Optimal business model adaptation plan for a company under a transition scenario

Elisa Ndiaye, Antoine Bezat, Emmanuel Gobet, Céline Guivarch, Ying Jiao

Optimal business model adaptation plan for a company under a transition scenario*

Elisa NDIAYE[†1,2,3], Antoine BEZAT[1], Emmanuel GOBET[2], Céline GUIVARCH[3], and Ying JIAO[4]

[1]STFS, STMM, BNP Paribas, Paris, France
[2]CMAP, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, France
[3]CIRED, CNRS, Ecole des Ponts Paris Tech, Paris, France
[4]ISFA, Laboratoire SAF, Université Claude Bernard - Lyon 1, France

August 30, 2024

Abstract

Tackling climate change is one of the biggest challenges of today. Limiting climate change translates to drastically cutting carbon emissions to net zero as soon as possible. More and more commitments have been made by various authorities and companies to mitigate their GHG emissions accordingly, notably the Paris Agreement in 2015 that sets the *'well-below 2°C'* target. These energy targets generate the so-called *'transition risks'* and has impulsed a new type of financial risks assessment exercise: Climate Stress-Tests. However, the tools for these Stress-Tests remain limited. We propose a model that accounts for companies' business model evolution in a given transition scenario for credit risk stress testing. Our model represents a single firm's business model employing probabilistic modeling. We use stochastic control to derive the company's intensity reduction strategy, as well as the resulting sales revenues and total emissions. We solve the minimization program using a numerical resolution method that we call Backward Sampling. We find that the intensity reduction strategy that would consist in following the same decrease rate as the sector inflates the company's costs (up to 15.7% more expensive than the optimal strategy). Moreover, we show that investing the same amount as the total carbon cost paid at a given date is limited by its lack of a forward-looking feature, making it unable to provide a buffer for future carbon shocks in a disorderly transition scenario.

*This research is supported by *Association Nationale Recherche Technologie (ANRT)* in the context of the first author PhD CIFRE, and it also benefits from the support of *Chair Stress Test, RISK Management and Financial Steering* of the Foundation Ecole Polytechnique.

[†]Corresponding author: `elisa.ndiaye@polytechnique.edu`

1 Introduction

The context of climate change. Tackling climate change is one of the biggest challenges of today. It is caused by human activities such as the use of fossil fuels, livestock breeding, and logging activities [SK14]. Indeed, all these activities emit considerable quantities of greenhouse gases (GHG) in the atmosphere that then intensify the greenhouse effect. Most often referred to as carbon, CO_2, CO_{2_e}, or even Kyoto Gases; the GHG are "*any gas in the atmosphere which absorbs and re-emits heat and thereby keeps the planet's atmosphere warmer than it otherwise would be*" ([Bra12]).
Limiting climate change translates to drastically cutting carbon emissions to net zero as soon as possible. More and more commitments have been made worldwide by various authorities and companies to mitigate their GHG emissions accordingly, notably the Paris Agreement in 2015 that has set the *'well-below 2°C'* target. For some countries, commitments are formalized by the law, detailing the objectives and means of the national climate policy.
Nonetheless, despite the implemented actions and announced objectives, global GHG emissions are still increasing [SSR+22]. Without immediate and drastic cuts in carbon emissions across all sectors, the target of less than 1.5°C increase in mean global temperature by the end of the century cannot be reached. Major transitions are required, particularly the switch from fossil fuels-based energy to low-carbon one, improved energy efficiency, and the use of alternative fuels.

Transition Risks and their challenges for financial risk management. This energy transition generates the so-called 'transition risks' as described by M. Carney (see [Car15]): "the financial risks which could result from the process of adjustment towards a lower-carbon economy". The question of the impact of climate risks on the financial sphere has been well-studied thanks to a close collaboration between financial regulators and academia [GKL+19, ADCG+20, BDN18], with papers dating as far as 2009 [LBR09]. The creation of the Network For Greening the Financial System (NGFS) in 2017 has confirmed the interest of the financial authorities in this matter, notably through its conception of several energy transition scenarios [NfGtFS19, BGL+22]. This has impulsed the launch of a new type of financial risks assessment exercise: *Climate Stress-Tests* [BoE19, AdCPedR20].
Banks and insurers are well familiar with stress tests, which consist of assessing the impact of adverse but plausible short-term macroeconomic scenarios on banks or insurers' portfolios, in particular, their credit portfolios. However, the tools used for these frequent risk management practices remain limited. [HRTK19] explain why traditional stress tests are not appropriate for capturing the impact of a delayed and sudden energy transition, which is the case that generates the most transition risk.
Top-down climate stress tests of the Euro Area banking sector have already been conducted [BMM+17, GHK+21, Bor20] and show that major Euro Area banks could suffer a loss of up

to 27.91% of their assets [BMM+17]. They do not detail the intrinsic dynamics within the macro-system nor the possible adaptation of the bank to the new market structure. However, the thorough study from [BCoBS21] shows that transition risks can be transmitted to both the macro and micro economy through various channels, namely, policy (e.g. carbon tax, subsidies...), technology (emergence of low-carbon technology), and lastly, consumer and investor sentiment. These different drivers should be taken into account; nevertheless, there is no historical data allowing us to compute statistical estimates of the relationships; moreover, they are expected to mutate over time due to the energy transition. A finer granularity is a mean to bypass this issue. Therefore, climate stress tests raise new modeling challenges, whether it be the longer time horizon considered, the dynamicity of the bank's balance sheet, or the change of granularity needed in order to bypass the lack of historical data on all the drivers of transition risks.

State of the art. The question of modeling the impact of transition risks on credit risk has been raised in recent literature. [GGG22] refines the widely accepted asymptotic single risk factor model for credit risk stress testing in the case of climate risks by adding a climate factor. However, it does not inform on how this factor should be accurately calibrated in practice, nor does it allow for changing correlations in the credit portfolio throughout the scenario. Hence, it is not fit for the finer granularity, dynamicity and longer-term horizon scenarios required in climate-risks stress testing.
Early works by [BLG20] propose to assess the sensitivity of a company's credit risk to short-term carbon price increases, relying solely on a linear relationship between the carbon price and the company's revenues, set in a framework often used for credit risk stress-testing, i.e. the Merton model [Mer74]. [BKM22] conduct a similar analysis by assessing the sensitivity of the Euro Area banking system to carbon price shocks, with an arbitrary and one-size-fits-all reduction of firm-level absolute emissions. Despite being adequate for short-term studies of the impact of a given carbon price shock, and providing the right granularity, neither of these methods is appropriate for climate risk stress testing, as the scenarios are usually longer. Moreover, they use static positions for the firm and do not account for correlation effects that are visible in a credit portfolio, thus, despite answering the need for finer granularity, they do not meet the requirements for longer scenarios and dynamicity.
[ADE+21, EFK+23] carry on a full credit risk climate-stress test of different Euro Area banks' portfolios, by making use of a highly granular bottom-up approach. They project individually each company in the portfolio conditionally to a given energy transition scenario, allowing for companies' business model adaptation under a constant market share assumption, then gather results to compute the overall portfolio credit risk. The projections are done using econometric regressions with poor R-squared (as low as 0.11). This confirms the unfitness of statistical models based on historical data for such tasks. [Bar21] also describes a fully deterministic approach for conducting a credit risk climate-stress of a bank corporate loan portfolio, under

a constant market share assumption. Their methodology is sector-dependent and consists of applying the different moving rates of key scenario variables to account for company transition, or applying the company's transition engagement, if available and deemed *feasible* according to their own criteria. Most methods of modeling adaptation in a given transition energy scenario for credit risk assessment are thus deterministic and do not put the cost of fulfilling these targets as a central factor of choice, which is another source of credit risk. Furthermore, they provide the same adaptation strategy for all companies, with their own characteristics only coming into play at the starting point.

Our contribution. Henceforth, we propose a model to improve the representation of companies' adaptation in a given transition scenario for credit stress testing in climate stress tests. Our model is a scenario and sector invariant model for a firm's business model based on probabilistic modelling. There are several combinations of individual behaviors that can lead to the same result when aggregated, thus, probabilistic modelling helps us encompass this randomness. Following microeconomics theory logic, we use optimization to model the firm's adaptation, precisely, we use stochastic control to derive the best intensity reduction strategy according to our definition, as well as the inherent sales revenues and total emissions. Because of the complexity of the stochastic control problem at hand, we solve the minimization program using a specific and novel numerical resolution method that we call *Backward Sampling*: in particular, this permits to recycle past samplings to use them for updated controls. We find that frequently used strategies in corporate credit risk stress testing, consisting in following the same decrease rate as the sector, or investing the same amount as the total carbon cost, are non-optimal in terms of involved costs, which could lead to a misevaluation of credit risk for both high-carbon and low-carbon companies. Precisely, such strategies are not viable in terms of investment requirements, leading to inflated debt stock and carbon costs, regardless of the firm's starting level carbon efficiency. Indeed, applying the same decrease rate as the scenario would require up to thrice as much investment more than the optimal strategy. This number increases up to four when investing the same amount as the carbon cost. Moreover, in a disorderly transition, brown companies need to catch up with the new market requirements, which leads to greater optimal intensity reduction rates than their sector average. In contrast, incentives for carbon mitigation barely reach green companies in orderly transitions, resulting in their relative emissions coming together with those of their pairs. The addition of anticipations regarding the future of the climate policy to the strategy's design enables better managed investment in terms of scheduling and amount. This will in turn lessen credit risk indicators.

Organization of the paper. The paper is organized as follows: Section 1 outlines the state of the art. Section 2 showcases a model for the firm's business model, as well as the minimization problem that we aim to solve. In Section 3, the existence of a solution for our

problem is proven based on the dynamic programming principle. We then solve numerically our stochastic control problem using the *Backward Sampling* algorithm and present some results based on fictitious companies with different levels of transition risk vulnerability in Section 4.

2 Definition of the model

2.1 The Model

We consider the following discrete-time model. We assume a management period defined by $[0, T]$ where T is the end of the scenario. Over this period, the company may choose at N fixed dates $i = 0, 1, \ldots, N$ with a constant time step $\delta := T/N$, whether to invest in its own carbon emissions mitigation plan. Between i and $i+1$, the company's cumulated emissions in $\mathtt{tCO_2}$ and sales revenues in $\mathtt{USD}$ are denoted by E_i and S_i respectively. The indicator for the current state of a company's business model is the sales revenues emissions intensity expressed in $\mathtt{tCO_2/USD}$. In practice, it is obtained by dividing the emissions by the sales revenues, i.e.

$$I_i = \frac{E_i}{S_i}. \tag{2.1}$$

Here, we model it as the next stochastic process:

$$I_{i+1} = I_i e^{-\gamma_i \delta} \times e^{\sigma_I \Delta\varepsilon_i^I - \psi_i^I(\sigma_I)}, \quad I_0 > 0 \tag{2.2}$$

with $\gamma_i \in \mathbb{R}^+$ the intensity reduction effort rate at date i and σ_I is a positive constant. Here $\Delta\varepsilon_i^I$ should be viewed as the time-increment of a random process $\Delta\varepsilon_i^I$ which models the uncertainty in the evolution of the emission intensity: writing it as a time-increment allows to get consistency in the modeling across different time steps. The uncertainty modeling can be made by choosing ε^I as a Brownian motion or as a compound Poisson process for instance. We assume that $(\Delta\varepsilon_i^I)_i$ are independent and bounded for technical reasons. The additional factor $\psi_i^I(\sigma_I)$ is the log moment generating function of $\Delta\varepsilon_i^I$ at point σ_I:

$$e^{\psi_i^I(\sigma_I)} = \mathbb{E}\left[e^{\sigma_I \Delta\varepsilon_i^I}\right],$$

which ensures that the uncertainty factor in (2.2) has unit mean. By definition of the sales revenues emissions intensity (2.1), we get the following identity:

$$E_i = I_i S_i.$$

We define the sales revenues evolution as:

$$S_{i+1} = S_i \frac{\bar{S}_{i+1}}{\bar{S}_i} e^{-\kappa\left(I_i - I_i^{\mathrm{ref}}\right)\delta} \times e^{\sigma_S \Delta\varepsilon_i^S - \psi_i^S(\sigma_S)}, \quad S_0 > 0 \tag{2.3}$$

where $\bar{S}$ are the sales revenues of the reference market given by the scenario in $\mathtt{USD}$, and I^{ref} the market reference for the intensity. We assume that the market shares are allocated with

respect to the company's carbon efficiency relatively to its market: this is embodied by the first exponential, where $\kappa \geq 0$ is the relative sensitivity to the company's intensity. As for (2.2), σ_S is a constant real number and the disturbances $(\Delta\varepsilon_i^S)_i$ are independent random variables with log moment generating $\psi_i^S(.)$:

$$e^{\psi_i^S(\sigma_S)} = \mathbb{E}\left[e^{\sigma_S \Delta\varepsilon_i^S}\right].$$

We also assume that $\Delta\varepsilon_i^S$ is bounded for technical reasons. To keep the modelling consistent with the definition of the intensity as the ratio of emissions over sales revenues, we add the following assumption:

$$\mathbb{C}\mathrm{ov}(\Delta\varepsilon_i^I, \Delta\varepsilon_i^S) < 0.$$

At each date i, the company chooses whether to invest ($\gamma_i > 0$) or not ($\gamma_i = 0$) in its own carbon emissions mitigation plan for the next period $[i, i+1]$. It does so by selecting its sales revenues emission intensity reduction strategy denoted by $\pi = (\gamma_i)_{i=0,\dots,N-1}$ where each $\gamma_i \in \mathbb{R}^+$. However, this investment is costly for the company, and the marginal cost of emissions reduction is an increasing function of the amount of reduction [GETBM93, M$^+$13]. We also allow for decreasing costs with time, to model the gradually diminishing costs of low-carbon technology. Thus, the investment costs are a positive, increasing and convex function of the emissions reduction, and decreasing function of time, that we denote by I^C :

$$I^C(i, S, \gamma) = S \times c \times \alpha^{i\delta} \times \frac{\left(1 - e^{-\gamma\delta}\right)^\beta}{\beta}, \tag{2.4}$$

where $c > 0$ is the unit cost in $\mathtt{USD}$, $\alpha \in [0.95, 1]$ the factor of autonomous cost decrease over time, and $\beta \geq 2$ the exponent of the emissions reduction rate $1 - e^{-\gamma\delta}$. Indeed, the investments are used to green the already existing production chain. It aims at answering "how much is needed to produce S_i with a lesser intensity ?". Thus, we compute the total emissions reduction with S being kept constant. We obtain:

$$\Delta E_i = S_i I_i - S_i I_i e^{-\gamma_i\delta} = I_i S_i (1 - e^{-\gamma_i\delta}).$$

Finally, let us introduce the unitary carbon price in $\mathtt{USD/tCO_2}$ at date i, denoted $\mathtt{cp}_i$, which is a deterministic function of time, provided by the scenario. The company must pay the carbon price for each ton of $\mathtt{CO_2}$ emitted, meaning it pays $C_i^c = \mathtt{cp}_i \times E_i$ at each date i.

2.2 Statement of the problem

We define our controlled system as $X := \{(I_i, S_i)\}_{i=0,\dots,N}$: we will write as $X^{(1)} = I$ and $X^{(2)} = S$ for the two components of X. Consider a filtered probability space $(\Omega, (\mathcal{F}_i)_{0\leq i\leq N}, \mathbb{P})$, with $\mathcal{F}_i$ the canonical filtration of X_i, defined as $\mathcal{F}_i = \sigma(X_s, \forall 0 \leq s \leq i)$ for all i.
Similarly to [LH22], we assume that the best investment strategy for the company is such that it will minimize its future carbon cost in the cheapest way over the time horizon $[0, T]$.

Because of the possible lack of necessary technology, either due to technological progress or constrained number of available units, some large investment strategies are not feasible. We thus add an upper bound for γ denoted $\gamma_{\max} > 0$ to embody this constraint.
At the end of the time horizon, we assume the company may no longer reduce its carbon cost and has to endure some uncontrollable discounted carbon cost until the end of times. Since targets set in scenarios are also defined in terms of time, meaning they aim at answering the following question: "When should we attain our emissions goal ?", this assumption is in line with the essence of energy transition scenarios. We illustrate this with some non-negative function of our controlled system denoted by $\mathcal{C}_N$: in this work, we assume a mild condition of local Lipschitzness on $\mathcal{C}_N$, i.e.

$$|\mathcal{C}_N(x) - \mathcal{C}_N(y)| \leq C_{(2.5)}|x-y|(1+|x|+|y|), \qquad \forall (x,y) \in (\mathbb{R}^+)^2 \times (\mathbb{R}^+)^2. \tag{2.5}$$

An example of such a function is given in (5.1). With initial state space $X_0 = x = (I, S)$ (for intensity and sales level), and given a random intensity reduction strategy $\pi = (\gamma_i)_{i=0,\dots,N-1}$, the total expected cost of the strategy π is:

$$\begin{aligned} J_\pi(0,x) &:= \mathbb{E}\left[\sum_{i=0}^{N-1} \frac{\mathtt{cp}_i X_i^{(1)} X_i^{(2)} + I^C(i, X_i^{(2)}, \gamma_i)}{(1+r\delta)^i} + \frac{\mathcal{C}_N(X_N)}{(1+r\delta)^N} \mid X_0 = x\right] \\ &= \mathbb{E}\left[\sum_{i=0}^{N-1} \frac{\mathcal{C}_i(X_i, \gamma_i)}{(1+r\delta)^i} + \frac{\mathcal{C}_N(X_N)}{(1+r\delta)^N} \mid X_0 = x\right] \end{aligned} \tag{2.6}$$

where r is the risk-free rate and $\mathcal{C}_i(I, S, \gamma)$ is the one-stage cost function defined by:

$$\mathcal{C}_i(X, \gamma) = \mathtt{cp}_i X^{(1)} X^{(2)} + I^C(i, X^{(2)}, \gamma) \qquad \text{if} \quad i < N. \tag{2.7}$$

The company selects $\pi = (\gamma_i)_{i=0,\dots,N-1}$ such that:

$$J^*(0,x) := \inf_{\pi \in \Pi} J_\pi(0,x), \tag{2.8}$$

where Π is the set of admissible strategies, meaning $\mathcal{F}$-adapted and such that $\gamma_i \in [0, \gamma_{\max}] \quad \forall i = 0, \dots, N-1$. This is a finite horizon discrete time stochastic control problem, with control γ and controlled system X; the next section discusses the existence of an optimal control for this problem. Note that for simplicity of notations, we have chosen a constant interest rate. However, a deterministic rate curve could be easily used with no impact on the methodology.

3 Writing the dynamic programming algorithm

3.1 The dynamic programming principle

For pedagogical reasons, we describe the backward induction process in an informal framework to build the dynamic programming equation (DPE), also known as Bellman's equation. A

rigorous treatment is undertaken in [BS78, Chapter 8], assumptions and conditions needed are detailed in the next subsection. This algorithm based on DPE enables us to characterize the optimal control by first showing that an optimal strategy exists for all N sub-minimization problems starting in $i = N-1, \cdots, 0$. There are several ways to write the DPE, here our exposition is aimed at presenting the ingredients and representations that are suitable for our subsequent numerical experiments.

The stochastic system evolves from date i to $i+1$ via Equations (2.2) and (2.3). Start with an initial state $X_0 \in (\mathbb{R}^+)^2$. Chronologically, the company will first pick the control γ_0 in $[0, \gamma_{\max}]$ at date 0 while knowing X_0. At date $i = 1$, the value for X_0 is still known, and X_1 is realized on $(\mathbb{R}^+)^2$ according to its probability law conditionally to the value of X_0, and the value of γ_0 chosen by the firm. This process is then repeated until date N with the last chosen control being γ_{N-1}. At date i, the company's goal is to select each γ_i dependent on the history $(X_0, \gamma_0, \ldots, X_{i-1}, \gamma_{i-1}, X_i)$. Let us formalize the definition of a strategy:

Definition 3.1. *A strategy is a sequence $\pi = (\gamma_0, \gamma_1, \ldots, \gamma_{N-1})$ where, each γ_i is a universally measurable stochastic kernel of the history $(X_0, \gamma_0, \ldots, \gamma_{i-1}, X_i) \in (\mathbb{R}^+)^2 \times \mathbb{R}^+ \times \cdots \times \mathbb{R}^+ \times (\mathbb{R}^+)^2$ with values in $\mathbb{R}^+$, such that:*

$$\gamma_i([0, \gamma_{\max}] \mid X_0, \gamma_0, \ldots, \gamma_{i-1}, X_i) = 1,$$

where $\gamma_k, k = 0, \cdots, i-1$ are the past realizations of the control process.

In other words, the conditional probability γ_i has a support in $[0, \gamma_{\max}]$. We may consider subsets of the above strategies. This is the object of the next definitions.

Definition 3.2. *A Markovian strategy is a sequence $\pi = (\gamma_0, \gamma_1, \ldots, \gamma_{N-1})$ such that, for each $i = 0, \ldots, N-1$, γ_i is a universally measurable stochastic kernel on $(\mathbb{R}^+)^2$ satisfying:*

$$\gamma_i([0, \gamma_{\max}] \mid X_i) = 1.$$

Definition 3.3. *A non randomized strategy is a sequence $\pi = (\gamma_0, \gamma_1, \ldots, \gamma_{N-1})$ such that, for each $i = 0, \ldots, N-1$, γ_i is a universally measurable stochastic kernel on $(\mathbb{R}^+)^2 \times \mathbb{R}^+ \times \cdots \times \mathbb{R}^+ \times (\mathbb{R}^+)^2$ that assigns mass one to some $\tilde{\gamma} \in [0, \gamma_{\max}]$ for each each $(X_0, \gamma_0, \ldots, X_i)$.*

Depending on the above type of strategies, the notation γ_i can represent either a conditional probability depending on the entire past, or a Markovian kernel, or a deterministic function respectively; we will use the same notation as it will be clear from the context.

Since the choice of each control at date i impacts the entire trajectory of the state variable after date i, and since we seek to minimize the total expected costs over the entire period (given by (2.6)), we will use backward induction, a.k.a. dynamic programming principle (DPE), to solve this minimization problem: we will gradually learn from $N-1$ to 0 the best trajectory for our problem.

To define the control problem starting from date i instead of date 0 (as in (2.6)), we need a few extra notations. Denote by $X_s(i,x)$ the trajectory of $(X_s)_{i\leq s\leq N}$ with starting value x at date i. Consider a strategy such that $\pi_i = (\gamma_i, \dots, \gamma_{N-1}) \in \Pi_i, i = 0, \dots, N-1$ such as in Definition 3.1 with Π_i the set of $\mathcal{F}^i$-adapted strategies with each $\gamma_j \in [0, \gamma_{\max}]$, $j = i, \dots, N-1$, and $\mathcal{F}^i$ the filtration such that $\mathcal{F}^i_j = \sigma(X_s(i,x), i \leq s \leq j)$. For simplicity of notations, we denote the strategy by π for $i = 0$. We introduce the expected cost function of the strategy evaluated in $i = 0, \dots, N-1$ with starting state $x = (I, S)$:

$$J_{\pi_i}(i,x) = \mathbb{E}\left[\sum_{j=i}^{N-1} \frac{\mathcal{C}_j(X_j(i,x), \gamma_j)}{(1+r\delta)^{j-i}} + \frac{\mathcal{C}_N(X_N(i,x))}{(1+r\delta)^{N-i}}\right]. \tag{3.1}$$

The optimal expected cost function evaluated is then defined by:

$$J^*(i,x) = \inf_{\pi_i \in \Pi_i} J_{\pi_i}(i,x). \tag{3.2}$$

Now our objective is to prove that a non-randomized Markov optimal strategy

$$\gamma_i^* : (\mathbb{R}^+)^2 \longrightarrow [0, \gamma_{\max}],$$
$$\gamma_i^* : X \longmapsto \gamma_i^*(X)$$

exists for this stochastic control problem, then to characterize it. To identify the optimal control, start with a non-randomized Markovian strategy and using the tower property, write

$$\begin{aligned} J_{\pi_i}(i,x) &= \mathcal{C}_i(x, \gamma_i(x)) \\ &\quad + \frac{1}{(1+r\delta)}\mathbb{E}_\gamma\left[\mathbb{E}\left[\sum_{j=i+1}^{N-1} \frac{\mathcal{C}_j(X_j(i,x), \gamma_j(X_j(i,x)))}{(1+r\delta)^{j-i}} + \frac{\mathcal{C}_N(X_N(i,x))}{(1+r\delta)^{N-i}} \mid X_{i+1}\right] \mid X_i = x\right] \\ &= \mathcal{C}_i(x, \gamma_i(x)) + \frac{1}{(1+r\delta)}\mathbb{E}_\gamma\left[J_{\pi_{i+1}}(i+1, X_{i+1}(i,x)) \mid X_i = x\right], \end{aligned}$$

with $\mathbb{E}_\gamma[. \mid X_i]$ the expectation with respect to the transition density of X_{i+1} conditionally to X_i and $\gamma_i = \gamma$. In the above, we take $i = 0, \dots, N$ with the convention $J_{\pi_{N+1}} \equiv 0$ and $\mathcal{C}_N(x, \gamma_N(x)) = \mathcal{C}_N(x)$. This shows that the optimal control γ_i^* should be such as

$$\gamma_i^* : x \mapsto \arg\min_{\gamma \in [0, \gamma_{\max}]} \left\{\mathcal{C}_i(x,\gamma) + \frac{1}{(1+r\delta)}\mathbb{E}_\gamma\left[J^*(i+1, X_{i+1}(i,x)) \mid X_i = x\right]\right\}. \tag{3.3}$$

In addition, $J^*(i+1,x)$ can be represented as an expectation. Recall we work by backward induction, we thus have already computed $J^*(j,x)$ for all $j > i$ and the corresponding optimal strategy π_j^*. Indeed, since each function γ_j^* is known for all $j > i$, one can rewrite J^* as:

$$J^*(i+1,x) = \mathbb{E}\left[\sum_{j=i+1}^{N-1} \frac{\mathcal{C}_j(X_j^*, \gamma_j^*(X_j^*))}{(1+r\delta)^{j-(i+1)}} + \frac{\mathcal{C}_N(X_N^*)}{(1+r\delta)^{N-(i+1)}} \mid X_{i+1} = x\right]. \tag{3.4}$$

Here $(X_j^*)_{j=i+1,\dots,N}$ is the best trajectory for the state variable and $\gamma_j^*(X_j^*)$ is the best control computed along the best trajectory for X.
Note that one could plug the above representation (3.4) into (3.3) to obtain the expression (3.5) below, however it is not as convenient as (3.3)-(3.4) for the numerical resolution:

$$\gamma_i^*(x) = \arg \min_{\gamma \in [0,\gamma_{\max}]} \Bigg\{ \mathcal{C}_i(x,\gamma) + \frac{1}{(1+r\delta)} \mathbb{E}_\gamma \left[\sum_{j=i+1}^{N-1} \frac{\mathcal{C}_j(X_j^*, \gamma_j^*(X_j^*))}{(1+r\delta)^{j-(i+1)}} + \frac{\mathcal{C}_N(X_N^*)}{(1+r\delta)^{N-(i+1)}} \mid X_i = x \right] \Bigg\}. \quad (3.5)$$

All in all, we have characterized a non-randomized Markov policy $\pi^*(x) = (\gamma_i^*(x), \dots, \gamma_{N-1}^*(.)), i = 0, \dots, N-1$ where each $\gamma_i^*(.)$ is a deterministic function given by the dynamic programming equation, i.e. by measurably selecting for each $i = N-1, \cdots, 0$ a function as given by (3.3) or (3.5), with $J^*(N,x) = \mathcal{C}_N(x)$ by convention. In the following subsection, we prove that the above policy exists and is optimal.

3.2 Existence of an optimal control

Recall the stochastic control problem (2.8). Theorem 3.4 states the conditions for which an optimal nonrandomized Markov intensity reduction strategy exists, and that such a strategy is given by the DPE.

Theorem 3.4. *1. If the following assumption is satisfied:*

$$\mathbb{E}\left[\max\left(0, -\mathcal{C}_i(X,\gamma)\right)\right] < \infty, \quad \forall i = 0, \dots, N, \quad \gamma \in [0, \gamma_{\max}], \qquad (F^+)$$

2. and if the infimum in:

$$\inf_{\gamma \in [0,\gamma_{\max}]} \left\{ \mathcal{C}_i(x,\gamma) + \frac{1}{1+r\delta} \mathbb{E}_\gamma \left[J^*(i+1, X_{i+1}) \mid X_i = x \right] \right\}, \quad i = 0, \dots, N-1, \quad x \in (\mathbb{R}^+)^2, \qquad (3.6)$$

is attained for all $x \in (\mathbb{R}^+)^2$ with $J^*(N,x) = \mathcal{C}_N(x)$,
then an optimal nonrandomized Markov strategy $\pi^*(x)$ exists. This strategy is given by the dynamic programming algorithm, i.e. by measurably selecting for each $x \in (\mathbb{R}^+)^2$ a control that achieves the minimum in (3.6).

Proof. This is an application of [BS78, Proposition 8.5]. □

Theorem 3.9 below states that both (F^+) and (3.6) hold. Thus, the results of Theorem 3.4 apply to prove the existence of an optimal intensity reduction strategy for our model. Before enunciating Theorem 3.9, we need to introduce the following three technical lemmas, which proofs are postponed to Section B.

Lemma 3.5. *We have*

$$\sup_{0\le i\le j\le N,\,\omega\in\Omega} \frac{X_j^{(1)}(i,x)}{X_i^{(1)}(i,x)}(\omega) =: K_1 < +\infty, \tag{3.7}$$

$$\sup_{0\le i\le j\le N,\,\omega\in\Omega} \frac{X_j^{(2)}(i,x)}{X_i^{(2)}(i,x)}(\omega) =: K_2 < +\infty. \tag{3.8}$$

Lemma 3.6. *For any $i = 0,\dots,N-1$, $j \ge i$ and $(x,y) \in (\mathbb{R}^+)^2 \times (\mathbb{R}^+)^2$, we have:*

$$|X_j^{(1)}(i,x) - X_j^{(1)}(i,y)| \le K_3|x^{(1)} - y^{(1)}|, \tag{3.9}$$

$$|X_j^{(2)}(i,x) - X_j^{(2)}(i,y)| \le K_4[(|x^{(2)}| + |y^{(2)}|)|x^{(1)} - y^{(1)}| + |x^{(2)} - y^{(2)}|], \tag{3.10}$$

$$|X_j^{(1)}(i,x)X_j^{(2)}(i,x) - X_j^{(1)}(i,y)X_j^{(2)}(i,y)| \le K_5|x-y|(|x|+|y|)(1+|x|+|y|), \tag{3.11}$$

where K_3, K_4 and K_5 are finite non-negative constants.

Lemma 3.7. *For any $i = 0,\dots,N-1$, $J_{\pi_i}(i,x)$ is locally Lipschitz, uniformly in π_i and γ:*

$$|J_{\pi_i}(i,x) - J_{\pi_i}(i,y)| \le K_6|x-y|[1 + (|x|+|y|)(1+|x|+|y|)], \quad (x,y) \in (\mathbb{R}^+)^2 \times (\mathbb{R}^+)^2.$$

Here K_6 is a non-negative constant, uniform in π_i and j.

Lemma 3.8. *For any $i = 0,\dots,N-1$, $J^*(i,x)$ is locally Lipschitz and uniformly in π_i:*

$$|J^*(i,x) - J^*(i,y)| \le K_6|x-y|[1 + (|x|+|y|)(1+|x|+|y|)], \quad (x,y) \in (\mathbb{R}^+)^2 \times (\mathbb{R}^+)^2.$$

Proof. This is a easy consequence of the general inequality

$$\Big|\inf_{\pi_i\in\Pi_i} J_{\pi_i}(i,x) - \inf_{\pi_i\in\Pi_i} J_{\pi_i}(i,y)\Big| \le \sup_{\pi_i\in\Pi_i} \Big|J_{\pi_i}(i,x) - J_{\pi_i}(i,y)\Big|$$

applied to the definition (3.2) of J^* and using Lemma 3.7. □

Theorem 3.9. *The following statements are true:*

1. *The assumption (F^+) is satisfied.*
2. *The infimum in (3.6) is attained for all $i = 0,\dots,N-1$, $x \in (\mathbb{R}^+)^2$ with $J^*(N,x) = \mathcal{C}_N(x)$.*

Proof. 1. Since the one-stage cost function is non negative, (F^+) is trivially satisfied.

2. $\mathcal{C}_i$ is continuous with respect to $\gamma \in [0,\gamma_{\max}]$ for any $i = 0,\dots,N-1$. Let us now prove the continuity of $\mathbb{E}_\gamma\left[J^*(i+1,X_{i+1}) \mid X_i = x\right]$ with respect to γ.

(a) Recall the dynamics of the state variable with control γ at time i:

$$X_{i+1}^{(1)}(i,x,\gamma) = x^{(1)} e^{-\gamma\delta} \times e^{\sigma_I \Delta\varepsilon_i^I - \psi_i^I(\sigma_I)},$$
$$X_{i+1}^{(2)}(i,x,\gamma) = x^{(2)} \frac{\bar{S}_{i+1}}{\bar{S}_i} e^{-\kappa\left(x^{(1)} - I_i^{\text{ref}}\right)\delta} \times e^{\sigma_S \Delta\varepsilon_i^S - \psi_i^S(\sigma_S)}.$$

This clearly shows that $\gamma \longmapsto X_{i+1}(i,x,\gamma)$ is continuous. Here we have written explicitly the dependence in γ in the state variable to highlight its impact.

(b) From Lemma 3.8, J^* is locally Lipschitz, thus $J^*(i+1,x)$ is continuous in $x \in (\mathbb{R}^+)^2$ for any $i = 0,\dots,N-1$. Taking (2a), we can extend this result to the continuity of $J^*(i+1, X_{i+1}(i,x,\gamma))$ with respect to γ.

(c) From Lemma 3.5, we deduce that $X_{i+1}(i,x,\gamma)$ and therefore $J^*(i+1, X_{i+1}(i,x,\gamma))$ are uniformly bounded, as a function γ and ω (for any given x).

(d) From (2c), we can thus apply the dominated convergence theorem to write

$$\begin{aligned}
\lim_{\gamma \longrightarrow \tilde{\gamma}} \mathbb{E}_\gamma \left[J^*(i+1, X_{i+1}) \mid X_i = x \right] &= \lim_{\gamma \longrightarrow \tilde{\gamma}} \mathbb{E}\left[J^*(i+1, X_{i+1}(i,x,\gamma)) \right] \\
&= \mathbb{E}\left[\lim_{\gamma \longrightarrow \tilde{\gamma}} J^*(i+1, X_{i+1}(i,x,\gamma)) \right] \\
&= \mathbb{E}\left[J^*(i+1, X_{i+1}(i,x,\tilde{\gamma})) \right] \\
&= \mathbb{E}_{\tilde{\gamma}} \left[J^*(i+1, X_{i+1}) \mid X_i = x \right].
\end{aligned}$$

Ergo, $\gamma \longmapsto \mathbb{E}_\gamma \left[J^*(i+1, X_{i+1}) \mid X_i = x \right]$ is continuous, and the infimum in (3.6) is attained on the compact set $[0, \gamma_{\max}]$ for all $x \in (\mathbb{R}^+)^2$, $i = 0,\dots,N-1$.

□

Finally, Corollary 3.10 characterizes the optimal intensity reduction strategy for the company.

Corollary 3.10. *The optimal strategy is generated by the dynamic programming algorithm, i.e.* $\pi^*(x) = (\gamma_0^*(x), \dots, \gamma_{N-1}^*(.))$ *an admissible Markov strategy such that, given* $x \in (\mathbb{R}^+)^2$, *we have for all* $i = 0,\dots,N-1$*:*

$$\gamma_i^* : (\mathbb{R}^+)^2 \longrightarrow [0, \gamma_{\max}],$$
$$x \longmapsto \arg \min_{\gamma \in [0,\gamma_{\max}]} \left\{ \mathcal{C}_i(x,\gamma) + \frac{1}{1+r\delta} \mathbb{E}_\gamma \left[\sum_{j=i+1}^{N-1} \frac{\mathcal{C}_j(X_j^*, \gamma_j^*(X_j^*))}{(1+r\delta)^{j-i-1}} + \frac{\mathcal{C}_N(X_N^*)}{(1+r\delta)^{N-i-1}} \mid X_i = x \right] \right\}. \tag{3.12}$$

Proof. We can apply Theorem 3.4 because its assumptions are fulfilled thanks to Theorem 3.9. Therefore, the Markovian representation (3.3) is valid, and the continuation value (3.4) too. Replacing the latter in (3.3) gives the announced formula (3.12). □

We have now proven the existence, of an optimal intensity reduction strategy, as well as specified its form. Given the complexity of our model, we cannot use the dynamic programming equation to derive a closed-form solution for this minimization problem. We therefore propose in the next section a resolution method based on a numerical procedure that we call *backward sampling*.

4 Methodology for the Numerical Resolution

We use the dynamic programming equation stated in Corollary 3.10 to learn the optimal strategy. This amounts to selecting at each $i = N-1, \dots, 0$ a process such that it minimizes Equation (3.12). Working by backward induction, this involves computing the conditional expectation of the optimal future costs at each step i. The backward sampling algorithm that we propose consists of gradually learning the optimal control process and the optimal trajectory for the state variable at each date $i = N-1, \dots, 0$ using simulations of the future state of the controlled variable (i.e. $j > i$). A key feature of our approach is to avoid resampling the paths once the control is learned. This requires to shift from the trajectory computed with a non optimal strategy to the optimal trajectory at each date i. For this, we need to introduce the following notations for all $0 \leq i < j \leq N$:

- Let $\pi_{i:j} := (\gamma_i, \dots, \gamma_{j-1})$ be an admissible strategy between i and $j-1$.
- Let $X_i^{\pi_{0:i}^{(0)}}(0, X_0) := \left(I_i^{\pi_{0:i}^{(0)}}(0, X_0), S_i^{\pi_{0:i}^{(0)}}(0, X_0)\right)$ be the value for the state variable at i starting in X_0 at date 0 and computed with a reference Markovian strategy $\pi_{0:i}^{(0)}$: it could be related to an uncontrolled intensity $\pi_{0:i} = \mathbf{0}_i$, or to any other prescribed intensity (hence non optimal) chosen by the user. We will see later that this reference strategy will be updated a few times to globally improve the efficiency of the learning algorithm. For the sake of notational simplicity, whenever unambiguous, we shall write $X_i^{\pi_{0:i}^{(0)}}$ instead of $X_i^{\pi_{0:i}^{(0)}}(0, X_0)$.
- Let $X_j^{\pi_{i:j}}(i, X_i^{\pi_{0:i}^{(0)}}) := \left(I_j^{\pi_{i:j}}(i, X_i^{\pi_{0:i}^{(0)}}), S_j^{\pi_{i:j}}(i, X_i^{\pi_{0:i}^{(0)}})\right)$ be the value for the state variable at j starting in $X_i^{\pi_{0:i}^{(0)}}$ at date i and computed with strategy $\pi_{i:j}$.

Using the exponential form of the controlled variable,
it is possible to get from the controlled state with reference strategy $\pi_{i:j}^{(0)} = (\gamma_q^{(0)})_{q=i,\dots,j-1}$ to the one with learnt strategy $\widehat{\pi}_{i:j} = (\widehat{\gamma}_q)_{q=i,\dots,j-1}$ at any date $0 \leq i < j \leq N$.Here, we understand that $\gamma_q^{(0)} = \gamma_q^{(0)}(X_q^{\pi_{0:q}^{(0)}})$. Precisely, the first coordinate for $X_j^{\pi_{i:j}^{(0)}}(i, X_i^{\pi_{0:i}^{(0)}})$ and

$X_j^{\widehat{\pi}_{i:j}}(i, X_i^{\pi_{0:i}^{(0)}})$ are respectively given by:

$$I_j^{\pi_{i:j}^{(0)}}(i, X_i^{\pi_{0:i}^{(0)}}) = I_i^{\pi_{0:i}^{(0)}}(0, X_0) \prod_{q=i}^{j-1} \exp\left(-\gamma_q^{(0)}\delta + \sigma_I \Delta\varepsilon_q^I - \psi_q^I(\sigma_I)\right), \tag{4.1}$$

$$I_j^{\widehat{\pi}_{i:j}}(i, X_i^{\pi_{0:i}^{(0)}}) = I_i^{\pi_{0:i}^{(0)}}(0, X_0) \prod_{q=i}^{j-1} \exp\left(-\widehat{\gamma}_q\delta + \sigma_I \Delta\varepsilon_q^I - \psi_q^I(\sigma_I)\right). \tag{4.2}$$

By dividing (4.2) by (4.1), we get:

$$I_j^{\widehat{\pi}_{i:j}}(i, X_i^{\pi_{0:i}^{(0)}}) = I_j^{\pi_{i:j}^{(0)}}(i, X_i^{\pi_{0:i}^{(0)}}) \prod_{q=i}^{j-1} \exp\left(-(\widehat{\gamma}_q - \gamma_q^{(0)})\delta\right).$$

Using the same logic, we get the following result for the second component of $X_j^{\widehat{\pi}_{i:j}}(i, X_i^{\pi_{0:i}^{(0)}})$:

$$S_j^{\widehat{\pi}_{i:j}}(i, X_i^{\pi_{0:i}^{(0)}}) = S_j^{\pi_{i:j}^{(0)}}(i, X_i^{\pi_{0:i}^{(0)}}) \exp\left(-\kappa \sum_{q=i}^{j-1} \left(I_q^{\widehat{\pi}_{i:q}}(i, X_i^{\pi_{0:i}^{(0)}}) - I_q^{\pi_{i:q}^{(0)}}(i, X_i^{\pi_{0:i}^{(0)}})\right)\delta\right).$$

The algorithm goes as follows:

1. We simulate M paths for $(\Delta\varepsilon_i^I, \Delta\varepsilon_i^S)_{i=1:N}$ to get controlled trajectories with an arbitrarily chosen admissible deterministic strategy $\pi_{0:N}^{(0)}$. The strategy chosen for our experiments is described in (5.2).

2. In view of the dynamic programming equation (3.12) and (3.4), it is needed to compute the following quantity for each $i = N-1, \ldots, 0$:

$$R_\gamma(i, x) := \mathbb{E}_\gamma\left[J^*(i+1, X_{i+1}) \mid X_i = x\right].$$

3. We first compute an approximation of J^* as an explicit function of γ using backward sampling:

$$J^*(i+1, X_{i+1}^{\pi_{0:i+1}^{(0)}}) \approx \mathbb{E}\Bigg[\sum_{j=i+1}^{N-1} \frac{\mathcal{C}_j\left(X_j^{\widehat{\pi}_{i+1:j}}(i+1, X_{i+1}^{\pi_{0:i+1}^{(0)}}), \widehat{\gamma}_j(X_j^{\widehat{\pi}_{i+1:j}}(i+1, X_{i+1}^{\pi_{0:i+1}^{(0)}}))\right)}{(1+r\delta)^{j-i-1}} + \frac{\mathcal{C}_N\left(X_N^{\widehat{\pi}_{i+1:N}}(i+1, X_{i+1}^{\pi_{0:i+1}^{(0)}})\right)}{(1+r\delta)^{N-i-1}} \mid X_{i+1}^{\pi_{0:i+1}^{(0)}}\Bigg].$$

It yields a representation of $J^*(i+1,\cdot)$ as a conditional expectation with explicit quantities, thus a regression-based approach can be used. We approximate this scalar function using L basis functions $\phi = \{\phi_l\}_{l=1,\dots,L}$,

$$\hat{J}^*(i+1,\cdot) := \sum_{l=1}^{L} \hat{\alpha}_l^i \phi_l(\cdot) = \hat{\alpha}^i \cdot \phi(\cdot), \tag{4.3}$$

where $\hat{\alpha}_l^i, l = 1,\dots,L$ are coefficients obtained using least squares over the M simulations:

$$\hat{\alpha}^i = \arg\min_{\alpha} \sum_{m=1}^{M} \left[\sum_{j=i+1}^{N-1} \frac{\mathcal{C}_j\left(X_j^{\widehat{\pi}_{i+1:j},m}(i+1, X_{i+1}^{\pi^{(0)}_{0:i+1},m}), \widehat{\gamma}_j(X_j^{\widehat{\pi}_{i+1:j},m}(i+1, X_{i+1}^{\pi^{(0)}_{0:i+1},m}))\right)}{(1+r\delta)^{j-i-1}} \right.$$
$$\left. + \frac{\mathcal{C}_N\left(X_N^{\widehat{\pi}_{i+1:N},m}(i+1, X_{i+1}^{\pi^{(0)}_{0:i+1},m})\right)}{(1+r\delta)^{N-i-1}} - \alpha \cdot \phi\big(X_{i+1}^{\pi^{(0)}_{0:i+1},m}\big) \right]^2,$$

with $\widehat{\gamma}_j$ and $X_j^{\widehat{\pi}_{i+1:j},m}(i+1, X_{i+1}^{\pi^{(0)}_{0:i+1},m})$ respectively the best known to date i intensity reduction function process for any date $j > i$ and its corresponding state variable for path m.

4. It remains to estimate the conditional expectation $R_\gamma(i,x) = \mathbb{E}_\gamma\left[J^*(i+1, X_{i+1}) \mid X_i = x\right]$. It is tempting to apply an extra regression: however, it does not lead to a representation of $R_\gamma(i,x)$ in a form suitable for the minimization in (3.12). This is why we prefer to use an integral representation of the conditional expectation $R_\gamma(i,x)$. By definition of the conditional expectation, R_γ rewrites as:

$$R_\gamma(i,x) = \mathbb{E}_\gamma\left[J^*(i+1, X_{i+1}) \mid X_i = x\right]$$
$$= \iint_{\mathcal{G}} J^*\left(i+1, \left(x^{(1)} e^{-\gamma\delta} e^{\sigma_I z - \psi_i^I(\sigma_I)}, x^{(2)} \frac{\bar{S}_{i+1}}{\bar{S}_i} e^{-\kappa\left(x^{(1)} - I_i^{\text{ref}}\right)\delta} e^{\sigma_S y - \psi_i^S(\sigma_S)}\right)\right) \mathbb{P}_{\Delta\varepsilon_i^I, \Delta\varepsilon_i^S}(\mathrm{d}z, \mathrm{d}y),$$

where $\mathcal{G}$ and $\mathbb{P}_{\Delta\varepsilon_i^I, \Delta\varepsilon_i^S}$ are respectively the support and the joint probability of $(\Delta\varepsilon_i^I, \Delta\varepsilon_i^S)$. R_γ is then estimated by substituting J^* by (4.3) and applying a numerical integration to the above double integral:

$$R_\gamma(i,x) \approx \sum_{q_1,q_2} w_{q_1,q_2} \hat{\alpha}^i \cdot \phi\left(x^{(1)} e^{-\gamma\delta} e^{\sigma_I z_{q_1} - \psi_i^I(\sigma_I)}, x^{(2)} \frac{\bar{S}_{i+1}}{\bar{S}_i} e^{-\kappa\left(x^{(1)} - I_i^{\text{ref}}\right)\delta} e^{\sigma_S y_{q_2} - \psi_i^S(\sigma_S)}\right)$$
$$=: \widehat{R}_\gamma(i,x),$$

where the points $(z_{q_1}, y_{q_2})_{q_1,q_2}$ and weights $(w_{q_1,q_2})_{q_1,q_2}$ are related to an integration rule on the set $\mathcal{G}$ w.r.t. the distribution $\mathbb{P}_{\Delta\varepsilon_i^I, \Delta\varepsilon_i^S}$.

5. For each $x \in \{X_i^{\pi_{0:i}^{(0)},m}\}_{m=1,\dots,M}$, we solve the following with a one-dimensional minimization algorithm:

$$\gamma_i^{\widehat{\pi},m} = \arg \min_{\gamma \in [0,\gamma_{\max}]} \left[\mathcal{C}_i(x,\gamma) + \frac{1}{1+r\delta} \widehat{R}_\gamma(i,x) \right].$$

6. Then, we get $\gamma_i^{\widehat{\pi}}(\cdot)$ by linearly regressing $\{\gamma_i^{\widehat{\pi},m}\}_{m=1,\dots,M}$ on $\{X_i^{\pi_{0:i}^{(0)},m}\}_{m=1,\dots,M}$. Note that we could consider more general features for the regression, however in our tests we have not observed any significant improvement in doing so compared to the current features; if the regression R-squared is too small ($<50\%$), an interpolation could be alternatively used.

7. We repeat this process using the obtained strategy $\widehat{\pi}_{0:N}$ as the new starting point strategy $\pi_{0:N}^{(0)}$, until we achieve stability of results and set $\pi_{0:N}^* = (\gamma_i^*(\cdot))_{i=0,\dots,N}$ as the last $\widehat{\pi}_{0:N}$. This strategy update step allows the statistical learning part of the algorithm to learn on a dataset closer to the one with the optimal strategy, than if we had stayed with the initial reference strategy; this contributes to better results. In practice, the number of updates remains low (a few units are enough in our tests).

We summarize this method in Algorithm 1.

Algorithm 1: Resolution of the Dynamic Programming Algorithm by Backward Sampling

Input: *Number of Monte Carlo Simulations:* $M \gg 1$,
Number of epochs: $K \in \{1, 2, ...\}$,
Set of basis functions: $\{\phi_l\}_{l=1}^{L}$,
Starting strategy: $\pi^{(0)}$.

Output: *Optimal process function:* $\gamma_i^*(\cdot) \quad \forall i = 1, \dots, N-1$.

Sample M Monte Carlo paths for $\{(X_i^{\pi^{(0)},m})_{i=0,\dots,N}\}_{m=1:M}$ with strategy $\pi^{(0)}$.

\# Iteration over reference strategies

Set $\widehat{\pi}^{(0)} := \pi^{(0)}$.

for $k = 1 : K$ **do**

  Set $X_N^{\widehat{\pi}^{(k)},m} := X_N^{\widehat{\pi}^{(k-1)},m}$ for all simulated paths.

  \# Backward iteration in time

  for $i = N-1, \dots, 0$ **do**

    \# Computation of sample data for regression

    for $m = 1, \dots, M$ **do**

      Compute

$$J_{i+1}^{\widehat{\pi}^{(k)},m} = \sum_{j=i+1}^{N-1} \frac{\mathcal{C}_j\Big(X_j^{\widehat{\pi}^{(k)}_{i+1:j},m}(i+1, X_{i+1}^{\widehat{\pi}^{(k-1)}_{0:i+1},m}), \widehat{\gamma}_j^{(k)}\big(X_j^{\widehat{\pi}^{(k)}_{i+1:j},m}(i+1, X_{i+1}^{\widehat{\pi}^{(k-1)}_{0:i+1},m})\big)\Big)}{(1+r\delta)^{j-i-1}} + \frac{\mathcal{C}_N\Big(X_N^{\widehat{\pi}^{(k)}_{i+1:N},m}(i+1, X_{i+1}^{\widehat{\pi}^{(k-1)}_{0:i+1},m})\Big)}{(1+r\delta)^{N-i-1}}.$$

    Compute $\hat{\alpha}^i = \arg\min_\alpha \sum_{m=1}^{M} \Big[J_{i+1}^{\widehat{\pi}^{(k)},m} - \alpha \cdot \phi\Big(X_{i+1}^{\widehat{\pi}^{(k-1)}_{0:i+1},m}\Big)\Big]^2$.

    Cross-validation: Choose the number $L = 1, \dots, 20$ of basis functions that maximizes the R^2-score of the least squares regression.

    \# Numerical integration of reward function

    Set $\widehat{R}_\gamma(i,x) := \sum_{q_1,q_2 \in Q} w_{q_1,q_2} \times \hat{\alpha}^i \cdot \phi\Big(x^{(1)} e^{-\gamma\delta} e^{\sigma_I z_{q_1} - \psi_i^I(\sigma_I)}, x^{(2)} \frac{\bar{S}_{i+1}}{\bar{S}_i} e^{-\kappa\left(x^{(1)} - I_i^{\text{ref}}\right)\delta} e^{\sigma_S y_{q_2} - \psi_i^S(\sigma_S)}\Big)$.

    \# Optimal control for each sample

    for *m=1,...,M* **do**

      Solve $\gamma_i^{\widehat{\pi}^{(k)},m} = \arg\min_{\gamma \in [0, \gamma_{\max}]} \left\{\mathcal{C}_i(X_i^{\widehat{\pi}^{(k-1)}_{0:i},m}, \gamma) + \frac{1}{1+r\delta} \widehat{R}_\gamma(i, X_i^{\widehat{\pi}^{(k-1)}_{0:i},m})\right\}$.

    \# Learning the control function

    Compute $\widehat{\beta^i} = \arg\min_{\beta^i = (\beta_0^i, \beta_1^i)} \sum_{m=1}^{M} \left[\gamma_i^{\widehat{\pi}^{(k)},m} - (\beta_0^i + \beta_1^i \cdot X_i^{\widehat{\pi}^{(k-1)}_{0:i},m})\right]^2$.

    Update $\{(X_j^{\widehat{\pi}^{(k)}_{i:j},m}(i, X_i^{\widehat{\pi}^{(k-1)}_{0:i},m}))_{j=i,\dots,N}\}_{m=1,\dots,M}$, using the same noises $\Delta\varepsilon^I$ and $\Delta\varepsilon^S$.

5 Application to fictitious companies

5.1 Description of scenarios

We have based our application on two NGFS phase III energy transition scenarios [RBK+22], namely, *Below 2 Degrees*, and *Delayed Transition* (see Table 1). Both show different pathways running from 2020 to 2050. They provide different trajectories for many macroeconomic variables, such as gross domestic product (GDP), inflation, and energy consumption, on different geographic scales (from region to country) using three IAMs (Remind-MAgPIE, GCAM, and MESSAGEix-GLOBIOM). They propose sectoral pathways based on the NACE classification [Eur08] up to level 2. We have focused on a transition vulnerable sector 'D35: Electricity, gas, steam, and air conditioning supply' for France.
For our analysis, we have extracted three variables' trajectories from these scenarios (Remind-MAgPIE): the carbon price, the sectoral emissions, and the sectoral sales revenues (Fig. 1). Note that the sales revenues are not one of the main variables of the NGFS scenarios. We have inferred the trajectories using the general equilibrium model from *Banque de France* [DL20] similarly to the methodology employed for the 2020 French banking authority climate stress test [AdCPedR20].

Below 2 Degrees (B2C). B2C is a central scenario in which climate regulators take action as soon as 2020 to achieve the *'well below two Celsius degrees'* as laid out in the 2015 Paris Agreement. The stringency of climate policies gradually increases, leaving enough time for the economic actors to adapt and plan their adaptation to the new low-carbon economy. Worldwide Net-zero CO_2 emissions are achieved after 2070, taming both physical and transition risks. This scenario has an estimated chance of 67% of limiting global warming to below two Celsius degrees. Because this scenario is less disruptive and allows us to achieve transition goals, it will be used as our benchmark scenario for the remainder of the analysis.

Delayed Transition (DT). In DT, no more policies than the ones currently in place are introduced before 2030. Until then, global emissions have grown, making the need for a transition even higher and more arduous to fulfill. Incisive policies are thus required to limit warming to below two degrees. This scenario still leads to a successful transition, i.e. with global warming limited to two degrees. However, disruptive climate policies are employed to achieve it. The level of action differs across countries and regions. Carbon Dioxide Removal (CDR) technologies are scarce, negative emissions are limited, leading to a more rapid decline in carbon emissions than in the B2C scenario from 2030 onwards.

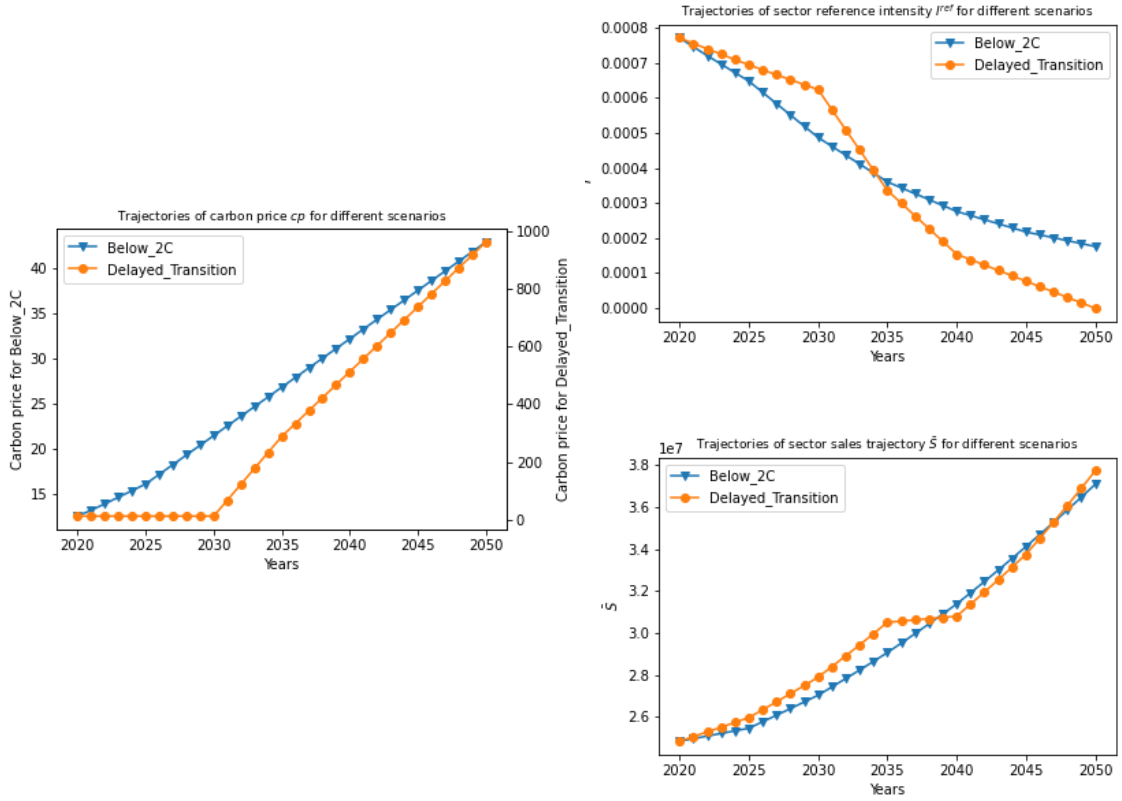


Figure 1: Trajectories of carbon price ($\mathtt{cp}_i$), reference intensity (I^{ref}) and sector sales revenues ($\bar{S}$) extracted from the NGFS phase 2 scenarios for sector D35 for France. Starting point: $\mathtt{cp}_0 = 12.56\mathtt{USD/tCO_2}$, $I_0^{ref} = 7.71E{-}4\mathtt{tCO_2/USD}$, $\bar{S}_0 = 2,48E7\mathtt{USD}$.

5.2 Choice of parameters and hyperparameters

Model parameters. We have considered a 6 dates problem ($N = 6, \delta = 5$), and have ran tests on three fictitious companies that only differ in starting point intensity: $I_0^1 = 1.16{-}3$ (brown), $I_0^2 = 7.71E{-}4$ (yellow), $I_0^3 = 3.86E{-}4$ (green). Note that $I_0^{ref} = I_0^2$ for all scenarios. Table 2 summarizes all other parameters used. Because of the steady feature of companies' revenues and intensity, we have preferred low volatility levels (2%). The disturbance process is $(\Delta\varepsilon_i^I, \Delta\varepsilon_i^S) = (\sqrt{\delta}\min(\max(Z^1, -3), 3); \sqrt{\delta}\min(\max(Z^2, -3), 3))$ for all i where (Z^1, Z^2) is a centered and normalized 2-dimensional Gaussian variable such that $cov(Z^1, Z^2) = -0.3$.
The parameters chosen for the investment cost function are the same as those proposed by default in [Nor17]. We have set $\gamma_{\max}$ such that the maximum yearly decrease rate of intensity is no greater than 40%. This is quite a high upper bond, which is not hit in most cases. We have assumed a constant carbon cost $\mathtt{cp}_N I_N S_N$ for any $j \geq N$ that the company will pay for an infinite number of dates after the end of the scenario. This quantity is then discounted,

Scenario	Temperature increase in 2100	Policy type	Technology	CDR	Regional Policy Variation
B2C	1.6°C	Immediate and smooth	Moderate change	Medium-high	Low variation
DT	1.6°C	Delayed	Slow then Fast change	Low-medium	High variation

Table 1: Summary of the different NGFS Scenarios studied for the analysis. CDR means carbon Dioxide Removal technologies (Source: NGFS).

which leads to:

$$\mathcal{C}_N(I_N, S_N) := \mathtt{cp}_N I_N S_N \frac{1 + r\delta}{r\delta}. \tag{5.1}$$

σ_I	S_0	σ_S	κ	α	β	c	r	$\gamma_{\max}$
0.02	8.4×10^5	0.02	30	0.95	2.8	1.26	6%	$1 - e^{-\gamma_{\max}} = 0.4$

Table 2: Values of model parameters for the toy example.

Algorithm Parameters. The starting point strategy (recall π^0 in Algorithm 1) is designed such that γ_i provides the same decrease rate as the sectoral intensity between year i and $i+1$, capped at $\gamma_{\max}$. The elected basis function family is polynomials, including cross products of the intensity and sales. It provided satisfactory results (R^2 of regression equal to 90% on average) with small numbers of functions (up to degree 4). We have used SVD to solve the least square problem in order to bypass possible multi-colinearity issues raised by the chosen basis functions. We have computed the double integral using Simpson's method over a product grid of 10×10 points uniformly distributed on $[-3, 3] \times [-3, 3]$. Despite the small number of points, the precision was more than satisfactory. We used BFGS' method with a tolerance set to 10^{-6} for the minimization step. The optimal function is inferred using *statsmodels* OLS function when the R-squared is sufficient, otherwise with *griddata* function from the *scipy interpolate function*. All tests were run on Python 3.7. with Intel Xeon CPU Gold 6230 20 cores @ 2.1 Ghz.

5.3 Results and interpretation

5.3.1 Description of the strategies

For each scenario, we have computed the results for three benchmark strategies. The first one is an *exogenous* (exo) strategy consisting in selecting the same reduction rate as the reference intensity, omitting the disturbance. Obviously, the reduction rate cannot be greater than 40%. It is the same as the strategy $\pi^{(0)}$ described before. The second one, called the *myopic* (myo)

strategy, consists of selecting the reduction rate such that the investment at date i is the same as the cost of paying the carbon tax at date i:

$$\pi^{\text{exo}} := (\gamma_i^{\text{exo}})_{i=0,\dots,N}\,, \qquad e^{-\gamma_i^{\text{exo}}\delta} = \max\left(\frac{I_{i+1}^{ref}}{I_i^{ref}}, e^{-\gamma_{\max}\delta}\right), \tag{5.2}$$

$$\pi^{\text{myo}} := (\gamma_i^{\text{myo}}(I,S))_{i=0,\dots,N}\,, \qquad \gamma_i^{\text{myo}}(I,S) = \min\left(-\frac{1}{\delta}\ln\left(1-\left(\beta\frac{\mathtt{cp}_i I}{c\alpha^i}\right)^{\frac{1}{\beta}}\right), \gamma_{\max}\right),$$

for all $i = 0, \dots, N-1$.

Note that our definition of the exogenous strategy gets around scenarios where the reference intensity falls to 0. Moreover, the myopic strategy is not a direct function of the firm's sales, but it depends on its intensity level. In some cases, the unitary carbon cost of the company may be too high for a solution to exist, we cap it to $\gamma_{\max}$. Eventually, we also consider the uncontrolled strategy, i.e. with $\pi := \mathbf{0}_N$ denoted by the exponent 0. Finally, the optimal strategy π^* is the one obtained by solving the minimization problem defined in (2.8) using the backward sampling algorithm. Table 4 summarizes all different scenarios, firms, and strategies considered.

5.3.2 Main results for different companies

As expected, not controlling the intensity (i.e. $\pi = \mathbf{0}_N$) largely decreases the companies sales revenues and inflates the carbon tax. This happens even for the green company in an orderly transition setting (Fig. 4) where the sum of total discounted sales revenues is reduced by 3.3%. More importantly, the uncontrolled strategy leads to a decreasing trend in sales in all scenarios (Figs. 2 & 5). The gap widens with time, which makes long-term forward looking credit risk metrics with a constant intensity assumption (such as in [BLG20, BKM22]) largely degraded.

In contrast, all other strategies yield similar total discounted sales revenues levels over the scenario (see Table 5). For instance, for the yellow company in scenario B2C this amounts to $\mathtt{USD}3.40E6$ for the optimal strategy, $\mathtt{USD}3.446$ for the exogenous strategy and $\mathtt{USD}3.49E6$ for the myopic strategy. Regarding other variables, the exogenous strategy gives results that are close to the myopic strategy both in terms of carbon tax in the central scenario (e.g. scenario B2C Green company : exogenous strategy = $\mathtt{USD}18,056$ vs myopic = $\mathtt{USD}17,467$). However, this is reversed when the scenario displays shocks because the investment decision of the exogenous strategy are independent from the carbon price and other scenario's aspects. Particularly, it can be seen in Figure 6 that the exogenous strategy prevents the company to set buffer for the 2030 carbon price shock (e.g. total expected discounted carbon cost : exogenous strategy = $\mathtt{USD}111,387$, myopic strategy = $\mathtt{USD}89,055$, optimal strategy = $\mathtt{USD}72,693$). The myopic strategy does act earlier than the exogenous one, however its investments are poorly planned due to lack of foresight, leading to a bump in investment efforts in 2035. The forward-looking

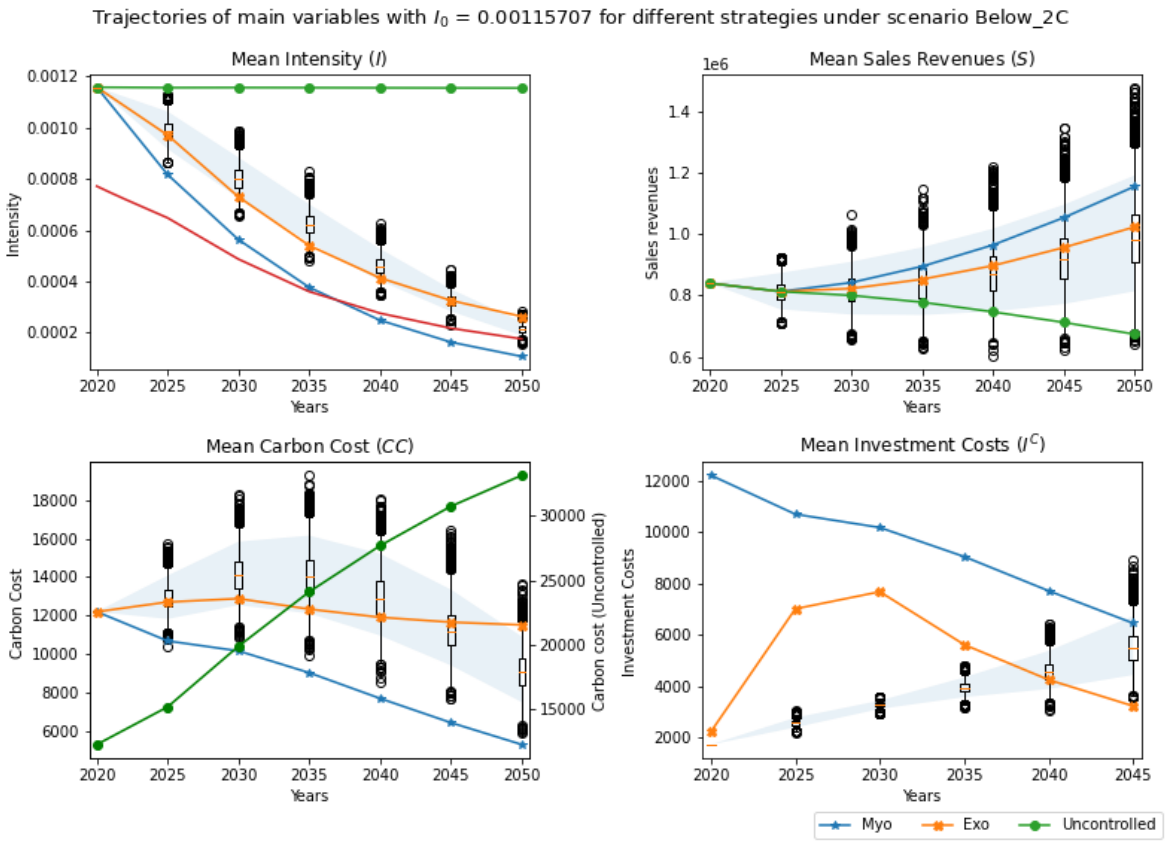


Figure 2: Simulated results for the **brown company** with the optimal strategy (boxplots) and benchmark strategies (myopic = stars, exogenous = squares, uncontrolled = dots) in scenario B2C and reference intensity I^{ref} (red line). Shaded area : 90% confidence interval for trajectories under the optimal strategy. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs. Carbon cost for the uncontrolled strategy to be read on the RHS y-axis.

feature of the optimal strategy allows not only to time the investment accordingly but also to arbitrage between paying more carbon tax and less in the carbon reduction strategy.
The optimal strategy takes advantage of both the discount factor and the autonomous decreasing feature of the investment costs to time investments when they are less costly for the brown company (Fig. 2 & 5). In particular, it waits until the last period before the 2030 carbon price shock to invest in its intensity mitigation plan to benefit from the drop in green technologies price (Figures 5, 6 & 7). For this reason, it systematically supersedes the other strategies, both in terms of total carbon cost and investment costs (Table 5). Note that, in scenario B2C, the brown company would rather pay more carbon tax than attempting to fol-

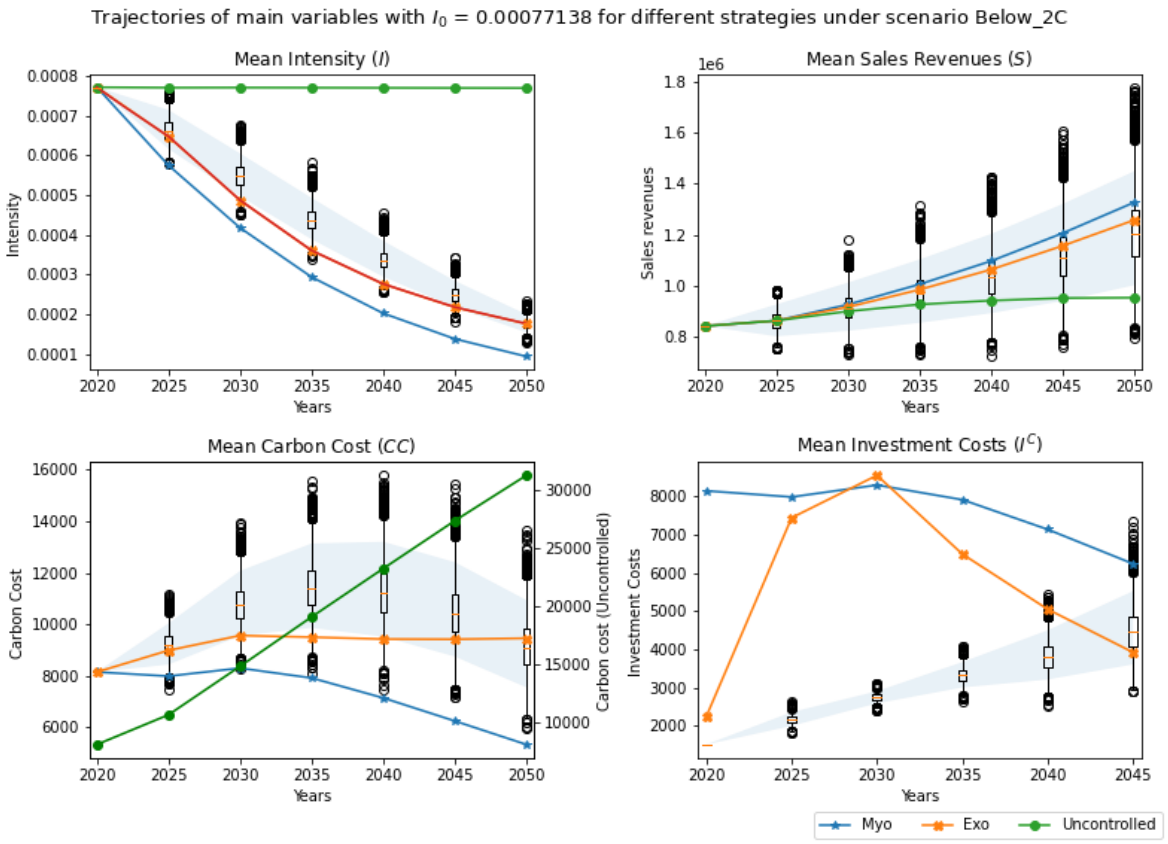


Figure 3: Simulated results for the **yellow company** with the optimal strategy (boxplots) and benchmark strategies (myopic = stars, exogenous = squares, uncontrolled = dots) in scenario B2C and reference intensity I^{ref} (red line). Shaded area : 90% confidence interval for trajectories under the optimal strategy. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs. Carbon cost for the uncontrolled strategy to be read on the RHS y-axis.

low its sectoral average intensity because of the convexity of the investment cost in a smooth transition (Fig. 2). However, the paradigm flips in a disorderly transition setting due to the shoot up of carbon price in 2030 (Fig. 5).

For the green company, the optimal intensity is higher than the other strategies in scenario B2C (Fig. 4) and is gradually caught with the sectoral average. The need for carbon mitigation is weaker since the company starts with a low intensity. This leads to small investment levels: $\texttt{USD}5,459$ total expected discounted investment costs over the scenario, against $\texttt{USD}20,698$ for the exogenous strategy and $\texttt{USD}16,505$ for the myopic strategy. However, the smaller investment costs are partially compensated by the augmented total expected discounted carbon

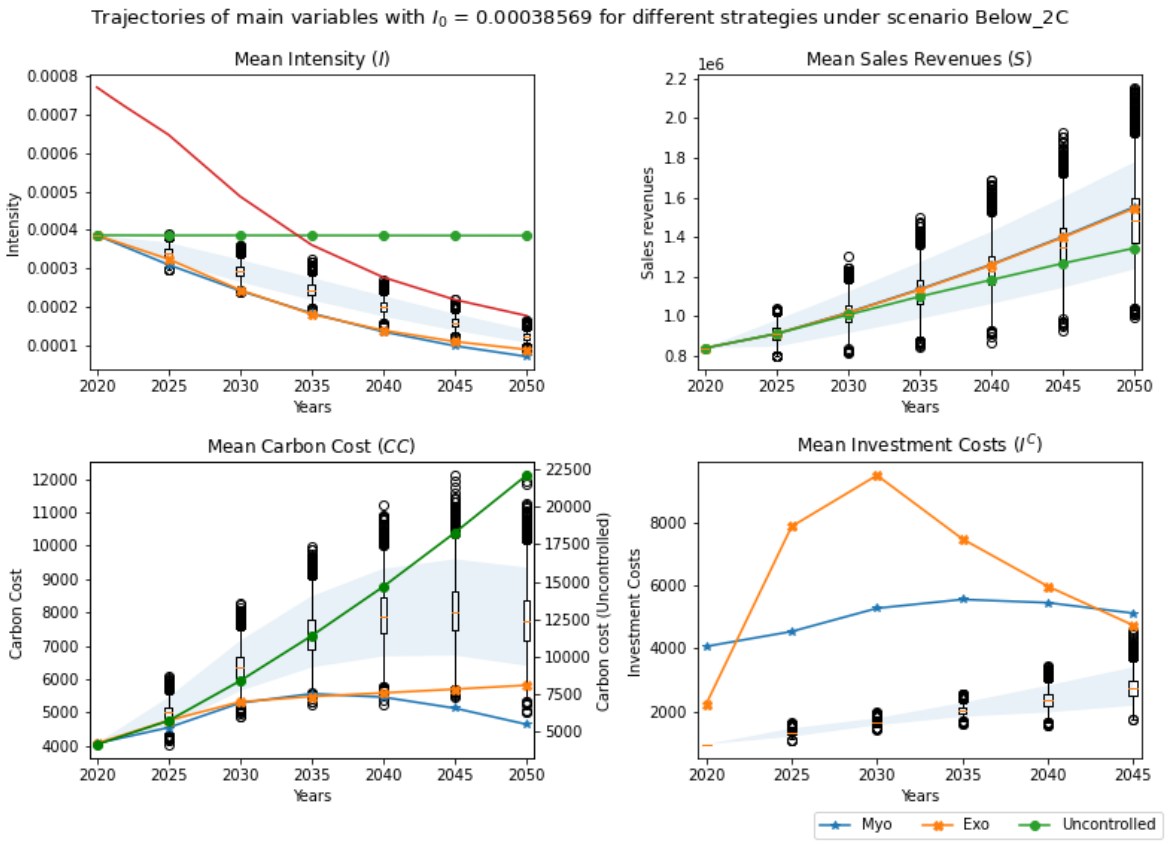


Figure 4: Simulated results for the **green company** with the optimal strategy (boxplots) and benchmark strategies (myopic = stars, exogenous = squares, uncontrolled = dots) in scenario B2C and reference intensity I^{ref} (red line). Shaded area : 90% confidence interval for trajectories under the optimal strategy. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs. Carbon cost for the uncontrolled strategy to be read on the RHS y-axis.

costs (scenario B2C: optimal = $\texttt{USD}21,583$, exogenous = $\texttt{USD}18,056$, myopic = $\texttt{USD}17,467$) with little difference in total expected discounted sales revenues (optimal = $\texttt{USD}3.76E6$, exogenous = $\texttt{USD}3.80E6$, myopic = $\texttt{USD}3.80E6$). Again, this facet is reversed in scenario DT because the company braces against the 2030 carbon price shock. It invests earlier than with the other strategies, allowing lower carbon costs (optimal = $\texttt{USD}54,168$, exogenous = $\texttt{USD}64,864$, myopic = $\texttt{USD}69,458$). Then, the optimal intensity converges towards that of other strategies.
The yellow company displays two different behaviors. On the one hand, in B2C, the optimal strategy stays slightly above the sector average because of manageable carbon costs and to avoid unnecessary convexly increasing investment costs (Fig. 3). This allows to divide

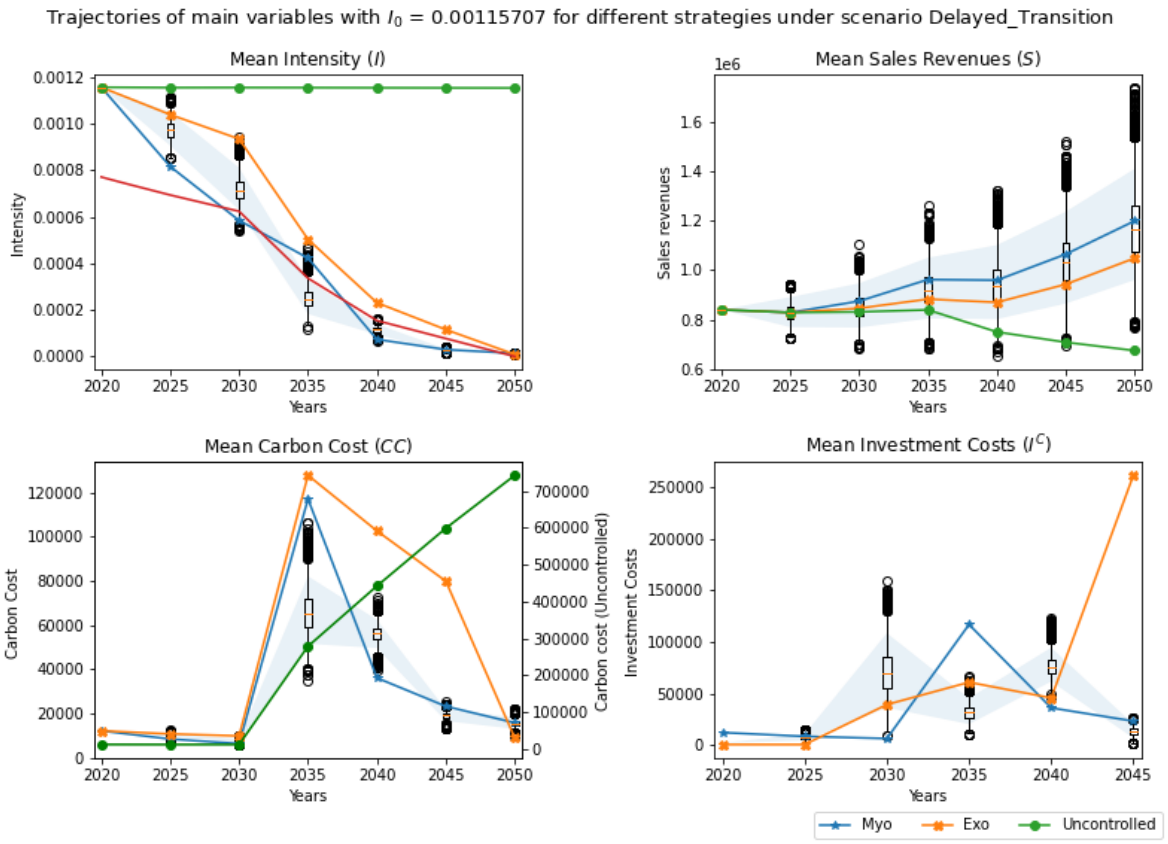


Figure 5: Simulated results for the **brown company** with the optimal strategy (boxplots) and benchmark strategies (myopic = stars, exogenous = squares, uncontrolled = dots) in scenario DT and reference intensity I^{ref} (red line). Shaded area : 90% confidence interval for trajectories under the optimal strategy. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs. Carbon cost for the uncontrolled strategy to be read on the RHS y-axis.

investment costs by more than 2 (optimal = USD8,869, exogenous = USD18,789, myopic = USD26,948) with little difference in both sales and carbon costs. On the other hand, similarly to the green company, the optimal intensity is cut much lower and earlier than the other strategies in order to dodge the carbon price shock (Fig. 6).

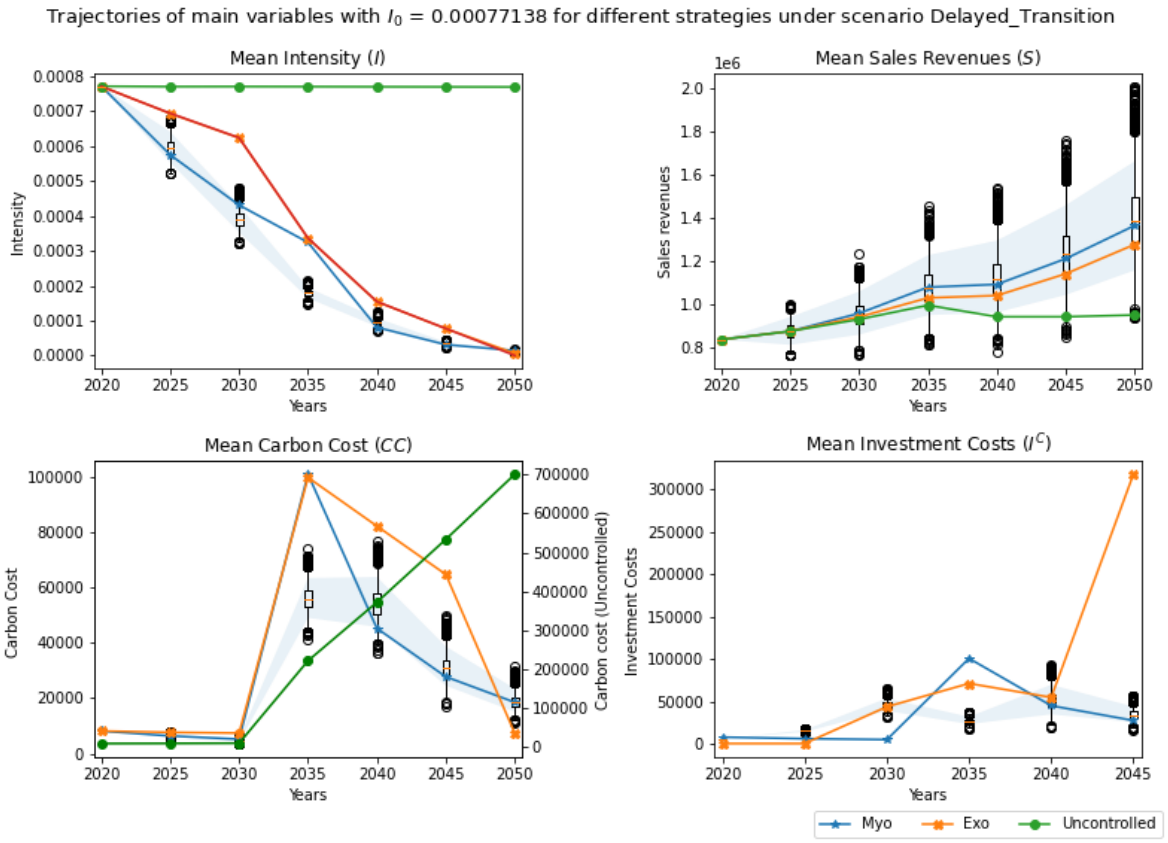


Figure 6: Simulated results for the **yellow company** with the optimal strategy (boxplots) and benchmark strategies (myopic = stars, exogenous = squares, uncontrolled = dots) in scenario DT and reference intensity I^{ref} (red line). Shaded area : 90% confidence interval for trajectories under the optimal strategy. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs. Carbon cost for the uncontrolled strategy to be read on the RHS y-axis.

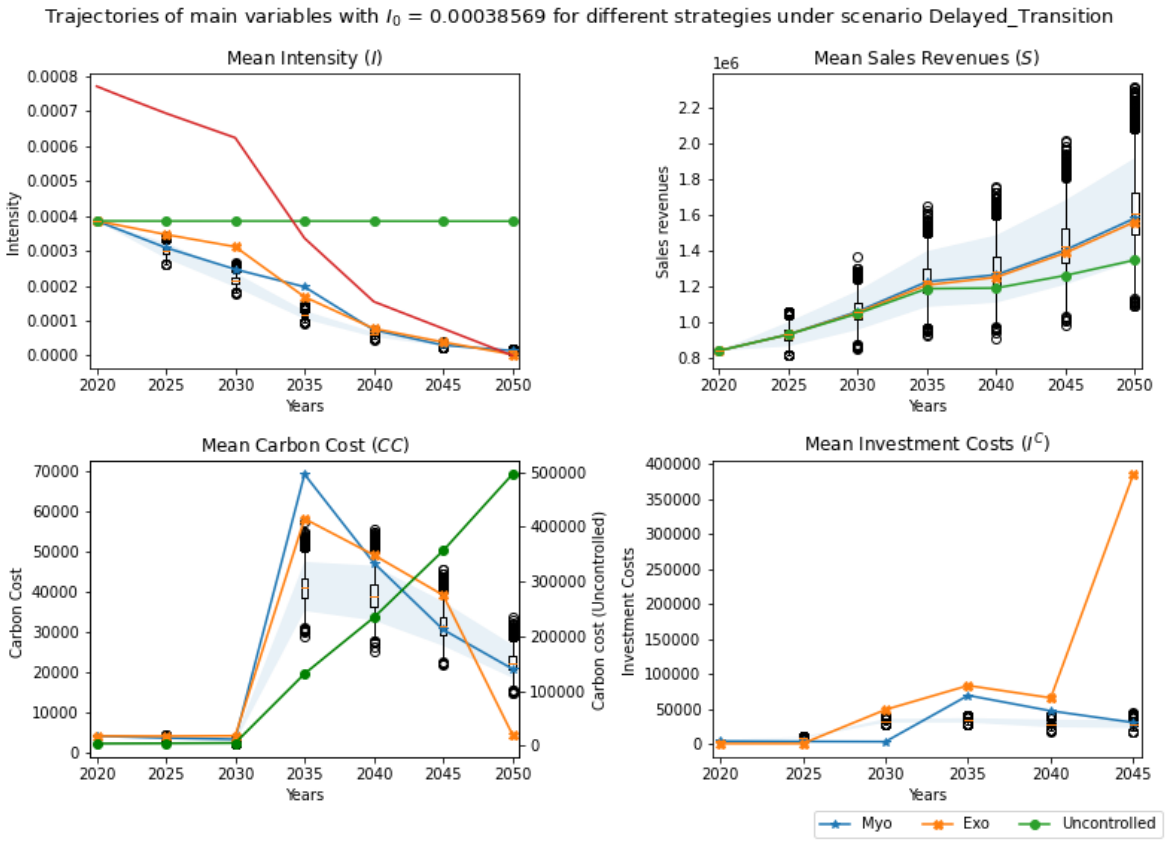


Figure 7: Simulated results for the **green company** with the optimal strategy (boxplots) and benchmark strategies (myopic = stars, exogenous = squares, uncontrolled = dots) in scenario DT and reference intensity I^{ref} (red line). Shaded area : 90% confidence interval for trajectories under the optimal strategy. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs. Carbon cost for the uncontrolled strategy to be read on the RHS y-axis.

5.4 Numerical sensitivity analysis of the optimal trajectories to the model parameters

5.4.1 Methodology of the sensitivity analysis

We have numerically studied the sensitivity of the optimal trajectories to some of the model parameters, namely : α, β, c, κ and r. For each model parameter considered, we have run 100 simulations with values taken in a carefully chosen interval, with constant step. The corresponding intervals are disclosed in Table 3. We have focused on the yellow company (i.e. such that $I_0 = I_0^{ref}$) and the scenario B2C because they are the most central. For this matter,

α	β	c	κ	r
$[0.95; 1]$	$[2; 3.6]$	$[0.63; 1.89]$	$[0; 100]$	$[0; 0.1]$

Table 3: Ranges of values used for the numerical sensitivity analysis of the optimal trajectories to the model parameters.

we have considered a 6-dates noise-free scenario, i.e. with $\sigma_I = \sigma_S = 0$. Since the value of the sales revenues does not depend on the control process at date i in this particular setting, the new dynamic programming equation thus becomes:

$$\gamma_i^*(x) = \arg \min_{\gamma \in [0, \gamma_{\max}]} \left(\mathcal{C}_i(x, \gamma) + \frac{1}{1 + r\delta} J^*(i+1, X_{i+1}(x)) \right), \quad \forall i = N-1, \dots, 0, \quad x \in (\mathbb{R}^+)^2,$$

$J^*(N, x) = \mathcal{C}_N(x)$.

Working our way from $N-1$ towards 0, we solve the dynamic programming equation at each date i by computing the trajectories over a grid of 500 possible values for $\gamma \in [0, \gamma_{\max}]$ and by selecting the value conferring the lowest cost. We repeat this step for 500 values for $X^{(1)} \in (0, I_0]$ since the intensity can only decrease asymptotically towards 0. We then learn the optimal process function by regressing the optimal values for γ on the values we have used for $X^{(1)}$. The results are detailed in the following paragraphs.

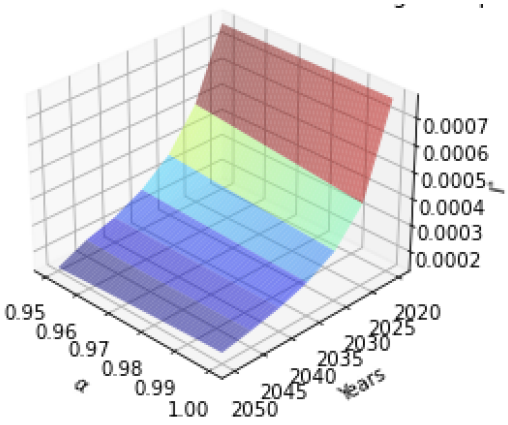
5.4.2 Sensitivity to α

We study the effects of different values for the autonomous factor of investment cost decrease $\alpha \in [0.95, 1]$ as defined in (2.4) on the optimal trajectories. We see in Fig. 8a that the increase in α yields a decrease in the optimal intensity in the beginning of the scenario, whereas a slight increase by the end. Indeed, for 2025, there is a 14.3% drop in the relative emissions between $\alpha = 0.95$ and between $\alpha = 1$, compared to a 20.8% raise at the end of the scenario. Since α is the autonomous factor of decrease of the investment costs, this is economically sound. The higher α is, and the less beneficial it is for the company to wait before investing in reducing its relative emissions. Thus, there is a more significant cut in relative emissions in the first years

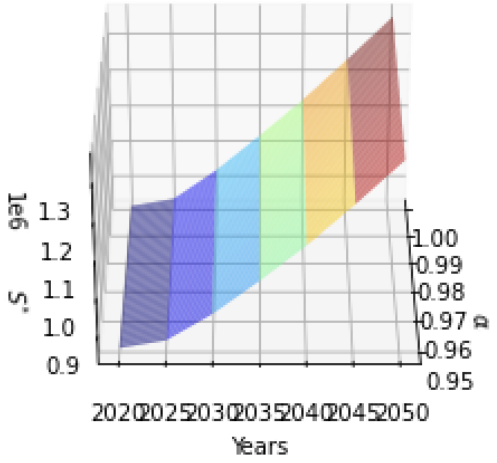
of the scenario when α tends to one. This pattern is confirmed by the investment costs (Fig. 8d) with yearly investments increasing with α in 2020, and decreasing for all years after 2030. Similarly to the optimal intensity, the investments costs increase with α in the first part of the scenario, however this trend is compensated by the end of the scenario (from $\texttt{USD}18,093$ for $\alpha = 0.95$ in 2020 to $\texttt{USD}36,079$ for $\alpha = 1$).
Moreover, an augmented α factor implies inflated unitary intensity mitigation costs by the end of the scenario meaning that a same reduction effort will be more onerous. Thus, delaying the investment becomes less and less beneficial, leading to a shoot up of total investment costs in 2020 (Fig. 8d) and a slight increase in carbon tax with respect to α in 2050 (Fig 8c).
Overall, the sales revenues are hardly altered (Fig. 8b). However, the results regarding the sales revenues are dependent on the value of κ we have used as we will see in Section 5.4.5. In conclusion, higher levels of α lead to lower optimal γ because of augmented investment costs, thus higher optimal intensity as well as higher investment costs, with negligible changes in sales, thus higher carbon cost.

5.4.3 Sensitivity to β

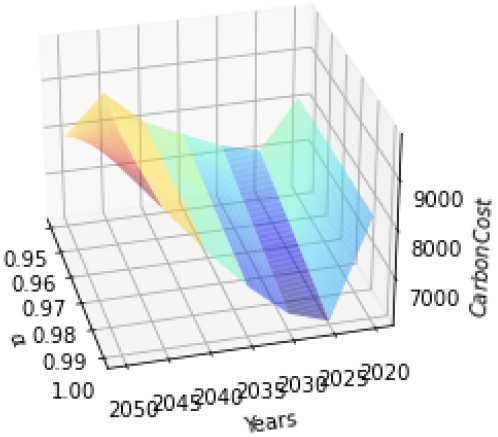
The study of the impacts the convexity factor of the investment costs β with respect to the intensity reduction rate displays interesting results. Contrarily to α (see Section 5.4.2), an increase in β results in exponentially decaying optimal intensity for all dates (Fig. 9a). This is because β is also in the denominator of the investment costs function (see (2.4)); and the intensity reduction rate is in $(0, 1]$, thus $\dfrac{\partial I^C}{\partial \beta}(.) < 0$ *ceteris paribus*. Clearly, an increase in β actually implies lower unitary investment costs which encourages higher emissions mitigation efforts. Precisely, the investment costs are convexly increasing with β and decreasing with time, with a peak displayed in 2020 with $\beta = 3.6$ at $\texttt{USD}44,563$ (Fig. 9d). Slashing the firm's relative emissions pushes sales up, with a convex feature (Fig. 9b). Nonetheless, carbon costs go down despite convexly increasing sales revenues with respect to β (Fig. 9c) thanks to the deep cut in the intensity at the beginning of the scenario.
The change in β actually reverses the curvature of the carbon cost with respect to the scenario: it goes from concave for lower values for β to convex when β gets closer to 3.6. This is explained by the interaction of different features of the model. First of all, a decrease in intensity *ceteris paribus* leads to a drop in carbon costs. However, for any $\kappa > 0$, a decrease in intensity implies a convex increase in sales, which intervenes in the carbon cost computation. The final result on carbon cost will depend on which variable moves faster. With the intensity being concave and the sales being convex with respect to β, the intensity effect will gradually lose the upper hand. In conclusion, higher β leads to lower unitary investment, thus, lower intensity levels that drive sales up, but this surge is compensated by the relative emissions drop in the carbon cost computations.



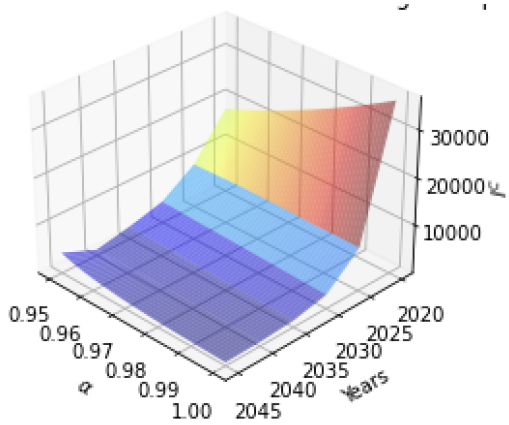
(a) I^* through years for different values of α.



(b) S^* through years for different values of α.



(c) C^C through years for different values of α.

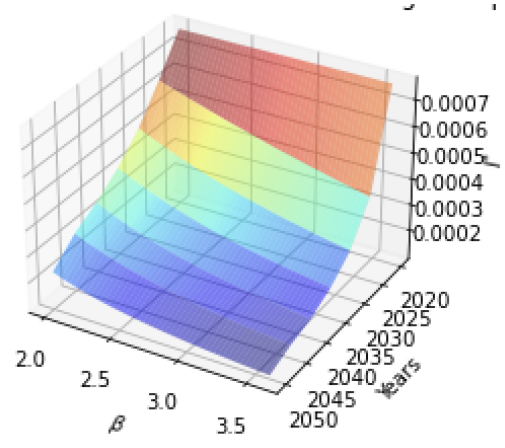


(d) $I^C(., S^*, \gamma^*)$ through years for different values of α.

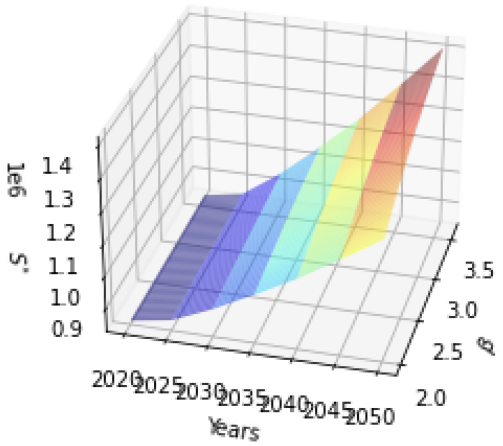
Figure 8: Sensitivity analysis of the optimal trajectories to $\alpha \in [0.95; 1]$ computed with the yellow company in scenario B2C. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs.

5.4.4 Sensitivity to c

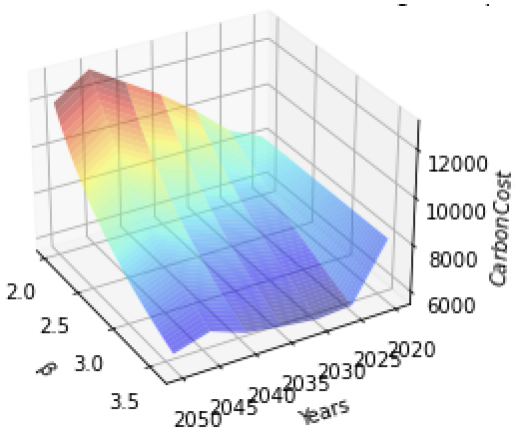
As expected, the augmentation of the unit cost of relative CO_2 reduction c as defined in (2.4) induces larger unitary abatement costs. Therefore, the optimal intensity heightens concavely with respect to c (Fig. 10a). For $c = 0.63$, the intensity starts declining in 2030 and hits its lowest levels. Naturally, this leads to convexly declining optimal sales (Fig. 10b). The surge in intensity overrules the sales constriction in the computation of the carbon tax, causing



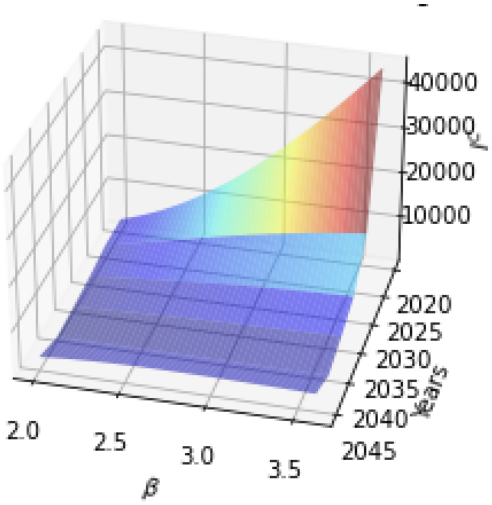
(a) I^* through years for different values of β.



(b) S^* through years for different values of β.



(c) C^C through years for different values of β.

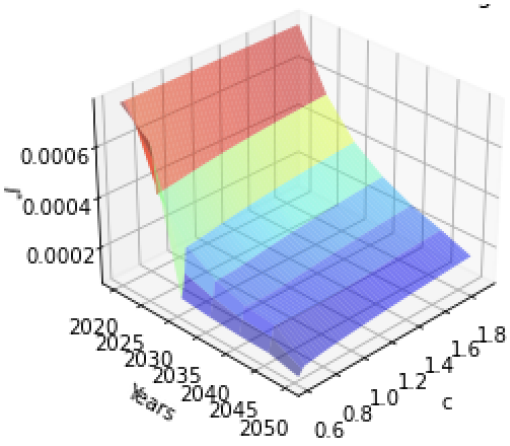


(d) $I^C(., S^*, \gamma^*)$ through years for different values of β.

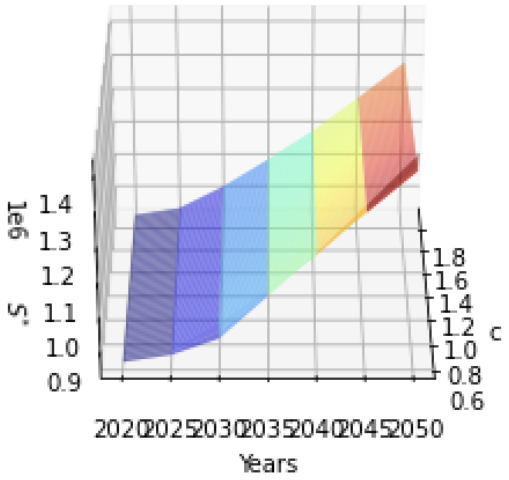
Figure 9: Sensitivity analysis of the optimal trajectories to $\beta \in [2; 3.6]$ computed with the yellow company in scenario B2C. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs.

convexly increasing carbon costs with respect to c (Fig. 10c). Similarly to α (see Section 5.4.2) changes in c reverse the curvature of the carbon costs with respect to time. Regarding the investments costs, both the timing and quantity of investment effort are impacted by c. We can see in Fig. 10d that $I^C(.)$ peaks when c is the lowest in 2025 ($\texttt{USD}29,517$). This is in line with the aforementioned results regarding the optimal intensity. The overall investments levels remain stable otherwise. Because c is growing, this actually means decaying intensity

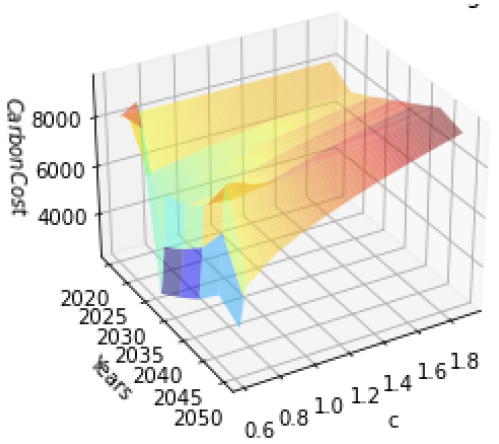
reduction efforts, which explains the concavity of I^* with respect to c. in Figure 10a. In turn, the carbon cost goes up with c (Fig. 10c) with a minimum attained by 2030 for all values of c when the augmentation of the carbon price cp_i catches up with the intensity mitigation. In a nutshell, augmenting the unit carbon mitigation cost c reduces the optimal investment efforts, thus increasing the optimal intensity and carbon costs, as well as slowing down the company's sales revenues.



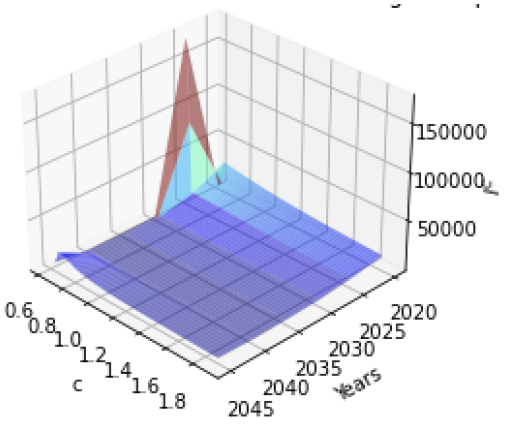
(a) I^* through years for different values of c.



(b) S^* through years for different values of c.



(c) C^C through years for different values of c.



(d) $I^C(., S^*, \gamma^*)$ through years for different values of c.

Figure 10: Sensitivity analysis of the optimal trajectories to $c \in [0.63; 1.89]$ computed with the yellow company in scenario B2C. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs.

5.4.5 Sensitivity to κ

The increase in the sensitivity of the sales revenues to the relative intensity κ (see (2.3)) leads to higher optimal intensity up until $\kappa \approx 60$. Precisely, the optimal intensity is increasing and slightly convex with respect to κ on $[0, 60]$(Fig. 11a). However, this results in concavely augmenting optimal sales on the same interval (Fig. 11b). Then, $\kappa \mapsto I^*$ becomes decreasing convex on $(60, 74]$, hits a minimum for $\kappa = 74$, then becomes increasing concave on $(74, 100]$ for all years.

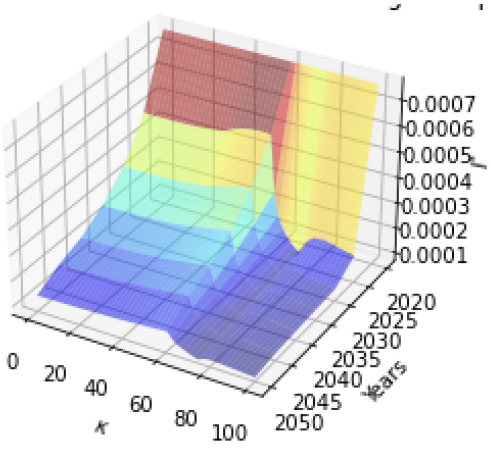
The takeaway is that, for lower values of κ raising the sensitivity to the reference market intensity does not incentivize to further decrease the company's intensity. On the contrary, increasing values of κ up to 74 lead to accelerating revenues for the same mitigation effort due to the convexity of the sales revenues with respect to the company's relative intensity (i.e. $I^{\text{ref}} - I^*$). Hence, despite higher optimal intensity, the company sales grow convexly with respect to κ because the reward is better than the penalty. Moreover, convexly increasing investments costs further disincentivize the company to have a much lower intensity than the sectoral average. Note that these results stand for an average company at starting point. Then, the paradigm shifts because increasing sales as well as intensity both result in growing carbon costs (Fig. 11c). It is more beneficial to lower the intensity, this is why investment costs peak for $\kappa > 74$ (Fig. 11d), to avoid augmented overall costs.

In conclusion, higher κ leads the company to gradually give up on mitigating its relative CO_2 emissions, in spite of lower sales revenues, to avoid convexly increasing investments costs for lower values of κ. Then, because of increasing carbon cost due to both inflated intensity and sales revenues, the company rather slash its relative emissions by largely investing in the beginning of the scenario.

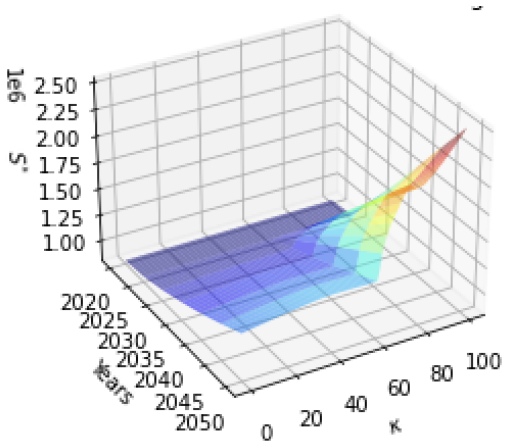
5.4.6 Sensitivity to r

The analysis of the sensitivity of the optimal trajectories with respect to the risk-free interest rate r yields interesting results. When $r = 0$, the optimal intensity stays constant for all years, meaning it is optimal to not invest at all. Then, the function $r \mapsto I_i^*$ reaches its maximum around $r = 0.4\%$ for all dates except $i = 2050$ where it is a minimum (Fig. 12a). It is then convexly decreasing for all $r > 0.4\%$ (and increasing in 2050). For small values of $r > 0$, the discount factors for costs are almost 1. Thus, there is no need for spreading the investments costs throughout the scenario. Thus, the bulk of the investment is completed when it is the less costly i.e. at the end of the scenario, due to the autonomous factor of abatement costs decrease α.

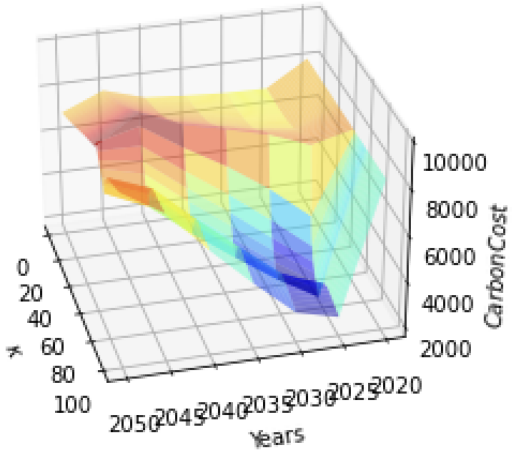
Then, the growth of the risk-free interest rate reduces the net present value of the sum of the future costs. It means that further cost will weight less and less in the decision process. Thus, as expected, the investment effort will be higher while r grows, specifically in the beginning of the scenario (Fig. 12d). Thus, the optimal intensity decreases (Fig. 12a), boosting the



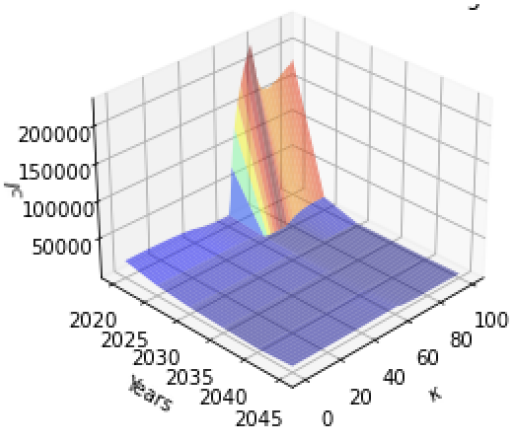
(a) I^* through years for different values of κ.



(b) S^* through years for different values of κ.



(c) C^C through years for different values of κ.

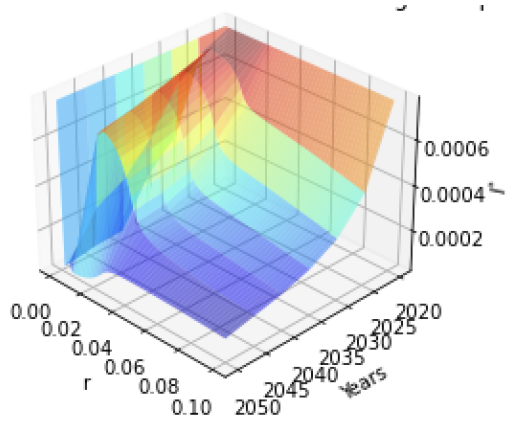


(d) $I^C(., S^*, \gamma^*)$ through years for different values of κ.

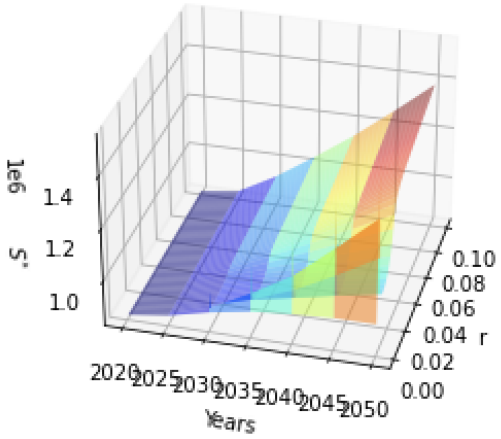
Figure 11: Sensitivity analysis of the optimal trajectories to $\kappa \in [0; 100]$ computed with the yellow company in scenario B2C. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs.

optimal sales revenues up (Fig. 12b). The drop in optimal intensity causes the carbon cost to slightly decrease despite higher sales revenues. In conclusion, on the one hand, when $r > 0$ tends to 0, the majority of the investment is planned at the end of the scenario because of the autonomous factor of cost decrease α. This causes the intensity to stay constant until it drastically falls in 2050, further leading to lower sales and lower carbon costs. On the other hand, when r grows afar from 0, future costs matter less. In turn, more and more investment is scheduled at the beginning of the scenario as r grows. In terms of optimal intensity, this

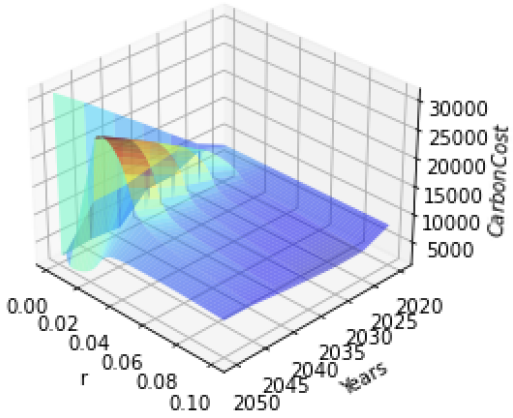
means decreasing intensity, then increasing sales, and lower carbon costs. The sales revenues and carbon cost follow the pattern of the intensity.



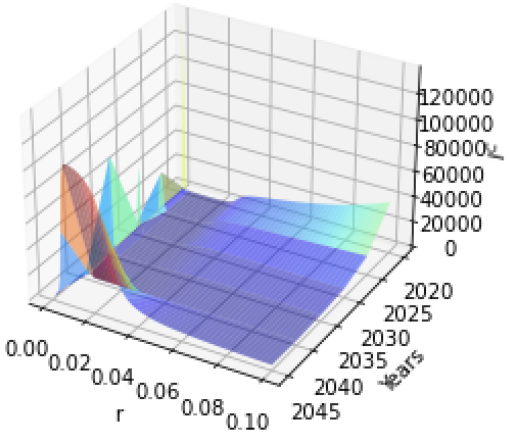
(a) I^* through years for different values of r.



(b) S^* through years for different values of r.



(c) C^C through years for different values of r.



(d) $I^C(., S^*, \gamma^*)$ through years for different values of r.

Figure 12: Sensitivity analysis of the optimal trajectories to $r \in [0; 0.1]$ computed with the yellow company in scenario B2C. From top LHS to bottom RHS : Intensity, Sales Revenues, Carbon Cost and Investment Costs.

6 Conclusion

We have proposed a model to assess how a company will adapt its business model to the energy transition in a given transition scenario. We have done so using a cost logic in a probability theory framework, to model the company's business model conditionally on a transition

scenario. We have used stochastic control to infer the best business model adaptation strategy to the energy transition, symbolized by the company's sales revenues emissions intensity, and its sales revenues. The best strategy is defined as the one that minimizes the total sum of expected discounted future carbon and mitigation investment costs over the period. This way, the trade-off between paying the carbon cost and avoiding carbon taxation is exploited. After defining the stochastic control problem, we have proven that a solution does exist in Section 3. Then, we have proposed an algorithm to solve numerically the aforementioned problem. Finally, we have employed our model on three fictitious companies in two different transition scenarios and computed benchmark deterministic carbon mitigation strategies. The key takeaways are that, in spite of a cost logic, overall intensity decrease leads to a cut in absolute emissions, without harming sales revenues. As expected, paying the full price carbon tax (i.e. with no attempt at reducing relative emissions) is more onerous than investing in carbon mitigation capabilities in absolute terms. Usual intensity reduction strategies, such as following the same decrease as the sector, or investing the same amount as the total carbon cost, are non-optimal according to our definition. They tend to overestimate the investments needed, or the cost of carbon emissions, which could lead to a misevaluation of inherent credit risk measure in credit risk stress testing models. Moreover, the high carbon emissions companies will try to catch up with their sectoral average to stay afloat, whereas low carbon companies will rest on their laurels, which implies an uneven but fair distribution of transition efforts and market shares within a sector. Sectors with high sensitivity to consumer and investor sentiment could lead carbon-inefficient companies to a drastic contraction of their sales rather than a change of production intensity, in order to cut their emissions. Finally, diminishing investment costs with time would lead to a delay of intensity reduction efforts, but they also induce higher absolute emissions cuts at the end of the scenario.

References

[ADCG+20] Thomas Allen, Stéphane Dées, Carlos Mateo Caicedo Graciano, Valérie Chouard, Laurent Clerc, Annabelle de Gaye, Antoine Devulder, Sebastien Diot, Noëmie Lisack, Fulvio Pegoraro, Marie Rabate, Romain Svartzman, and Lucas Vernet. Climate-related scenarios for financial stability assessment: An application to France. *Banque de France Working Paper,* https://ssrn.com/abstract=3653131, 774, July 2020.

[AdCPedR20] Autorité de Contrôle Prudentiel et de Résolution. Scénarios et hypothèses principales de l'exercice pilote climatique. https://acpr.banque-france.fr/sites/default/files/medias/documents/20200525_principales_hypotheses_pour_lexercice_pilote_climatique_vf.pdf, May 2020.

[ADE+21] Spyros Alogoskoufis, Nepomuk Dunz, Tina Emambakhsh, Tristan Hennig, Michiel Kaijser, Charalampos Kouratzoglou, Manuel A Muñoz, Laura Parisi, and Carmelo Salleo. ECB economy-wide climate stress test: Methodology and results. https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op281~05a7735b1c.en.pdf, 2021.

[Bar21] Barclays PLC. Corporate transition risk forecast model. https://home.barclays/content/dam/home-barclays/documents/citizenship/ESG/2021/Corporate-Transition-Forecast-Model-2021.pdf, 2021.

[BCoBS21] Basel Committee on Banking Supervision. Climate-related risk drivers and their transmission channels. https://www.bis.org/bcbs/publ/d517.pdf, April 2021.

[BDN18] Bank De Nederlandsche. An energy transition risk stress test for the financial system of the Netherlands. https://www.dnb.nl/media/pdnpdalc/201810_nr-_7_-2018-_an_energy_transition_risk_stress_test_for_the_financial_system_of_the_netherlands.pdf, 2018.

[BGL+22] Antoine Boirard, David Gayle, Theresa Löber, Laura Parisi, Clement Payerols, Edo Schets, Martina Spaggiari, Antoine Bavandi, Christoph Bertram, Matthies Darracq Paries, et al. NGFS scenarios for Central Banks and supervisors. https://www.ngfs.net/sites/default/files/medias/documents/ngfs_climate_scenarios_for_central_banks_and_supervisors_.pdf.pdf, 2022.

[BKM22] Marco Belloni, Friderike Kuik, and Luca Mingarelli. Euro area banks' sensitivity to changes in carbon price. https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp2654~9a537f810a.en.pdf, 03 2022.

[BLG20] Vincent Bouchet and Théo Le Guenedal. Credit risk sensitivity to carbon price. *SSRN Electronic Journal,* https://ssrn.com/abstract=3574486, April 2020.

[BMM+17] Stefano Battiston, Antoine Mandel, Irene Monasterolo, Franziska Schütze, and Gabriele Visentin. A climate stress-test of the financial system. *Nature Climate Change*, 7(4):283–288, 2017.

[BoE19] Bank of England. Discussion paper: the 2021 biennial exploratory scenario on the financial risks from climate change. https://www.bankofengland.co.uk/-/media/boe/files/paper/2019/the-2021-biennial-exploratory-scenario-on-the-financial-risks-from-climate-chang pdf, December 2019.

[Bor20] Eszter Boros. Risks of climate change and credit institution stress tests. *Financial and Economic Review*, 19(4):107–131, 2020.

[Bra12] Matthew Brander. Greenhouse Gases, CO2, CO2e, and carbon: What do all these terms mean? https://ecometrica.com/assets/GHGs-CO2-CO2e-and-Carbon-What-Do-These-Mean-v2.1.pdf, 2012.

[BS78] Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic Optimal Control: The discrete time case*. Elsevier, 1978.

[Car15] Mark Carney. Breaking the tragedy of the horizon: Climate change and financial stability. https://www.bankofengland.co.uk/-/media/boe/files/speech/2015/breaking-the-tragedy-of-the-horizon-climate-change-and-financial-stability.pdf?la=en&hash=7C67E785651862457D99511147C7424FF5EA0C1A, 2015.

[DL20] Antoine Devulder and Noemie Lisack. Carbon tax in a production network: Propagation and sectoral incidence. https://publications.banque-france.fr/sites/default/files/medias/documents/wp760.pdf, April 2020.

[EFK+23] Tina Emambakhsh, Maximilian Fuchs, Simon Kordel, Charalampos Kouratzoglou, Chiara Lelli, Riccardo Pizzeghello, Carmelo Salleo, and Martina Spaggiari. The road to Paris: stress testing the transition towards a net-zero economy. https://ssrn.com/abstract=4564374, 2023.

[Eur08] Europa Eurostat. NACE Rev. 2 - statistical classification of economic activities. https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF.pdf/dd5443f5-b886-40e4-920d-9df03590ff91?t=1414781457000, 2008.

[GETBM93] Michael Grubb, Jae Edmonds, Patrick Ten Brink, and Michael Morrison. The costs of limiting fossil-fuel CO2 emissions: a survey and analysis. *Annual Review of Energy and the environment*, 18(1):397–478, 1993.

[GGG22] Josselin Garnier, Jean-Baptiste Gaudemet, and Anne Gruz. The Climate Extended Risk Model (CERM). https://arxiv.org/abs/2103.03275, 2022.

[GHK+21] Martin Guth, Jannika Hesse, Csilla Königswieser, Gerald Krenn, Christian Lipp, Benjamin Neudorfer, Martin Schneider, Philipp Weiss, et al. OeNB climate risk stress test–modeling a carbon price shock for the Austrian banking sector. *Financial Stability Report*, 42:27–45, 2021.

[GKL+19] Margherita Giuzio, Dejan Krušec, Anouk Levels, Ana Sofia Melo, Katri Mikkonen, and Petya Radulova. Climate change and financial stability. *Financial Stability Review*, 1, 2019.

[HRTK19] Michael Hayne, Soline Ralite, Jakob Thomä, and Daan Koopman. Factoring transition risks into regulatory stress-tests: The case for a standardized framework for climate stress testing and measuring impact tolerance to abrupt late and sudden economic decarbonization. *ACRN Journal of Finance and Risk Perspectives*, 8(1):206–222, 2019.

[LBR09] Ausrine Lakstutiene, Aiste Breiteryte, and Dalia Rumsaite. Stress testing of credit risk Lithuania Banks under simulated economical crisis environment conditions. *Engineering Economics*, 65(5), 2009.

[LH22] Jin Liang and Wenlin Huang. Optimal control model of an enterprise for single and inheriting periods of carbon emission reduction. *Mathematics and Financial Economics*, 16:89–123, 2022.

[M+13] GI McKinsey et al. Pathways to a low-carbon economy: Version 2 of the global greenhouse gas abatement cost curve. *McKinsey Company: Stockholm, Sweden,* https://www.mckinsey.com/capabilities/sustainability/our-insights/pathways-to-a-low-carbon-economy#/, September 2013.

[Mer74] Robert C Merton. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, 29(2):449–470, 1974.

[NfGtFS19] Network for Greening the Financial System. A call for action Climate change as a source of financial risk. https://www.ngfs.net/sites/default/files/medias/documents/ngfs_first_comprehensive_report_-_17042019_0.pdf, April 2019.

[Nor17] William D. Nordhaus. Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, 114(7):1518–1523, 2017.

[RBK+22] Oliver Richters, Christoph Bertram, Elmar Kriegler, Jacob Anz, Thessa Beck, David N. Bresch, Molly Charles, Leon Clarke, and Ryna Cui. NGFS climate scenario database: Technical documentation v2. 2. https://data.ece.iiasa.ac.at/ngfs-phase-2/#/downloads, 2022.

[SK14] David I Stern and Robert K Kaufmann. Anthropogenic and natural causes of climate change. *Climatic change*, 122:257–269, 2014.

[SSR+22] P.R. Shukla, J. Skea, A. Reisinger, R. Slade, R. Fradera, M. Pathak, A. Al Khourdajie, M. Belkacemi, R. van Diemen, A. Hasija, G. Lisboa, S. Luz, J. Malley, D. McCollum, S. Some, and P. Vyas. *Summary for Policymakers*. Cambridge University Press, United Kingdom, 2022.

A Additional tables

Scenarios	**Firms**	**Strategies**
Below 2 Degrees	Brown : $I_0^1 = 1.16E{-}3$	Uncontrolled
Delayed Transition	Average : $I_0^2 = 7.71E{-}4$	Exogenous 'Exo'
-	Green : $I_0^3 = 3.86E{-}4$	Myopic 'Myo'
-	-	Optimal

Table 4: Summary of the scenarios, firms and strategies studied.

	Scenario	Below 2C			Delayed Transition		
Variable	**Strategy**	**Brown**	**Yellow**	**Green**	**Brown**	**Yellow**	**Green**
C^C	Exo	44,923	32,806	18,056	143,859	111,387	64,864
	Myo	36,110	28,050	17,467	98,000	89,055	69,458
	Optimal	46,371	35,437	21,583	82,577	72,693	54,168
	Uncontrolled	71,399	55,842	33,174	625,565	539,560	351,051
I^C	Exo	17,107	18,789	20,698	138,622	164,253	194,834
	Myo	35,007	26,948	16,505	94,721	85,225	65,142
	Optimal	10,581	8,869	5,459	94,734	85,093	65,229
S	Exo	3,123,585	3,435,621	3,798,401	3,155,879	3,482,481	3,863,772
	Myo	3,231,128	3,490,649	3,804,383	3,304,202	3,569,804	3,889,211
	Optimal	3,094,170	3,400,693	3,761,548	3,255,305	3,592,938	3,915,797
	Uncontrolled	2,885,709	3,237,470	3,674,218	,2945,044	3,305,190	3,751,660

Table 5: Total discounted values of main variables (Sales Revenues S, Carbon Cost C^C and Investment Costs I^C) over both scenarios B2C and DT for all strategies and all companies.

	Below 2C			Delayed Transition		
Strategy	**Brown**	**Yellow**	**Green**	**Brown**	**Yellow**	**Green**
Exo	64,886	52,337	36,196	213,818	191,328	156,042
Myo	74,794	58,670	37,176	203,651	187,046	148,984
Optimal	63,251	50,616	32,416	188,125	170,838	134,900
Uncontrolled	94,297	77,440	48,457	1,138,517	1,023,380	693,430

Table 6: Total sum of discounted cost ($J_\pi(0, X_0)$) for each company, each strategy for both scenarios.

B Proof of technical results

B.1 Proof of Lemma 3.5

In view of (2.2), we get by induction:

$$X_j^{(1)}(i,x) = X_i^{(1)}(i,x) \times \prod_{q=i}^{j-1} e^{-\gamma_q \delta} \times e^{\sigma_I \Delta\varepsilon_q^I - \psi_q^I(\sigma_I)}. \tag{B.1}$$

Since γ_q is non-negative and $\Delta\varepsilon_q^I$ is bounded, we have $\frac{X_j^{(1)}(i,x)}{X_i^{(1)}(i,x)}(\omega) \leq K_1$, for a constant K_1 uniform in $0 \leq i \leq j \leq N$. We have proved (3.7).
We repeat the same reasoning with $X^{(2)}$, and we get by induction (from (2.3)):

$$X_j^{(2)}(i,x) = X_i^{(2)}(i,x)\frac{\bar{S}_j}{\bar{S}_i} \times \prod_{q=i}^{j-1} e^{-\kappa\left(X_q^{(1)}(i,x) - I_q^{\text{ref}}\right)\delta} \times e^{\sigma_S \Delta\varepsilon_q^S - \psi_q^S(\sigma_S)}.$$

Using that $(\bar{S}_i)_{i=0,\dots,N}$ and $(I_i^{\text{ref}})_{i=0,\dots,N}$ are deterministic bounded functions, that $X_q^{(1)}(i,x) \geq 0$ for any i,q,x with $\kappa \geq 0$, that $\Delta\varepsilon_q^S$ is bounded, we easily obtain (3.8). □

B.2 Proof of Lemma 3.6

1. First, let us prove (3.9). From (B.1), $X_j^{(1)}(i,x)$ is linear in $x^{(1)}$ for any $j = i+1 : N$. A direct application of Lemma 3.5 gives the desired result with $K_3 = K_1$.
2. We have:

$$X_j^{(2)}(i,x) - X_j^{(2)}(i,y) = \left(x^{(2)}e^{-A(x)} - y^{(2)}e^{-A(y)}\right) \times \frac{\bar{S}_j}{\bar{S}_i} \times e^{\sum_{q=i}^{j-1} \kappa I_q^{\text{ref}}\delta + \sigma_S \Delta\varepsilon_q^S - \psi_q^S(\sigma_S)},$$

$$\text{with} \quad A(x) = \kappa \sum_{q=i}^{j-1} X_q^{(1)}(i,x)\delta \geq 0 \quad \text{and} \quad A(y) = \kappa \sum_{q=i}^{j-1} X_q^{(1)}(i,y)\delta \geq 0.$$

Since $(\Delta\varepsilon_i^S)_{i=0,\dots,N}$ are bounded, I^{ref} and $\bar{S}$ are deterministic, the final terms of the previous quantity are therefore bounded. Hence, we have (for a constant $k_1 \in \mathbb{R}^+$):

$$\begin{aligned} |X_j^{(2)}(i,x) - X_j^{(2)}(i,y)| \leq & k_1 |x^{(2)}e^{-A(x)} - y^{(2)}e^{-A(y)}| \\ \leq & k_1 |x^{(2)}(e^{-A(x)} - e^{-A(y)})| + k_1 |(x^{(2)} - y^{(2)})e^{-A(y)}|. \end{aligned}$$

For any $u, v \geq 0$ we have $|e^{-u} - e^{-v}| = |(u-v)\int_0^1 e^{-v - z(u-v)}\mathrm{d}z| \leq |u-v|$. Adding results from (3.9), we get:

$$|X_j^{(2)}(i,x) - X_j^{(2)}(i,y)| \leq K_4 x^{(2)} |A(x) - A(y)| + K_4 |x^{(2)} - y^{(2)}|$$

$$\leq K_4 \left[|x^{(2)}| \, |x^{(1)} - y^{(1)}| + |x^{(2)} - y^{(2)}| \right]$$

for a non-negative constant K_4 changing from line to line. Inequality (3.10) is proved.
3. Finally, let us prove (3.11):

$$\begin{aligned} &X_j^{(1)}(i,x)X_j^{(2)}(i,x) - X_j^{(1)}(i,y)X_j^{(2)}(i,y) \\ &= (X_j^{(1)}(i,x) - X_j^{(1)}(i,y))X_j^{(2)}(i,x) + (X_j^{(2)}(i,x) - X_j^{(2)}(i,y))X_j^{(1)}(i,y). \end{aligned}$$

We know from Lemma 3.5, Inequalities (3.9) and (3.10) that:

$$\begin{aligned} &|X_j^{(1)}(i,x) - X_j^{(1)}(i,y)|X_j^{(2)}(i,x) \leq K_2K_3x^{(2)}|x^{(1)} - y^{(1)}|, \\ &|X_j^{(2)}(i,x) - X_j^{(2)}(i,y)|X_j^{(1)}(i,y) \leq K_1K_4y^{(1)}[(|x^{(2)}| + |y^{(2)}|)|x^{(1)} - y^{(1)}| + |x^{(2)} - y^{(2)}|]. \end{aligned}$$

Thus, (3.11) is proven with $K_5 = (K_2K_3) \vee (K_1K_4)$. □

B.3 Proof of Lemma 3.7

Recall from (3.1) that $J_{\pi_i}(i,x) = \mathbb{E}\left[\sum_{j=i}^{N-1} \frac{\mathcal{C}_j(X_j(i,x),\gamma_j)}{(1+r\delta)^{j-i}} + \frac{\mathcal{C}_N(X_N(i,x))}{(1+r\delta)^{N-i}}\right]$. From (2.4)-(2.7), we have for $i = 0, \ldots, N-1$:

$$\mathcal{C}_i(X,\gamma) = \mathtt{cp}_i X^{(1)}X^{(2)} + X^{(2)}c\alpha^{i\delta}\frac{(1-e^{-\gamma\delta})^\beta}{\beta},$$

and $\mathcal{C}_N$ is a locally Lipschitz function (see the condition (2.5)). Given that $\mathtt{cp}_i$ is deterministic and that $\alpha^{i\delta}\frac{(1-e^{-\gamma\delta})^\beta}{\beta} \in [0,1]$, owing to Lemma 3.6 we have for any $j \geq i$, any $\gamma \in [0, \gamma_{\max}]$ and $(x,y) \in (\mathbb{R}^+)^2 \times (\mathbb{R}^+)^2$:

$$|\mathcal{C}_j(X_j(i,x),\gamma) - \mathcal{C}_j(X_j(i,y),\gamma)| \leq K_6|x-y|[1 + (|x| + |y|)(1 + |x| + |y|)]$$

for a non-negative constant K_6. Thus, for any strategy $\pi_i = (\gamma_i, \ldots, \gamma_{N-1}) \in \Pi_i$, we have:

$$\sum_{j=i}^{N} \left| \frac{\mathcal{C}_j(X_j(i,x),\gamma_j) - \mathcal{C}_j(X_j(i,y),\gamma_j)}{(1+r\delta)^{j-i}} \right| \leq K_6|x-y|[1 + (|x| + |y|)(1 + |x| + |y|)],$$

with a new constant K_6. By taking the expectation and using the triangular inequality, we obtain:

$$\begin{aligned} |J_{\pi_i}(i,x) - J_{\pi_i}(i,y)| &\leq \mathbb{E}\left[\sum_{j=i}^{N} \left| \frac{\mathcal{C}_j(X_j(i,x),\gamma_j) - \mathcal{C}_j(X_j(i,y),\gamma_j)}{(1+r\delta)^{j-i}} \right| \right] \\ &\leq K_6|x-y|[1 + (|x| + |y|)(1 + |x| + |y|)]. \end{aligned}$$

□