Eurozone

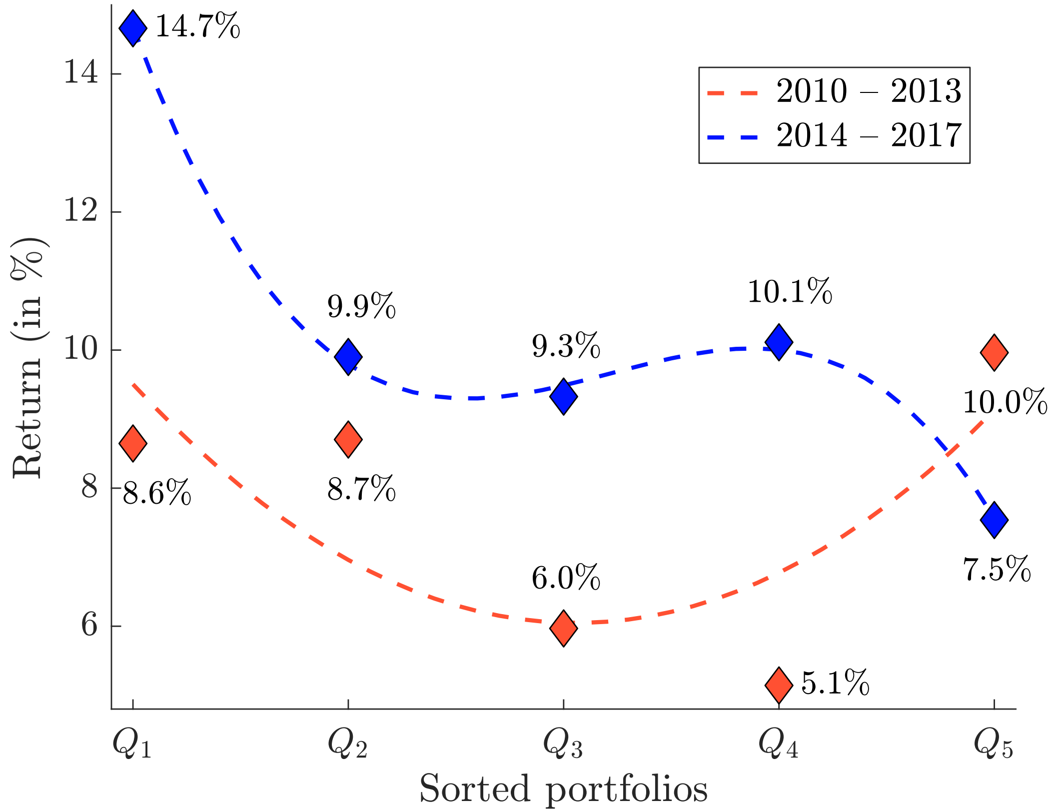


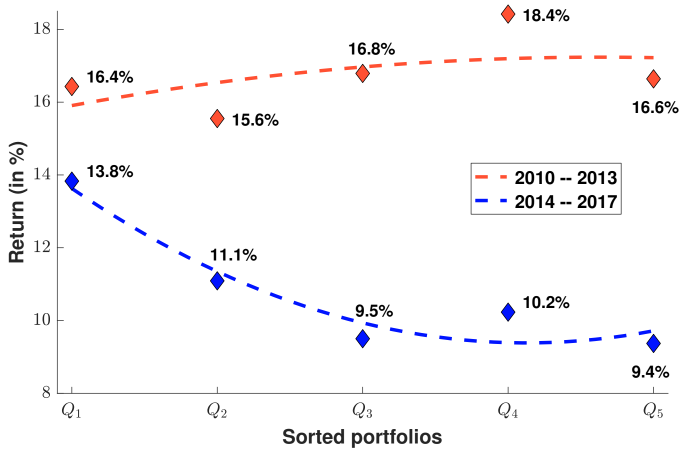
Figure 98: Annualized return of **ESG** sorted portfolios (Eurozone)

Source: Amundi Quantitative Research (2018)

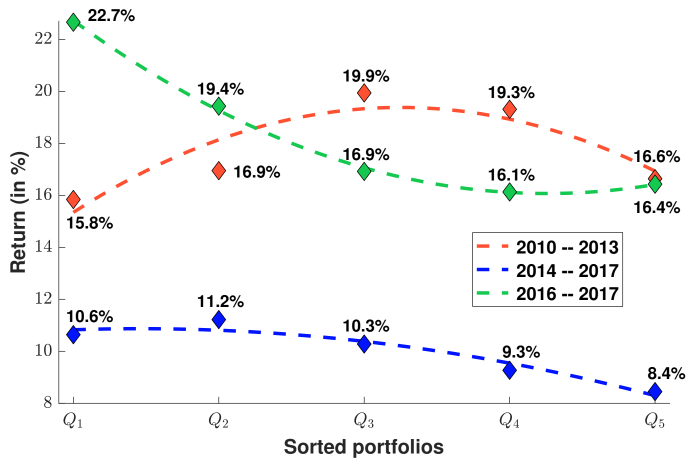
Performance of ESG active management (2010 – 2017)

North America

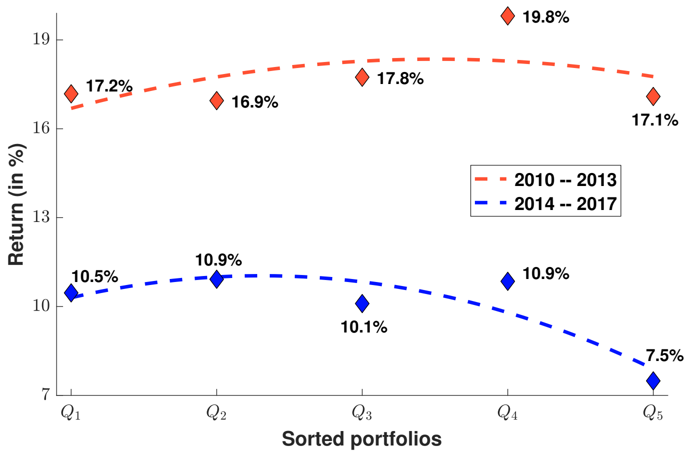
Environmental



Social



Governance

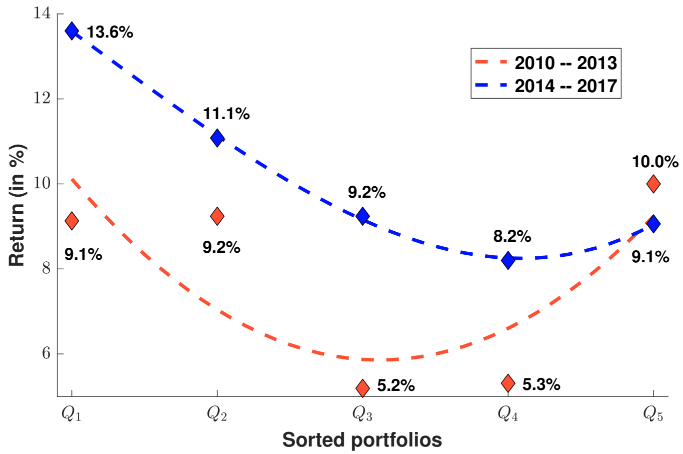


Source: Amundi Quantitative Research (2018)

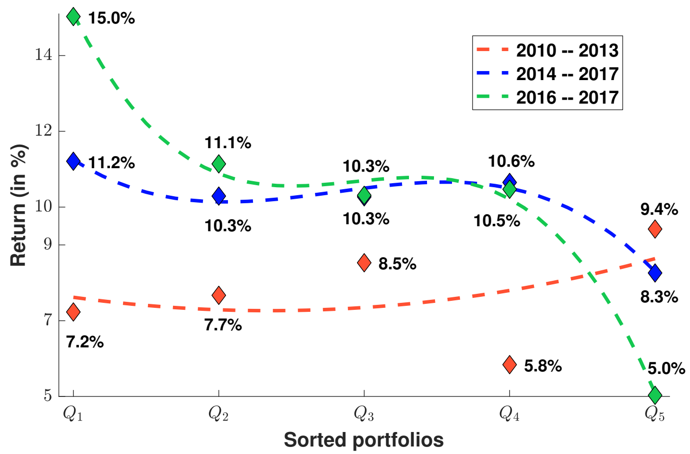
Performance of ESG active management (2010 – 2017)

Eurozone

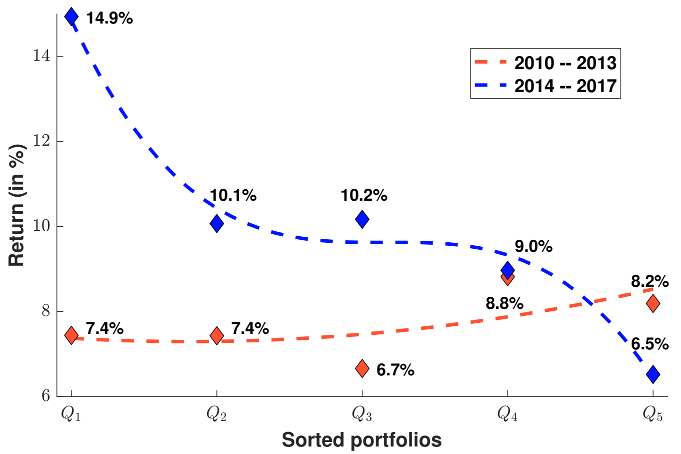
Environmental



Social



Governance



Source: Amundi Quantitative Research (2018)

Performance of ESG active management (2010 – 2017)

The 2014 break

Table 71: Summary of the results

Before 2014					
Factor	North America	Eurozone	Europe ex-EMU	Japan	World DM
ESG	−−	−	0	+	0
E	−	0	+	−	0
S	−	−	0	−	−
G	−	0	+	0	+
Since 2014					
Factor	North America	Eurozone	Europe ex-EMU	Japan	World DM
ESG	++	++	0	−	+
E	++	++	−	+	++
S	+	+	0	0	+
G	+	++	0	+	++

Source: Amundi Quantitative Research (2018)

The 2014 break

How to explain the 2014 break?

1. **The intrinsic value of ESG screening** or **the materiality of ESG**

   "Since we observe a feedback loop between extra-financial risks and asset pricing, we may also wonder whether the term 'extra' is relevant, because ultimately, we can anticipate that these risks may no longer be extra-financial, but simply financial" (Bennani et al., 2018).

   ESG risks $\Rightarrow$ Asset pricing

2. **The extrinsic value of ESG investing** or **the supply/demand imbalance**

   Investment flows matter!

The steamroller of ESG for institutional investors

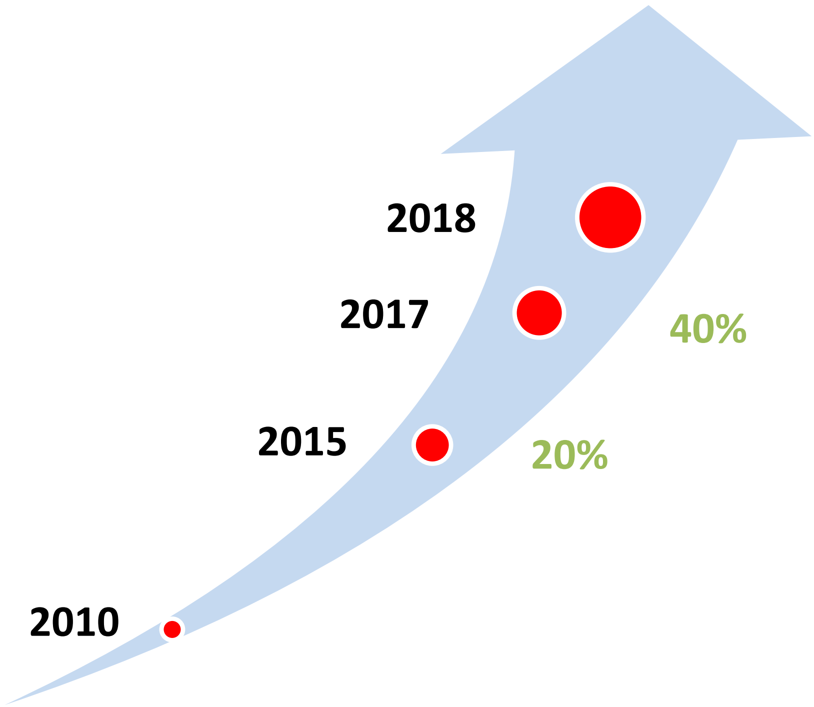


Figure 99: Frequency of institutional RFPs that require ESG filters

Source: Based on RFPs received at Amundi.

- In some countries, 100% of RFPs require ESG filters
- For some institutional investors, 100% of RFPs require ESG filters (public, para-public and insurance investors)
- For some strategies, 100% of RFPs require ESG filters (index tracking)

Passive management (optimized portfolios)

Portfolio optimization with a benchmark

We consider the following optimization problem[17]:

$$x^{\star}(\gamma) = \arg\min \frac{1}{2}\sigma^2(x \mid b) - \gamma \mathcal{S}(x \mid b)$$

where $\sigma(x \mid b)$ is the ex-ante tracking error (TE) of Portfolio x with respect to the benchmark b:

$$\sigma(x \mid b) = \sqrt{(x - b)^\top \Sigma (x - b)}$$

and $\mathcal{S}(x \mid b)$ is the excess score (ES) of Portfolio x wrt the benchmark b:

$$\begin{aligned}\mathcal{S}(x \mid b) &= (x - b)^\top \mathcal{S} \\ &= \mathcal{S}(x) - \mathcal{S}(b)\end{aligned}$$

[17] We note b the benchmark, $\mathcal{S}$ the vector of scores and Σ the covariance matrix.

Passive management (optimized portfolios)

Portfolio optimization with a benchmark

The objective is to find the optimal portfolio with the minimum TE for a given ESG excess score

This is a standard γ-problem where the expected returns are replaced by the ESG scores (see Lecture 1)

Performance of ESG passive management (2010-2017)

Arbitrage between ESG and TE

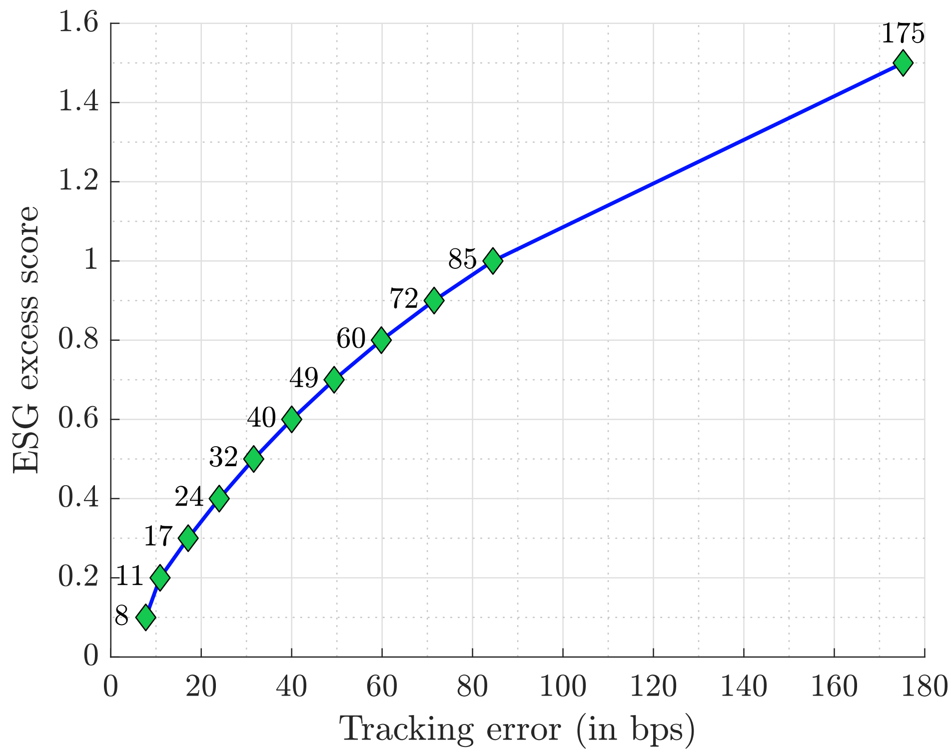


Figure 100: Efficient frontier of **ESG** optimized portfolios (World DM)

Source: Amundi Quantitative Research (2018)

No free lunch: **ESG investing implies to take a tracking-error risk!**

Performance of ESG passive management (2010-2017)

Performance of optimized portfolios

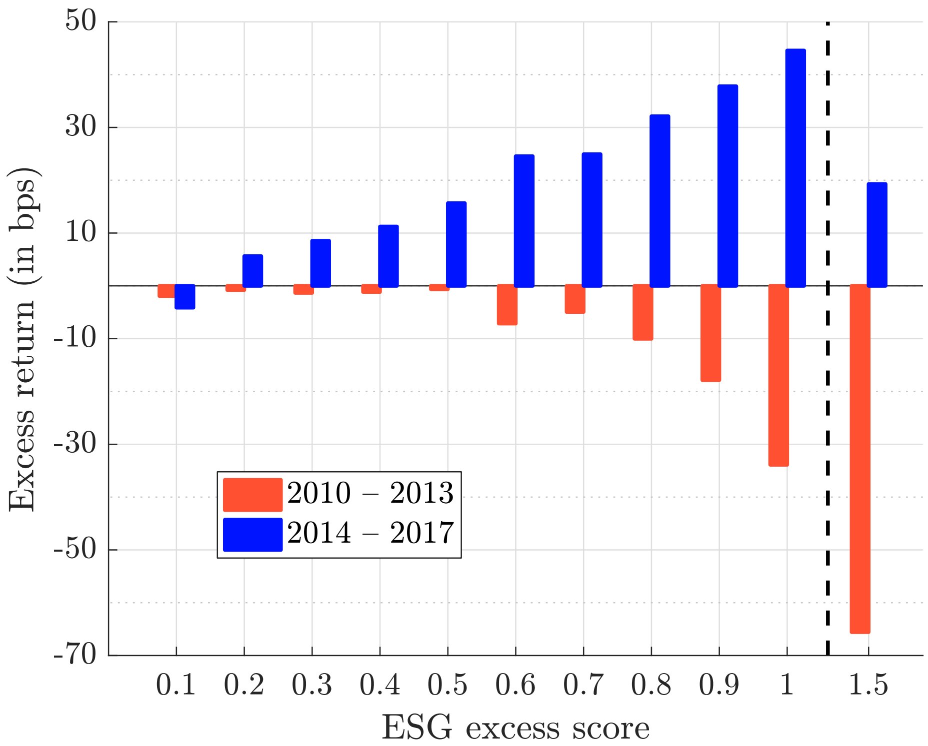


Figure 101: Annualized excess return of **ESG** optimized portfolios (World DM)

Source: Amundi Quantitative Research (2018)

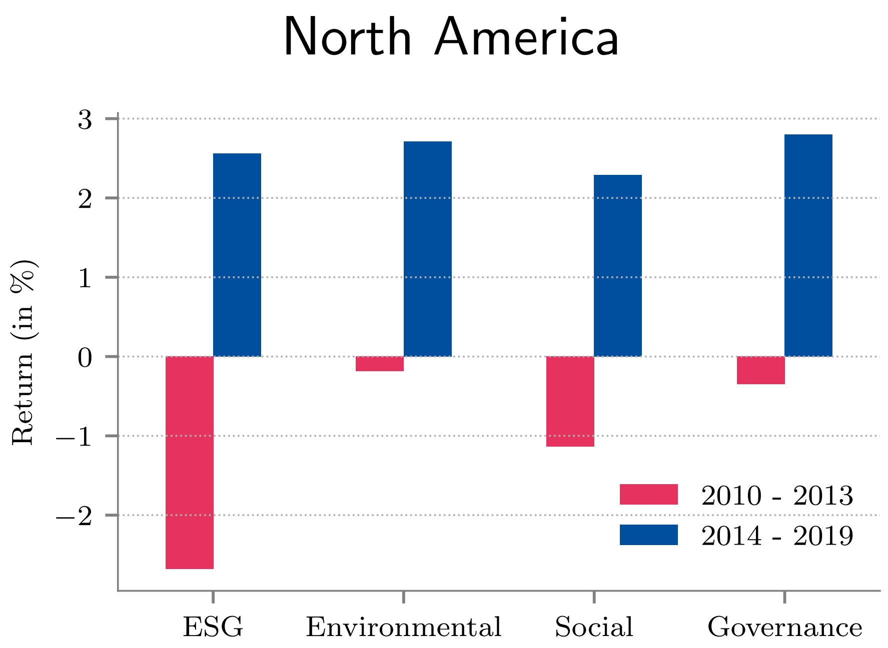
ESG investing & diversification: **Mind the gap**

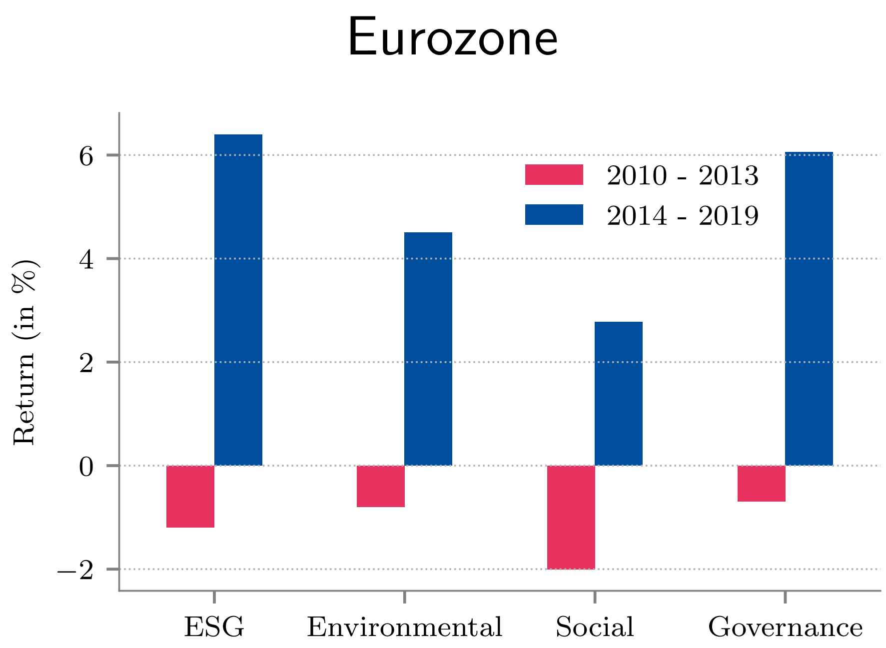
Performance of ESG active management (2018-2019)

On the road again

Main result

The 2018 – 2019 period seems to be a continuity of the 2014 – 2017 period rather than another distinctive phase





Source: Amundi Quantitative Research (2020)

Performance of ESG active management (2018-2019)

New findings in the stock market

1. **The transatlantic divide**

$$\text{Eurozone} \succ \text{North America}$$

2. **Social: from laggard to leader**[18]

$$\textbf{S} \succ \textbf{E}, \textbf{G}$$

3. **ESG investing: growing in complexity**

   Beyond worst-in-class exclusion and best-in-class selection strategies

[18] In the Eurozone: 2010 – 2013: **E**, then 2014 – 2017: **G**, then 2018 – 2019: **S**
In North America: 2010 – 2013: **G**, then 2014 – 2017: **E**, then 2018 – 2019: **S**

Performance of ESG active management (2018-2019)

The transatlantic divide: the case of the Eurozone

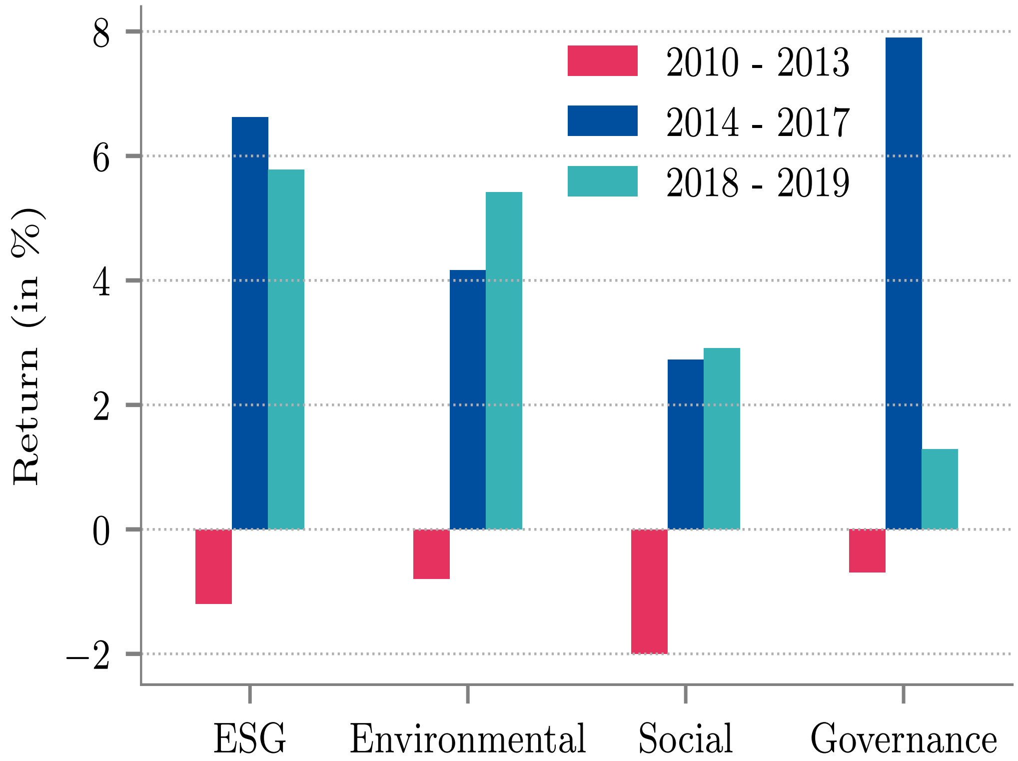


Figure 102: Annualized return of long/short $Q_1 - Q_5$ sorted portfolios

Source: Amundi Quantitative Research (2020)

$\Rightarrow$ Performance remains highly positive, and is improved for **E** and **S**

Performance of ESG active management (2018-2019)

The transatlantic divide: the case of North America

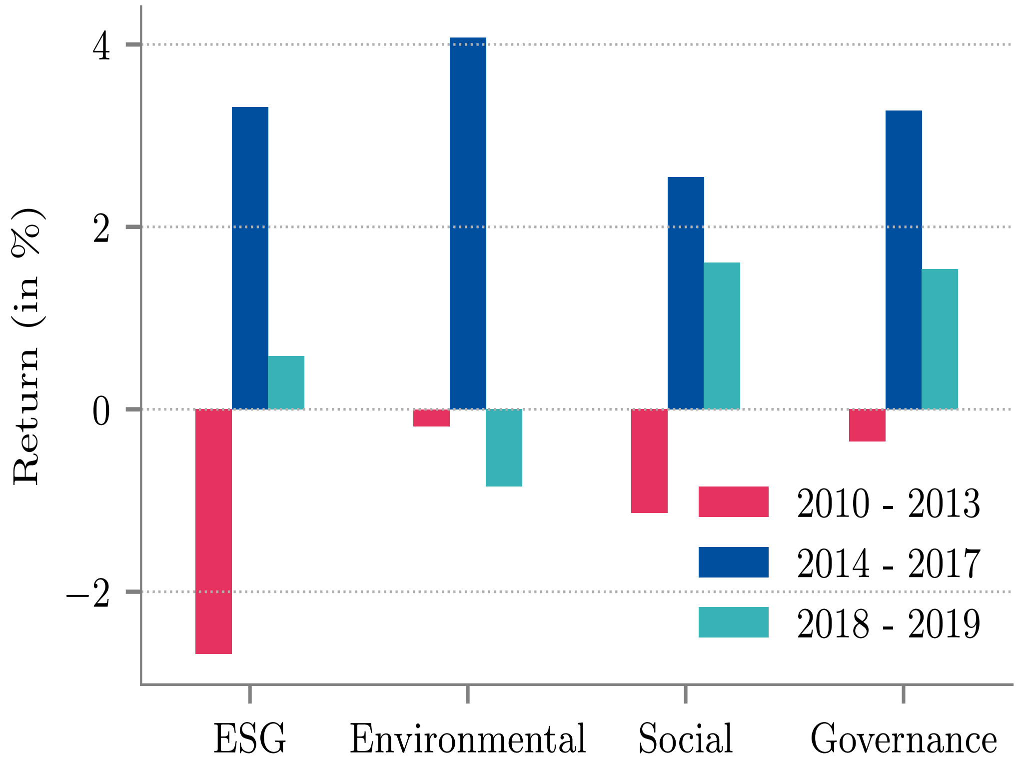


Figure 103: Annualized return of long/short $Q_1 - Q_5$ sorted portfolios

Source: Amundi Quantitative Research (2020)

$\Rightarrow$ Performance is positive, but reduced for **S** and **G**, whereas **E** is negative

Performance of ESG active management (2018-2019)

How to explain the American setback?

The regulatory value of ESG investing (or the intrinsic value *revisited*)

- Trump election effect
- Regulatory environment

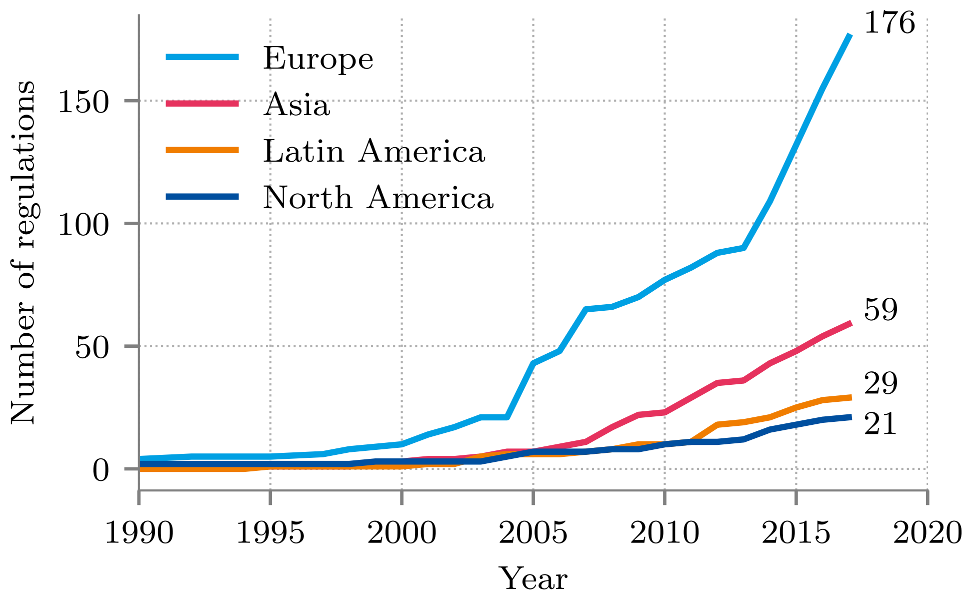


Figure 104: Number of ESG regulations

- ESG regulations are increasing, with a strong momentum in Europe but a weaker one in North America
- US withdrawal from Paris Climate Agreement

Source: PRI, responsible investment regulation database, 2019.

Performance of ESG active management (2018-2019)

How to explain the American setback?

The extrinsic value of ESG investing

- The 2014 break
  - November 2013: Responsible Investment and the Norwegian Government Pension Fund Global (2013 Strategy Council)
  - Strong mobilization of the largest institutional European investors: NBIM, APG, PGGM, ERAFP, FRR, etc.
  - They are massively invested in European stocks and America stocks:

$$\text{NBIM} \succ \text{CalPERS} + \text{CalSTRS} + \text{NYSCRF for U.S. stocks}$$

- The 2018-2019 period
  - Implication of U.S. investors continues to be weak
  - Strong mobilization of medium (or tier two) institutional European investors, that have a low exposure on American stocks
  - Mobilization of European investors is not sufficient

$\Rightarrow$ The extrinsic value of ESG investing is temporary, and a new equilibrium will be found on the long run

Performance of ESG active management (2018-2019)

Social is strong in Eurozone

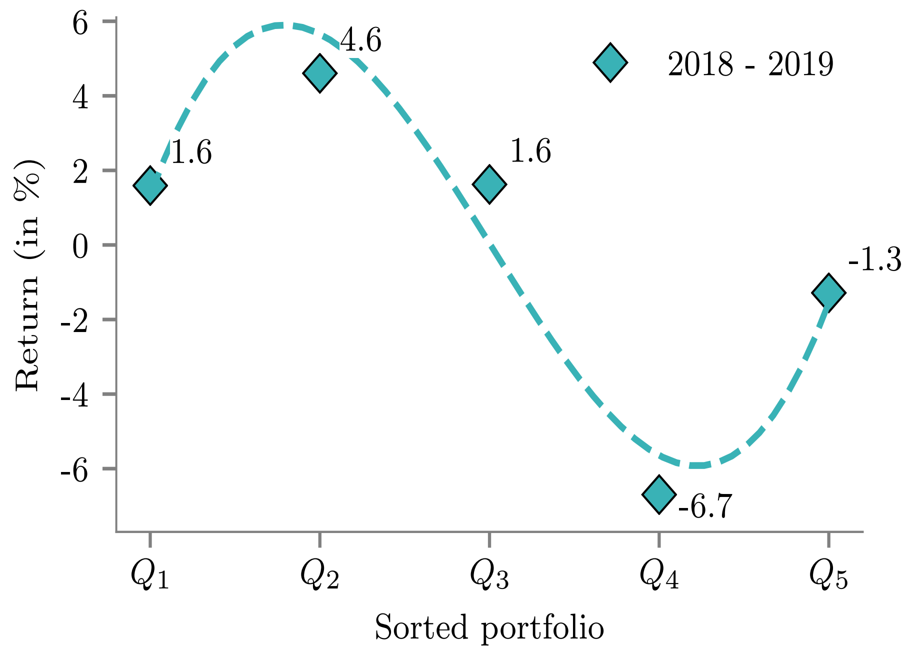


Figure 105: Sorted portfolios

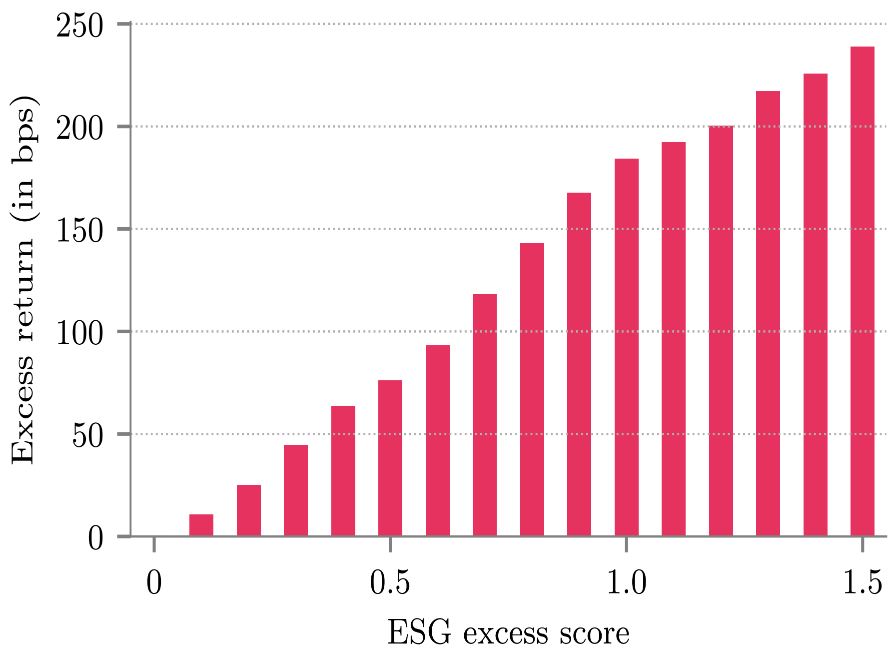


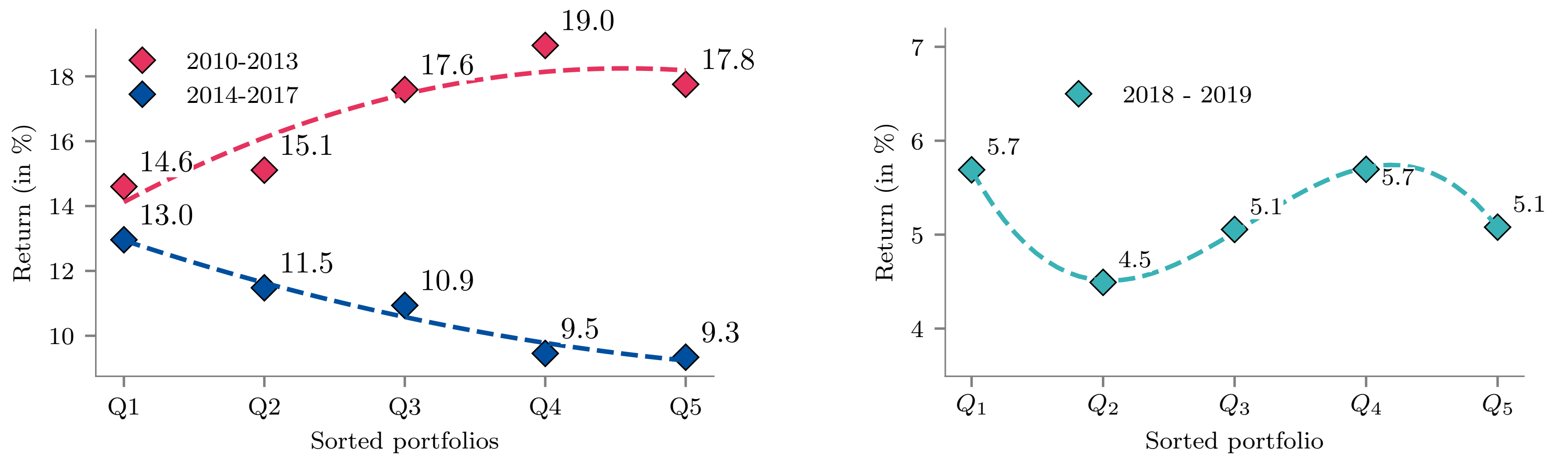
Figure 106: Optimized portfolios

Source: Amundi Quantitative Research (2020)

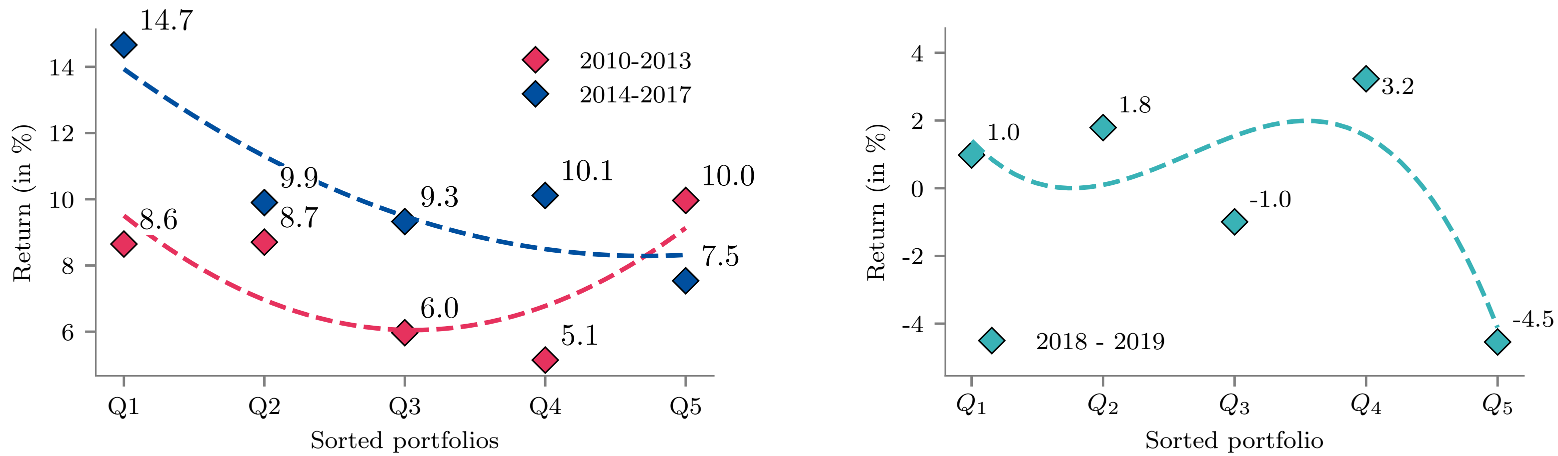
Performance of ESG active management (2018-2019)

ESG investing: growing in complexity

North America, **ESG**-Sorted portfolios, 2010 – 2019



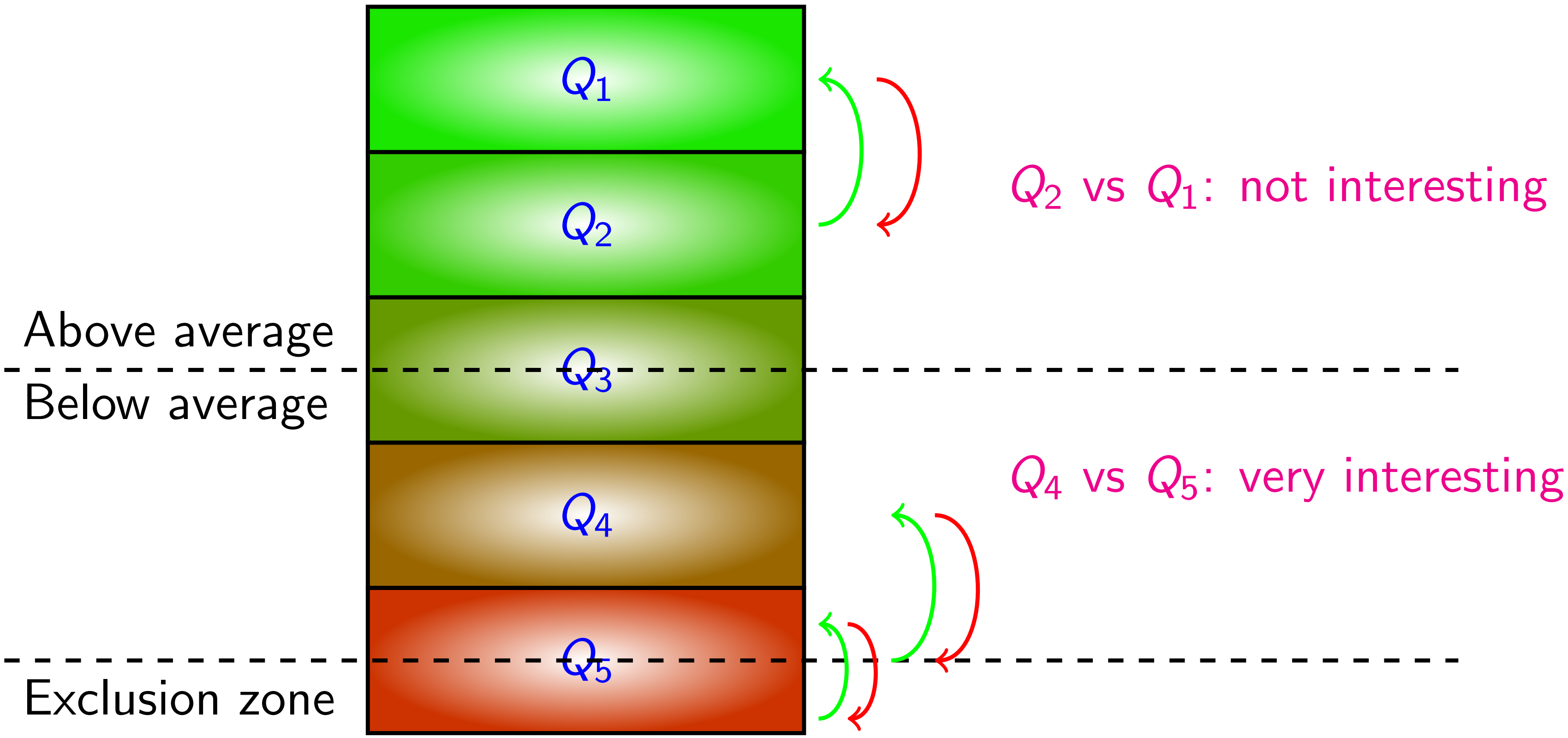
Eurozone, **ESG**-Sorted portfolios, 2010 – 2019



Performance of ESG active management (2018-2019)

The dynamic view of ESG investing

Figure 107: How to play ESG momentum?



ESG and factor investing

Single-factor model

Regression model

We have:

$$R_{i,t} = \alpha_i + \beta_i^j \mathcal{F}_{j,t} + \varepsilon_{i,t}$$

where $\mathcal{F}_{j,t}$ can be: market, size, value, momentum, low-volatility, quality or ESG.

ESG and factor investing

Single-factor model

Table 72: Results of cross-section regressions with long-only risk factors (average R^2)

Factor	North America		Eurozone	
	2010 – 2013	2014 – 2019	2010 – 2013	2014 – 2019
Market	40.8%	28.6%	42.8%	36.3%
Size	39.3%	26.1%	37.1%	23.3%
Value	38.9%	26.7%	41.6%	33.6%
Momentum	39.6%	26.3%	40.8%	34.1%
Low-volatility	35.8%	25.1%	38.7%	33.4%
Quality	39.1%	26.6%	42.4%	34.6%
ESG	40.1%	27.4%	42.6%	35.3%

Source: Amundi Quantitative Research (2020)

- Specific risk has increased during the period 2014 – 2019
- Since 2014, we find that:
  - ESG $\succ$ Value $\succ$ Quality $\succ$ Momentum $\succ \ldots$ (North America)
  - ESG $\succ$ Quality $\succ$ Momentum $\succ$ Value $\succ \ldots$ (Eurozone)

ESG and factor investing

Multi-factor model

Regression model

We have:

$$R_{i,t} = \alpha_i + \sum_{j}^{n_{\mathcal{F}}} \beta_i^j \mathcal{F}_{j,t} + \varepsilon_{i,t}$$

- 1F = market
- 5F = size + value + momentum + low-volatility + quality
- 6F = 5F + ESG

ESG and factor investing

Multi-factor model

Table 73: Results of cross-section regressions with long-only risk factors (average R^2)

Factor	North America		Eurozone	
	2010 – 2013	2014 – 2019	2010 – 2013	2014 – 2019
Market	40.8%	28.6%	42.8%	36.3%
5F model	46.1%	38.4%	49.5%	45.0%
6F model (5F + ESG)	46.7%	39.7%	50.1%	45.8%

Source: Amundi Quantitative Research (2020)

***p-value statistic for the MSCI Index (time-series, 2014 – 2019):

- 6F = ~~Size~~, Value, Momentum, Low-volatility, Quality, ~~ESG~~ (North America)
- 6F = Size, Value, Momentum, ~~Low-volatility~~, Quality, ESG (Eurozone)

ESG and factor investing

Factor selection

Least absolute shrinkage and selection operator (lasso)

The lasso regression is defined by:

$$\frac{y_i - \bar{y}}{\sigma_y} = \sum_{k=1}^{K} \beta_k \left(\frac{x_{i,k} - \bar{x}_k}{\sigma_{x_k}} \right) + \varepsilon_i$$

$$\text{s.t.} \quad \sum_{k=1}^{K} |\beta_k| \leq \tau$$

We note $\hat{\beta}^{\text{lasso}}(\tau)$ the lasso estimator. We have $\hat{\beta}^{\text{lasso}}(\infty) = \hat{\beta}^{\text{ols}}$ and $\hat{\beta}^{\text{lasso}}(0) = \mathbf{0}_K$.

ESG and factor investing

Factor selection

In the two-asset case, we have:

$$\mathrm{RSS}\left(\beta_1, \beta_2\right) = \sum_{i=1}^{n} \left(\tilde{y}_i - \beta_1 \tilde{x}_{i,1} - \beta_2 \tilde{x}_{2,1}\right)^2$$

If we consider the equation $\mathrm{RSS}\left(\beta_1, \beta_2\right) = c$, we obtain the following cases:

$c < \mathrm{RSS}\left(\hat{\beta}_1^{\mathrm{ols}}, \hat{\beta}_2^{\mathrm{ols}}\right)$	$c = \mathrm{RSS}\left(\hat{\beta}_1^{\mathrm{ols}}, \hat{\beta}_2^{\mathrm{ols}}\right)$	$c > \mathrm{RSS}\left(\hat{\beta}_1^{\mathrm{ols}}, \hat{\beta}_2^{\mathrm{ols}}\right)$
No solution	One solution $\left(\hat{\beta}_1^{\mathrm{ols}}, \hat{\beta}_2^{\mathrm{ols}}\right)$	An ellipsoid

What does this result become when imposing the lasso constraint $\left|\beta_1\right| + \left|\beta_2\right| \leq \tau$?

Sparsity property

$$\exists\, \eta > 0 : \forall \tau < \eta,\ \min\left(\left|\hat{\beta}_1^{\mathrm{lasso}}\right|, \left|\hat{\beta}_2^{\mathrm{lasso}}\right|\right) = 0$$

ESG and factor investing

Factor selection

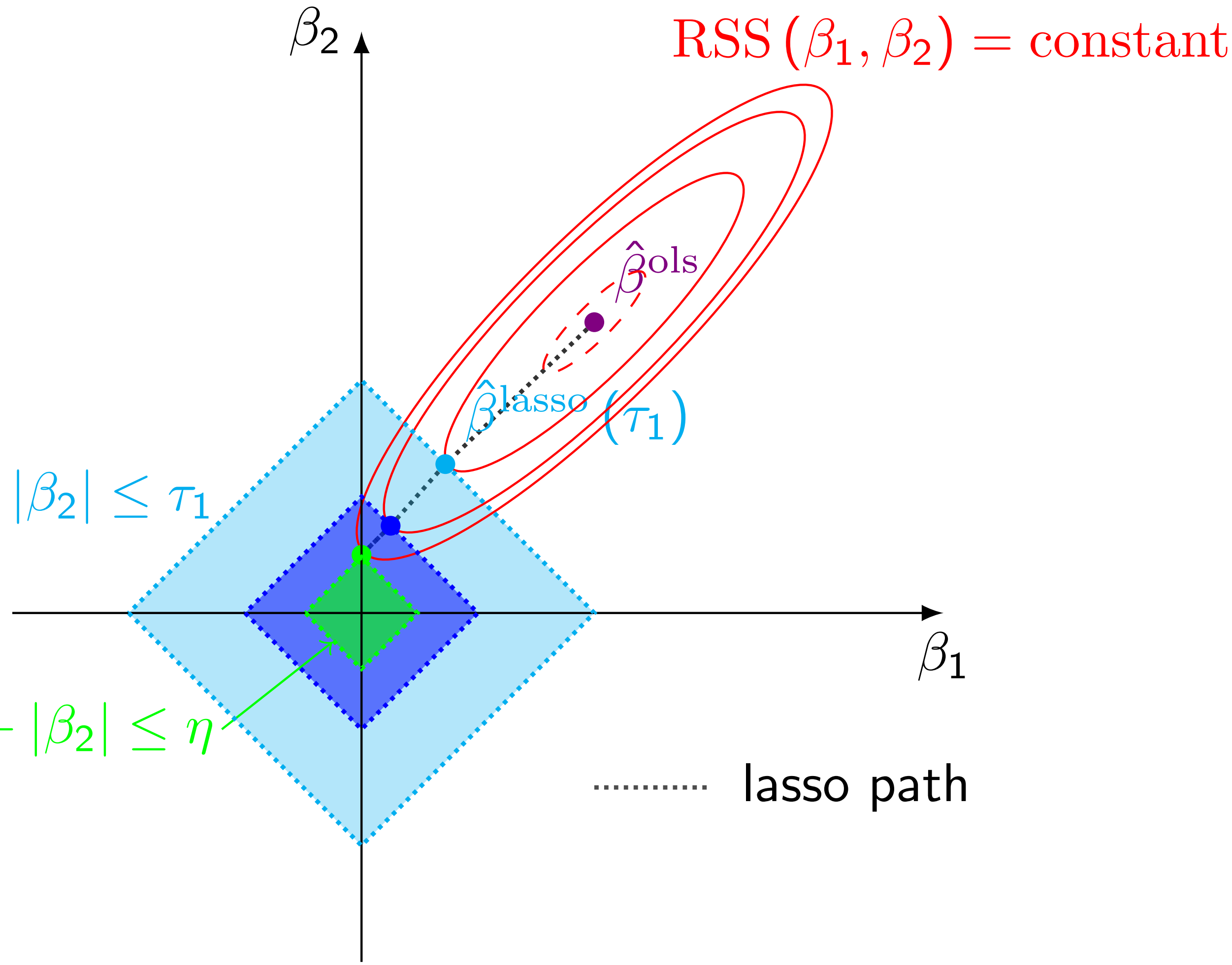


Figure 108: Interpretation of the lasso regression

Factor selection

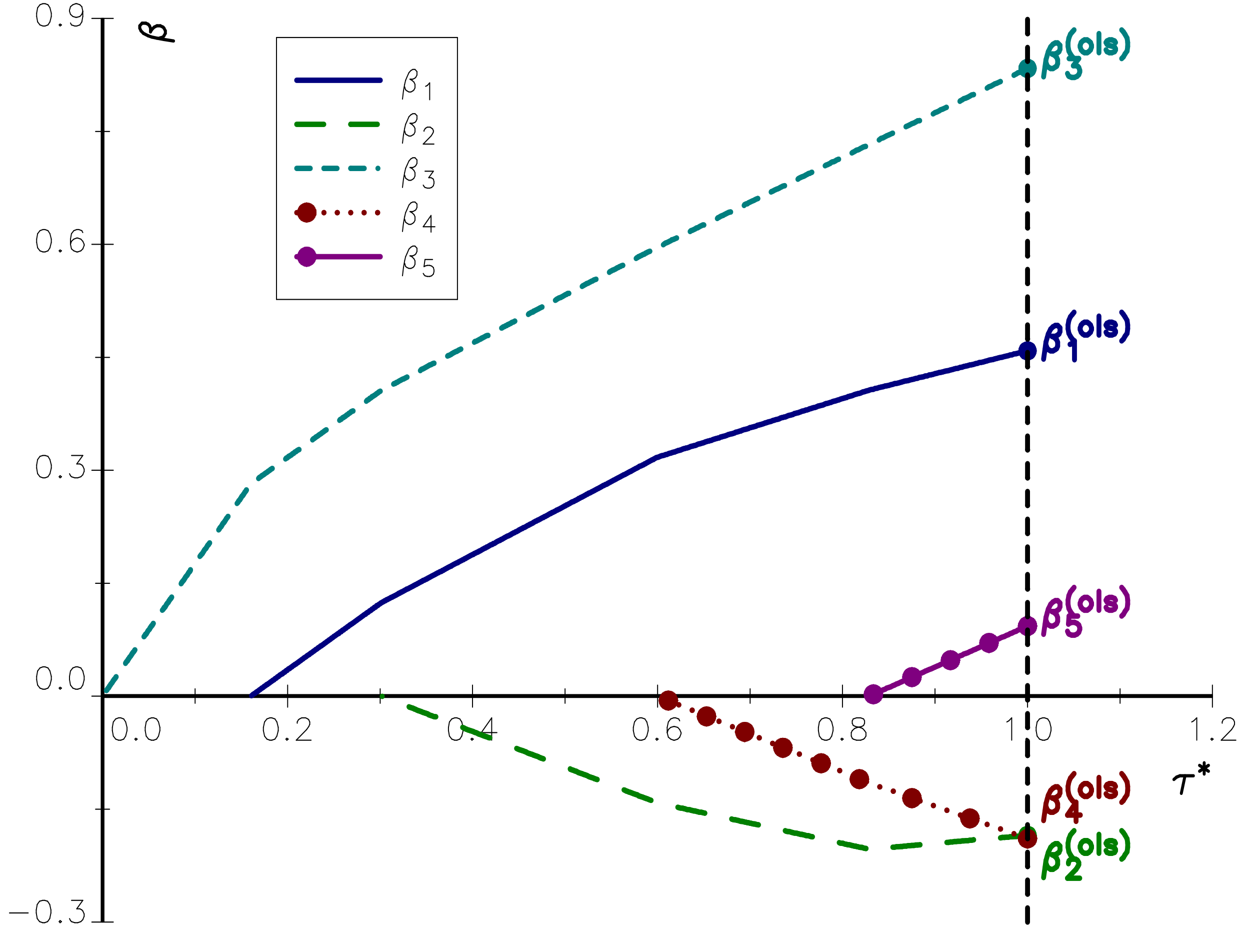


Figure 109: Variable selection with the lasso method (variable ordering: $x_3 \succ x_1 \succ x_2 \succ x_4 \succ x_5$)

ESG and factor investing

ESG as an alpha strategy

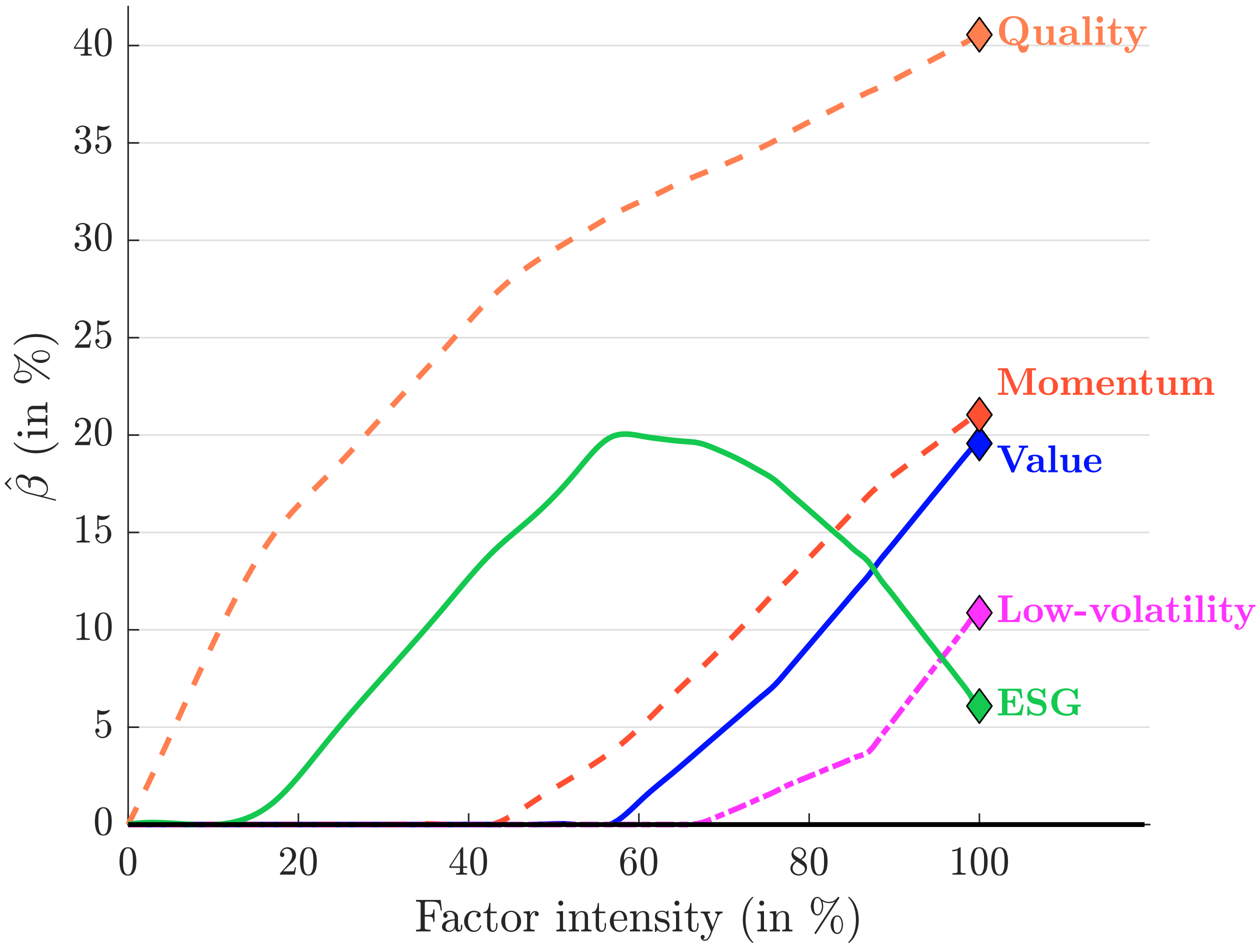


Figure 110: Factor selection (North America)

Source: Amundi Quantitative Research (2020)

ESG and factor investing

ESG as a beta strategy

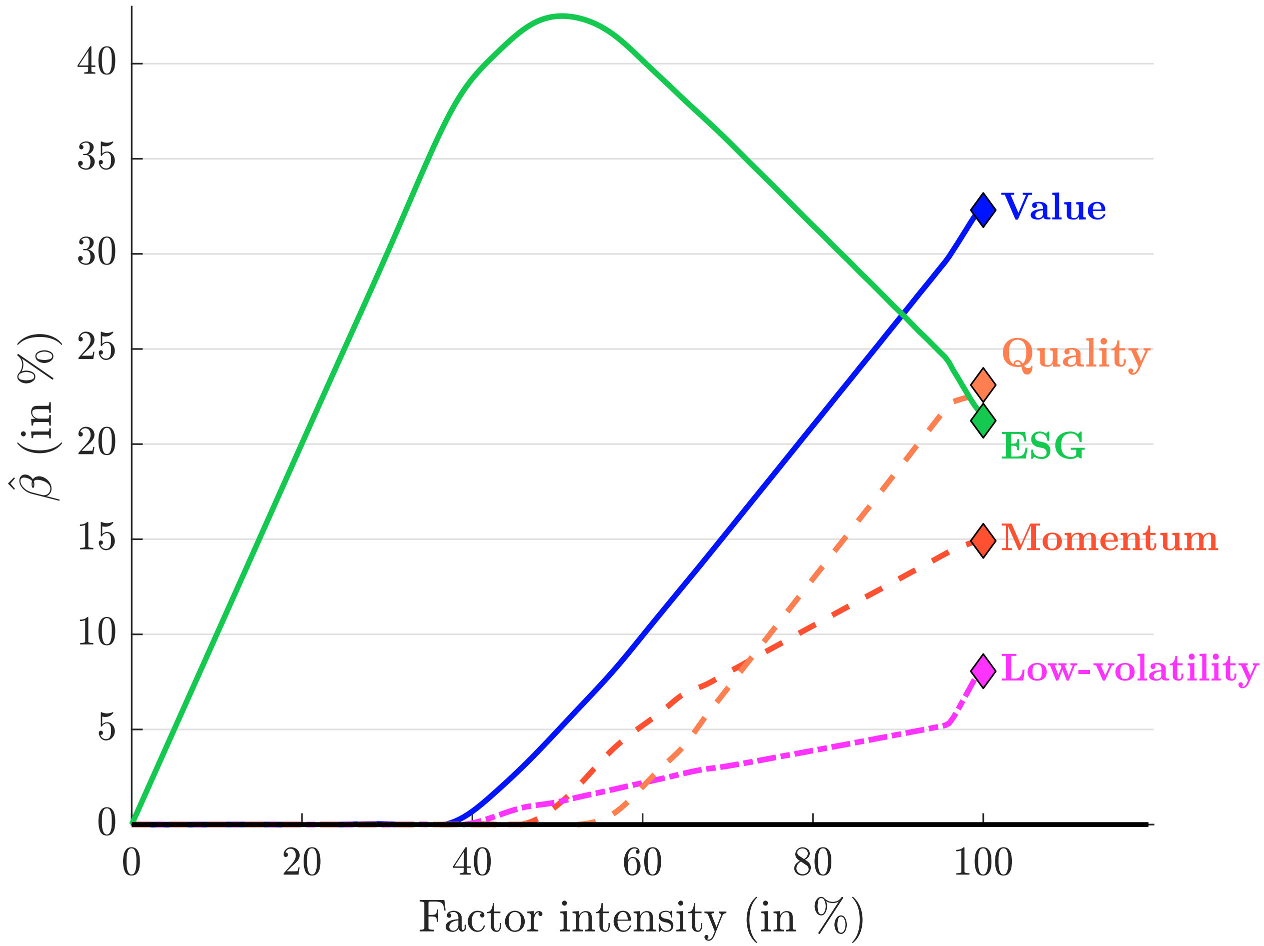


Figure 111: Factor selection (Eurozone)

Source: Amundi Quantitative Research (2020)

ESG and factor investing

What is the difference between alpha and beta?

α or β?

"[...] When an alpha strategy is massively invested, it has an enough impact on the structure of asset prices to become a risk factor.
[...] Indeed, an alpha strategy becomes a common market risk factor once it represents a significant part of investment portfolios and explains the cross-section dispersion of asset returns" (Roncalli, 2020)

- ESG remains an alpha strategy in North America
- ESG becomes a beta strategy (or a risk factor) in Europe

Why ESG investing in bond markets is different than ESG investing in stock markets

Stocks

- ESG scoring is incorporated in portfolio management
- ESG = long-term business risk $\Rightarrow$ strongly impacts the equity
- Portfolio integration
- Managing the business risk

Bonds

- ESG integration is generally limited to exclusions
- ESG lowly impacts the debt
- Portfolio completion
- Fixed income = impact investing
- Development of pure play ESG securities (green and social bonds)

$\Rightarrow$ Stock holders are more ESG sensitive than bond holders because of the capital structure

Why ESG investing in bond markets is different than ESG investing in stock markets

ESG investment flows affect asset pricing differently

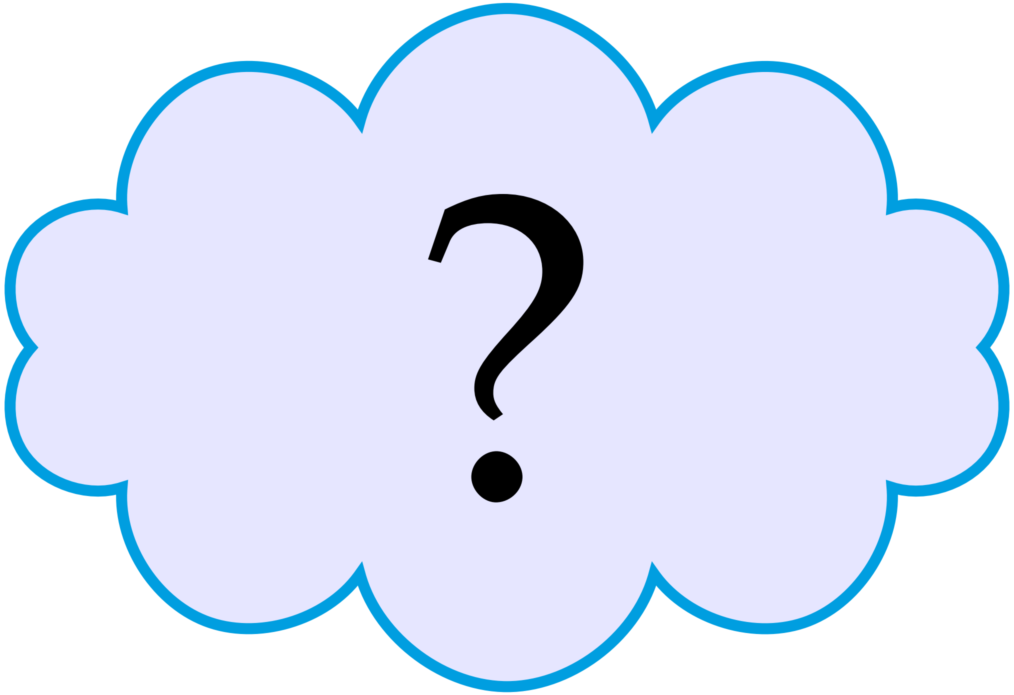
- Impact on carry (coupon effect)?
- Impact on price dynamics (credit spread/mark-to-market effect)?
- Buy-and-hold portfolios $\neq$ managed portfolios

The distinction between IG and HY bonds

- ESG and credit ratings are correlated
- There are more worst-in-class issuers in the HY universe, and best-in-class issuers in the IG universe
- Non-neutrality of the bond universe (bonds $\neq$ stocks)

What is the performance of ESG investing?

Academic findings



What is the performance of ESG investing?

We consider the two studies conducted by Amundi Quantitative Research:

- Ben Slimane, M., Le Guenedal, T., Roncalli, T., and Sekine T. (2020), ESG Investing in Corporate Bonds: Mind the Gap, Amundi Working Paper, WP-94-2019, `https://research-center.amundi.com`
- Ben Slimane, M., Brard, E., Le Guenedal, T., Roncalli, T., and Sekine T. (2020), ESG Investing in Fixed Income: It's Time To Cross the Rubicon, Amundi Discussion Paper, DP-45-2019, `https://research-center.amundi.com`

Sorted portfolio methodology

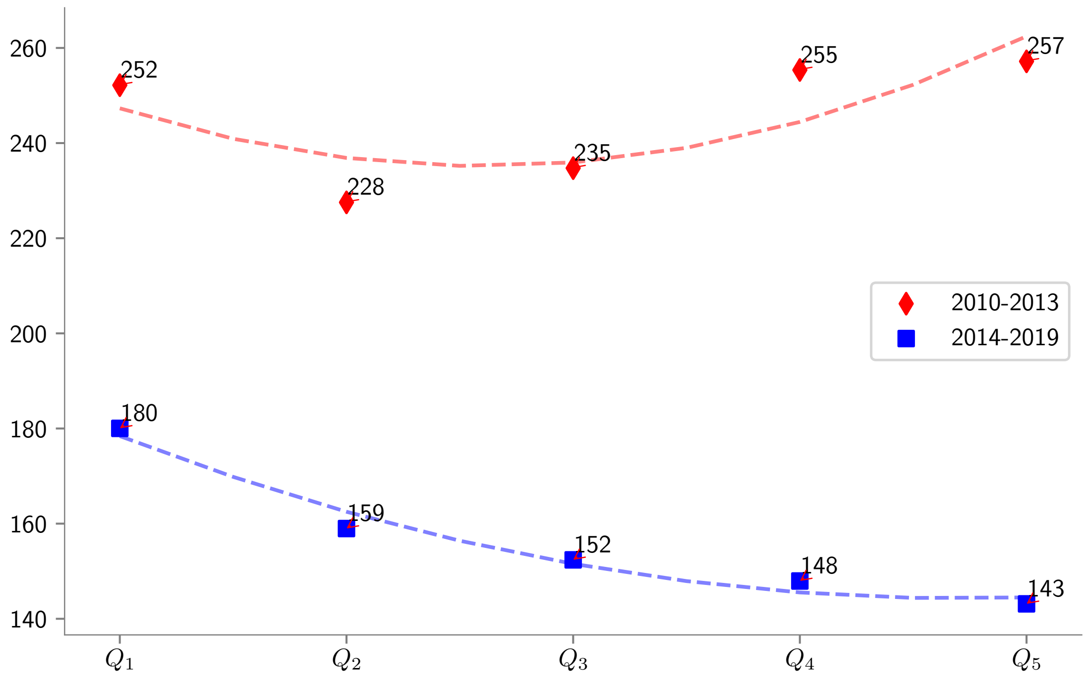
Sorted-portfolio approach

- Sorted-based approach of Fama-French (1992)
- At each rebalancing date t, we rank the bonds according to their Amundi **ESG** z-score
- We form the five quintile portfolios Q_i for $i = 1, \ldots, 5$
- The portfolio Q_i is invested during the period $]t, t+1]$:
  - Q_1 corresponds to the best-in-class portfolio (best scores)
  - Q_5 corresponds to the worst-in-class portfolio (worst scores)
- Monthly rebalancing
- Universe: ICE (BofAML) Large Cap IG EUR Corporate Bond
- Sector-weighted and sector-neutral portfolio
- Within a sector, bonds are equally-weighted

What is the performance of ESG investing?

Sorted portfolios

Figure 112: Annualized credit return in bps of **ESG** sorted portfolios (EUR IG, 2010 – 2019)



Source: Amundi Quantitative Research (2020)

Table 74: Carry statistics (in bps)

Period	Q_1	Q_5
2010-2013	175	192
2014-2019	113	128

- Negative carry (coupon level)
- Positive mark-to-market (dynamics of credit spreads and bond prices)

Bond portfolio optimization

We consider the following optimization problem:

$$x^{\star}(\gamma) = \arg\min \mathcal{R}(x \mid b) - \gamma \cdot \mathcal{S}(x \mid b)$$

where:

$$\mathcal{R}(x \mid b) = \frac{1}{2}\mathcal{R}_{\mathrm{MD}}(x \mid b) + \frac{1}{2}\mathcal{R}_{\mathrm{DTS}}(x \mid b)$$

and:

- $\mathcal{R}_{\mathrm{MD}}(x \mid b)$ and $\mathcal{R}_{\mathrm{DTS}}(x \mid b)$ are the interest rate and credit **active risk** measures wrt the benchmark b
- $\mathcal{S}(x \mid b)$ is the ESG excess score of Portfolio x wrt the benchmark b

The objective is to find the optimal portfolio minimizing interest rate and credit active risk for a given ESG excess score

What is the performance of ESG investing?

Optimized portfolios

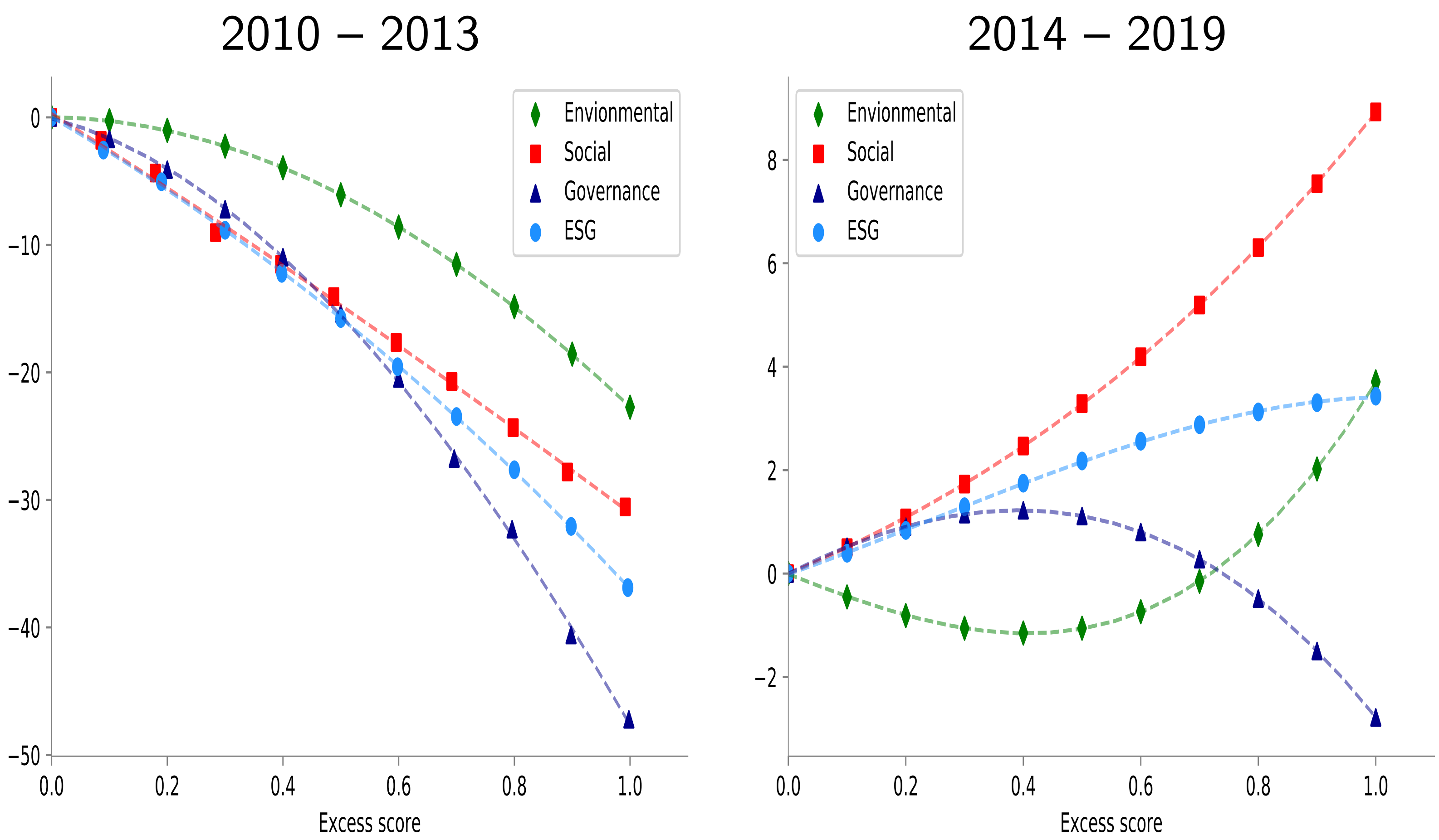


Figure 113: Excess credit return in bps of optimized portfolios (EUR IG)

Source: Amundi Quantitative Research (2020)

What is the performance of ESG investing?

Optimized portfolios

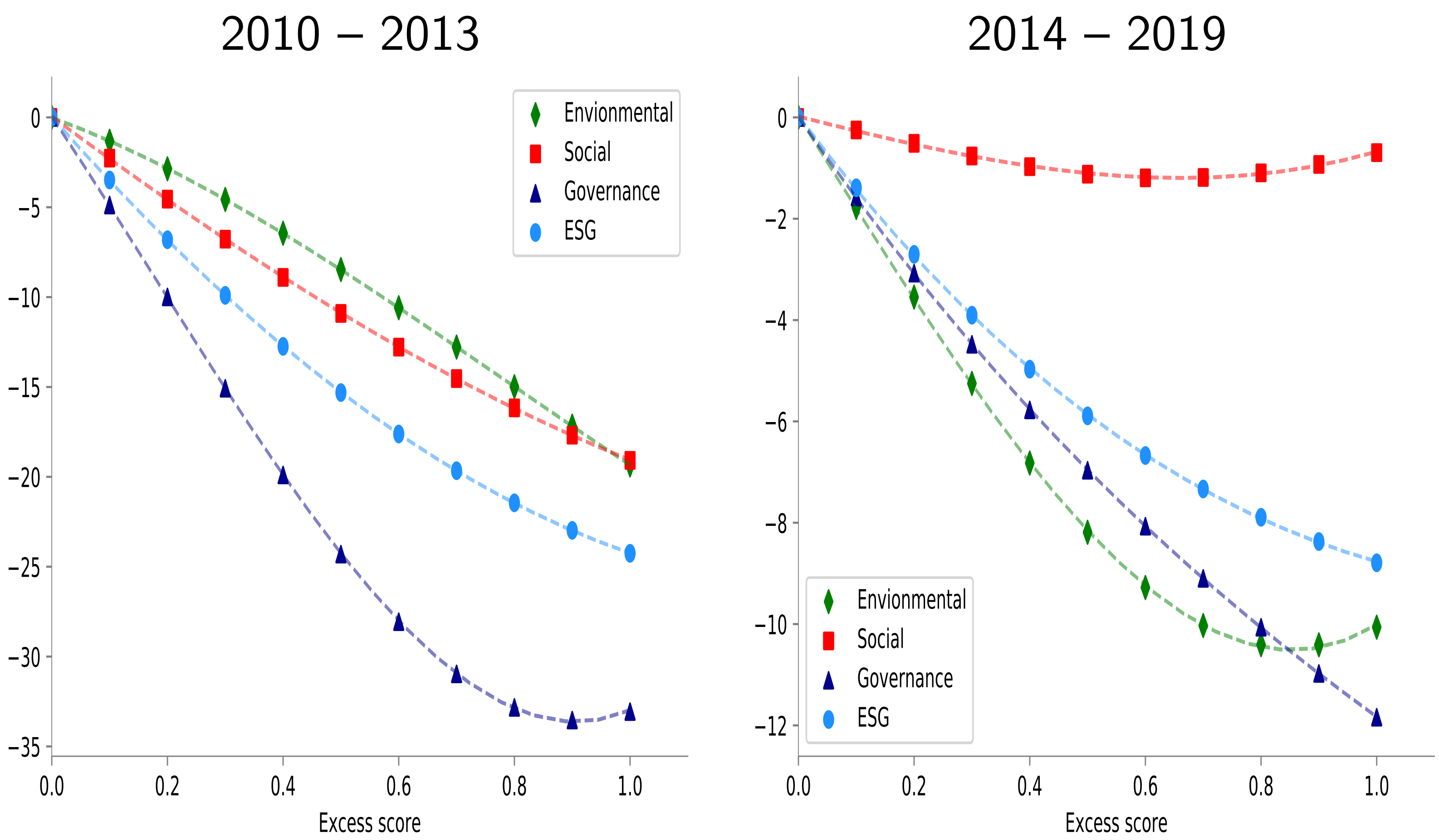


Figure 114: Excess credit return in bps of optimized portfolios (USD IG)

Source: Amundi Quantitative Research (2020)

The impact of ESG on the funding cost

An integrated Credit-ESG model

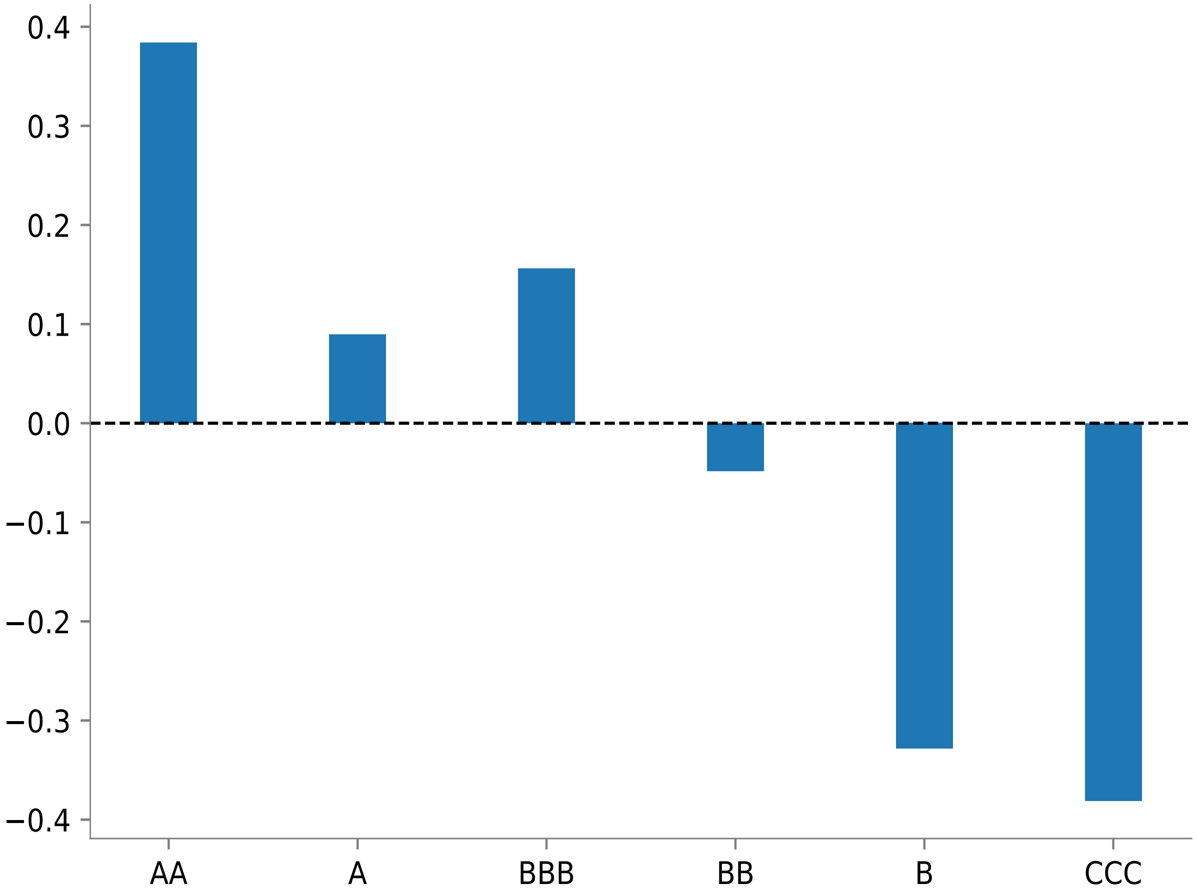


Figure 115: Average **ESG** score with respect to the credit rating (2010 – 2019)

Source: Amundi Quantitative Research (2020)

The impact of ESG on the funding cost

An integrated Credit-ESG model

We consider the following regression model:

$$\begin{aligned}\ln \mathrm{OAS}_{i,t} &= \alpha_t + \beta_{esg} \cdot \mathcal{S}_{i,t} + \beta_{md} \cdot \mathrm{MD}_{i,t} + \sum_{j=1}^{N_{\mathcal{S}ector}} \beta_{\mathcal{S}ector}(j) \cdot \mathcal{S}ector_{i,t}(j) + \\ & \beta_{sub} \cdot \mathrm{SUB}_{i,t} + \sum_{k=1}^{N_{\mathcal{R}ating}} \beta_{\mathcal{R}ating}(k) \cdot \mathcal{R}ating_{i,t}(k) + \varepsilon_{i,t}\end{aligned}$$

where:

- $\mathcal{S}_{i,t}$ is the **ESG** z-score of Bond i at time t
- $\mathrm{SUB}_{i,t}$ is a dummy variable accounting for subordination of the bond
- $\mathrm{MD}_{i,t}$ is the modified duration
- $\mathcal{S}ector_{i,t}(j)$ is a dummy variable for the j^{th} sector
- $\mathcal{R}ating_{i,t}(k)$ is a dummy variable for the k^{th} rating

The impact of ESG on the funding cost

An integrated Credit-ESG model

Table 75: Results of the panel data regression model (EUR IG, 2010 – 2019)

	2010–2013				2014–2019			
	ESG	**E**	**S**	**G**	**ESG**	**E**	**S**	**G**
R^2	60.0%	59.4%	59.5%	60.3%	66.3%	65.0%	65.2%	64.6%
Excess R^2 of ESG	0.6%	0.0%	0.2%	1.0%	4.0%	2.6%	2.9%	2.3%
$\hat{\beta}_{esg}$	-0.05	-0.01	-0.02	-0.07	-0.09	-0.08	-0.08	-0.08
t-statistic	-32	-7	-16	-39	-124	-98	-104	-92

Source: Amundi Quantitative Research (2020)

The assumption $\mathcal{H}_0 : \beta_{esg} < 0$ is not rejected

The impact of ESG on the funding cost

ESG cost of capital with min/max score bounds

We calculate the difference between:

(1) the funding cost of **the worst-in-class issuer** and

(2) the funding cost of **the best-in-class issuer**

by assuming that:

- the two issuers have the same credit rating;
- the two issuers belong to the same sector;
- the two issuers have the same capital structure;
- the two issuers have the same debt maturity.

$\Rightarrow$ Two approaches:

1. Theoretical approach: ESG scores are set to -3 and $+3$ (not realistic)
2. Empirical approach: ESG scores are set to observed min/max score bounds (e.g. min/max $= -2.0/+1.9$ for Consumer Cyclical A-rated EUR, $-2.1/+3.2$ for Banking A-rated EUR, etc.)

The impact of ESG on the funding cost

ESG cost of capital with min/max score bounds

Table 76: **ESG** cost of capital (IG, 2014 – 2019)

	EUR				USD			
	AA	A	BBB	Average	AA	A	BBB	Average
Banking	23	45	67	45	11	19	33	21
Basic	9	25	44	26	5	15	34	18
Capital Goods	8	32	42	27	6	15	26	16
Communication		26	48	37	5	11	23	13
Consumer Cyclical	3	26	43	28	2	8	17	10
Consumer Non-Cyclical	15	29	31	25	6	12	19	12
Utility & Energy	12	32	56	33	9	14	31	18
Average	12	31	48	**31**	7	13	26	**15**

Source: Amundi Quantitative Research (2020)

ESG investing versus ESG financing

- Markowitz, H. (1952), Portfolio Selection, *Journal of Finance*, 7(1), pp. 77-91.
- Modigliani, F., and Miller, M.H. (1958), The Cost of Capital, Corporation Finance and the Theory of Investment, *American Economic Review*, 48(3), pp. 261-297.

⇒ Two misunderstandings:

1. Capital allocation & asset allocation
2. Cost of capital & asset (stock/bond) return

Prologue

"There is no Plan B, because there is no Planet B"

Ban Ki-moon, UN Secretary-General, September 2014

Is it a question of climate-related issues?

In fact, it is more an economic growth issue

"The Golden Rule of Accumulation: A Fable for Growthmen"

Edmund Phelps, *American Economic Review*, 1961
Nobel Prize in Economics, 2006

Climate risks and financial losses

Climate risks transmission channels to financial stability

- The **physical risks** that arise from the increased frequency and severity of climate and weather related events that damage property and disrupt trade
- The **liability risks** stemming from parties who have suffered loss from the effects of climate change seeking compensation from those they hold responsible
- The **transition risks** that can arise through a sudden and disorderly adjustment to a low carbon economy

Speech by Mark Carney at the International Climate Risk Conference for Supervisors, Amsterdam, April 6, 2018

Climate risks and financial risks

Risks are traversal to financial risks

- **Carbon risk** (reputational and regulation risks) $\Rightarrow$ economic, market and credit risks
- **Climate risk** (extreme weather events, natural disasters) $\Rightarrow$ economic, operational, credit and market risks

Some definitions

Climate risk

Climate Risks include transition risk and physical risks:

- Transition risk is defined as the financial risk associated with the transition to a low-carbon economy. It include policy changes, reputational impacts, and shifts in market preferences, norms and technology
- Physical risk is defined as the financial losses due to extreme weather events and climate disasters like flooding, sea level rise, wildfires, droughts and storms

Some definitions

Global warming (≈ climate change)

Global warming is the long-term heating of Earth's climate system observed since the pre-industrial period (between 1850 and 1900) due to human activities, primarily fossil fuel burning

NASA Global Climate Change — `https://climate.nasa.gov`

Some definitions

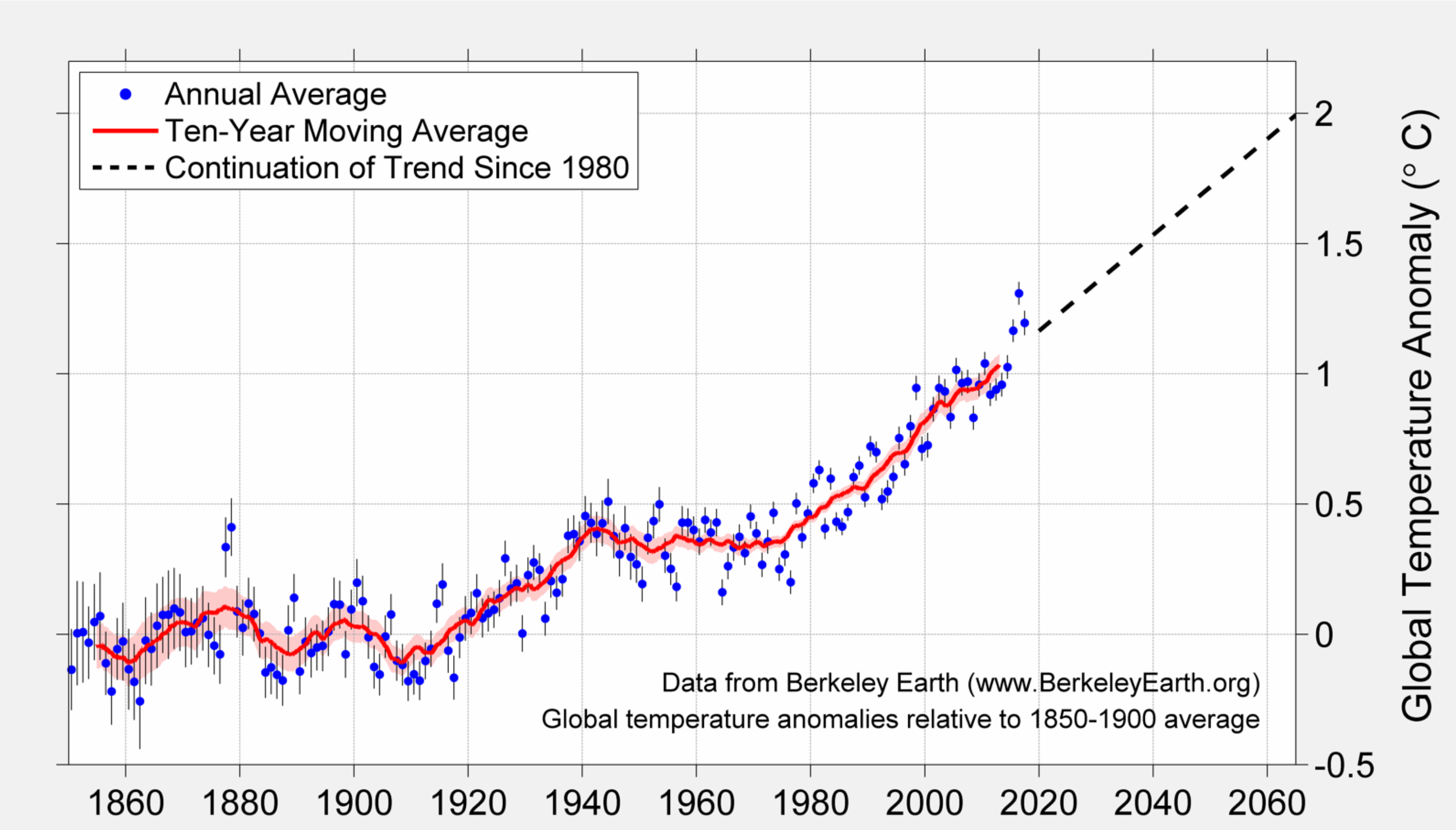


Figure 116: Global temperature anomaly

Source: Berkeley Earth (2018), `http://berkeleyearth.org`

Some definitions

Carbon risk

Carbon risks correspond to the potential financial losses due to greenhouse gas (or GHG) emissions, mainly CO_2 emissions (in a strengthening regulatory context)

Some definitions

GHG

Greenhouse gases absorb and emit radiation energy, causing the greenhouse effect[a]:

1. water vapour (H_2O)
2. Carbon dioxide (CO_2)
3. Methane (CH_4)
4. Nitrous oxide (N_2O)
5. Ozone (O_3)

---

[a] Without greenhouse effect, the average temperature of Earth's surface would be about -18°C. With greenhouse effect, the current temperature of Earth's surface is about $+15^\circ$C.

Some definitions

Carbon equivalent

Carbon dioxide equivalent (or CO_2e) is a term for describing different GHG in a common unit

- A quantity of GHG can be expressed as CO_2e by multiplying the amount of the GHG by its global warming potential (GWP)
- 1 kg of methane corresponds to 25 kg of CO_2
- 1 kg of Nitrous oxide corresponds to 310 kg of CO_2

CO_2 emissions

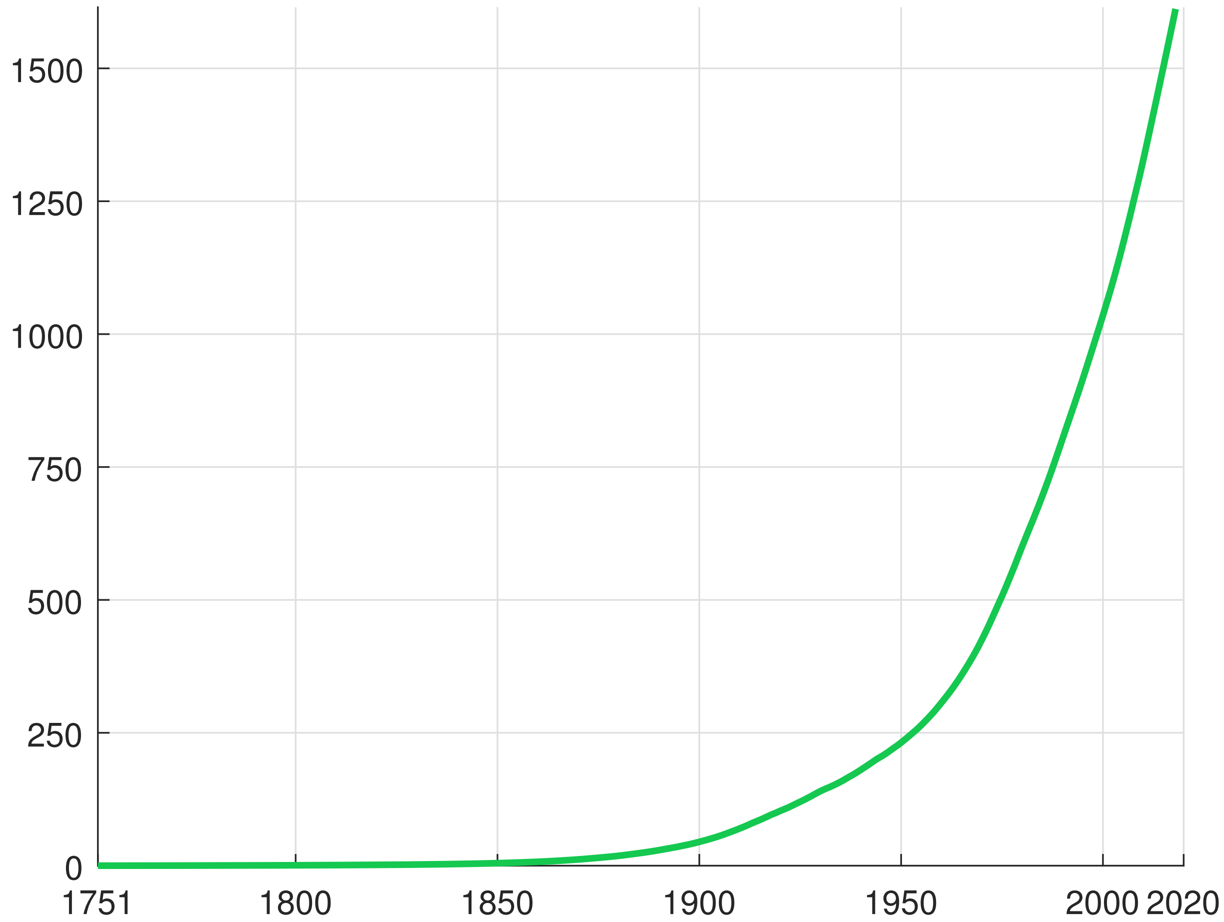


Figure 117: Cumulative CO_2 emissions (in BT)

Source: Data on CO_2 and GHG Emissions by Our World in Data (https://github.com/owid/co2-data)

CO_2 emissions

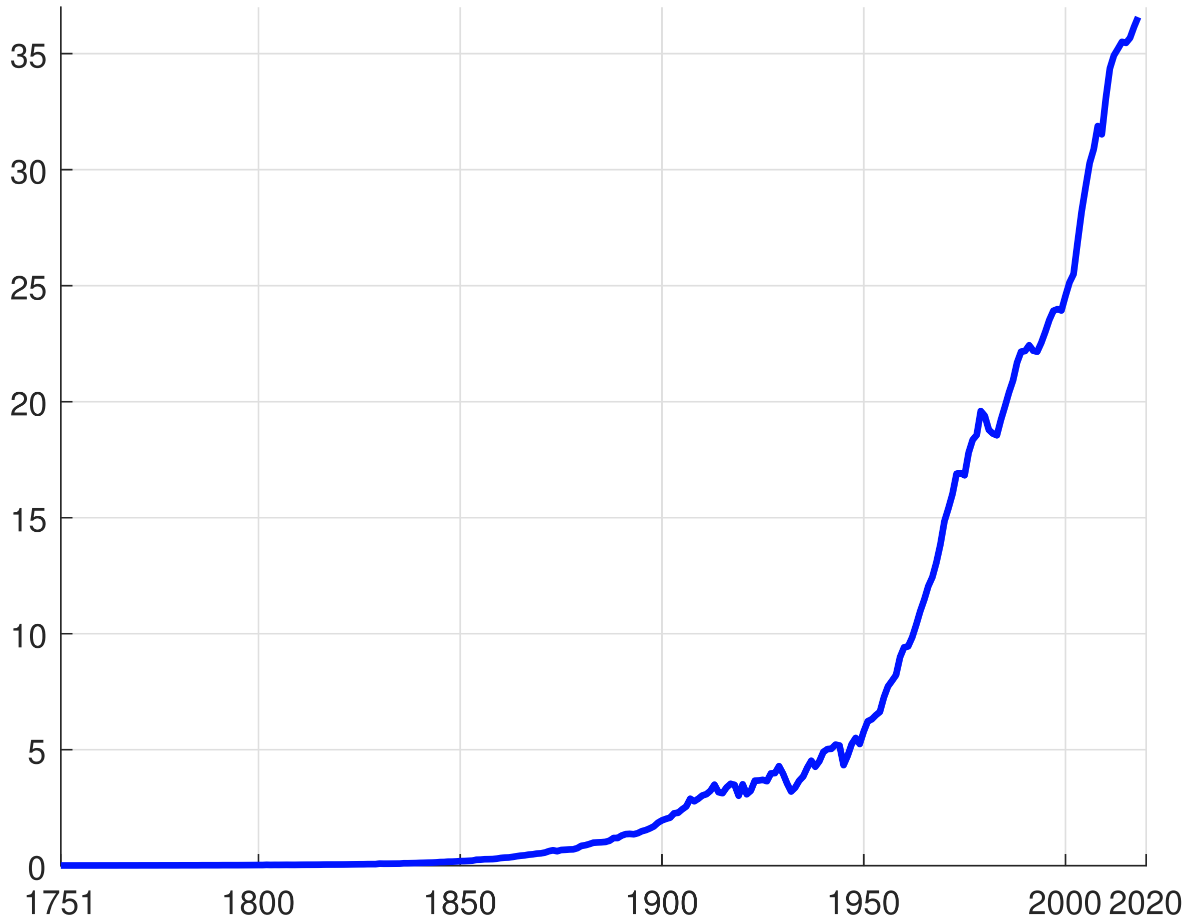


Figure 118: Annual CO_2 emissions (in BT)

Source: Data on CO_2 and GHG Emissions by Our World in Data (https://github.com/owid/co2-data)

CO_2 emissions

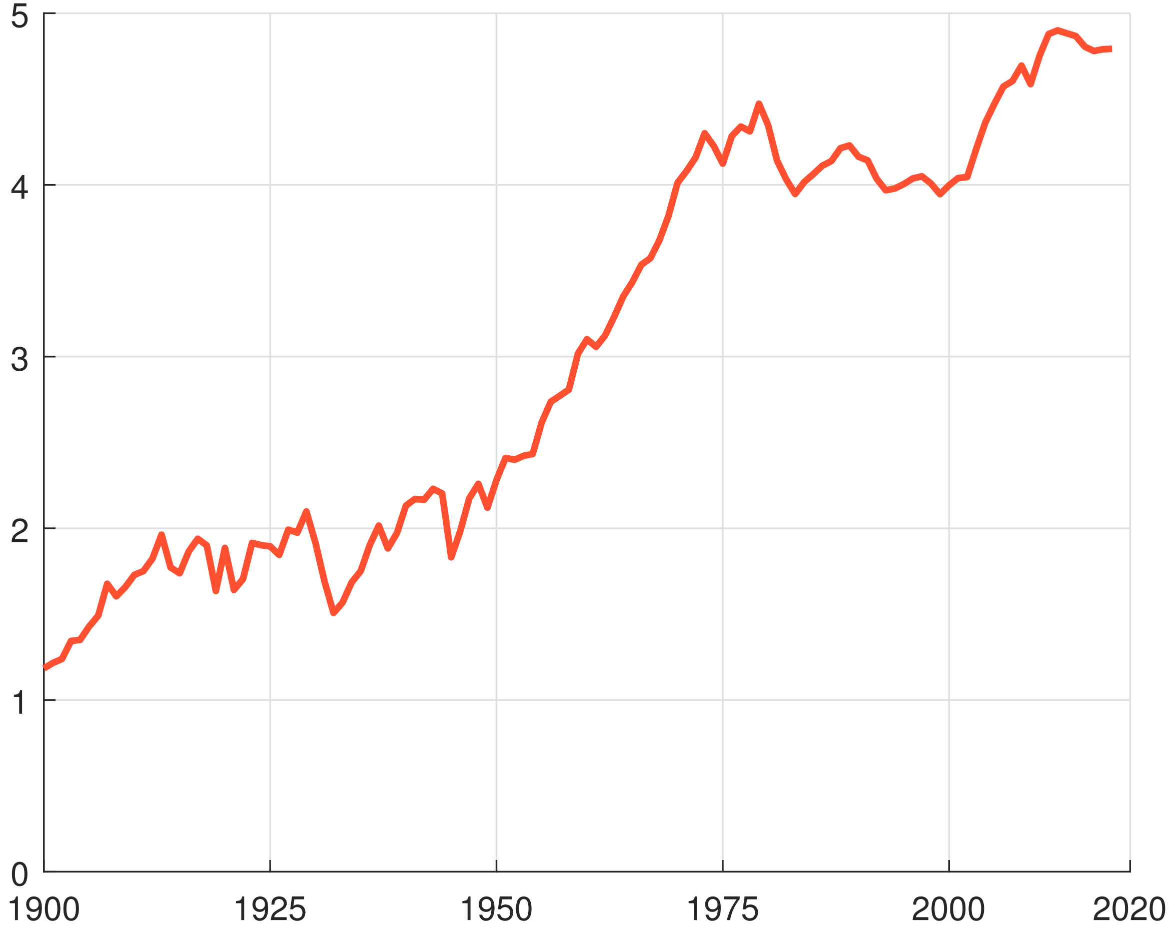


Figure 119: CO_2 emissions per capita (in MT)

Source: Data on CO_2 and GHG Emissions by Our World in Data (https://github.com/owid/co2-data)

CO_2 emissions

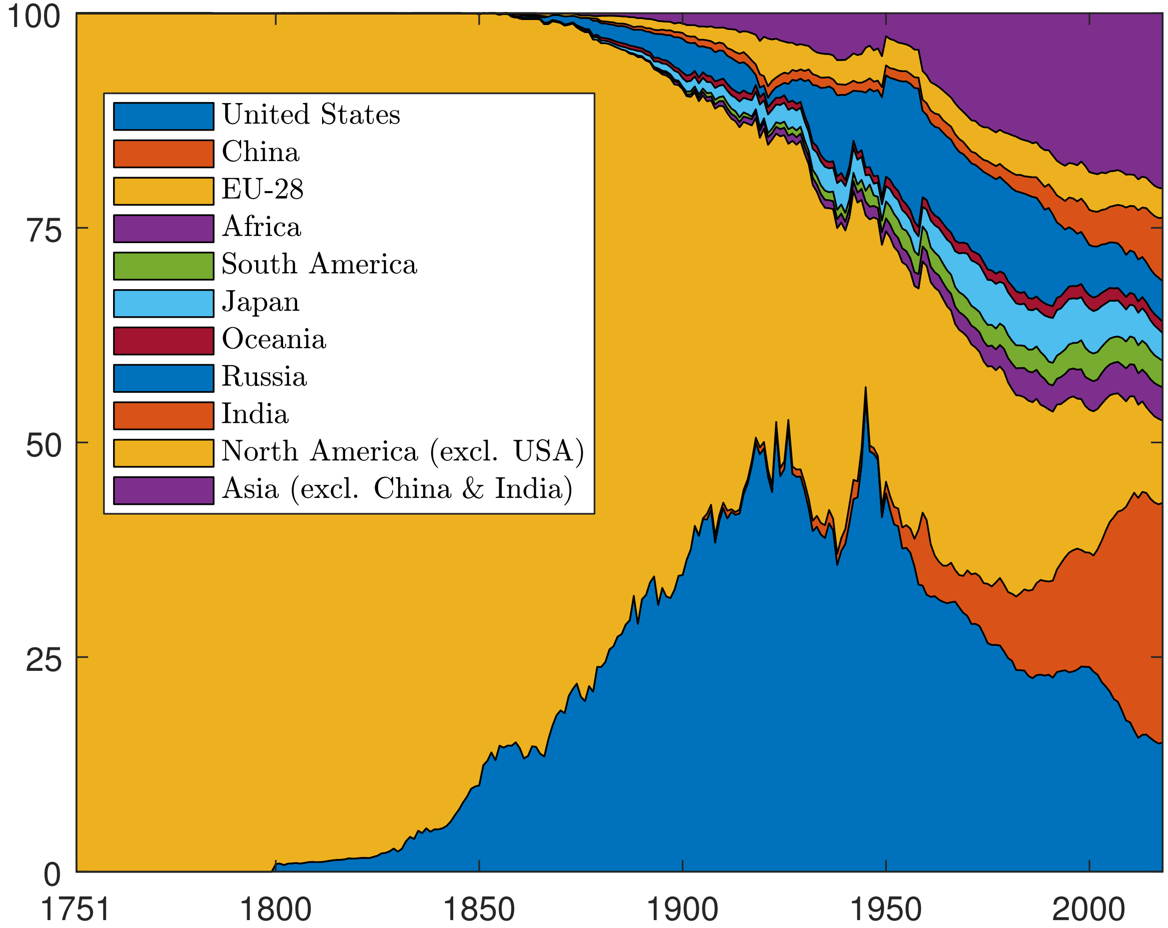
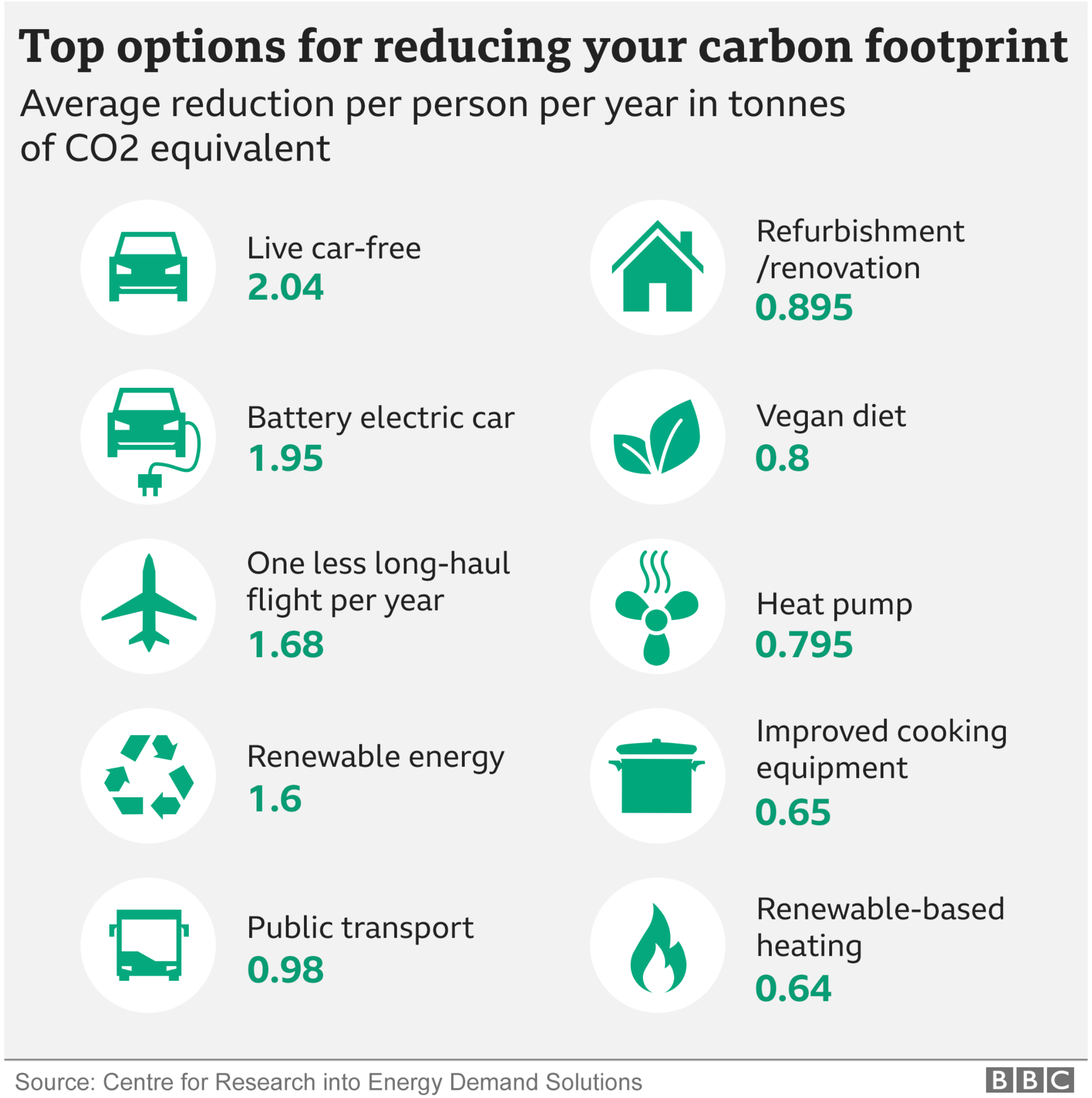


Figure 120: Share of CO_2 emissions (in %)

Source: Data on CO_2 and GHG Emissions by Our World in Data (`https://github.com/owid/co2-data`)

CO_2 emissions



IPCC

- The Intergovernmental Panel on Climate Change (IPCC) is the United Nations body for assessing the science related to climate change
- The IPCC was created to provide policymakers with regular scientific assessments on climate change, its implications and potential future risks, as well as to put forward adaptation and mitigation options
- Website: `https://www.ipcc.ch`

Remark

IPCC is known as "*Groupe d'experts intergouvernemental sur l'évolution du climat*" (GIEC)

IPCC

IPCC working groups

- The IPCC Working Group I (WGI) examines the physical science underpinning past, present, and future climate change
- The IPCC Working Group II (WGII) assesses the impacts, adaptation and vulnerabilities related to climate change
- The IPCC Working Group II (WGIII) focuses on climate change mitigation, assessing methods for reducing greenhouse gas emissions, and removing greenhouse gases from the atmosphere

IPCC

Some famous reports

- IPCC Fifth Assessment Report (AR5): Climate Change 2014 — `www.ipcc.ch/report/ar5`
- Global Warming of 1.5°C — `www.ipcc.ch/sr15`
- IPCC Sixth Assessment Report (AR6): Climate Change 2022 — `www.ipcc.ch/report/sixth-assessment-report-cycle`

IPCC scenarios

- Website: `https://www.ipcc.ch/data`
- The IPCC AR5 scenarios database comprises 31 models and in total 1 184 scenarios
- 4 reference scenarios: **representative concentration pathways** (RCP)
- Each RCP represents one possible evolution profile of GHG concentrations
    - RCP 2.6: CO_2 emissions start declining by 2020 and go to zero by 2100
    - RCP 4.5: CO_2 emissions peak around 2040, then decline
    - RCP 6.0: CO_2 emissions peak around 2080, then decline
    - RCP 8.5: CO_2 emissions continue to rise throughout the 21st century
- For each RCP, socio-economic development scenarios and various adaptation and mitigation strategies are associated
- They are called the **shared socioeconomic pathways** (SSP)

IPCC scenarios

RCP	Model	Contact
RCP 2.6	IMAGE	Detlef van Vuuren (detlef.vanvuuren@pbl.nl)
RCP 4.5	MiniCAM	Katherine Calvin (katherine.calvin@pnnl.gov)
RCP 6.0	AIM	Toshihiko Masui (masui@nies.go.jp)
RCP 8.5	MESSAGE	Keywan Riahi (riahi@iiasa.ac.at)

Table 77: Associated model for each RCP

IPCC scenarios

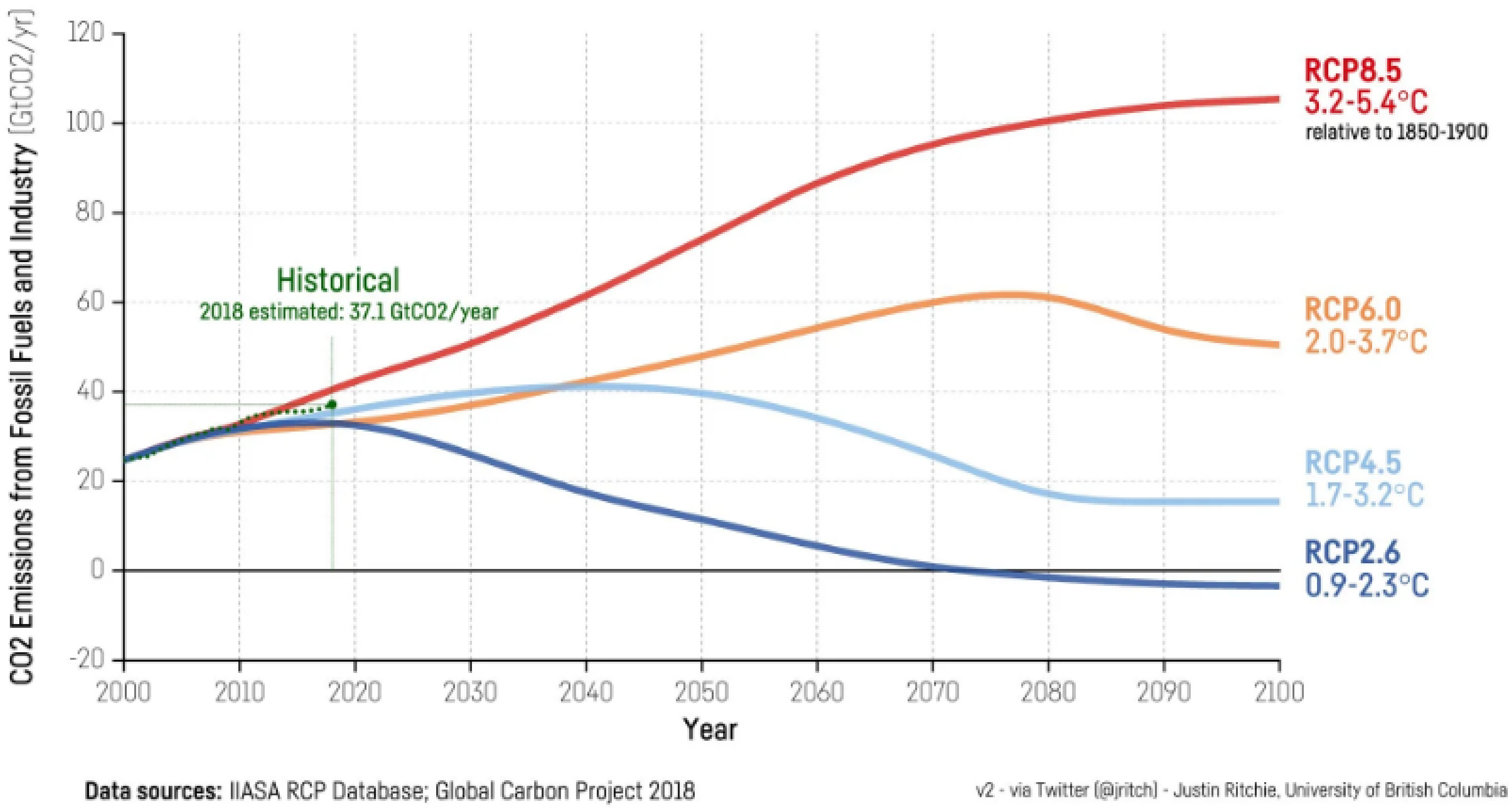


Figure 121: IPCC RCP scenarios: CO_2 emissions from fossil fuels and industry

Carbon neutrality

Carbon neutrality (or net zero) means that any CO2 released into the atmosphere from human activity is balanced by an equivalent amount being removed

Apple Commits to Become Carbon Neutral to by 2030 (https://www.bbc.com/news/technology-53485560)

Carbon dioxide removal

Carbon dioxide removal (CDR)

1. Nature-based solutions
   - Afforestation
   - Reforestation
   - Restoration of peat bogs
   - Restoration of coastal and marine habitats
2. Enhanced natural processes
   - Land management and no-till agriculture, which avoids carbon release through soil disturbance
   - Better wildfire management
   - Ocean fertilisation to increase its capacity to absorb CO2
3. Technology solutions
   - Bioenergy with carbon capture and storage (BECCS)
   - Direct air capture (DAC)

The shared socioeconomic pathways

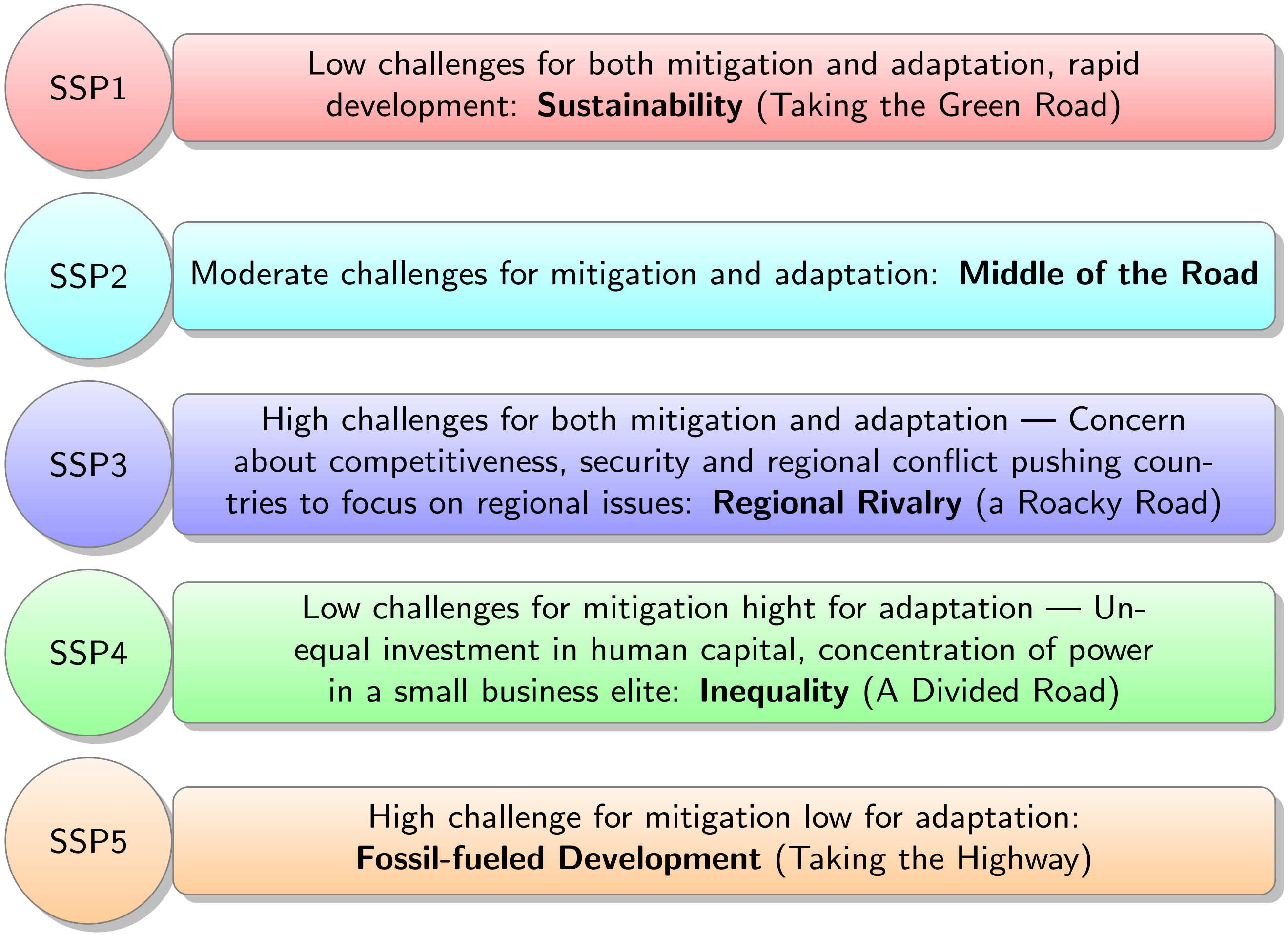


Figure 122: The shared socioeconomic pathways

Source: O'Neill *et al.* (2016)

The shared socioeconomic pathways

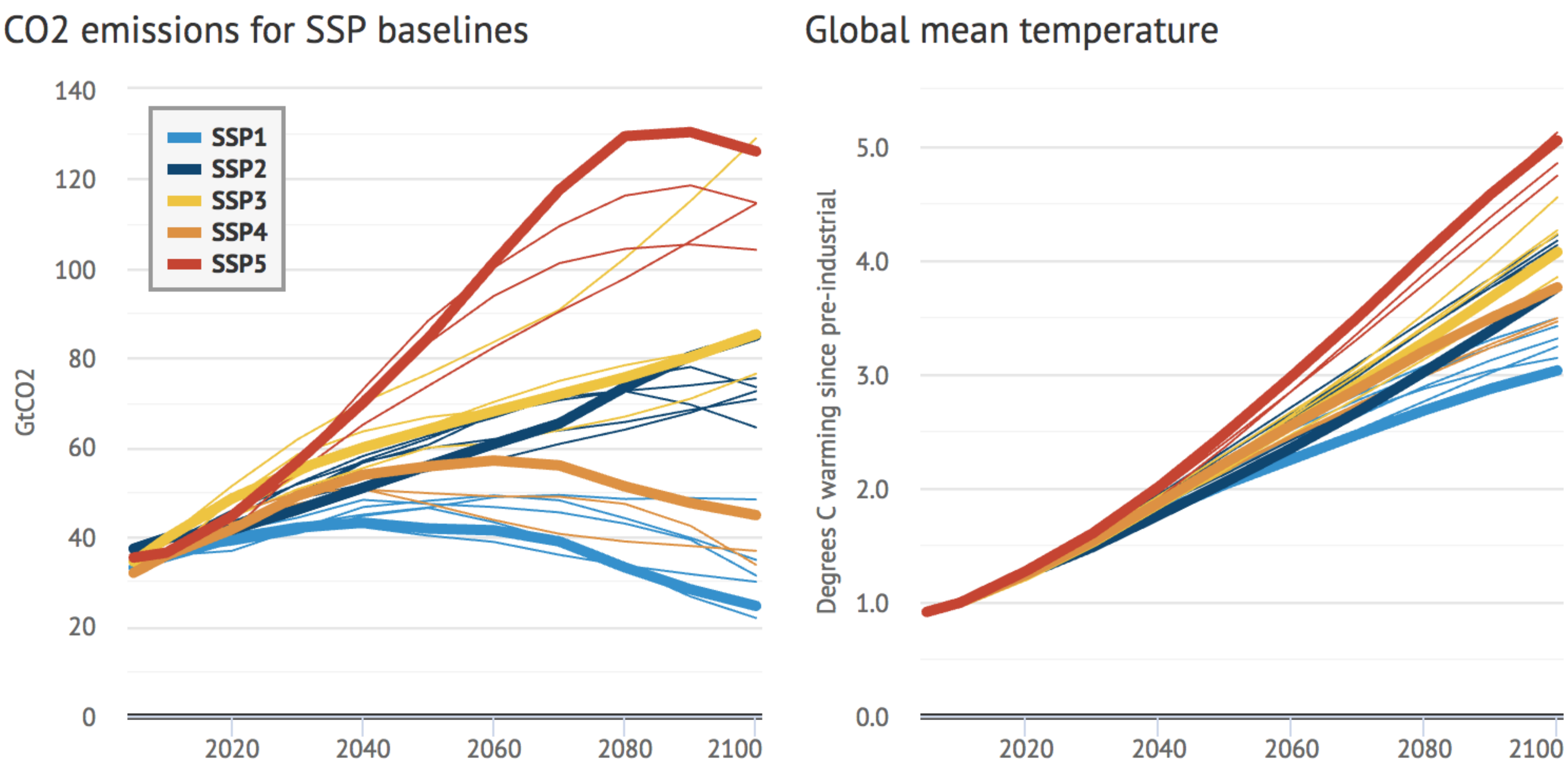


Figure 123: Projections of CO_2 emissions and temperatures across SSP

Source: https://www.carbonbrief.org/explainer-how-shared-socioeconomic-pathways-explore-future-climate-change

The shared socioeconomic pathways

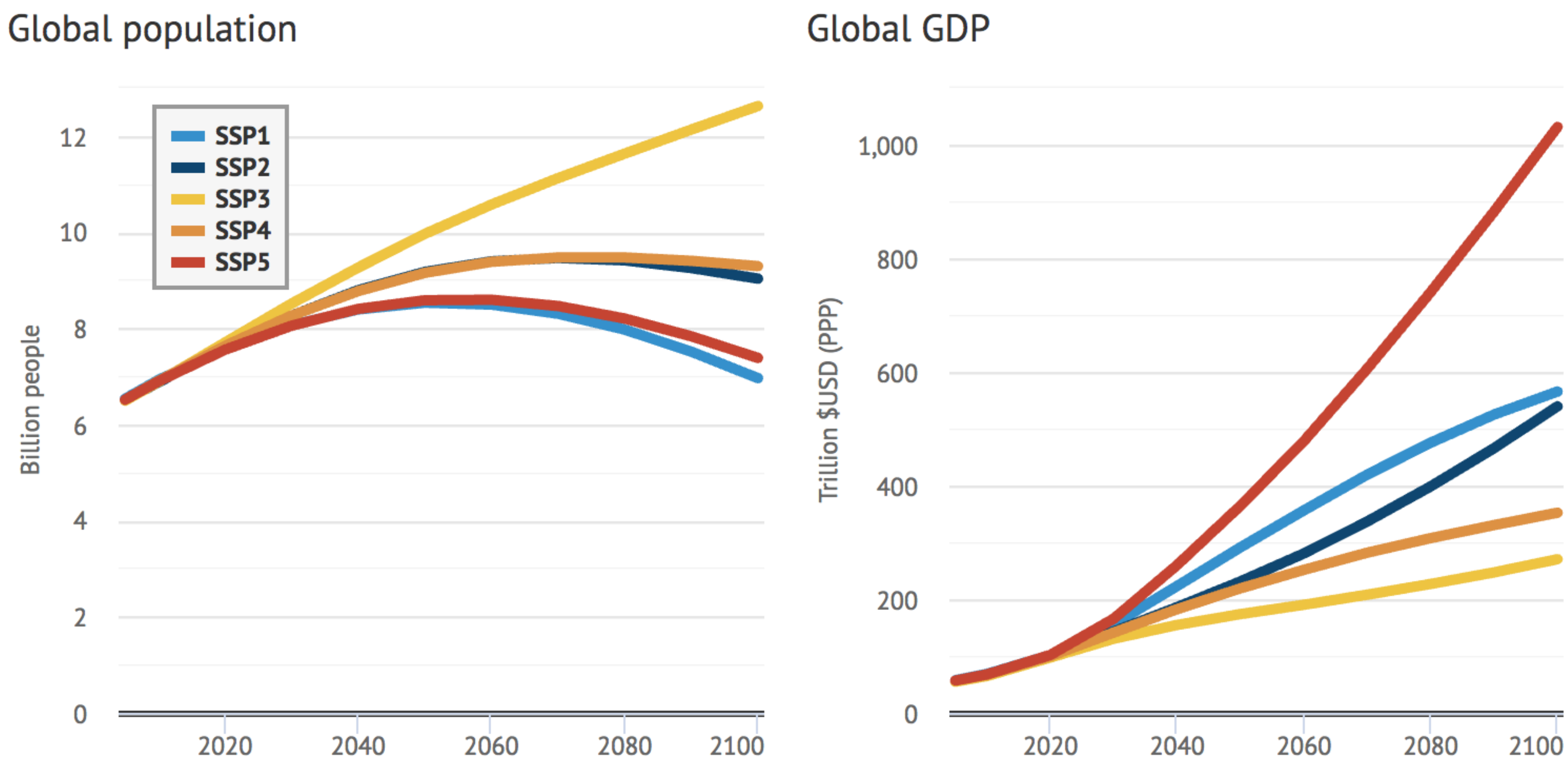


Figure 124: Projections of population and economic growth across SSP

Source: `https://www.carbonbrief.org/explainer-how-shared-socioeconomic-pathways-explore-future-climate-change`

Climate risk and missing factors

The example of permafrost

- The permafrost contains 1.700 billion tons of carbon, almost double the amount of carbon that is currently in the atmosphere.
- Arctic permafrost holds roughly 15 million gallons of mercury – at least twice the amount contained in the oceans, atmosphere and all other land combined.
- A global temperature rise of 1.5°C above current levels would be enough to start the thawing of permafrost in Siberia.
- The global warming will become out-of-control after this tipping point.
- The thawing of the permafrost also threatens to unlock disease-causing viruses long trapped in the ice.

$\Rightarrow$ The survival of Humanity becomes uncertain if the tipping point is reached

Climate risk modeling

Remark

In what follows, we use the survey and the simulations of Le Guenedal (2019)

Climate risk modeling

The Solow growth model

The model

- Production function:

$$Y(t) = F(K(t), A(t)L(t))$$

  where $K(t)$ is the capital, $L(t)$ is the labor and $A(t)$ is the knowledge factor

- Law of motion for the capital per unit of effective labor $k(t) = K(t) / (A(t)L(t))$:

$$\frac{\mathrm{d}k(t)}{\mathrm{d}t} = s f(k(t)) - (g_L + g_A + \delta_K) k(t)$$

  where s is the saving rate, δ_K is the depreciation rate of capital and g_A and g_L are the productivity and labor growth rates

Climate risk modeling

The golden rule

Golden rule with the Cobb-Douglas production and Hicks neutrality

The equilibrium to respect the '*fairness*' between generations is:

$$k^\star = \left(\frac{s}{g_L + g_A + \delta_K} \right)^{\dfrac{1}{1-\alpha}}$$

"*Each generation in a boundless golden age of natural growth will prefer the same investment ratio, which is to say the same natural growth path*" (Phelps, 1961, page 640).

"*By a golden age I shall mean a dynamic equilibrium in which output and capital grow exponentially at the same rate so that the capital-output ratio is stationary over time*" (Phelps, 1961, page 639).

Climate risk modeling

Golden rule and climate risk

What is economic growth and what is the balanced growth path?

- There is a saving rate that maximizes consumption over time and between generations ("**the fair rate to preserve future generations**")
- Economic growth corresponds to the exponential growth of capital and output to answer the needs of the growing population
- Introducing human and natural capitals add constraints and therefore **reduce growth**!

$$\text{Economic growth} \Rightarrow \begin{cases} \text{productivity} \nearrow \text{ and labor} \nearrow \\ \text{maximization of } \textbf{consumption-based utility} \text{ function} \end{cases}$$

Climate risk modeling
Extension to natural capital

What are the effects of environmental constraints on growth?

Introducing a decreasing natural capital (Romer, 2006)

The balanced growth path $g_Y^\star$ is equal to:

$$g_Y^\star = g_L + g_A - \frac{g_L + g_A + \delta_{N_c}}{1 - \alpha} \vartheta$$

where δ_{N_c} is the depreciation rate of natural capital and ϑ is the elasticity of output with respect to (normalized) natural capital $N_c(t)$

"*The static-equilibrium type of economic theory which is now so well developed is plainly inadequate for an industry in which the indefinite maintenance of a steady rate of production is a physical impossibility, and which is therefore bound to decline*" (Hotteling, 1931, page 138-139)

Accounting for environment... changes the definition of economic growth

Climate risk modeling

Inter-temporal utility functions

Preferences modeling (Ramsey model)

- ρ is the discount rate (time preference)
- $c(t)$ is the consumption per capita and u is the CRRA utility function:

$$u(c(t)) = \begin{cases} \dfrac{1}{1-\theta} c(t)^{1-\theta} & \text{if } \theta > 0, \quad \theta \neq 1 \\ \ln c(t) & \text{if } \theta = 1 \end{cases}$$

  where θ is the risk aversion parameter
- Maximization of the welfare function:

$$\int_t^{\infty} e^{-\rho t} u(c(t)) \, \mathrm{d}t$$

Climate risk modeling

The discounting issue

Does the golden rule of saving rates hold in a Keynesian approach with discounted maximization of consumption?

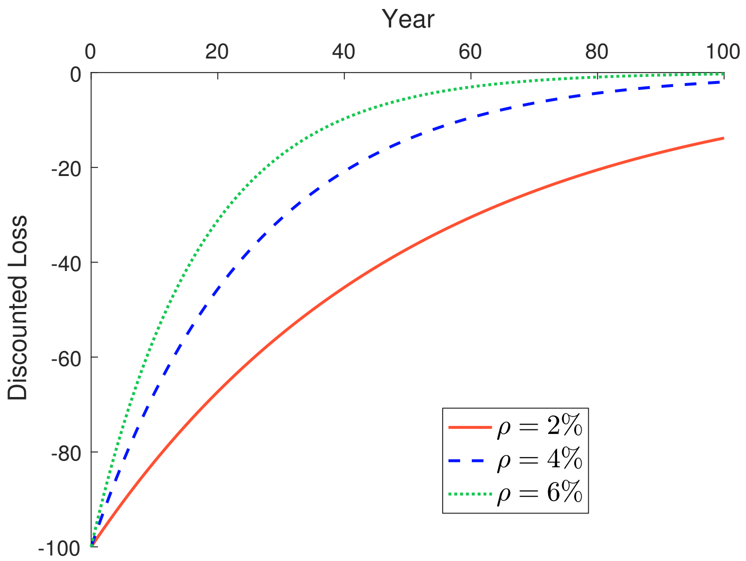


Figure 125: Discounted value of $100 loss

- "*There is still time to avoid the worst impacts of climate change, if we take strong action now*" (Stern, 2007)
- "*I got it wrong on climate change – it's far, far worse*" (Stern, 2013)

The value of a loss in 100 years almost disappears... while it is only the next generation!

Climate risk modeling

Does consumption maximization make sense?

How many planets do we need?

To achieve the current levels of consumption for the world population, we need:

- US: 5 planets
- France: 3 planets
- India: 0.6 planet

Source: Global Footprint Network, `http://www.footprintcalculator.org`

Climate risk modeling

Fairness between generations

Keynes

"In the long run, we are all dead"

John Maynard Keynes[a], *A Tract on Monetary Reform*, 1923.

[a] *"Men will not always die quietly"*, *The Economic Consequences of the Peace*, 1919.

Carney

"The Tragedy of the Horizon"

Mark Carney, Chairman of the Financial Stability Board, 2015

⇒ Back to the Golden Rule and the Fable for Growthmen...

Integrated assessment model (IAM)

Definition

Main categories

- **Optimization models**
  The inputs of these models are parameters and assumptions about the structure of the relationships between variables. The outputs provided by optimization process are scenarios depending on a set of constraints
- **Evaluation models**
  Based on exogenous scenarios, the outputs provide results from partial equilibriums between variables

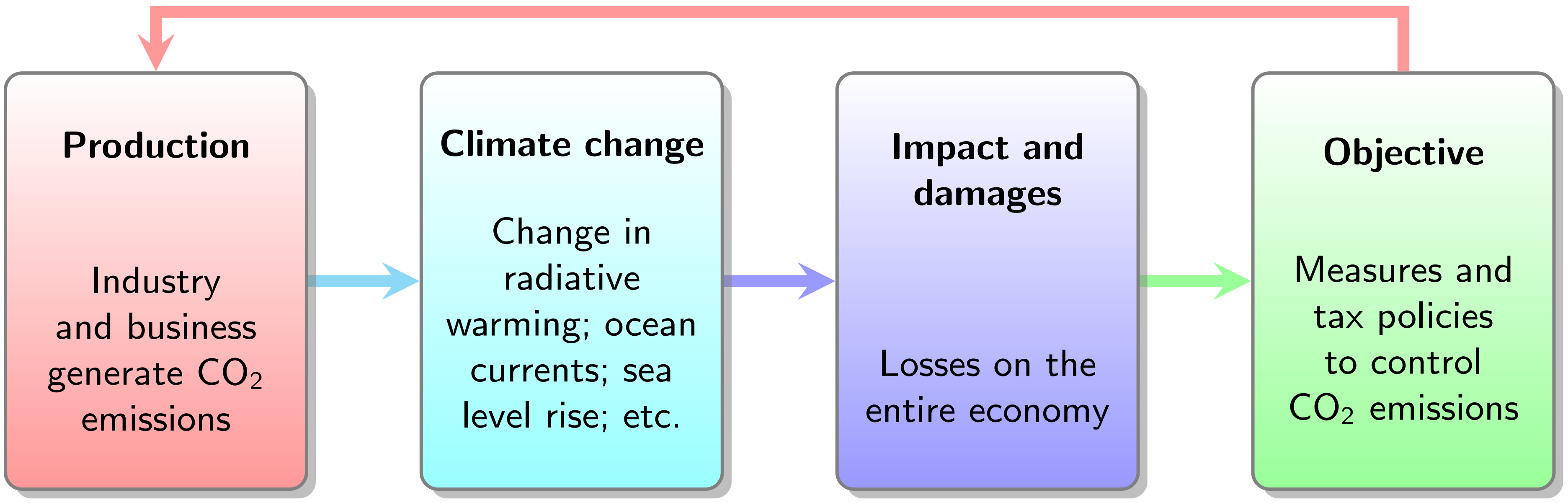
Three main components of IAMs

1. Economic growth relationships
2. Dynamics of climate emissions
3. Objective function

Integrated assessment model (IAM)

Modeling framework

Figure 126: Economic models of climate risk



Integrated assessment model (IAM)

Modeling framework

1. Economic module
    1. Production function $\Longrightarrow$ GDP
    2. Impact of the climate risk on GDP (damage losses, mitigation and adaptation costs)
    3. The climate loss function depends on the temperature
2. Climate module
    1. Dynamics of GHG emissions
    2. Modeling of Atmospheric and lower ocean temperatures
3. Optimal control problem
    1. Maximization of the utility function
    2. We can test many variants

Integrated assessment model (IAM)

Modeling framework

The most famous IAM is the **Dynamic Integrated model of Climate and the Economy** (or DICE) developed by Nordhaus[19] (1993)

The RICE model (Regional Integrated Climate-Economy model) is a variant of the DICE model

[19] 2018 Nobel Laureate

Integrated assessment model (IAM)

Production and output

- The **gross output** is equal to:

$$Y(t) = A_{\mathrm{TFP}}(t) K(t)^{\alpha} L(t)^{1-\alpha}$$

where:

$$\left\{\begin{array}{l} A_{\mathrm{TFP}}(t) = (1 + g_A(t)) A_{\mathrm{TFP}}(t-1) \\ K(t) = (1 - \delta_K) K(t-1) + I(t) \\ L(t) = (1 + g_L(t)) L(t-1) \end{array}\right.$$

- Climate change impacts the **net output**:

$$Q(t) = \Omega_{\mathrm{Climate}}(t) Y(t)$$

- We also have $Q(t) = C(t) + I(t)$ and $C(t) = (1 - s(t)) Q(t)$

Integrated assessment model (IAM)

The loss (or damage) function

- The loss function is given by:

$$\Omega_{\text{Climate}}(t) = \Omega_D \cdot \Omega_\Lambda = \frac{1}{1 + D(t)} \cdot (1 - \Lambda(t))$$

where $D(t)$ and $\Lambda(t)$ measure climate damages[20] and abatement costs[21]

- Climate damages are assumed to be quadratic:

$$D(t) = a_1 \mathcal{T}_{\text{AT}}(t) + a_2 \mathcal{T}_{\text{AT}}(t)^2$$

where $\mathcal{T}_{AT}(t)$ is the atmospheric temperature, while abatement costs depend on the control rate $\mu(t)$:

$$\Lambda(t) = b_1 \mu(t)^{b_2}$$

[20] The climate damage coefficient $\Omega_D(t) = (1 + D(t))^{-1}$ represents the fraction of GDP lost because of the temperature increase

[21] It includes costs of reduction of greenhouse gases emission, abatement and mitigation costs

Integrated assessment model (IAM)

GHG emissions, concentrations and radiative forcing

- The total emission of green house gases $\mathcal{E}(t)$ is given by:

$$\mathcal{E}(t) = (1 - \mu(t))\,\sigma(t)\,Y(t) + \mathcal{E}_{\mathrm{Land}}(t)$$

  where mitigation policies are translated by the control rate $\mu(t)$, $\mathcal{E}_{\mathrm{Land}}(t)$ represents exogenous land-use emissions and $\sigma(t)$ is the uncontrolled ratio of green house gases emissions to output

- The evolution of the GHG concentration $\mathcal{C} = (\mathcal{C}_{\mathrm{AT}}, \mathcal{C}_{\mathrm{UP}}, \mathcal{C}_{\mathrm{LO}})$ is given by:

$$\mathcal{C}(t) = \Phi_{\mathcal{C},\Delta}\mathcal{C}(t-1) + B_{\mathcal{C},\Delta}\mathcal{E}(t)$$

- The increase of radiative forcing $\mathcal{F}_{\mathrm{RAD}}(t)$ depends on the GHG concentration in the atmosphere:

$$\mathcal{F}_{\mathrm{RAD}}(t) = \eta \ln_2\left(\frac{\mathcal{C}_{\mathrm{AT}}(t)}{\mathcal{C}_{\mathrm{AT}}(1750)}\right) + \mathcal{F}_{\mathrm{EX}}(t)$$

Integrated assessment model (IAM)

Temperatures

Atmospheric and lower ocean temperatures are given by:

$$
\begin{aligned}
C_{\mathrm{AT}} \frac{\mathrm{d}\mathcal{T}_{\mathrm{AT}}(t)}{\mathrm{d}t} &= \mathcal{F}_{\mathrm{RAD}}(t) - \lambda \mathcal{T}_{\mathrm{AT}}(t) - \gamma(\mathcal{T}_{\mathrm{LO}}(t) - \mathcal{T}_{\mathrm{AT}}(t)) \\
C_{\mathrm{LO}} \frac{\mathrm{d}\mathcal{T}_{\mathrm{LO}}(t)}{\mathrm{d}t} &= \gamma(\mathcal{T}_{\mathrm{LO}}(t) - \mathcal{T}_{\mathrm{AT}}(t))
\end{aligned}
$$

where γ is the heat exchange coefficient and λ is the climate feedback parameter.

Integrated assessment model (IAM)

The optimal control problem

Simplified version of the DICE model (Nordhaus, 1993)

$$\{\mu^\star(t), s^\star(t)\} = \arg\max \sum_{t=0}^{T} \frac{u(c(t), L(t))}{(1+\rho)^t}$$

$$\text{s.t.} \quad \begin{cases} Y(t) = A_{\mathrm{TFP}}(t) K(t)^\alpha L(t)^{1-\alpha} \\ A_{\mathrm{TFP}}(t) = (1 + g_A(t)) A_{\mathrm{TFP}}(t-1) \\ K(t) = (1-\delta_K) K(t-1) + I(t) \\ L(t) = (1 + g_L(t)) L(t-1) \\ Q(t) = \Omega_{\mathrm{Climate}}(t) Y(t) \\ C(t) = (1 - s(t)) Q(t) \\ \mathcal{E}(t) = (1 - \mu(t)) \sigma(t) Y(t) + \mathcal{E}_{\mathrm{Land}}(t) \\ \mathcal{C}(t) = \Phi_{\mathcal{C},\Delta} \mathcal{C}(t-1) + B_{\mathcal{C},\Delta} \mathcal{E}(t) \\ \mathcal{F}_{\mathrm{RAD}}(t) = \eta \log_2 \left(\frac{\mathcal{C}_{\mathrm{AT}}(t)}{\mathcal{C}_{\mathrm{AT}}(1750)} \right) + \mathcal{F}_{\mathrm{EX}}(t) \\ \mathcal{T}(t) = \Phi_{\mathcal{T},\Delta} \mathcal{T}(t-1) + B_{\mathcal{T},\Delta} \mathcal{F}_{\mathrm{RAD}}(t) \end{cases}$$

Integrated assessment model (IAM)

Scenario analysis

The process of building scenarios is the same in every model

1. Choice of the structure
   - Optimization or evaluation?
   - Optimization function?
   - Complexity or simplicity?
2. Calibration
   - Choice for the discount rate (Nordhaus vs Stern)
   - Calibration of energy prices and substitution (etc.)
3. Applications
   - Compare baseline scenario of the different models
   - Compute the 2°C scenario, the optimal welfare scenario, etc.

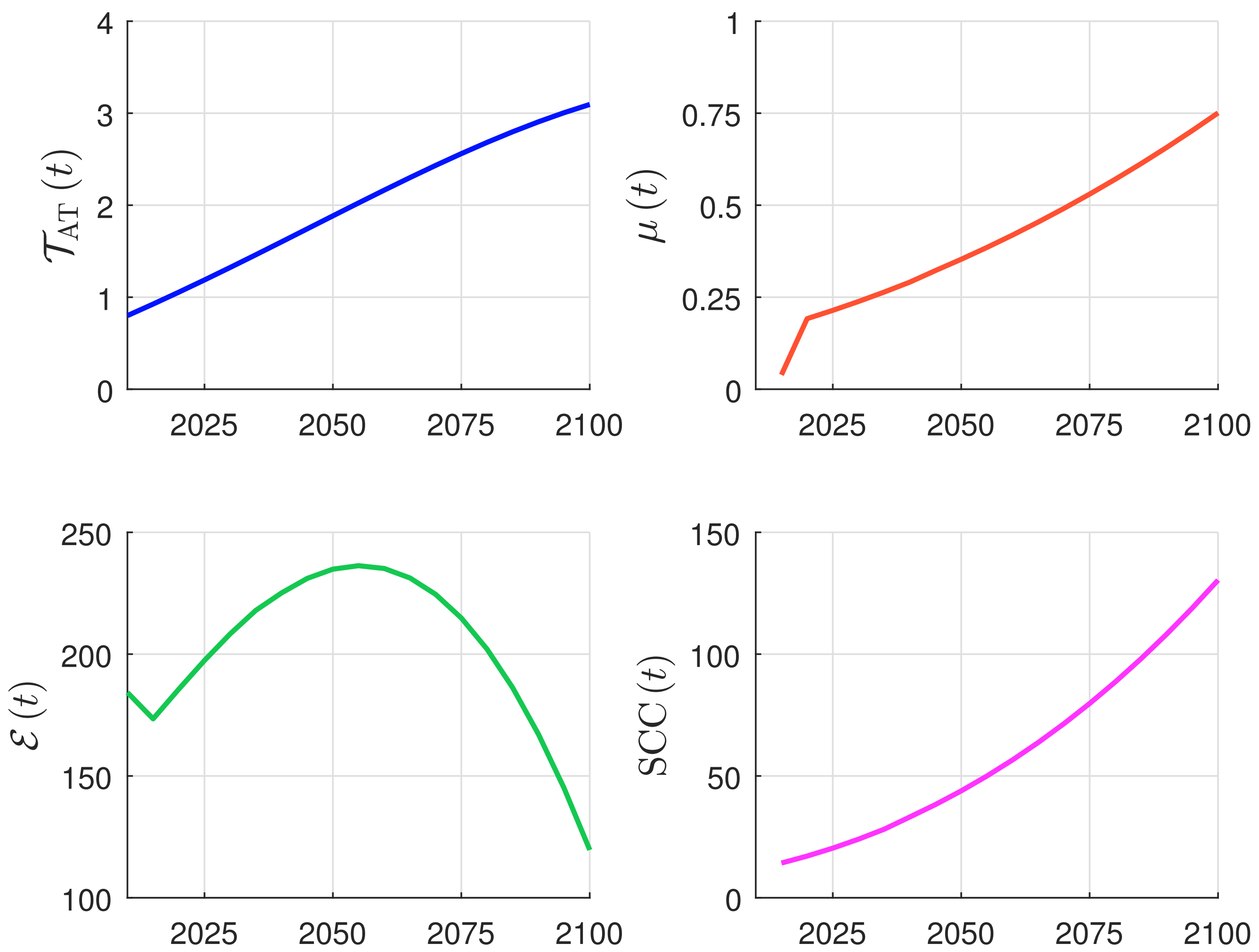
Integrated assessment model (IAM)

Important variables

- $\mathcal{T}_{\mathrm{AT}}(t)$ — Atmospheric temperature
- $\mu(t)$ — Control rate (mitigation policies)
- $\mathcal{E}(t)$ — Total emissions of GHG
- $\mathrm{SCC}(t)$ — Social cost of carbon

Integrated assessment model (IAM)

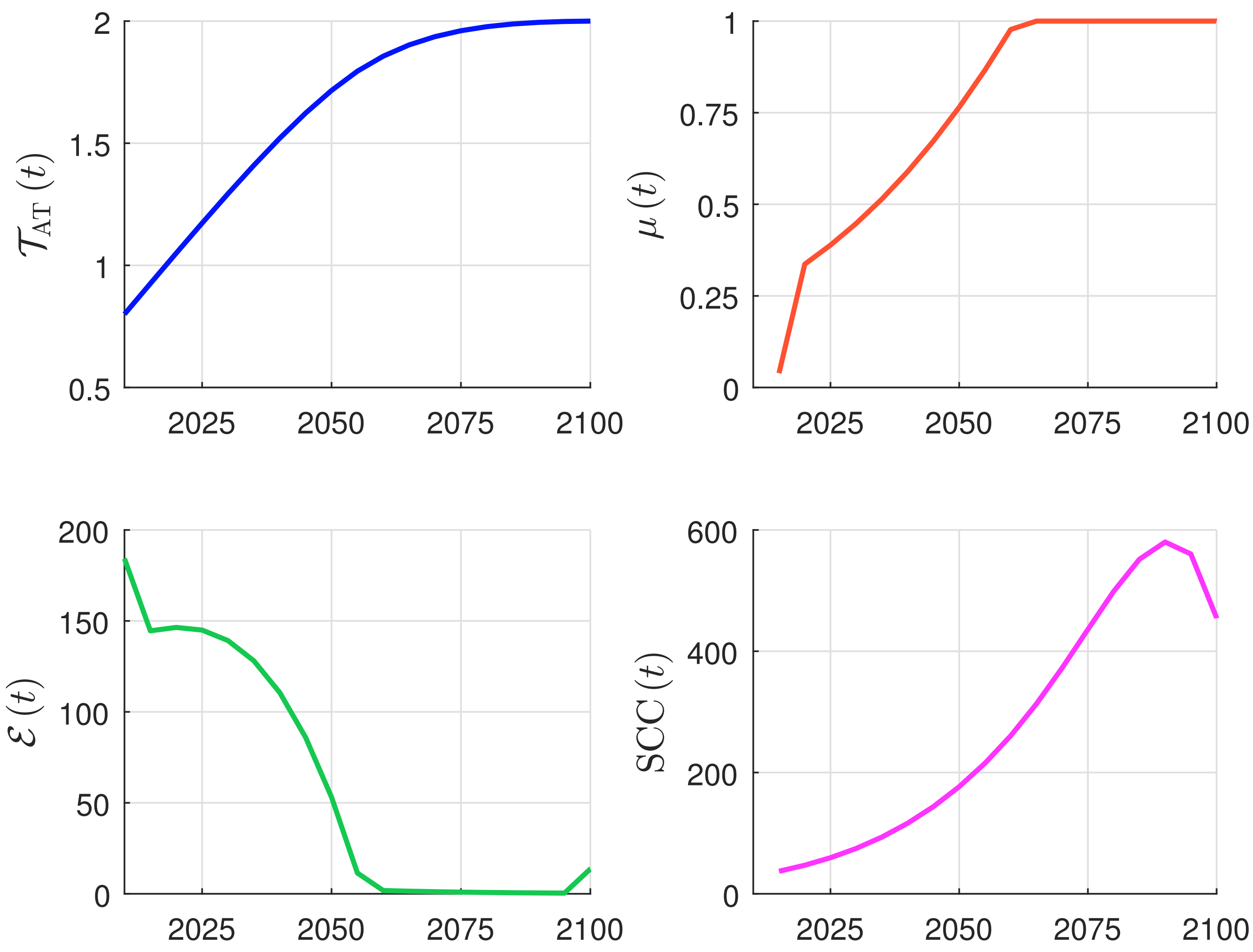
2013 DICE optimal welfare scenario



Source: Le Guenedal (2019)

Integrated assessment model (IAM)

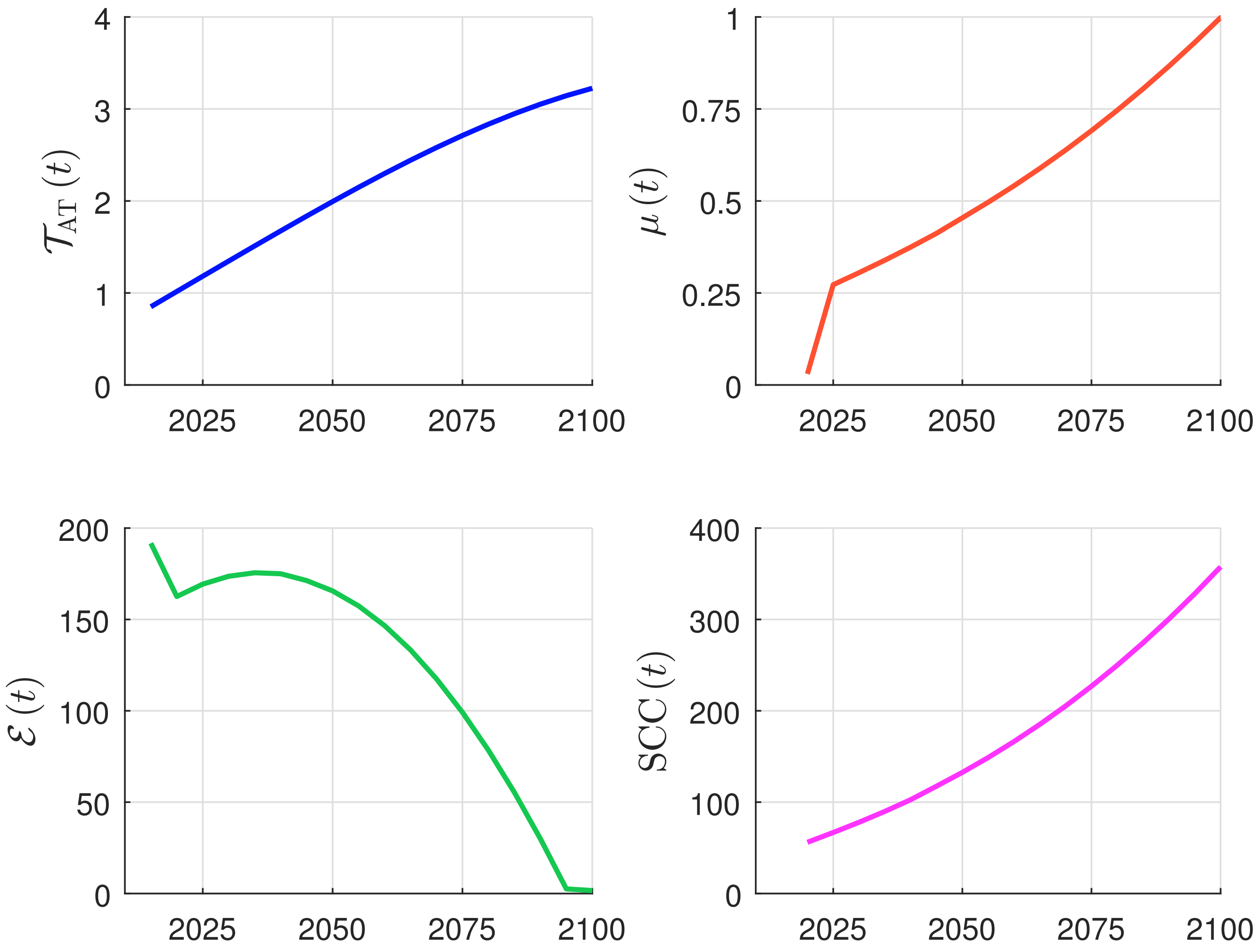
2013 DICE 2°C scenario



Source: Le Guenedal (2019)

Integrated assessment model (IAM)

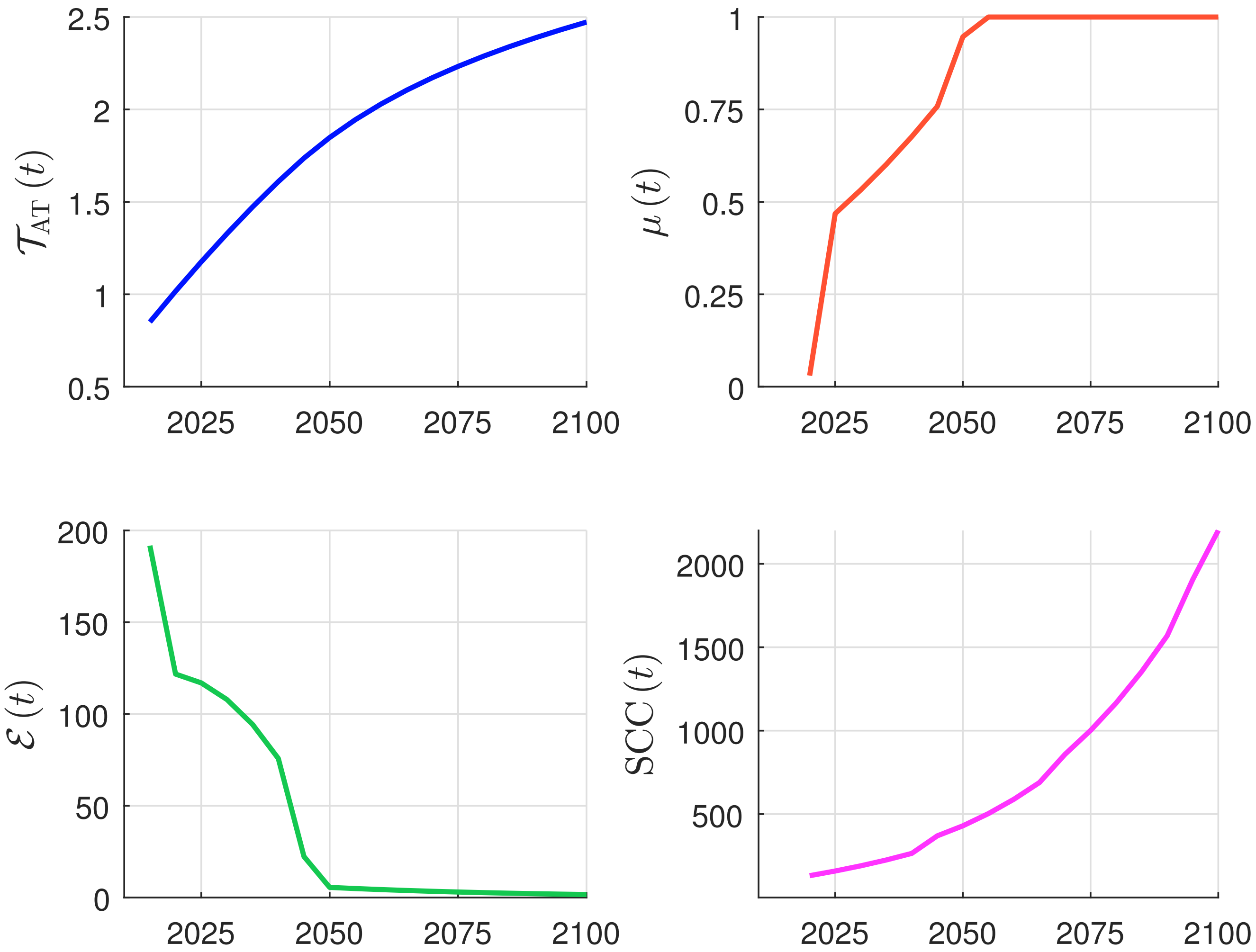
2016 DICE optimal welfare scenario



Source: Le Guenedal (2019)

Integrated assessment model (IAM)

2016 DICE 2°C scenario



Source: Le Guenedal (2019)

Integrated assessment model (IAM)

The tragedy of the horizon

Achieving the 2°C scenario

- In 2013, the DICE model suggested to reduce drastically CO_2 emissions...
- Since 2016, **the 2°C trajectory is no longer feasible!** (minimum ≈ 2.6°C)
- For many models, we now have:

$$\mathbb{P}(\Delta T > 2^{\circ} C) > 95\%$$

Integrated assessment model (IAM)

Malthusianism and climate risk

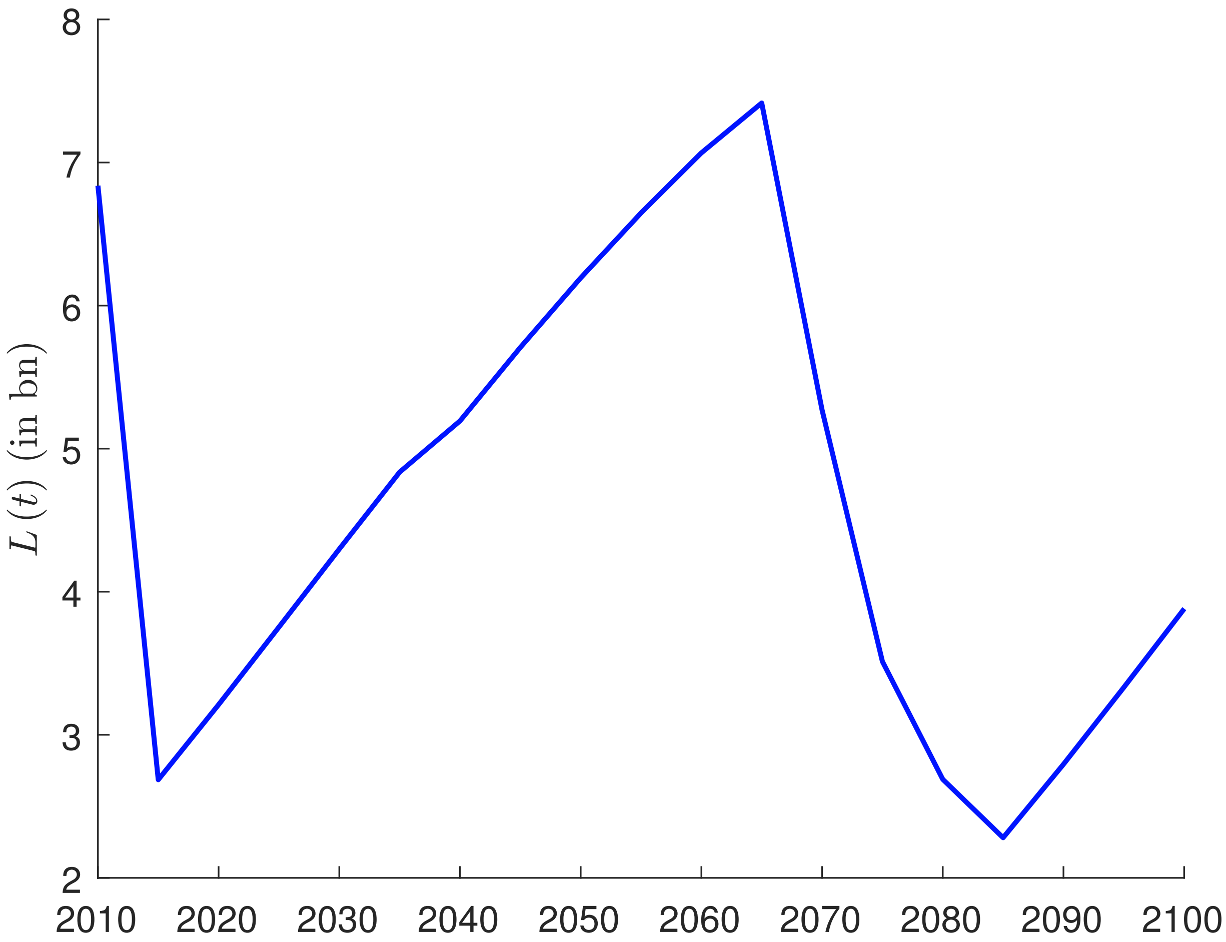


Figure 127: Optimal control on population growth rate (2°C scenario)

Integrated assessment model (IAM)

Social cost of carbon (SCC)

"This concept represents the economic cost caused by an additional ton of carbon dioxide emissions (or more succinctly carbon) or its equivalent. [...] In the language of mathematical programming, the SCC is the shadow price of carbon emissions along a reference path of output, emissions, and climate change" (Nordhaus, 2011).

Mathematical definition

We have:

$$\mathrm{SCC}(t) = \frac{\partial\, W^{\star} / \partial\, \mathcal{E}(t)}{\partial\, W^{\star} / \partial\, C(t)} = \frac{\partial\, C(t)}{\partial\, \mathcal{E}(t)}$$

Integrated assessment model (IAM)

Debate around the social cost of carbon

We have:

- \$266/tCO_2 for Stern (2007)
- \$57/tCO_2 for Golosov *et al.* (2014)
- \$31.2/tCO_2 for Nordhaus (2018) in the case of optimal welfare
- \$229/tCO_2 for Nordhaus (2018) in the case of the 2.5°C scenario
- \$125/tCO_2 for Daniel *et al.* (2018)

Integrated assessment model (IAM)

Limits of IAMs

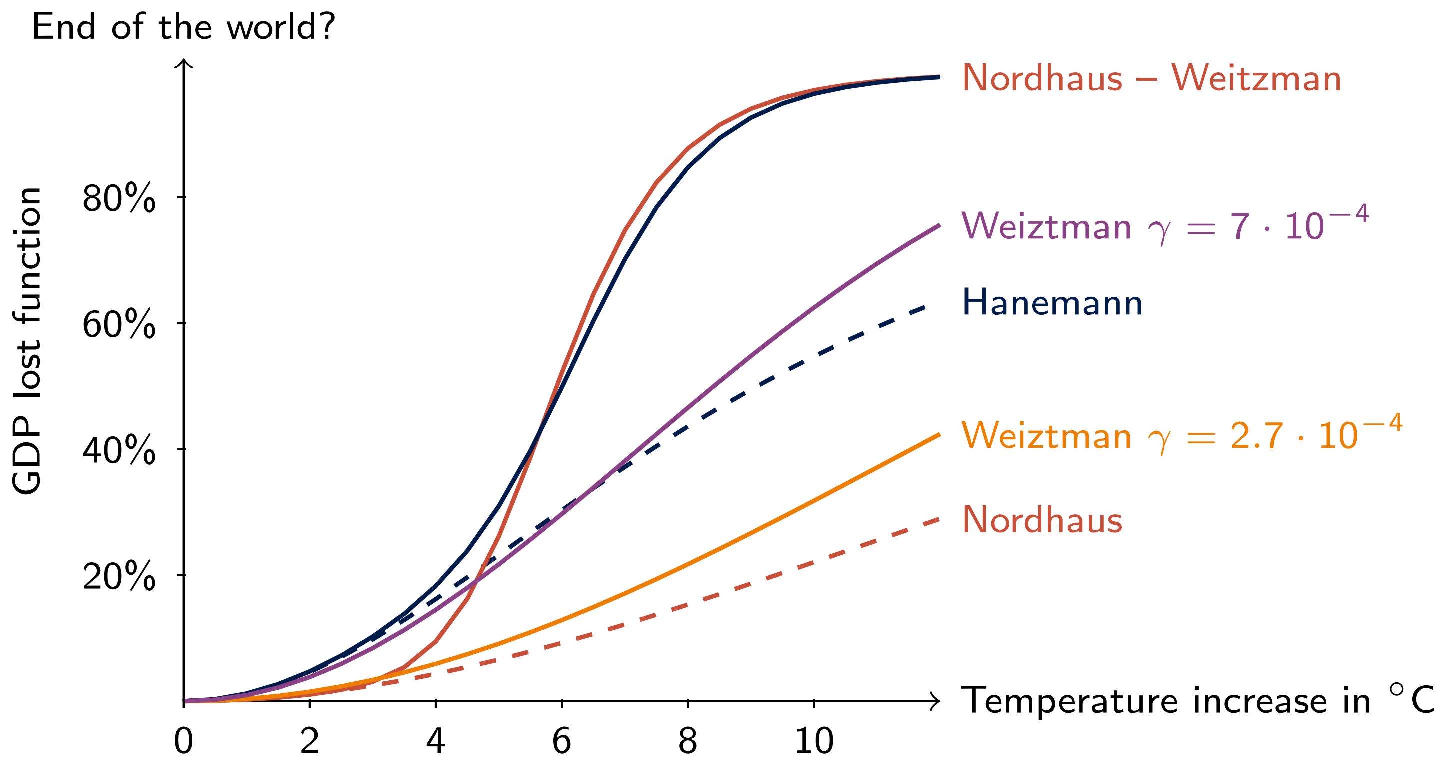


Figure 128: Damage functions

$\Rightarrow$ There is high uncertainty above 2°C and **financial models cannot be based on damage functions**

Integrated assessment model (IAM)

Limits of IAMs

- Financial models do not account for portfolio contribution to the technical change (**adaptation/mitigation**)
- The direct exposure to an **optimal tax** (regulation risk) may be approached by using optimization models of policy makers. However, each model leads to a different carbon price...
- Interconnectedness and systemic risks
- First round losses $\neq$ second round losses
- Stranded assets

Integrated assessment model (IAM)

Limits of IAMs

- AIM
- DICE/RICE
- FUND
- GCAM
- IMACLIM (CIRED)
- IMAGE
- MESSAGE
- MiniCAM
- PAGE
- REMIND
- RESPONSE (CIRED)
- WITCH

Regulation of climate risk

- UN, international bodies & coalitions
- Countries
- Cities
- Industry self-regulation
- Non-governmental organizations (NGO)
- Financial regulators

Hard regulation $\neq$ **soft regulation**

Regulation of climate risk

UN

United Nations Climate Change Conference

- Conference of the Parties (COP)
- Dealing with climate change
- COP 1: Berlin (1995)
- COP 3: Kyoto (1997) $\Rightarrow$ Kyoto Protocol (CMP)
- COP 21: Paris (2015) $\Rightarrow$ Paris Agreement (CMA)
- COP 26: Glasgow (2022)

Regulation of climate risk

UN

The **Kyoto Protocol** is an international treaty that commits state parties to reduce GHG emissions, based on the scientific consensus that:

1. **Global warming is occurring**
2. It is likely that **human-made CO_2 emissions have caused it**

Regulation of climate risk

UN

The **Paris Agreement** is an international treaty with the following goals:

1. Keep a global temperature rise this century well below 2°C above the pre-industrial levels
2. Pursue efforts to limit the temperature increase to 1.5°C
3. Increase the ability of countries to deal with the impacts of climate change
4. Make finance flows consistent with low GHG emissions and climate-resilient pathways

$\Rightarrow$ Nationally determined contributions (NDC)

Regulation of climate risk

UN

Table 78: CO_2 emissions by country

Rank	Country	CO_2 emissions Total (in GT)	Share	CO_2 emissions Per capita (in MT)
1	China	10.06	28%	7.2
2	USA	5.41	15%	15.5
3	India	2.65	7%	1.8
4	Russia	1.71	5%	12.0
5	Japan	1.16	3%	8.9
6	Germany	0.75	2%	8.8
7	Iran	0.72	2%	8.3
8	South Korea	0.72	2%	12.1
9	Saudi Arabia	0.72	2%	17.4
10	Indonesia	0.72	2%	2.2
11	Canada	0.56	2%	15.1
15	Turkey	0.42	1%	4.7
17	United Kingdom	0.37	1%	5.8
19	France	0.33	1%	4.6
17	Italy	0.33	1%	5.3

Source: Earth System Science Data, `https://earth-system-science-data.net`

World Bank Open Data, `https://data.worldbank.org/topic/climate-change`

Regulation of climate risk

UN

Paris Agreement: where we are?

- 194 states have signed the Agreement
- They represent about 80% of GHG emissions
- USA, Iran and Turkey have not signed the Agreement

Regulation of climate risk
UN

Figure 129: Paris Agreement assessments of aviation and shipping

Source: Climate Action Tracker (CAT), https://climateactiontracker.org

Regulation of climate risk

Coalitions

- **The Coalition of Finance Ministers for Climate Action**

  `www.financeministersforclimate.org`

  - Commitment to implement fully the Paris Agreement
  - Santiago Action Plan
  - Helsinki principles (1. align, 2. share, 3. promote, 4. mainstream, 5. mobilize, 6. engage)

Regulation of climate risk

Coalitions

- **One Planet Summit**

  www.oneplanetsummit.fr

- **One Planet Sovereign Wealth Funds (OPSWF)**
  - Funding members: Abu Dhabi Investment Authority (ADIA), Kuwait Investment Authority (KIA), NZ Superannuation Fund (NZSF), Public Invesment Fund (PIF), Qatar Investment Authority (QIA)
  - New members: Bpifrance, CDP Equity, COFIDES, FONSIS, ISIF, KIC, Mubadala IC, NIIF, NIC NBK
- **One Planet Asset Managers**
  - Funding members: Amundi AM, BlackRock, BNP PAM, GSAM, HSBC Global AM, Natixis IM, Northern Trust AM, SSGA
  - New members: AXA IM, Invesco, Legal & General IM, Morgan Stanley IM, PIMCO UBS AM
- **One Planet Private Equity Funds**
  - Members: Ardian, Carlyle Group, Global Infrastructure Partners, Macquarie Infrastructure and Real Assets (MIRA), SoftBank IA

Regulation of climate risk

Countries

The example of France

- August 2015: French Energy Transition for Green Growth Law (or Energy Transition Law)
- Roadmap to mitigate climate change and diversify the energy mix

Regulation of climate risk
Countries

Article 173 of the French Energy Transition Law

- The annual report of listed companies must include:
  - Financial risks related to the effects of climate change
  - The measures adopted by the company to reduce them
  - The consequences of climate change on the company's activities
- New requirements for investors:
  - Disclosure of climate (and ESG) criteria into investment decision making process
  - Disclosure of the contribution to the energy transition and the global warming limitation international objective
  - Reporting on climate change-related risks (including both physical risks and transition risks), and GHG emissions of assets
- Banks and credit providers shall conduct climate stress testing

Regulation of climate risk

Carbon pricing

- Polluter pays principle
  - A carbon price is a cost applied to carbon pollution to encourage polluters to reduce the amount of GHG they emit into the atmosphere
  - Negative externality
- Two instruments of carbon pricing
  1. **Carbon tax**
  2. **Cap-and-trade** (CAT) or emissions trading scheme (ETS)
- Some examples
  1. EU emissions trading system (2005) — `https://ec.europa.eu/clima/policies/ets_en`
  2. New Zealand ETS (2008)
  3. Chinese national carbon trading scheme (2017)

Regulation of climate risk

Carbon pricing

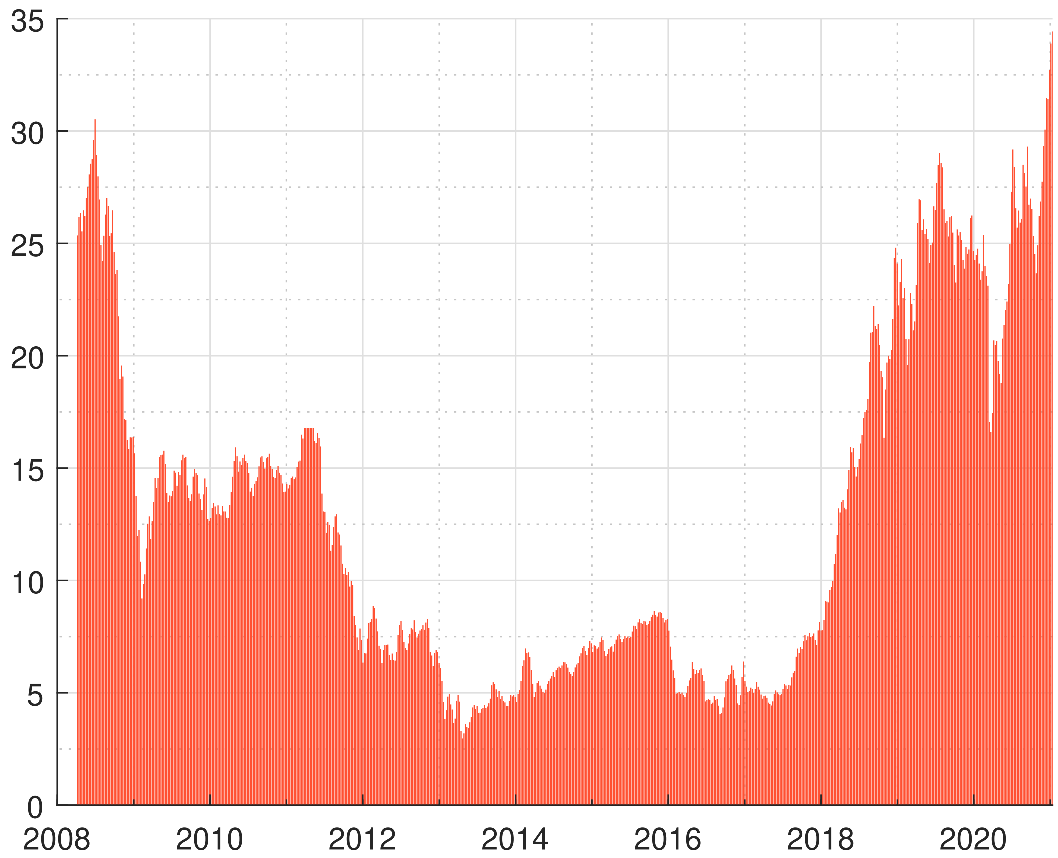


Figure 130: EU ETS carbon price* (in €/tCO_2)

(*)The carbon price reaches 34.43 euros a tonne on Monday 11, 2021

Regulation of climate risk

Carbon pricing

Table 79: Carbon tax (in $/tCO_2$)

Country	2018	2019	2020	Country	2018	2019	2020
Sweden	139.11	126.78	133.26	Latvia	5.58	5.06	10.49
Liechtenstein	100.90	96.46	105.69	South Africa			7.38
Switzerland	100.90	96.46	104.65	France	55.30	50.11	6.98
Finland	76.87	69.66	72.24	Argentina		6.24	5.94
Norway	64.29	59.22	57.14	Chile	5.00	5.00	5.00
Ireland	24.80	22.47	30.30	Colombia	5.67	5.17	4.45
Iceland	35.71	31.34	30.01	Singapore		3.69	3.66
Denmark	28.82	26.39	27.70	Mexico	3.01	2.99	2.79
Portugal	8.49	14.31	27.52	Japan	2.74	2.60	2.76
United Kingdom	25.46	23.59	23.23	Estonia	2.48	2.25	2.33
Slovenia	21.45	19.44	20.16	Ukraine	0.02	0.37	0.35
Spain	24.80	16.85	17.48	Poland	0.09	0.08	0.08

Source: World Bank Carbon Pricing Dashboard, https://carbonpricingdashboard.worldbank.org

Regulation of climate risk
Stranded assets

- Stranded Assets are assets that have suffered from unanticipated or premature write-downs, devaluations or conversion to liabilities
- For example, a 2°C alignment implies to keep a large proportion of existing fossil fuel reserves in the ground (30% of oil reserves, 50% of gas reserves and 80% of coal)
- Risk factors: Regulations, carbon prices, change in demand, social pressure, etc.
- Example of the covid-19 crisis $\Rightarrow$ air travel

Regulation of climate risk

Financial regulation

- Financial Stability Board (FSB)
- European Central Bank (ECB)
- The French Prudential Supervision and Resolution Authority (ACPR)
- The Prudential Regulation Authority (PRA)
- Network for Greening the Financial System (NGFS)
- Etc.

Regulation of climate risk

Financial regulation

Bolton, P., Despres, M., Pereira Da Silva, L.A., Samama, F. and Svartzman, R. (2020), *The Green Swan — Central Banking and Financial Stability in the Age of Climate Change*, BIS Publication, https://www.bis.org/publ/othp31.htm

Regulation of climate risk

Financial regulation

Task Force on Climate-related Financial Disclosures (TCFD)

- Established by the FSB in 2015 to develop a set of voluntary, consistent disclosure recommendations for use by companies in providing information to investors, lenders and insurance underwriters about their climate-related financial risks
- Website: `www.fsb-tcfd.org`
- Chairman: Michael R. Bloomberg (founder of Bloomberg L.P.)
- 31 members
- June 2017: Publication of the "*Recommendations of the Task Force on Climate-related Financial Disclosures*"
- October 2020: Publication of the 2020 "*Status Report: Task Force on Climate-related Financial Disclosures*"

Regulation of climate risk

Financial regulation

Recommendation	ID	Recommended Disclosure
Governance	1	Board oversight
	2	Management's role
Strategy	3	Risks and opportunities
	4	Impact on organization
	5	Resilience of strategy
Risk management	6	Risk ID and assessment processes
	7	Risk management processes
	8	Integration into overall risk management
Metrics and targets	9	Climate-related metrics
	10	Scope 1, 2, 3 GHG emissions
	11	Climate-related targets

Table 80: The 11 recommended disclosures (TCFD, 2017)

Regulation of climate risk

Financial regulation

Some key findings of the 2020 Status Report (TCFD, 2020):

- Disclosure of climate-related financial information has increased since 2017, but continuing progress is needed
- Average level of disclosure across the Task Force's 11 recommended disclosures was 40% for energy companies and 30% for materials and buildings companies
- Asset manager and asset owner reporting to their clients and beneficiaries, respectively, is likely insufficient

Climate stress testing

- ACPR (2020): Climate Risk Analysis and Supervision[22]
- Bank of England (2021): Climate Biennial Exploratory Scenario (June 2021)

Top-down approach $\neq$ bottom-up approach

Stress of risk-weighted asset: Bouchet and Le Guenedal (2020).

[22] https://acpr.banque-france.fr/en/scenarios-and-main-assumptions-acpr-pilot-climate-exercise

Climate capital requirements

Green supporting factor

- Risk weights may depend on the green/brown nature of the credit
- Green loans
- Green supporting factor $\neq$ Brown penalising factor

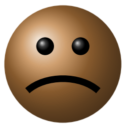
Similar idea: Green Quantitative Easing (GQE)

Climate risk measurement

- Climate risk = risk factor for long-term investors, because of its impacts on asset prices
- Managing climate risk in a portfolio first requires to measure it

Physical risk

- More an operational risk than a business risk
- Measuring physical risk is a difficult task
- Strong impact on real estate & insurance sectors
- Low impact on stock prices?



Transition risk

- A business risk
- Measuring transition risk is a difficult task
- Impact on many sectors (energy, materials, industrials, utilities, etc.)
- High impact on stock prices?

Climate risk measurement

Physical risk and tropical cyclone damage modeling

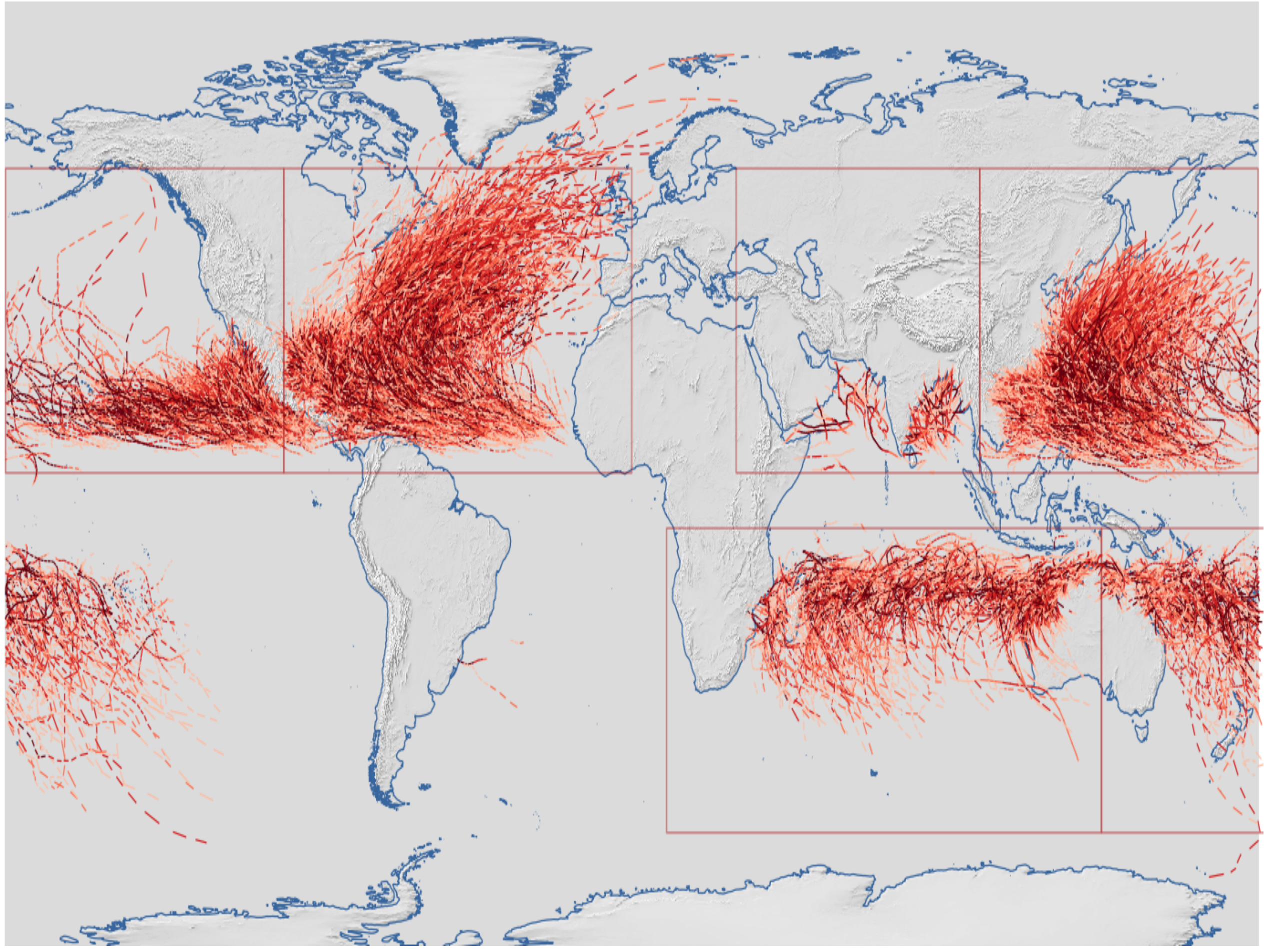


Figure 131: Sample of storms (ERA-5 climate data)

Climate risk measurement

Physical risk and tropical cyclone damage modeling

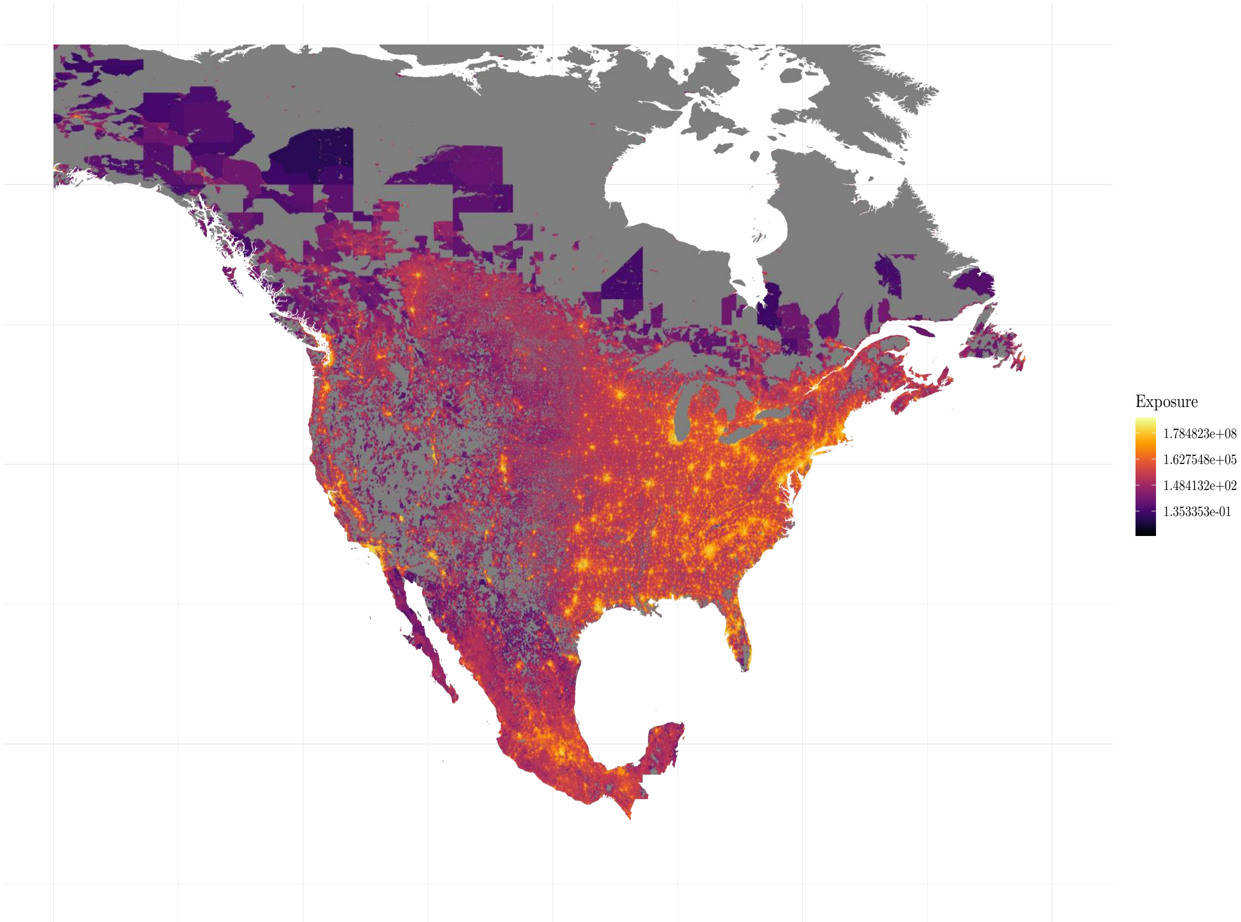
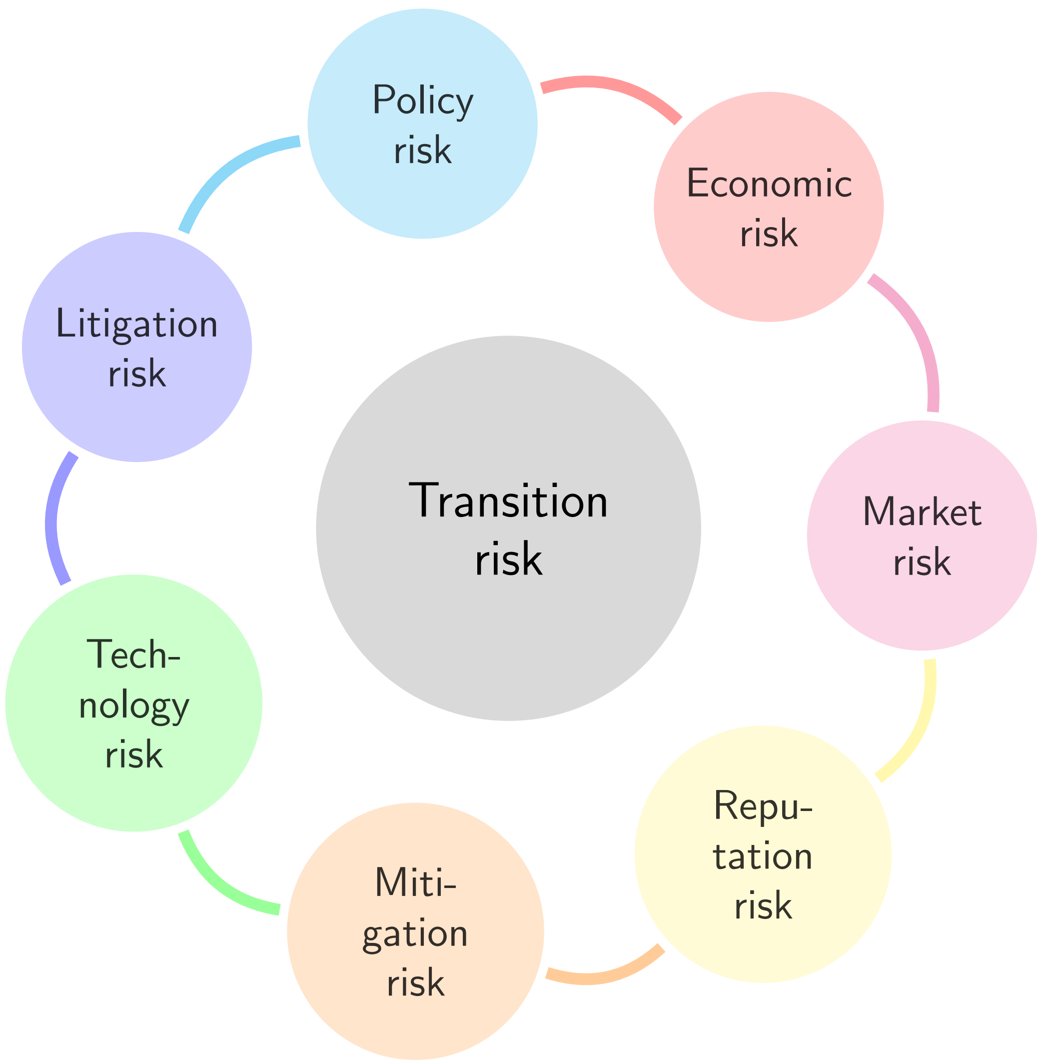


Figure 132: GDP decomposition of North America (or physical asset values) (Litpop database)

Climate risk measurement

Transition risk



Carbon risk measurement

Main assumption

Transition risk can be measured (or approximated) by carbon risk

Carbon risk can be measured by **current** carbon emissions

Carbon risk measurement

The GHG Protocol corporate standard classifies a company's greenhouse gas emissions in three scopes:

- **Scope 1**: direct GHG emissions from all direct GHG emissions by the company
- **Scope 2**: indirect GHG emissions from the consumption of purchased energy (electricity, heat, steam, etc.)
- **Scope 3**: other indirect GHG emissions (not included in Scope 2) that occur in the value chain of the reporting company, including both upstream and downstream emissions (extraction and production of purchased materials and fuels, transport-related activities in vehicles not owned or controlled by the reporting entity, electricity-related activities not covered in Scope 2, outsourced activities, waste disposal, etc.)

Carbon risk measurement

Remark

Scopes 1 and 2 are mandatory to report, whereas scope 3 is voluntary (and harder to measure and monitor)

Carbon risk measurement

- **Carbon intensity** is the amount of GHG emissions per unit of another variable such as gross domestic product (sovereign) or revenue (corporate):

$$\text{Carbon intensity} = \frac{\text{Carbon scope}}{\text{Revenue}}$$

- Carbon scopes are measured in tCO_2e
- Carbon intensities are measured in $tCO_2e/\$$ (or $tCO_2e/\$$ mn)

Carbon footprint $\approx$ Carbon scope

Carbon footprint $\approx$ Carbon intensity

Carbon risk measurement

How to find data of carbon emission and intensity?

- **Carbon Disclosure Project (CDP)** is a not-for-profit charity that runs the global disclosure system for investors, companies, cities, states and regions to manage their environmental impacts
  `https://www.cdp.net`
- **Trucost** was established to provide the data, tools and insights needed by companies, investors and policy makers to deliver the transition to a low carbon, resource efficient economy[23]
  `https://www.trucost.com`
- ESG rating agencies: ISS ESG, MSCI, Sustainalytics, Thomson Reuters, etc.

[23] Trucost is now part of S&P Global

Carbon risk measurement

Table 81: Some carbon variables of the Trucost database

Company	Carbon-Direct[1]
Financial Year	Carbon-First Tier Indirect[1]
Trucost Sector Name	Carbon-Direct + First Tier Indirect[1]
Trucost Sector	Carbon Intensity-Direct[2]
Country	Carbon Intensity-First Tier Indirect[2]
Carbon-Scope 1[1]	Carbon Intensity-Direct + First Tier Indirect[2]
Carbon-Scope 2[1]	GHG-Direct ($ mn)
Carbon-Scope 3[1]	GHG-Indirect ($ mn)
Carbon Intensity-Scope 1[2]	GHG-Total ($ mn)
Carbon Intensity-Scope 2[2]	GHG-Direct Impact Ratio (%)
Carbon Intensity-Scope 3[2]	GHG-Indirect Impact Ratio (%)
Carbon Disclosure	GHG-Total Impact Ratio (%)
Carbon-Weighted Disclosure (%)	Revenue ($ mn)

Source: Trucost Database (2021).

(1) in t CO_2e

(1) in t CO_2e/$ mn

Carbon risk measurement

Table 82: Examples of carbon data (2019)

Company	Carbon emissions (tCO_2e)			Carbon Intensity (tCO_2e/$ mn)			Carbon
	Scope 1	Scope 2	Scope 3	Scope 1	Scope 2	Scope 3	Disclosure
Apple Inc.	50 463	862 127	27 618 943	0.194	3.314	106.156	CDP
Microsoft Corporation	113 414	3 556 553	5 977 488	0.901	28.262	47.500	CDP
Danone SA	722 122	944 877	28 969 780	25.509	33.378	1 023.365	CDP
Nestle SA	3 291 303	3 206 495	61 262 078	35.332	34.422	657.647	CDP
Sanofi	559 422	417 689	3 470 724	13.833	10.328	85.819	CDP
Pfizer Inc.	715 631	762 286	4 669 554	13.829	14.730	90.233	CDP
LVMH-Moet Vuitton	67 613	262 609	11 853 749	1.125	4.371	197.291	CDP
L'Oreal	49 511	160 393	5 556 670	1.480	4.796	166.154	CDP
BP p.l.c.	49 199 999	5 200 000	103 840 194	177.714	18.783	375.077	Env./CSR
TOTAL SE	40 909 129	3 596 127	49 893 263	204.097	17.941	248.920	CDP
Tesla Inc.	327 159	273 116	6 471 521	13.311	11.112	263.305	Estimated
Volkswagen AG	4 494 066	5 973 894	65 335 372	15.890	21.123	231.016	CDP

Source: Trucost Database (2021).

In 2019, there are 12 989 companies in the Trucost data.

Carbon risk measurement

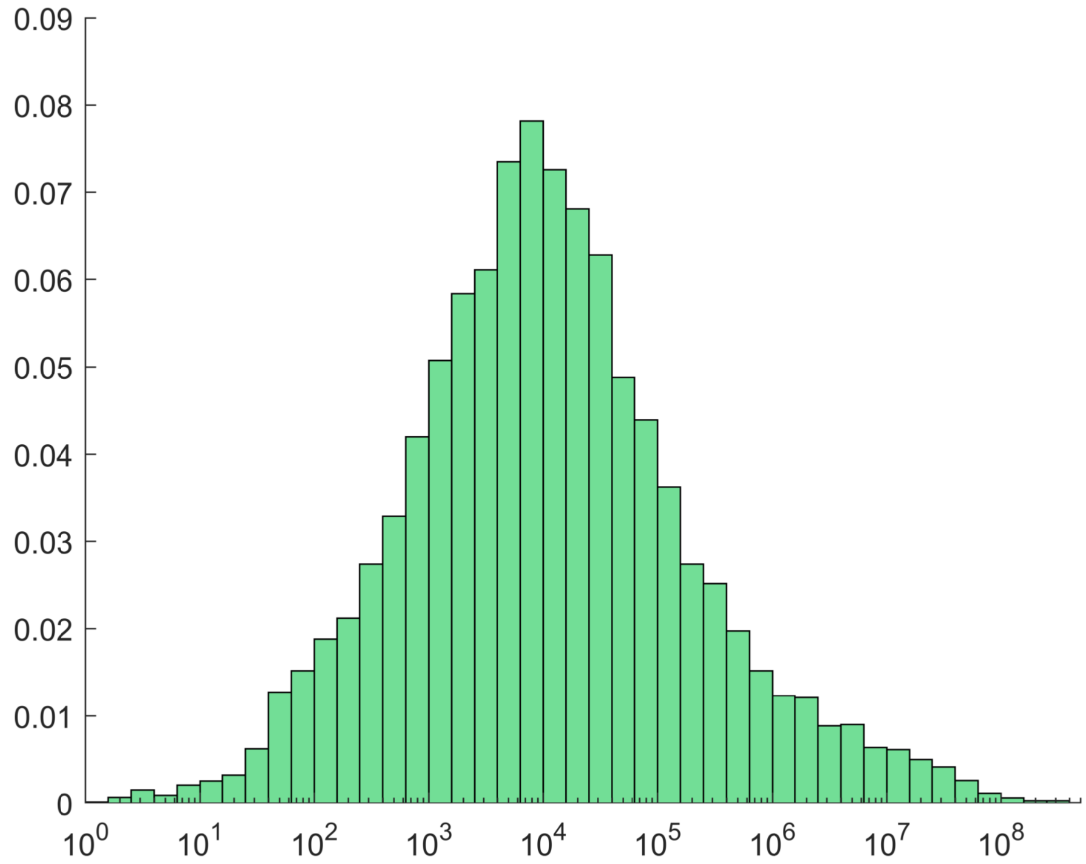


Figure 133: Histogram of carbon emissions (Scope 1, tCO_2e)

Source: Trucost Database (2021) & author's calculations.

Carbon risk measurement

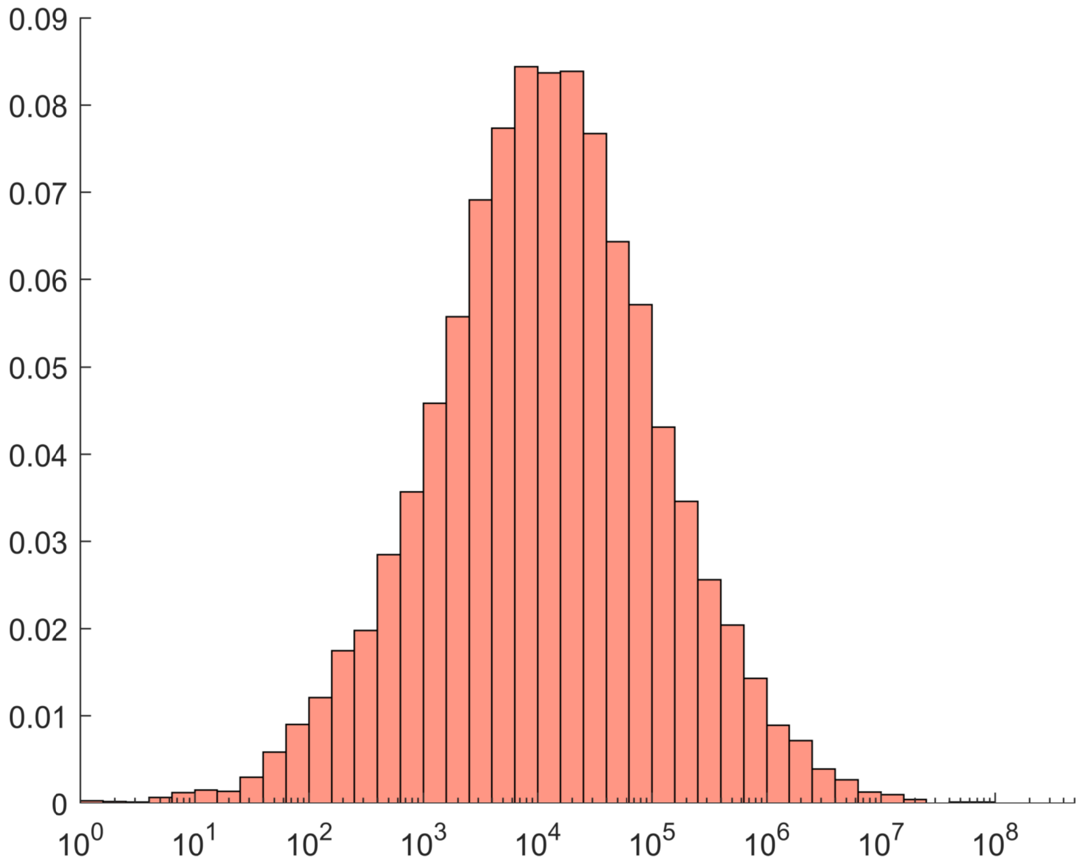


Figure 134: Histogram of carbon emissions (Scope 2, tCO_2e)

Source: Trucost Database (2021) & author's calculations.

Carbon risk measurement

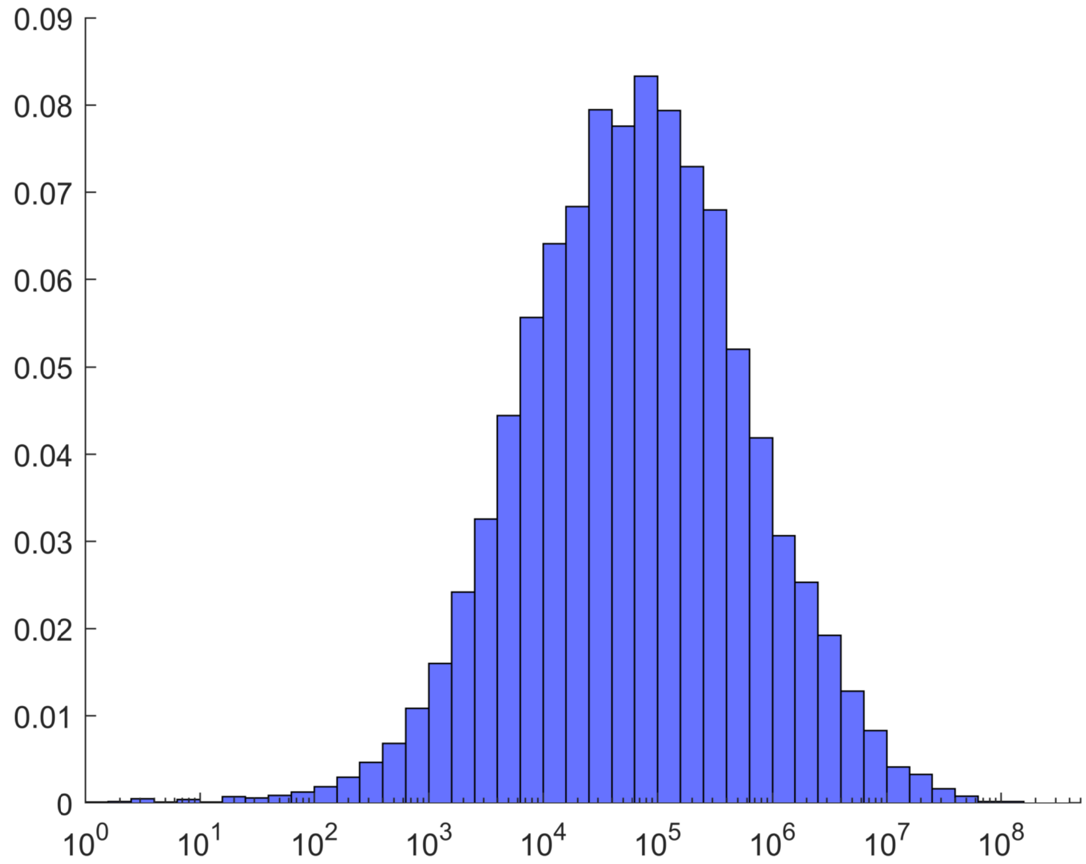


Figure 135: Histogram of carbon emissions (Scope 3, tCO_2e)

Source: Trucost Database (2021) & author's calculations.

Carbon risk measurement

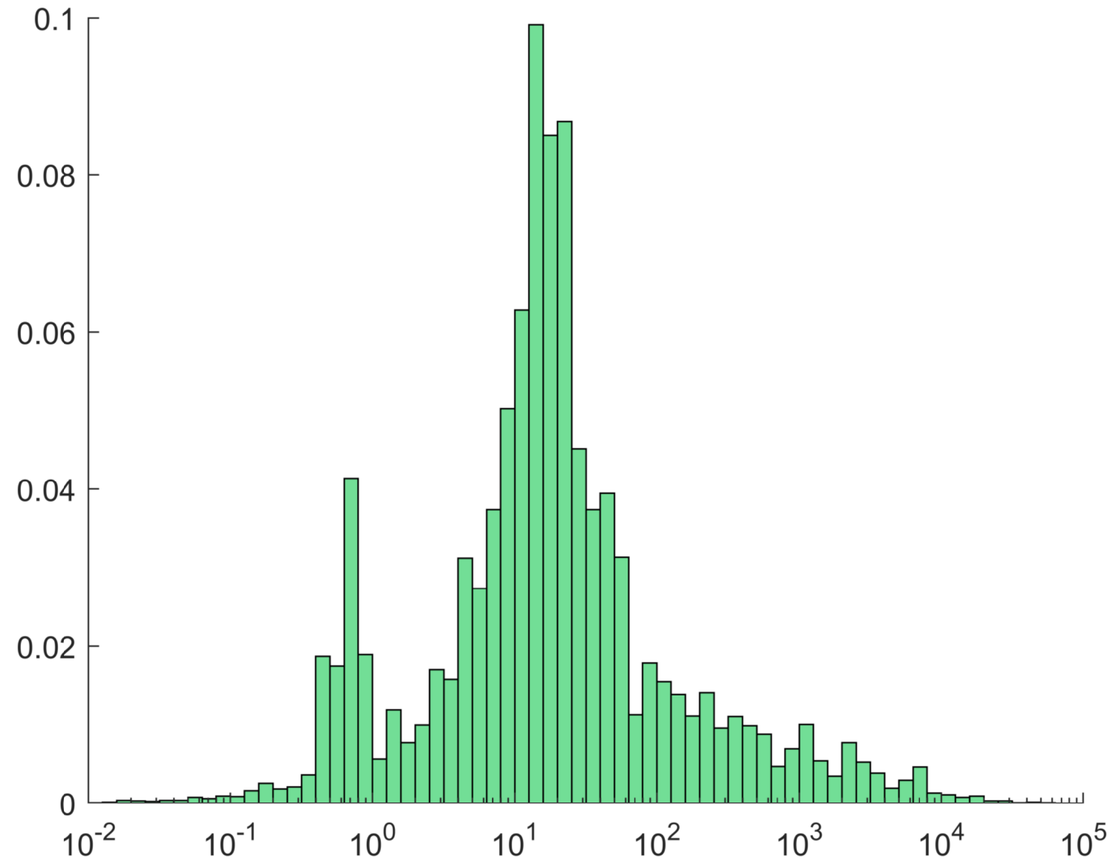


Figure 136: Histogram of carbon intensity (Scope 1, tCO_2e/$ mn)

Source: Trucost Database (2021) & author's calculations.

Carbon risk measurement

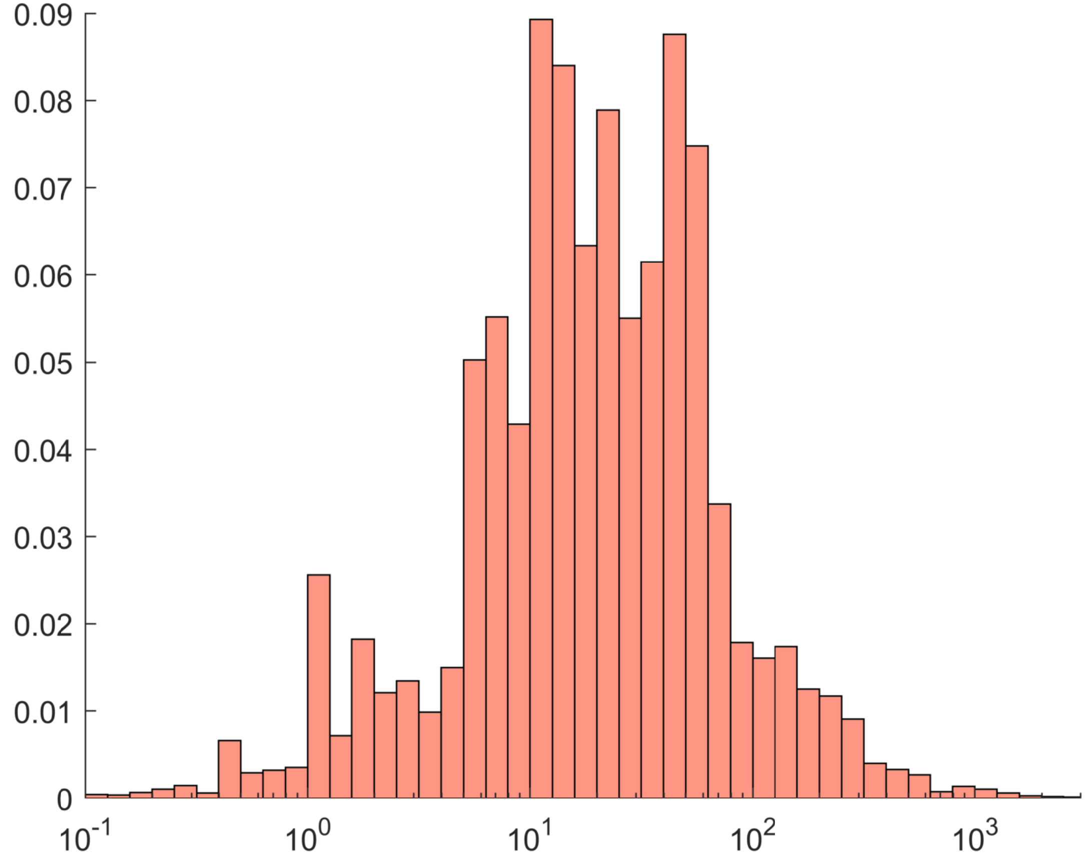


Figure 137: Histogram of carbon intensity (Scope 2, tCO_2e/$ mn)

Source: Trucost Database (2021) & author's calculations.

Carbon risk measurement

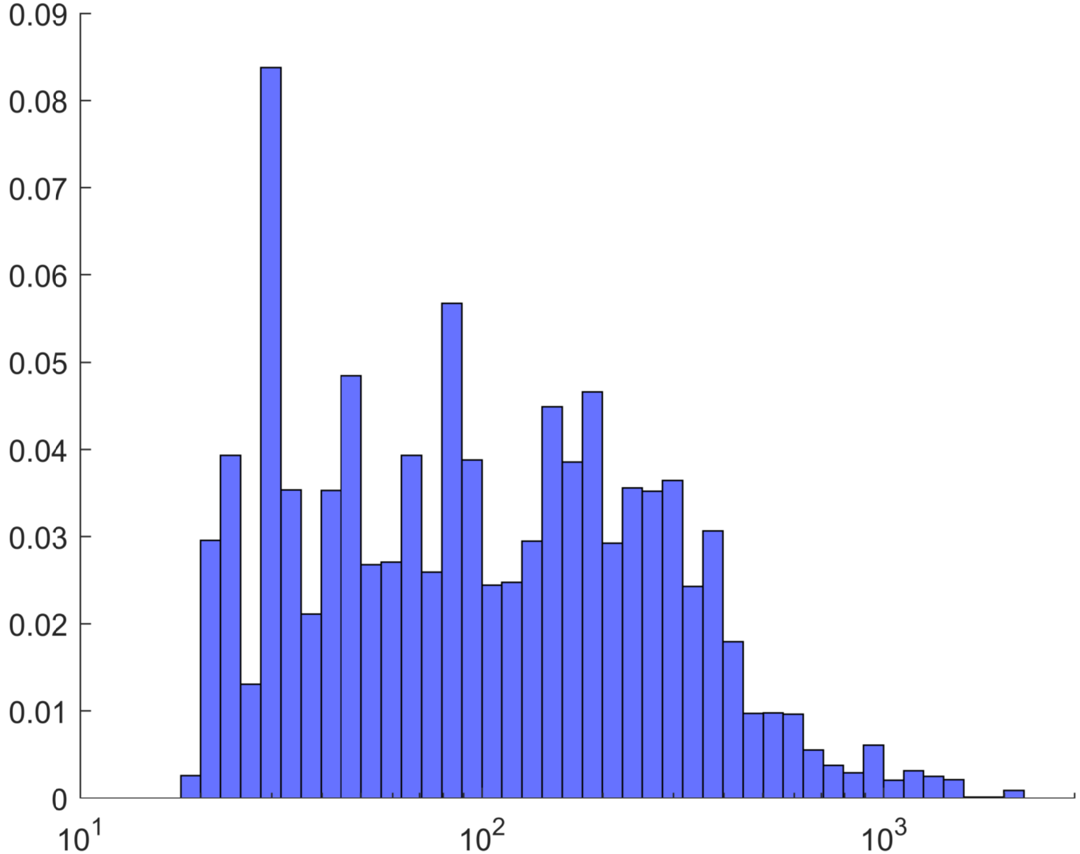


Figure 138: Histogram of carbon intensity (Scope 3, tCO_2e/$ mn)

Source: Trucost Database (2021) & author's calculations.

Portfolio optimization with a benchmark

The γ-optimization problem is:

$$
\begin{aligned}
x^{\star} & = \arg\min \frac{1}{2}\sigma^{2}(x \mid b) - \gamma x^{\top}\mu(x \mid b) \\
\text{u.c.} & \begin{cases} \mathbf{1}_n^{\top} x = 1 \\ \mathbf{0}_n \leq x \leq \mathbf{1}_n & \text{(no short selling)} \\ x \in \Omega \end{cases}
\end{aligned}
$$

where $x = (x_1, \ldots, x_n)$ is the portfolio, $b = (b_1, \ldots, b_n)$ is the benchmark, $\sigma(x \mid b) = \sqrt{(x - b)^{\top} \Sigma (x - b)}$ is the volatility of the tracking error, $\mu(x \mid b) = (x - b)^{\top} \mu$ is the expected excess return and $x \in \Omega$ corresponds to additional constraints

Remark

We remind that the objective function can be cast into a QP problem:

$$
x^{\star} = \arg\min \frac{1}{2} x^{\top} \Sigma x - x^{\top} (\gamma\mu + \Sigma b)
$$

Quadratic programming problem

Reminder (Lecture 1)

The formulation of a standard QP problem is:

$$
\begin{aligned}
x^{\star} & = & \arg\min \frac{1}{2} x^{\top} Q x - x^{\top} R \\
& & \text{u.c.} \quad \left\{ \begin{array}{l} Ax = B \\ Cx \leq D \\ x^{-} \leq x \leq x^{+} \end{array} \right.
\end{aligned}
$$

Portfolio decarbonization

We note $\mathcal{CI}_i$ the carbon intensity[24] associated to asset i

- The carbon intensity of the benchmark is equal to:

$$\mathcal{CI}(b) = \sum_{i=1}^{n} b_i \cdot \mathcal{CI}_i = b^\top \mathcal{CI}$$

  where $\mathcal{CI} = (\mathcal{CI}_1, \ldots, \mathcal{CI}_n)$ is the vector of carbon intensities

- The carbon intensity of the portfolio is equal to:

$$\mathcal{CI}(x) = x^\top \mathcal{CI}$$

  $\mathcal{CI}(x)$ is also called the weighted average carbon intensity (WACI)

- The objective is to reduce the carbon intensity of the benchmark by a factor $\pi_{\mathcal{CI}}$:

$$\mathcal{CI}(x) \leq \mathcal{CI}^\star = \pi_{\mathcal{CI}} \cdot \mathcal{CI}(b)$$

[24] It corresponds to the carbon intensity of the company i

Portfolio decarbonization

- We deduce that the optimization problem is:

$$
\begin{aligned}
x^{\star} & = \frac{1}{2}\sigma^{2}\left(x \mid b\right) \\
\text{u.c.} & \left\{
\begin{array}{l}
\mathbf{1}_n^{\top} x = 1 \\
\mathbf{0}_n \leq x \leq \mathbf{1}_n \\
\mathcal{CI}\left(x\right) \leq \pi_{\mathcal{CI}} \cdot \mathcal{CI}\left(b\right)
\end{array}
\right.
\end{aligned}
$$

- The underlying idea is to obtain a decarbonized portfolio $x^{\star}$ such that the tracking error with respect to the benchmark b is the lowest
- The benchmark b can be a current portfolio (active management) or an index portfolio (passive management)

Portfolio decarbonization

- Since the constraint on the carbon intensity is equivalent to:

$$\mathcal{CI}^\top x \leq \pi_{\mathcal{CI}} \cdot \left(b^\top \mathcal{CI}\right)$$

We obtain the following QP problem:

$$\begin{aligned} x^\star &= \frac{1}{2} x^\top \Sigma x - x^\top \Sigma b \\ &\text{u.c.} \quad \left\{ \begin{array}{l} \mathbf{1}_n^\top x = 1 \\ \mathcal{CI}^\top x \leq \pi_{\mathcal{CI}} \cdot \left(b^\top \mathcal{CI}\right) \\ \mathbf{0}_n \leq x \leq \mathbf{1}_n \end{array} \right. \end{aligned}$$

Portfolio decarbonization

- We have the following QP correspondences:

$$\begin{aligned} Q &= \Sigma \\ R &= \Sigma b \\ A &= \mathbf{1}_n^\top \\ B &= 1 \\ C &= \mathcal{CI}^\top \\ D &= \mathcal{CI}^\star = \pi_{\mathcal{CI}} \cdot \left(b^\top \mathcal{CI} \right) \\ x^- &= \mathbf{0}_n \\ x^+ &= \mathbf{1}_n \end{aligned}$$

Portfolio decarbonization

Example 1

We consider a capitalization-weighted equity index, which is composed of 8 stocks. The weights are equal to 23%, 19%, 17%, 13%, 9%, 8%, 6% and 5%. We assume that their volatilities are equal to 22%, 20%, 25%, 18%, 35%, 23%, 13% and 29%. The correlation matrix is given by:

$$\rho = \begin{pmatrix} 100\% & & & & & & & \\ 80\% & 100\% & & & & & & \\ 70\% & 75\% & 100\% & & & & & \\ 60\% & 65\% & 80\% & 100\% & & & & \\ 70\% & 50\% & 70\% & 85\% & 100\% & & & \\ 50\% & 60\% & 70\% & 80\% & 60\% & 100\% & & \\ 70\% & 50\% & 70\% & 75\% & 80\% & 50\% & 100\% & \\ 60\% & 65\% & 70\% & 75\% & 65\% & 70\% & 80\% & 100\% \end{pmatrix}$$

The carbon intensities (expressed in tCO_2e/\$ mn) are respectively equal to: 100.5, 57.2, 250.4, 352.3, 27.1, 54.2, 78.6 and 426.7.

Portfolio decarbonization

Table 83: Optimal decarbonization portfolios (max-threshold approach)

$\pi_{\mathcal{CI}}$	1.00	0.90	0.80	0.70	0.60	0.50
$x_1^\star$	23.00	20.98	18.97	16.95	14.91	11.96
$x_2^\star$	19.00	21.15	23.30	25.46	28.25	33.40
$x_3^\star$	17.00	16.79	16.59	16.38	14.79	9.05
$x_4^\star$	13.00	9.12	5.24	1.36	0.00	0.00
$x_5^\star$	9.00	10.33	11.67	13.00	14.51	16.92
$x_6^\star$	8.00	9.18	10.37	11.55	12.63	13.59
$x_7^\star$	6.00	8.20	10.40	12.59	14.21	15.06
$x_8^\star$	5.00	4.23	3.47	2.70	0.70	0.00
$\sigma(x^\star)$ (in bps)	0.00	19.32	38.64	57.96	84.74	141.97
$\mathcal{CI}(x)$	155.18	139.66	124.14	108.62	93.11	77.59

- The carbon intensity of the index is equal to 155.18 tCO_2/\$ mn
- The tracking error of the portfolio is equal to 141.97 bps if we target a 50% reduction of the carbon intensity

Portfolio decarbonization

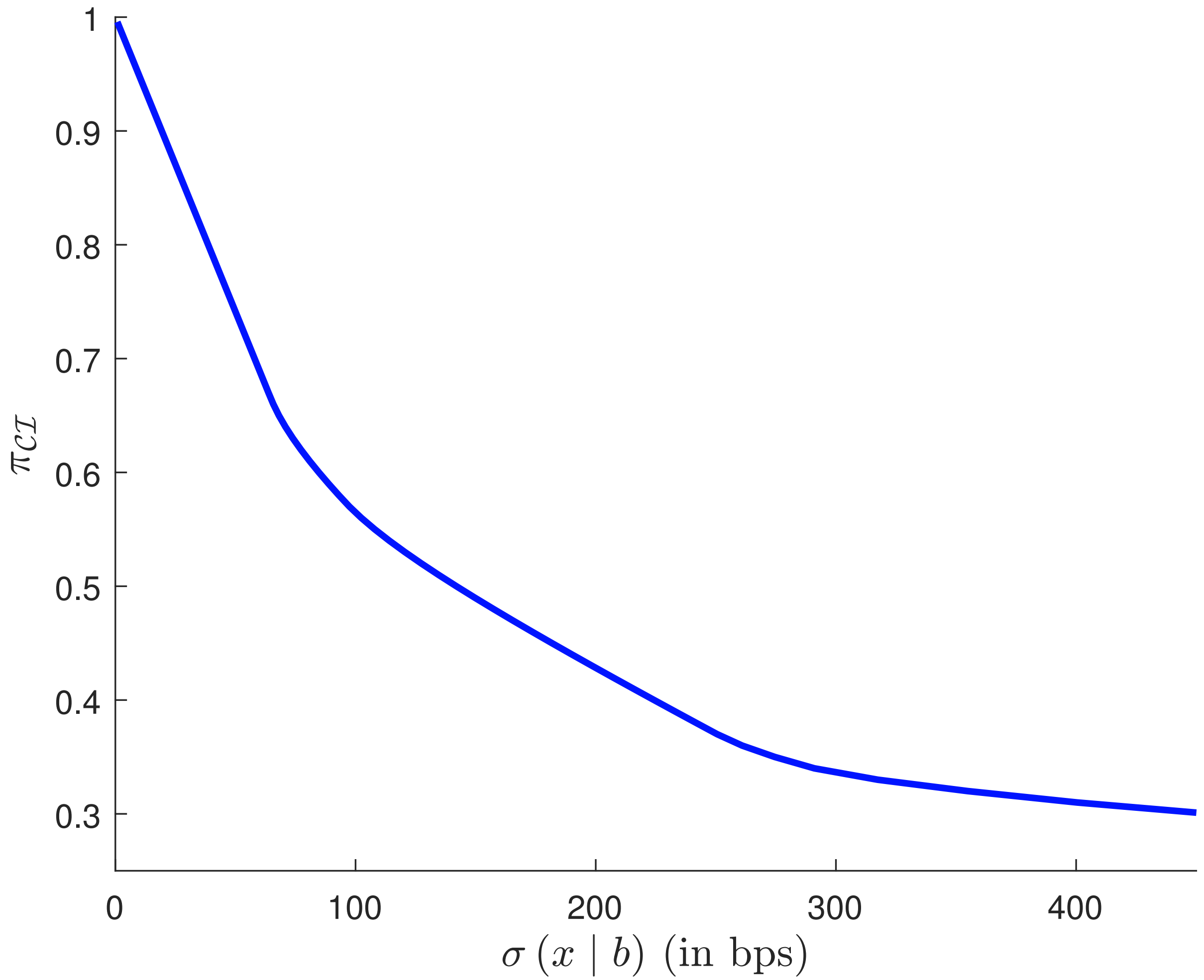


Figure 139: The efficient frontier of optimal decarbonization portfolios

Portfolio decarbonization

Andersson *et al.* (2016) propose a second approach of portfolio decarbonization by eliminating the m worst performer assets in terms of carbon intensity

- We note $\mathcal{CI}_{i:n}$ the order statistics of $\mathcal{CI} = (\mathcal{CI}_1, \ldots, \mathcal{CI}_n)$:

  $$\min \mathcal{CI}_i = \mathcal{CI}_{1:n} \leq \mathcal{CI}_{2:n} \leq \cdots \leq \mathcal{CI}_{i:n} \leq \cdots \leq \mathcal{CI}_{n-1:n} \leq \mathcal{CI}_{n:n} = \max \mathcal{CI}_i$$

- The carbon intensity threshold $\mathcal{CI}^{(m,n)}$ is defined as:

  $$\mathcal{CI}^{(m,n)} = \mathcal{CI}_{n-m+1:n}$$

  where $\mathcal{CI}_{n-m+1:n}$ is the $(n-m+1)$-th order statistic of $\mathcal{CI}$

- Eliminating the m worst performer assets is equivalent to:

  $$\mathcal{CI}_i \geq \mathcal{CI}^{(m,n)} \Rightarrow x_i = 0$$

Portfolio decarbonization

- The optimization problem becomes:

$$x^{\star} = \frac{1}{2}x^{\top}\Sigma x - x^{\top}\Sigma b$$

$$\text{u.c.} \quad \left\{ \begin{array}{l} \mathbf{1}_n^{\top} x = 1 \\ x_i \in \left\{ \begin{array}{ll} [0,1] & \text{if } \mathcal{CI}_i < \mathcal{CI}^{(m,n)} \\ \{0\} & \text{if } \mathcal{CI}_i \geq \mathcal{CI}^{(m,n)} \end{array} \right. \end{array} \right.$$

- The last constraint can be written as:

$$\mathbf{0}_n \leq x \leq x^{+}$$

where:

$$x_i^{+} = \mathbb{1}\left\{\mathcal{CI}_i < \mathcal{CI}^{(m,n)}\right\}$$

We obtain again a QP problem

Portfolio decarbonization

Table 84: Optimal decarbonization portfolios (order-statistic approach)

m	0	1	2	3	4	5	6	7	$\mathcal{CI}$
$x_1^\star$	23.00	18.68	15.94	14.00	**0.00**	0.00	0.00	0.00	100.5
$x_2^\star$	19.00	23.54	26.26	35.84	45.65	56.44	**0.00**	0.00	57.2
$x_3^\star$	17.00	17.46	17.50	**0.00**	0.00	0.00	0.00	0.00	250.4
$x_4^\star$	13.00	6.50	**0.00**	0.00	0.00	0.00	0.00	0.00	352.3
$x_5^\star$	9.00	11.88	13.63	17.98	21.18	26.14	34.73	100.00	27.1
$x_6^\star$	8.00	10.85	12.44	15.84	13.20	17.42	65.27	**0.00**	54.2
$x_7^\star$	6.00	11.11	14.23	16.34	19.98	**0.00**	0.00	0.00	78.6
$x_8^\star$	5.00	**0.00**	0.00	0.00	0.00	0.00	0.00	0.00	426.7
$\sigma(x^\star)$ (in bps)	0.00	77.78	84.51	240.71	278.40	400.71	11.4%	21.6%	
$\mathcal{CI}(x)$	155.18	116.66	96.48	60.87	54.70	48.81	44.79	27.10	

- The reduction of carbon intensity is equal to 24.82% if we eliminate the worst performer
- In this case, we obtain a tracking error of 77.78 bps

Portfolio decarbonization

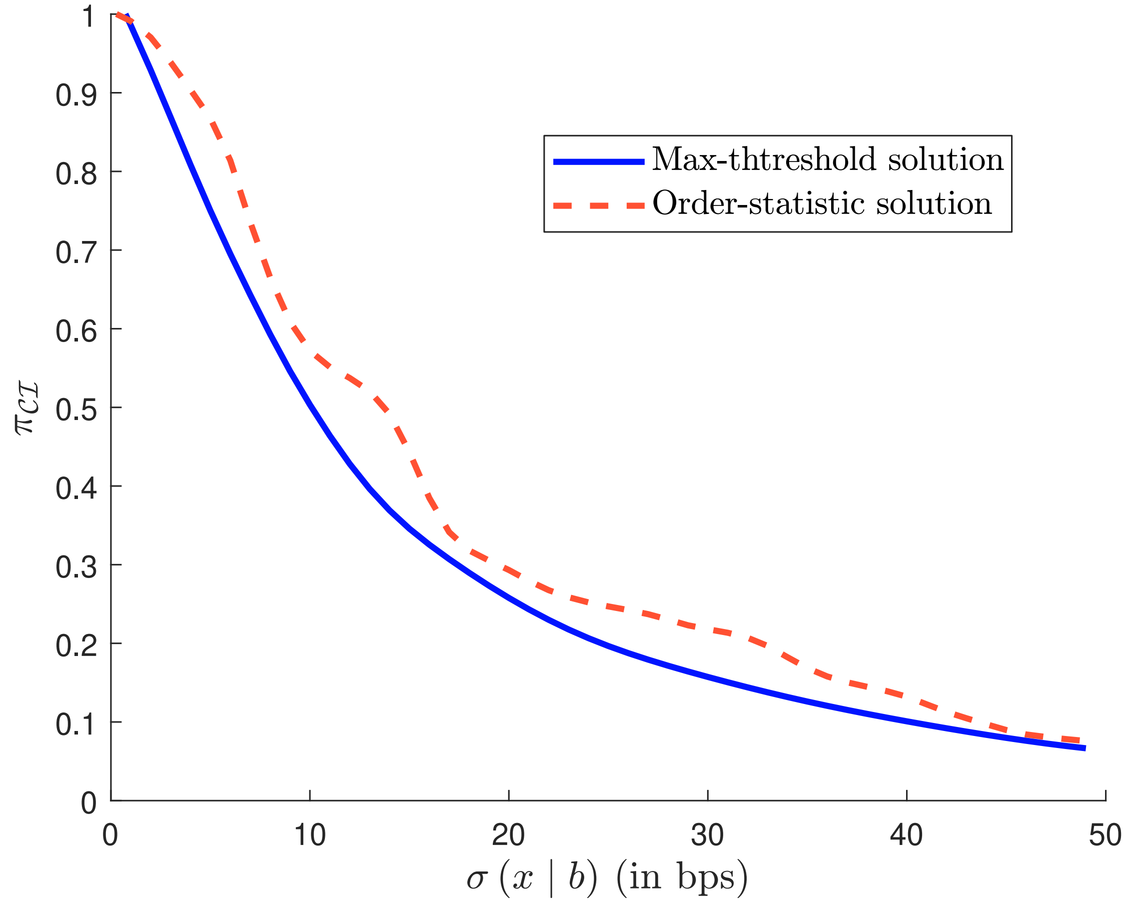


Figure 140: The efficient frontier of optimal decarbonization portfolios (S&P 500 Index, January 2021, Scope 1)

Carbon intensity and the size bias

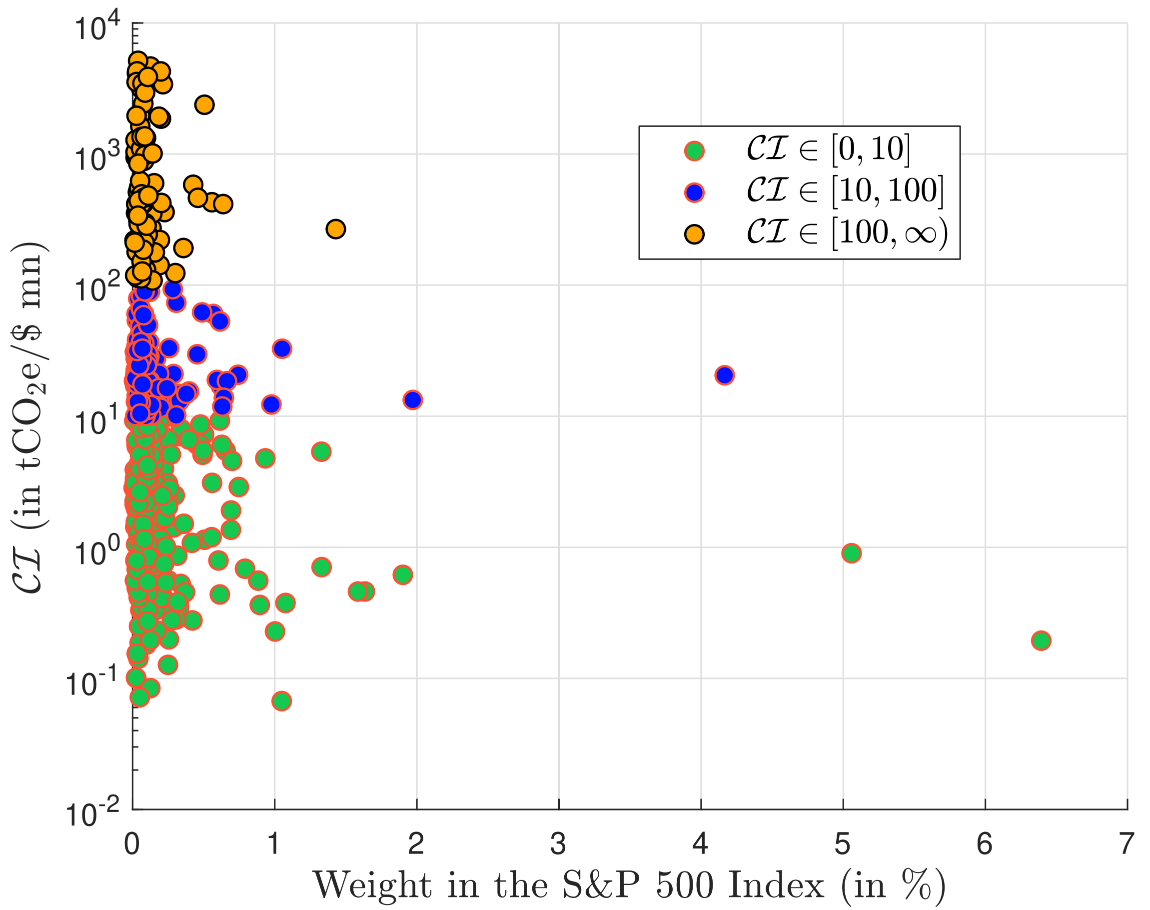


Figure 141: Scatterplot between the index weights b_i and the carbon intensity $\mathcal{CI}_i$
(S&P 500 Index, January 2021, Scope 1)

Carbon intensity and the size bias

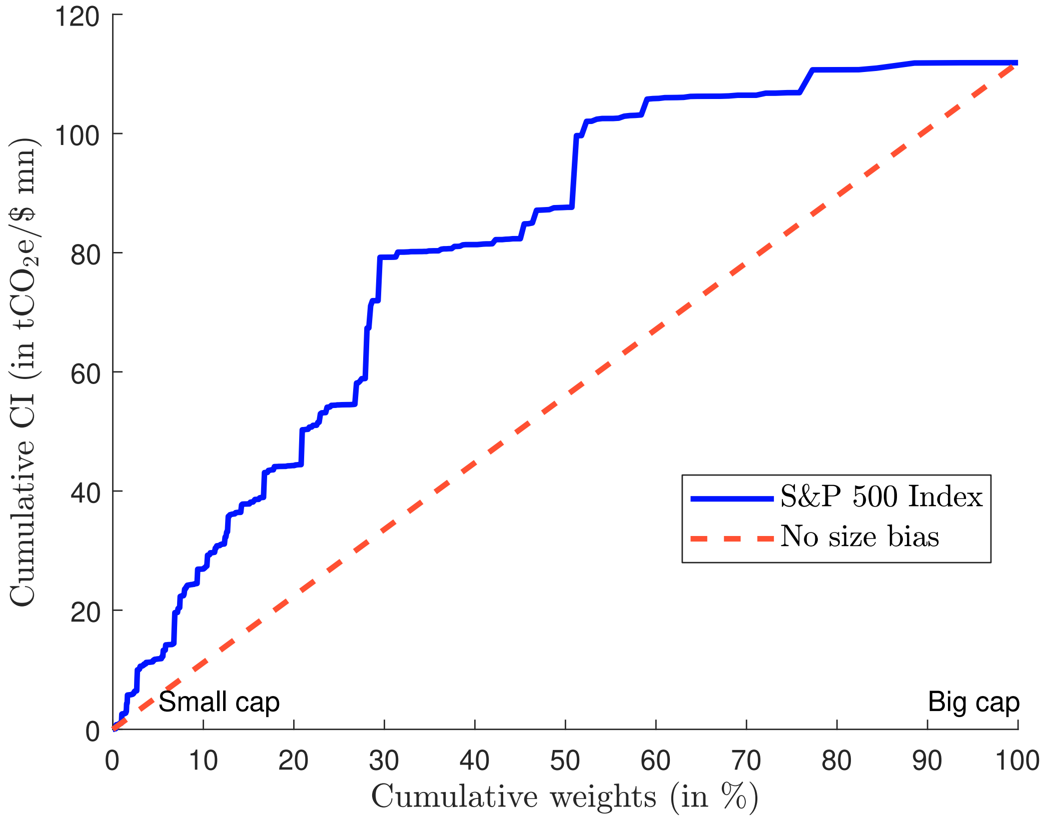


Figure 142: Lorenz curve of the carbon intensity contributions (S&P 500 Index, January 2021, Scope 1)

In January 2021, the Carbon intensity of the S&P 500 Index is equal to 111.89 tCO_2e/$ mn.

Climate changes indexes

- MSCI Climate Change Indexes — `www.msci.com/climate-change-indexes`
- MSCI Climate Paris Aligned Indexes — `www.msci.com/esg/climate-paris-aligned-indexes`
- FTSE Global Climate Index Series — `www.ftserussell.com/products/indices/global-climate`
- FTSE TPI Climate Transition Index Series — `www.ftserussell.com/products/indices/tpi-climate-transition`
- FTSE Climate Risk-Adjusted Government Bond Index Series — `www.ftserussell.com/products/indices/climate-wgbi`
- S&P Climate Indices — `www.spglobal.com/spdji/en/index-family/equity/esg/climate`
- STOXX Climate Transition Benchmark (CTB) and STOXX Paris-Aligned Benchmark (PAB) Indices — `qontigo.com/solutions/climate-indices`

Climate changes indexes

- Most of the climate change indices use the following weighting scheme:

$$x_i = \frac{s_i \times b_i}{\sum_{j=1}^{n} s_j \times b_j}$$

  where s_i is the climate change score of the company and b_i is the weight of the company in the parent index (or benchmark)

- The climate change score is generally a combined score based on:
  1. Carbon emission score
  2. Asset stranding score
  3. Climate management score
  4. Green revenue score
  5. Etc.

Financial risk of climate change

The previous approach assumes that
the **climate-related market risk** of a company
is measured by its **current carbon intensity**

...But the **market perception**
of the climate change may be different

Financial risk of climate change

The following analysis is based on the following papers:

- Görgen, M., Jacob, A., Nerlinger, M., Riordan, R., Rohleder, M., and Wilkens, M. (2019), Carbon Risk, *SSRN*, `https://www.ssrn.com/abstract=2930897`.
- Roncalli, T., Le Guenedal, T., Lepetit, F., Roncalli, T., and Sekine, T. (2020), Measuring and Managing Carbon Risk in Investment Portfolios, *Amundi Working Paper*, WP-99-2020, `www.research-center.amundi.com`.

Financial risk of climate change

Goal

The main objective is to define a **market** measure of carbon risk

Three-step approach

- Defining a brown green score (BGS) for each stock (scoring model)
- Building a brown minus green factor (Fama-French approach)
- Estimating the carbon beta of a stock with respect to the BMG factor (Multi-factor regression analysis)

Carbon beta = **market** measure of carbon risk

$\neq$

Carbon intensity = **fundamental** measure of carbon risk

Financial risk of climate change

The example of carbon intensity

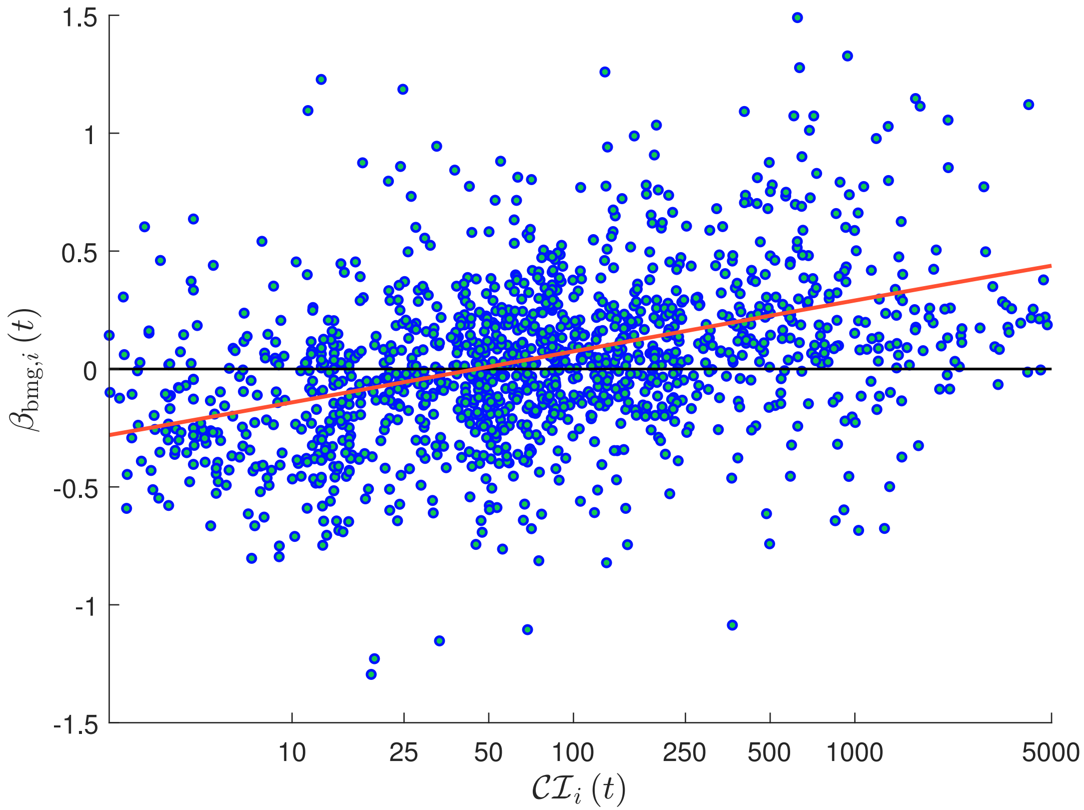


Figure 143: Market-based vs fundamental-based measures of carbon risk

⇒ The market perception of a carbon risk measure depends on several dimensions: sector, country, etc.

Which carbon risk?

Systematic carbon risk

- Common risk
- Carbon beta

Market measure ($\approx$ general carbon risk exposure, e.g. market repricing risk)

Idiosyncratic carbon risk

- Specific risk
- Carbon intensity

Fundamental measure ($\approx$ specific carbon risk exposure, e.g. reputational risk)

Construction of the BMG factor

Risk factor approach (Fama-French)

	Green	Neutral	Brown
Small	SG	SN	SB
Big	BG	BN	BB

The BMG factor return $R_{\mathrm{bmg}}(t)$ is derived from the Fama-French method:

$$R_{\mathrm{bmg}}(t) = \frac{1}{2}\left(R_{\mathrm{SB}}(t) + R_{\mathrm{BB}}(t)\right) - \frac{1}{2}\left(R_{\mathrm{SG}}(t) + R_{\mathrm{BG}}(t)\right)$$

where the returns of each portfolio $R_j(t)$ (small green SG, big green BG, small brown SB, big brown BB) is value-weighted by the market capitalisation

$\Rightarrow$ The BMG factor is a Fama-French risk factor based on a scoring system (brown green score or BGS)

Construction of the BGS

The CARIMA approach

- Carbon Risk Management (CARIMA)
- Project sponsored by the German Federal Ministry of Education and Research
- They publish the carbon risk factor Brown-Minus-Green (BMG)
- They also provide an excel tool
- Contact: Martin Nerlinger (martin.nerlinger@wiwi.uni-augsburg.de)

https://carima-project.de/en/downloads

Construction of the BGS – The CARIMA approach

Görgen *et al.* (2019) use 55 proxy variables to define the brown green score:

- Value chain (impact of a climate policy or a cap-and-trade system on the different activities of a firm) — VC
- Public perception (external environmental image of a firm) — PP
- Adaptability (capacity of the firm to shift towards a low carbon strategy without strong efforts and losses) — PP

A brown green score (BGS) is created for each stock:

$$\begin{aligned}\mathrm{BGS}_i(t) &= \frac{2}{3}\left(0.7\cdot\mathrm{VC}_i(t)+0.3\cdot\mathrm{PP}_i(t)\right)+ \\ &\quad \frac{\mathrm{NA}_i(t)}{3}\left(0.7\cdot\mathrm{VC}_i(t)+0.3\cdot\mathrm{PP}_i(t)\right)\end{aligned}$$

where VC_i is the value chain score of stock i, PP_i is the public perception score of stock i and NA_i is the non-adaptability score of stock i

Cumulative performance of the BMG factor

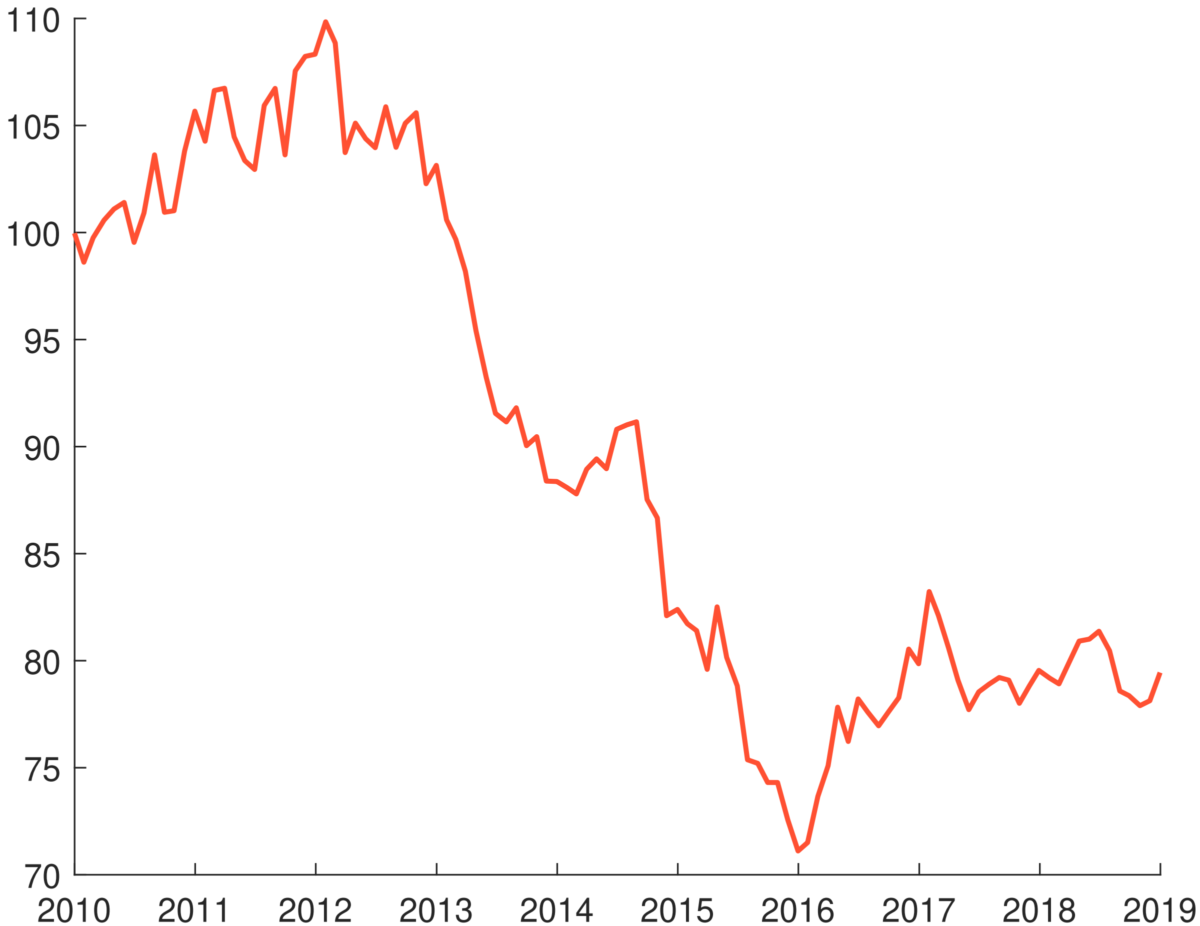


Figure 144: Cumulative performance of the BMG factor

Source: Görgen *et al.* (2019)

Correlation between BMG and other risk factors

Table 85: Correlation matrix of factor returns (in %)

Factor	MKT	SMB	HML	WML	BMG
MKT	100.00***				
SMB	1.41	100.00***			
HML	11.51	−8.93	100.00***		
WML	−14.59	3.87	−41.43***	100.00***	
BMG	5.33	20.33**	27.41***	−21.28**	100.00***

Source: Roncalli *et al.* (2020)

- No significant correlation between market and carbon factors
- Size, value and momentum-specific effects in the BMG factor

Multi-factor analysis

- CAPM

$$R_i(t) = \alpha_i + \beta_{\mathrm{mkt},i} R_{\mathrm{mkt}}(t) + \varepsilon_i(t)$$

- Fama-French 3F model (FF)

$$R_i(t) = \alpha_i + \beta_{\mathrm{mkt},i} R_{\mathrm{mkt}}(t) + \beta_{\mathrm{smb},i} R_{\mathrm{smb}}(t) + \beta_{\mathrm{hml},i} R_{\mathrm{hml}}(t) + \varepsilon_i(t)$$

- MKT+BMG model

$$R_i(t) = \alpha_i + \beta_{\mathrm{mkt},i} R_{\mathrm{mkt}}(t) + \beta_{\mathrm{bmg},i} R_{\mathrm{bmg}}(t) + \varepsilon_i(t)$$

- Extended Fama-French model (FF+BMG)

$$\begin{aligned} R_i(t) &= \alpha_i + \beta_{\mathrm{mkt},i} R_{\mathrm{mkt}}(t) + \beta_{\mathrm{smb},i} R_{\mathrm{smb}}(t) + \beta_{\mathrm{hml},i} R_{\mathrm{hml}}(t) + \\ &\quad \beta_{\mathrm{bmg},i} R_{\mathrm{bmg}}(t) + \varepsilon_i(t) \end{aligned}$$

Multi-factor analysis

- Carhart model (4F)

$$\begin{aligned} R_i(t) &= \alpha_i + \beta_{\mathrm{mkt},i} R_{\mathrm{mkt}}(t) + \beta_{\mathrm{smb},i} R_{\mathrm{smb}}(t) + \beta_{\mathrm{hml},i} R_{\mathrm{hml}}(t) + \\ & \beta_{\mathrm{wml},i} R_{\mathrm{wml}}(t) + \varepsilon_i(t) \end{aligned}$$

- Extended Carhart model (4F+BMG)

$$\begin{aligned} R_i(t) &= \alpha_i + \beta_{\mathrm{mkt},i} R_{\mathrm{mkt}}(t) + \beta_{\mathrm{smb},i} R_{\mathrm{smb}}(t) + \beta_{\mathrm{hml},i} R_{\mathrm{hml}}(t) + \\ & \beta_{\mathrm{wml},i} R_{\mathrm{wml}}(t) + \beta_{\mathrm{bmg},i} R_{\mathrm{bmg}}(t) + \varepsilon_i(t) \end{aligned}$$

$\Rightarrow$ These models are estimated using OLS and stocks that compose the MSCI World Index from January 2010 to December 2018

Relevance of the BMG factor

Table 86: Comparison of cross-section regressions (in %)

	Adjusted $\mathfrak{R}^2$ difference	F-test 10%	F-test 5%	F-test 1%
CAPM vs FF	1.74	34.6	25.5	13.5
CAPM vs MKT+BMG	1.74	21.2	15.6	9.2
FF vs FF+BMG	1.73	22.5	17.5	9.7
FF vs FF+WML	0.22	6.6	3.0	0.8
4F vs 4F+BMG	1.76	23.6	18.6	10.0

Source: Roncalli *et al.* (2020)

⇒ The effect on the explanatory power is at the same level for the SMB and HML factors together and the BMG factor alone

Sectorial analysis

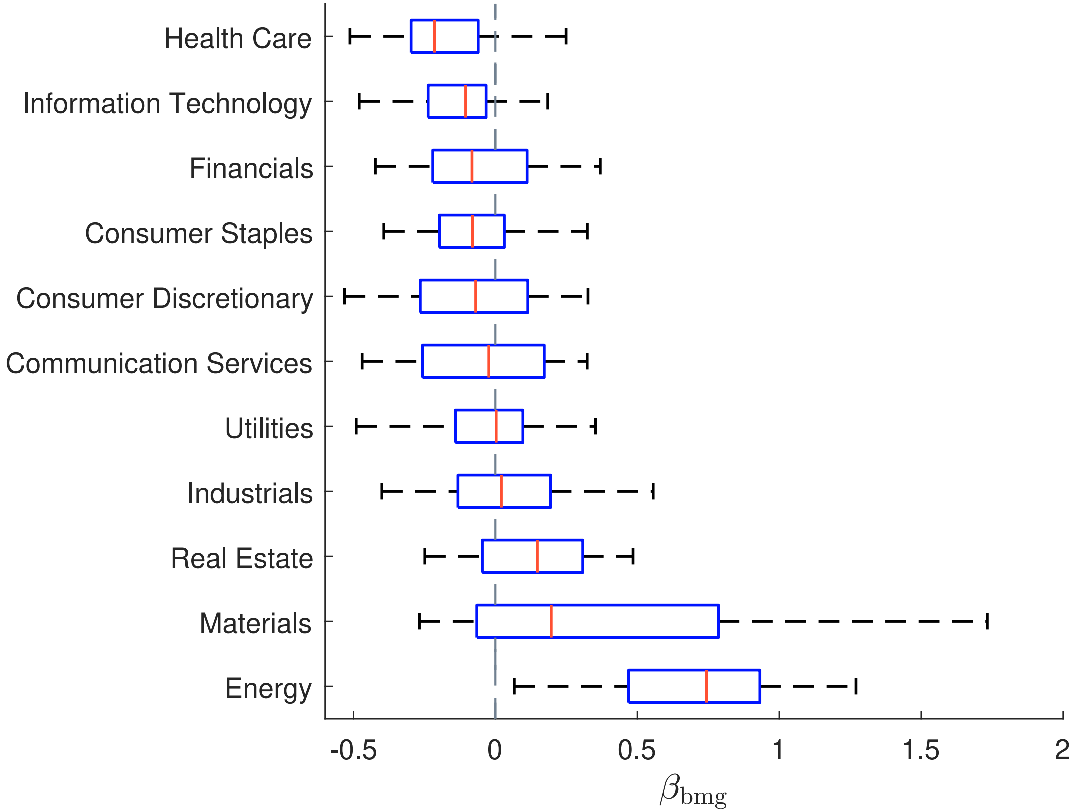


Figure 145: Box plots of the carbon sensitivities[25]

Source: Roncalli *et al.* (2020)

[25] The box plots provide the median, the quartiles and the 5% and 95% quantiles

Absolute versus relative carbon risk

Relative carbon risk

- The right measure is β_{bmg}
- Sign matters
- **Negative exposure** is preferred

Absolute carbon risk

- The right measure is $|\beta_{\mathrm{bmg}}|$
- Sign doesn't matter
- **Zero exposure** is preferred

Two examples

1. We consider three portfolios with a carbon beta of -0.30, -0.05 and $+0.30$ respectively
2. We consider two portfolios with the following characteristics:
   - The value of the carbon beta is $+0.10$ and the stock dispersion of carbon beta is 0.20
   - The value of the carbon beta is -0.30 and the stock dispersion of carbon beta is 1.50

$\Rightarrow$ Impact of portfolio management and theory

Dynamic estimation of β_{bmg}

We use the following dynamic common factor model:

$$R_i(t) = R(t)^\top \beta_i(t) + \varepsilon_i(t)$$

where $R(t) = (1, R_{\mathrm{mkt}}(t), R_{\mathrm{bmg}}(t))$ is the vector of factor returns, $\beta_i(t) = (\alpha_i(t), \beta_{\mathrm{mkt},i}(t), \beta_{\mathrm{bmg},i}(t))$ is the vector of factor betas and $\varepsilon_i(t)$ is a white noise.

Assumption

The state vector $\beta_i(t)$ follows a random walk process:

$$\beta_i(t) = \beta_i(t-1) + \eta_i(t)$$

where $\eta_i(t) \sim \mathcal{N}(\mathbf{0}_3, \Sigma_{\beta,i})$ is the white noise vector and $\Sigma_{\beta,i}$ is the diagonal covariance matrix of the white noise.

$\Rightarrow$ The model is estimated with the Kalman filter

Dynamic estimation of β_{bmg}

State space model (SSM)

- The measurement equation defines the relationship between an observable system y_t and state variables α_t:

$$y_t = Z_t \alpha_t + d_t + \epsilon_t$$

  where y_t is a n-dimensional time series, Z_t is a $n \times m$ matrix, d_t is a $n \times 1$ vector
- The state vector α_t is generated by a Markov linear process:

$$\alpha_t = T_t \alpha_{t-1} + c_t + R_t \eta_t$$

  where α_t is a $m \times 1$ vector, T_t is a $m \times m$ matrix, c_t is a $m \times 1$ vector and R_t is a $m \times p$ matrix
- $\eta_t \sim \mathcal{N}\left(\mathbf{0}_p, Q_t\right)$ and $\epsilon_t \sim \mathcal{N}\left(\mathbf{0}_n, H_t\right)$ are independent white noise processes of dimension p and n with covariance matrices Q_t and H_t

Dynamic estimation of β_{bmg}

- $\alpha_0 \sim \mathcal{N}(\hat{\alpha}_0, P_0)$ is the initial position of the state vector
- We note $\hat{\alpha}_{t|t}$ (or $\hat{\alpha}_t$) and $\hat{\alpha}_{t|t-1}$ the optimal estimators of α_t given the available information until time t and $t-1$:

$$\begin{aligned}\hat{\alpha}_{t|t} &= \mathbb{E}\left[\alpha_t \mid \mathcal{F}_t\right] \\ \hat{\alpha}_{t|t-1} &= \mathbb{E}\left[\alpha_t \mid \mathcal{F}_{t-1}\right]\end{aligned}$$

- $P_{t|t}$ (or P_t) and $P_{t|t-1}$ are the covariance matrices associated to $\hat{\alpha}_{t|t}$ and $\hat{\alpha}_{t|t-1}$:

$$\begin{aligned}P_{t|t} &= \mathbb{E}\left[\left(\hat{\alpha}_{t|t} - \alpha_t\right)\left(\hat{\alpha}_{t|t} - \alpha_t\right)^\top\right] \\ P_{t|t-1} &= \mathbb{E}\left[\left(\hat{\alpha}_{t|t-1} - \alpha_t\right)\left(\hat{\alpha}_{t|t-1} - \alpha_t\right)^\top\right]\end{aligned}$$

Dynamic estimation of β_{bmg}

Kalman filter

- These different quantities are calculated thanks to the Kalman filter, which consists in the following recursive algorithm:

$$\left\{\begin{array}{l} \hat{\alpha}_{t|t-1} = T_t \hat{\alpha}_{t-1|t-1} + c_t \\ P_{t|t-1} = T_t P_{t-1|t-1} T_t^\top + R_t Q_t R_t^\top \\ \hat{y}_{t|t-1} = Z_t \hat{\alpha}_{t|t-1} + d_t \\ v_t = y_t - \hat{y}_{t|t-1} \\ F_t = Z_t P_{t|t-1} Z_t^\top + H_t \\ \hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-1} + P_{t|t-1} Z_t^\top F_t^{-1} v_t \\ P_{t|t} = \left(I_m - P_{t|t-1} Z_t^\top F_t^{-1} Z_t\right) P_{t|t-1} \end{array}\right.$$

where:

- $\hat{y}_{t|t-1} = \mathbb{E}\left[y_t \mid \mathcal{F}_{t-1}\right]$ is the best estimator of y_t given the available information until time $t-1$
- $v_t \sim \mathcal{N}\left(\mathbf{0}_n, F_t\right)$ is the innovation process

Dynamic estimation of β_{bmg}

- The time-varying risk factor model can be written as a state space model:

$$\left\{ \begin{array}{l} y(t) = x(t)^\top \beta(t) + \varepsilon(t) \\ \beta(t) = \beta(t-1) + \eta(t) \end{array} \right.$$

  where $\varepsilon(t) \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$, $\eta(t) \sim \mathcal{N}\left(\mathbf{0}_{K+1}, \Sigma_\beta\right)$ and K is the number of risk factors

- In the case of the MKT+BMG model, $y(t)$ corresponds to the asset return $R_i(t)$, $x(t)$ is a 3×1 vector, whose elements are 1, $R_{\mathrm{mkt}}(t)$ and $R_{\mathrm{bmg}}(t)$ and:

$$\beta(t) = \left(\begin{array}{c} \alpha_i(t) \\ \beta_{\mathrm{mkt},i}(t) \\ \beta_{\mathrm{bmg},i}(t) \end{array} \right)$$

Dynamic estimation of β_{bmg}

- $\beta(0) \sim \mathcal{N}(\beta_0, P_0)$ is the initial position of the state vector
- We note $\hat{\beta}(t \mid t-1) = \mathbb{E}[\beta(t) \mid \mathcal{F}(t-1)]$ and $\hat{\beta}(t \mid t) = \mathbb{E}[\beta(t) \mid \mathcal{F}(t)]$ as the optimal estimators of $\beta(t)$ given the available information until time $t-1$ and t
- $P(t \mid t-1)$ and $P(t \mid t)$ are the covariance matrices associated with $\hat{\beta}(t \mid t-1)$ and $\hat{\beta}(t \mid t)$
- The estimate of $y(t)$ is equal to:

$$\hat{y}(t \mid t-1) = x(t)^\top \hat{\beta}(t \mid t-1)$$

- The innovation process $v(t) = y(t) - \hat{y}(t \mid t-1)$ is equal to:

$$\begin{aligned} v(t) &= x(t)^\top \beta(t) + \varepsilon(t) - x(t)^\top \hat{\beta}(t \mid t-1) \\ &= -x(t)^\top \left(\hat{\beta}(t \mid t-1) - \beta(t)\right) + \varepsilon(t) \end{aligned}$$

- The variance $F(t)$ of the innovation process $v(t)$ is then equal to:

$$F(t) = x(t)^\top P(t \mid t-1)\, x(t) + \sigma_\varepsilon^2$$

Dynamic estimation of β_{bmg}

- The Kalman filter becomes:

$$\begin{cases}\hat{\beta}(t \mid t-1) = \hat{\beta}(t-1 \mid t-1) \\ P(t \mid t-1) = P(t-1 \mid t-1) + \Sigma_\beta \\ v(t) = y(t) - x(t)^\top \hat{\beta}(t \mid t-1) \\ F(t) = x(t)^\top P(t \mid t-1) x(t) + \sigma_\varepsilon^2 \\ \hat{\beta}(t \mid t) = \hat{\beta}(t \mid t-1) + \left(\dfrac{P(t \mid t-1)}{F(t)}\right) x(t) v(t) \\ P(t \mid t) = \left(I_{K+1} - \left(\dfrac{P(t \mid t-1)}{F(t)}\right) x(t) x(t)^\top\right) P(t \mid t-1)\end{cases}$$

Dynamic estimation of β_{bmg}

- In this model, the parameters σ_{ε}^2 and Σ_{β} are unknown and can be estimated by the method of maximum likelihood
- Since $v(t) \sim \mathcal{N}(0, F(t))$, the log-likelihood function is equal to:

$$\ell(\theta) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln F(t) + \frac{v^2(t)}{F(t)}\right)$$

where $\theta = \left(\sigma_{\varepsilon}^2, \Sigma\right)$

- Maximizing the log-likelihood function requires specifying the initial conditions β_0 and P_0, which are not necessarily known. In this case, we use the linear regression $y(t) = x(t)^{\top}\beta + \varepsilon(t)$, and the OLS estimates $\hat{\beta}_{\mathrm{ols}}$ and $\hat{\sigma}_{\varepsilon}^2\left(X^{\top}X\right)^{-1}$ to initialize β_0 and P_0

Regional analysis of the relative carbon risk

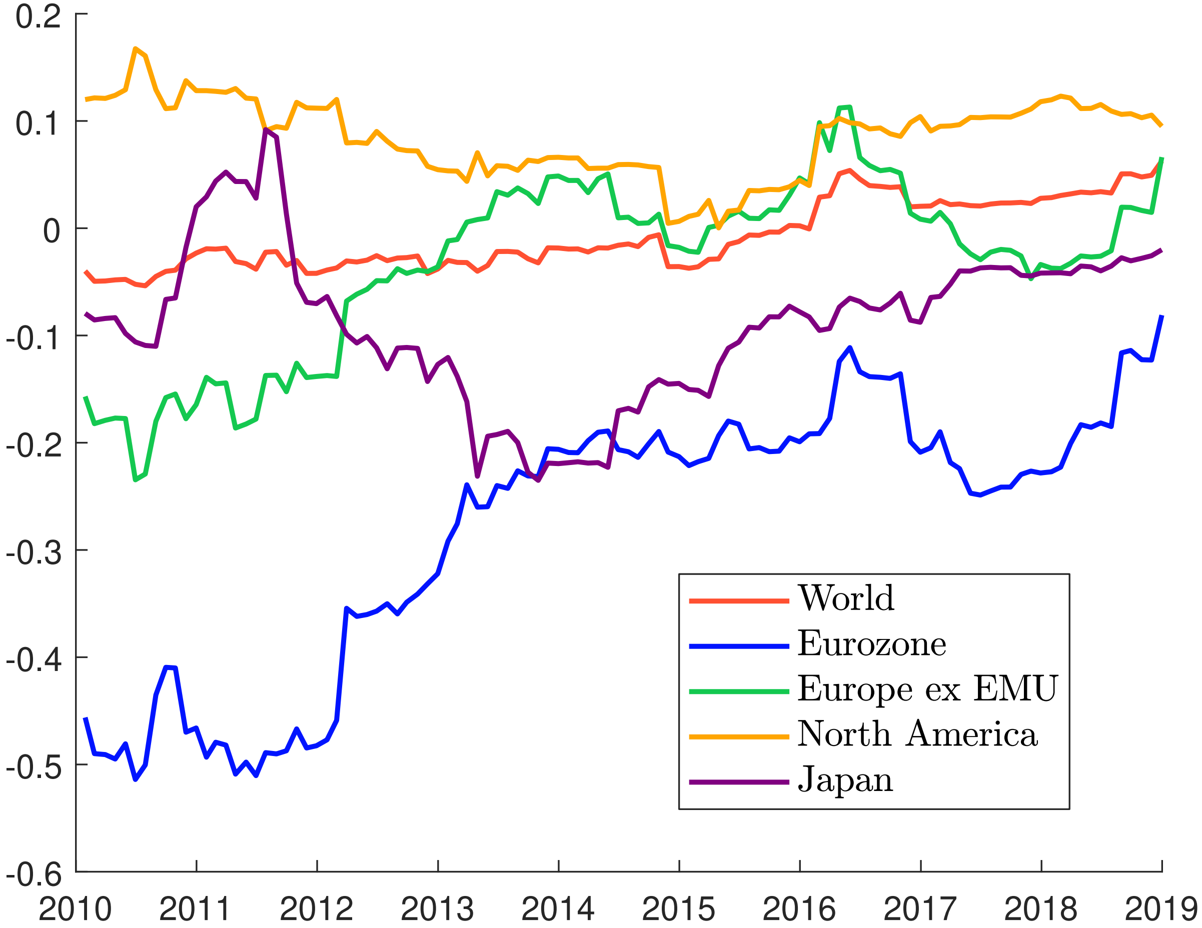


Figure 146: Dynamics of the average relative carbon risk $\beta_{\mathrm{bmg},\mathcal{R}}(t)$ by region

Source: Roncalli *et al.* (2020)

Regional analysis of the absolute carbon risk

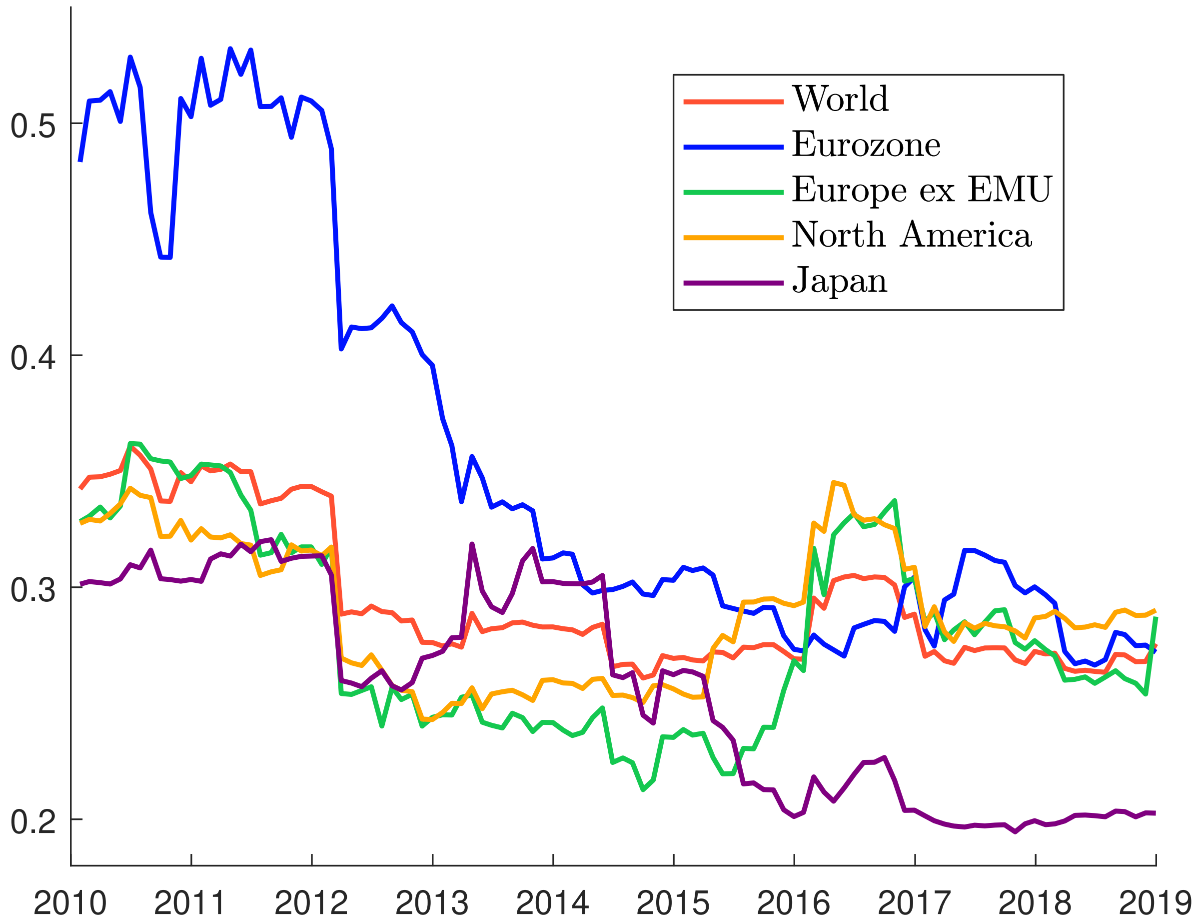


Figure 147: Dynamics of the average absolute carbon risk $|\beta|_{\mathrm{bmg},\mathcal{R}}(t)$ by region

Source: Roncalli *et al.* (2020)

Sectorial analysis

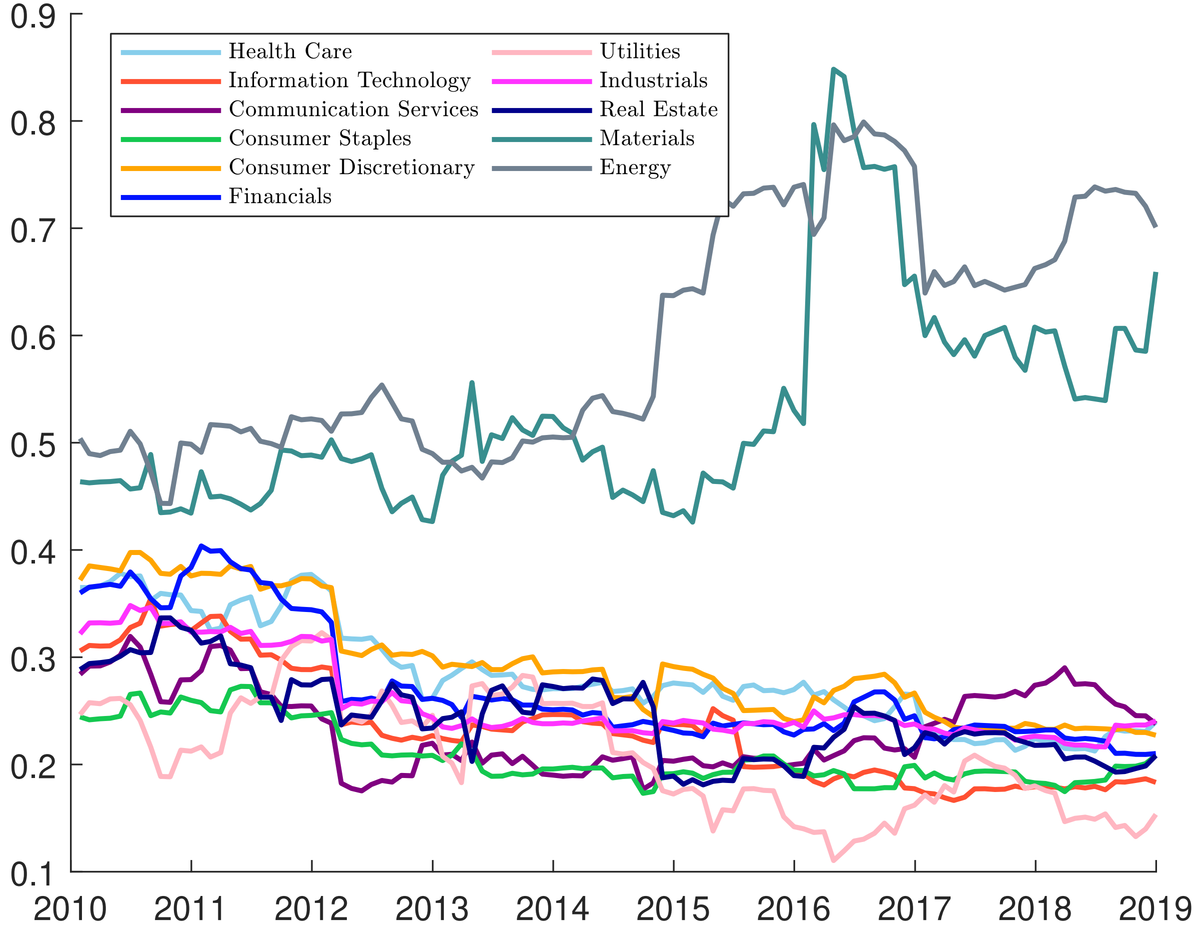


Figure 148: Dynamics of the average absolute carbon risk $|\beta|_{\text{bmg},\mathcal{S}}(t)$ by sector

Source: Roncalli *et al.* (2020)

Advantages and limits of the Carima factor

Advantages

- Biases in the databases are offset because the BGS scores are derived from several databases
- No significant country-specific and sector-specific effects
- No problem of extreme values
- Encompass a lot of climate change-relevant information

Limits

- No differentiation between values near and far the median of a variable
- No rebalancing schemes
- Correlation between BMG factor and some other factors
- Double counting problems
- Not only carbon risk dimension

$\Rightarrow$ Some variables can create more noise than information

Which climate change-related dimensions are the more priced by the market?

Alternative risk factors

We consider the following dimensions

1. Carbon intensity
2. Carbon emissions exposure
3. Carbon emissions management
4. Carbon emissions (exposure + management)
5. Climate change
6. Environmental

Differences with the CARIMA factor

1. Equally-weighted portfolio
2. Integration of the financials sector
3. Rebalancing
4. One variable $\Rightarrow$ no double counting problems

Alternative risk factors

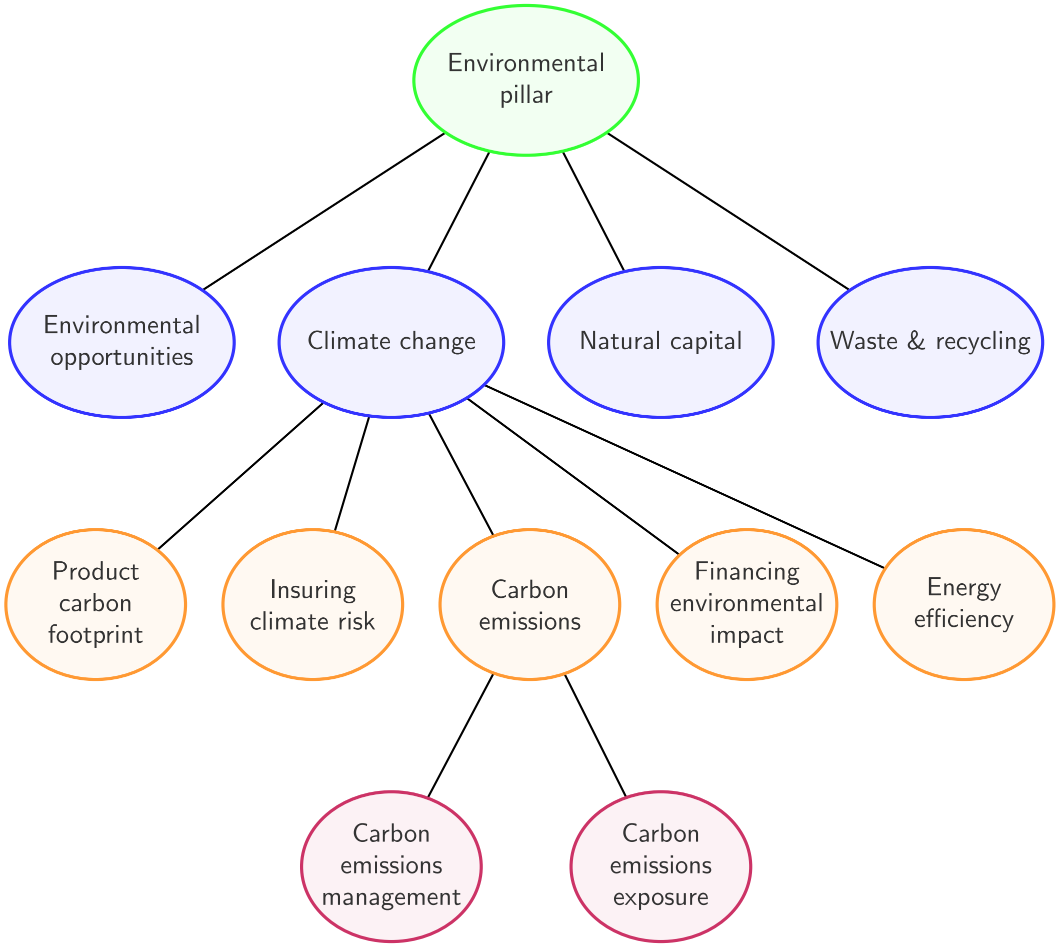


Figure 149: Dimension hierarchy in the environmental pillar (MSCI methodology)

Source: MSCI (2020)

Alternative risk factors

Exposure to carbon costs

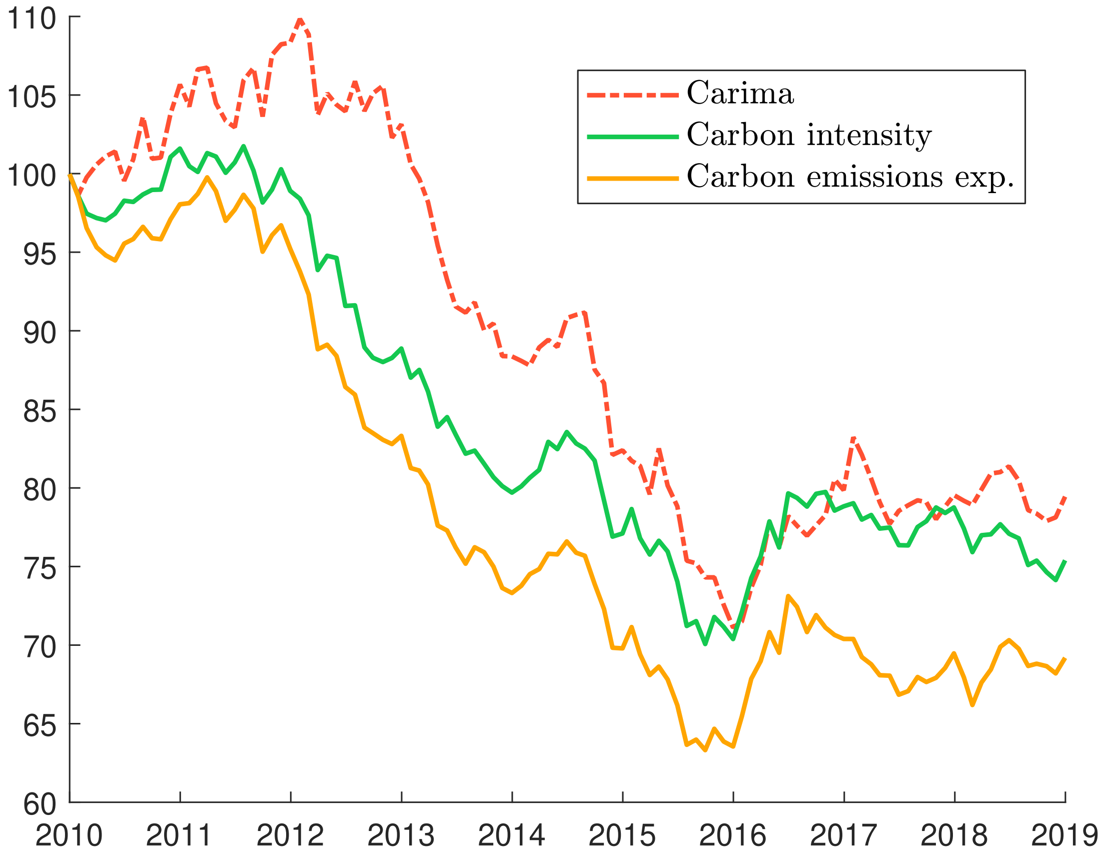


Figure 150: Cumulative performance of the carbon exposure factors

Source: Roncalli *et al.* (2020)

Alternative risk factors

Exposure to carbon costs

Is carbon intensity the unique carbon dimension priced by the market?

Alternative risk factors

Environmental, climate and carbon dimensions

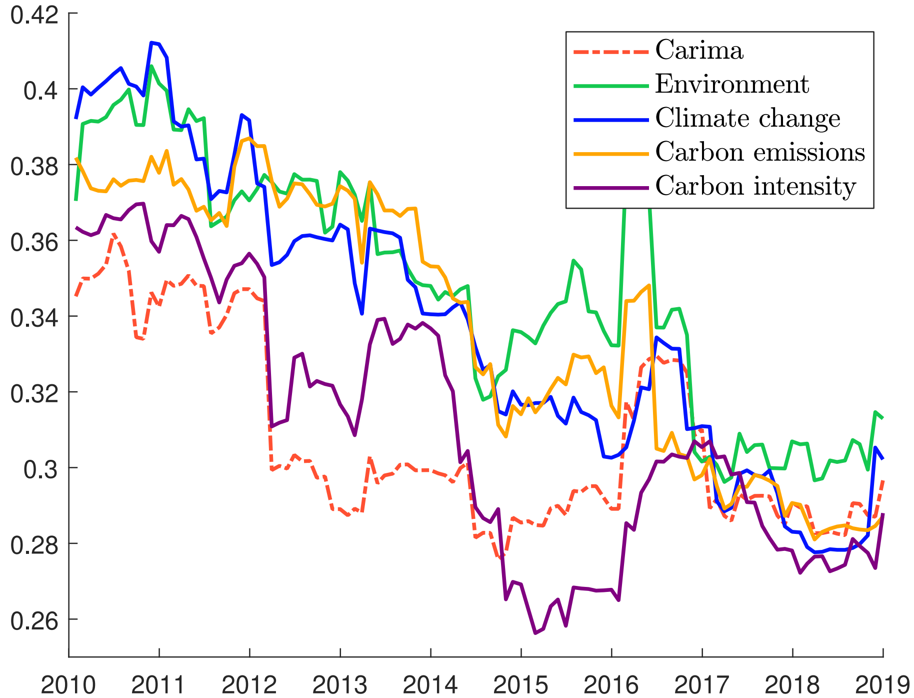


Figure 151: Dynamics of the average absolute carbon risk $|\beta|_{\text{bmg},i}(t)$

Source: Roncalli *et al.* (2020)

Each carbon factor is standardized such that its monthly volatility is equal to the monthly volatility of the market risk

Alternative risk factors

Comparison of the explanatory power

Table 87: Adjusted $\mathfrak{R}^2$ difference

	Full period	1st subperiod	2nd subperiod
Carima	1.74	1.16	2.21
Carbon intensity	1.77	1.43	2.53
Carbon emissions	2.00	2.18	2.39
Climate change	1.58	1.98	1.83
Environment	1.63	1.35	2.17
Carbon intensity*	2.06	1.25	3.13
Carbon emissions*	1.91	1.41	2.42

Source: Roncalli *et al.* (2020)

* *means that the carbon factor is based on the quintile methodology* Q_5 - Q_1

Alternative risk factors

Factor correlations

Table 88: Correlation matrix of factor returns (in %)

Factor	MKT	SMB	HML	WML	BMG
Carbon intensity	−6.46	13.71	8.71	−3.04	58.40***
Carbon emissions exp.	−6.71	14.95	4.03	−4.03	64.02***
Carbon emissions mgmt.	−17.93*	24.16**	−20.91**	20.93**	38.66***
Carbon emissions	1.22	25.85***	−0.23	5.15	72.36***
Climate change	−15.02	16.30*	11.43	2.07	61.11***
Environment	−28.20***	21.16**	−0.33	3.70	68.53***
Carbon intensity*	−18.69*	7.79	−3.64	8.24	54.13***
Carbon emissions*	10.04	27.94***	22.15**	−17.92*	81.42***

Source: Roncalli *et al.* (2020)

Market-specific effect for carbon emissions management, environmental and carbon intensity* factors $\Rightarrow$ bias in a minimum variance portfolio

Portfolio optimization with climate risk

Risk factor model

- We consider the MKT+BMG risk factor model:

$$R_i(t) = \alpha_i + \beta_{\mathrm{mkt},i} R_{\mathrm{mkt}}(t) + \beta_{\mathrm{bmg},i} R_{\mathrm{bmg}}(t) + \varepsilon_i(t)$$

- We assume that $R_{\mathrm{mkt}}(t)$ and $R_{\mathrm{bmg}}(t)$ are uncorrelated
- The covariance matrix is:

$$\Sigma = \beta_{\mathrm{mkt}} \beta_{\mathrm{mkt}}^\top \sigma_{\mathrm{mkt}}^2 + \beta_{\mathrm{bmg}} \beta_{\mathrm{bmg}}^\top \sigma_{\mathrm{bmg}}^2 + D$$

where β_{mkt} and β_{bmg} are the vector of MKT and BMG betas respectively, σ_{mkt}^2 and σ_{bmg}^2 are the variance of the market and carbon portfolios and $D = \mathrm{diag}\left(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_n^2\right)$ is the diagonal matrix of idiosyncratic risks

Application to the minimum variance portfolio

Analytical model

We consider the GMV portfolio:

$$\begin{aligned} x^{\star} &= \arg\min \frac{1}{2} x^{\top} \Sigma x \\ \text{s.t.} & \quad \mathbf{1}_n^{\top} x = 1 \end{aligned}$$

where x is the vector of portfolio weights and Σ is the covariance matrix of stock returns

Application to the minimum variance portfolio

Analytical model

Reminder (Lecture 3)

The solution is equal to:

$$x^{\star} = \frac{\Sigma^{-1}\mathbf{1}_n}{\mathbf{1}_n^{\top}\Sigma^{-1}\mathbf{1}_n}$$

Application to the minimum variance portfolio

Analytical model

Sherman-Morrison-Woodbury (SMW) formula

Suppose u and v are two $n \times 1$ vectors and A is an invertible $n \times n$ matrix. We can show that:

$$\left(A + uv^\top\right)^{-1} = A^{-1} - \frac{1}{1 + v^\top A^{-1} u} A^{-1} uv^\top A^{-1}$$

Application to the minimum variance portfolio

Analytical model

Extended SMW formula

Roncalli *et al.* (2020) show that:

$$\left(A + u_1 v_1^\top + u_2 v_2^\top\right)^{-1} = A^{-1} - A^{-1} U S^{-1} V^\top A^{-1}$$

where $U = \begin{pmatrix} u_1 & u_2 \end{pmatrix}$, $V = \begin{pmatrix} v_1 & v_2 \end{pmatrix}$ and:

$$S = \begin{pmatrix} 1 + v_1^\top A^{-1} u_1 & v_1^\top A^{-1} u_2 \\ v_2^\top A^{-1} u_1 & 1 + v_2^\top A^{-1} u_2 \end{pmatrix}$$

Application to the minimum variance portfolio
Analytical model

In order to compute Σ^{-1}, we apply the extended SMW formula with:

- $A = D$
- $u_1 = v_1 = \sigma_{\mathrm{mkt}} \beta_{\mathrm{mkt}}$
- $u_2 = v_2 = \sigma_{\mathrm{bmg}} \beta_{\mathrm{bmg}}$

It follows that the inverse of the covariance matrix is equal to:

$$\Sigma^{-1} = D^{-1} - D^{-1} U S^{-1} V^{\top} D^{-1}$$

where:

$$U = V = \left(\begin{array}{cc} \sigma_{\mathrm{mkt}} \beta_{\mathrm{mkt}} & \sigma_{\mathrm{bmg}} \beta_{\mathrm{bmg}} \end{array} \right)$$

and:

$$S = \left(\begin{array}{cc} 1 + \sigma_{\mathrm{mkt}}^2 \beta_{\mathrm{mkt}}^{\top} D^{-1} \beta_{\mathrm{mkt}} & \sigma_{\mathrm{mkt}} \sigma_{\mathrm{bmg}} \beta_{\mathrm{mkt}}^{\top} D^{-1} \beta_{\mathrm{bmg}} \\ \sigma_{\mathrm{mkt}} \sigma_{\mathrm{bmg}} \beta_{\mathrm{mkt}}^{\top} D^{-1} \beta_{\mathrm{bmg}} & 1 + \sigma_{\mathrm{bmg}}^2 \beta_{\mathrm{bmg}}^{\top} D^{-1} \beta_{\mathrm{bmg}} \end{array} \right)$$

Application to the minimum variance portfolio

Analytical model

Reminder (Lecture 3)

In the case of the MKT risk factor model, the solution of the GMV portfolio is equal to:

$$x_i^\star = \frac{\sigma^2(x^\star)}{\tilde{\sigma}_i^2}\left(1 - \frac{\beta_{\mathrm{mkt},i}}{\beta_{\mathrm{mkt}}^\star}\right)$$

where $\beta_{\mathrm{mkt}}^\star$ is a threshold value

In the case of the MKT+BMG risk factor model, the solution becomes:

$$x_i^\star = \frac{\sigma^2(x^\star)}{\tilde{\sigma}_i^2}\left(1 - \frac{\beta_{\mathrm{mkt},i}}{\beta_{\mathrm{mkt}}^\star} - \frac{\beta_{\mathrm{bmg},i}}{\beta_{\mathrm{bmg}}^\star}\right)$$

where $\beta_{\mathrm{mkt}}^\star$ and $\beta_{\mathrm{bmg}}^\star$ are two threshold values

Application to the minimum variance portfolio

Analytical model

We consider the long-only MV portfolio:

$$\begin{array}{rcl} x^{\star} & = & \arg\min \frac{1}{2} x^{\top} \Sigma x \\ & \text{s.t.} & \left\{ \begin{array}{l} \mathbf{1}_n^{\top} x = 1 \\ \mathbf{0}_n \leq x \leq \mathbf{1}_n \\ x \in \Omega \end{array} \right. \end{array}$$

where x is the vector of portfolio weights and Σ is the covariance matrix of stock returns

Application to the minimum variance portfolio
Analytical model

In the case of long-only portfolios, we obtain the following formula:

$$x_i^\star = \begin{cases} \dfrac{\sigma^2\left(x^\star\right)}{\tilde{\sigma}_i^2}\left(1 - \dfrac{\beta_{\mathrm{mkt},i}}{\beta_{\mathrm{mkt}}^\star} - \dfrac{\beta_{\mathrm{bmg},i}}{\beta_{\mathrm{bmg}}^\star}\right) & \text{if } \dfrac{\beta_{\mathrm{mkt},i}}{\beta_{\mathrm{mkt}}^\star} + \dfrac{\beta_{\mathrm{bmg},i}}{\beta_{\mathrm{bmg}}^\star} \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

where $\beta_{\mathrm{mkt}}^\star$ is a positive threshold and $\beta_{\mathrm{bmg}}^\star$ may be a positive or negative threshold. The MV portfolio selects assets that present a low market beta value but the impact of $\beta_{\mathrm{bmg},i}$ is more complex

Application to the minimum variance portfolio

Analytical model

Low beta, low volatility and negative correlation

$$\rho_{i,j} = \frac{\beta_{\mathrm{mkt},i}\beta_{\mathrm{mkt},j}\sigma^2_{\mathrm{mkt}} + \beta_{\mathrm{bmg},i}\beta_{\mathrm{bmg},j}\sigma^2_{\mathrm{bmg}}}{\sigma_i \sigma_j}$$

where $\beta_{\mathrm{mkt},i}\beta_{\mathrm{mkt},j}$ is generally positive and $\beta_{\mathrm{bmg},i}\beta_{\mathrm{bmg},j}$ is positive or negative. By considering BMG contributions, there is no coherency between low volatility and low correlated assets

Application to the minimum variance portfolio

Applications

Example 2

We consider an investment universe of five assets. Their beta is respectively equal to 0.9, 0.8, 1.2, 0.7 and 1.3 whereas their specific volatility is 4%, 12%, 5%, 8% and 5%. We also assume that the market portfolio volatility is equal to 25%

Parameter set #1

We assume that the BMG sensitivities are respectively equal to -0.5, 0.7, 0.2, 0.9 and -0.3, whereas the volatility of the BMG factor is set to 10%

Application to the minimum variance portfolio

Applications

Table 89: Composition of the minimum variance portfolio (parameter set #1)

Asset	$\beta_{\mathrm{mkt},i}$	$\beta_{\mathrm{bmg},i}$	CAPM GMV	CAPM MV	MKT+BMG GMV	MKT+BMG MV
1	0.90	−0.50	147.33	0.00	166.55	33.54
2	0.80	0.70	24.67	9.45	21.37	1.46
3	1.20	0.20	−49.19	0.00	−58.80	0.00
4	0.70	0.90	74.20	90.55	65.06	64.99
5	1.30	−0.30	−97.01	0.00	−94.18	0.00
	$\beta^{\star}_{\mathrm{mkt}}$		1.0972	0.8307	1.0906	0.8667
	$\beta^{\star}_{\mathrm{bmg}}$				19.7724	9.7394

Source: Roncalli *et al.* (2020)

Application to the minimum variance portfolio

Applications

Example 3

We consider an investment universe of five assets. Their beta is respectively equal to 0.9, 0.8, 1.2, 0.7 and 1.3 whereas their specific volatility is 4%, 12%, 5%, 8% and 5%. We also assume that the market portfolio volatility is equal to 25%

Parameter set #2

We assume that the BMG sensitivities are respectively equal to -1.5, -0.5, 3.0, -1.2 and -0.9, whereas the volatility of the BMG factor is set to 10%

Parameter set #2

We assume that the BMG sensitivities are respectively equal to 1.5, 0.5, -3.0, 1.2 and 0.9, whereas the volatility of the BMG factor is set to 10%

Application to the minimum variance portfolio

Applications

Table 90: Composition of the minimum variance portfolio (parameter sets #2 and #3)

Asset	$\beta_{\text{mkt},i}$	Parameter set #2			Parameter set #3		
		$\beta_{\text{bmg},i}$	GMV	MV	$\beta_{\text{bmg},i}$	GMV	MV
1	0.90	−1.50	105.46	0.00	1.50	105.46	0.00
2	0.80	−0.50	27.88	19.48	0.50	27.88	19.48
3	1.20	3.00	40.19	13.61	−3.00	40.19	13.61
4	0.70	−1.20	76.77	66.91	1.20	76.77	66.91
5	1.30	−0.90	−150.30	0.00	0.90	−150.30	0.00
$\beta^{\star}_{\text{mkt}}$			1.0982	0.9070		1.0982	0.9070
$\beta^{\star}_{\text{bmg}}$			−19.4470	−9.0718		−19.4470	−9.0718

Source: Roncalli *et al.* (2020)

Application to the minimum variance portfolio

Applications

- MSCI World Index
- December 2018

Remark

The BMG factor is rescaled in order to have the same volatility than the MKT factor $\Rightarrow$ does not change the results, but β and β are now comparable!

Application to the minimum variance portfolio

Absolute carbon risk management

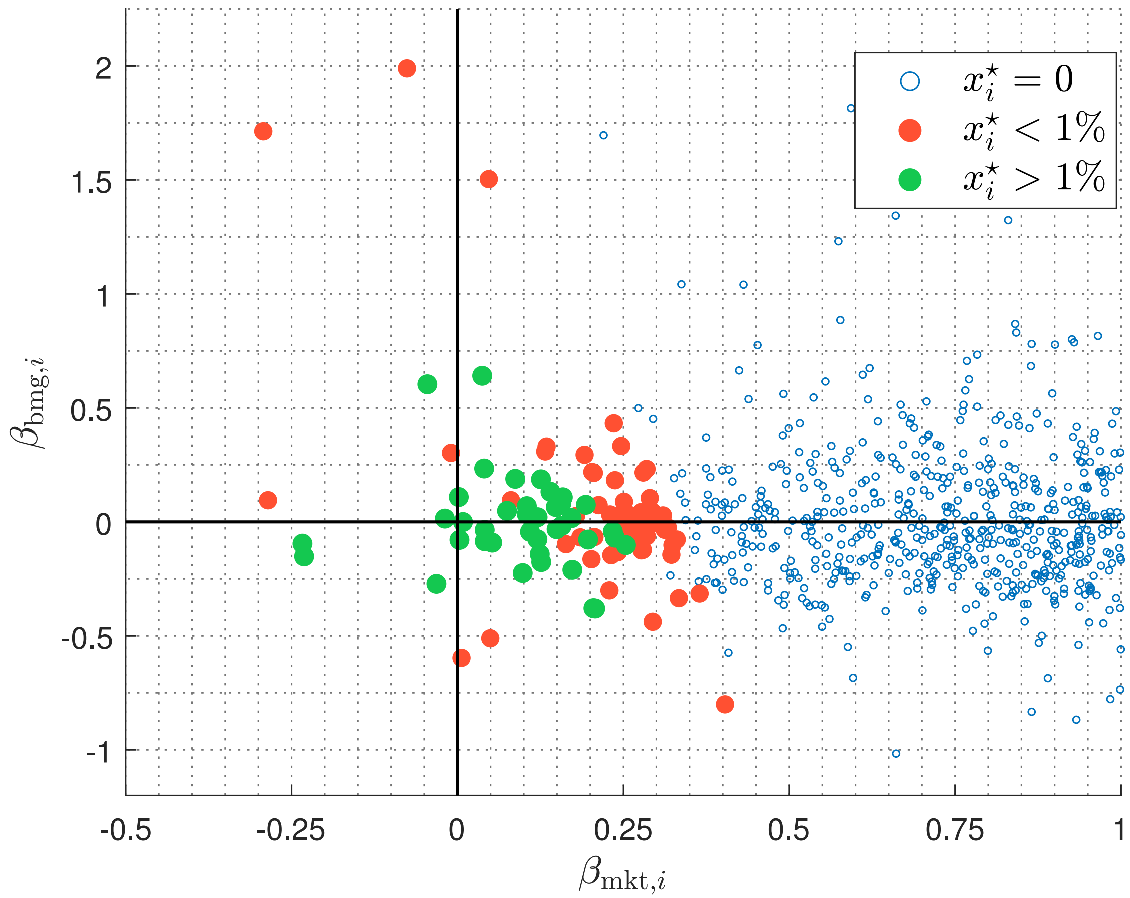


Figure 152: Weights of the MV portfolio (MSCI World Index, Dec. 2018)

Source: Roncalli *et al.* (2021)

Application to the minimum variance portfolio

Absolute carbon risk management

No need to set the constraint:

$$\Omega = \left\{ x \in \mathbb{R}^n : \left| \beta_{\mathrm{bmg}}^\top x \right| \leq |\beta|_{\mathrm{bmg}}^+ \right\}$$

where $|\beta|_{\mathrm{bmg}}^+$ is the maximum absolute carbon risk threshold

The minimum variance portfolio reduces naturally the absolute carbon risk without constraint. Indeed, the portfolio's carbon risk is:

$$\beta_{\mathrm{bmg}}^\top x = 0.016$$

The market risk of a stock determine whether it takes into account in the MV portfolio whereas the carbon risk adjusts the weights of the asset

Application to the minimum variance portfolio

Relative carbon risk management

The optimization program becomes:

$$
\begin{aligned}
x^{\star} & = \arg\min \frac{1}{2} x^{\top} \Sigma x \\
\text{s.t.} & \quad \left\{
\begin{array}{l}
\mathbf{1}_n^{\top} x = 1 \\
\beta_{\mathrm{bmg}}^{\top} x \leq \beta_{\mathrm{bmg}}^{+} \\
x \geq \mathbf{0}_n
\end{array}
\right.
\end{aligned}
$$

where β_{bmg}^{+} is the maximum tolerance of the investor with respect to the relative BMG risk

Application to the minimum variance portfolio

Relative carbon risk management

Table 91: Composition of the constrained MV portfolio ($\beta_{\mathrm{bmg}}^{+} = 0$)

Asset	$\beta_{\mathrm{mkt},i}$	Parameter set #1		Parameter set #2		Parameter set #3	
		$\beta_{\mathrm{bmg},i}$	MV	$\beta_{\mathrm{bmg},i}$	MV	$\beta_{\mathrm{bmg},i}$	MV
1	0.90	−0.50	64.29	−1.50	0.00	1.50	0.00
2	0.80	0.70	0.00	−0.50	19.48	0.50	16.11
3	1.20	0.20	0.00	3.00	13.61	−3.00	25.89
4	0.70	0.90	35.71	−1.20	66.91	1.20	58.00
5	1.30	−0.30	0.00	−0.90	0.00	0.90	0.00

Source: Roncalli *et al.* (2020)

Relative carbon risk management

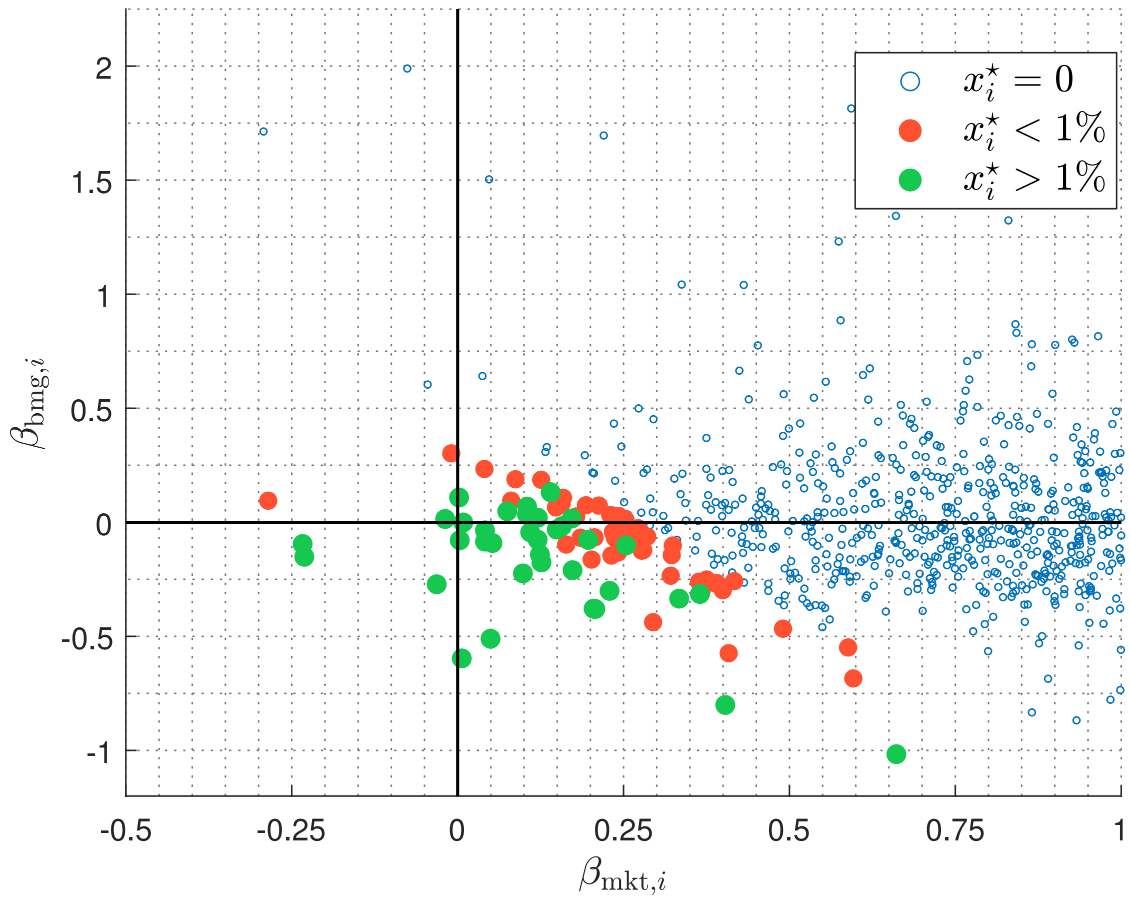


Figure 153: Weights of the constrained MV portfolio ($\beta^{+}_{\mathrm{bmg}} = -0.25$)

Source: Roncalli *et al.* (2021)

Application to the minimum variance portfolio

Managing both systematic and idiosyncratic carbon risks

- Market-based risk management
  - Absolute carbon risk

$$\left|\sum_{i=1}^{n} x_i \times \beta_{\mathrm{bmg},i}\right| \approx 0$$

  - Relative carbon risk

$$\beta_{\mathrm{bmg}}(x) = \sum_{i=1}^{n} x_i \times \beta_{\mathrm{bmg},i} \leq \beta_{\mathrm{bmg}}^{+}$$

- Fundamental-based risk management
  - Individual threshold

$$x_i = 0 \quad \text{if} \quad \mathcal{CI}_i \leq \mathcal{CI}^{+}$$

    where $\mathcal{CI}_i$ is the carbon intensity of stock i
  - Portfolio threshold

$$\mathcal{CI}(x) = \sum_{i=1}^{n} x_i \times \mathcal{CI}_i \leq \mathcal{CI}^{\star}$$

    where $\mathcal{CI}(x)$ is the weighted average carbon intensity (WACI) of portfolio x

Application to the minimum variance portfolio

Managing both systematic and idiosyncratic carbon risks

- $\beta_{\mathrm{bmg}}(x)$ is the carbon beta of portfolio x
- $\mathcal{CI}(x)$ is the carbon intensity of portfolio x
- $\mathcal{CI}(x)$ is the number of holdings of portfolio x
- β^{+}_{bmg} is the maximum tolerance of the investor with respect to the relative carbon risk of the portfolio
- $\mathcal{CI}^{+}$ is the maximum tolerance of the investor with respect to the carbon intensity of individual assets
- $\mathcal{CI}^{\star}$ is the maximum tolerance of the investor with respect to the carbon intensity of the portfolio
- $\mathrm{WO}(x)$ is the portfolio's weight overlap with respect to the optimized portfolio based only on the CI constraint

Application to the minimum variance portfolio

Managing both systematic and idiosyncratic carbon risks

Table 92: Minimum variance portfolios with a relative carbon beta constraint (MSCI World Index, December 2018)

β^{+}_{bmg}	$\beta_{\mathrm{bmg}}(x)$	$\mathcal{CI}(x)$	$\mathcal{N}(x)$
	1.43%	538	105
-10.00%	-10.00%	501	100
-20.00%	-20.00%	422	89
-40.00%	-40.00%	289	70

Source: Roncalli *et al.* (2021)

Application to the minimum variance portfolio

Managing both systematic and idiosyncratic carbon risks

Table 93: Minimum variance portfolios with a carbon intensity constraint (MSCI World Index, December 2018)

$\mathcal{CI}^\star$	$\mathcal{CI}(x)$	$\beta_{\mathrm{bmg}}(x)$	$\mathcal{N}(x)$
500	500	1.43%	105
250	250	1.37%	103
100	100	1.36%	98
50	50	1.33%	82

Source: Roncalli *et al.* (2021)

Application to the minimum variance portfolio

Managing both systematic and idiosyncratic carbon risks

⇒ it makes sense to combine the approaches by imposing two constraints:

$$\begin{cases} \mathcal{CI}(x) \leq \mathcal{CI}^{\star} \\ \beta_{\mathrm{bmg}}(x) \leq \beta_{\mathrm{bmg}}^{+} \end{cases}$$

Table 94: Minimum variance portfolios with carbon beta and intensity constraints — $\beta_{\mathrm{bmg}}^{+} = -20\%$ (MSCI World Index, December 2018)

$\mathcal{CI}^{\star}$	$\mathcal{CI}(x)$	$\beta_{\mathrm{bmg}}(x)$	$\mathcal{N}(x)$	$\mathrm{WO}(x)$
500	430	-20.00%	111	74.65%
250	250	-20.00%	86	75.26%
100	100	-20.00%	79	74.87%
50	50	-20.00%	74	74.99%

Source: Roncalli *et al.* (2021)

Application to enhanced index portfolios

Several optimization approaches

1. Max-threshold optimization solution (integration policy)
2. Order-statistic optimization solution (exclusion policy)
3. Zero-inflated optimization solution (exclusion policy)
4. Neutral-absolute optimization solution (hedging policy)

Application to enhanced index portfolios
Several optimization approaches

The generic optimization problem is:

$$
\begin{aligned}
x^{\star} &= \arg\min \frac{1}{2} (x-b)^{\top} \Sigma (x-b) \\
\text{s.t.} & \left\{
\begin{array}{l}
\mathbf{1}_n^{\top} x = 1 \\
x \geq \mathbf{0}_n \\
{\color{red} x \in \Omega}
\end{array}
\right.
\end{aligned}
$$

Application to enhanced index portfolios

Several optimization approaches

1. Max-threshold optimization solution
   - Without a benchmark

$$\Omega = \left\{ x \in \mathbb{R}^n : \beta_{\mathrm{bmg}}^\top x \leq \beta_{\mathrm{bmg}}^+ \right\}$$

   - With a benchmark

$$\Omega = \left\{ x \in \mathbb{R}^n : \beta_{\mathrm{bmg}}^\top (x - b) \leq -\Delta_{\mathrm{bmg}} \right\}$$

Application to enhanced index portfolios

Several optimization approaches

2. Order-statistic optimization solution
   This approach consists in excluding the first m stocks that present the largest carbon beta:

$$\Omega = \left\{ x \in \mathbb{R}^n : x_i = 0 \text{ if } \beta_{\mathrm{bmg},i} \geq \beta_{\mathrm{bmg}}^{(m,n)} \right\}$$

   where $\beta_{\mathrm{bmg}}^{(m,n)} = \beta_{\mathrm{bmg},n-m+1:n}$ is the $(n-m+1)$-*th* order statistic of $(\beta_{\mathrm{bmg},1}, \ldots, \beta_{\mathrm{bmg},n})$

Application to enhanced index portfolios

Several optimization approaches

3. Zero-inflated optimization solution
This approach exclude the assets with both high weight and high carbon beta:

$$\Omega = \left\{ x \in \mathbb{R}^n : x_i = 0 \text{ if } b_i \beta_{\mathrm{bmg},i} \geq \left(b \odot \beta_{\mathrm{bmg}} \right)^{(m,n)} \right\}$$

where $\left(b \odot \beta_{\mathrm{bmg}} \right)^{(m,n)} = \left(b \odot \beta_{\mathrm{bmg}} \right)_{n-m+1:n}$ is the $(n-m+1)$-*th* order statistic of the vector $\left(b_1 \beta_{\mathrm{bmg},1}, \ldots, b_n \beta_{\mathrm{bmg},n} \right)$

Application to enhanced index portfolios

Several optimization approaches

4. Neutral-absolute optimization solution
   In this approach, we consider the following constraint:

$$\Omega = \left\{ x \in \mathbb{R}^n : \left| \beta_{\mathrm{bmg}}^\top x \right| \leq |\beta|_{\mathrm{bmg}}^{+} \right\}$$

where $|\beta|_{\mathrm{bmg}}^{+}$ is the maximum sensitivity to absolute carbon risk

Application to enhanced index portfolios

Max-threshold optimization problem

- Δ_{bmg} is the difference between the benchmark's carbon risk and the portfolio's carbon risk
- $\sigma(x \mid b)$ is the tracking error
- $\mathrm{AS}(x \mid b)$ is the active share
- $\mathcal{N}_0(x \mid b)$ is the number of excluding stocks
- $\mathrm{WACI}(x)$ is the weighted average carbon intensity

Application to enhanced index portfolios

Max-threshold optimization problem

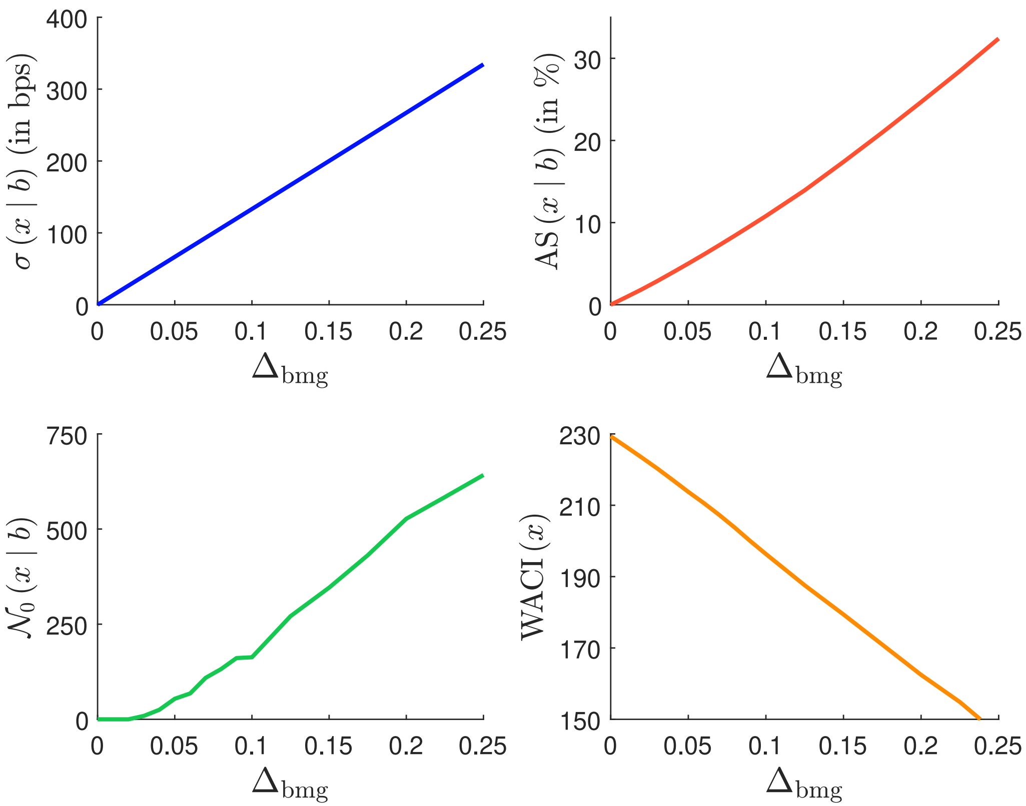


Figure 154: Solution of the max-threshold optimization problem (MSCI World Index, Dec. 2018)

Source: Roncalli *et al.* (2021)

Application to enhanced index portfolios

Order-statistic optimization problem

Remark

The order-statistic (or zero-inflated) optimization problem is less efficient than the max-threshold optimization problem

SRI Investment funds

- Investment vehicles
  - Mutual funds
  - ETFs
  - Mandates & dedicated funds
- Investment strategies
  - Thematic strategies (e.g. water, social, wind energy, climate, plastic, etc.)
  - ESG-tilted strategies (e.g. exclusion, negative screening, best-in-class, enhanced ESG score, controlled TE, etc.)
  - Climate strategies (e.g. low carbon, 2° alignment, activity exclusions[26], etc.)
  - Sustainability-linked securities (e.g. green bonds, social bonds, etc.)

Both α and β management

[26] e.g. coal exploration, oil exploration, electricity generation with a high GHG intensity

SRI Investment funds

Some examples

Mutual funds

- Amundi Climate Transition
- Amundi ARI European Credit SRI
- AXA World Funds – Euro Bonds SRI
- CPR Invest Social Impact
- Fidelity U.S. Sustainability Index
- Fidelity Sustainable Water & Waste
- Natixis ESG Dynamic Fund
- Vanguard FTSE Social Index
- Etc.

ETFs

- Amundi Index MSCI Europe SRI UCITS ETF
- Amundi MSCI Emerging ESG Leaders UCITS ETF
- Amundi EURO ISTOXX Climate Paris Aligned PAB UCITS ETF
- Lyxor New Energy UCITS ETF
- Lyxor World Water UCITS ETF
- SPDR S&P 500 ESG
- First Trust Global Wind Energy ETF
- Invesco S&P 500 ESG UCITS ETF
- Etc.

SRI Investment funds

Regulation

The big issue for an investor is:

How to avoid Greenwashing (& ESG washing)?

Greenwash (also greenwashing)

- Activities by a company or an organization that are intended to make people think that it is concerned about the environment, even if its real business actually harms the environment
- A common form of greenwash is to publicly claim a commitment to the environment while quietly lobbying to avoid regulation

Source: Oxford English Dictionary (2020), `https://www.oed.com`

In finance, greenwashing is understood as making misleading claims about environmental practices, performance or products

SRI Investment funds

Regulation

European sustainable finance labels

- Novethic label (pioneer label in 2009, suspended in 2016)
- French SRI label — https://www.lelabelisr.fr
- FNG label (Germany) — https://fng-siegel.org
- Towards Sustainability label (Belgium) — https://www.towardssustainability.be
- LuxFLAG label (Luxembourg) — https://www.luxflag.org
- Nordic Swan Ecolabel (Nordic countries) — https://www.nordic-ecolabel.org
- Umweltzeichen Ecolabel (Austria) — https://www.umweltzeichen.at/en
- French Greenfin label — https://www.ecologie.gouv.fr/label-greenfin

SRI Investment funds

Regulation

Remark

According to Novethic (2020), 806 funds had a label at the end of December 2019. Nine months later, this number has increased by 392 and the AUM has be multiplied by 3.2!

SRI Investment funds

Regulation

"Today it is difficult for consumers, companies and other market actors to make sense of the many environmental labels and initiatives on the environmental performance of products and companies. There are more than 200 environmental labels active in the EU, and more than 450 active worldwide; there are more than 80 widely used reporting initiatives and methods for carbon emissions only. Some of these methods and initiatives are reliable, some not; they are variable in the issues they cover" (European Commission, 2020).

Source: `https://ec.europa.eu/environment/eussd/index.htm`

SRI Investment funds

Regulation

The High Level Expert Group (HLEG) on Sustainable Finance was created in October 2016 by the European Commission

HLEG 2018 report

- Definition of a taxonomy for sustainable assets
- Inclusion of sustainability and ESG Duties of investors
- Disclosure of ESG metrics
- EU label for green investment funds
- EU standard for green bonds
- Sustainability as part of the mandates of European Supervisory Authorities (ESA)

SRI Investment funds

Regulation

ESMA

- Final report on integrating sustainability risks and factors in the UCITS Directive and the AIFMD (May 2019)
- Final report on integrating sustainability risks and factors in the MIFID II (May 2019)

Green bonds

Definition

Green bonds (or green loans/green debt instruments) are debt instruments where the proceeds will be exclusively applied to finance or re-finance, in part or in full, new and/or existing eligible green projects, and which is aligned with the four core components of the Green Bond Principles (GBP) or the Green Loan Principles.

Source: CBI (2019), `https://www.climatebonds.net`

⇒ Green bonds are "*regular*" bonds[27] aiming at funding projects with positive environmental and/or climate benefits

[27] A regular bond pays regular interest to bondholders

Green bonds

Standardization is strongly required by investors and regulators

- Green Bond Principles[28] (ICMA, 2018)
- Climate Bonds Standard (CBI)
- EU Green Bond Standard[29]
- China's Green Bond Standards[30] (PBOC, 2015)

[28] The first version is published in 2014

[29] The European Green Deal Investment Plan of 14 January 2020 announced that the European Commission will establish a GBS based on the report of the Technical Expert Group on Sustainable Finance (TEG)

[30] See CBI (2020), *China Green Bond Market 2019 Research Report*, https://www.climatebonds.net/resources/reports/china-green-bond-market-2019-research-report

Green bonds

Green Bonds Principles

Green Bonds Principles (GBP)

The 4 core components of the GBP are:

1. Use of proceeds
    1. Pollution prevention and control
    2. Biodiversity conservation
    3. Climate change adaptation
2. Process for project evaluation and selection
3. Management of proceeds
4. Reporting

https://www.icmagroup.org/sustainable-finance/the-principles-guidelines-and-handbooks

Green bonds

Green Bonds Principles

The use of proceeds includes:

- Renewable energy
- Energy efficiency
- Pollution prevention (e.g. GHG control, soil remediation, waste recycling)
- Sustainable management of living natural resources (e.g. sustainable agriculture, sustainable forestry, restoration of natural landscapes)
- Terrestrial and aquatic biodiversity conservation (e.g. protection of coastal, marine and watershed environments)
- Clean transportation
- Sustainable water management
- Climate change adaptation
- Eco-efficient products
- Green buildings

Green bonds

Green Bonds Principles

With respect to the **process for project evaluation and selection** (component 2), the issuer of a green bond should clearly communicate:

- the environmental sustainability objectives
- the eligible projects
- the related eligibility criteria

The **management of proceeds** (component 3) includes:

- The tracking of the "*balance sheet*" and the allocation of funds[31]
- An external review (not mandatory but highly recommended)

[31] The proceeds should be credited to a sub-account

Green bonds

Green Bonds Principles

The **reporting** (component 4) must be based on the following pillars:

- Transparency
- Description of the projects, allocated amounts and expected impacts
- Qualitative performance indicators
- Quantitative performance measures (e.g. energy capacity, electricity generation, GHG emissions reduced/avoided, number of people provided with access to clean power, decrease in water use, reduction in the number of cars required)

Types of debt instruments

Asset-linked bond structures

- Regular bond
- Revenue bond
- Project bond
- Green loans

Asset-backed bond structures

- Securitized bond
- Project bond
- ABS/MBS/CLO/CDO
- Covered bond

The green bond market

- Solar bond by the City of San Francisco in 2001
- Equity-linked climate awareness bond by the European Investment Bank (EIB) in 2007
- First green bond issued by the World Bank (in collaboration with Skandinaviska Enskilda Banken) in November 2008

The green bond market

Green bond issuers

- Sovereigns (agencies, municipals, governments)
- Multilateral development banks (MDB)
- Energy and utility companies
- Banks
- Other corporates

Green bond investors

- Pension funds
- Sovereign wealth funds
- Insurance companies
- Asset managers
- Retail investors (e.g. employee savings plans)

Strong imbalance between supply and demand

The green bond market

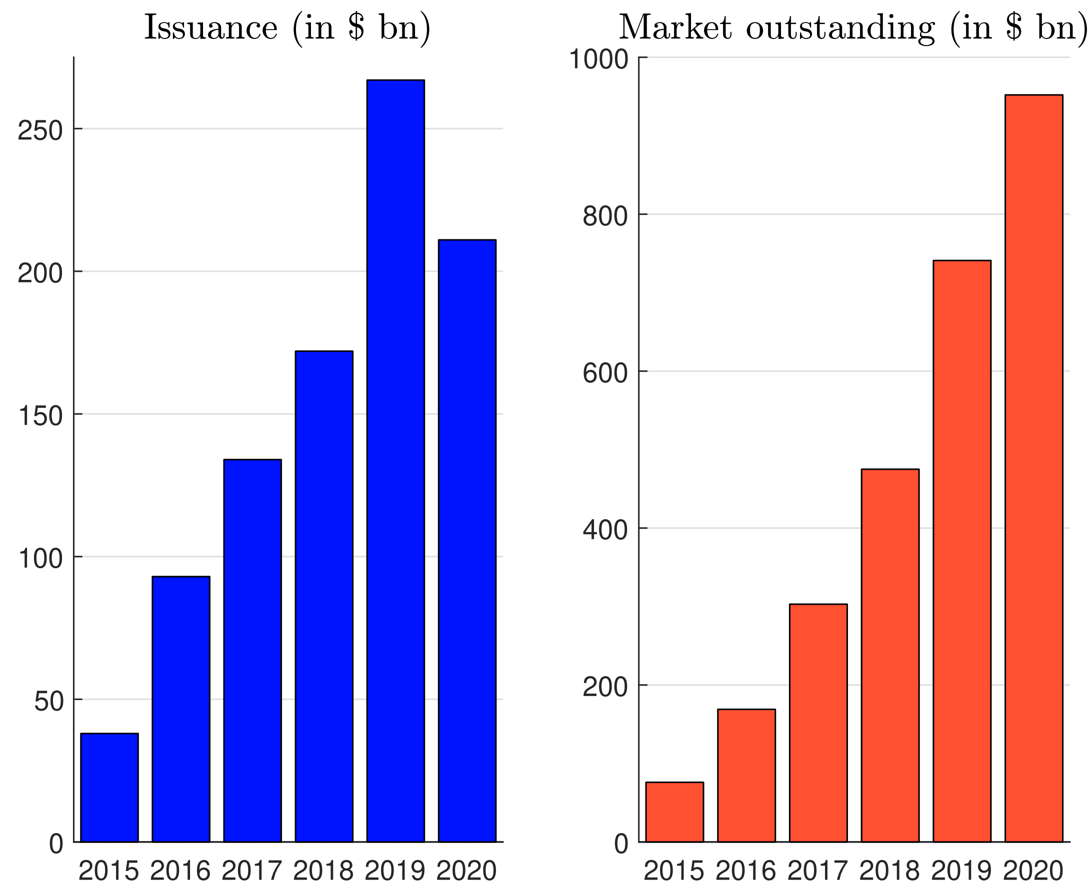


Figure 155: The green bond market

Source: CBI (2020), `https://www.climatebonds.net/market`

The green bond market

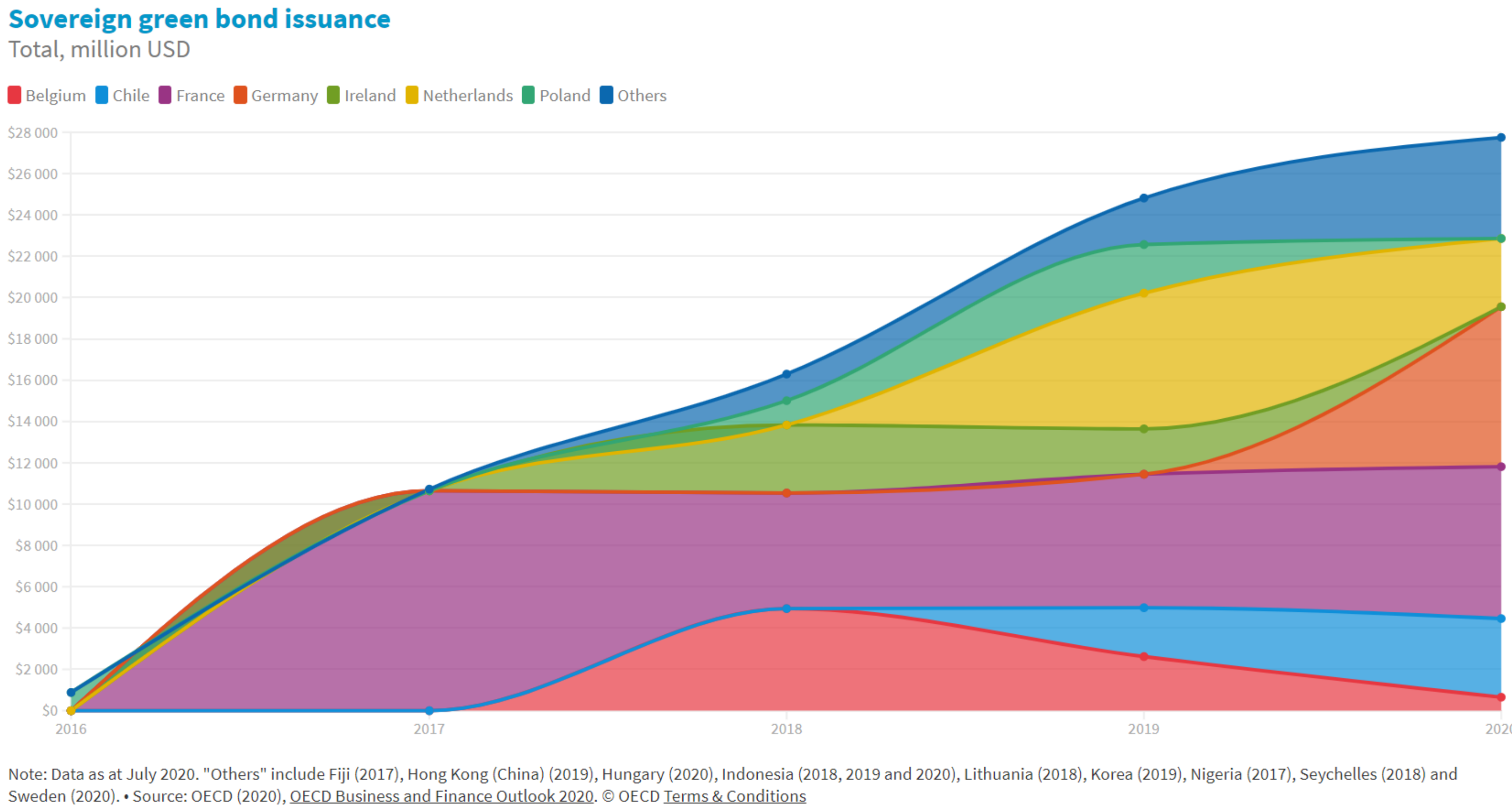


Figure 156: Growing momentum for sovereign green bonds (OECD, Sep. 2020)

Investing in green bonds
Active management

Example of green bond funds:

- Amundi Planet Emerging Green One (EGO), in collaboration with IFC (World Bank)
- Amundi ARI Impact Green Bonds
- AXA WF Global Green Bonds
- BNP Paribas Green Bond
- Mirova Global Green Bond Fund
- Etc.

Investing in green bonds

Passive management

List of green bond indices:

- Bloomberg Barclays MSCI Global Green Bond Index
- S&P Green Bond Index
- Solactive Green Bond Index
- ChinaBond China Climate-Aligned Bond Index:
- ICE BofA Green Index

$\Rightarrow$ ETF and index funds (e.g. Lyxor Green Bond UCITS ETF, iShares Green Bond Index Fund)

The green bond premium

Definition

The green bond premium (or greenium) is the difference in pricing between green bonds and regular bonds

The green bond premium

The greenium debate is a hot topic

You can read the article of the Wall Street Journal written by Matt Wirz[32]:

Why Going Green Saves Bond Borrowers Money

[32] The article is available on the following webpage: https://www.wsj.com/articles/why-going-green-saves-bond-borrowers-money-11608201002

The green bond premium

Table 95: Overview of GB pricing

Study	Market	#GBs	Universe	Period	Method	Greenium
Bachelet *et al.*(2019)	Secondary	89	Global	2013 - 2017	OLS model	2.1/5.9
Bour (2019)	Secondary	95	Global	2014 - 2018	Fixed effects model	−23.2
Ehlers and Packer (2017)	Primary	21	EUR & USD	2014 - 2017	Yield comparison	−18
Fatica *et al.* (2019)	Primary	1 397	Global	2007 - 2018	OLS model	
Hachenberg and Sciereck (2018)	Secondary	63	Global	August 2016	Panel data regression	NS
Hyun *et al* (2020)	Secondary	60	Global	2010 - 2017	Fixed effects GLS model	NS
Karpf and Mandel (2018)	Secondary	1 880	US Municipals	2010 - 2016	Oaxaca-Blinder decomposition	+7.8
Larcker and Watts (2019)	Secondary	640	US Municipals	2013 - 2018	Matching & Yield comparison	NS
Lau *et al.* (2020)	Secondary	267	Global	2013 - 2017	Two-way Fixed effects model	−1.2
Nanayakkara and Colombage (2019)	Secondary	43	Global	2016 - 2017	Panel data with hybrid model	−62.7
Ostlund (2015)	Secondary	28	Global	2011 - 2015	Yield comparison	NS
Preclaw and Bakshi (2015)	Secondary	Index	Global	2014 - 2015	OLS model	−16.7
Schmitt (2017)	Secondary	160	Global	2015 - 2017	Fixed effects model	−3.2
Zerbib (2019)	Secondary	110	Global	2013 - 2017	Fixed effects model	−1.8
Baker *et al.* (2018)	Secondary	2 083	US Municipals	2010 - 2016	OLS model	−7.6/−5.5
		19	US Corporates	2014 - 2016		
Gianfrate and Peri (2019)	Primary	121	EUR	2013 - 2017	Propensity score matching	−18
	Secondary	70/118		3 dates in 2017		−11/−5
Kapraun and Scheins (2019)	Primary	1 513	Global	2009 - 2018	Fixed effects model	−18
	Secondary	769				+10
Partridge and Medda (2018)	Primary	521	US Municipals	2013 - 2018	Yield curve analysis	−4
	Secondary					NS

Source: Ben Slimane *et al.* (2020)

The green bond premium

- From the issuer's point of view, a green bond issuance is more expensive than a conventional issuance due to the need for external review, regular reporting and impact assessments
- From the investor's point of view, there is no fundamental difference between a green bond and a conventional bond, meaning that one should consider a negative green bond premium as a market anomaly

The green bond premium

Ben Slimane *et al.* (2020) test two approaches:

1. Top-down approach
   - Compare a green bond index portfolio to a conventional bond index portfolio
   - Same characteristics in terms of currency, sector, credit quality and maturity
2. Bottom-up approach
   - Compares the green bond of an issuer with a synthetic conventional bond of the same issuer
   - Same characteristics in terms of currency, seniority and duration.

The green bond premium

Main result (Ben Slimane *et al.*, 2020)

The greenium is negative between -5 and -2 bps on average

Other results:

- Differences between sectors, currencies, maturities, regions and ratings
- Transatlantic divided between US and Europe
- The volatility of green bond portfolios are lower than the volatility of conventional bond portfolios $\Rightarrow$ identical Sharpe ratio since the last four years
- Time-varying property of the greenium

The green bond premium

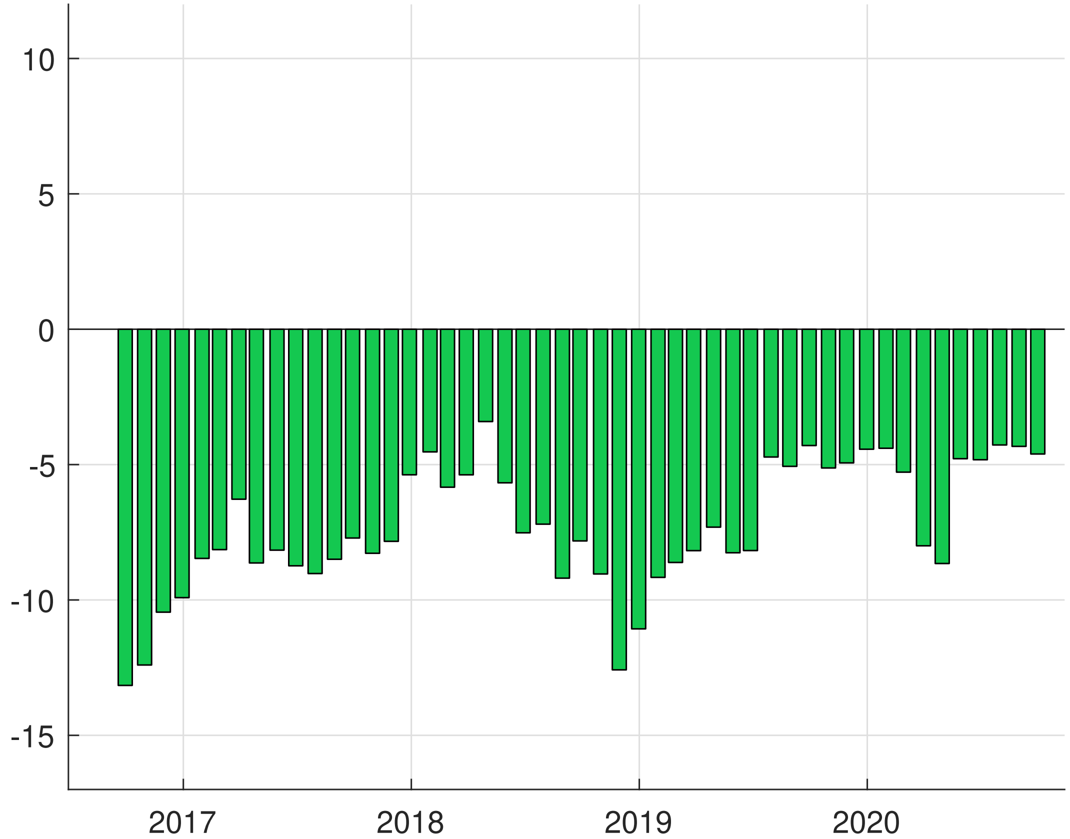


Figure 157: Evolution of the EUR greenium

Source: Ben Slimane *et al.* (2020)

The green bond premium

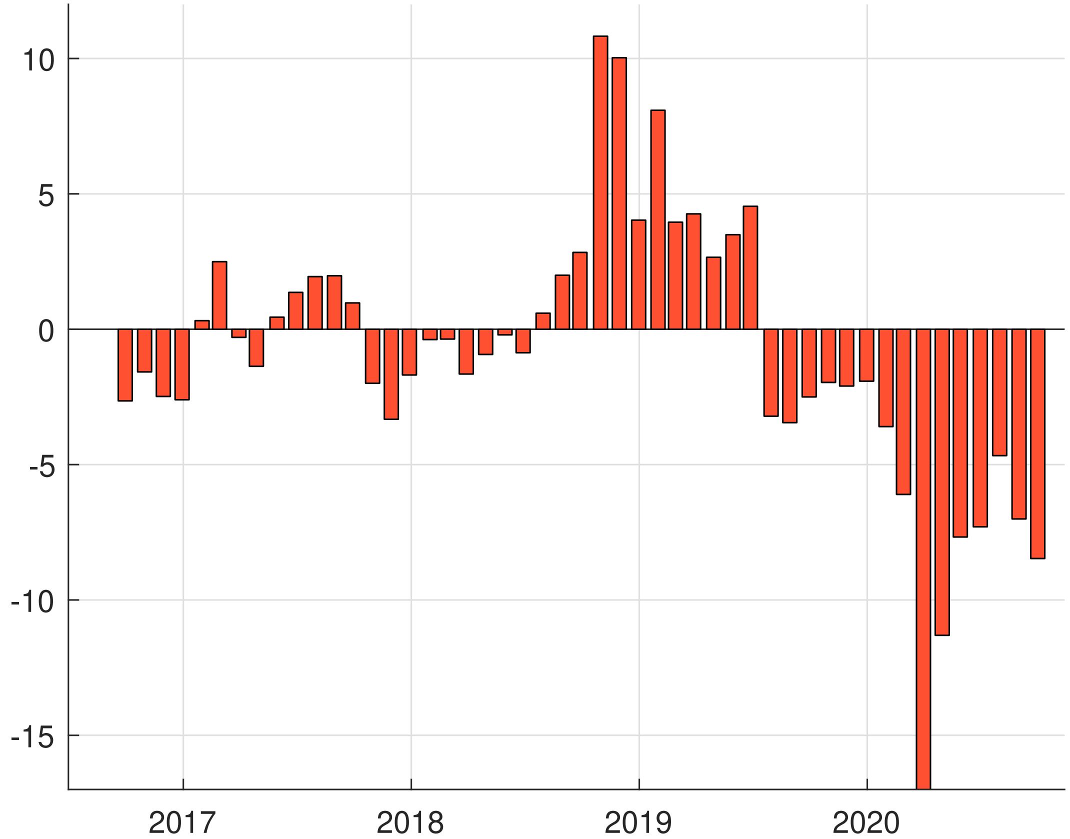


Figure 158: Evolution of the USD greenium

Source: Ben Slimane *et al.* (2020)

The green bond premium

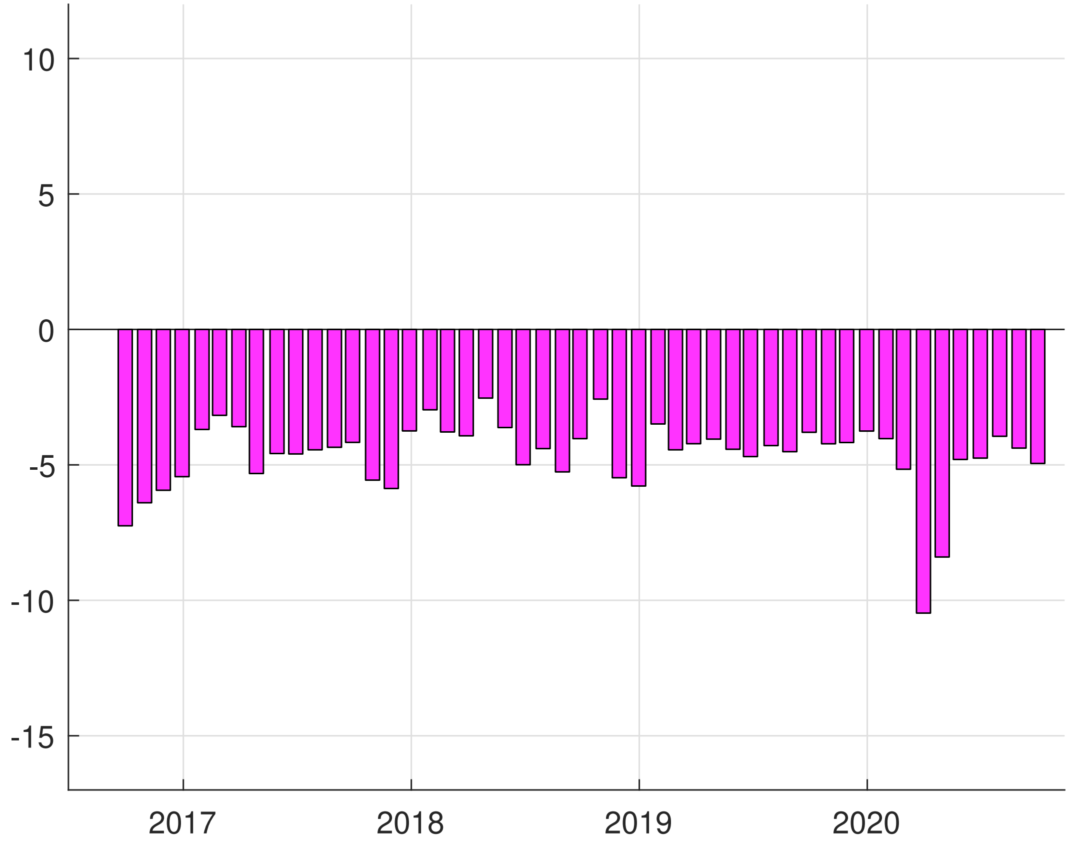


Figure 159: Evolution of the green bond premium (all currencies)

Source: Ben Slimane *et al.* (2020)

The green bond premium

Green financing $\Leftrightarrow$ green investing

1. Bond issuers have a competitive advantage to finance their environmental projects using green bonds instead of conventional bonds
2. Another premium? the "green bond issuer premium"

Social bonds

Definition

Social Bonds are any type of bond instrument where the proceeds will be exclusively applied to finance or re-finance in part or in full new and/or existing eligible Social Projects and which are aligned with the four core components of the Social Bonds Principles (SBP).

Source: ICMA (2020), `https://www.icmagroup.org/sustainable-finance`

Social bonds

Social Bonds Principles

Social Bonds Principles (SBP)

The 4 core components of the SBP are:

1. Use of proceeds
    1. Eligible social project categories
    2. Target populations
2. Process for project evaluation and selection
3. Management of proceeds
4. Reporting

https://www.icmagroup.org/sustainable-finance/the-principles-guidelines-and-handbooks

Social bonds

Social Bonds Principles

The **eligible social projects categories** (component 1) are:

- Affordable basic infrastructure (e.g. clean drinking water, sanitation, clean energy)
- Access to essential services (e.g. health, education)
- Affordable housing (e.g. sustainable cities)
- Employment generation (e.g. pandemic crisis)
- Food security and sustainable food systems (e.g. nutritious and sufficient food, resilient agriculture)
- Socioeconomic advancement and empowerment (e.g. income inequality, gender inequality)
- Etc.

Social bonds

Social Bonds Principles

The **target populations** (component 1) are:

- Living below the poverty line
- Excluded and/or marginalised populations/communities
- People with disabilities
- Migrants and /or displaced persons
- Undereducated
- Unemployed
- Women and/or sexual and gender minorities
- Aging populations and vulnerable youth
- Etc.

Social bonds

Social Bonds Principles

With respect to the **process for project evaluation and selection** (component 2), the issuer of a social bond should clearly communicate:

- the social objectives
- the eligible projects
- the related eligibility criteria

The **management of proceeds** (component 3) includes:

- The tracking of the "*balance sheet*" and the allocation of funds[33]
- An external review (not mandatory but highly recommended)

[33] The proceeds should be credited to a sub-account

Social bonds

Social Bonds Principles

The **reporting** (component 4) must be based on the following pillars:

- Transparency
- Description of the projects, allocated amounts and expected impacts
- Qualitative performance indicators
- Quantitative performance measures (e.g. number of beneficiaries)

Social bonds

Examples

You can download the *Green, Social and Sustainability bonds database* at the following webpage:

`https://www.icmagroup.org/sustainable-finance/green-social-and-sustainability-bonds-database`

You can download the market information template of the social project "*Women's Livelihood Bond 2 (WLB 2) — Singapore*" at the following address:

`https://www.icmagroup.org/Emails/icma-vcards/WLB2_Market%20Information%20Template.pdf`

The social bond market

- The tremendous growth of the social bond market

  "Of the $\$1,280$ bn in cumulative sustainable fixed-income issuance, social bonds account for around 14% of the total, amounting to \$180bn [...] This overall expansion trend has intensified during the pandemic. In fact, the growth of the social bond market in 2020, i.e. $+374\%$ with respect to 2019 levels, dwarf both the green and sustainability bonds markets' expansion, respectively $+37\%$ and $+100\%$" (Laugel and Vic-Philippe, 2020)

- The pandemic has increased the popularity of social bonds
- Investors focus more on the **S** pillar of ESG

Other sustainability-linked strategies

- Sustainable bonds
- Sustainable loans
- Green notes
- Green ABCP notes
- Financing renewables
- Green infrastructure funds
- ESG private equity funds
- Etc.

Definition

Definition

The key elements of impact investing are:

1. Intentionality
   The intention of an investor to generate a positive and measurable social and environmental impact
2. Additionality
   Fulfilling a positive impact beyond the provision of private capital
3. **Measurement**
   Being able to account for in a transparent way on the financial, social and environmental performance of investments

Source: Eurosif (2019)

The investor must be able to measure its impact from a quantitative point of view

GIIN

Figure 160: Global Impact Investing Network (GIIN)

https://thegiin.org

The example of social impact bonds

Social impact bond (SIB) = pay-for-success bond ($\approx$ call option)

The Peterborough SIB

- On 18 March 2010, the UK Secretary of State for Justice announced a six-year SIB pilot scheme that will see around 3 000 short term prisoners from Peterborough prison, serving less than 12 months, receiving intensive interventions both in prison and in the community
- Funding from investors will be initially used to pay for the services
- If reoffending is not reduced by at least 7.5%, the investors will receive no recompense

Measurement tools

Impact assessment and metrics

- Avoided CO_2 emissions in tons per $M invested
- Amount of clean water produced by the project
- Number of children who are less obese

Sustainable development goals (SDG)

The sustainable development goals are a collection of 17 interlinked global goals designed to be a "*blueprint to achieve a better and more sustainable future for all*"

https://sdgs.un.org

Sustainable development goals (SDG)

Figure 161: The map of sustainable development goals

Sustainable development goals (SDG)

Figure 162: Mapping the SDGs across E, S and G

Sustainable development goals (SDG)

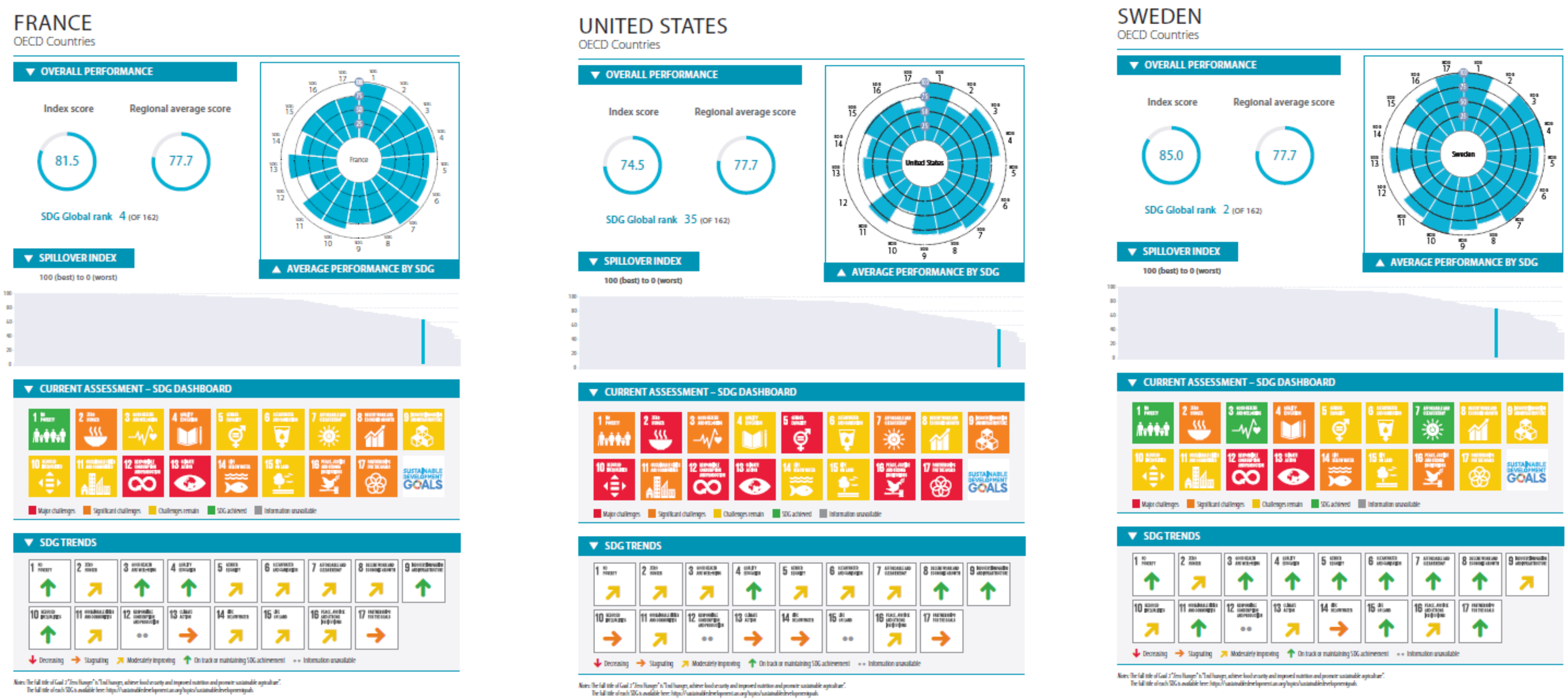


Figure 163: Examples of sovereign SDG reports

Source: Sustainable Development Report 2019, `https://dashboards.sdgindex.org`

Shareholder activism

Shareholder activism can take various forms

1. Exit (sell shares, take an offsetting bet)
2. Vote (form coalition/express dissent/call back lent shares)
3. Engage behind the scene with management and the board
4. Voice displeasure publicly (in the media)
5. Propose resolutions (shareholder proposals)
6. Initiate a takeover (acquire a sizable equity share)

Source: Bekjarovski and Brière (2018)

ESG engagement policies

- On-going engagement
  - Meet companies in order to better understand sectorial ESG challenges
  - Encourage companies to adopt best ESG practices
  - Challenge companies on ESG risks
- Engagement for influence
  - Make recommendations
  - Measure companies ESG progress
- AGM[34] engagement
  - Exercise on voting rights
  - Discuss with companies any resolution items that the investor may vote against

[34] Annual General Meeting

The challenge of reporting

- Impact reporting and investment standards (IRIS) proposed by GIIN
- EU taxonomy on sustainable finance
- Non-financial reporting directive 2014/95/EU (NFRD)
- Carbon accounting

Main references I

- Bennani, L., Le Guenedal, T., Lepetit, F., Ly, L., Mortier, V., Roncalli, T., and Sekine, T. (2018)
  How ESG Investing Has Impacted the Asset Pricing in the Equity Market, Amundi Discussion Paper, DP-39-2018,
  www.research-center.amundi.com.
- Bouchet, V., and Le Guenedal, T. (2020)
  Credit Risk Sensitivity to Carbon Price, *SSRN*,
  https://www.ssrn.com/abstract=3574486.
- Eurosif (2019)
  European SRI Study 2018, http://www.eurosif.org.
- GSIA (2019)
  Global Sustainable Investment Review 2018,
  http://www.gsi-alliance.org.

Main references II

- LAUGEL, E., and VIC-PHILIPPE, I. (2020)
  Social Bonds: Financing the Recovery and Long-Term Inclusive Growth, *Amundi Insights Paper*, www.research-center.amundi.com.

- LE GUENEDAL, T. (2019)
  Economic Modeling of Climate Risks, *Amundi Working Paper*, WP-83-2019, www.research-center.amundi.com.

- RONCALLI, T., LE GUENEDAL, T., LEPETIT, F., RONCALLI, T., and SEKINE, T. (2020)
  Measuring and Managing Carbon Risk in Investment Portfolios, *Amundi Working Paper*, WP-99-2020, www.research-center.amundi.com.

References I

- ANDERSSON, M., BOLTON, P., and SAMAMA, F. (2016)
  Hedging Climate Risk, *Financial Analysts Journal*, 72(3), pp. 13-32.
- BEKJAROVSKI, F., and BRIÈRE, M. (2018)
  Shareholder Activism: Why Should Investors Care?, *Amundi Discussion Paper*, DP-30-2018, www.research-center.amundi.com.
- BEN SLIMANE, M., DA FONSECA, D., and MAHTANI, V. (2020)
  Facts and Fantasies About the Green Bond Premium, *Amundi Working Paper*, WP-102-2020, www.research-center.amundi.com.
- BEN SLIMANE, M., LE GUENEDAL, T., RONCALLI, T., and SEKINE, T. (2020)
  ESG Investing in Corporate Bonds: Mind the Gap, *Amundi Working Paper*, WP-94-2019, www.research-center.amundi.com.

References II

Ben Slimane, M., Brard, E., Le Guenedal, T., Roncalli, T., and Sekine, T. (2020)
ESG Investing in Fixed Income: It's Time To Cross the Rubicon, *Amundi Discussion Paper*, DP-45-2019, www.research-center.amundi.com.

Bolton, P., Despres, M., Pereira Da Silva, L.A., Samama, F. and Svartzman, R. (2020)
The Green Swan — Central Banking and Financial Stability in the Age of Climate Change, BIS Publication, https://www.bis.org/publ/othp31.htm.

Climate Bonds Initiative — CBI (2019)
Climate Bonds Standard — Version 3.0, December 2019

References III

COOPER, C., GRAHAM, C., and HIMICK, D. (2016)
Social Impact Bonds: The Securitization of the Homeless, *Accounting, Organizations and Society*, 55, pp. 63-82.

DREI, A., LE GUENEDAL, T., LEPETIT, F., MORTIER, V., RONCALLI, T., and SEKINE, T. (2020)
ESG Investing in Recent Years: New Insights from Old Challenges, *Amundi Discussion Paper*, DP-42-2019, `www.research-center.amundi.com`.

FREEMAN, R.E. (2010)
Strategic Management: A Stakeholder Approach, Cambridge University Press.

References IV

Friedman, M. (2007)
The Social Responsibility of Business Is to Increase Its Profits, in Zimmerli, W.C., Holzinger, M., and Richter, K. (Eds), *Corporate Ethics and Corporate Governance*, Springer, pp. 173-178 (written in 1970 for the New York Times).

Görgen, M., Jacob, A., Nerlinger, M., Riordan, R., Rohleder, M., and Wilkens, M. (2019)
Carbon Risk, *SSRN*, https://www.ssrn.com/abstract=2930897.

GSIA (2017)
Global Sustainable Investment Review 2016, http://www.gsi-alliance.org.

International Capital Market Association — ICMA (2020)
Social Bond Principles — Voluntary Process Guidelines for Issuing Social Bonds, June 2020.

References V

- International Capital Market Association — ICMA (2018)
  Geen Bond Principles — Voluntary Process Guidelines for Issuing Social Bonds, June 2018.
- Le Guenedal, T., Girault, J., Jouanneau, M., Lepetit, F., and Sekine, T. (2020)
  Trajectory Monitoring in Portfolio Management and Issuer intentionality Scoring, *Amundi Working Paper*, 97, www.research-center.amundi.com.
- MSCI (2020)
  MSCI ESG Ratings Methodology, MSCI ESG Research, April 2020.
- Nordhaus, W.D. (1993)
  Rolling the DICE: An Optimal Transition Path for Controlling Greenhouse, *Resource and Energy Economics*, 15, pp. 27-50.

References VI

O'NEILL, B.C., KRIEGLER, E., EBI, K.L., *et al.* (2017)
The Roads Ahead: Narratives for Shared Socioeconomic Pathways Describing World Futures in the 21st Century, *Global Environmental Change*, 42, pp. 169-180.

PINDYCK, R.S. (2017)
The Use and Misuse of Models for Climate Policy, *Review of Environmental Economics and Policy*, 11(1), pp. 100-114.

RONCALLI, T. (2017)
ESG & Factor Investing: A New Stage Has Been Reached, *Amundi ViewPoint*, www.research-center.amundi.com.

References VII

- Roncalli, T., Le Guenedal, T., Lepetit, F., Roncalli, T., and Sekine, T. (2021)
The Market Measure of Carbon Risk and its Impact on the Minimum Variance Portfolio, *Amundi Working Paper*, WP-105-2021, www.research-center.amundi.com.

- So, I., and Staskevicius, A. (2015)
Measuring the 'Impact' in Impact Investing, *Harvard Business School Report*.

- Stern, N. (2007)
The Economics of Climate Change: The Stern Review, Cambridge University Press.

- Tortorice, D.L., Bloom, D.E., Kirby, P., and Regan, J. (2020)
A Theory of Social Impact Bonds, *NBER Working Paper*, w27527.

Tutorial Exercises

Tutorial exercise 1

Probability distribution of an ESG score

Question 1

We consider an investment universe of 8 issuers with the following ESG scores:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
E	−2.80	−1.80	−1.75	0.60	0.75	1.30	1.90	2.70
S	−1.70	−1.90	0.75	−1.60	1.85	1.05	0.90	0.70
G	0.30	−0.70	−2.75	2.60	0.45	2.35	2.20	1.70

Tutorial exercise 1

Probability distribution of an ESG score

Question 1.a

Calculate the ESG score of the issuers if we assume the following weighting scheme: 40% for E, 40% for S and 20% for G.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$s_i^{(\mathrm{ESG})} = 0.4 \times s_i^{(\mathrm{E})} + 0.4 \times s_i^{(\mathrm{S})} + 0.2 \times s_i^{(\mathrm{G})}$$

- We obtain the following results:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
$s_i^{(\mathrm{E})}$	−2.80	−1.80	−1.75	0.60	0.75	1.30	1.90	2.70
$s_i^{(\mathrm{S})}$	−1.70	−1.90	0.75	−1.60	1.85	1.05	0.90	0.70
$s_i^{(\mathrm{G})}$	0.30	−0.70	−2.75	2.60	0.45	2.35	2.20	1.70
$s_i^{(\mathrm{ESG})}$	−1.74	−1.62	−0.95	0.12	1.13	1.41	1.56	1.70

Tutorial exercise 1

Probability distribution of an ESG score

Question 1.b

Calculate the ESG score of the equally-weighted portfolio x_{ew}.

Tutorial exercise 1

Probability distribution of an ESG score

- We obtain:

$$\begin{aligned} \mathcal{S}^{(\mathrm{ESG})}(x_{\mathrm{ew}}) &= \sum_{i=1}^{8} x_{\mathrm{ew},i} \times \mathcal{S}_i^{(\mathrm{ESG})} \\ &= 0.2013 \end{aligned}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 2

We assume that the ESG scores are *iid* and follow a standard Gaussian distribution:

$$s_i \sim \mathcal{N}(0, 1)$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 2.a

We note $x_{\mathrm{ew}}^{(n)}$ the equally-weighted portfolio composed of n issuers.
Calculate the distribution of the ESG score $\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right)$ of the portfolio $x_{\mathrm{ew}}^{(n)}$.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$\begin{aligned} \mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right) &= \sum_{i=1}^{n} x_{\mathrm{ew},i}^{(n)} \times \mathcal{S}_i \\ &= \frac{1}{n} \sum_{i=1}^{n} \mathcal{S}_i \end{aligned}$$

We deduce that $\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right)$ follows a Gaussian distribution.

Tutorial exercise 1
Probability distribution of an ESG score

- Its mean is equal to:

$$\mathbb{E}\left[\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right)\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\mathcal{S}_i\right] = 0$$

- Its standard deviation is equal to:

$$\begin{aligned}\sigma\left(\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right)\right) &= \sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2\left(\mathcal{S}_i\right)} \\ &= \frac{1}{\sqrt{n}}\end{aligned}$$

- Finally, we obtain:

$$\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right) \sim \mathcal{N}\left(0, \frac{1}{n}\right)$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 2.b

What is the ESG score of a well-diversified portfolio?

Tutorial exercise 1

Probability distribution of an ESG score

- The behavior of a well-diversified portfolio is close to an equally-weighted portfolio with n sufficiently large. Therefore, the ESG score is close to zero because we have:

$$\lim_{n \to \infty} \mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right) = 0$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 2.c

We note $T \sim \mathbf{F}_{\alpha}$ where $\mathbf{F}_{\alpha}(t) = t^{\alpha}$, $t \in [0, 1]$ and $\alpha \geq 0$. Draw the graph of the probability density function $f_{\alpha}(t)$ when α is respectively equal to 0.5, 1.5, 2.5 and 70. What do you notice?

Tutorial exercise 1

Probability distribution of an ESG score

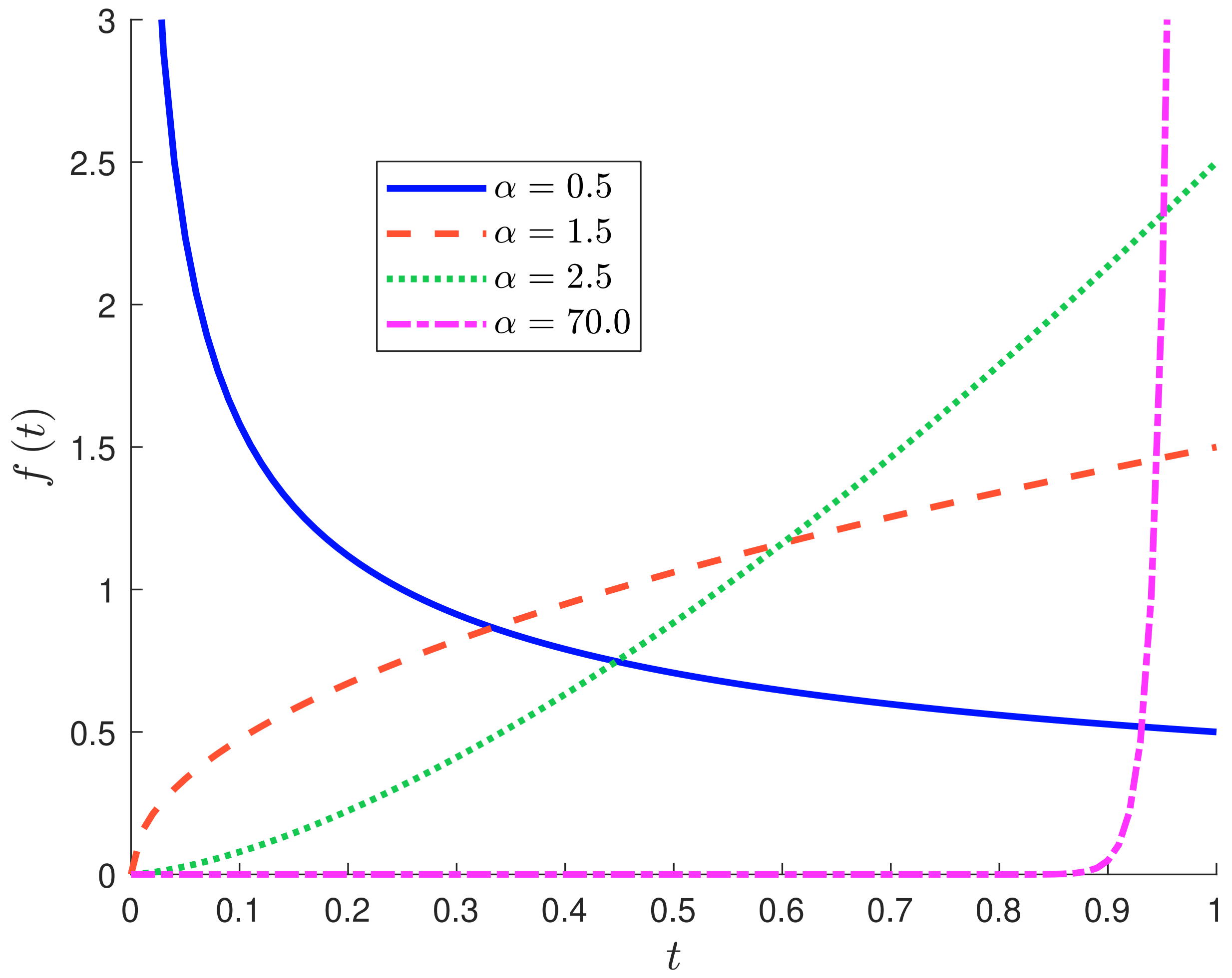


Figure 164: Probability density function $f_\alpha(t)$

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$f_{\alpha}(t) = \alpha t^{\alpha - 1}$$

- We notice that the function $f_{\alpha}(t)$ tends to the dirac delta function when α tends to infinity:

$$\lim_{\alpha \to \infty} f_{\alpha}(t) = \delta_1(t) = \begin{cases} 0 & \text{if } t \neq 1 \\ +\infty & \text{if } t = 1 \end{cases}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 2.d

We assume that the weights of the portfolio $x = (x_1, \ldots, x_n)$ follow a power-law distribution $\mathbf{F}_\alpha$:

$$x_i \sim c T_i$$

where $T_i \sim \mathbf{F}_\alpha$ are *iid* random variables and c is a normalization constant. Explain how to simulate the portfolio weights $x = (x_1, \ldots, x_n)$. Represent one simulation of the portfolio x for the previous values of α. Comment on these results. Deduce the relationship between the Herfindahl index $\mathcal{H}_\alpha(x)$ of the portfolio weights x and the parameter α.

Remark

We use $n = 50$ in the rest of the exercise.

Tutorial exercise 1

Probability distribution of an ESG score

- To simulate T_i, we use the property of the probability integral transform:

$$U_i = \mathbf{F}_\alpha(T_i) \sim \mathcal{U}_{[0,1]}$$

We deduce that:

$$\begin{aligned} T_i &= \mathbf{F}_\alpha^{-1}(U_i) \\ &= U_i^{1/\alpha} \end{aligned}$$

Tutorial exercise 1

Probability distribution of an ESG score

The algorithm for simulating the portfolio x is then the following:

1. We simulate n independent uniform random numbers $(u_1, \ldots, u_n)$.
2. We compute the random variates $(t_1, \ldots, t_n)$ where:

$$t_i = u_i^{1/\alpha}$$

3. We calculate the normalization constant:

$$c = \left(\sum_{i=1}^{n} t_i\right)^{-1} = \left(\sum_{i=1}^{n} u_i^{1/\alpha}\right)^{-1}$$

4. We deduce the portfolio weights $x = (x_1, \ldots, x_n)$:

$$x_i = c \cdot t_i = c \cdot u_i^{1/\alpha} = \frac{u_i^{1/\alpha}}{\sum_{j=1}^{n} u_j^{1/\alpha}}$$

Tutorial exercise 1

Probability distribution of an ESG score

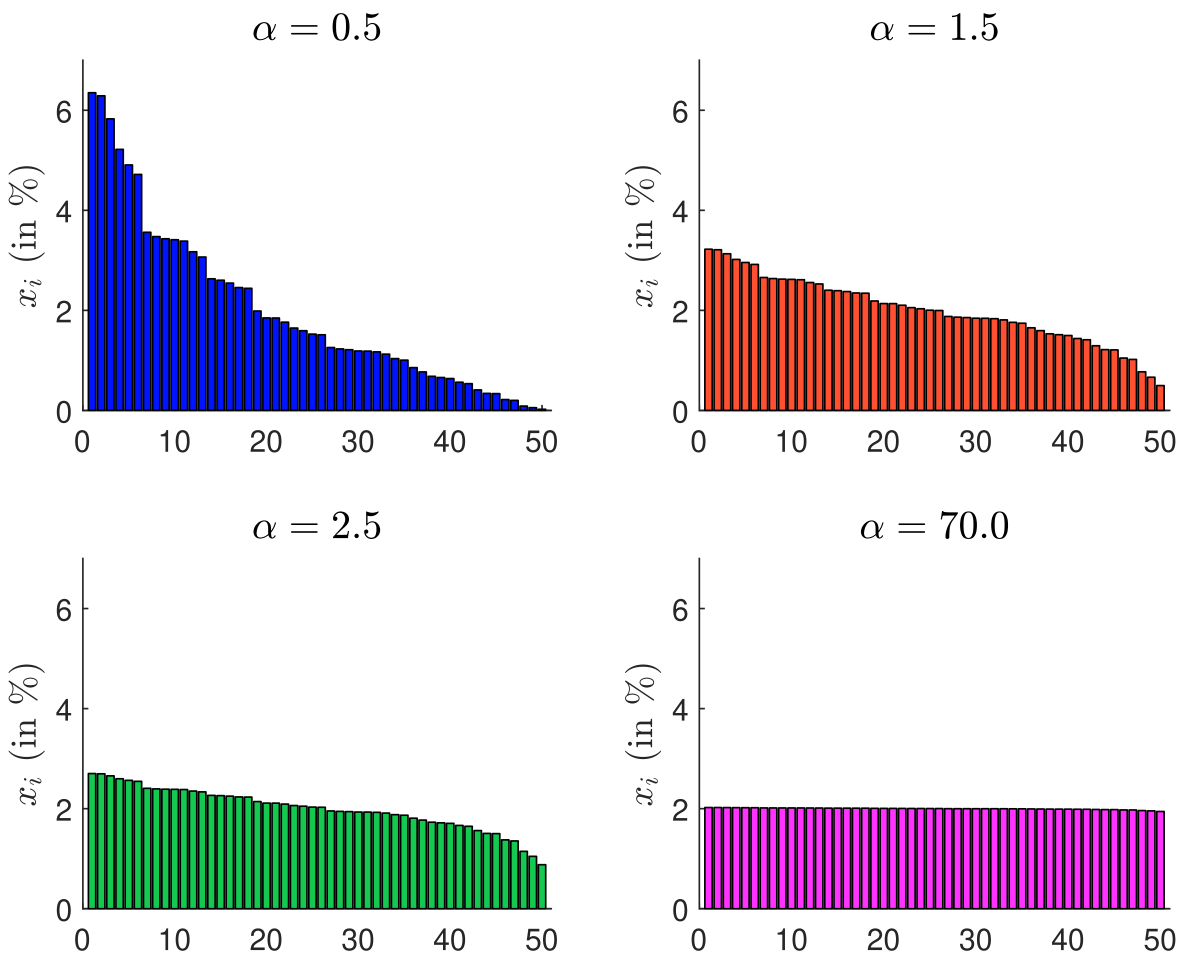


Figure 165: Repartition of the portfolio weights in descending order

Tutorial exercise 1

Probability distribution of an ESG score

- In Figure 165, we have represented the composition of the portfolio x for the 4 values of α. The weights are ranked in descending order. We deduce that the portfolio x is uniform when $\alpha \to \infty$. The parameter α controls the concentration of the portfolio. Indeed, when α is small, the portfolio is highly concentrated. It follows that the Herfindahl index $\mathcal{H}_{\alpha}(x)$ of the portfolio weights is a decreasing function of the parameter α.

Tutorial exercise 1

Probability distribution of an ESG score

Question 2.e

We assume that the weight x_i and the ESG score $\mathcal{S}_i$ of the issuer i are independent. How to simulate the portfolio ESG score $\mathcal{S}(x)$? Using 50 000 replications, estimate the probability distribution function of $\mathcal{S}(x)$ by the Monte Carlo method. Comment on these results.

Tutorial exercise 1

Probability distribution of an ESG score

- We simulate $x = (x_1, \ldots, x_n)$ using the previous algorithm. The vector of ESG scores $\mathcal{S} = (\mathcal{S}_1, \ldots, \mathcal{S}_n)$ is generated with normally-distributed random variables since we have $\mathcal{S}_i \sim \mathcal{N}(0, 1)$. We deduce that the simulated value of the portfolio ESG score $\mathcal{S}(x)$ is equal to:

$$\mathcal{S}(x) = \sum_{i=1}^{n} x_i \cdot \mathcal{S}_i$$

- We replicate the simulation of $\mathcal{S}(x)$ 50 000 times and draw the corresponding histogram in Figure 166. We also report the fitted Gaussian distribution. We observe that the portfolio ESG score $\mathcal{S}(x)$ is equal to zero on average, and its variance is an increasing function of the portfolio concentration.

Tutorial exercise 1

Probability distribution of an ESG score

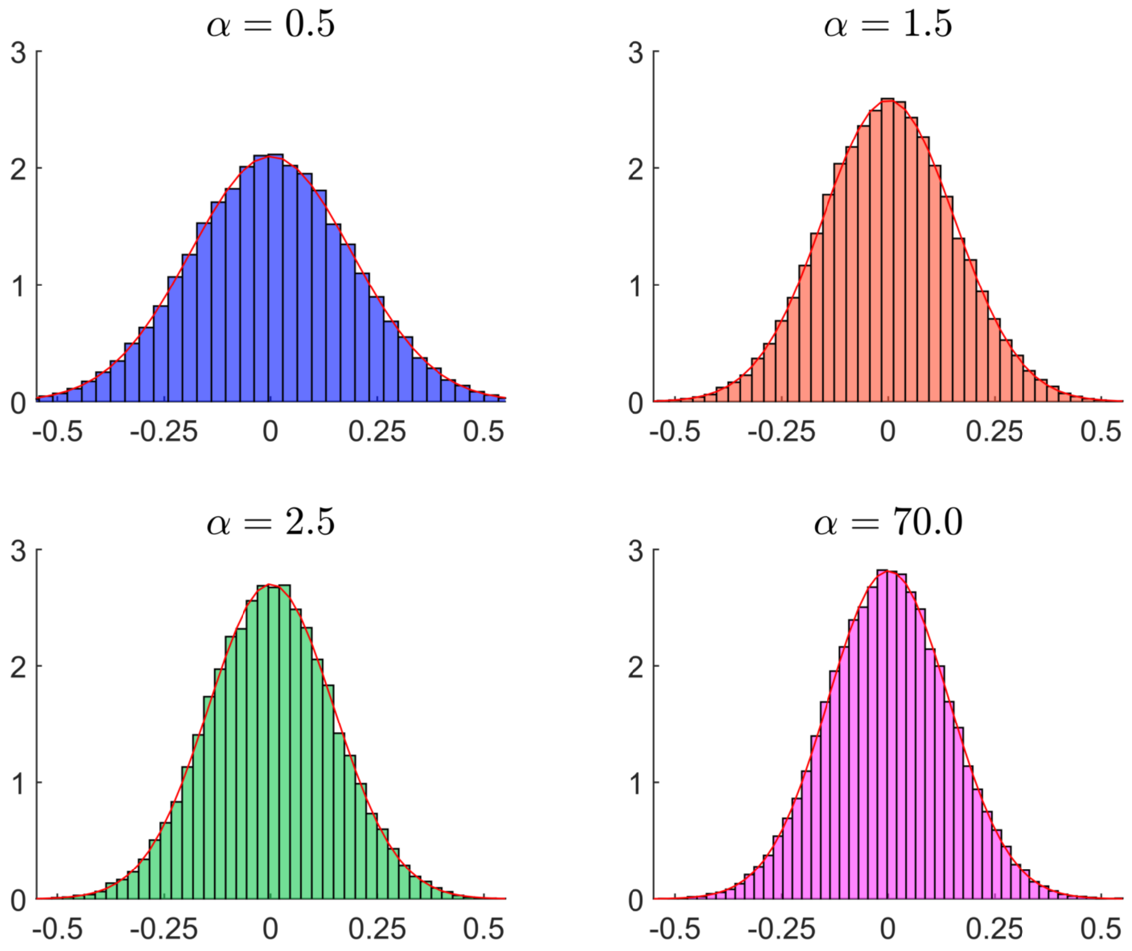


Figure 166: Histogram of the portfolio ESG score $\mathcal{S}(x)$

Tutorial exercise 1

Probability distribution of an ESG score

Question 2.f

We now assume that the weight x_i and the ESG score $\mathcal{S}_i$ of the issuer i are positively correlated. More precisely, the dependence function between x_i and $\mathcal{S}_i$ is the Normal copula function with parameter ρ. Show that this is also the copula function between T_i and $\mathcal{S}_i$. Deduce an algorithm to simulate $\mathcal{S}(x)$.

Tutorial exercise 1

Probability distribution of an ESG score

- Since $x_i \sim cT_i$, x_i is an increasing function of T_i. We deduce that the copula function of $(T_i, \mathcal{S}_i)$ is the same as the copula function of $(x_i, \mathcal{S}_i)$.
- To simulate the Normal copula function $\mathbf{C}(u, v)$, we use the transformation algorithm based on the Cholesky decomposition:

$$\left\{ \begin{array}{l} u_i = \Phi\left(g_i'\right) \\ v_i = \Phi\left(\rho g_i' + \sqrt{1-\rho^2} g_i''\right) \end{array} \right.$$

where g_i' and g_i'' are two independent random numbers from the probability distribution $\mathcal{N}(0, 1)$.

Tutorial exercise 1

Probability distribution of an ESG score

Here is the algorithm to simulate the ESG portfolio score $\mathcal{S}(x)$:

1. We simulate n independent normally-distributed random numbers g_i' and g_i'' and we compute (u_i, v_i):

$$\begin{cases} u_i = \Phi(g_i') \\ v_i = \Phi\left(\rho g_i' + \sqrt{1-\rho^2} g_i''\right) \end{cases}$$

2. We compute the random variates $(t_1, \ldots, t_n)$ where $t_i = u_i^{1/\alpha}$
3. We deduce the vector of weights $x = (x_1, \ldots, x_n)$:

$$x_i = t_i \Big/ \sum_{j=1}^n t_j$$

4. We simulate the vector of scores $\mathcal{S} = (\mathcal{S}_1, \ldots, \mathcal{S}_n)$:

$$\mathcal{S}_i = \Phi^{-1}(v_i) = \rho g_i' + \sqrt{1-\rho^2} g_i''$$

5. We calculate the portfolio score:

$$\mathcal{S}(x) = \sum_{i=1}^n x_i \cdot \mathcal{S}_i$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 2.g

Using 50 000 replications, estimate the probability distribution function of $s(x)$ by the Monte Carlo method when the correlation parameter ρ is set to 50%. Comment on these results.

Tutorial exercise 1

Probability distribution of an ESG score

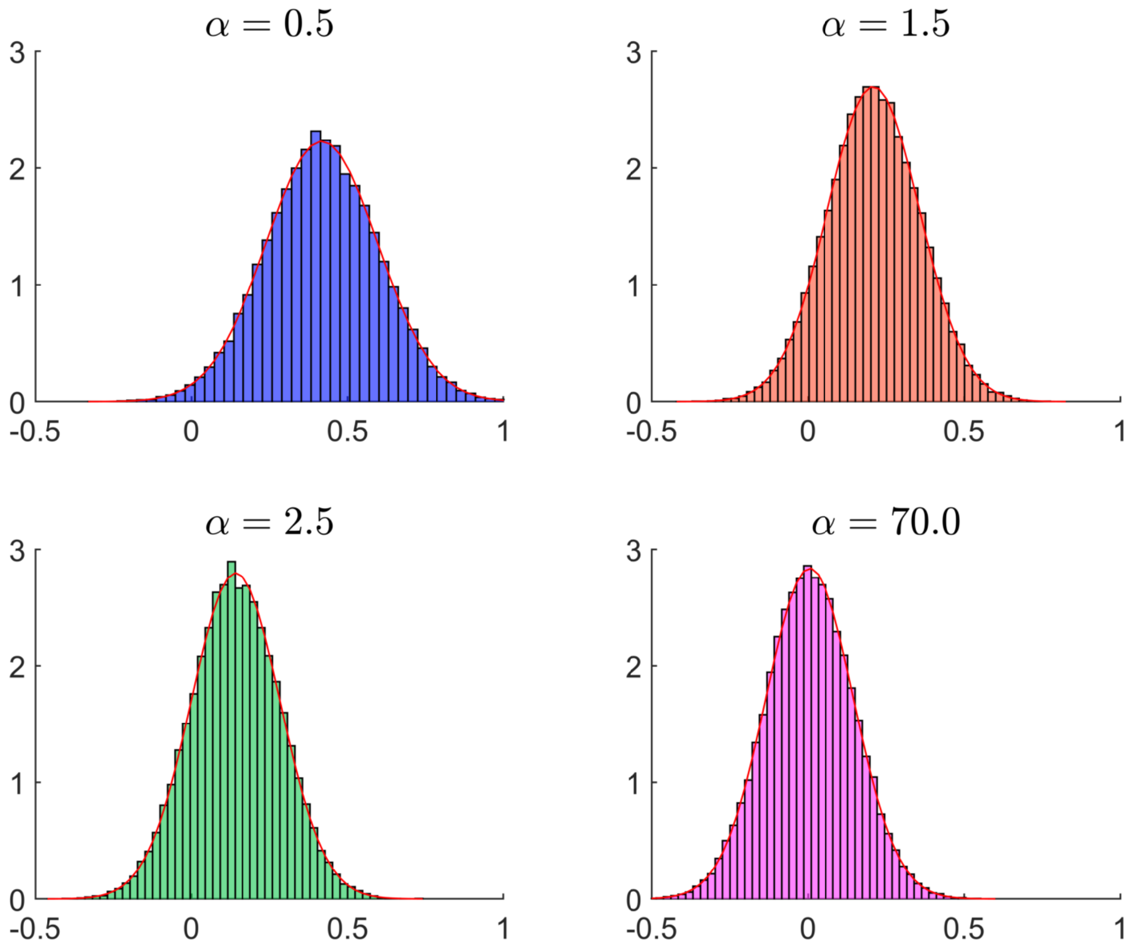


Figure 167: Histogram of the portfolio ESG score $\mathcal{S}(x)$ ($\rho = 50\%$)

Tutorial exercise 1

Probability distribution of an ESG score

- In the independent case, we found that $\mathbb{E}[s(x)] = 0$. In Figure 167, we notice that $\mathbb{E}[s(x)] \neq 0$ when ρ is equal to 50%. Indeed, we obtain:

$$\mathbb{E}[s(x)] = \begin{cases} 0.418 & \text{if } \alpha = 0.5 \\ 0.210 & \text{if } \alpha = 1.5 \\ 0.142 & \text{if } \alpha = 2.5 \\ 0.006 & \text{if } \alpha = 70.0 \end{cases}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 2.h

Estimate the relationship between the correlation parameter ρ and the expected ESG score $\mathbb{E}\left[s\left(x\right)\right]$ of the portfolio x. Comment on these results.

Tutorial exercise 1

Probability distribution of an ESG score

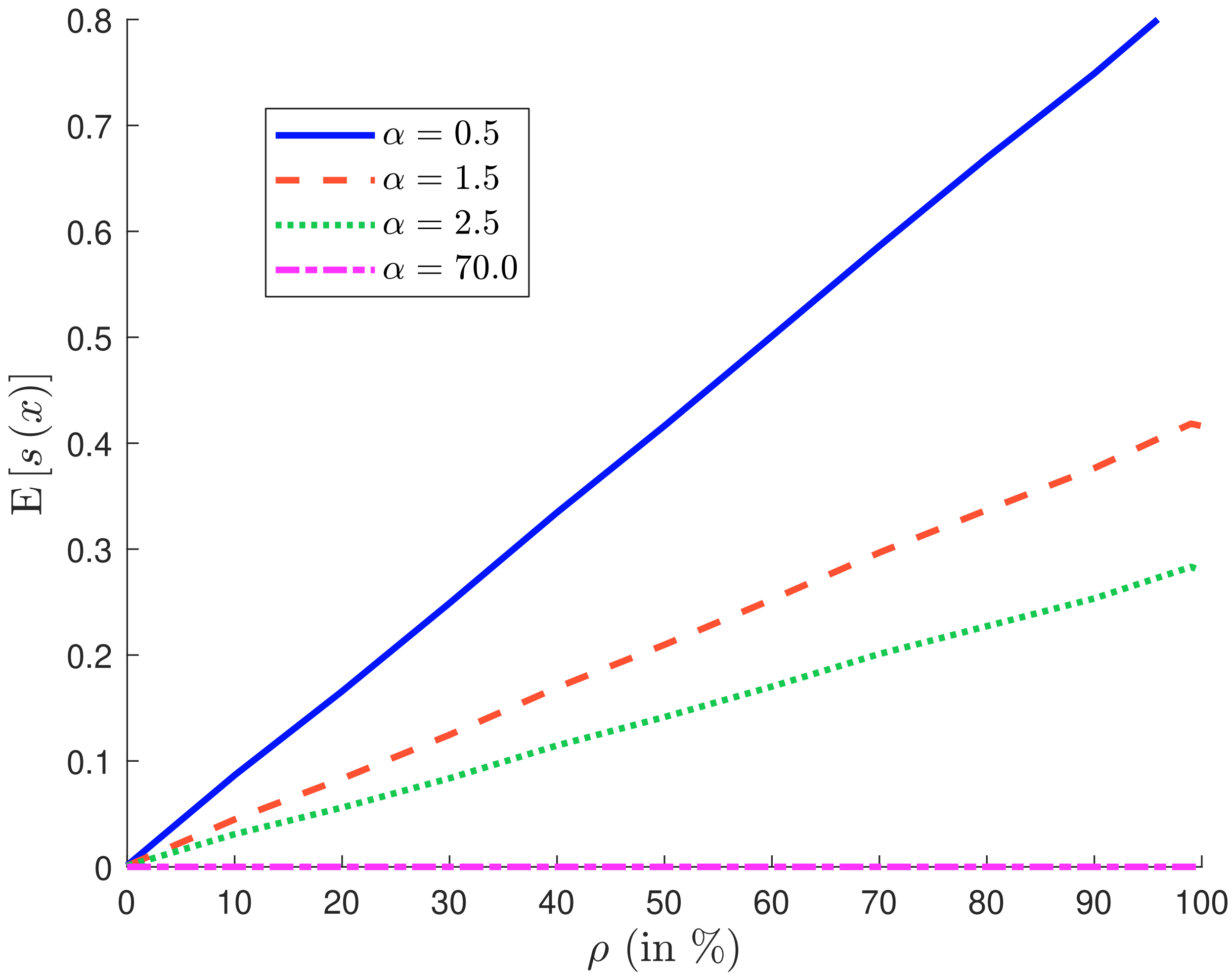


Figure 168: Relationship between ρ and $\mathbb{E}[s(x)]$

Tutorial exercise 1
Probability distribution of an ESG score

- We notice that there is a positive relationship between ρ and $\mathbb{E}[\mathcal{S}(x)]$ and the slope increases with the concentration of the portfolio.

Tutorial exercise 1

Probability distribution of an ESG score

Question 2.i

How are the previous results related to the size bias of ESG scoring?

Tutorial exercise 1

Probability distribution of an ESG score

- Big cap companies have more (financial and human) resources to develop an ESG policy than small cap companies.
- Therefore, we observe a positive correlation between the market capitalization and the ESG score of an issuer.
- It follows that ESG portfolios have generally a size bias. For instance, we generally observe that cap-weighted indexes have an ESG score which is greater than the average of ESG scores.
- In the previous questions, we verify that $\mathbb{E}[\mathcal{S}(x)] \geq \mathbb{E}[\mathcal{S}]$ when the Herfindahl index of the portfolio x is high and the correlation between x_i and $\mathcal{S}_i$ is positive.

Tutorial exercise 1

Probability distribution of an ESG score

Question 3

Let $\mathcal{S}$ be the ESG score of the issuer. We assume that the ESG score follows a standard Gaussian distribution:

$$\mathcal{S} \sim \mathcal{N}(0, 1)$$

The ESG score $\mathcal{S}$ is also converted into an ESG rating $\mathcal{R}$, which can take the values **A**, **B**, **C** and **D** — **A** is the best rating and **D** is the worst rating.

Tutorial exercise 1

Probability distribution of an ESG score

Question 3.a

We assume that the breakpoints of the rating system are -1.5, 0 and $+1.5$. Compute the frequencies of the ratings.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$\begin{aligned}\Pr\{\mathcal{R} = \mathbf{A}\} &= \Pr\{\mathcal{s} \geq 1.5\} \\ &= 1 - \Phi(1.5) \\ &= 6.68\%\end{aligned}$$

and:

$$\begin{aligned}\Pr\{\mathcal{R} = \mathbf{B}\} &= \Pr\{0 \leq \mathcal{s} < 1.5\} \\ &= \Phi(1.5) - \Phi(0) \\ &= 43.32\%\end{aligned}$$

- Since the Gaussian distribution is symmetric around 0, we also have:

$$\Pr\{\mathcal{R} = \mathbf{C}\} = \Pr\{\mathcal{R} = \mathbf{B}\} = 43.32\%$$

and:

$$\Pr\{\mathcal{R} = \mathbf{D}\} = \Pr\{\mathcal{R} = \mathbf{A}\} = 6.68\%$$

Tutorial exercise 1

Probability distribution of an ESG score

- The mapping function is:

$$\mathcal{M}_{\text{appring}}(s) = \begin{cases} \textbf{A} & \text{if } s < -1.5 \\ \textbf{B} & \text{if } -1.5 \leq s < 0 \\ \textbf{C} & \text{if } 0 \leq s < 1.5 \\ \textbf{D} & \text{if } s \geq 1.5 \end{cases}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 3.b

We would like to build a rating system such that each category has the same frequency. Find the mapping function.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

  $$\Pr\{\mathcal{R}(t) = \mathbf{A}\} = \Pr\{\mathcal{R}(t) = \mathbf{B}\} = \Pr\{\mathcal{R}(t) = \mathbf{C}\} = \Pr\{\mathcal{R}(t) = \mathbf{D}\}$$

  and:

  $$\Pr\{\mathcal{R}(t) = \mathbf{A}\} + \Pr\{\mathcal{R}(t) = \mathbf{B}\} + \Pr\{\mathcal{R}(t) = \mathbf{C}\} + \Pr\{\mathcal{R}(t) = \mathbf{D}\} = 1$$

  We deduce that:

  $$\Pr\{\mathcal{R}(t) = \mathbf{A}\} = \frac{1}{4} = 25\%$$

  and $\Pr\{\mathcal{R}(t) = \mathbf{B}\} = \Pr\{\mathcal{R}(t) = \mathbf{C}\} = \Pr\{\mathcal{R}(t) = \mathbf{D}\} = 25\%$.
- We want to find the breakpoints (s_1, s_2, s_3) such that:

  $$\left\{\begin{array}{l} \Pr\{\mathcal{S} < s_1\} = 25\% \\ \Pr\{s_1 \leq \mathcal{S} < s_2\} = 25\% \\ \Pr\{s_2 \leq \mathcal{S} < s_3\} = 25\% \\ \Pr\{\mathcal{S} \geq s_3\} = 25\% \end{array}\right.$$

Tutorial exercise 1

Probability distribution of an ESG score

- We deduce that:

$$\begin{cases} s_1 = \Phi^{-1}(0.25) = -0.6745 \\ s_2 = \Phi^{-1}(0.50) = 0 \\ s_3 = \Phi^{-1}(0.75) = +0.6745 \end{cases}$$

- The mapping function is:

$$\mathcal{M}_{\text{appring}}(s) = \begin{cases} \textbf{A} & \text{if } s < -0.6745 \\ \textbf{B} & \text{if } -0.6745 \leq s < 0 \\ \textbf{C} & \text{if } 0 \leq s < 0.6745 \\ \textbf{D} & \text{if } s \geq 0.6745 \end{cases}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 3.c

We would like to build a rating system such that the frequency of the median ratings **B** and **C** is 40% and the frequency of the extreme ratings **A** and **D** is 10%. Find the mapping function.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$\begin{cases} s_1 = \Phi^{-1}(0.10) = -1.2816 \\ s_2 = \Phi^{-1}(0.50) = 0 \\ s_3 = \Phi^{-1}(0.90) = +1.2816 \end{cases}$$

- The mapping function is:

$$\mathcal{M}_{\text{appring}}(\mathcal{s}) = \begin{cases} \textbf{A} & \text{if } \mathcal{s} < -1.2816 \\ \textbf{B} & \text{if } -1.2816 \leq \mathcal{s} < 0 \\ \textbf{C} & \text{if } 0 \leq \mathcal{s} < 1.2816 \\ \textbf{D} & \text{if } \mathcal{s} \geq 1.2816 \end{cases}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 4

Let $s(t)$ be the ESG score of the issuer at time t. The ESG scoring system is evaluated every month. The index time t corresponds to the current month, whereas the previous month is $t-1$. We assume that:

- The ESG score at time $t-1$ follows a standard Gaussian distribution:

$$s(t-1) \sim \mathcal{N}(0,1)$$

- The variation of the ESG score is Gaussian between two months:

$$\Delta s(t) = s(t) - s(t-1) \sim \mathcal{N}(0,\sigma^2)$$

- The ESG score $s(t-1)$ and the variation $\Delta s(t)$ are independent.

Tutorial exercise 1

Probability distribution of an ESG score

Question 4

The ESG score $\mathcal{S}(t)$ is converted into an ESG rating $\mathcal{R}(t)$, which can take following grades:

$$\mathcal{R}_1 \prec \mathcal{R}_2 \prec \cdots \prec \mathcal{R}_k \prec \cdots \prec \mathcal{R}_{K-1} \prec \mathcal{R}_K$$

We assume that the breakpoints of the rating system are $(s_1, s_2, \ldots, s_{K-1})$. We also note $s_0 = -\infty$ and $s_K = +\infty$.

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.a

Compute the bivariate probability distribution of the random vector $(s(t-1), \Delta s(t))$.

Tutorial exercise 1

Probability distribution of an ESG score

- The joint distribution of $(s(t-1), \Delta s(t))$ is:

$$\begin{pmatrix} s(t-1) \\ \Delta s(t) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \sigma^2 \end{pmatrix}\right)$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.b

Compute the bivariate distribution of the random vector $(s(t-1), s(t))$.

Tutorial exercise 1

Probability distribution of an ESG score

- Since we have:

$$s(t) = s(t-1) + \Delta s(t)$$

we deduce that:

$$\begin{pmatrix} s(t-1) \\ s(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} s(t-1) \\ \Delta s(t) \end{pmatrix}$$

We conclude that $(s(t-1), s(t))$ is a Gaussian random vector.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$\operatorname{var}\left(\mathcal{s}\left(t\right)\right) = 1 + \sigma^2$$

and:

$$\begin{aligned}\operatorname{cov}\left(\mathcal{s}\left(t-1\right), \mathcal{s}\left(t\right)\right) &= \mathbb{E}\left[\mathcal{s}\left(t-1\right) \cdot \mathcal{s}\left(t\right)\right] \\ &= \mathbb{E}\left[\mathcal{s}^2\left(t-1\right) + \mathcal{s}\left(t-1\right) \cdot \Delta \mathcal{s}\left(t\right)\right] \\ &= 1\end{aligned}$$

Tutorial exercise 1

Probability distribution of an ESG score

- It follows that:

$$\begin{pmatrix} \mathcal{S}(t-1) \\ \mathcal{S}(t) \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}_2, \Sigma_\sigma\right)$$

where Σ_σ is the covariance matrix:

$$\Sigma_\sigma = \begin{pmatrix} 1 & 1 \\ 1 & 1+\sigma^2 \end{pmatrix}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.c

Compute the probability $p_k = \Pr\{\mathcal{R}(t-1) = \mathcal{R}_k\}$.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$\begin{aligned}\Pr\left\{\mathcal{R}(t-1)=\mathcal{R}_k\right\} &= \Pr\left\{s_{k-1}\leq \mathcal{S}(t-1)<s_k\right\} \\ &= \Phi(s_k)-\Phi(s_{k-1})\end{aligned}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.d

Compute the joint probability $\Pr\left\{\mathcal{R}(t) = \mathcal{R}_k, \mathcal{R}(t-1) = \mathcal{R}_j\right\}$.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$
\begin{aligned}
(*) &= \Pr\left\{\mathcal{R}(t) = \mathcal{R}_k, \mathcal{R}(t-1) = \mathcal{R}_j\right\} \\
&= \Pr\left\{s_{k-1} \leq \mathcal{s}(t) < s_k, s_{j-1} \leq \mathcal{s}(t-1) < s_j\right\} \\
&= \Phi_2\left(s_j, s_k; \Sigma_\sigma\right) - \Phi_2\left(s_{j-1}, s_k; \Sigma_\sigma\right) - \\
&\quad \Phi_2\left(s_j, s_{k-1}; \Sigma_\sigma\right) + \Phi_2\left(s_{j-1}, s_{k-1}; \Sigma_\sigma\right)
\end{aligned}
$$

where $\Phi_2\left(x, y; \Sigma_\sigma\right)$ is the bivariate Normal cdf with covariance matrix Σ_σ.

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.e

Compute the transition probability
$p_{j,k} = \Pr \{\mathcal{R}(t) = \mathcal{R}_k \mid \mathcal{R}(t-1) = \mathcal{R}_j\}$.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$\begin{aligned} p_{j,k} &= \Pr\left\{\mathcal{R}(t) = \mathcal{R}_k \mid \mathcal{R}(t-1) = \mathcal{R}_j\right\} \\ &= \frac{\Pr\left\{\mathcal{R}(t) = \mathcal{R}_k, \mathcal{R}(t-1) = \mathcal{R}_j\right\}}{\Pr\left\{\mathcal{R}(t-1) = \mathcal{R}_j\right\}} \\ &= \frac{\Phi_2\left(s_j, s_k; \Sigma_\sigma\right) + \Phi_2\left(s_{j-1}, s_{k-1}; \Sigma_\sigma\right)}{\Phi\left(s_j\right) - \Phi\left(s_{j-1}\right)} - \\ &\quad \frac{\Phi_2\left(s_{j-1}, s_k; \Sigma_\sigma\right) + \Phi_2\left(s_j, s_{k-1}; \Sigma_\sigma\right)}{\Phi\left(s_j\right) - \Phi\left(s_{j-1}\right)} \end{aligned}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.f

Compute the monthly turnover $\mathcal{T}(\mathcal{R}_k)$ of the ESG rating $\mathcal{R}_k$.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$\begin{aligned}
\mathcal{T}\left(\mathcal{R}_k\right) &= \Pr\left\{\mathcal{R}\left(t\right) \neq \mathcal{R}_k \mid \mathcal{R}\left(t-1\right) = \mathcal{R}_k\right\} \\
&= 1 - \Pr\left\{\mathcal{R}\left(t\right) = \mathcal{R}_k \mid \mathcal{R}\left(t-1\right) = \mathcal{R}_k\right\} \\
&= 1 - p_{k,k}
\end{aligned}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.g

Compute the monthly turnover $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$ of the ESG rating system.

Tutorial exercise 1

Probability distribution of an ESG score

- We have:

$$\begin{aligned}\mathcal{T}\left(\mathcal{R}_1,\ldots,\mathcal{R}_K\right) &= \sum_{k=1}^{K}\Pr\left\{\mathcal{R}\left(t-1\right)=\mathcal{R}_k\right\}\cdot\mathcal{T}\left(\mathcal{R}_k\right) \\ &= \sum_{k=1}^{K}\Pr\left\{\mathcal{R}\left(t\right)\neq\mathcal{R}_k,\mathcal{R}\left(t-1\right)=\mathcal{R}_k\right\}\end{aligned}$$

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.h

For each rating system given in Questions 3.a, 3.b and 3.c, determine the corresponding ESG migration matrix and the monthly turnover of the rating system if we assume that σ is equal to 10%. What is the best ESG rating system if we would like to control the turnover of ESG ratings?

Tutorial exercise 1

Probability distribution of an ESG score

Table 96: ESG rating migration matrix (Question 3.a)

Rating	s_k	p_k	Transition probability $p_{j,k}$				$\mathcal{T}(\mathcal{R}_k)$
D		6.68%	92.96%	7.04%	0.00%	0.00%	7.04%
	-1.50						
C		43.32%	1.31%	95.03%	3.66%	0.00%	4.97%
	0.00						
B		43.32%	0.00%	3.66%	95.03%	1.31%	4.97%
	1.50						
A		6.68%	0.00%	0.00%	7.04%	92.96%	7.04%
$\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$							5.25%

Tutorial exercise 1

Probability distribution of an ESG score

Table 97: ESG rating migration matrix (Question 3.b)

Rating	s_k	p_k	Transition probability $p_{j,k}$				$\mathcal{T}(\mathcal{R}_k)$
D		25.00%	95.15%	4.85%	0.00%	0.00%	4.85%
C	−0.67	25.00%	5.27%	88.38%	6.35%	0.00%	11.62%
B	0.00	25.00%	0.00%	6.35%	88.38%	5.27%	11.62%
A	0.67	25.00%	0.00%	0.00%	4.85%	95.15%	4.85%
$\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$							8.23%

Tutorial exercise 1

Probability distribution of an ESG score

Table 98: ESG rating migration matrix (Question 3.c)

Rating	s_k	p_k	Transition probability $p_{j,k}$				$\mathcal{T}(\mathcal{R}_k)$
D		10.00%	93.54%	6.46%	0.00%	0.00%	6.46%
C	-1.28	40.00%	1.89%	94.14%	3.97%	0.00%	5.86%
B	0.00	40.00%	0.00%	3.97%	94.14%	1.89%	5.86%
A	1.28	10.00%	0.00%	0.00%	6.46%	93.54%	6.46%
$\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$							5.98%

Tutorial exercise 1

Probability distribution of an ESG score

The ESG rating system defined in Question 3.a is the best rating system if we would like to reduce the monthly turnover of ESG ratings.

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.i

Draw the relationship between the parameter σ and the turnover $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$ for the three ESG rating systems.

Tutorial exercise 1

Probability distribution of an ESG score

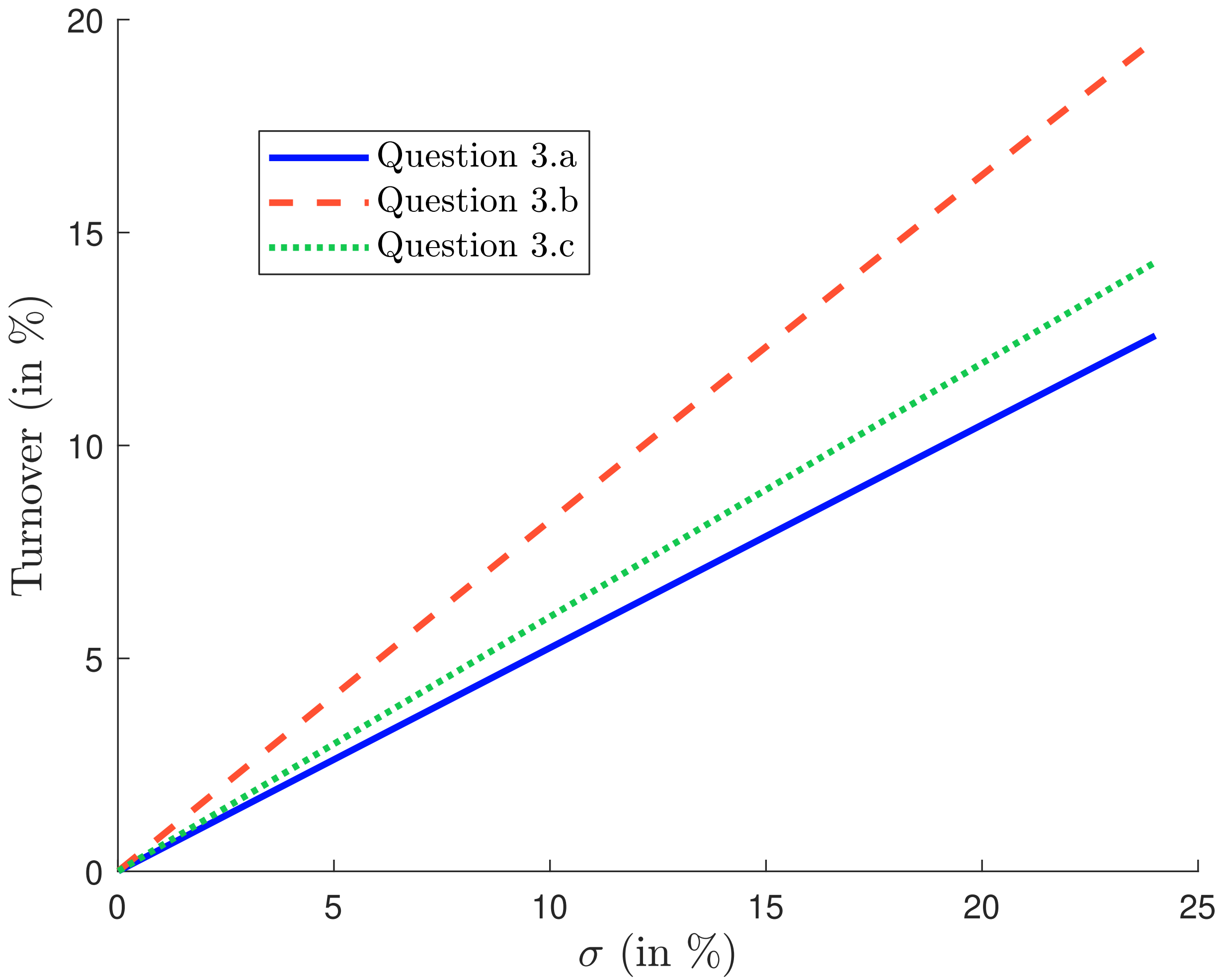


Figure 169: Relationship between σ and $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.j

We consider a uniform ESG rating system where:

$$\Pr\left\{\mathcal{R}(t-1)=\mathcal{R}_k\right\}=\frac{1}{K}$$

Draw the relationship between the number of notches K and the turnover $\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$ when the parameter σ takes the values 5%, 10% and 25%.

Tutorial exercise 1

Probability distribution of an ESG score

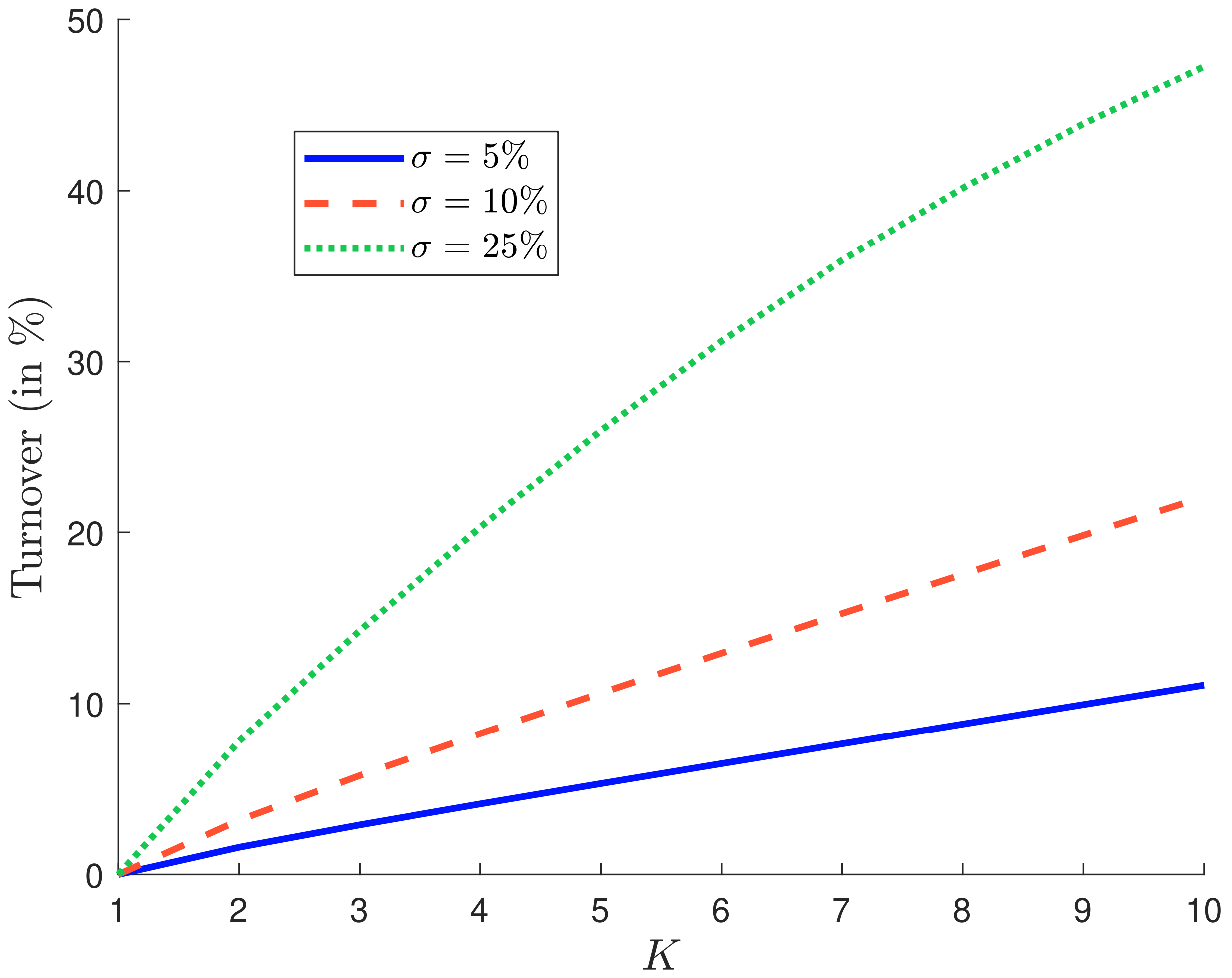


Figure 170: Relationship between K and $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$

Tutorial exercise 1

Probability distribution of an ESG score

Question 4.k

Why is an ESG rating system different than a credit rating system? What do you conclude from the previous analysis? What is the issue of ESG exclusion policy and negative screening?

Tutorial exercise 1

Probability distribution of an ESG score

- An ESG rating system is mainly quantitative and highly depends on the mapping function. This is not the case of a credit rating system, which is mainly qualitative and discretionary.
- This explains that the turnover of an ESG rating system is higher than the turnover of a credit rating system.
- The stabilization of the ESG rating system implies to reduce the turnover $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$, which depends on:
  1. The number of notches[35] K;
  2. The volatility σ of score changes
  3. The design of the ESG rating system $(s_1, \ldots, s_{K-1})$
- The turnover $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$ has a big impact on an ESG exclusion (or negative screening) policy, because it creates noisy short-term entry/exit positions that do not necessarily correspond to a decrease or increase of the long-term ESG risks.

[35] This is why ESG rating systems have less notches than credit rating systems

Tutorial exercise 2

Enhanced ESG score & tracking error control

Exercise

We consider a capitalization-weighted equity index, which is composed of 8 stocks. Their weights, volatilities and ESG scores are the following:

Stock	#1	#2	#3	#4	#5	#6	#7	#8
CW weight	0.23	0.19	0.17	0.13	0.09	0.08	0.06	0.05
Volatility	0.22	0.20	0.25	0.18	0.35	0.23	0.13	0.29
ESG score	−1.20	0.80	2.75	1.60	−2.75	−1.30	0.90	−1.70

The correlation matrix is given by:

$$\rho = \begin{pmatrix} 100\% & & & & & & & \\ 80\% & 100\% & & & & & & \\ 70\% & 75\% & 100\% & & & & & \\ 60\% & 65\% & 80\% & 100\% & & & & \\ 70\% & 50\% & 70\% & 85\% & 100\% & & & \\ 50\% & 60\% & 70\% & 80\% & 60\% & 100\% & & \\ 70\% & 50\% & 70\% & 75\% & 80\% & 50\% & 100\% & \\ 60\% & 65\% & 70\% & 75\% & 65\% & 70\% & 80\% & 100\% \end{pmatrix}$$

Tutorial exercise 2

Enhanced ESG score & tracking error control

Question 1

Calculate the ESG score of the benchmark.

Tutorial exercise 2

Enhanced ESG score & tracking error control

- We note b_i and $\mathcal{S}_i$ the weight in the benchmark and the ESG score of Stock i
- The ESG score of the benchmark is equal to:

$$\mathcal{S}(b) = \sum_{i=1}^{8} b_i \cdot \mathcal{S}_i = 0.1690$$

Tutorial exercise 2

Enhanced ESG score & tracking error control

Question 2

We consider the EW and ERC portfolios. Calculate the ESG score of these two portfolios. Define the ESG excess score with respect to the benchmark. Comment on these results.

Tutorial exercise 2

Enhanced ESG score & tracking error control

- The composition of the EW portfolio is $x_i = 12.5\%$ and we have:

$$\mathcal{S}(x_{\mathrm{ew}}) = \sum_{i=1}^{8} \frac{\mathcal{S}_i}{8} = -0.1125$$

- The composition of the ERC portfolio is $x_1 = 12.42\%$, $x_2 = 14.03\%$, $x_3 = 10.17\%$, $x_4 = 13.79\%$, $x_5 = 7.59\%$, $x_6 = 12.34\%$, $x_7 = 20.61\%$ and $x_8 = 9.06\%$. We have:

$$\mathcal{S}(x_{\mathrm{erc}}) = \sum_{i=1}^{8} x_i \cdot \mathcal{S}_i = 0.1259$$

Tutorial exercise 2

Enhanced ESG score & tracking error control

- The ESG excess score with respect to the benchmark is:

$$\mathcal{S}(x \mid b) = \mathcal{S}(x) - \mathcal{S}(b)$$

We have:

$$\begin{aligned} \mathcal{S}(x_{\mathrm{ew}} \mid b) &= -0.1125 - 0.1690 = -0.2815 \\ \mathcal{S}(x_{\mathrm{erc}} \mid b) &= 0.1259 - 0.1690 = -0.0431 \end{aligned}$$

- The two portfolios are riskier than the benchmark portfolio in terms of ESG risk

Tutorial exercise 2

Enhanced ESG score & tracking error control

Question 3

Write the γ-problem of the ESG optimized portfolio when the goal is to improve the ESG score of the benchmark and control at the same time the tracking error volatility. Give the QP objective function.

Tutorial exercise 2

Enhanced ESG score & tracking error control

- We have:

$$x^{\star} = \arg\min \frac{1}{2}\sigma^2(x \mid b) - \gamma \mathcal{S}(x \mid b)$$
$$\text{u.c.} \quad \begin{cases} \mathbf{1}_n^\top x = 1 \\ \mathbf{0}_n \leq x \leq \mathbf{1}_n \\ x \in \Omega \end{cases}$$

- Since $\sigma^2(x \mid b) = (x - b)^\top \Sigma (x - b)$ and $\mathcal{S}(x \mid b) = (x - b)^\top \mathcal{S}$, we deduce that the QP objective function is:

$$x^{\star} = \arg\min \frac{1}{2} x^\top \Sigma x - x^\top (\gamma \mathcal{S} + \Sigma b)$$

Tutorial exercise 2

Enhanced ESG score & tracking error control

Question 4

Draw the efficient frontier between the tracking error volatility and the ESG excess score[a].

[a] We notice that $\gamma \in [0, 1.2\%]$ is sufficient for drawing the efficient frontier.

Tutorial exercise 2

Enhanced ESG score & tracking error control

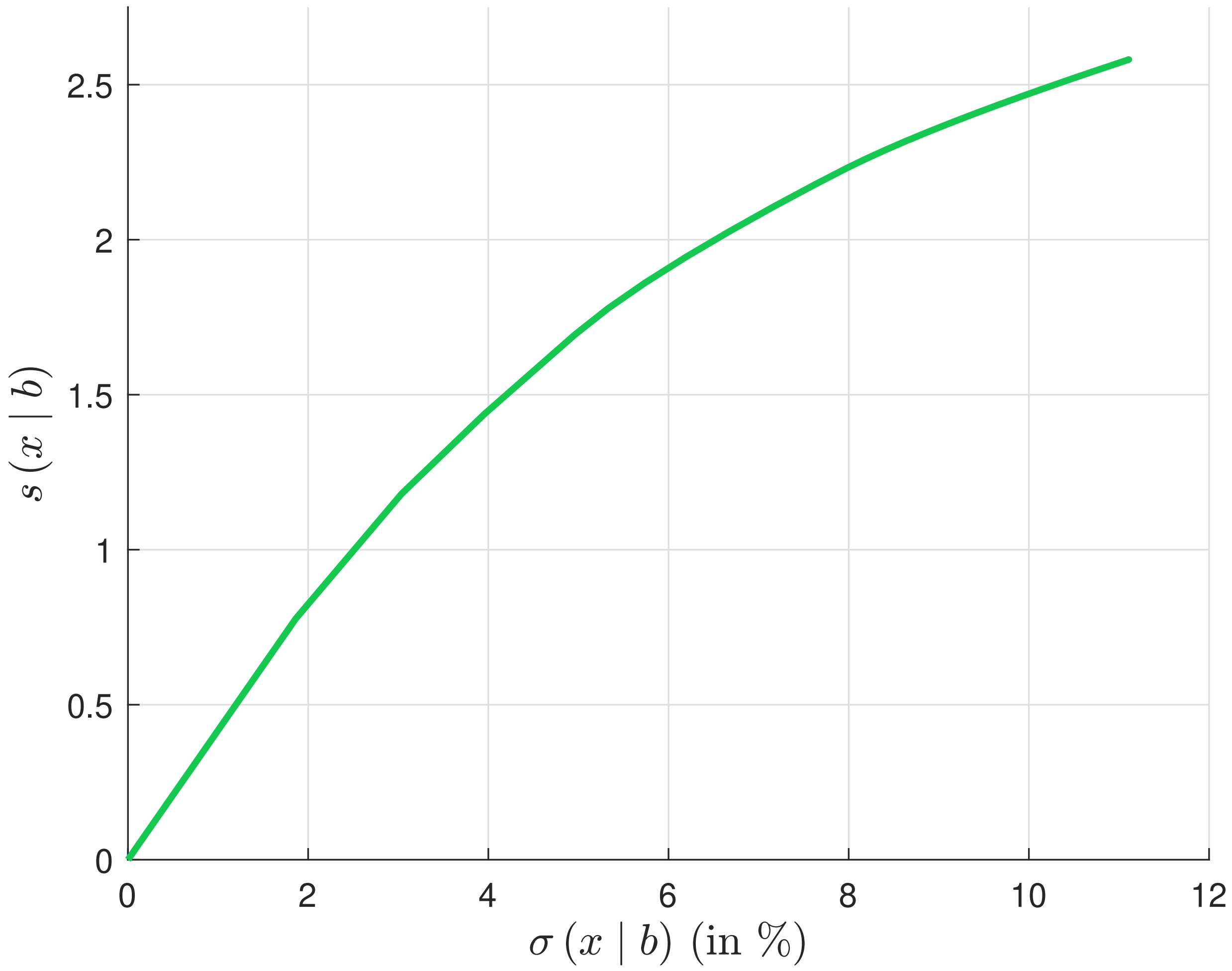


Figure 171: ESG efficient frontier

Tutorial exercise 2

Enhanced ESG score & tracking error control

Question 5

Using the bisection algorithm, find the optimal portfolio if we would like to improve the ESG score of the benchmark by 0.5. Give the optimal value of γ. Compute the tracking error volatility $\sigma(x \mid b)$.

Tutorial exercise 2

Enhanced ESG score & tracking error control

- The solution is equal to:

Stock	s_i	b_i	$x_i^\star$
#1	-1.200	23.000	25.029
#2	0.800	19.000	14.251
#3	2.750	17.000	21.947
#4	1.600	13.000	27.305
#5	-2.750	9.000	3.718
#6	-1.300	8.000	1.339
#7	0.900	6.000	1.675
#8	-1.700	5.000	4.736

- The optimal value of γ is 0.02768%
- The tracking error volatility is equal to 1.17636%

Tutorial exercise 2

Enhanced ESG score & tracking error control

Question 6

Same question if we would like to improve the ESG score of the benchmark by 1.0.

Tutorial exercise 2

Enhanced ESG score & tracking error control

- The solution is equal to:

Stock	s_i	b_i	$x_i^\star$
#1	-1.200	23.000	21.699
#2	0.800	19.000	12.443
#3	2.750	17.000	28.739
#4	1.600	13.000	33.555
#5	-2.750	9.000	0.002
#6	-1.300	8.000	0.000
#7	0.900	6.000	2.433
#8	-1.700	5.000	1.129

- The optimal value of γ is 0.07276%
- The tracking error volatility is equal to 2.48574%

Tutorial exercise 2

Enhanced ESG score & tracking error control

Question 7

We impose that the portfolio weights can not be greater than 30%. Find the optimal portfolio if we would like to improve the ESG score of the benchmark by 1.0.

Tutorial exercise 2

Enhanced ESG score & tracking error control

- The solution is equal to:

Stock	s_i	b_i	$x_i^\star$
#1	-1.200	23.000	20.116
#2	0.800	19.000	14.082
#3	2.750	17.000	29.481
#4	1.600	13.000	30.000
#5	-2.750	9.000	0.644
#6	-1.300	8.000	0.000
#7	0.900	6.000	4.662
#8	-1.700	5.000	1.015

- The optimal value of γ is 0.07355%
- The tracking error volatility is equal to 2.50317%

Tutorial exercise 2

Enhanced ESG score & tracking error control

Question 8

Comment on these results.

Tutorial exercise 2

Enhanced ESG score & tracking error control

- We notice that the evolution of the weights is not necessarily monotonous with respect to the ESG excess score $\mathcal{S}(x \mid b)$. For instance, if we target an improvement of 0.5, the weight of Stock #1 increases ($23\% \Rightarrow 25.029\%$). If we target an improvement of 1.0, the the weight of Stock #1 decreases ($25.029\% \Rightarrow 21.699\%$)
- Generally, the optimiser reduces the weight of stocks with low ESG scores and increases the weight of stocks with high ESG scores
- Nevertheless, the weight differences are not ranked in the same order than the ESG scores. For instance, if we target an improvement of 0.5, the largest variation is observed for Stock #4, which has an ESG score of 1.6. This is not the largest ESG score, since Stock #3 has an ESG score of 2.75
- This is due to the structure of the covariance matrix (Stock #3 is riskier than Stock #4)

Asset Management
Lecture 5. Machine Learning in Asset Management

Thierry Roncalli*

*University of Paris-Saclay

January 2021

Agenda

- Lecture 1: Portfolio Optimization
- Lecture 2: Risk Budgeting
- Lecture 3: Smart Beta, Factor Investing and Alternative Risk Premia
- Lecture 4: Green and Sustainable Finance, ESG Investing and Climate Risk
- **Lecture 5: Machine Learning in Asset Management**

Prologue

- Machine learning is a hot topic in asset management (and more generally in finance)
- Machine learning and data mining are two sides of the same coin

backtesting performance $\neq$ **live performance**

- Reaching for the stars: a complex/complicated process does not mean a good solution

Don't forget the 3 rules in asset management

1. It is difficult to make money
2. It is difficult to make money
3. It is difficult to make money

Prologue

- In this lecture, we focus on ML optimization algorithms, because they have proved their worth
- We have no time to study classical ML methods that can be used by quants to build investment strategies[36]

[36] Don't believe that they are always significantly better than standard statistical approaches!!!

Standard optimization algorithms

- Gradient descent methods
- Conjugate gradient (CG) methods (Fletcher–Reeves, Polak–Ribiere, etc.)
- Quasi-Newton (QN) methods (NR, BFGS, DFP, etc.)
- Quadratic programming (QP) methods
- Sequential QP methods
- Interior-point methods

Standard optimization algorithms

- We consider the following unconstrained minimization problem:

$$x^\star = \arg\min_x f(x) \tag{7}$$

  where $x \in \mathbb{R}^n$ and $f(x)$ is a continuous, smooth and convex function
- In order to find the solution $x^\star$, optimization algorithms use iterative algorithms:

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \Delta x^{(k)} \\ &= x^{(k)} - \eta^{(k)} D^{(k)} \end{aligned}$$

  where:
  - $x^{(0)}$ is the vector of starting values
  - $x^{(k)}$ is the approximated solution of Problem (7) at the k^{th} iteration
  - $\eta^{(k)} > 0$ is a scalar that determines the step size
  - $D^{(k)}$ is the direction

Standard optimization algorithms

- Gradient descent:

$$D^{(k)} = \nabla f\left(x^{(k)}\right) = \frac{\partial\, f\left(x^{(k)}\right)}{\partial\, x}$$

- Newton-Raphson method:

$$D^{(k)} = \left(\nabla^2 f\left(x^{(k)}\right)\right)^{-1} \nabla f\left(x^{(k)}\right) = \left(\frac{\partial^2 f\left(x^{(k)}\right)}{\partial\, x\, \partial\, x^\top}\right)^{-1} \frac{\partial\, f\left(x^{(k)}\right)}{\partial\, x}$$

- Quasi-Newton method:

$$D^{(k)} = H^{(k)} \nabla f\left(x^{(k)}\right)$$

where $H^{(k)}$ is an approximation of the inverse of the Hessian matrix

Standard optimization algorithms

What are the issues?

1. How to solve large-scale optimization problems?
2. How to solve optimization problems where there are multiple solutions?
3. How to just find an "*acceptable*" solution?

The case of neural networks and deep learning

⇒ Standard approaches are not well adapted

Machine learning optimization algorithms

Machine learning problems

- Non-smooth objective function
- Non-unique solution
- Large-scale dimension

Optimization in machine learning requires to reinvent numerical optimization

Machine learning optimization algorithms

We consider 4 methods:

- Cyclical coordinate descent (CCD)
- Alternative direction method of multipliers (ADMM)
- Proximal operators (PO)
- Dykstra's algorithm (DA)

Coordinate descent methods

The fall and the rise of the steepest descent method

In the 1980s:

- Conjugate gradient methods (Fletcher–Reeves, Polak–Ribiere, etc.)
- Quasi-Newton methods (NR, BFGS, DFP, etc.)

In the 1990s:

- Neural networks
- Learning rules: Descent, Momentum/Nesterov and Adaptive learning methods

In the 2000s:

- Gradient descent (by **observations**): Batch gradient descent (BGD), Stochatic gradient descent (SGD), Mini-batch gradient descent (MGD)
- Gradient descent (by **parameters**): Coordinate descent (CD), cyclical coordinate descent (CCD), Random coordinate descent (RCD)

Coordinate descent methods

Descent method

The descent algorithm is defined by the following rule:

$$x^{(k+1)} = x^{(k)} + \Delta x^{(k)} = x^{(k)} - \eta^{(k)} D^{(k)}$$

At the k^{th} Iteration, the current solution $x^{(k)}$ is updated by going in the opposite direction to $D^{(k)}$ (generally, we set $D^{(k)} = \partial_x f\left(x^{(k)}\right)$)

Coordinate descent method

Coordinate descent is a modification of the descent algorithm by minimizing the function along one coordinate at each step:

$$x_i^{(k+1)} = x_i^{(k)} + \Delta x_i^{(k)} = x_i^{(k)} - \eta^{(k)} D_i^{(k)}$$

$\Rightarrow$ The coordinate descent algorithm becomes a scalar problem

Coordinate descent methods

Choice of the variable i

1. Random coordinate descent (RCD)
   We assign a random number between 1 and n to the index i (Nesterov, 2012)
2. Cyclical coordinate descent (CCD)
   We cyclically iterate through the coordinates (Tseng, 2001):

$$x_i^{(k+1)} = \arg\min_x f\left(x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}, x, x_{i+1}^{(k)}, \ldots, x_n^{(k)}\right)$$

Cyclical coordinate descent (CCD)

Example 1

We consider the following function:

$$f(x_1, x_2, x_3) = (x_1 - 1)^2 + x_2^2 - x_2 + (x_3 - 2)^4 e^{x_1 - x_2 + 3}$$

We have:

$$\begin{aligned} D_1 = \frac{\partial f(x_1, x_2, x_3)}{\partial x_1} &= 2(x_1 - 1) + (x_3 - 2)^4 e^{x_1 - x_2 + 3} \\ D_2 = \frac{\partial f(x_1, x_2, x_3)}{\partial x_2} &= 2x_2 - 1 - (x_3 - 2)^4 e^{x_1 - x_2 + 3} \\ D_3 = \frac{\partial f(x_1, x_2, x_3)}{\partial x_3} &= 4(x_3 - 2)^3 e^{x_1 - x_2 + 3} \end{aligned}$$

Cyclical coordinate descent (CCD)

The CCD algorithm is defined by the following iterations:

$$\left\{\begin{array}{l}
x_1^{(k+1)} = x_1^{(k)} - \eta^{(k)} \left(2\left(x_1^{(k)} - 1\right) + \left(x_3^{(k)} - 2\right)^4 e^{x_1^{(k)} - x_2^{(k)} + 3}\right) \\
x_2^{(k+1)} = x_2^{(k)} - \eta^{(k)} \left(2x_2^{(k)} - 1 - \left(x_3^{(k)} - 2\right)^4 e^{x_1^{(k+1)} - x_2^{(k)} + 3}\right) \\
x_3^{(k+1)} = x_3^{(k)} - \eta^{(k)} \left(4\left(x_3^{(k)} - 2\right)^3 e^{x_1^{(k+1)} - x_2^{(k+1)} + 3}\right)
\end{array}\right.$$

We have the following scheme:

$$\begin{array}{ccl}
\left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}\right) & \rightarrow & x_1^{(1)} \rightarrow \left(x_1^{(1)}, x_2^{(0)}, x_3^{(0)}\right) \rightarrow x_2^{(1)} \rightarrow \left(x_1^{(1)}, x_2^{(1)}, x_3^{(0)}\right) \rightarrow x_3^{(1)} \rightarrow \\
\left(x_1^{(1)}, x_2^{(1)}, x_3^{(1)}\right) & \rightarrow & x_1^{(2)} \rightarrow \left(x_1^{(2)}, x_2^{(1)}, x_3^{(1)}\right) \rightarrow x_2^{(2)} \rightarrow \left(x_1^{(2)}, x_2^{(2)}, x_3^{(1)}\right) \rightarrow x_3^{(2)} \rightarrow \\
\left(x_1^{(2)}, x_2^{(2)}, x_3^{(2)}\right) & \rightarrow & x_1^{(3)} \rightarrow \ldots
\end{array}$$

Cyclical coordinate descent (CCD)

Table 99: Solution obtained with the CCD algorithm ($\eta^{(k)} = 0.25$)

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$D_1^{(k)}$	$D_2^{(k)}$	$D_3^{(k)}$
0	1.0000	1.0000	1.0000			
1	-4.0214	0.7831	1.1646	20.0855	0.8675	-0.6582
2	-1.5307	0.8834	2.2121	-9.9626	-0.4013	-4.1902
3	-0.2663	0.6949	2.1388	-5.0578	0.7540	0.2932
4	0.3661	0.5988	2.0962	-2.5297	0.3845	0.1703
5	0.6827	0.5499	2.0758	-1.2663	0.1957	0.0818
6	0.8412	0.5252	2.0638	-0.6338	0.0989	0.0480
7	0.9205	0.5127	2.0560	-0.3172	0.0498	0.0314
8	0.9602	0.5064	2.0504	-0.1588	0.0251	0.0222
9	0.9800	0.5033	2.0463	-0.0795	0.0126	0.0166
∞	1.0000	0.5000	2.0000	0.0000	0.0000	0.0000

The lasso revolution

Least absolute shrinkage and selection operator (lasso)

The lasso method consists in adding a ℓ_1 penalty function to the least square problem:

$$\begin{aligned} \hat{\beta}^{\text{lasso}}(\tau) &= \arg\min \frac{1}{2}(Y - X\beta)^\top (Y - X\beta) \\ \text{s.t.} \quad & \|\beta\|_1 = \sum_{j=1}^{m} |\beta_j| \leq \tau \end{aligned}$$

This problem is equivalent to:

$$\hat{\beta}^{\text{lasso}}(\lambda) = \arg\min \frac{1}{2}(Y - X\beta)^\top (Y - X\beta) + \lambda \|\beta\|_1$$

We have:

$$\tau = \left\| \hat{\beta}^{\text{lasso}}(\lambda) \right\|_1$$

Solving the lasso regression problem

We introduce the parametrization:

$$\beta = \begin{pmatrix} I_m & -I_m \end{pmatrix} \begin{pmatrix} \beta^+ \\ \beta^- \end{pmatrix} = \beta^+ - \beta^-$$

under the constraints $\beta^+ \geq \mathbf{0}_m$ and $\beta^- \geq \mathbf{0}_m$. We deduce that:

$$\left\| \beta \right\|_1 = \sum_{j=1}^{m} \left| \beta_j^+ - \beta_j^- \right| = \sum_{j=1}^{m} \left| \beta_j^+ \right| + \sum_{j=1}^{m} \left| \beta_j^- \right| = \mathbf{1}_m^\top \beta^+ + \mathbf{1}_m^\top \beta^-$$

Solving the lasso regression problem

Augmented QP program of the lasso regression (λ-problem)

The augmented QP program is specified as follows:

$$\begin{aligned} \hat{\theta} &= \arg\min \frac{1}{2}\theta^\top Q\theta - \theta^\top R \\ &\text{s.t. } \theta \geq \mathbf{0}_{2m} \end{aligned}$$

where $\theta = (\beta^+, \beta^-)$, $\tilde{X} = \begin{pmatrix} X & -X \end{pmatrix}$, $Q = \tilde{X}^\top \tilde{X}$ and $R = \tilde{X}^\top Y + \lambda \mathbf{1}_{2m}$. If we denote $T = \begin{pmatrix} I_m & -I_m \end{pmatrix}$, we obtain:

$$\hat{\beta}^{\mathrm{lasso}}(\lambda) = T\hat{\theta}$$

Solving the lasso regression problem

Augmented QP program of the lasso regression (τ-problem)

If we consider the τ-problem, we obtain another augmented QP program:

$$\begin{aligned} \hat{\theta} & = & \arg\min \frac{1}{2}\theta^\top Q\theta - \theta^\top R \\ & & \text{s.t.} \quad \left\{ \begin{array}{l} C\theta \leq D \\ \theta \geq \mathbf{0}_{2m} \end{array} \right. \end{aligned}$$

where $Q = \tilde{X}^\top \tilde{X}$, $R = \tilde{X}^\top Y$, $C = \mathbf{1}_{2m}^\top$ and $D = \tau$. Again, we have:

$$\hat{\beta}(\tau) = T\hat{\theta}$$

Solving the lasso regression problem

We consider the linear regression:

$$Y = X\beta + \varepsilon$$

where Y is a $n \times 1$ vector, X is a $n \times m$ matrix and β is a $m \times 1$ vector. The optimization problem is:

$$\hat{\beta} = \arg\min f(\beta) = \frac{1}{2}(Y - X\beta)^\top (Y - X\beta)$$

Since we have $\partial_\beta f(\beta) = -X^\top (Y - X\beta)$, we deduce that:

$$\begin{aligned}
\frac{\partial f(\beta)}{\partial \beta_j} &= x_j^\top (X\beta - Y) \\
&= x_j^\top \left(x_j\beta_j + X_{(-j)}\beta_{(-j)} - Y\right) \\
&= x_j^\top x_j\beta_j + x_j^\top X_{(-j)}\beta_{(-j)} - x_j^\top Y
\end{aligned}$$

where x_j is the $n \times 1$ vector corresponding to the j^{th} variable and $X_{(-j)}$ is the $n \times (m-1)$ matrix (without the j^{th} variable)

Solving the lasso regression problem

At the optimum, we have $\partial_{\beta_j} f(\beta) = 0$ or:

$$\beta_j = \frac{x_j^\top Y - x_j^\top X_{(-j)}\beta_{(-j)}}{x_j^\top x_j} = \frac{x_j^\top \left(Y - X_{(-j)}\beta_{(-j)}\right)}{x_j^\top x_j}$$

CCD algorithm for the linear regression

We have:

$$\beta_j^{(k+1)} = \frac{x_j^\top \left(Y - \sum_{j'=1}^{j-1} x_{j'}\beta_{j'}^{(k+1)} - \sum_{j'=j+1}^{m} x_{j'}\beta_{j'}^{(k)}\right)}{x_j^\top x_j}$$

$\Rightarrow$ Introducing pointwise constraints is straightforward

Solving the lasso regression problem

The objective function becomes:

$$\begin{aligned} f(\beta) &= \frac{1}{2}(Y - X\beta)^\top (Y - X\beta) + \lambda \|\beta\|_1 \\ &= f_{\mathrm{OLS}}(\beta) + \lambda \|\beta\|_1 \end{aligned}$$

Since the norm is separable — $\|\beta\|_1 = \sum_{j=1}^m |\beta_j|$, the first-order condition is:

$$\frac{\partial\, f_{\mathrm{OLS}}(\beta)}{\partial\, \beta_j} + \lambda \partial |\beta_j| = 0$$

or:

$$\underbrace{\left(x_j^\top x_j\right)}_{c} \beta_j - \underbrace{x_j^\top \left(Y - X_{(-j)}\beta_{(-j)}\right)}_{v} + \lambda \partial\, |\beta_j| = 0$$

Derivation of the soft-thresholding operator

We consider the following equation:

$$c\beta_j - v + \lambda \partial \left|\beta_j\right| \in \{0\}$$

where $c > 0$ and $\lambda > 0$. Since we have $\partial \left|\beta_j\right| = \operatorname{sign}\left(\beta_j\right)$, we deduce that:

$$\beta_j^\star = \begin{cases} c^{-1}\left(v + \lambda\right) & \text{if } \beta_j^\star < 0 \\ 0 & \text{if } \beta_j^\star = 0 \\ c^{-1}\left(v - \lambda\right) & \text{if } \beta_j^\star > 0 \end{cases}$$

If $\beta_j^\star < 0$ or $\beta_j^\star > 0$, then we have $v + \lambda < 0$ or $v - \lambda > 0$. This is equivalent to set $|v| > \lambda > 0$. The case $\beta_j^\star = 0$ implies that $|v| \leq \lambda$. We deduce that:

$$\beta_j^\star = c^{-1} \cdot \mathcal{S}\left(v; \lambda\right)$$

where $\mathcal{S}\left(v; \lambda\right)$ is the soft-thresholding operator:

$$\begin{aligned} \mathcal{S}\left(v; \lambda\right) &= \begin{cases} 0 & \text{if } |v| \leq \lambda \\ v - \lambda \operatorname{sign}\left(v\right) & \text{otherwise} \end{cases} \\ &= \operatorname{sign}\left(v\right) \cdot \left(|v| - \lambda\right)_+ \end{aligned}$$

Solving the lasso regression problem

CCD algorithm for the lasso regression

We have:

$$\beta_j^{(k+1)} = \frac{1}{x_j^\top x_j} \mathcal{S}\left(x_j^\top \left(Y - \sum_{j'=1}^{j-1} x_{j'} \beta_{j'}^{(k+1)} - \sum_{j'=j+1}^{m} x_{j'} \beta_{j'}^{(k)} \right) ; \lambda \right)$$

where $\mathcal{S}(v; \lambda)$ is the **soft-thresholding operator**:

$$\mathcal{S}(v; \lambda) = \operatorname{sign}(v) \cdot (|v| - \lambda)_+$$

Solving the lasso regression problem

Table 100: Matlab code

```
for k = 1:nIters
    for j = 1:m
        x_j = X(:,j);
        X_j = X;
        X_j(:,j) = zeros(n,1);
        if lambda > 0
            v = x_j'*(Y - X_j*beta);
            beta(j) = max(abs(v) - lambda,0) * sign(v) / (x_j'*x_j);
        else
            beta(j) = x_j'*(Y - X_j*beta) / (x_j'*x_j);
        end
    end
end
```

Solving the lasso regression problem

Example 2

We consider the following data:

i	y	x_1	x_2	x_3	x_4	x_5
1	3.1	2.8	4.3	0.3	2.2	3.5
2	24.9	5.9	3.6	3.2	0.7	6.4
3	27.3	6.0	9.6	7.6	9.5	0.9
4	25.4	8.4	5.4	1.8	1.0	7.1
5	46.1	5.2	7.6	8.3	0.6	4.5
6	45.7	6.0	7.0	9.6	0.6	0.6
7	47.4	6.1	1.0	8.5	9.6	8.6
8	−1.8	1.2	9.6	2.7	4.8	5.8
9	20.8	3.2	5.0	4.2	2.7	3.6
10	6.8	0.5	9.2	6.9	9.3	0.7
11	12.9	7.9	9.1	1.0	5.9	5.4
12	37.0	1.8	1.3	9.2	6.1	8.3
13	14.7	7.4	5.6	0.9	5.6	3.9
14	−3.2	2.3	6.6	0.0	3.6	6.4
15	44.3	7.7	2.2	6.5	1.3	0.7

Solving the lasso regression problem

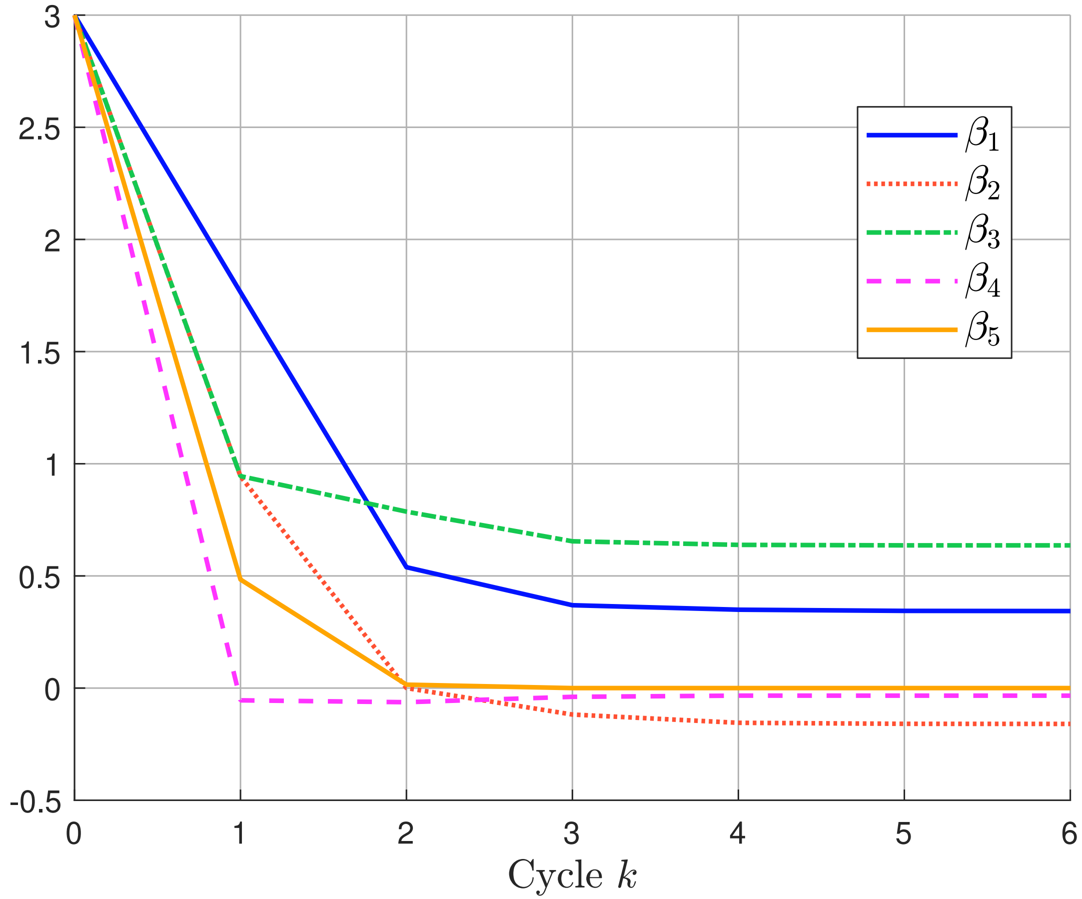


Figure 172: Convergence of the CCD algorithm (lasso regression, $\lambda = 2$)

Note: we start the CCD algorithm with $\beta_j^{(0)} = 0$ (don't forget to standardize the data!)

Solving the lasso regression problem

1. The dimension problem is $(2m, 2m)$ for QP and $(1, 0)$ for CCD!
2. CCD is faster for lasso regression than for linear regression (because of the soft-thresholding operator)!

Suppose $n = 50\,000$ and $m = 1\,000\,000$ (DNA sequence problem!)

Solving the lasso regression problem

Example 3

- We consider an experiment with $n = 100\,000$ observations and $m = 50$ variables.
- The design matrix X is built using the uniform distribution while the residuals are simulated using a Gaussian distribution and a standard deviation of 20%.
- The beta coefficients are distributed uniformly between -3 and $+3$ except four coefficients that take a larger value.
- We then standardize the data of X and Y.
- For initializing the coordinates, we use uniform random numbers between -1 and $+1$.

Solving the lasso regression problem

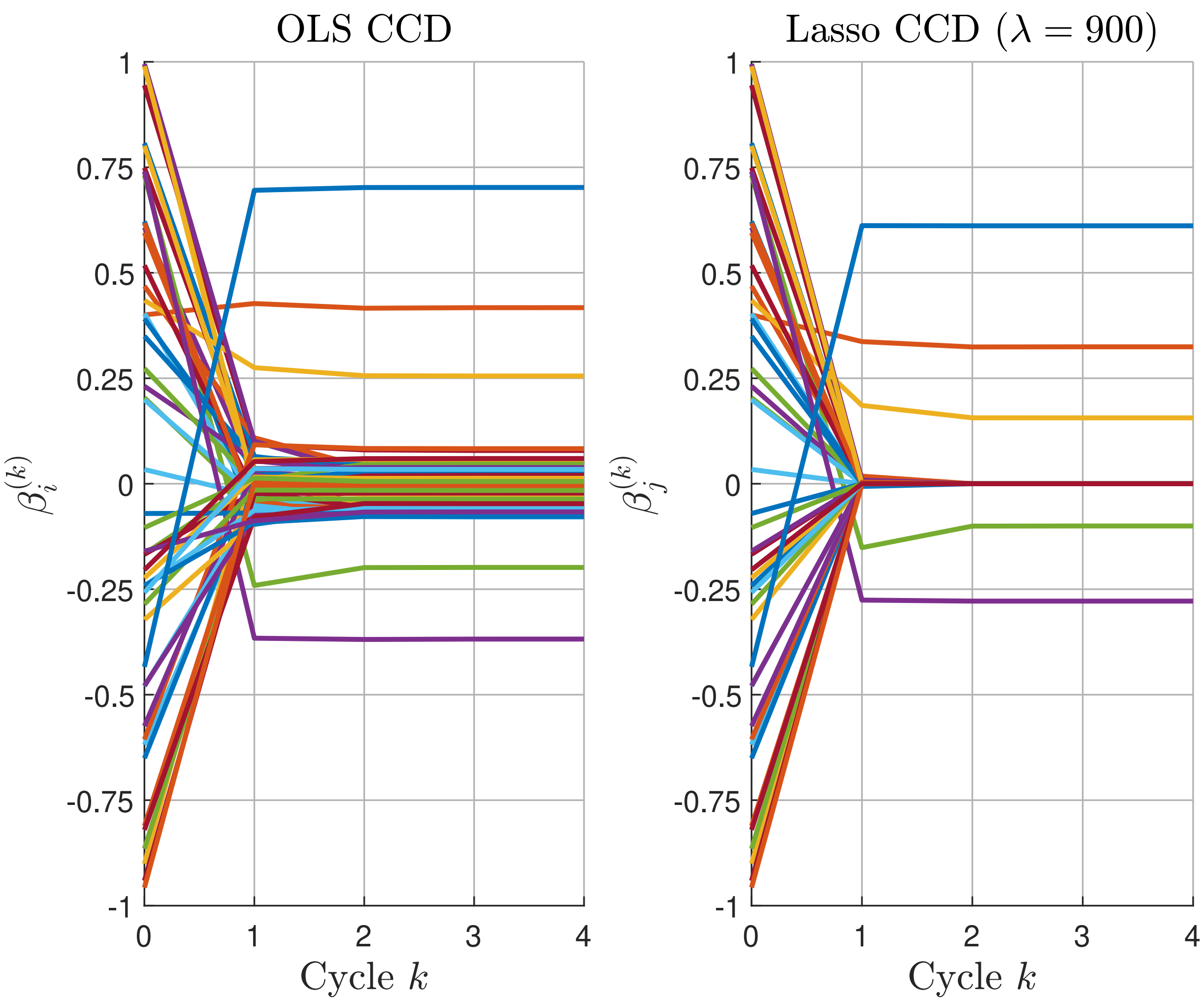


Figure 173: Convergence of the CCD algorithm (lasso vs linear regression)

Alternative direction method of multipliers

Definition

The alternating direction method of multipliers (ADMM) is an algorithm introduced by Gabay and Mercier (1976) to solve optimization problems which can be expressed as:

$$\begin{aligned} \{x^\star, y^\star\} &= \arg\min_{(x,y)} f_x(x) + f_y(y) \\ \text{s.t.} &\quad Ax + By = c \end{aligned}$$

The algorithm is:

$$\begin{aligned} x^{(k+1)} &= \arg\min_x \left\{ f_x(x) + \frac{\varphi}{2} \left\| Ax + By^{(k)} - c + u^{(k)} \right\|_2^2 \right\} \\ y^{(k+1)} &= \arg\min_y \left\{ f_y(y) + \frac{\varphi}{2} \left\| Ax^{(k+1)} + By - c + u^{(k)} \right\|_2^2 \right\} \\ u^{(k+1)} &= u^{(k)} + \left(Ax^{(k+1)} + By^{(k+1)} - c \right) \end{aligned}$$

Alternative direction method of multipliers

What is the underlying idea?

- Minimizing $f_x(x) + f_y(y)$ with respect to (x, y) is a difficult task
- Minimizing
$$g_x(x) = f_x(x) + \frac{\varphi}{2} \|Ax + By - c\|_2^2$$
with respect to x and minimizing
$$g_y(y) = f_y(y) + \frac{\varphi}{2} \|Ax + By - c\|_2^2$$
with respect to y is easier

Alternative direction method of multipliers

We use the following notations:

- $f_x^{(k+1)}(x)$ is the objective function of the x-update step:

$$f_x^{(k+1)}(x) = f_x(x) + \frac{\varphi}{2} \left\| Ax + By^{(k)} - c + u^{(k)} \right\|_2^2$$

- $f_y^{(k+1)}(y)$ is the objective function of the y-update step:

$$f_y^{(k+1)}(y) = f_y(y) + \frac{\varphi}{2} \left\| Ax^{(k+1)} + By - c + u^{(k)} \right\|_2^2$$

Alternative direction method of multipliers

When $A = I_n$ and $B = -I_n$, we have:

1. $$Ax + By^{(k)} - c + u^{(k)} = x - y^{(k)} - c + u^{(k)} = x - v_x^{(k+1)}$$

   where:

   $$v_x^{(k+1)} = y^{(k)} + c - u^{(k)}$$

2. $$Ax^{(k+1)} + By - c + u^{(k)} = x^{(k+1)} - y - c + u^{(k)} = v_y^{(k+1)} - y$$

   where:

   $$v_y^{(k+1)} = x^{(k+1)} - c + u^{(k)}$$

3. $$\begin{aligned} f_x^{(k+1)}(x) &= f_x(x) + \frac{\varphi}{2} \left\| x - v_x^{(k+1)} \right\|_2^2 \\ f_y^{(k+1)}(y) &= f_y(y) + \frac{\varphi}{2} \left\| y - v_y^{(k+1)} \right\|_2^2 \end{aligned}$$

Alternative direction method of multipliers

- We consider a problem of the form:

$$x^{\star} = \arg\min_{x} g(x)$$

The idea is then to write $g(x)$ as a separable function:

$$g(x) = g_1(x) + g_2(x)$$

and to consider the following equivalent ADMM problem:

$$\begin{aligned} \{x^{\star}, y^{\star}\} &= \arg\min_{(x,y)} f_x(x) + f_y(y) \\ \text{s.t.} & \quad x = y \end{aligned}$$

where $f_x(x) = g_1(x)$ and $f_y(y) = g_2(y)$

Alternative direction method of multipliers

- We consider a problem of the form:

$$\begin{array}{rcl} x^{\star} & = & \arg\min_{x} g(x) \\ & \text{s.t.} & x \in \Omega \end{array}$$

We have:

$$\begin{array}{rcl} \{x^{\star}, y^{\star}\} & = & \arg\min_{(x,y)} f_x(x) + f_y(y) \\ & \text{s.t.} & x = y \end{array}$$

where $f_x(x) = g(x)$, $f_y(y) = \mathbb{1}_{\Omega}(y)$ and:

$$\mathbb{1}_{\Omega}(y) = \begin{cases} 0 & \text{if} \quad y \in \Omega \\ +\infty & \text{if} \quad y \notin \Omega \end{cases}$$

Alternative direction method of multipliers

Special case

$$\Omega = \left\{ x : x^{-} \leq x \leq x^{+} \right\}$$

By setting $\varphi = 1$, the y-step becomes:

$$\begin{aligned} y^{(k+1)} &= \arg\min \left\{ \mathbb{1}_{\Omega}(y) + \frac{1}{2} \left\| x^{(k+1)} - y + u^{(k)} \right\|_2^2 \right\} \\ &= \mathbf{prox}_{f_y}\left(x^{(k+1)} + u^{(k)} \right) \end{aligned}$$

where the proximal operator is the box projection or the truncation operator:

$$\begin{aligned} \mathbf{prox}_{f_y}(v) &= x^{-} \odot \mathbb{1}\left\{ v < x^{-} \right\} + \\ & \quad v \odot \mathbb{1}\left\{ x^{-} \leq v \leq x^{+} \right\} + \\ & \quad x^{+} \odot \mathbb{1}\left\{ v > x^{+} \right\} \\ &= \mathcal{T}\left(v; x^{-}, x^{+} \right) \end{aligned}$$

Alternative direction method of multipliers

Special case

$$\Omega = \left\{x : x^{-} \leq x \leq x^{+}\right\}$$

The ADMM algorithm is then:

$$\begin{aligned}
x^{(k+1)} &= \arg\min \left\{ g(x) + \frac{1}{2} \left\| x - y^{(k)} + u^{(k)} \right\|_2^2 \right\} \\
y^{(k+1)} &= \mathbf{prox}_{f_y}\left(x^{(k+1)} + u^{(k)} \right) \\
u^{(k+1)} &= u^{(k)} + \left(x^{(k+1)} - y^{(k+1)} \right)
\end{aligned}$$

$\Rightarrow$ Solving the constrained optimization problem consists in solving the unconstrained optimization problem, applying the box projection and iterating these steps until convergence

Alternative direction method of multipliers

Lasso regression

The λ-problem of the lasso regression has the following ADMM formulation:

$$\begin{aligned}\left\{\beta^{\star},\bar{\beta}^{\star}\right\} &= \arg\min \frac{1}{2}(Y - X\beta)^{\top}(Y - X\beta) + \lambda\|\bar{\beta}\|_1 \\ \text{s.t.} &\quad \beta - \bar{\beta} = \mathbf{0}_m\end{aligned}$$

We have:

$$\begin{aligned}f_x(\beta) &= \frac{1}{2}(Y - X\beta)^{\top}(Y - X\beta) \\ &= \frac{1}{2}\beta^{\top}\left(X^{\top}X\right)\beta - \beta^{\top}\left(X^{\top}Y\right) + \frac{1}{2}Y^{\top}Y\end{aligned}$$

and:

$$f_y\left(\bar{\beta}\right) = \lambda\|\bar{\beta}\|_1$$

Alternative direction method of multipliers

The x-step is:

$$\beta^{(k+1)} = \arg\min_{\beta} \left\{ \frac{1}{2}\beta^\top \left(X^\top X\right)\beta - \beta^\top \left(X^\top Y\right) + \frac{\varphi}{2}\left\|\beta - \bar{\beta}^{(k)} + u^{(k)}\right\|_2^2 \right\}$$

Since we have:

$$\begin{aligned} \frac{\varphi}{2}\left\|\beta - \bar{\beta}^{(k)} + u^{(k)}\right\|_2^2 &= \frac{\varphi}{2}\beta^\top\beta - \varphi\beta^\top\left(\bar{\beta}^{(k)} - u^{(k)}\right) + \\ &\quad \frac{\varphi}{2}\left(\bar{\beta}^{(k)} - u^{(k)}\right)^\top\left(\bar{\beta}^{(k)} - u^{(k)}\right) \end{aligned}$$

we deduce that the x-update is a standard QP problem where:

$$f_x^{(k+1)}(\beta) = \frac{1}{2}\beta^\top\left(X^\top X + \varphi I_m\right)\beta - \beta^\top\left(X^\top Y + \varphi\left(\bar{\beta}^{(k)} - u^{(k)}\right)\right)$$

It follows that the solution is:

$$\begin{aligned} \beta^{(k+1)} &= \arg\min f_x^{(k+1)}(\beta) \\ &= \left(X^\top X + \varphi I_m\right)^{-1}\left(X^\top Y + \varphi\left(\bar{\beta}^{(k)} - u^{(k)}\right)\right) \end{aligned}$$

Alternative direction method of multipliers

The y-step is:

$$\begin{aligned}\bar{\beta}^{(k+1)} &= \arg\min_{\bar{\beta}} \left\{ \lambda \|\bar{\beta}\|_1 + \frac{\varphi}{2} \left\| \beta^{(k+1)} - \bar{\beta} + u^{(k)} \right\|_2^2 \right\} \\ &= \arg\min \left\{ \frac{1}{2} \left\| \bar{\beta} - \left(\beta^{(k+1)} + u^{(k)} \right) \right\|_2^2 + \frac{\lambda}{\varphi} \|\bar{\beta}\|_1 \right\}\end{aligned}$$

We recognize the soft-thresholding problem with $v = \beta^{(k+1)} + u^{(k)}$. We have:

$$\bar{\beta}^{(k+1)} = \mathcal{S}\left(\beta^{(k+1)} + u^{(k)}; \varphi^{-1}\lambda \right)$$

where:

$$\mathcal{S}(v; \lambda) = \operatorname{sign}(v) \cdot (|v| - \lambda)_+$$

Alternative direction method of multipliers

ADMM-Lasso algorithm (Boyd *et al.*, 2011)

Finally, the ADMM algorithm is made up of the following steps:

$$\left\{\begin{array}{l} \beta^{(k+1)} = \left(X^\top X + \varphi I_m\right)^{-1}\left(X^\top Y + \varphi\left(\bar{\beta}^{(k)} - u^{(k)}\right)\right) \\ \bar{\beta}^{(k+1)} = \mathcal{S}\left(\beta^{(k+1)} + u^{(k)}; \varphi^{-1}\lambda\right) \\ u^{(k+1)} = u^{(k)} + \left(\beta^{(k+1)} - \bar{\beta}^{(k+1)}\right) \end{array}\right.$$

Alternative direction method of multipliers

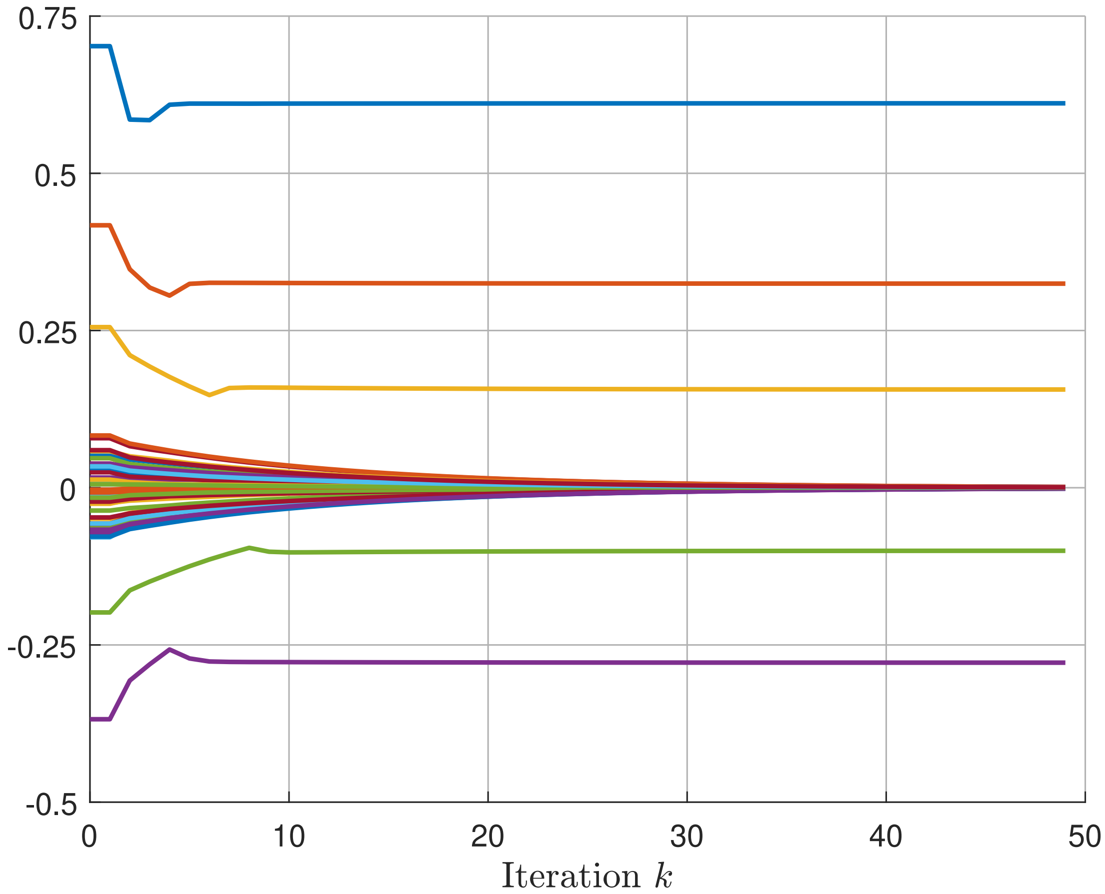


Figure 174: Convergence of the ADMM algorithm (Example 3, $\lambda = 900$)

Note: the initial values are the OLS estimates and we set $\varphi = \lambda$

Alternative direction method of multipliers

In practice, we use a time-varying parameter $\varphi^{(k)}$ (see Perrin and Roncalli, 2020).

Proximal operator

Definition

The proximal operator $\mathbf{prox}_f(v)$ of the function $f(x)$ is defined by:

$$\mathbf{prox}_f(v) = x^\star = \arg\min_x \left\{ f_v(x) = f(x) + \frac{1}{2}\|x - v\|_2^2 \right\}$$

Proximal operator

Example 4

We consider the scalar-valued logarithmic barrier function $f(x) = -\lambda \ln x$

Proximal operator

We have:

$$\begin{aligned} f_v(x) &= -\lambda \ln x + \frac{1}{2}(x-v)^2 \\ &= -\lambda \ln x + \frac{1}{2}x^2 - xv + \frac{1}{2}v^2 \end{aligned}$$

The first-order condition is $-\lambda x^{-1} + x - v = 0$. We obtain two roots with opposite signs:

$$x' = \frac{v - \sqrt{v^2 + 4\lambda}}{2} \text{ and } x'' = \frac{v + \sqrt{v^2 + 4\lambda}}{2}$$

Since the logarithmic function is defined for $x > 0$, we deduce that:

$$\mathbf{prox}_f(v) = \frac{v + \sqrt{v^2 + 4\lambda}}{2}$$

Proximal operator

In the case where $f(x) = \mathbb{1}_{\Omega}(x)$, we have:

$$\begin{aligned}
\mathbf{prox}_f(v) &= \arg\min_{x} \left\{ \mathbb{1}_{\Omega}(x) + \frac{1}{2}\left\| x - v \right\|_2^2 \right\} \\
&= \arg\min_{x \in \Omega} \left\{ \left\| x - v \right\|_2^2 \right\} \\
&= \mathcal{P}_{\Omega}(v)
\end{aligned}$$

where $\mathcal{P}_{\Omega}(v)$ is the standard projection of v onto Ω

Proximal operator

Table 101: Projection for some simple polyhedra

Notation	Ω	$\mathcal{P}_{\Omega}(v)$
$\mathcal{A}_{ffineset}[A, B]$	$Ax = B$	$v - A^{\dagger}(Av - B)$
$\mathcal{H}_{yperplane}[a, b]$	$a^{\top}x = b$	$v - \dfrac{(a^{\top}v - b)}{\lVert a \rVert_2^2} a$
$\mathcal{H}_{alfspace}[c, d]$	$c^{\top}x \leq d$	$v - \dfrac{(c^{\top}v - d)_+}{\lVert c \rVert_2^2} c$
$\mathcal{B}_{ox}[x^-, x^+]$	$x^- \leq x \leq x^+$	$\mathcal{T}(v; x^-, x^+)$

Source: Parikh and Boyd (2014)

Note: $A^{\dagger}$ is the Moore-Penrose pseudo-inverse of A, and $\mathcal{T}(v; x^-, x^+)$ is the truncation operator

Remark: No analytical formula for the (multi-dimensional) inequality constraint $Cx \leq D \Rightarrow$ it may be solved using the Dykstra's algorithm

Proximal operator

Separable sum

If $f(x) = \sum_{i=1}^{n} f_i(x_i)$ is fully separable, then the proximal of $f(v)$ is the vector of the proximal operators applied to each scalar-valued function $f_i(x_i)$:

$$\mathbf{prox}_f(v) = \begin{pmatrix} \mathbf{prox}_{f_1}(v_1) \\ \vdots \\ \mathbf{prox}_{f_n}(v_n) \end{pmatrix}$$

Proximal operator

If $f(x) = -\lambda \ln x$, we have:

$$\mathbf{prox}_f(v) = \frac{v + \sqrt{v^2 + 4\lambda}}{2}$$

In the case of the vector-valued logarithmic barrier $f(x) = -\lambda \sum_{i=1}^{n} \ln x_i$, we deduce that:

$$\mathbf{prox}_f(v) = \frac{v + \sqrt{v \odot v + 4\lambda}}{2}$$

Proximal operator

Moreau decomposition theorem

We have:

$$\mathbf{prox}_f(v) + \mathbf{prox}_{f^*}(v) = v$$

where f^* is the convex conjugate of f.

Application

If $f(x)$ is a ℓ_q-norm function, then $f^*(x) = \mathbb{1}_{\mathcal{B}_p}(x)$ where $\mathcal{B}_p$ is the ℓ_p unit ball and $p^{-1} + q^{-1} = 1$. Since we have $\mathbf{prox}_{f^*}(v) = \mathcal{P}_{\mathcal{B}_p}(v)$, we deduce that:

$$\mathbf{prox}_f(v) + \mathcal{P}_{\mathcal{B}_p}(v) = v$$

The proximal of the ℓ_p-ball can be deduced from the proximal operator of the ℓ_q-norm function.

Proximal operator

Table 102: Proximal of the ℓ_p-norm function $f(x) = \|x\|_p$

p	$\mathbf{prox}_{\lambda f}(v)$
$p = 1$	$\mathcal{S}(v;\lambda) = \operatorname{sign}(v) \odot (\lvert v\rvert - \lambda \mathbf{1}_n)_+$
$p = 2$	$\left(1 - \dfrac{\lambda}{\max(\lambda, \lVert v\rVert_2)}\right) v$
$p = \infty$	$\operatorname{sign}(v) \odot \mathbf{prox}_{\lambda \max x}(\lvert v\rvert)$

We have:

$$\mathbf{prox}_{\lambda \max x}(v) = \min(v, s^\star)$$

where $s^\star$ is the solution of the following equation:

$$s^\star = \left\{ s \in \mathbb{R} : \sum_{i=1}^{n} (v_i - s)_+ = \lambda \right\}$$

Proximal operator

Table 103: Proximal of the ℓ_p-ball $\mathcal{B}_p(c,\lambda) = \left\{x \in \mathbb{R}^n : \|x - c\|_p \leq \lambda\right\}$ when c is equal to $\mathbf{0}_n$

p	$\mathcal{P}_{\mathcal{B}_p(\mathbf{0}_n,\lambda)}(v)$	q
$p = 1$	$v - \operatorname{sign}(v) \odot \mathbf{prox}_{\lambda \max x}(\lvert v \rvert)$	$q = \infty$
$p = 2$	$v - \mathbf{prox}_{\lambda \|x\|_2}(v)$	$q = 2$
$p = \infty$	$\mathcal{T}(v; -\lambda, \lambda)$	$q = 1$

Proximal operator

Scaling and translation

Let us define $g(x) = f(ax + b)$ where $a \neq 0$. We have:

$$\mathbf{prox}_g(v) = \frac{\mathbf{prox}_{a^2 f}(av + b) - b}{a}$$

Application

We can use this property when the center c of the ℓ_p ball is not equal to $\mathbf{0}_n$. Since we have $\mathbf{prox}_g(v) = \mathbf{prox}_f(v - c) + c$ where $g(x) = f(x - c)$ and the equivalence $\mathcal{B}_p(\mathbf{0}_n, \lambda) = \{x \in \mathbb{R}^n : f(x) \leq \lambda\}$ where $f(x) = \|x\|_p$, we deduce that:

$$\mathcal{P}_{\mathcal{B}_p(c,\lambda)}(v) = \mathcal{P}_{\mathcal{B}_p(\mathbf{0}_n,\lambda)}(v - c) + c$$

Application to the τ-problem of the lasso regression

We have:

$$\begin{aligned} \hat{\beta}(\tau) &= \arg\min_{\beta} \frac{1}{2}(Y - X\beta)^\top (Y - X\beta) \\ \text{s.t.} &\quad \|\beta\|_1 \leq \tau \end{aligned}$$

The ADMM formulation is:

$$\begin{aligned} \left\{\beta^\star, \bar{\beta}^\star\right\} &= \arg\min_{(\beta,\bar{\beta})} \frac{1}{2}(Y - X\beta)^\top (Y - X\beta) + \mathbb{1}_\Omega\left(\bar{\beta}\right) \\ \text{s.t.} &\quad \beta = \bar{\beta} \end{aligned}$$

where $\Omega = \mathcal{B}_1(\mathbf{0}_m, \tau)$ is the centered ℓ_1 ball with radius τ

Application to the τ-problem of the lasso regression

1. The x-update is:

$$\begin{aligned}\beta^{(k+1)} &= \arg\min_{\beta}\left\{\frac{1}{2}\left(Y-X\beta\right)^{\top}\left(Y-X\beta\right)+\frac{\varphi}{2}\left\|\beta-\bar{\beta}^{(k)}+u^{(k)}\right\|_2^2\right\} \\ &= \left(X^{\top}X+\varphi I_m\right)^{-1}\left(X^{\top}Y+\varphi\left(\bar{\beta}^{(k)}-u^{(k)}\right)\right)\end{aligned}$$

where $v_x^{(k+1)}=\bar{\beta}^{(k)}-u^{(k)}$

Application to the τ-problem of the lasso regression

2. The y-update is:

$$
\begin{aligned}
\bar{\beta}^{(k+1)} &= \arg\min_{\bar{\beta}} \left\{ \mathbb{1}_{\Omega}\left(\bar{\beta}\right) + \frac{\varphi}{2}\left\| \beta^{(k+1)} - \bar{\beta} + u^{(k)} \right\|_2^2 \right\} \\
&= \mathbf{prox}_{f_y}\left(\beta^{(k+1)} + u^{(k)}\right) \\
&= \mathcal{P}_{\Omega}\left(v_y^{(k+1)}\right) \\
&= v_y^{(k+1)} - \operatorname{sign}\left(v_y^{(k+1)}\right) \odot \mathbf{prox}_{\tau \max x}\left(\left|v_y^{(k+1)}\right|\right)
\end{aligned}
$$

where $v_y^{(k+1)} = \beta^{(k+1)} + u^{(k)}$

Application to the τ-problem of the lasso regression

3. The u-update is:

$$u^{(k+1)} = u^{(k)} + \beta^{(k+1)} - \bar{\beta}^{(k+1)}$$

Application to the τ-problem of the lasso regression

ADMM-Lasso algorithm

The ADMM algorithm is :

$$\left\{ \begin{array}{l} \beta^{(k+1)} = \left(X^\top X + \varphi I_m\right)^{-1} \left(X^\top Y + \varphi \left(\bar{\beta}^{(k)} - u^{(k)}\right)\right) \\ \bar{\beta}^{(k+1)} = \left\{ \begin{array}{ll} \mathcal{S}\left(\beta^{(k+1)} + u^{(k)}; \varphi^{-1}\lambda\right) & (\lambda\text{-problem}) \\ \mathcal{P}_{\mathcal{B}_1(\mathbf{0}_m,\tau)}\left(\beta^{(k+1)} + u^{(k)}\right) & (\tau\text{-problem}) \end{array} \right. \\ u^{(k+1)} = u^{(k)} + \left(\beta^{(k+1)} - \bar{\beta}^{(k+1)}\right) \end{array} \right.$$

Remark

The ADMM algorithm is similar for λ- and τ-problems since the only difference concerns the y-step. However, the τ-problem is easier to solve with the ADMM algorithm from a practical point of view, because the y-update of the τ-problem is independent of the penalization parameter φ.

Derivation of the soft-thresholding operator

We consider the following equation:

$$cx - v + \lambda \partial |x| \in 0$$

where $c > 0$ and $\lambda > 0$. Since we have $\partial |x| = \operatorname{sign}(x)$, we deduce that:

$$x^\star = \begin{cases} c^{-1}(v + \lambda) & \text{if } x^\star < 0 \\ 0 & \text{if } x^\star = 0 \\ c^{-1}(v - \lambda) & \text{if } x^\star > 0 \end{cases}$$

If $x^\star < 0$ or $x^\star > 0$, then we have $v + \lambda < 0$ or $v - \lambda > 0$. This is equivalent to set $|v| > \lambda > 0$. The case $x^\star = 0$ implies that $|v| \leq \lambda$. We deduce that:

$$x^\star = c^{-1} \cdot \mathcal{S}(v; \lambda)$$

where $\mathcal{S}(v; \lambda)$ is the soft-thresholding operator:

$$\begin{aligned} \mathcal{S}(v; \lambda) &= \begin{cases} 0 & \text{if } |v| \leq \lambda \\ v - \lambda \operatorname{sign}(v) & \text{otherwise} \end{cases} \\ &= \operatorname{sign}(v) \cdot (|v| - \lambda)_+ \end{aligned}$$

Derivation of the soft-thresholding operator

We use the result on the separable sum

Remark

If $f(x) = \lambda \|x\|_1$, we have $f(x) = \lambda \sum_{i=1}^{n} |x_i|$ and $f_i(x_i) = \lambda |x_i|$. We deduce that the proximal operator of $f(x)$ is the vector formulation of the soft-thresholding operator:

$$\mathbf{prox}_{\lambda\|x\|_1}(v) = \begin{pmatrix} \mathrm{sign}(v_1) \cdot (|v_1| - \lambda)_+ \\ \vdots \\ \mathrm{sign}(v_n) \cdot (|v_n| - \lambda)_+ \end{pmatrix} = \mathrm{sign}(v) \odot (|v| - \lambda \mathbf{1}_n)_+$$

The soft-thresholding operator is the proximal operator of the ℓ_1-norm $f(x) = \|x\|_1$. Indeed, we have $\mathbf{prox}_f(v) = \mathcal{S}(v; 1)$ and $\mathbf{prox}_{\lambda f}(v) = \mathcal{S}(v; \lambda)$.

Dykstra's algorithm

We consider the following optimization problem:

$$\begin{aligned} x^{\star} &= \arg\min f_x(x) \\ \text{s.t.} &\quad x \in \Omega \end{aligned}$$

where Ω is a complex set of constraints:

$$\Omega = \Omega_1 \cap \Omega_2 \cap \cdots \Omega_m$$

We set $y = x$ and $f_y(y) = \mathbb{1}_{\Omega}(y)$. The ADMM algorithm becomes

$$\begin{aligned} x^{(k+1)} &= \arg\min \left\{ f_x(x) + \frac{\varphi}{2} \left\| x - y^{(k)} + u^{(k)} \right\|_2^2 \right\} \\ v^{(k)} &= x^{(k+1)} + u^{(k)} \\ y^{(k+1)} &= \mathcal{P}_{\Omega}\left(v^{(k)}\right) \\ u^{(k+1)} &= u^{(k)} + \left(x^{(k+1)} - y^{(k+1)}\right) \end{aligned}$$

How to compute $\mathcal{P}_{\Omega}(v)$?

Dykstra's algorithm

More generally, we consider the proximal optimization problem where the function $f(x)$ is the convex sum of basic functions $f_j(x)$:

$$x^{\star} = \arg\min_{x} \left\{ \sum_{j=1}^{m} f_j(x) + \frac{1}{2} \left\| x - v \right\|_2^2 \right\}$$

and the proximal of each basic function is known.

How to find the solution $x^{\star}$?

Dykstra's algorithm

The case $m = 2$

- We know the proximal solution of the ℓ_1-norm function $f_1(x) = \lambda_1 \|x\|_1$
- We know the proximal solution of the logarithmic barrier function $f_2(x) = \lambda_2 \sum_{i=1}^{n} \ln x_i$
- We don't know how to compute the proximal operator of $f(x) = f_1(x) + f_2(x)$:

$$
\begin{aligned}
x^\star &= \arg\min_x f_1(x) + f_2(x) + \frac{1}{2}\|x - v\|_2^2 \\
&= \mathbf{prox}_f(v)
\end{aligned}
$$

Dykstra's algorithm

The case $m = 2$

The Dykstra's algorithm consists in the following iterations:

$$\left\{\begin{array}{l} x^{(k+1)} = \mathbf{prox}_{f_1}\left(y^{(k)} + p^{(k)}\right) \\ p^{(k+1)} = y^{(k)} + p^{(k)} - x^{(k+1)} \\ y^{(k+1)} = \mathbf{prox}_{f_2}\left(x^{(k+1)} + q^{(k)}\right) \\ q^{(k+1)} = x^{(k+1)} + q^{(k)} - y^{(k+1)} \end{array}\right.$$

where $x^{(0)} = y^{(0)} = v$ and $p^{(0)} = q^{(0)} = \mathbf{0}_n$

Dykstra's algorithm

The case $m = 2$

This algorithm is related to the Douglas-Rachford splitting framework:

$$\left\{\begin{array}{l} x^{\left(k+\frac{1}{2}\right)} = \mathbf{prox}_{f_1}\left(x^{(k)} + p^{(k)}\right) \\ p^{(k+1)} = p^{(k)} - \Delta_{1/2} x^{\left(k+\frac{1}{2}\right)} \\ x^{(k+1)} = \mathbf{prox}_{f_2}\left(x^{\left(k+\frac{1}{2}\right)} + q^{(k)}\right) \\ q^{(k+1)} = q^{(k)} - \Delta_{1/2} x^{(k+1)} \end{array}\right.$$

where $\Delta_h x^{(k)} = x^{(k)} - x^{(k-h)}$

Dykstra's algorithm

The case $m = 2$

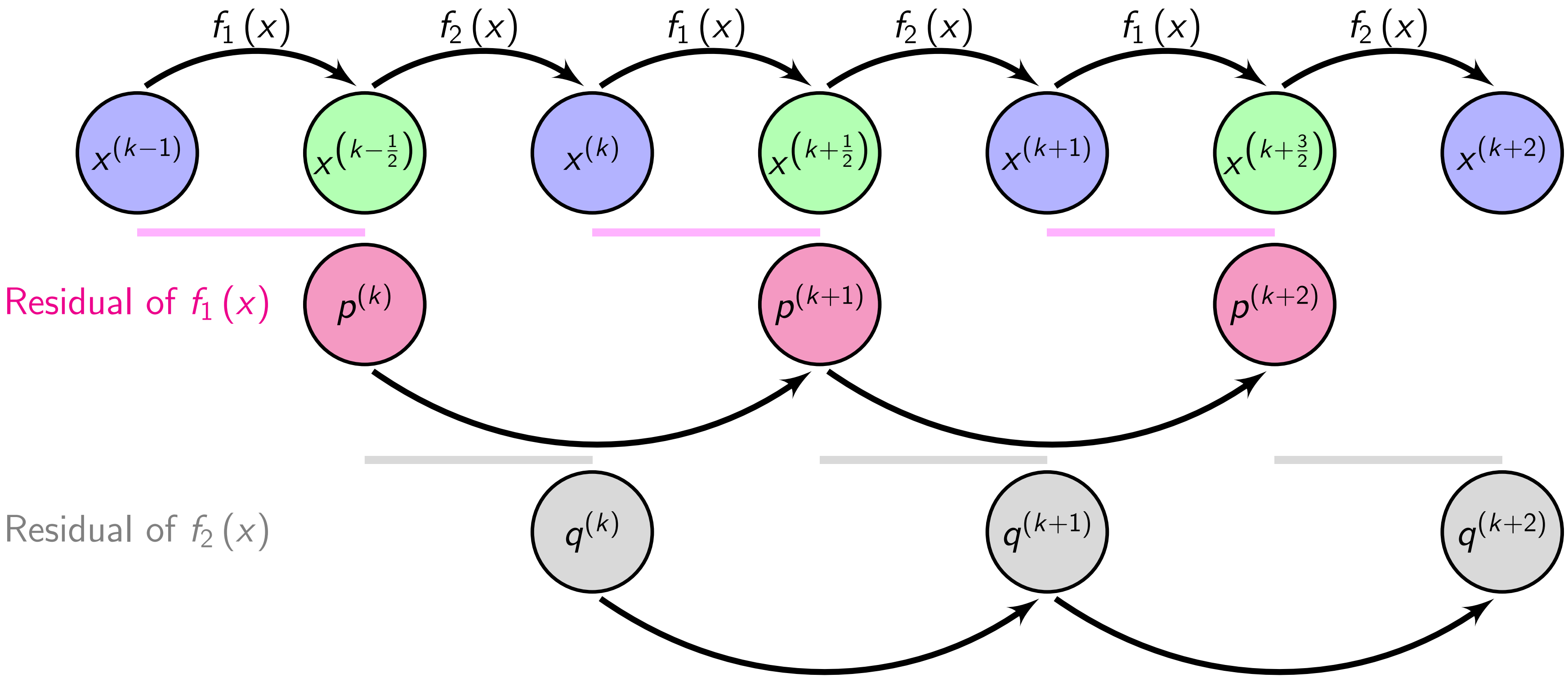


Figure 175: Splitting method of the Dykstra's algorithm

Dykstra's algorithm

The case $m > 2$

The case $m > 2$ is a generalization of the previous algorithm by considering m residuals:

1. The x-update is:

$$x^{(k+1)} = \mathbf{prox}_{f_{j(k)}} \left(x^{(k)} + z^{(k+1-m)} \right)$$

2. The z-update is:

$$z^{(k+1)} = x^{(k)} + z^{(k+1-m)} - x^{(k+1)}$$

where $x^{(0)} = v$, $z^{(k)} = \mathbf{0}_n$ for $k < 0$ and $j(k) = \mathrm{mod}(k+1, m)$ denotes the modulo operator taking values in $\{1, \ldots, m\}$

Remark

The variable $x^{(k)}$ is updated at each iteration while the residual $z^{(k)}$ is updated every m iterations. This implies that the basic function $f_j(x)$ is related to the residuals $z^{(j)}$, $z^{(j+m)}$, $z^{(j+2m)}$, etc.

Dykstra's algorithm

The case $m > 2$

Tibshirani (2017) proposes to write the Dykstra's algorithm by using two iteration indices k and j. The main index k refers to the cycle, whereas the sub-index j refers to the constraint number

The Dykstra's algorithm becomes:

1. The x-update is:

$$x^{(k+1,j)} = \mathbf{prox}_{f_j}\left(x^{(k+1,j-1)} + z^{(k,j)}\right)$$

2. The z-update is:

$$z^{(k+1,j)} = x^{(k+1,j-1)} + z^{(k,j)} - x^{(k+1,j)}$$

where $x^{(1,0)} = v$, $z^{(k,j)} = \mathbf{0}_n$ for $k = 0$ and $x^{(k+1,0)} = x^{(k,m)}$

Dykstra's algorithm

The case $m > 2$

The Dykstra's algorithm is particularly efficient when we consider the projection problem:

$$x^{\star} = \mathcal{P}_{\Omega}\left(v\right)$$

where:

$$\Omega = \Omega_1 \cap \Omega_2 \cap \cdots \cap \Omega_m$$

Indeed, the Dykstra's algorithm becomes:

1. The x-update is:

$$x^{(k+1,j)} = \mathbf{prox}_{f_j}\left(x^{(k+1,j-1)} + z^{(k,j)}\right) = \mathcal{P}_{\Omega_j}\left(x^{(k+1,j-1)} + z^{(k,j)}\right)$$

2. The z-update is:

$$z^{(k+1,j)} = x^{(k+1,j-1)} + z^{(k,j)} - x^{(k+1,j)}$$

where $x^{(1,0)} = v$, $z^{(k,j)} = \mathbf{0}_n$ for $k = 0$ and $x^{(k+1,0)} = x^{(k,m)}$

Dykstra's algorithm

Successive projections of $\mathcal{P}_{\Omega_j}\left(x^{(k+1,j-1)}\right)$ do not work!

Successive projections of $\mathcal{P}_{\Omega_j}\left(x^{(k+1,j-1)} + z^{(k,j)}\right)$ do work!

Dykstra's algorithm

Table 104: Solving the proximal problem with linear inequality constraints

The goal is to compute the solution $x^{\star} = \mathbf{prox}_f(v)$ where $f(x) = \mathbb{1}_{\Omega}(x)$ and $\Omega = \{x \in \mathbb{R}^n : Cx \leq D\}$
We initialize $x^{(0,m)} \leftarrow v$
We set $z^{(0,1)} \leftarrow \mathbf{0}_n, \ldots, z^{(0,m)} \leftarrow \mathbf{0}_n$
$k \leftarrow 0$
repeat
  $x^{(k+1,0)} \leftarrow x^{(k,m)}$
  for $j = 1 : m$ **do**
    The x-update is:

$$x^{(k+1,j)} = x^{(k+1,j-1)} + z^{(k,j)} - \frac{\left(c_{(j)}^{\top} x^{(k+1;j-1)} + c_{(j)}^{\top} z^{(k,j)} - d_{(j)}\right)_+}{\left\| c_{(j)} \right\|_2^2} c_{(j)}$$

    The z-update is:

$$z^{(k+1,j)} = x^{(k+1,j-1)} + z^{(k,j)} - x^{(k+1,j)}$$

  end for
  $k \leftarrow k + 1$
until Convergence
return $x^{\star} \leftarrow x^{(k,m)}$

Dykstra's algorithm

Table 105: Solving the proximal problem with general linear constraints

The goal is to compute the solution $x^{\star} = \mathbf{prox}_f(v)$ where $f(x) = \mathbb{1}_{\Omega}(x)$, $\Omega = \Omega_1 \cap \Omega_2 \cap \Omega_3$, $\Omega_1 = \{x \in \mathbb{R}^n : Ax = B\}$, $\Omega_2 = \{x \in \mathbb{R}^n : Cx \leq D\}$ and $\Omega_3 = \{x \in \mathbb{R}^n : x^- \leq x \leq x^+\}$

We initialize $x_m^{(0)} \leftarrow v$

We set $z_1^{(0)} \leftarrow \mathbf{0}_n$, $z_2^{(0)} \leftarrow \mathbf{0}_n$ and $z_3^{(0)} \leftarrow \mathbf{0}_n$

$k \leftarrow 0$

repeat

$x_0^{(k+1)} \leftarrow x_m^{(k)}$

$x_1^{(k+1)} \leftarrow x_0^{(k+1)} + z_1^{(k)} - A^{\dagger}\left(Ax_0^{(k+1)} + Az_1^{(k)} - B\right)$

$z_1^{(k+1)} \leftarrow x_0^{(k+1)} + z_1^{(k)} - x_1^{(k+1)}$

$x_2^{(k+1)} \leftarrow \mathcal{P}_{\Omega_2}\left(x_1^{(k+1)} + z_2^{(k)}\right)$

$z_2^{(k+1)} \leftarrow x_1^{(k+1)} + z_2^{(k)} - x_2^{(k+1)}$

$x_3^{(k+1)} \leftarrow \mathcal{T}\left(x_2^{(k+1)} + z_3^{(k)}; x^-, x^+\right)$

$z_3^{(k+1)} \leftarrow x_2^{(k+1)} + z_3^{(k)} - x_3^{(k+1)}$

$k \leftarrow k + 1$

until Convergence

return $x^{\star} \leftarrow x_3^{(k)}$

▶ Previous algorithm

Dykstra's algorithm

Remark

Since we have:

$$\frac{1}{2}\left\| x - v \right\|_2^2 = \frac{1}{2} x^\top x - x^\top v + \frac{1}{2} v^\top v$$

the two previous problems can be cast into a QP problem:

$$\begin{aligned} x^\star & = & \arg\min_x \frac{1}{2} x^\top I_n x - x^\top v \\ & \text{s.t.} & x \in \Omega \end{aligned}$$

Dykstra's algorithm

Dykstra's algorithm versus QP algorithm

- The vector v is defined by the elements $v_i = \ln\left(1 + i^2\right)$
- The set of constraints is:

$$\Omega = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i \leq \frac{1}{2}, \sum_{i=1}^n e^{-i} x_i \geq 0 \right\}$$

- Using a Matlab implementation, we find that the computational time of the Dykstra's algorithm when n is equal to 10 million is equal to the QP algorithm when n is equal to 12 500!
- The QP algorithm requires to store the matrix I_n — impossible when $n > 10^5$. For instance, the size of I_n is equal to 7450.6 GB when $n = 10^6$

Application to portfolio allocation

Table 106: Some objective functions used in portfolio optimization

Item	Portfolio	$f(x)$	Reference
(1)	MVO	$\frac{1}{2}x^\top \Sigma x - \gamma x^\top \mu$	Markowitz (1952)
(2)	GMV	$\frac{1}{2}x^\top \Sigma x$	Jagganathan and Ma (2003)
(3)	MDP	$\ln\left(\sqrt{x^\top \Sigma x}\right) - \ln\left(x^\top \sigma\right)$	Choueifaty and Coignard (2008)
(4)	KL	$\sum_{i=1}^n x_i \ln\left(x_i / \tilde{x}_i\right)$	Bera and Park (2008)
(5)	ERC	$\frac{1}{2}x^\top \Sigma x - \lambda \sum_{i=1}^n \ln x_i$	Maillard *et al.* (2010)
(6)	RB	$\mathcal{R}(x) - \lambda \sum_{i=1}^n \mathcal{RB}_i \cdot \ln x_i$	Roncalli (2015)
(7)	RQE	$\frac{1}{2}x^\top D x$	Carmichael *et al.* (2018)

Application to portfolio allocation

Table 107: Some regularization penalties used in portfolio optimization

Item	Regularization	$\mathfrak{R}(x)$	Reference
(8)	Ridge	$\lambda \left\Vert x - \tilde{x} \right\Vert_2^2$	DeMiguel *et al.* (2009)
(9)	Lasso	$\lambda \left\Vert x - \tilde{x} \right\Vert_1$	Brodie *at al.* (2009)
(10)	Log-barrier	$-\sum_{i=1}^n \lambda_i \ln x_i$	Roncalli (2013)
(11)	Shannon's entropy	$\lambda \sum_{i=1}^n x_i \ln x_i$	Yu *et al.* (2014)

Application to portfolio allocation

Table 108: Some constraints used in portfolio optimization

Item	Constraint	Ω
(12)	No cash and leverage	$\sum_{i=1}^{n} x_i = 1$
(13)	No short selling	$x_i \geq 0$
(14)	Weight bounds	$x_i^- \leq x_i \leq x_i^+$
(15)	Asset class limits	$c_j^- \leq \sum_{i \in \mathcal{C}_j} x_i \leq c_j^+$
(16)	Turnover	$\sum_{i=1}^{n} \lvert x_i - \tilde{x}_i \rvert \leq \boldsymbol{\tau}^+$
(17)	Transaction costs	$\sum_{i=1}^{n} \left(c_i^- \left(\tilde{x}_i - x_i \right)_+ + c_i^+ \left(x_i - \tilde{x}_i \right)_+ \right) \leq \boldsymbol{c}^+$
(18)	Leverage limit	$\sum_{i=1}^{n} \lvert x_i \rvert \leq \mathcal{L}^+$
(19)	Long/short exposure	$-\mathcal{LS}^- \leq \sum_{i=1}^{n} x_i \leq \mathcal{LS}^+$
(20)	Benchmarking	$\sqrt{\left(x - \tilde{x} \right)^\top \Sigma \left(x - \tilde{x} \right)} \leq \sigma^+$
(21)	Tracking error floor	$\sqrt{\left(x - \tilde{x} \right)^\top \Sigma \left(x - \tilde{x} \right)} \geq \sigma^-$
(22)	Active share floor	$\frac{1}{2} \sum_{i=1}^{n} \lvert x_i - \tilde{x}_i \rvert \geq \mathcal{AS}^-$
(23)	Number of active bets	$\left(x^\top x \right)^{-1} \geq \mathcal{N}^-$

Application to portfolio allocation

Most of portfolio optimization problems are a combination of:

1. an objective function (Table 106)
2. one or two regularization penalty functions (Table 107)
3. some constraints (Table 108)

Perrin and Roncalli (2020) solve **all these problems** using CCD, ADMM, Dykstra and the appropriate proximal functions. For that, they derive:

- the semi-analytical solution of the x-step for all objective functions
- the proximal solution of the y-step for all regularization penalty functions and constraints

Herfindahl-MV optimization

Formulation of the mathematical problem

- The second generation of minimum variance strategies uses a global diversification constraint
- The most popular solution is based on the Herfindahl index:

$$\mathcal{H}(x) = \sum_{i=1}^{n} x_i^2$$

- The effective number of bets is the inverse of the Herfindahl index:

$$\mathcal{N}(x) = \mathcal{H}(x)^{-1}$$

- The optimization program is:

$$\begin{aligned} x^{\star} & = & \arg\min_{x} \frac{1}{2} x^{\top} \Sigma x \\ & \text{s.t.} & \left\{ \begin{array}{l} \mathbf{1}_n^{\top} x = 1 \\ \mathbf{0}_n \leq x \leq x^{+} \\ \mathcal{N}(x) \geq \mathcal{N}^{-} \end{array} \right. \end{aligned}$$

where $\mathcal{N}^{-}$ is the minimum number of effective bets.

Herfindahl-MV optimization

The QP solution

- The Herfindhal constraint is equivalent to:

$$\begin{aligned} \mathcal{N}(x) \geq \mathcal{N}^- & \Leftrightarrow & \left(x^\top x\right)^{-1} \geq \mathcal{N}^- \\ & \Leftrightarrow & x^\top x \leq \frac{1}{\mathcal{N}^-} \end{aligned}$$

- The QP problem is:

$$\begin{aligned} x^\star(\lambda) &= \arg\min_x \frac{1}{2} x^\top \Sigma x + \lambda x^\top x = \frac{1}{2} x^\top \left(\Sigma + 2\lambda I_n\right) x \\ & \text{s.t.} \quad \left\{ \begin{array}{l} \mathbf{1}_n^\top x = 1 \\ \mathbf{0}_n \leq x \leq x^+ \end{array} \right. \end{aligned}$$

  where $\lambda \geq 0$ is a scalar
- We have $\mathcal{N}(x) \in [\mathcal{N}(x^\star(0)), n]$
- The optimal value $\lambda^\star$ is found using the bi-section algorithm such that $\mathcal{N}(x^\star(\lambda)) = \mathcal{N}^-$

Herfindahl-MV optimization

The ADMM solution (first version)

- The ADMM form is:

$$\begin{aligned}\{x^{\star}, y^{\star}\} &= \arg\min_{(x,y)} \frac{1}{2}x^{\top}\Sigma x + \mathbb{1}_{\Omega_1}(x) + \mathbb{1}_{\Omega_2}(y) \\ \text{s.t.} &\quad x = y\end{aligned}$$

  where $\Omega_1 = \left\{x \in \mathbb{R}^n : \mathbf{1}_n^{\top}x = 1, \mathbf{0}_n \leq x \leq x^+\right\}$ and $\Omega_2 = \mathcal{B}_2\left(\mathbf{0}_n, \sqrt{\frac{1}{\mathcal{N}^-}}\right)$

- The x-update is a QP problem:

$$x^{(k+1)} = \arg\min_{x}\left\{\frac{1}{2}x^{\top}\left(\Sigma + \varphi I_n\right)x - \varphi x^{\top}\left(y^{(k)} - u^{(k)}\right) + \mathbb{1}_{\Omega_1}(x)\right\}$$

- The y-update is:

$$y^{(k+1)} = \frac{x^{(k+1)} + u^{(k)}}{\max\left(1, \sqrt{\mathcal{N}^-}\left\|x^{(k+1)} + u^{(k)}\right\|_2\right)}$$

Herfindahl-MV optimization

The ADMM solution (second version)

- A better approach is to write the problem as follows:

$$\begin{aligned} \{x^\star, y^\star\} &= \arg\min_{(x,y)} \frac{1}{2} x^\top \Sigma x + \mathbb{1}_{\Omega_3}(x) + \mathbb{1}_{\Omega_4}(y) \\ \text{s.t.} &\quad x = y \end{aligned}$$

where $\Omega_3 = \mathcal{H}_{yperplane}\left[\mathbf{1}_n, 1\right]$ and $\Omega_4 = \mathcal{B}_{ox}\left[\mathbf{0}_n, x^+\right] \cap \mathcal{B}_2\left(\mathbf{0}_n, \sqrt{\frac{1}{\mathcal{N}^-}}\right)$

- The x-update is:

$$x^{(k+1)} = \left(\Sigma + \varphi I_n\right)^{-1} \left(\varphi\left(y^{(k)} - u^{(k)}\right) + \frac{1 - \mathbf{1}_n^\top \left(\Sigma + \varphi I_n\right)^{-1} \varphi\left(y^{(k)} - u^{(k)}\right)}{\mathbf{1}_n^\top \left(\Sigma + \varphi I_n\right)^{-1} \mathbf{1}_n} \mathbf{1}_n\right)$$

- The y-update is:

$$y^{(k+1)} = \mathcal{P}_{\mathcal{B}\text{ox}-\mathcal{B}\text{all}}\left(x^{(k+1)} + u^{(k)}; \mathbf{0}_n, x^+, \mathbf{0}_n, \sqrt{\frac{1}{\mathcal{N}^-}}\right)$$

where $\mathcal{P}_{\mathcal{B}\text{ox}-\mathcal{B}\text{all}}$ corresponds to the Dykstra's algorithm given by Perrin and Roncalli (2020)

Herfindahl-MV optimization

Remark

If we compare the computational time of the three approaches, we observe that the best method is the second version of the ADMM algorithm:

$$\mathcal{CT}\left(\mathrm{QP}; n = 1000\right) = 50 \times \mathcal{CT}\left(\mathrm{ADMM}_2; n = 1000\right)$$

$$\mathcal{CT}\left(\mathrm{ADMM}_1; n = 1000\right) = 400 \times \mathcal{CT}\left(\mathrm{ADMM}_2; n = 1000\right)$$

Herfindahl-MV optimization

The QP solution

Example 5

We consider an investment universe of eight stocks. We assume that their volatilities are 21%, 20%, 40%, 18%, 35%, 23%, **7%** and 29%. The correlation matrix is defined as follows:

$$\rho = \begin{pmatrix} 100\% & & & & & & & \\ 80\% & 100\% & & & & & & \\ 70\% & 75\% & 100\% & & & & & \\ 60\% & 65\% & 90\% & 100\% & & & & \\ 70\% & 50\% & 70\% & 85\% & 100\% & & & \\ 50\% & 60\% & 70\% & 80\% & 60\% & 100\% & & \\ 70\% & 50\% & 70\% & 75\% & 80\% & 50\% & 100\% & \\ 60\% & 65\% & 70\% & 75\% & 65\% & 70\% & 80\% & 100\% \end{pmatrix}$$

Herfindahl-MV optimization

Table 109: Minimum variance portfolios (in %)

$\mathcal{N}^-$	1.00	2.00	3.00	4.00	5.00	6.00	6.50	7.00	7.50	8.00
$x_1^\star$	0.00	3.22	9.60	13.83	15.18	15.05	14.69	14.27	13.75	12.50
$x_2^\star$	0.00	12.75	14.14	15.85	16.19	15.89	15.39	14.82	14.13	12.50
$x_3^\star$	0.00	0.00	0.00	0.00	0.00	0.07	2.05	4.21	6.79	12.50
$x_4^\star$	0.00	10.13	15.01	17.38	17.21	16.09	15.40	14.72	13.97	12.50
$x_5^\star$	0.00	0.00	0.00	0.00	0.71	5.10	6.33	7.64	9.17	12.50
$x_6^\star$	0.00	5.36	8.95	12.42	13.68	14.01	13.80	13.56	13.25	12.50
$x_7^\star$	100.00	68.53	52.31	40.01	31.52	25.13	22.92	20.63	18.00	12.50
$x_8^\star$	0.00	0.00	0.00	0.50	5.51	8.66	9.41	10.14	10.95	12.50
$\lambda^\star$ (in %)	0.00	1.59	3.10	5.90	10.38	18.31	23.45	31.73	49.79	∞

Note: the upper bound x^+ is set to $\mathbf{1}_n$. The solutions are those found by the ADMM algorithm. We also report the value of $\lambda^\star$ found by the bi-section algorithm when we use the QP algorithm.

ERC portfolio optimization

We recall that:

$$x^{\star} = \arg\min_{x} \frac{1}{2} x^{\top} \Sigma x - \lambda \sum_{i=1}^{n} \ln x_i$$

and:

$$x_{\mathrm{erc}} = \frac{x^{\star}}{\mathbf{1}_n^{\top} x^{\star}}$$

ERC portfolio optimization

The CCD solution

- The first-order condition $(\Sigma x)_i - \lambda x_i^{-1} = 0$ implies that:

$$x_i^2 \sigma_i^2 + x_i \sigma_i \sum_{j \neq i} x_j \rho_{i,j} \sigma_j - \lambda = 0$$

- The CCD algorithm is:

$$x_i^{(k+1)} = \frac{-v_i^{(k+1)} + \sqrt{\left(v_i^{(k+1)}\right)^2 + 4\lambda\sigma_i^2}}{2\sigma_i^2}$$

where:

$$v_i^{(k+1)} = \sigma_i \sum_{j<i} x_j^{(k+1)} \rho_{i,j} \sigma_j + \sigma_i \sum_{j>i} x_j^{(k)} \rho_{i,j} \sigma_j$$

ERC portfolio optimization

The ADMM solution

- In the case of the ADMM algorithm, we set:

$$\begin{aligned} f_x(x) &= \frac{1}{2} x^\top \Sigma x \\ f_y(y) &= -\lambda \sum_{i=1}^{n} \ln y_i \\ x &= y \end{aligned}$$

- The x-update step is:

$$x^{(k+1)} = (\Sigma + \varphi I_n)^{-1} \varphi \left(y^{(k)} - u^{(k)} \right)$$

- The y-update step is:

$$y_i^{(k+1)} = \frac{1}{2} \left(\left(x_i^{(k+1)} + u_i^{(k)} \right) + \sqrt{\left(x_i^{(k+1)} + u_i^{(k)} \right)^2 + 4\lambda\varphi^{-1}} \right)$$

RB portfolio optimization

The RB portfolio is equal to:

$$x_{\mathrm{rb}} = \frac{x^{\star}}{\mathbf{1}_n^{\top} x^{\star}}$$

where $x^{\star}$ is the solution of the logarithmic barrier problem:

$$x^{\star} = \arg\min_{x} \mathcal{R}(x) - \lambda \sum_{i=1}^{n} \mathcal{RB}_i \cdot \ln x_i$$

λ is any positive scalar and $\mathcal{RB}_i$ is the risk budget allocated to Asset i

RB portfolio optimization

The CCD solution (SD risk measure)

- In the case of the standard deviation-based risk measure:

$$\mathcal{R}(x) = -x^\top (\mu - r) + \xi \sqrt{x^\top \Sigma x}$$

the first-order condition for defining the CCD algorithm is:

$$-(\mu_i - r) + \xi \frac{(\Sigma x)_i}{\sqrt{x^\top \Sigma x}} - \lambda \frac{\mathcal{RB}_i}{x_i} = 0$$

- It follows that $\xi x_i (\Sigma x)_i - (\mu_i - r) x_i \sigma(x) - \lambda \sigma(x) \cdot \mathcal{RB}_i = 0$ or equivalently:

$$\alpha_i x_i^2 + \beta_i x_i + \gamma_i = 0$$

where $\alpha_i = \xi \sigma_i^2$, $\beta_i = \xi \sigma_i \sum_{j \neq i} x_j \rho_{i,j} \sigma_j - (\mu_i - r) \sigma(x)$ and $\gamma_i = -\lambda \sigma(x) \cdot \mathcal{RB}_i$

RB portfolio optimization

The CCD solution (SD risk measure)

- The CCD algorithm is:

$$x_i^{(k+1)} = \frac{-\beta_i^{(k+1)} + \sqrt{\left(\beta_i^{(k+1)}\right)^2 - 4\alpha_i^{(k+1)}\gamma_i^{(k+1)}}}{2\alpha_i^{(k+1)}}$$

where:

$$\left\{\begin{array}{l}
\alpha_i^{(k+1)} = \xi\sigma_i^2 \\
\beta_i^{(k+1)} = \xi\sigma_i\left(\sum_{j<i} x_j^{(k+1)}\rho_{i,j}\sigma_j + \sum_{j>i} x_j^{(k)}\rho_{i,j}\sigma_j\right) - (\mu_i - r)\,\sigma_i^{(k+1)}(x) \\
\gamma_i^{(k+1)} = -\lambda\sigma_i^{(k+1)}(x)\cdot\mathcal{RB}_i \\
\sigma_i^{(k+1)}(x) = \sqrt{\chi^\top\Sigma\chi} \\
\chi = \left(x_1^{(k+1)},\ldots,x_{i-1}^{(k+1)},x_i^{(k)},x_{i+1}^{(k)}\ldots,x_n^{(k)}\right)
\end{array}\right.$$

RB portfolio optimization

The ADMM solution (convex risk measure)

- We have:

$$\begin{aligned} \{x^\star, y^\star\} &= \arg\min_{x,y} \mathcal{R}(x) - \lambda \sum_{i=1}^{n} \mathcal{RB}_i \cdot \ln y_i \\ \text{s.t.} &\quad x = y \end{aligned}$$

- The ADMM algorithm is:

$$\left\{ \begin{array}{l} x^{(k+1)} = \mathbf{prox}_{\varphi^{-1}\mathcal{R}(x)} \left(y^{(k)} - u^{(k)} \right) \\ v_y^{(k+1)} = x^{(k+1)} + u^{(k)} \\ y^{(k+1)} = \frac{1}{2} \left(v_y^{(k+1)} + \sqrt{v_y^{(k+1)} \odot v_y^{(k+1)} + 4\lambda\varphi^{-1} \cdot \mathcal{RB}} \right) \\ u^{(k+1)} = u^{(k)} + x^{(k+1)} - y^{(k+1)} \end{array} \right.$$

Tips and tricks of portfolio optimization

- Full allocation — $\sum_{i=1}^{n} x_i = 1$:

$$\Omega = \mathcal{H}_{yperplane} \left[\mathbf{1}_n, 1 \right]$$

We have:

$$\mathcal{P}_{\Omega}\left(v\right) = v - \left(\frac{\mathbf{1}_n^\top v - 1}{n} \right) \mathbf{1}_n$$

- Cash neutral — $\sum_{i=1}^{n} x_i = 0$:

$$\Omega = \mathcal{H}_{yperplane} \left[\mathbf{1}_n, 0 \right]$$

We have:

$$\mathcal{P}_{\Omega}\left(v\right) = v - \left(\frac{\mathbf{1}_n^\top v}{n} \right) \mathbf{1}_n$$

Tips and tricks of portfolio optimization

- No short selling — $x \geq \mathbf{0}_n$:

$$\Omega = \mathcal{B}_{ox}\left[\mathbf{0}_n, \infty\right]$$

  We have:

$$\mathcal{P}_\Omega(v) = \mathcal{T}\left(v; \mathbf{0}_n, \infty\right)$$

- Weight bounds — $x^- \leq x \leq x^+$:

$$\Omega = \mathcal{B}_{ox}\left[x^-, x^+\right]$$

  We have:

$$\mathcal{P}_\Omega(v) = \mathcal{T}\left(v; x^-, x^+\right)$$

Tips and tricks of portfolio optimization

- μ-problem — $\mu(x) \geq \mu^{\star}$:

$$\Omega = \mathcal{H}_{alfspace}\left[-\mu, -\mu^{\star}\right]$$

We have:

$$\mathcal{P}_{\Omega}(v) = v + \frac{\left(\mu^{\star} - \mu^{\top} v\right)_{+}}{\left\|\mu\right\|_2^2} \mu$$

Tips and tricks of portfolio optimization

- σ-problem — $\sigma(x) \leq \sigma^\star$:

$$\Omega = \left\{ x : \sqrt{x^\top \Sigma x} \leq \sigma^\star \right\}$$

We have:

$$\begin{aligned} \sqrt{x^\top \Sigma x} \leq \sigma^\star & \Leftrightarrow & \sqrt{x^\top \left(L L^\top\right) x} \leq \sigma^\star \\ & \Leftrightarrow & \left\| y^\top y \right\|_2 \leq \sigma^\star \\ & \Leftrightarrow & y \in \mathcal{B}_2\left(\mathbf{0}_n, \sigma^\star\right) \end{aligned}$$

where $y = L^\top x$ and L is the Cholesky decomposition of Σ. It follows that the proximal of the y-update is the projection onto the ℓ_2 ball $\mathcal{B}_2\left(\mathbf{0}_n, \sigma^\star\right)$:

$$\begin{aligned} \mathcal{P}_\Omega(v) &= v - \mathbf{prox}_{\sigma^\star \|x\|_2}(v) \\ &= v - \left(1 - \frac{\sigma^\star}{\max\left(\sigma^\star, \|v\|_2\right)}\right) v \end{aligned}$$

Tips and tricks of portfolio optimization

- Leverage management — $\sum_{i=1}^{n} |x_i| \leq \mathcal{L}^{+}$:

$$\begin{aligned} \Omega &= \left\{ x : \|x\|_1 \leq \mathcal{L}^{+} \right\} \\ &= \mathcal{B}_1 \left(\mathbf{0}_n, \mathcal{L}^{+} \right) \end{aligned}$$

The proximal of the y-update is the projection onto the ℓ_1 ball $\mathcal{B}_1 \left(\mathbf{0}_n, \mathcal{L}^{+} \right)$:

$$\mathcal{P}_{\Omega} \left(v \right) = v - \operatorname{sign} \left(v \right) \odot \mathbf{prox}_{\mathcal{L}^{+} \max x} \left(|v| \right)$$

Tips and tricks of portfolio optimization

- Leverage management — $\mathcal{LS}^{-} \leq \sum_{i=1}^{n} x_i \leq \mathcal{LS}^{+}$:

$$\Omega = \mathcal{H}_{alfspace}\left[\mathbf{1}_n, \mathcal{LS}^{+}\right] \cap \mathcal{H}_{alfspace}\left[-\mathbf{1}_n, -\mathcal{LS}^{-}\right]$$

  The proximal of the y-update is obtained with the Dykstra's algorithm by combining the two half-space projections.

- Leverage management — $\left|\sum_{i=1}^{n} x_i\right| \leq \mathcal{L}^{+}$:

$$\Omega = \left\{x : \left|\mathbf{1}_n^{\top} x\right| \leq \mathcal{L}^{+}\right\}$$

  This is a special case of the previous result where $\mathcal{LS}^{+} = \mathcal{L}^{+}$ and $\mathcal{LS}^{-} = -\mathcal{L}^{+}$:

$$\Omega = \mathcal{H}_{alfspace}\left[\mathbf{1}_n, \mathcal{L}^{+}\right] \cap \mathcal{H}_{alfspace}\left[-\mathbf{1}_n, \mathcal{L}^{+}\right]$$

Tips and tricks of portfolio optimization

- Concentration management[37]
  Portfolio managers can also use another constraint concerning the sum of the k largest values:

$$f(x) = \sum_{i=n-k+1}^{n} x_{(i:n)} = x_{(n:n)} + \ldots + x_{(n-k+1:n)}$$

where $x_{(i:n)}$ is the order statistics of x: $x_{(1:n)} \leq x_{(2:n)} \leq \cdots \leq x_{(n:n)}$.
Beck (2017) shows that:

$$\mathbf{prox}_{\lambda f(x)}(v) = v - \lambda \mathcal{P}_{\Omega}\left(\frac{v}{\lambda}\right)$$

where:

$$\Omega = \left\{x \in [0,1]^n : \mathbf{1}_n^\top x = k\right\} = \mathcal{B}_{ox}\left[\mathbf{0}_n, \mathbf{1}_n\right] \cap \mathcal{H}_{yperlane}\left[\mathbf{1}_n, k\right]$$

[37] An example is the 5/10/40 UCITS rule: A UCITS fund may invest no more than 10% of its net assets in transferable securities or money market instruments issued by the same body, with a further aggregate limitation of 40% of net assets on exposures of greater than 5% to single issuers.

Tips and tricks of portfolio optimization

- Entropy portfolio management
  Bera and Park (2008) propose using a cross-entropy measure as the objective function:

$$
\begin{aligned}
x^{\star} & = & \arg\min_{x} \mathrm{KL}\left(x \mid \tilde{x}\right) \\
& \text{s.t.} & \left\{\begin{array}{l}
\mathbf{1}_n^\top x = 1 \\
\mathbf{0}_n \leq x \leq \mathbf{1}_n \\
\mu\left(x\right) \geq \mu^{\star}, \sigma\left(x\right) \leq \sigma^{\star}
\end{array}\right.
\end{aligned}
$$

  where $\mathrm{KL}\left(x \mid \tilde{x}\right)$ is the Kullback-Leibler measure:

$$
\mathrm{KL}\left(x \mid \tilde{x}\right) = \sum_{i=1}^{n} x_i \ln\left(x_i / \tilde{x}_i\right)
$$

  and $\tilde{x}$ is a reference portfolio

Tips and tricks of portfolio optimization

- Entropy portfolio management
  We have:

$$\mathbf{prox}_{\lambda\,\mathrm{KL}(v|\tilde{x})}(v) = \lambda \begin{pmatrix} W\left(\lambda^{-1}\tilde{x}_1 e^{\lambda^{-1}v_1 - \tilde{x}_1^{-1}}\right) \\ \vdots \\ W\left(\lambda^{-1}\tilde{x}_n e^{\lambda^{-1}v_n - \tilde{x}_n^{-1}}\right) \end{pmatrix}$$

  where $W(x)$ is the Lambert W function

Tips and tricks of portfolio optimization

Remark

Since the Shannon's entropy is equal to $\mathrm{SE}(x) = -\mathrm{KL}(x \mid \mathbf{1}_n)$, we deduce that:

$$\mathbf{prox}_{\lambda \, \mathrm{SE}(x)}(v) = \lambda \begin{pmatrix} W\left(\lambda^{-1} e^{\lambda^{-1} v_1 - 1}\right) \\ \vdots \\ W\left(\lambda^{-1} e^{\lambda^{-1} v_n - 1}\right) \end{pmatrix}$$

Tips and tricks of portfolio optimization

- Active share constraint — $\mathcal{AS}(x \mid \tilde{x}) \geq \mathcal{AS}^{-}$:

$$\mathcal{AS}(x \mid \tilde{x}) = \frac{1}{2}\sum_{i=1}^{n} |x_i - \tilde{x}_i| \geq \mathcal{AS}^{-}$$

We use the projection onto the complement $\bar{\mathcal{B}}_1(c, r)$ of the ℓ_1 ball and we obtain:

$$\mathcal{P}_{\Omega}(v) = v + \operatorname{sign}(v - \tilde{x}) \odot \frac{\max\left(2\mathcal{AS}^{-} - \left\| v - \tilde{x} \right\|_1, 0\right)}{n}$$

Tips and tricks of portfolio optimization

- Tracking error volatility — $\sigma(x \mid \tilde{x}) \leq \sigma^\star$:

$$\begin{aligned}\sigma(x \mid \tilde{x}) \leq \sigma^\star \quad &\Leftrightarrow \quad \sqrt{(x-\tilde{x})^\top \Sigma (x-\tilde{x})} \leq \sigma^\star \\ &\Leftrightarrow \quad \|y\|_2 \leq \sigma^\star \\ &\Leftrightarrow \quad y \in \mathcal{B}_2(\mathbf{0}_n, \sigma^\star)\end{aligned}$$

where $y = L^\top x - L^\top \tilde{x}$. It follows that $Ax + By = c$ where $A = L^\top$, $B = -I_n$ and $c = L^\top \tilde{x}$. It follows that the proximal of the y-update is the projection onto the ℓ_2 ball $\mathcal{B}_2(\mathbf{0}_n, \sigma^\star)$:

$$\begin{aligned}\mathcal{P}_\Omega(v) &= v - \mathbf{prox}_{\sigma^\star \|x\|_2}(v) \\ &= v - \left(1 - \frac{\sigma^\star}{\max(\sigma^\star, \|v\|_2)}\right) v\end{aligned}$$

Tips and tricks of portfolio optimization

- Bid-ask transaction cost management:

$$\boldsymbol{c}\left(x \mid x_0\right) = \lambda \sum_{i=1}^{n} \left(c_i^- \left(x_{0,i} - x_i\right)_+ + c_i^+ \left(x_i - x_{0,i}\right)_+\right)$$

where c_i^- and c_i^+ are the bid and ask transaction costs. We have:

$$\mathbf{prox}_{\boldsymbol{c}(x|x_0)}\left(v\right) = x_0 + \mathcal{S}\left(v - x_0; \lambda c^-, \lambda c^+\right)$$

where $\mathcal{S}\left(v; \lambda_-, \lambda_+\right) = \left(v - \lambda_+\right)_+ - \left(v + \lambda_-\right)_-$ is the two-sided soft-thresholding operator.

Tips and tricks of portfolio optimization

- Turnover management:

$$\Omega = \left\{x \in \mathbb{R}^n : \left\|x - x_0\right\|_1 \leq \boldsymbol{\tau}^+\right\}$$

The proximal operator is:

$$\mathcal{P}_\Omega\left(v\right) = v - \operatorname{sign}\left(v - x_0\right) \odot \min\left(\left|v - x_0\right|, s^\star\right)$$

where $s^\star = \left\{s \in \mathbb{R} : \sum_{i=1}^n \left(\left|v_i - x_{0,i}\right| - s\right)_+ = \boldsymbol{\tau}^+\right\}$.

Pattern learning and self-automated strategies

Table 110: What works / What doesn't

	Bond Scoring	Stock Picking	Trend Filtering	Mean Reverting	Index Tracking	HF Tracking	Stock Classification	Technical Analysis
Lasso		☺	☺	☺	☹	☺		
NMF							☺	☹
Boosting		☺				☺		
Bagging		☺				☺		
Random forests	☺			☹				☹
Neural nets	☺					☹		
SVM	☺	☹	☹				☹	
Sparse Kalman					☹	☺		
K-NN	☹							
K-means	☺						☺	
Testing protocols[38]	☺	☺	☺	☺		☺		

Source: Roncalli (2014), Big Data in Asset Management, ESMA/CEMA/GEA meeting, Madrid.

[38] Cross-validation, training/test/probe sets, K-fold, etc.

Pattern learning and self-automated strategies

$$2021 \neq 2014$$

The evolution of machine learning in finance is fast, very fast!

Pattern learning and self-automated strategies

Some examples

- Natural Language Processing (NLP)
- Deep learning (DL)
- Reinforcement learning (RL)
- Gaussian process (GP) and Bayesian optimization (BO)
- Learning to rank (MLR)
- Etc.

Some applications

- Robo-advisory
- Stock classification
- $Q_1 - Q_5$ long/short strategy
- Trend-following strategies
- Mean-reverting strategies
- Scoring models
- Sentiment and news analysis
- Etc.

Market generators

- The underlying idea is to simulate artificial multi-dimensional financial time series, whose statistical properties are the same as those observed in the financial markets

  $\approx$ **Monte Carlo simulation of the financial market**

- 3 main approaches:
  1. Restricted Boltzmann machines (RBM)
  2. Generative adversarial networks (GAN)
  3. Convolutional Wasserstein models (W-GAN)
- The goal is to:
  - improve the the risk management of quantitative investment strategies
  - avoid the over-fitting bias of backtesting

The current research shows that results are disappointed until now

Portfolio optimization with CCD and ADMM algorithms

Question 1

We consider the following optimization program:

$$x^{\star} = \arg\min \frac{1}{2} x^{\top} \Sigma x - \lambda \sum_{i=1}^{n} b_i \ln x_i$$

where Σ is the covariance matrix, b is a vector of positive budgets and x is the vector of portfolio weights.

Portfolio optimization with CCD and ADMM algorithms

Question 1.a

Write the first-order condition with respect to the coordinate x_i and show that the solution $x^\star$ corresponds to a risk-budgeting portfolio.

Portfolio optimization with CCD and ADMM algorithms

We have:

$$\mathcal{L}(x;\lambda) = \arg\min \frac{1}{2} x^\top \Sigma x - \lambda \sum_{i=1}^{n} b_i \ln x_i$$

The first-order condition is:

$$\frac{\partial\, \mathcal{L}(x;\lambda)}{\partial\, x_i} = (\Sigma x)_i - \lambda \frac{b_i}{x_i} = 0$$

or:

$$x_i \cdot (\Sigma x)_i = \lambda b_i$$

Portfolio optimization with CCD and ADMM algorithms

If we assume that the risk measure is the portfolio volatility:

$$\mathcal{R}(x) = \sqrt{x^\top \Sigma x}$$

the risk contribution of Asset i is equal to:

$$\mathcal{RC}_i(x) = \frac{x_i \cdot (\Sigma x)_i}{\sqrt{x^\top \Sigma x}}$$

We deduce that the optimization problem defines a risk budgeting portfolio:

$$\frac{x_i \cdot (\Sigma x)_i}{b_i} = \frac{x_j \cdot (\Sigma x)_j}{b_j} = \lambda \Leftrightarrow \frac{\mathcal{RC}_i(x)}{b_i} = \frac{\mathcal{RC}_j(x)}{b_j}$$

where the risk measure is the portfolio volatility and the risk budgets are $(b_1, \ldots, b_n)$.

Portfolio optimization with CCD and ADMM algorithms

Question 1.b

Find the optimal value $x_i^\star$ when we consider the other coordinates $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ as fixed.

Portfolio optimization with CCD and ADMM algorithms

The first-order condition is equivalent to:

$$x_i \cdot (\Sigma x)_i - \lambda b_i = 0$$

We have:

$$(\Sigma x)_i = x_i \sigma_i^2 + \sigma_i \sum_{j \neq i} x_j \rho_{i,j} \sigma_j$$

It follows that:

$$x_i^2 \sigma_i^2 + x_i \sigma_i \sum_{j \neq i} x_j \rho_{i,j} \sigma_j - \lambda b_i = 0$$

Portfolio optimization with CCD and ADMM algorithms

We obtain a second-degree equation:

$$\alpha_i x_i^2 + \beta_i x_i + \gamma_i = 0$$

where:

$$\begin{cases} \alpha_i = \sigma_i^2 \\ \beta_i = \sigma_i \sum_{j \neq i} x_j \rho_{i,j} \sigma_j \\ \gamma_i = -\lambda b_i \end{cases}$$

1. The polynomial function is convex because we have $\alpha_i = \sigma_i^2 > 0$
2. The product of the roots is negative:

$$x_i' x_i'' = \frac{\gamma_i}{\alpha_i} = -\frac{\lambda b_i}{\sigma_i^2} < 0$$

3. The discriminant is positive:

$$\Delta = \beta_i^2 - 4\alpha_i\gamma_i = \left(\sigma_i \sum_{j \neq i} \rho_{i,j} \sigma_j y_j \right)^2 + 4\lambda b_i \sigma_i^2 > 0$$

Portfolio optimization with CCD and ADMM algorithms

We always have two solutions with opposite signs. We deduce that the solution is the positive root of the second-degree equation:

$$\begin{aligned} x_i^\star &= x_i'' = \frac{-\beta_i + \sqrt{\beta_i^2 - 4\alpha_i\gamma_i}}{2\alpha_i} \\ &= \frac{-\sigma_i \sum_{j\neq i} x_j \rho_{i,j}\sigma_j + \sqrt{\sigma_i^2 \left(\sum_{j\neq i} x_j \rho_{i,j}\sigma_j\right)^2 + 4\lambda b_i \sigma_i^2}}{2\sigma_i^2} \end{aligned}$$

Portfolio optimization with CCD and ADMM algorithms

Question 1.c

We note $x_i^{(k)}$ the value of the i^{th} coordinate at the k^{th} iteration. Deduce the corresponding CCD algorithm. How to find the RB portfolio x_{rb}?

Portfolio optimization with CCD and ADMM algorithms

The CCD algorithm consists in iterating the following formula:

$$x_i^{(k)} = \frac{-\beta_i^{(k)} + \sqrt{\left(\beta_i^{(k)}\right)^2 - 4\alpha_i^{(k)}\gamma_i^{(k)}}}{2\alpha_i^{(k)}}$$

where:

$$\left\{\begin{array}{l} \alpha_i^{(k)} = \sigma_i^2 \\ \beta_i^{(k)} = \sigma_i \left(\sum_{j<i} \rho_{i,j}\sigma_j x_j^{(k)} + \sum_{j>i} \rho_{i,j}\sigma_j x_j^{(k-1)}\right) \\ \gamma_i^{(k)} = -\lambda b_i \end{array}\right.$$

The RB portfolio is the scaled solution:

$$x_{\mathrm{rb}} = \frac{x^\star}{\sum_{i=1}^n x_i^\star}$$

Portfolio optimization with CCD and ADMM algorithms

Question 1.d

We consider a universe of three assets, whose volatilities are equal to 20%, 25% and 30%. The correlation matrix is equal to:

$$\rho = \begin{pmatrix} 100\% & & \\ 50\% & 100\% & \\ 60\% & 70\% & 100\% \end{pmatrix}$$

We would like to compute the ERC portfolio[a] using the CCD algorithm. We initialize the CCD algorithm with the following starting values $x^{(0)} = (33.3\%, 33.3\%, 33.3\%)$. We assume that $\lambda = 1$.

[a] This means that:

$$b_i = \frac{1}{3}$$

Portfolio optimization with CCD and ADMM algorithms

Question 1.d.i

Starting from $x^{(0)}$, find the optimal coordinate $x_1^{(1)}$ for the first asset.

Portfolio optimization with CCD and ADMM algorithms

We have:

$$\left\{\begin{array}{l} \alpha_1^{(1)} = 0.2^2 = 4\% \\ \beta_1^{(1)} = 0.02033 \\ \gamma_i^{(1)} = -0.333\% \end{array}\right.$$

We obtain:

$$x_1^{(1)} = 2.64375$$

Portfolio optimization with CCD and ADMM algorithms

Question 1.d.ii

Compute then the optimal coordinate $x_2^{(1)}$ for the second asset.

Portfolio optimization with CCD and ADMM algorithms

We have:

$$\left\{\begin{array}{l}\alpha_2^{(1)} = 0.25^2 = 6.25\% \\ \beta_2^{(1)} = 0.08359 \\ \gamma_2^{(1)} = -0.333\%\end{array}\right.$$

We obtain:

$$x_2^{(1)} = 1.73553$$

Portfolio optimization with CCD and ADMM algorithms

Question 1.d.iii

Compute then the optimal coordinate $x_3^{(1)}$ for the third asset.

Portfolio optimization with CCD and ADMM algorithms

We have:

$$\left\{\begin{array}{l}\alpha_3^{(1)} = 0.3^2 = 9\% \\ \beta_3^{(1)} = 0.18629 \\ \gamma_3^{(1)} = -0.333\%\end{array}\right.$$

We obtain:

$$x_3^{(1)} = 1.15019$$

Portfolio optimization with CCD and ADMM algorithms

Question 1.d.iv

Give the CCD coordinates $x_i^{(k)}$ for $k = 1, \ldots, 10$.

Portfolio optimization with CCD and ADMM algorithms

Table 111: CCD coordinates ($k = 1, \ldots, 5$)

k	i	$\alpha_i^{(k)}$	$\beta_i^{(k)}$	$\gamma_i^{(k)}$	$x_i^{(k)}$	CCD coordinates x_1	x_2	x_3
0						0.33333	0.33333	0.33333
1	1	0.04000	0.02033	−0.33333	2.64375	2.64375	0.33333	0.33333
1	2	0.06250	0.08359	−0.33333	1.73553	2.64375	1.73553	0.33333
1	3	0.09000	0.18629	−0.33333	1.15019	2.64375	1.73553	1.15019
2	1	0.04000	0.08480	−0.33333	2.01525	2.01525	1.73553	1.15019
2	2	0.06250	0.11077	−0.33333	1.58744	2.01525	1.58744	1.15019
2	3	0.09000	0.15589	−0.33333	1.24434	2.01525	1.58744	1.24434
3	1	0.04000	0.08448	−0.33333	2.01782	2.01782	1.58744	1.24434
3	2	0.06250	0.11577	−0.33333	1.56202	2.01782	1.56202	1.24434
3	3	0.09000	0.15465	−0.33333	1.24842	2.01782	1.56202	1.24842
4	1	0.04000	0.08399	−0.33333	2.02183	2.02183	1.56202	1.24842
4	2	0.06250	0.11609	−0.33333	1.56044	2.02183	1.56044	1.24842
4	3	0.09000	0.15471	−0.33333	1.24821	2.02183	1.56044	1.24821
5	1	0.04000	0.08395	−0.33333	2.02222	2.02222	1.56044	1.24821
5	2	0.06250	0.11609	−0.33333	1.56044	2.02222	1.56044	1.24821
5	3	0.09000	0.15472	−0.33333	1.24817	2.02222	1.56044	1.24817

Portfolio optimization with CCD and ADMM algorithms

Table 112: CCD coordinates ($k = 6, \ldots, 10$)

k	i	$\alpha_i^{(k)}$	$\beta_i^{(k)}$	$\gamma_i^{(k)}$	$x_i^{(k)}$	CCD coordinates x_1	x_2	x_3
0						0.33333	0.33333	0.33333
6	1	0.04000	0.08395	−0.33333	2.02223	2.02223	1.56044	1.24817
6	2	0.06250	0.11608	−0.33333	1.56045	2.02223	1.56045	1.24817
6	3	0.09000	0.15472	−0.33333	1.24816	2.02223	1.56045	1.24816
7	1	0.04000	0.08395	−0.33333	2.02223	2.02223	1.56045	1.24816
7	2	0.06250	0.11608	−0.33333	1.56046	2.02223	1.56046	1.24816
7	3	0.09000	0.15472	−0.33333	1.24816	2.02223	1.56046	1.24816
8	1	0.04000	0.08395	−0.33333	2.02223	2.02223	1.56046	1.24816
8	2	0.06250	0.11608	−0.33333	1.56046	2.02223	1.56046	1.24816
8	3	0.09000	0.15472	−0.33333	1.24816	2.02223	1.56046	1.24816
9	1	0.04000	0.08395	−0.33333	2.02223	2.02223	1.56046	1.24816
9	2	0.06250	0.11608	−0.33333	1.56046	2.02223	1.56046	1.24816
9	3	0.09000	0.15472	−0.33333	1.24816	2.02223	1.56046	1.24816
10	1	0.04000	0.08395	−0.33333	2.02223	2.02223	1.56046	1.24816
10	2	0.06250	0.11608	−0.33333	1.56046	2.02223	1.56046	1.24816
10	3	0.09000	0.15472	−0.33333	1.24816	2.02223	1.56046	1.24816

Portfolio optimization with CCD and ADMM algorithms

Question 1.d.v

Deduce the ERC portfolio.

Portfolio optimization with CCD and ADMM algorithms

The CCD algorithm has converged to the following solution:

$$x^\star = \begin{pmatrix} 2.02223 \\ 1.56046 \\ 1.24816 \end{pmatrix}$$

Since $\sum_{i=1}^{3} x_i^\star = 4.83085$, we deduce that:

$$x_{\mathrm{erc}} = \frac{1}{4.83085} \begin{pmatrix} 2.02223 \\ 1.56046 \\ 1.24816 \end{pmatrix} = \begin{pmatrix} 41.86076\% \\ 32.30189\% \\ 25.83736\% \end{pmatrix}$$

Portfolio optimization with CCD and ADMM algorithms

Question 1.d.vi

Compute the variance of the previous CCD solution. What do you notice? Explain this result.

Portfolio optimization with CCD and ADMM algorithms

We remind that the CCD solution is:

$$x^{\star} = \begin{pmatrix} 2.02223 \\ 1.56046 \\ 1.24816 \end{pmatrix}$$

We have:

$$\sigma^2 \left(x^{\star} \right) = x^{\star\top} \Sigma x^{\star} = 1$$

We notice that:

$$\sigma^2 \left(x^{\star} \right) = \lambda$$

Portfolio optimization with CCD and ADMM algorithms

At the optimum, we remind that:

$$\lambda = \frac{x_i^\star \cdot (\Sigma x^\star)_i}{b_i} = \frac{x_i^\star \cdot (\Sigma x^\star)_i}{n^{-1}}$$

We deduce that:

$$\begin{aligned}
\lambda &= \frac{1}{n} \sum_{i=1}^{n} \frac{x_i^\star \cdot (\Sigma x^\star)_i}{n^{-1}} \\
&= \sum_{i=1}^{n} x_i^\star \cdot (\Sigma x^\star)_i \\
&= x^{\star\top} \Sigma x^\star \\
&= \sigma^2 (x^\star)
\end{aligned}$$

It follows that the portfolio variance of the CCD solution is exactly equal to λ.

Portfolio optimization with CCD and ADMM algorithms

Question 1.d.vii

Verify that the CCD solution converges faster to the ERC portfolio when we assume that $\lambda = x_{\mathrm{erc}}^{\top} \Sigma x_{\mathrm{erc}}$.

Portfolio optimization with CCD and ADMM algorithms

We have:

$$\sigma\left(x_{\mathrm{erc}}\right) = \sqrt{x_{\mathrm{erc}}^{\top}\Sigma x_{\mathrm{erc}}} = 20.70029\%$$

and:

$$\sigma^{2}\left(x_{\mathrm{erc}}\right) = 4.28502\%$$

We obtain the results given in Table 113 when $\lambda = 4.28502\%$. If we compare with those given in Tables 111 and 112, it is obvious that the convergence is faster in the present case.

Portfolio optimization with CCD and ADMM algorithms

Table 113: CCD coordinates ($k = 1, \ldots, 5$)

k	i	$\alpha_i^{(k)}$	$\beta_i^{(k)}$	$\gamma_i^{(k)}$	$x_i^{(k)}$	CCD coordinates		
						x_1	x_2	x_3
0						0.33333	0.33333	0.33333
1	1	0.04000	0.02033	−0.01428	0.39521	0.39521	0.33333	0.33333
1	2	0.06250	0.02738	−0.01428	0.30680	0.39521	0.30680	0.33333
1	3	0.09000	0.03033	−0.01428	0.26403	0.39521	0.30680	0.26403
2	1	0.04000	0.01718	−0.01428	0.42027	0.42027	0.30680	0.26403
2	2	0.06250	0.02437	−0.01428	0.32133	0.42027	0.32133	0.26403
2	3	0.09000	0.03200	−0.01428	0.25847	0.42027	0.32133	0.25847
3	1	0.04000	0.01734	−0.01428	0.41893	0.41893	0.32133	0.25847
3	2	0.06250	0.02404	−0.01428	0.32295	0.41893	0.32295	0.25847
3	3	0.09000	0.03204	−0.01428	0.25835	0.41893	0.32295	0.25835
4	1	0.04000	0.01737	−0.01428	0.41863	0.41863	0.32295	0.25835
4	2	0.06250	0.02403	−0.01428	0.32302	0.41863	0.32302	0.25835
4	3	0.09000	0.03203	−0.01428	0.25837	0.41863	0.32302	0.25837
5	1	0.04000	0.01738	−0.01428	0.41861	0.41861	0.32302	0.25837
5	2	0.06250	0.02403	−0.01428	0.32302	0.41861	0.32302	0.25837
5	3	0.09000	0.03203	−0.01428	0.25837	0.41861	0.32302	0.25837

Portfolio optimization with CCD and ADMM algorithms

Question 2

We recall that the ADMM algorithm is based on the following optimization problem:

$$\begin{aligned} \{x^\star, y^\star\} &= \arg\min f_x(x) + f_y(y) \\ \text{s.t.} \quad & Ax + By = c \end{aligned}$$

Portfolio optimization with CCD and ADMM algorithms

Question 2.a

Describe the ADMM algorithm.

Portfolio optimization with CCD and ADMM algorithms

The ADMM algorithm consists in the following iterations:

$$\left\{\begin{array}{l} x^{(k+1)} = \arg\min_x \left\{ f_x\left(x\right) + \dfrac{\varphi}{2}\left\| Ax + By^{(k)} - c + u^{(k)} \right\|_2^2 \right\} \\ y^{(k+1)} = \arg\min_y \left\{ f_y\left(y\right) + \dfrac{\varphi}{2}\left\| Ax^{(k+1)} + By - c + u^{(k)} \right\|_2^2 \right\} \\ u^{(k+1)} = u^{(k)} + \left(Ax^{(k+1)} + By^{(k+1)} - c\right) \end{array}\right.$$

Portfolio optimization with CCD and ADMM algorithms

Question 2.b

We consider the following optimization problem:

$$
\begin{aligned}
w^{\star}(\gamma) &= \arg\min \frac{1}{2}(w-b)^{\top}\Sigma(w-b)-\gamma(w-b)^{\top}\mu \\
&\text{s.t.} \quad \begin{cases} \mathbf{1}_n^{\top} w = 1 \\ \sum_{i=1}^{n} |w_i - b_i| \leq \boldsymbol{\tau}^{+} \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \end{cases}
\end{aligned}
$$

Portfolio optimization with CCD and ADMM algorithms

Question 2.b.i

Give the meaning of the symbols w, b, Σ, and μ. What is the goal of this optimization program? What is the meaning of the constraint $\sum_{i=1}^{n} |w_i - b_i| \leq \boldsymbol{\tau}^{+}$?

Portfolio optimization with CCD and ADMM algorithms

- w is the vector of portfolio weights:

$$w = (w_1, \ldots, w_n)$$

- b is the vector of benchmark weights:

$$b = (b_1, \ldots, b_n)$$

- Σ is the covariance matrix of asset returns
- μ is the vector of expected returns

Portfolio optimization with CCD and ADMM algorithms

The goal of the optimization problem is to tilt a benchmark portfolio by controlling the volatility of the tracking error:

$$\sigma(w \mid b) = \sqrt{(w-b)^\top \Sigma (w-b)}$$

and improving the expected excess return:

$$\mu(w \mid b) = (w-b)^\top \mu$$

This is a typical γ-problem when there is a benchmark

Portfolio optimization with CCD and ADMM algorithms

We remind that the turnover between the benchmark b and the portfolio w is equal to:

$$\boldsymbol{\tau}\left(w \mid b\right) = \sum_{i=1}^{n} \left|w_i - b_i\right|$$

Therefore, we impose that the turnover is less than an upper limit:

$$\boldsymbol{\tau}\left(w \mid b\right) \leq \boldsymbol{\tau}^{+}$$

Portfolio optimization with CCD and ADMM algorithms

Question 2.b.ii

What is the best way to specify $f_x(x)$ and $f_y(y)$ in order to find numerically the solution. Justify your choice.

Portfolio optimization with CCD and ADMM algorithms

The best way to specify $f_x(x)$ and $f_y(y)$ is to split the QP problem and the turnover constraint:

$$\begin{aligned} \{x^\star, y^\star\} &= \arg\min_{x,y} f_x(x) + f_y(y) \\ \text{s.t.} &\quad x - y = \mathbf{0}_n \end{aligned}$$

where:

$$\begin{aligned} f_x(x) &= \frac{1}{2}(x-b)^\top \Sigma (x-b) - \gamma (x-b)^\top \mu + \mathbb{1}_{\Omega_1}(x) + \mathbb{1}_{\Omega_3}(x) \\ f_y(y) &= \mathbb{1}_{\Omega_2}(y) \\ \Omega_1(x) &= \left\{x : \mathbf{1}_n^\top x = 1\right\} \\ \Omega_2(y) &= \left\{y : \sum_{i=1}^n |y_i - b_i| \leq \boldsymbol{\tau}^+\right\} \\ \Omega_3(x) &= \{x : \mathbf{0}_n \leq x \leq \mathbf{1}_n\} \end{aligned}$$

Indeed, the x-update step is a standard QP problem whereas the y-update step is the projection onto the ℓ_1-ball $\mathcal{B}_1(b, \boldsymbol{\tau}^+)$.

Portfolio optimization with CCD and ADMM algorithms

Question 2.b.iii

Give the corresponding ADMM algorithm.

Portfolio optimization with CCD and ADMM algorithms

We have:

$$\begin{aligned}
(*) &= \frac{1}{2}(x-b)^\top \Sigma (x-b) - \gamma (x-b)^\top \mu \\
&= \frac{1}{2}x^\top \Sigma x - x^\top \Sigma b + \frac{1}{2} b^\top \Sigma b - \gamma x^\top \mu + \gamma b^\top \mu \\
&= \frac{1}{2}x^\top \Sigma x - x^\top (\Sigma b + \gamma \mu) + \underbrace{\left(\gamma b^\top \mu + \frac{1}{2} b^\top \Sigma b\right)}_{\text{constant}}
\end{aligned}$$

Portfolio optimization with CCD and ADMM algorithms

If we note $v_x^{(k+1)} = y^{(k)} - u^{(k)}$, we have:

$$\begin{aligned}
\left\| x - y^{(k)} + u^{(k)} \right\|_2^2 &= \left\| x - v_x^{(k+1)} \right\|_2^2 \\
&= \left(x - v_x^{(k+1)} \right)^\top \left(x - v_x^{(k+1)} \right) \\
&= x^\top I_n x - 2x^\top v_x^{(k+1)} + \underbrace{\left(v_x^{(k+1)} \right)^\top v_x^{(k+1)}}_{\text{constant}}
\end{aligned}$$

Portfolio optimization with CCD and ADMM algorithms

It follows that:

$$\begin{aligned} f_x^{(k+1)}(x) &= f_x(x) + \frac{\varphi}{2}\left\| x - y^{(k)} + u^{(k)} \right\|_2^2 \\ &= \frac{1}{2}(x-b)^\top \Sigma (x-b) - \gamma (x-b)^\top \mu + \\ &\quad \mathbb{1}_{\Omega_1}(x) + \mathbb{1}_{\Omega_3}(x) + \frac{\varphi}{2}\left\| x - y^{(k)} + u^{(k)} \right\|_2^2 \\ &= \frac{1}{2}x^\top (\Sigma + \varphi I_n) x - x^\top \left(\Sigma b + \gamma \mu + \varphi v_x^{(k+1)} \right) + \\ &\quad \mathbb{1}_{\Omega_1}(x) + \mathbb{1}_{\Omega_3}(x) + \text{constant} \end{aligned}$$

Portfolio optimization with CCD and ADMM algorithms

We have:

$$\begin{aligned} f_y^{(k+1)}(y) &= \mathbb{1}_{\Omega_2}(y) + \frac{\varphi}{2}\left\| x^{(k+1)} - y + u^{(k)} \right\|_2^2 \\ &= \mathbb{1}_{\Omega_2}(y) + \frac{\varphi}{2}\left\| y - v_y^{(k+1)} \right\|_2^2 \end{aligned}$$

where $v_y^{(k+1)} = x^{(k+1)} + u^{(k)}$. We deduce that:

$$\begin{aligned} y^{(k+1)} &= \arg\min_y f_y^{(k+1)}(y) \\ &= \mathcal{P}_{\Omega_2}\left(v_y^{(k+1)}\right) \end{aligned}$$

where:

$$\Omega_2 = \mathcal{B}_1\left(b, \tau^+\right)$$

Portfolio optimization with CCD and ADMM algorithms

We remind that:

$$\begin{aligned}
\mathcal{P}_{\mathcal{B}_1(c,\lambda)}(v) &= \mathcal{P}_{\mathcal{B}_1(\mathbf{0}_n,\lambda)}(v-c)+c \\
\mathcal{P}_{\mathcal{B}_1(\mathbf{0}_n,\lambda)}(v) &= v-\operatorname{sign}(v)\odot\mathbf{prox}_{\lambda\max x}(|v|) \\
\mathbf{prox}_{\lambda\max x}(v) &= \min(v,s^\star)
\end{aligned}$$

where $s^\star$ is the solution of the following equation:

$$s^\star=\left\{s\in\mathbb{R}:\sum_{i=1}^{n}(v_i-s)_+=\lambda\right\}$$

Portfolio optimization with CCD and ADMM algorithms

We deduce that:

$$\begin{aligned}
\mathcal{P}_{\Omega_2}\left(v_y^{(k+1)}\right) &= \mathcal{P}_{\mathcal{B}_1(b,\boldsymbol{\tau}^+)}\left(v_y^{(k+1)}\right) \\
&= \mathcal{P}_{\mathcal{B}_1(\mathbf{0}_n,\boldsymbol{\tau}^+)}\left(v_y^{(k+1)} - b\right) + b \\
&= v_y^{(k+1)} - \operatorname{sign}\left(v_y^{(k+1)} - b\right) \odot \mathbf{prox}_{\boldsymbol{\tau}^+ \max x}\left(\left|v_y^{(k+1)} - b\right|\right) \\
&= v_y^{(k+1)} - \operatorname{sign}\left(v_y^{(k+1)} - b\right) \odot \min\left(\left|v_y^{(k+1)} - b\right|, s^\star\right)
\end{aligned}$$

where $s^\star$ is the solution of the following equation:

$$s^\star = \left\{ s \in \mathbb{R} : \sum_{i=1}^{n} \left(\left|v_{y,i}^{(k+1)} - b_i\right| - s\right)_+ = \boldsymbol{\tau}^+ \right\}$$

Portfolio optimization with CCD and ADMM algorithms

The ADMM algorithm becomes:

$$\left\{\begin{array}{l} v_x^{(k+1)} = y^{(k)} - u^{(k)} \\ Q^{(k+1)} = \Sigma + \varphi I_n \\ R^{(k+1)} = \Sigma b + \gamma \mu + \varphi v_x^{(k+1)} \\ x^{(k+1)} = \arg\min_x \left\{ \frac{1}{2} x^\top Q^{(k+1)} x - x^\top R^{(k+1)} + \mathbb{1}_{\Omega_1}(x) + \mathbb{1}_{\Omega_3}(x) \right\} \\ v_y^{(k+1)} = x^{(k+1)} + u^{(k)} \\ s^\star = \left\{ s \in \mathbb{R} : \sum_{i=1}^n \left(\left| v_{y,i}^{(k+1)} - b_i \right| - s \right)_+ = \tau^+ \right\} \\ y^{(k+1)} = v_y^{(k+1)} - \operatorname{sign}\left(v_y^{(k+1)} - b \right) \odot \min\left(\left| v_y^{(k+1)} - b \right|, s^\star \right) \\ u^{(k+1)} = u^{(k)} + x^{(k+1)} - y^{(k+1)} \end{array}\right.$$

Portfolio optimization with CCD and ADMM algorithms

Question 2.c

We consider the following optimization problem:

$$
\begin{aligned}
w^{\star} &= \arg\min \left\| w - \tilde{w} \right\|_1 \\
\text{s.t.} &\quad \left\{ \begin{array}{l} \mathbf{1}_n^\top w = 1 \\ \sqrt{(w-b)^\top \Sigma (w-b)} \leq \sigma^+ \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \end{array} \right.
\end{aligned}
$$

Portfolio optimization with CCD and ADMM algorithms

Question 2.c.i

What is the meaning of the objective function $\left\|w-\tilde{w}\right\|_1$? What is the meaning of the constraint $\sqrt{\left(w-b\right)^\top \Sigma \left(w-b\right)} \leq \sigma^+$?

Portfolio optimization with CCD and ADMM algorithms

The objective function $\left\| w - \tilde{w} \right\|_1$ is the turnover between a given portfolio $\tilde{w}$ and the optimized portfolio w

The constraint $\sqrt{\left(w - b\right)^\top \Sigma \left(w - b\right)} \leq \sigma^+$ is a tracking error limit with respect to a benchmark b

Portfolio optimization with CCD and ADMM algorithms

Question 2.c.ii

Propose an equivalent optimization problem such that $f_x(x)$ is a QP problem. How to solve the y-update?

Portfolio optimization with CCD and ADMM algorithms

The optimization problem is equivalent to solve the following program:

$$\begin{aligned} w^{\star} &= \arg\min \frac{1}{2}(w-b)^{\top}\Sigma(w-b)+\lambda\left\|w-\tilde{w}\right\|_{1} \\ & \text{s.t.} \quad \left\{ \begin{array}{l} \mathbf{1}_n^{\top}w=1 \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \end{array} \right. \end{aligned}$$

Portfolio optimization with CCD and ADMM algorithms

We deduce that:

$$f_x(x) = \frac{1}{2}(x-b)^\top \Sigma (x-b) + \mathbb{1}_{\Omega_1}(x) + \mathbb{1}_{\Omega_2}(x)$$

where:

$$\Omega_1(x) = \left\{x : \mathbf{1}_n^\top x = 1\right\}$$

and:

$$\Omega_2(x) = \{x : \mathbf{0}_n \leq x \leq \mathbf{1}_n\}$$

Portfolio optimization with CCD and ADMM algorithms

We have:

$$f_y(y) = \lambda \|w - \tilde{w}\|_1$$

We remind that:

$$\mathbf{prox}_{\lambda\|x\|_1}(v) = \mathcal{S}(v;\lambda) = \operatorname{sign}(v) \odot (|v| - \lambda \mathbf{1}_n)_+$$

and:

$$\mathbf{prox}_{f(x+b)}(v) = \mathbf{prox}_f(v+b) - b$$

The y-update step is then equal to:

$$\begin{aligned} y^{(k+1)} &= \mathbf{prox}_{\lambda\|w-\tilde{w}\|_1}\left(x^{(k+1)} + u^{(k)}\right) \\ &= \tilde{w} + \operatorname{sign}\left(x^{(k+1)} + u^{(k)} - \tilde{w}\right) \odot \left(\left|x^{(k+1)} + u^{(k)} - \tilde{w}\right| - \lambda \mathbf{1}_n\right)_+ \end{aligned}$$

because $f_y(y)$ is fully separable[39]

[39] Otherwise the scaling property does not work!

Regularized portfolio optimization

Exercise

We consider an investment universe with 6 assets. We assume that their expected returns are 4%, 6%, 7%, 8%, 10% and 10%,, and their volatilities are 6%, 10%, 11%, 15%, 15% and 20%. The correlation matrix is given by:

$$\rho = \begin{pmatrix} 100\% & & & & & \\ 50\% & 100\% & & & & \\ 20\% & 20\% & 100\% & & & \\ 50\% & 50\% & 80\% & 100\% & & \\ 0\% & -20\% & -50\% & -30\% & 100\% & \\ 0\% & 20\% & 30\% & 0\% & 0\% & 100\% \end{pmatrix}$$

Regularized portfolio optimization

Question 1

We restrict the analysis to long-only portfolios meaning that $\sum_{i=1}^{n} x_i = 1$ and $x_i \geq 0$.

Regularized portfolio optimization

Question 1.a

We consider the Herfindahl index $\mathcal{H}(x) = \sum_{i=1}^{n} x_i^2$. What are the two limit cases of $\mathcal{H}(x)$? What is the interpretation of the statistic $\mathcal{N}(x) = \mathcal{H}^{-1}(x)$?

Regularized portfolio optimization

We consider the following optimization problem:

$$\begin{aligned} x^{\star} &= \arg\min \mathcal{H}(x) \\ &\text{s.t. } \sum_{i=1}^{n} x_i = 1 \end{aligned}$$

We deduce that the Lagrange function is:

$$\begin{aligned} \mathcal{L}(x;\lambda) &= \mathcal{H}(x) - \lambda \left(\sum_{i=1}^{n} x_i = 1 \right) \\ &= x^{\top} x - \lambda \left(\mathbf{1}_n^{\top} x - 1 \right) \end{aligned}$$

Regularized portfolio optimization

The first-order condition is:

$$\frac{\partial \mathcal{L}(x;\lambda)}{\partial x} = x - \lambda \mathbf{1}_n = \mathbf{0}_n$$

Since we have $\mathbf{1}_n^\top x - 1 = 0$, we deduce that:

$$\lambda = \frac{1}{\mathbf{1}_n^\top \mathbf{1}_n} = \frac{1}{n}$$

We conclude that the lower bound is reached for the equally-weighted portfolio:

$$x_{\mathrm{ew}} = \frac{1}{n} \cdot \mathbf{1}_n$$

and we have:

$$\mathcal{H}(x_{\mathrm{ew}}) = \frac{1}{n^2} \cdot \mathbf{1}_n^\top \mathbf{1}_n = \frac{1}{n}$$

Regularized portfolio optimization

Since the weights are positive, we have:

$$\begin{aligned}\mathcal{H}(x) &= \sum_{i=1}^{n} x_i^2 \\ &\leq \left(\sum_{i=1}^{n} x_i\right)^2 \\ &\leq 1\end{aligned}$$

The upper bound is reached when the portfolio is concentrated on one asset:

$$\exists i : x_i = 1$$

Regularized portfolio optimization

We conclude that:

$$\frac{1}{n} \leq \mathcal{H}(x) \leq 1$$

The statistic $\mathcal{N}(x) = \mathcal{H}^{-1}(x)$ is the effective number of assets

Regularized portfolio optimization

Question 1.b

We consider the following optimization problem $(\mathcal{P}_1)$:

$$
\begin{aligned}
x^{\star}(\lambda) &= \arg\min \frac{1}{2} x^\top \Sigma x + \lambda x^\top x \\
&\text{s.t.} \quad \left\{ \begin{array}{l} \sum_{i=1}^{n} x_i = 1 \\ x_i \geq 0 \end{array} \right.
\end{aligned}
$$

What is the link between this constrained optimization program and the weight diversification based on the Herfindahl index?

Regularized portfolio optimization

The optimization problem $(\mathcal{P}_1)$ is equivalent to:

$$\begin{aligned} x^\star\left(\mathcal{H}^+\right) & = & \arg\min \frac{1}{2} x^\top \Sigma x \\ & \text{s.t.} & \left\{ \begin{array}{l} \sum_{i=1}^n x_i = 1 \\ x_i \geq 0 \\ x^\top x \leq \mathcal{H}^+ \end{array} \right. \end{aligned}$$

We obtain a long-only minimum variance portfolio with a diversification constraint based on the Herfindahl index:

$$\mathcal{H}(x) \leq \mathcal{H}^+$$

We have the following correspondance:

$$\mathcal{H}^+ = \mathcal{H}\left(x^\star(\lambda)\right) = x^\star(\lambda)^\top x^\star(\lambda)$$

Given a value of λ, we can then compute the implicit constraint $\mathcal{H}(x) \leq \mathcal{H}^+$.

Regularized portfolio optimization

Question 1.c

Solve Program $(\mathcal{P}_1)$ when λ is equal to respectively 0, 0.001, 0.01, 0.05, 0.10 and 10. Compute the statistic $\mathcal{N}(x)$. Comment on these results.

Regularized portfolio optimization

Table 114: Solution of the optimization problem $(\mathcal{P}_1)$

λ	0.000	0.001	0.010	0.050	0.100	10.000
$x_1^\star(\lambda)$ (in %)	44.60	35.66	23.97	18.71	17.76	16.68
$x_2^\star(\lambda)$ (in %)	9.12	14.60	18.10	17.08	16.89	16.67
$x_3^\star(\lambda)$ (in %)	25.46	26.57	19.96	16.89	16.71	16.67
$x_4^\star(\lambda)$ (in %)	0.00	0.00	7.64	14.46	15.52	16.65
$x_5^\star(\lambda)$ (in %)	20.40	22.11	22.38	19.31	18.21	16.69
$x_6^\star(\lambda)$ (in %)	0.43	1.07	7.94	13.55	14.92	16.65
$\mathcal{H}(x^\star(\lambda))$	0.3137	0.2680	0.1923	0.1693	0.1675	0.1667
$\mathcal{N}(x^\star(\lambda))$	3.19	3.73	5.20	5.91	5.97	6.00

Regularized portfolio optimization

Question 1.d

Using the bisection algorithm, find the optimal value of $\lambda^{\star}$ that satisfies:

$$\mathcal{N}\left(x^{\star}\left(\lambda^{\star}\right)\right) = 4$$

Give the composition of $x^{\star}\left(\lambda^{\star}\right)$. What is the interpretation of $x^{\star}\left(\lambda^{\star}\right)$?

Regularized portfolio optimization

The optimal solution is:

$$\lambda^{\star} = 0.002301$$

The optimal weights (in %) are equal to:

$$x^{\star} = \begin{pmatrix} 31.62\% \\ 17.24\% \\ 26.18\% \\ 0.00\% \\ 22.63\% \\ 2.33\% \end{pmatrix}$$

The effective number of bets $\mathcal{N}\left(x^{\star}\right)$ is equal to 4

Regularized portfolio optimization

Question 2

We consider long/short portfolios and the following optimization problem $(\mathcal{P}_2)$:

$$
\begin{aligned}
x^{\star}(\lambda) &= \arg\min \frac{1}{2} x^{\top} \Sigma x + \lambda \sum_{i=1}^{n} |x_i| \\
\text{s.t.} & \quad \sum_{i=1}^{n} x_i = 1
\end{aligned}
$$

Regularized portfolio optimization

Question 2.a

Solve Program $(\mathcal{P}_2)$ when λ is equal to respectively 0, 0.0001, 0.001, 0.01, 0.05, 0.10 and 10. Comment on these results.

Regularized portfolio optimization

Table 115: Solution of the optimization problem $(\mathcal{P}_2)$

λ	0.000	0.0001	0.001	0.010	0.050	0.100	10.000
$x_1^\star(\lambda)$ (in %)	35.82	37.17	44.50	44.60	44.60	44.60	44.60
$x_2^\star(\lambda)$ (in %)	33.08	30.26	11.48	9.12	9.12	9.12	9.12
$x_3^\star(\lambda)$ (in %)	77.62	71.77	31.28	25.46	25.46	25.46	25.46
$x_4^\star(\lambda)$ (in %)	−53.48	−47.97	−7.16	0.00	0.00	0.00	0.00
$x_5^\star(\lambda)$ (in %)	20.83	20.56	19.90	20.40	20.40	20.40	20.40
$x_6^\star(\lambda)$ (in %)	−13.87	−11.78	0.00	0.43	0.43	0.43	0.43
$\mathcal{L}(x)$ (in %)	234.69	219.50	114.33	100.00	100.00	100.00	100.00

Regularized portfolio optimization

Question 2.b

For each optimized portfolio, calculate the following statistic:

$$\mathcal{L}(x) = \sum_{i=1}^{n} |x_i|$$

What is the interpretation of $\mathcal{L}(x)$? What is the impact of Lasso regularization?

Regularized portfolio optimization

$\mathcal{L}(x) = \sum_{i=1}^{n} |x_i|$ is the leverage ratio. Their values are reported in Table 115.

Regularized portfolio optimization

Question 3

We assume that the investor holds an initial portfolio $x^{(0)}$ defined as follows:

$$x^{(0)} = \begin{pmatrix} 10\% \\ 15\% \\ 20\% \\ 25\% \\ 30\% \\ 0\% \end{pmatrix}$$

We consider the optimization problem $(\mathcal{P}_3)$:

$$\begin{aligned} x^{\star}(\lambda) &= \arg\min \frac{1}{2} x^{\top} \Sigma x + \lambda \sum_{i=1}^{n} \left| x_i - x_i^{(0)} \right| \\ & \text{s.t.} \quad \sum_{i=1}^{n} x_i = 1 \end{aligned}$$

Regularized portfolio optimization

Question 3.a

Solve Program ($\mathcal{P}_3$) when λ is equal respectively to 0, 0.0001, 0.001, 0.0015 and 0.01. Compute the turnover of each optimized portfolio. Comment on these results.

Regularized portfolio optimization

Table 116: Solution of the optimization problem $(\mathcal{P}_3)$

λ	0.000	0.000	0.001	0.002	0.010
$x_1^\star(\lambda)$ (in %)	35.82	35.55	27.90	24.28	10.00
$x_2^\star(\lambda)$ (in %)	33.08	30.61	15.00	15.00	15.00
$x_3^\star(\lambda)$ (in %)	77.62	72.35	33.36	22.86	20.00
$x_4^\star(\lambda)$ (in %)	−53.48	−48.00	−5.20	7.87	25.00
$x_5^\star(\lambda)$ (in %)	20.83	21.51	28.94	30.00	30.00
$x_6^\star(\lambda)$ (in %)	−13.87	−12.02	0.00	0.00	0.00
$\tau\left(x^\star(\lambda) \mid x^{(0)}\right)$ (in %)	203.04	187.02	62.51	34.27	0.00

Regularized portfolio optimization

Question 3.b

Using the bisection algorithm, find the optimal value of $\lambda^{\star}$ such that the two-way turnover is equal to 60%. Give the composition of $x^{\star}(\lambda^{\star})$.

Regularized portfolio optimization

The optimal solution is:

$$\lambda^{\star} = 0.00103$$

The optimal weights (in %) are equal to:

$$x^{\star} = \begin{pmatrix} 27.23\% \\ 15.00\% \\ 32.77\% \\ -4.30\% \\ 29.30\% \\ 0.00\% \end{pmatrix}$$

The turnover $\tau\left(x^{\star} \mid x^{(0)}\right)$ is equal to 60%

Regularized portfolio optimization

Question 3.c

Same question when the two-way turnover is equal to 50%.

Regularized portfolio optimization

The optimal solution is:

$$\lambda^{\star} = 0.00119$$

The optimal weights (in %) are equal to:

$$x^{\star} = \begin{pmatrix} 25.53\% \\ 15.00\% \\ 29.47\% \\ 0.00\% \\ 30.00\% \\ 0.00\% \end{pmatrix}$$

The turnover $\tau \left(x^{\star} \mid x^{(0)} \right)$ is equal to 50%

Regularized portfolio optimization

Question 3.d

What becomes the portfolio $x^{\star}(\lambda)$ when $\lambda \rightarrow \infty$? How do you explain this result?

Regularized portfolio optimization

We notice that:

$$\lim_{\lambda \to \infty} x^\star(\lambda) = x^{(0)}$$

This is normal since we have:

$$\begin{aligned} x^\star(\lambda) &= \arg\min \frac{1}{2} x^\top \Sigma x + \lambda \sum_{i=1}^n \left| x_i - x_i^{(0)} \right| \\ &\text{s.t.} \quad \sum_{i=1}^n x_i = 1 \end{aligned}$$

We deduce that:

$$\begin{aligned} x^\star(\infty) &= \arg\min \sum_{i=1}^n \left| x_i - x_i^{(0)} \right| \\ &\text{s.t.} \quad \sum_{i=1}^n x_i = 1 \end{aligned}$$

The solution is $x^\star(\infty) = x^{(0)}$

Main references

BECK, A. (2017)
First-Order Methods in Optimization, MOS-SIAM Series on Optimization, 25, SIAM.

COQUERET, G., and GUIDA, T. (2020)
Machine Learning for Factor Investing, Chapman and Hall/CRC Financial Mathematics Series.

PERRIN, S., and RONCALLI, T. (2020)
Machine Learning Algorithms and Portfolio Optimization, in Jurczenko, E. (Ed.), *Machine Learning in Asset Management: New Developments and Financial Applications*, Wiley, pp. 261-328, arxiv.org/abs/1909.10233.

References I

Bourgeron, T., Lezmi, E., and Roncalli, T. (2018)
Robust Asset Allocation for Robo-Advisors, *arXiv*, arxiv.org/abs/1902.07449.

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2010)
Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, *Foundations and Trends® in Machine learning*, 3(1), pp. 1-122.

Gabay, D., and Mercier, B. (1976)
A Dual Algorithm for the Solution of Nonlinear Variational Problems via Finite Element Approximation, *Computers & Mathematics with Applications*, 2(1), pp. 17-40.

References II

- Gonzalvez, J., Lezmi, E., Roncalli, T., and Xu, J. (2019)
  Financial Applications of Gaussian Processes and Bayesian Optimization, *arXiv*, arxiv.org/abs/1903.04841.
- Griveau-Billion, T., Richard, J-C., and Roncalli, T. (2013)
  A Fast Algorithm for Computing High-dimensional Risk Parity Portfolios, *SSRN*, www.ssrn.com/abstract=2325255.
- Jurczenko, E. (2020)
  Machine Learning in Asset Management: New Developments and Financial Applications, Wiley.
- Kondratyev, A., and Schwarz, C. (2020)
  The Market Generator, *SSRN*, www.ssrn.com/abstract=3384948.

References III

Lezmi, E., Roche, J., Roncalli, T., and Xu, J. (2020)
Improving the Robustness of Trading Strategy Backtesting with Boltzmann Machines and Generative Adversarial Networks, *arXiv*, https://arxiv.org/abs/2007.04838.

Parikh, N., and Boyd, S. (2014)
Proximal Algorithms, *Foundations and Trends® in Optimization*, 1(3), pp. 127-239.

Tibshirani, R. (1996)
Regression Shrinkage and Selection via the Lasso, *Journal of the Royal Statistical Society B*, 58(1), pp. 267-288.

References IV

Tibshirani, R.J. (2017)
Dykstra's Algorithm, ADMM, and Coordinate Descent: Connections, Insights, and Extensions, in Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R. (Eds), *Advances in Neural Information Processing Systems*, 30, pp. 517-528.

Tseng, P. (2001)
Convergence of a Block Coordinate Descent Method for Nondifferentiable Minimization, *Journal of Optimization Theory and Applications*, 109(3), pp. 475-494.