Course 2024–2025 in Sustainable Finance
Lecture 2. Exercise — Probability Distribution of an ESG Score

Thierry Roncalli*

*Amundi Asset Management[1]

*University of Paris-Saclay

November 2024

[1] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Lecture 1: Introduction
- **Lecture 2: ESG Scoring**
- Lecture 3: Impact of ESG Investing on Asset Prices and Portfolio Returns
- Lecture 4: Sustainable Financial Products
- Lecture 5: Impact Investing
- Lecture 6: Biodiversity
- Lecture 7: Engagement & Voting Policy
- Lecture 8: Extra-financial Accounting
- Lecture 9: Awareness of Climate Change Impacts
- Lecture 10: The Ecosystem of Climate Change
- Lecture 11: Economic Models & Climate Change
- Lecture 12: Climate Risk Measures
- Lecture 13: Transition Risk Modeling
- Lecture 14: Climate Portfolio Construction
- Lecture 15: Physical Risk Modeling
- Lecture 16: Climate Stress Testing & Risk Management

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

Question (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.

We have:

$$\mathcal{S}_i^{(\mathrm{ESG})} = 0.4 \times \mathcal{S}_i^{(\mathrm{E})} + 0.4 \times \mathcal{S}_i^{(\mathrm{S})} + 0.2 \times \mathcal{S}_i^{(\mathrm{G})}$$

We deduce the following results:

Issuer	#1	#2	#3	#4	#5	#6	#7	#8
$\mathcal{S}_i^{(\mathrm{E})}$	−2.80	−1.80	−1.75	0.60	0.75	1.30	1.90	2.70
$\mathcal{S}_i^{(\mathrm{S})}$	−1.70	−1.90	0.75	−1.60	1.85	1.05	0.90	0.70
$\mathcal{S}_i^{(\mathrm{G})}$	0.30	−0.70	−2.75	2.60	0.45	2.35	2.20	1.70
$\mathcal{S}_i^{(\mathrm{ESG})}$	−1.74	−1.62	−0.95	0.12	1.13	1.41	1.56	1.70

Question (b)

Calculate the ESG score of the equally-weighted portfolio x_{ew}.

We obtain:

$$\mathcal{S}^{(\mathrm{ESG})}(x_{\mathrm{ew}}) = \sum_{i=1}^{8} x_{\mathrm{ew},i} \times \mathcal{S}_i^{(\mathrm{ESG})} = 0.2013$$

Question 2

We assume that the ESG scores are *iid* and follow a standard Gaussian distribution:

$$\mathcal{S}_i \sim \mathcal{N}(0,1)$$

Question (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 $x_{\mathrm{ew}}^{(n)}$.

We have:

$$\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right) = \sum_{i=1}^{n} x_{\mathrm{ew},i}^{(n)} \times \mathcal{S}_i = \frac{1}{n}\sum_{i=1}^{n} \mathcal{S}_i$$

We deduce that $\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right)$ follows a Gaussian distribution. Its mean is equal to:

$$\mathbb{E}\left[\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right)\right] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\mathcal{S}_i\right] = 0$$

whereas its standard deviation is equal to:

$$\sigma\left(\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right)\right) = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n} \sigma^2\left(\mathcal{S}_i\right)} = \frac{1}{\sqrt{n}}$$

Finally, we deduce that:

$$\mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right) \sim \mathcal{N}\left(0, \frac{1}{n}\right)$$

Question (b)

What is the ESG score of a well-diversified portfolio?

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:

$$\lim_{n\to\infty} \mathcal{S}\left(x_{\mathrm{ew}}^{(n)}\right) = 0$$

Question (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?

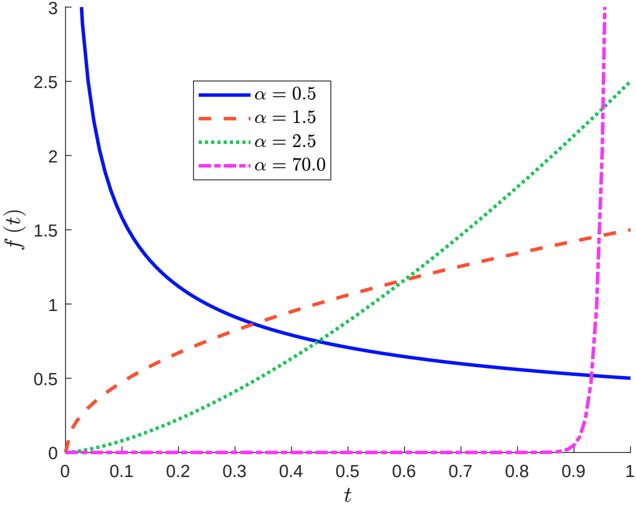
We have:

$$f_{\alpha}(t) = \alpha t^{\alpha - 1}$$

The probability density function $f_{\alpha}(t)$ is reported in Figure 1. We notice that the function $f_{\alpha}(t)$ tends to the dirac delta function when α tends to infinity:

$$\lim_{\alpha \to \infty} f_{\alpha}(t) = \delta_1(t) = \begin{cases} 0 & \text{if } t \neq 1 \\ +\infty & \text{if } t = 1 \end{cases}$$

Figure 1: Probability density function $f_\alpha(t)$



Question (d)

We assume that the weights of the portfolio[a] $x = (x_1, \ldots, x_n)$ follow a power-law distribution $\mathbf{F}_\alpha$:

$$x_i \sim cT_i$$

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.

[a] We use $n = 50$ in the rest of the exercise.

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:

$$T_i = \mathbf{F}_\alpha^{-1}(U_i) = U_i^{1/\alpha}$$

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:

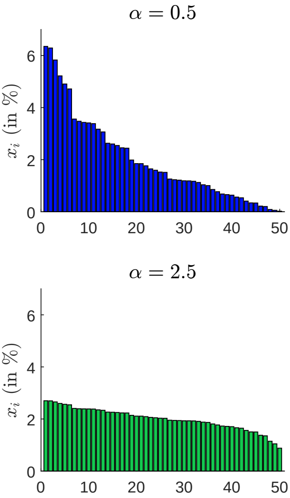
$$c = \left(\sum_{i=1}^n t_i\right)^{-1} = \left(\sum_{i=1}^n u_i^{1/\alpha}\right)^{-1}$$

4. We deduce the portfolio weights $x = (x_1, \ldots, x_n)$:

$$x_i = c \cdot t_i = c \cdot u_i^{1/\alpha} = \frac{u_i^{1/\alpha}}{\sum_{j=1}^n u_j^{1/\alpha}}$$

In Figure 2, 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.

Figure 2: Repartition of the portfolio weights in descending order



Question (e)

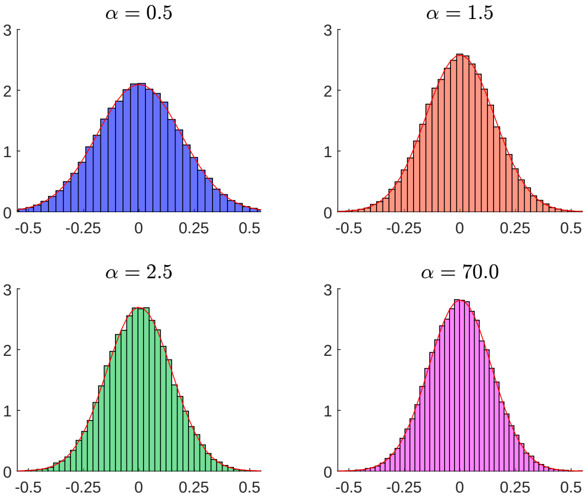
We assume that the weight x_i and the ESG score $\mathcal{S}_i$ of the issuer i are independent. How to simulate the portfolio's score $\mathcal{S}(x)$? Using 50 000 replications, estimate the probability distribution function of $\mathcal{S}(x)$ by the Monte Carlo method. Comment on these results.

We simulate $x = (x_1, \ldots, x_n)$ using the previous algorithm. The vector of ESG scores $\boldsymbol{\mathcal{S}} = (\boldsymbol{\mathcal{S}}_1, \ldots, \boldsymbol{\mathcal{S}}_n)$ is generated with normally-distributed random variables since we have $\boldsymbol{\mathcal{S}}_i \sim \mathcal{N}(0, 1)$. We deduce that the simulated value of the portfolio ESG score $\boldsymbol{\mathcal{S}}(x)$ is equal to:

$$\boldsymbol{\mathcal{S}}(x) = \sum_{i=1}^{n} x_i \cdot \boldsymbol{\mathcal{S}}_i$$

We replicate the simulation of $\boldsymbol{\mathcal{S}}(x)$ 50 000 times and draw the corresponding histogram in Figure 3. We also report the fitted Gaussian distribution. We observe that the portfolio's ESG score $\boldsymbol{\mathcal{S}}(x)$ is equal to zero on average, and its variance is an increasing function of the portfolio concentration.

Figure 3: Histogram of the portfolio ESG score $\mathcal{S}(x)$



Question (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 $\mathcal{S}(x)$.

Since $x_i \sim cT_i$, x_i is an increasing function of T_i. We deduce that the copula function of $(T_i, \boldsymbol{\mathcal{S}}_i)$ is the same as the copula function of $(x_i, \boldsymbol{\mathcal{S}}_i)$. To simulate the Normal copula function $\mathbf{C}(u, v)$, we use the transformation algorithm based on the Cholesky decomposition:

$$\begin{cases} u_i = \Phi(g_i') \\ v_i = \Phi\left(\rho g_i' + \sqrt{1-\rho^2} g_i''\right) \end{cases}$$

where g_i' and g_i'' are two independent random numbers from the probability distribution $\mathcal{N}(0,1)$.

Here is the algorithm to simulate the portfolios's ESG score $\mathcal{S}(x)$:

1. We simulate n independent normally-distributed random numbers g_i' and g_i'' and 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 $x = (x_1, \ldots, x_n)$:

$$x_i = t_i \left/ \sum_{j=1}^n t_j \right.$$

4. We simulate the vector of scores $\mathcal{S} = (\mathcal{S}_1, \ldots, \mathcal{S}_n)$:

$$\mathcal{S}_i = \Phi^{-1}(v_i) = \rho g_i' + \sqrt{1-\rho^2} g_i''$$

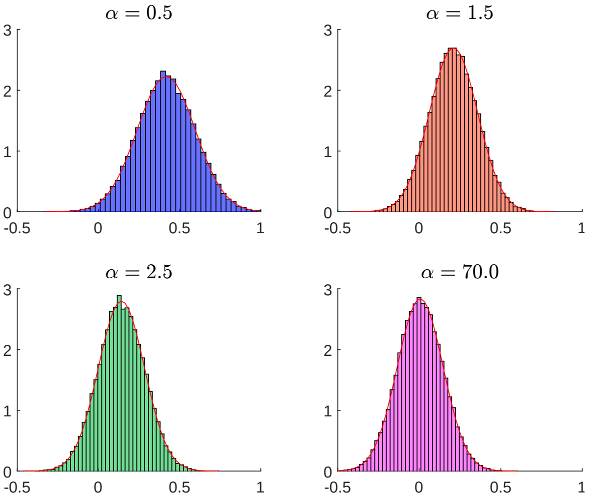
5. We calculate the portfolio score:

$$\mathcal{S}(x) = \sum_{i=1}^n x_i \cdot \mathcal{S}_i$$

Question (g)

Using 50 000 replications, estimate the probability distribution function of $\mathcal{S}(x)$ by the Monte Carlo method when the correlation parameter ρ is set to 50%. Comment on these results.

Figure 4: Histogram of the portfolio ESG score $\mathcal{S}(x)$ ($\rho = 50\%$)



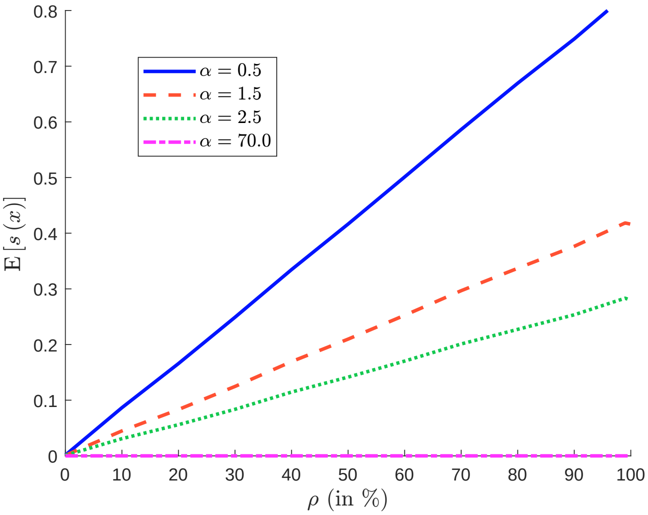
In the independent case, we found that $\mathbb{E}\left[\mathcal{S}\left(x\right)\right] = 0$. In Figure 4, we notice that $\mathbb{E}\left[\mathcal{S}\left(x\right)\right] \neq 0$ when ρ is equal to 50%. Indeed, we obtain:

$$\mathbb{E}\left[\mathcal{S}\left(x\right)\right] = \begin{cases} 0.418 & \text{if } \alpha = 0.5 \\ 0.210 & \text{if } \alpha = 1.5 \\ 0.142 & \text{if } \alpha = 2.5 \\ 0.006 & \text{if } \alpha = 70.0 \end{cases}$$

Question (h)

Estimate the relationship between the correlation parameter ρ and the expected ESG score $\mathbb{E}[\boldsymbol{\mathcal{S}}(x)]$ of the portfolio x. Comment on these results.

Figure 5: Relationship between ρ and $\mathbb{E}[\mathcal{S}(x)]$



In Figure 5, we notice that there is a positive relationship between ρ and $\mathbb{E}\left[\mathcal{S}\left(x\right)\right]$ and the slope increases with the concentration of the portfolio.

Question (i)

How are the previous results related to the size bias of ESG scoring?

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 $\mathbb{E}\left[\boldsymbol{\mathcal{S}}\left(x\right)\right] \geq \mathbb{E}\left[\boldsymbol{\mathcal{S}}\right]$ when the Herfindahl index of the portfolio x is high and the correlation between x_i and $\boldsymbol{\mathcal{S}}_i$ is positive.

Question 3

Let $\boldsymbol{\mathcal{S}}$ be the ESG score of the issuer. We assume that the ESG score follows a standard Gaussian distribution:

$$\boldsymbol{\mathcal{S}} \sim \mathcal{N}(0,1)$$

The ESG score $\boldsymbol{\mathcal{S}}$ is also converted into an ESG rating $\mathcal{R}$, which can take the values[a] **A**, **B**, **C** and **D**.

[a] **A** is the best rating and **D** is the worst rating.

Question (a)

We assume that the breakpoints of the rating system are -1.5, 0 and $+1.5$. Compute the frequencies of the ratings.

We have:

$$\begin{aligned}\Pr\{\mathcal{R}=\mathbf{A}\} &= \Pr\{\boldsymbol{\mathcal{S}}\geq 1.5\} \\ &= 1-\Phi(1.5) \\ &= 6.68\%\end{aligned}$$

and:

$$\begin{aligned}\Pr\{\mathcal{R}=\mathbf{B}\} &= \Pr\{0\leq\boldsymbol{\mathcal{S}}< 1.5\} \\ &= \Phi(1.5)-\Phi(0) \\ &= 43.32\%\end{aligned}$$

Since the Gaussian distribution is symmetric around 0, we also have:

$$\Pr\{\mathcal{R}=\mathbf{C}\}=\Pr\{\mathcal{R}=\mathbf{B}\}=43.32\%$$

and:

$$\Pr\{\mathcal{R}=\mathbf{D}\}=\Pr\{\mathcal{R}=\mathbf{A}\}=6.68\%$$

The mapping function is then equal to:

$$\mathcal{M}_{\text{appring}}(\mathcal{S}) = \begin{cases} \textbf{A} & \text{if } \mathcal{S} < -1.5 \\ \textbf{B} & \text{if } -1.5 \leq \mathcal{S} < 0 \\ \textbf{C} & \text{if } 0 \leq \mathcal{S} < 1.5 \\ \textbf{D} & \text{if } \mathcal{S} \geq 1.5 \end{cases}$$

Question (b)

We would like to build a rating system such that each category has the same frequency. Find the mapping function.

Since 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:

$$\begin{cases} \Pr\{\boldsymbol{\mathcal{S}} < s_1\} = 25\% \\ \Pr\{s_1 \leq \boldsymbol{\mathcal{S}} < s_2\} = 25\% \\ \Pr\{s_2 \leq \boldsymbol{\mathcal{S}} < s_3\} = 25\% \\ \Pr\{\boldsymbol{\mathcal{S}} \geq s_3\} = 25\% \end{cases}$$

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 then given by:

$$\mathcal{M}_{\text{appring}}(\boldsymbol{\mathcal{S}}) = \begin{cases} \mathbf{A} & \text{if } \boldsymbol{\mathcal{S}} < -0.6745 \\ \mathbf{B} & \text{if } -0.6745 \leq \boldsymbol{\mathcal{S}} < 0 \\ \mathbf{C} & \text{if } 0 \leq \boldsymbol{\mathcal{S}} < 0.6745 \\ \mathbf{D} & \text{if } \boldsymbol{\mathcal{S}} \geq 0.6745 \end{cases}$$

Question (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.

We have:

$$\begin{cases} s_1 = \Phi^{-1}(0.10) = -1.2816 \\ s_2 = \Phi^{-1}(0.50) = 0 \\ s_3 = \Phi^{-1}(0.90) = +1.2816 \end{cases}$$

We deduce that the mapping function is equal to:

$$\mathcal{M}_{\text{appring}}(\boldsymbol{\mathcal{S}}) = \begin{cases} \textbf{A} & \text{if } \boldsymbol{\mathcal{S}} < -1.2816 \\ \textbf{B} & \text{if } -1.2816 \leq \boldsymbol{\mathcal{S}} < 0 \\ \textbf{C} & \text{if } 0 \leq \boldsymbol{\mathcal{S}} < 1.2816 \\ \textbf{D} & \text{if } \boldsymbol{\mathcal{S}} \geq 1.2816 \end{cases}$$

Question 4

Let $\mathcal{S}(t)$ 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:

i. The ESG score at time $t-1$ follows a standard Gaussian distribution:

$$\mathcal{S}(t-1) \sim \mathcal{N}(0,1)$$

ii. The variation of the ESG score is Gaussian between two months:

$$\Delta\mathcal{S}(t) = \mathcal{S}(t) - \mathcal{S}(t-1) \sim \mathcal{N}\left(0,\sigma^2\right)$$

iii. The ESG score $\mathcal{S}(t-1)$ and the variation $\Delta\mathcal{S}(t)$ are independent.

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 $s_K = +\infty$.

Question (a)

Compute the bivariate probability distribution of the random vector $(\mathcal{S}(t-1), \Delta\mathcal{S}(t))$.

The joint distribution of $(\mathcal{S}(t-1), \Delta\mathcal{S}(t))$ is:

$$\left(\begin{array}{c}\mathcal{S}(t-1) \\ \Delta\mathcal{S}(t)\end{array}\right) \sim \mathcal{N}\left(\left(\begin{array}{c}0 \\ 0\end{array}\right), \left(\begin{array}{cc}1 & 0 \\ 0 & \sigma^2\end{array}\right)\right)$$

Question (b)

Compute the bivariate distribution of the random vector $(\mathcal{S}(t-1), \mathcal{S}(t))$.

Since we have:

$$\mathcal{S}(t) = \mathcal{S}(t-1) + \Delta\mathcal{S}(t)$$

we deduce that:

$$\begin{pmatrix} \mathcal{S}(t-1) \\ \mathcal{S}(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \mathcal{S}(t-1) \\ \Delta\mathcal{S}(t) \end{pmatrix}$$

We conclude that $(\mathcal{S}(t-1), \mathcal{S}(t))$ is a Gaussian random vector. We have:

$$\operatorname{var}(\mathcal{S}(t)) = 1 + \sigma^2$$

and:

$$\begin{aligned} \operatorname{cov}(\mathcal{S}(t-1), \mathcal{S}(t)) &= \mathbb{E}[\mathcal{S}(t-1) \cdot \mathcal{S}(t)] \\ &= \mathbb{E}\left[\mathcal{S}^2(t-1) + \mathcal{S}(t-1) \cdot \Delta\mathcal{S}(t)\right] \\ &= 1 \end{aligned}$$

It follows that:

$$\begin{pmatrix} \mathcal{S}(t-1) \\ \mathcal{S}(t) \end{pmatrix} \sim \mathcal{N}(\mathbf{0}_2, \Sigma_\sigma)$$

where Σ_σ is the covariance matrix:

$$\Sigma_\sigma = \begin{pmatrix} 1 & 1 \\ 1 & 1+\sigma^2 \end{pmatrix}$$

Question (c)

Compute the probability $p_k = \Pr\{\mathcal{R}(t-1) = \mathcal{R}_k\}$.

We have:

$$\Pr\left\{\mathcal{R}\left(t-1\right)=\mathcal{R}_k\right\}=\Pr\left\{s_{k-1}\leq\boldsymbol{\mathcal{S}}\left(t-1\right)<s_k\right\}=\Phi\left(s_k\right)-\Phi\left(s_{k-1}\right)$$

Question (d)

Compute the joint probability $\Pr\left\{\mathcal{R}\left(t\right)=\mathcal{R}_k,\mathcal{R}\left(t-1\right)=\mathcal{R}_j\right\}$.

We have:

$$\begin{aligned}
(*) &= \Pr\left\{\mathcal{R}(t) = \mathcal{R}_k, \mathcal{R}(t-1) = \mathcal{R}_j\right\} \\
&= \Pr\left\{s_{k-1} \leq \boldsymbol{\mathcal{S}}(t) < s_k, s_{j-1} \leq \boldsymbol{\mathcal{S}}(t-1) < s_j\right\} \\
&= \Phi_2\left(s_j, s_k; \Sigma_\sigma\right) - \Phi_2\left(s_{j-1}, s_k; \Sigma_\sigma\right) - \Phi_2\left(s_j, s_{k-1}; \Sigma_\sigma\right) + \\
&\quad\ \Phi_2\left(s_{j-1}, s_{k-1}; \Sigma_\sigma\right)
\end{aligned}$$

where $\Phi_2\left(x, y; \Sigma_\sigma\right)$ is the bivariate Normal cdf with covariance matrix Σ_σ.

Question (e)

Compute the transition probability
$p_{j,k} = \Pr\left\{\mathcal{R}(t) = \mathcal{R}_k \mid \mathcal{R}(t-1) = \mathcal{R}_j\right\}$.

We have:

$$
\begin{aligned}
p_{j,k} &= \Pr\left\{\mathcal{R}\left(t\right) = \mathcal{R}_k \mid \mathcal{R}\left(t-1\right) = \mathcal{R}_j\right\} \\
&= \frac{\Pr\left\{\mathcal{R}\left(t\right) = \mathcal{R}_k, \mathcal{R}\left(t-1\right) = \mathcal{R}_j\right\}}{\Pr\left\{\mathcal{R}\left(t-1\right) = \mathcal{R}_j\right\}} \\
&= \frac{\Phi_2\left(s_j, s_k; \Sigma_\sigma\right) + \Phi_2\left(s_{j-1}, s_{k-1}; \Sigma_\sigma\right)}{\Phi\left(s_j\right) - \Phi\left(s_{j-1}\right)} - \\
&\quad\; \frac{\Phi_2\left(s_{j-1}, s_k; \Sigma_\sigma\right) + \Phi_2\left(s_j, s_{k-1}; \Sigma_\sigma\right)}{\Phi\left(s_j\right) - \Phi\left(s_{j-1}\right)}
\end{aligned}
$$

Question (f)

Compute the monthly turnover $\mathcal{T}\left(\mathcal{R}_k\right)$ of the ESG rating $\mathcal{R}_k$.

We have:

$$\begin{aligned}
\mathcal{T}(\mathcal{R}_k) &= \Pr\{\mathcal{R}(t) \neq \mathcal{R}_k \mid \mathcal{R}(t-1) = \mathcal{R}_k\} \\
&= 1 - \Pr\{\mathcal{R}(t) = \mathcal{R}_k \mid \mathcal{R}(t-1) = \mathcal{R}_k\} \\
&= 1 - p_{k,k}
\end{aligned}$$

Question (g)

Compute the monthly turnover $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$ of the ESG rating system.

We have:

$$\begin{aligned}\mathcal{T}\left(\mathcal{R}_1, \ldots, \mathcal{R}_K\right) & = \sum_{k=1}^{K} \Pr\left\{\mathcal{R}\left(t-1\right) = \mathcal{R}_k\right\} \cdot \mathcal{T}\left(\mathcal{R}_k\right) \\ & = \sum_{k=1}^{K} \Pr\left\{\mathcal{R}\left(t\right) \neq \mathcal{R}_k, \mathcal{R}\left(t-1\right) = \mathcal{R}_k\right\}\end{aligned}$$

Question (h)

For each rating system given in Questions 3.a, 3.b and 3.c, compute the corresponding 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?

Table 1: ESG migration matrix (Question 3.a)

Rating	s_k	p_k	Transition probability $p_{j,k}$				$\mathcal{T}(\mathcal{R}_k)$
D	-1.50	6.68%	92.96%	7.04%	0.00%	0.00%	7.04%
C	0.00	43.32%	1.31%	95.03%	3.66%	0.00%	4.97%
B	1.50	43.32%	0.00%	3.66%	95.03%	1.31%	4.97%
A		6.68%	0.00%	0.00%	7.04%	92.96%	7.04%
$\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$							5.25%

Table 2: ESG migration matrix (Question 3.b)

Rating	s_k	p_k	Transition probability $p_{j,k}$				$\mathcal{T}(\mathcal{R}_k)$
D	-0.67	25.00%	95.15%	4.85%	0.00%	0.00%	4.85%
C	0.00	25.00%	5.27%	88.38%	6.35%	0.00%	11.62%
B	0.67	25.00%	0.00%	6.35%	88.38%	5.27%	11.62%
A		25.00%	0.00%	0.00%	4.85%	95.15%	4.85%
$\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$							8.23%

Table 3: ESG migration matrix (Question 3.c)

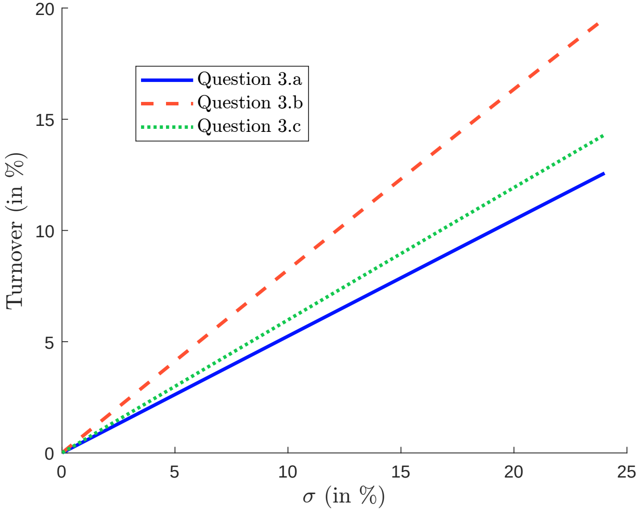
Rating	s_k	p_k	Transition probability $p_{j,k}$				$\mathcal{T}(\mathcal{R}_k)$
D		10.00%	93.54%	6.46%	0.00%	0.00%	6.46%
	-1.28						
C		40.00%	1.89%	94.14%	3.97%	0.00%	5.86%
	0.00						
B		40.00%	0.00%	3.97%	94.14%	1.89%	5.86%
	1.28						
A		10.00%	0.00%	0.00%	6.46%	93.54%	6.46%
$\mathcal{T}(\mathcal{R}_1,\ldots,\mathcal{R}_K)$							5.98%

The ESG migration matrices are given in Tables 1, 2 and 3. We deduce that 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.

Question (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.

Figure 6: Relationship between σ and $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$



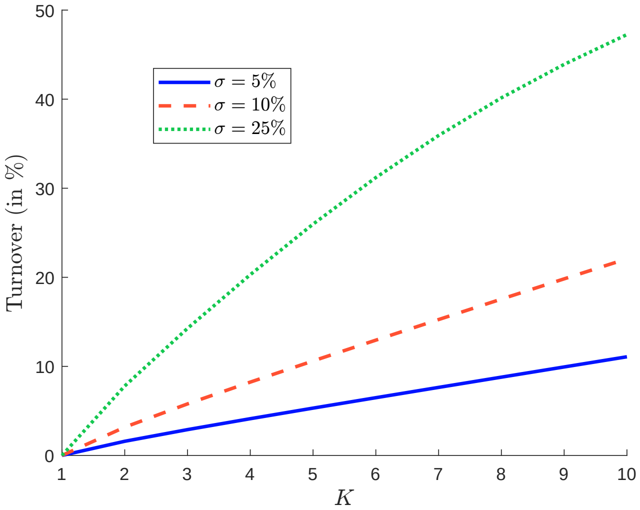
Question (j)

We consider a uniform ESG rating system where:

$$\Pr\left\{\mathcal{R}\left(t-1\right)=\mathcal{R}_k\right\}=\frac{1}{K}$$

Draw the relationship between the number of notches K and the turnover $\mathcal{T}\left(\mathcal{R}_1,\ldots,\mathcal{R}_K\right)$ when the parameter σ takes the values 5%, 10% and 25%.

Figure 7: Relationship between K and $\mathcal{T}(\mathcal{R}_1, \ldots, \mathcal{R}_K)$



Question (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?

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 three factors: (1) the number of notches[2] 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.

[2] This is why ESG rating systems have less notches than credit rating systems.