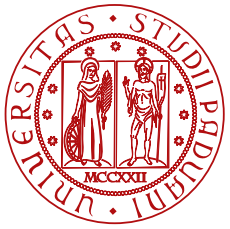


Università degli Studi di Padova

DIPARTIMENTO MATEMATICA
DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Dottorato in Scienze Matematiche
Curriculum Matematica
XXXVI ciclo

Multi-stage optimal control problems with stochastic switching times

Coordinatore

Prof. Giovanni Colombo

Supervisore

Prof. Alessandra Buratto

Co-supervisore

Prof. Stefan Wrzaczek

Dottoranda

Maddalena Muttoni

Anno Accademico
2022/2023

Abstract

This thesis develops a theory for two-stage optimal control problems with a stochastic switching time and applies it to socioeconomic models. A regime switch, occurring at a random time, divides the problem into two stages. The switch timing depends on a hazard rate function, possibly dependent on endogenous variables.

Two solution techniques are explored: a backward approach, solving the second stage first, and a vintage-structure approach, solving both stages simultaneously. Featured applications include optimizing lockdown and vaccination strategies in health economics, managing savings and emissions for climate change, and planning marketing and production strategies amid potential disruptions.

The research provides insights into decision-making under uncertainty, demonstrating in the practical applications of the developed theoretical framework the importance of an anticipative behaviour rather than a myopic one.

Sommario

Questa tesi sviluppa una teoria per problemi di controllo ottimo in due periodi con un istante di transizione stocastico e ne propone alcune applicazioni in ambito socioeconomico. Un cambio di regime, che si verifica in un momento casuale, divide il problema in due periodi. L'istante di transizione dipende da una funzione di tasso di rischio (*hazard rate*), che può dipendere da variabili endogene.

Sono esplorate due tecniche di soluzione: un approccio a ritroso (*backward*), in cui viene risolto prima il secondo periodo, e un recente approccio con struttura vintage, in cui vengono risolti entrambi i periodi simultaneamente. Le applicazioni trattate includono l'ottimizzazione delle strategie di lockdown e vaccinazione in economia sanitaria, la gestione degli investimenti e delle emissioni per il cambiamento climatico e la pianificazione delle strategie di marketing e produzione in caso di potenziali eventi che provochino cambiamenti nel sistema produttivo.

La ricerca fornisce approfondimenti sul processo decisionale in condizioni di incertezza, evidenziando nei vari contesti applicativi l'importanza di un approccio al problema anticipativo anziché miope.

Contents

Introduction

This thesis covers the theory of two-stage optimal control problems with a stochastic switching time, as well as its application to models related to various contexts. Specifically, we incorporate a regime switch in an optimal control framework.

Regime switches (see Haunschmied et al. [60]), also referred to as *regime shifts*, are instantaneous changes in a model's dynamics with possible change of the state variables as well as of the objective function, which may be externally or internally driven (exogenous or endogenous, respectively).

We consider a problem where a single regime switch is expected to occur at a random time, dividing the problem into two stages, and such that the dynamics and objective function before and after the switch are deterministic. The distribution of the switching time is governed by the *hazard rate* function, which – at every time – describes the probability of switching in the next infinitesimal interval: the higher the hazard rate, the sooner the switch is expected. If the hazard rate function depends on the state and/or the control, then the regime switch is considered endogenous (although not fully controllable due to its stochastic nature), otherwise it is exogenous.

Assuming the point of view of a decision-maker who is aware of the impending regime switch, of its effects, and of its hazard rate function, we aim to maximize the expectation of the total payoff over all possible occurrences of the switching time. Moreover, acknowledging the strength of these assumptions, we also aim to evaluate the performance of planners who do not have access to the same level of information. This type of problem can be solved with two different techniques: a *backward approach*, which requires solving the second stage first and including its optimal value as a scrap value for stage one, and a *vintage-structure approach*, which involves an organic simultaneous representation of both stages. In the first case, Dynamic Programming is employed to

compute the value function of the second stage; in the latter, a heterogeneous version of Pontryagin's Maximum Principle provides necessary conditions for the optimal solution. The convenience of the second method is also evident from the standpoint of application: the solution algorithm that relies on the vintage-structure PMP ensures the numerical tractability of high-dimensional, non-autonomous, complex problems.

The general framework of two-stage optimal control problems with a stochastic switching time is applied to models in various fields in economics, ranging from health economics to climate change economics, to marketing and production management. The purpose of this work is to characterize the optimal solutions for these types of problems (assuming a solution exists), by exploiting the peculiarities of the solution technique that is most suitable for the specific problem being considered.

This thesis is divided into four chapters: the first one presents the theoretical framework along with the two aforementioned solving techniques. The other three chapters are devoted to the application of the theoretical approach to different contexts, specifically: Health Economics, Climate Change Economics, and Marketing & Production, respectively. At the beginning of every applied work, we provide a brief overview of the theoretical setting of the model and list the specific effects of the switching time, as outlined in the classification defined in the first theoretical chapter.

Chapter 1 deals with the theoretical aspects of optimal control with a stochastic switching time. First, it explains how to turn a stochastic optimization problem into a deterministic optimal control problem at the cost of introducing auxiliary state variables. Then it describes and correlates the two solution methods, namely the backward approach and the vintage-structure (or heterogeneous) approach, evaluating the benefits and drawbacks of both. The main difference between them is that in the backward approach the second stage is solved for every possible initial state, and then its optimal value enters the first stage as a scrap value, whereas in the heterogeneous approach the two stages are solved simultaneously – the first stage's state entering the second stage's initial conditions and the second stage's co-states entering the first stage's co-state equations.

After studying the solution for a decision-maker who is well-informed about the properties of the impending regime switch and updates the optimal strategy after the switch, this chapter proceeds to consider other types of planner. These planners may be unaware of the threat of a regime switch (*myopic*) or may lack

the flexibility to adjust their strategy to the new regime (*unresponsive*). It is worth analyzing the different problems that those types of decision-maker solve to compute their optimal strategy and comparing the respective performances in terms of their actual expected payoff.

Finally, particular emphasis is placed on the first stage's co-state equations and the insight that can be gathered from analyzing the interplay of their components.

The main contributions of this chapter lie in the theoretical findings concerning the two-stage optimal control problem with a stochastic *switching* time with a *finite* time horizon, which, to the best of our knowledge, had not been previously explored in the literature until the work of Buratto et al. [18]. Reed [92] addressed a similar finite-horizon problem using Pontryagin's Maximum Principle, however, his model involves a stochastic *stopping* time instead of a switching time. Regarding the *infinite*-horizon scenario, employing the backward approach, one would refer to the theory of piecewise deterministic problems as outlined by Dockner et al. [33, ch.8]. However, it is worth noting that, in that theory, the hazard rate remains independent of the control variable. While Boukas et al. [11] shows that the Maximum Principle holds for infinite-horizon optimal control problems with a random *stopping* time, formal theoretical results on the vintage-structure Maximum Principle for infinite-horizon problems with a stochastic *switching* time have yet to be published. Nevertheless, after claiming that its validity carries over (due to the particular structure of the problem) it has been employed by Wrzaczek et al. [118], Kuhn and Wrzaczek [70], and Buratto et al. [20].

Most of this chapter's results are based on the article

- A. Buratto, L. Grosset, M. Muttoni. *Two different solution techniques for an optimal control problem with a stochastic switching time.* WSEAS Transactions on Mathematics (2023) [18]

to which the author of this thesis, M. Muttoni contributed with: conceptualization, methodology, formal analysis, and writing the original draft.

Chapter 2 is dedicated to the field of health economics, and it is divided in two sections.

The first one presents an SIRV model, in line with the epidemiological compartment models such as Kermack and McKendrick [67], where the policymaker contrasts the spread of the disease with the aid of two tools: lockdown and vaccination. In the initial stage, lockdown measures are complemented by research

investments aimed toward the discovery of an effective vaccine. Due to the unpredictability of the vaccine's discovery, its timing is modeled as a positive random variable with endogenous hazard rate: The greater the intensity of the research effort, the more likely a breakthrough is to occur sooner. The subsequent implementation of a systematic vaccination campaign (at an exogenous rate) marks the beginning of the second stage. Throughout the planning horizon, the policymaker balances the tradeoff between the costs associated with the infected individuals (in terms of health care expenditures and deaths) and the costs entailed by the lockdown measures (in terms of economic loss). The main findings are that

- a planner who is anticipating the discovery of a vaccine will implement a stricter lockdown in the first stage, compared to a myopic planner who is not taking this possibility into consideration;
- upon the vaccine's discovery, the optimal lockdown will be relaxed or intensified depending on the vaccination rate: if it progresses rapidly, the lockdown intensity decreases; if it progresses slowly, the lockdown intensity spikes (only to decrease faster thanks to the resolving epidemic).

The results are based on the article

- A. Buratto, M. Muttoni, S. Wrzaczek, M. Freiberger. *Should the COVID-19 lockdown be relaxed or intensified in case a vaccine becomes available?* PLOS ONE (2022) [20]

to which the author of this thesis, M. Muttoni contributed with: conceptualization, methodology, software, and formal analysis.

The second part of Chapter 2, unlike the rest of the thesis, presents a differential game without regime switches, where the healthcare system and a pharmaceutical firm implement pro-vaccine communication campaigns to contrast anti-vax action. We compute the Markovian Nash equilibrium strategies and perform a sensitivity analysis of the strategies and steady state with respect to the parameters. Despite the deviation from the primary theoretical framework of my thesis, the inclusion of this work in the chapter concerning health economics applications serves to broaden the interdisciplinary perspective, offering valuable insights into the broader context of strategic decision-making processes in health care.

The main findings are that, at the Markovian Nash equilibrium,

- the healthcare system's investment in vaccine communication increases with the strength of anti-vax word of mouth;

- the firms pro-vaccine communication campaign is not affected by the anti-vax word of mouth;
- upon variation of most parameters, the two players act as strategic substitutes: the smaller the healthcare system's campaign, the higher the firms one.

The results are based on the article

- A. Buratto, R. Cesaretto, M. Muttoni. *Communication strategies to contrast anti-vax action: a differential game approach.* Central European Journal of Operations Research (2024) [17]

to which the author of this thesis, M. Muttoni contributed with: methodology and formal analysis.

Chapter 3 addresses the topic of climate change economics. It features a continuous-time adaptation of the well-established DICE model by W. Nordhaus in its 2016 version, with the addition of a climate tipping point. DICE (Dynamic Integrated Climate-Economy model) is a complex model that simulates the dynamic interactions between the global economic production and the geophysical processes that drive global warming, such as the carbon cycle and the greenhouse effect. The core of the problem is that, while economic production increases consumption (from which utility is gained), it is also responsible for CO_2 emissions and the subsequent rise in global temperature. A climate tipping point is an abrupt and irreversible regime shift in the earth's geophysical dynamics. This study outlines four scenarios: two are strict climate tipping points (melting of the Arctic sea ice and transformation of a carbon sink into a carbon source), while the other two are broader disruptive events (instant destruction of a fraction of the capital due to a climate-related catastrophe and GDP's increased sensitivity to temperature). In all four cases, the hazard rate is assumed to increase with the global temperature.

The policymaker controls the savings fraction of the GDP – which is directed at the accumulation of capital while the rest constitutes the consumption – and the abatement, i.e., the curbing of a fraction of carbon emissions. The study provides the analytical conditions for the optimal savings and emission abatement policy, and a decomposition of the growth rate of the social cost of carbon into terms deriving from the capital dynamics, terms deriving from the carbon dynamics, and terms deriving from the anticipation of a tipping point.

The main findings are

- the importance of prevention: compared to a myopic abatement strategy, in all four tipping scenarios some degree of preventive abatement is employed to fend off the tipping point hazard;
- different tipping point scenarios entail different effects on the social cost of carbon, both in anticipation and in adaptation to the switch; these effects can be better understood through the decomposition of the SCC, from both the analytical and the numerical perspective.

The results are based on the monograph

- M. Muttoni. *How to prepare for and adapt to a climate tipping point*. IIASA YSSP report (2023) [83]

Chapter 4 is devoted to marketing and production applications, and it is divided in two sections.

The first one features an analysis of the implications of myopia in a dynamic marketing problem based on the goodwill model developed by Nerlove and Arrow [85]. Assuming constant price, the demand for the product is an increasing function of the goodwill, i.e., the perceived desirability of the product among the consumers; hence, the purpose of the marketing campaign is to increase the sales by increasing the goodwill. Adding to the original tradeoff between the profit from higher sales and the advertising costs, the model integrates the risk that a high demand may trigger a sudden rise in production costs (e.g., due to a shortage in raw materials, supply, or labor). We compare the marketing strategies and the resulting payoffs of three planners, one of whom is aware of the impending risk and plans the strategy accordingly, while the other two, oblivious to such a risk, act myopically. Of the latter two, one reacts to the switch adapting to the new regime, whereas the other persists with the originally determined strategy.

The main findings are that

- the marketing intensity is lower in anticipation, in order to both fend off the risk of switching and have a lower goodwill at the time the production costs increase;
- the marketing intensity in the second period is lower than in the first period: this is due to the higher production costs, which lower the profit margin;
- the anticipating planner's payoff is greater than the myopic planner's payoff, *on average* (i.e., in terms of expectation w.r.t. the switching time τ);
- however, if the switch occurs late in the time horizon, the myopic planner might be lucky and obtain a higher payoff *realization* compared to the anticipating planner. Nevertheless, this event occurs with a relatively low probability.

The results are based on the manuscript

- A. Buratto, L. Grosset, M. Muttoni, B. Viscolani. *The Cost of Myopia with Respect to a Switching Time in an Advertising Model* in "The unaffordable price of myopia in economics and management" by F. El Ouardighi and G. Feichtinger. (2024) (To appear)

to which the author of this thesis, M. Muttoni contributed with: conceptualization, methodology, software, and formal analysis.

The second part of Chapter 4 presents some preliminary results from the analysis of a firm's optimal offshoring and reshoring strategy under the exogenous threat of social and/or economic disruption. Inspired by recent events, including the COVID-19 pandemic, several war outbreaks, and other minor disruptions such as the Suez Canal blockage of 2021, the switch entails a sudden rise in the costs associated with offshoring (production, transportation, etc.). Motivated by the rising awareness about offshore labor conditions and dubious sustainability, another possible effect of the switch is the amplification of the *Made-in* effect, i.e., the consumers' negative perception of products that are made in certain countries. The *Made-in* effect in this work acts by decreasing the sales at a rate that is proportional to the offshore production. Finally, to represent the damages that the firm could face in case of, e.g., a war outbreak in the offshoring country, or the costs that would be entailed to reassess the convenience of offshoring in case of, e.g., a change in regulations, the switch also imposes an immediate penalty on the firm, assumed to increase with the extent of the firm's offshore production at the time of the disruption.

The model is very simple, analytically solvable, and produces basic results. However, the framework has significant potential for enhancement towards a more realistic representation and more insightful outcomes. For now, the main findings are that

- the anticipating firm offshores more cautiously than the myopic firm, meaning that the condition on the parameters that allows positive offshoring before the switch is stricter in the anticipative case than in the myopic one;
- if there is no threat of disruption, such as after its occurrence and in the myopic planner's belief, the optimal offshoring is either to offshore the whole production or none of it, depending on the parameters;
- in the anticipating phase, the optimal fraction of production that should be made offshore is again constant, but could be any value in $[0, 1]$ depending on the parameters.

The results are based on the manuscript

- A. Buratto, M. Muttoni. *Offshoring and reshoring under social and economic uncertainty.* EURO (2024) (In progress)

to which the author of this thesis, M. Muttoni contributed with: conceptualization, methodology, formal analysis, and writing the original draft.

1

Theory

When planning an optimal policy, a farsighted decision-maker should account for the possibile occurrence of disruptive events over the course of the time horizon. For example, when planning the optimal emission abatement policy, account for a possible climate catastrophe; when planning industrial production, account for an unpredictable disruption that may affect the producers profit.

We account for this type of uncertainty by introducing a *stochastic switching time* in the optimal control framework (see, e.g., Haunschmied et al. [60]). A stochastic switching time is a random instant, modeled as a positive random variable, which marks a *regime shift* – i.e., an abrupt and irreversible change in the system – splitting the planning horizon into two stages. The shift may affect the payoff and/or the state trajectory in several ways, all of which are covered in this chapter's analysis. This study features a single switching time only.

In search for the optimal policy under this kind of uncertainty, two methods are featured in the literature: the well-established *backward* approach and the novel *vintage-structure* one. The two methods will be described and compared.

When examining such problems, we are particularly interested in how the optimal strategy adapts to the regime shift and how it differs in anticipation of it, compared to the optimal strategy arising when no shift is expected.

This framework finds numerous applications across various domains, including health economics, climate change, production, and marketing, among others. And indeed, the concept of two-stage optimal control problems with a stochastic switching time is well-established in the literature. These problems,

essentially a specific case of *piecewise deterministic* models (see Dockner et al. [33], ch.8) have been addressed using the backward approach for decades. Examples of such studies can be found in works like Tsur and Zemel [105, 106] and more recently Polasky et al. [90], Tsur and Zemel [107, 108], van der Ploeg and de Zeeuw [110, 111].

However, a novel approach has emerged recently, known as the vintage-structure approach, thanks to the contributions of Wrzaczek et al. [118], Kuhn and Wrzaczek [70], and Buratto et al. [20]. This innovative method, where the solutions are characterized through a suitable version of Pontryagin's Maximum Principle, offers a broader analytical perspective and ensures numerical feasibility even for high-dimensional models with intricate dynamics.

Two-stage problems with a random switching time are an extension of problems with a random stopping time. The latter have been studied by Yaari [119], Reed [92], and Boukas et al. [11] in the context of optimal control, and by Gromov and Gromova [53] in the context of differential games.

Remark 1.1. The aim of this chapter is to characterize the optimal solutions of the two-stage optimal control problem with a stochastic switching time, using Dynamic Programming (within the backward approach) and Pontryagin's Maximum Principle (within the vintage-structure method). The problem will be formulated under the most general hypotheses, with no concern on the conditions for its well-posedness or the existence of a solution, and stricter hypotheses will be introduced when necessary for the employment of the solution techniques. In all the applications that are investigated in the following chapters, the data satisfy all the regularity assumptions that are required for the implementation of the theoretical results.

1.1 The simple optimal control problem

An optimal control problem is a dynamic optimization problem where, at every time in a given programming interval $[0, T]$, the agent sets the value of the control variable $u(t)$ from a given control set $U \subseteq \mathbb{R}^k$. For now, the only assumption on the control function, or *strategy*, $u : [0, T] \to U$ is that it is measurable. The strategy enters the state dynamics

$$\dot{x}(t) = f\big(t, x(t), u(t)\big) \quad \text{for } t \in [0, T] \tag{1.1}$$

influencing the evolution of the state variable $x(t) \in \mathbb{R}^n$, whose initial value is a given $x_0 \in \mathbb{R}^n$. A control strategy $u : [0,T] \to U$ is *feasible* if there exists a unique solution the corresponding Cauchy problem, in the sense that there exists a unique function $x : [0,T] \to \mathbb{R}^n$ such that, for all $t \in [0,T]$,

$$x(t) = x_0 + \int_0^t f\big(\theta, x(\theta), u(\theta)\big)\, d\theta. \tag{1.2}$$

One can ensure control feasibility by imposing conditions which guarantee that the composition $(t,x) \mapsto f\big(t,x,u(t)\big)$ satisfies the hypotheses for the existence and uniqueness of (1.2). Such conditions, for which we refer to Hartman [59], consist of regularity assumptions on f and on the set of admissible control functions $u(\cdot)$. The integral above is well defined if and only if the composite function $t \mapsto f\big(t,x(t),u(t)\big)$ belongs to $L^1([0,T])$; this yields that the state *trajectory* $x : [0,T] \to \mathbb{R}^n$ is an absolutely continuous function.

Any pair (u,x), where $u(\cdot)$ is a feasible control strategy and $x(\cdot)$ is the corresponding state trajectory, is called a *process* (see Grass et al. [52]).

The planner's objective is to maximize the payoff

$$\int_0^T g\big(t,x(t),u(t)\big)\, dt + S\big(x(T)\big)$$

which is the sum of an inter-temporal term and a salvage value. The first is the integral of the profit flow g over time, which depends on the control strategy and on the corresponding state trajectory; the latter is a lump sum S which depends on the final state $x(T)$.

In absence of constraints on the final state $x(T)$, the planner solves the following problem:

$$\underset{u(t)\in U}{\text{maximize}} \left[\int_0^T g\big(t,x(t),u(t)\big)\, dt + S\big(x(T)\big) \right]$$

subject to:

$$\begin{cases} \dot{x}(t) = f\big(t,x(t),u(t)\big) & \text{for } t \in [0,T] \\ x(0) = x_0 \end{cases}$$

where:

- $g(t,x,u)$ is the running utility / profit flow function;

- $S(x)$ is the salvage value function;
- $f(t, x, u)$ is the state dynamics.

There are two solution approaches to this problem: *Dynamic Programming* and *Pontryagin's Maximum Principle (PMP)*.

Dynamic Programming relies on the general mathematical principle of embedding the original problem in a large class of problems, each starting at $t \in [0, T]$ from the initial state y. The corresponding "value function" can be defined as follows:

$$V(t, y) := \sup_{u(\theta) \in U} \left[\int_t^T g\big(\theta, x(\theta), u(\theta)\big)\, d\theta + S\big(x(T)\big) \right]$$

subject to:

$$\begin{cases} \dot{x}(\theta) = f\big(\theta, x(\theta), u(\theta)\big) & \text{for } \theta \in [t, T] \\ x(t) = y \end{cases}$$

Of course, if an optimal control exists, the sup is actually a max, as it is attained by the optimal process.

If the value function is continuously differentiable, it is the classical solution of the Hamilton-Jacobi-Bellman (HJB) equation:

$$-\partial_t V(t, x) = \max_{u \in U} \big\{ g(t, x, u) + \nabla_x V(t, x) \cdot f(t, x, u) \big\}. \tag{1.3}$$

and, regardless of its regularity, it satisfies the terminal condition:

$$V(T, x) = S(x). \tag{1.4}$$

The Verification Theorem, or sufficiency theorem (see Bardi and Capuzzo-Dolcetta [7, ch.1] or Dockner et al. [33, ch.3]), states that

- if $W(t, x)$ is a classical solution of the HJB equation (1.3) that satisfies the terminal condition (1.4), and
- if $u(t)$ is a feasible control, with associated state trajectory $x(t)$, such that $u(t) \in \Phi\big(t, x(t)\big)$ where

$$\Phi(t, x) = \arg\max_{u \in U} \big\{ g(t, x, u) + \nabla_x W(t, x) \cdot f(t, x, u) \big\};$$

then $W(t, x)$ is the value function, and the process (u, x) is optimal.

$\Phi(t, x)$ is called a *feedback* strategy, as it depends on the state of the system.

Remark 1.2. In general, the value function V is not necessarily differentiable everywhere. At the non-differentiable points, V is a *viscosity solution* of the HJB (see [7]). However, as in all the applications featured in this thesis we encounter smooth value functions, we restrict our discussion to the classical solutions of the HJB equation.

Pontryagin's Maximum Principle, instead, provides necessary conditions for the maxima of the problem. Let us assume for simplicity that $f(t,x,u)$, $g(t,x,u)$, and $S(x)$ are continuous w.r.t. all the variables, and continuously differentiable w.r.t. x. We define the Hamiltonian function as follows:

$$H(t,x,u,\lambda) = g(t,x,u) + \lambda \cdot f(t,x,u)$$

where variable $\lambda \in \mathbb{R}^n$ is the so-called adjoint, or co-state, variable. If (u,x) is an optimal process, then there exists an absolutely continuous co-state function

$$\lambda : [0,T] \to \mathbb{R}^n$$

such that, for almost every $t \in [0,T]$,

$$u(t) \in \arg\max_{v \in U} H\big(t, x(t), v, \lambda(t)\big),$$

and

$$\begin{cases} \dot{\lambda}(t) = -\nabla_x H\big(t, x(t), u(t), \lambda(t)\big) \\ \lambda(T) = \nabla_x S\big(x(T)\big). \end{cases}$$

1.2 The two-stage optimal control problem with a stochastic switching time

We are interested in integrating a stochastic switching time into the optimal control framework described in the previous paragraph.

We model the stochastic switching time τ as a random variable taking values in $[0, +\infty)$. Due to the finiteness of the time horizon, τ could occur either during the programming interval $[0,T]$, splitting it into a Stage 1 and a Stage 2, or after T, leaving the whole interval in Stage 1.

We describe the probability distribution of τ through a quantity called the

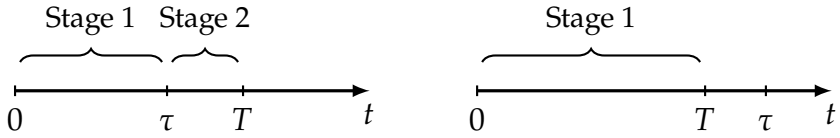


Figure 1.1: Case $\tau < T$ (left); Case $\tau \geq T$ (right)

hazard rate. We assume that there exists a hazard rate function $\eta(t, x, u)$ – possibly depending on the endogenous variables x and u – such that the hazard rate of τ at time t (LHS), for $t \in [0, T]$, is determined by evaluating the hazard rate function at time t (RHS):

$$\lim_{dt \to 0^+} \frac{\mathbb{P}\big(\tau \leq t + dt \,\big|\, \tau > t\big)}{dt} = \eta\big(t, x(t), u(t)\big). \tag{1.5}$$

It follows from the definition that $\eta \geq 0$.

Remark 1.3. Observe that, since equality (1.5) must hold, the limit in the LHS must exist and be finite. Computing the limit explicitly, we obtain the following:

$$\lim_{dt \to 0^+} \frac{1}{\mathbb{P}(\tau > t)} \cdot \frac{\mathbb{P}(\tau > t) - \mathbb{P}(\tau > t + dt)}{dt} = -\frac{1}{\mathbb{P}(\tau > t)} \frac{d}{dt} \mathbb{P}(\tau > t),$$

i.e., the tail distribution function must satisfy the ODE

$$\frac{d}{dt} \mathbb{P}(\tau > t) = -\eta\big(t, x(t), u(t)\big) \mathbb{P}(\tau > t).$$

Coupling this with the assumption that $\mathbb{P}(\tau > 0) = 1$, and solving the resulting Cauchy problem, we obtain that

$$\begin{aligned} \mathbb{P}(\tau > t) &= \exp\left(- \int_0^t \eta\big(s, x(s), u(s)\big)\, ds \right) \\ &= \mathbb{P}(\tau > T) + \int_t^T \eta\big(\theta, x(\theta), u(\theta)\big) \exp\left(- \int_0^\theta \eta\big(s, x(s), u(s)\big)\, ds \right) d\theta \end{aligned}$$

i.e., the restriction of the distribution of τ to the interval $[0, T]$ is given by the density function $f_\tau(t) = \eta\big(t, x(t), u(t)\big) e^{-\int_0^t \eta}$.

Remark 1.4. The random variable τ, representing the switching time, is assumed to take values in $[0, +\infty)$ even though the planning horizon ends at $T < +\infty$. This assumption is justified by the need to take into consideration cases where the disruptive event represented by τ, although possibly influenced by endogenous

variables during the programming interval $[0, T]$, may still occur with a certain hazard rate in the interval $(T, +\infty)$.

The hazard rate in $(T, +\infty)$ obviously cannot be instantly controlled (as the planning period is over), hence it will depend on the uncontrolled evolution in $(T, +\infty)$ of certain (unspecified) dynamic variables. We assume that the said dynamics in $(T, +\infty)$ is completely determined and known in T, possibly depending on the entire course of past events represented by the adopted strategy $u(\cdot)$. Therefore, the hazard rate of τ in $(T, +\infty)$ will be denoted as $\eta_{u(\cdot)}(t)$.

The observations and computations that we made above concerning the distribution of τ in $[0, T]$ hold for $(T, +\infty)$ as well, under the further assumption that $\eta_{u(\cdot)}$ has infinite mass, i.e.,

$$\int_T^{+\infty} \eta_{u(\cdot)}(t)\, dt = +\infty.$$

Nevertheless, while the distribution of τ in the interval $(T, +\infty)$ is needed if one is interested in computing, e.g., the ETA of τ $\mathbb{E}[\tau]$, it is irrelevant for the purpose of the solution of the optimal control problem over the programming interval $[0, T]$. For this reason, in this thesis we do not concern ourselves with the behavior of τ outside of the planning horizon.

1.2.1 Classification of possible effects of the switch

The switch may have one or more simultaneous effects on the system, hereafter listed. These changes will be recalled and discussed in the following chapters when describing each specific applied model.

(E1) Change in the control set, from $U_1 \subseteq \mathbb{R}^{k_1}$ to $U_2 \subseteq \mathbb{R}^{k_2}$;

(E2) Change in the state space, from $\mathbb{R}^{n_1}$ to $\mathbb{R}^{n_2}$;

(E3) Change in the state dynamics, from $f_1(t, x, u)$ to $f_2(\tau, t, x, u)$;

(E4) Change in the running utility, from $g_1(t, x, u)$ to $g_2(\tau, t, x, u)$;

(E5) Change in the salvage value (a lump sum at the end of the time horizon T), from $S_1\big(x(T)\big)$ to $S_2\big(\tau, x(T)\big)$;

(E6) Instantaneous switching costs $G\big(\tau, x(\tau), u(\tau)\big)$ at the switching time;

Note that $x(\tau)$ and $u(\tau)$ are the Stage 1 variables evaluated at the switching time, i.e., $x(\tau) = \lim_{t \to \tau^-} x(t) \in \mathbb{R}^{n_1}$ and $u(\tau) \in U_1$.

(E7) Jump discontinuity in the state trajectory, such that

$$x(\tau^+) = \varphi\big(\tau, x(\tau), u(\tau)\big) \in \mathbb{R}^{n_2}.$$

If the state is assumed to be continuous in τ, then $n_1 = n_2$ and $\varphi(\tau, x, u) = x$. In general, if the state space changes after the switching time, function φ describes the initial state for Stage 2 as a function of the Stage 1 state at the switch.

Be aware that the Stage 2 data f_2, g_2, S_2 may depend on the realization of τ. See table 1.1 for a schematic representation (and the respective notation) of the effects of the switch.

	Stage 1	**Switch**	**Stage 2**
Control set	$U_1 \subseteq \mathbb{R}^{k_1}$		$U_2 \subseteq \mathbb{R}^{k_2}$
State space	$\mathbb{R}^{n_1}$		$\mathbb{R}^{n_2}$
State dynamics	$f_1(t, x, u)$		$f_2(\tau, t, x, u)$
Running utility	$g_1(t, x, u)$		$g_2(\tau, t, x, u)$
Salvage value	$S_1\big(x(T)\big)$		$S_2\big(\tau, x(T)\big)$
State jump		$\varphi(\tau, x(\tau), u(\tau))$	
Switching costs		$G(\tau, x(\tau), u(\tau))$	

Table 1.1: Possible effects of τ

1.2.2 Information structure

In this context, we consider a planner who, in Stage 1, is aware that τ could occur at any time, knows the hazard rate function $\eta(t, x, u)$ defined in (1.5), and is informed about the effects that the switch will have on the system.

Because the planner cannot predict when τ will occur, they plan a Stage 1 strategy for the whole time horizon. Hence, a Stage 1 process is a pair (u_1, x_1) of functions of time, defined on the whole programming interval:

$$u_1(t) \in U_1, \quad x_1(t) \in \mathbb{R}^{n_1} \quad \text{for } t \in [0, T]$$

If the switching time occurs during the programming interval, we assume that the planner realizes that τ has occurred and when. In addition to this knowledge, we also assume that in the Stage 2 interval $[\tau, T]$ they are able to adjust the optimal strategy to the new regime, depending on that specific occurrence of τ. Therefore, we define a Stage 2 process (u_2, x_2) as a family of processes parameterized by the realization $s \in [0, T]$ of the switching time τ,

each of which is defined on the Stage 2 interval $[s, T]$:

$$u_2(s,t) \in U_2, \quad x_2(s,t) \in \mathbb{R}^{n_2} \quad \text{for } s \in [0,T],\ t \in [s,T]$$

Due to the stochasticity of τ, the planner can only aim to maximize the expectation of the total payoff, which takes different forms depending on the fact that τ occurs during the programming interval or afterwards. If $\tau > T$, then the payoff is simply

$$\int_0^T g_1\big(t, x_1(t), u_1(t)\big)\, dt + S_1\big(x_1(T)\big).$$

If otherwise $\tau \leq T$, the payoff is

$$\begin{aligned}&\int_0^\tau g_1\big(t, x_1(t), u_1(t)\big)\, dt + G\big(\tau, x_1(\tau), u_1(\tau)\big)\\ &+ \int_\tau^T g_2\big(\tau, t, x_2(\tau,t), u_2(\tau,t)\big)\, dt + S_2\big(\tau, x_2(\tau,T)\big).\end{aligned}$$

The most general formulation of the finite-time two-stage optimal control problem with a stochastic switch is the following:

$$\begin{aligned}\underset{\substack{u_1(t)\in U_1\\ u_2(s,t)\in U_2}}{\text{maximize}}\ \mathbb{E}\Bigg[&\chi_{\tau\leq T}\bigg\{\int_0^\tau g_1\big(t, x_1(t), u_1(t)\big)\, dt + G\big(\tau, x_1(\tau), u_1(\tau)\big)\\ &\qquad + \int_\tau^T g_2\big(\tau, t, x_2(\tau,t), u_2(\tau,t)\big)\, dt + S_2\big(\tau, x_2(\tau,T)\big)\bigg\}\\ &+ \chi_{\tau>T}\bigg\{\int_0^T g_1\big(t, x_1(t), u_1(t)\big)\, dt + S_1\big(x_1(T)\big)\bigg\}\Bigg]\end{aligned} \tag{1.6}$$

subject to:

$$\begin{cases}\dot{x}_1(t) = f_1\big(t, x_1(t), u_1(t)\big) & \text{for } t \in [0,T]\\ \qquad x_1(0) = x_0 & \\ \dot{x}_2(s,t) = f_2\big(s, t, x_2(s,t), u_2(s,t)\big) & \text{for } t \in [s,T]\\ \qquad x_2(s,s) = \varphi\big(s, x_1(s), u_1(s)\big) & \\ \text{Hazard rate of } \tau \text{ at time } t\text{: } \eta\big(t, x_1(t), u_1(t)\big) & \end{cases}$$

Observe the abuse of notation in the formulas above, that we will consistently

apply from now on: $\dot{x}_2(s,t) = \partial_t x_2(s,t)$.

It is worth noting that the initial condition of the Stage 2 problem is a function of the Stage 1 variables at the switch: $x_2(s,s) = \varphi\big(s, x_1(s), u_1(s)\big)$. This implies that the Stage 2 problem cannot be solved independently, *unless* one applies Dynamic Programming, where an optimal control problem is solved for every possible initial state.

Remark 1.5. In general, the Stage 2 data f_2, g_2, S_2 may also depend on the Stage 1 state variable at the switch, $x_1(s)$. With this purpose, let us introduce the auxiliary Stage 2 state variable $\tilde{x}_2$ such that $\tilde{x}_2(s,t) = x_1(s)$, i.e.,

$$\begin{cases} \partial_t \tilde{x}_2(s,t) = 0 & \text{for } t \in [s,T] \\ \tilde{x}_2(s,s) = x_1(s). \end{cases}$$

Adding $\tilde{x}_2$ to the Stage 2 state variables, and updating the Stage 2 dynamics and initial value with those of $\tilde{x}_2$, we can omit the dependence of f_2, g_2, S_2 on $x_1(s)$ without any loss of generality.

Dependence of switching costs and state jump on $u_1(\tau)$. Upon noticing that the switching costs G and state jump φ depend on $x_1(\tau)$ and $u_1(\tau)$, one might argue that, since the switch occurs at τ and the Stage 2 process starts at τ, such functions should preferably depend on the left limit of the Stage 1 process at τ, i.e., on $x_1(\tau^-)$ and $u_1(\tau^-)$. Now, because the state trajectory x_1 is continuous, $x_1(\tau^-)$ and $x_1(\tau)$ are interchangeable;[1] however, the observation sounds sensible in regard to the control (consider, e.g., a case where $u_1(\tilde{t}^-) \neq u_1(\tilde{t})$, and τ happens to occur exactly at $\tilde{t}$).

Indeed, what matters is the behavior of the control *leading up* to the switch, which in this case may be different from the value of the control *at* the switch. These functions' dependence on the left limit of the process at the switch would not only make more sense, but the reader might also worry that, if G and φ depend on $u_1(\tau)$ instead of $u_1(\tau^-)$, upon realizing the occurrence of the switch, the planner may tactically change the instantaneous value of the control $u_1(\tau)$ to optimize the combination of G and φ. Moreover, one might worry that the dependence on $u_1(\tau)$ of functions G (which is a one-time lump sum, much like a

[1]Except in the event that $\tau = 0$, in which case $x_1(\tau^-)$ does not exist, but this event by itself has probability zero, so it is irrelevant in terms of the payoff expectation.

salvage value) and φ (which determines the initial condition of the whole Stage 2 problem) may entail that two controls which differ on a zero-measure set (i.e., the singleton $\{\tau\} \subseteq [0,T]$) will generate different payoffs (an absolute aberration in standard optimal control theory).

Let us address these concerns:

1. The reason why we cannot use $u_1(\tau^-)$ is that the only regularity assumption on the control is measurability, which does not guarantee the existence of the limit $u(t^-)$ for every t. Because of this, $u_1(\tau^-)$ does not exist in general, so it would be quite problematic to involve it in the problem.

   The following are reasons why it is actually not wrong to use $u_1(\tau)$:

2. The Stage 1 control strategy $u_1(\cdot)$ is determined prior to the programming interval (hence before τ occurs), with the objective of maximizing the payoff expectation over all possible occurrences of the switching time. It is therefore independent of the realization of τ, which means that it cannot be "tactically changed" according to τ (that is a prerogative of the Stage 2 strategy $u_2(\tau, t)$).

3. In the aforementioned case where $u_1(\tilde{t}^-) \neq u_1(\tilde{t}^+)$, the event $\tau = \tilde{t}$ has probability zero, hence its occurrence is irrelevant when computing the expectation (this holds for any finite number of jump discontinuities).

4. A generalization of the previous point, which also addresses the last concern, is that two controls that differ on a zero-measure set – although leading to two different payoff *realizations* if τ occurs in that set – will lead to the same payoff *expectation*, which is the actual objective. Once again, this is due to the fact that the probability of τ occurring in a zero-measure set is zero.

The reader might find the measurability assumption on u_1 to be too loose, it being coupled with the state jump condition $x_2(s,s) = \varphi\big(s, x_1(s), u_1(s)\big)$: there may be doubts about the well-posedness of the condition and the regularity of x_2 with respect to s. First, let me stress that u_1 is assumed to be a measurable function from $[0,T]$ to $U_1 \subseteq \mathbb{R}^k$. As such, it is not intended as an equivalence class of functions: $u_1(t)$ is well defined for every t, and consequently so is $x_2(t,t)$. Of course this can lead to $x_2(s,t)$ being quite irregular with respect to s. However, due to the fact that the Stage 2 problem is parametrized by s and that there is no interaction between timelines (characteristic lines) originating from different instances of τ, the regularity of x_2 with respect to s is not a concern.

1.2.3 Reformulation

We compute the expectation with the aid of the auxiliary Stage 1 state variable $z_1(t) := \mathbb{P}(\tau > t)$, which is the probability of still being in Stage 1 at time t. To

view it as a state variable, we write its dynamics and initial value:

$$\begin{cases} \dot{z}_1(t) = -\eta\big(t, x(t), u(t)\big) z_1(t) \\ z_1(0) = 1 \end{cases}$$

where the dynamics is derived from the definition of hazard rate (1.5). Then, the probability density of τ at time t is:

$$f_\tau(t) = -\dot{z}_1(t) = \eta\big(t, x(t), u(t)\big) z_1(t). \tag{1.7}$$

Both the expressions $-\dot{z}_1$ and ηz_1 are employed as substitutes for f_τ when computing the expectation. After basic integral manipulations, which are reported in the Appendix to this chapter, the resulting objective is:

$$\begin{aligned} \underset{\substack{u_1(t)\in U_1 \\ u_2(s,t)\in U_2}}{\text{maximize}} \Bigg[\int_0^T z_1(t) \Big\{ & g_1\big(t, x_1(t), u_1(t)\big) + \eta\big(t, x_1(t), u_1(t)\big) G\big(t, x_1(t), u_1(t)\big) \\ & + \eta\big(t, x_1(t), u_1(t)\big) \Big[\int_t^T g_2\big(t, \theta, x_2(t,\theta), u_2(t,\theta)\big)\, d\theta + S_2\big(t, x_2(t,T)\big) \Big] \Big\}\, dt \\ & + z_1(T) S_1\big(x_1(T)\big) \Bigg] \end{aligned} \tag{1.8}$$

subject to:

$$\begin{cases} \dot{x}_1(t) = f_1\big(t, x_1(t), u_1(t)\big) & x_1(0) = x_0 \\ \dot{z}_1(t) = -\eta\big(t, x_1(t), u_1(t)\big) z_1(t) & z_1(0) = 1 \\ \dot{x}_2(s,t) = f_2\big(s, t, x_2(s,t), u_2(s,t)\big) & x_2(s,s) = \varphi\big(s, x_1(s), u_1(s)\big) \end{cases}$$

where we updated the Stage 1 dynamics with the new variable z_1.

This process of computing the expectation by introducing an auxiliary state variable is analogous to the one performed by Boukas et al. [11] and Reed [92, 93], although such variable is defined in a slightly different way.

Remark 1.6. From now on, we will omit the switching costs G, as they can be easily included by adding $+\eta(t,x,u)G(t,x,u)$ to the Stage 1 running utility $g_1(t,x,u)$.

1.3 Solution approaches

There are two possible ways of solving the two-stage optimal control problem defined above: the *backward* approach and the *vintage-structure* one.

The backward approach is based on Dynamic Programming (see Bardi and Capuzzo-Dolcetta [7], Dockner et al. [33]): it involves solving the Stage 2 problem for every possible occurrence of the switch and for every possible initial state, and then plugging the Stage 2 value function into the Stage 1 problem, which is solved as a simple optimal control problem, assuming optimal behavior in Stage 2. The two stages are solved separately (in reverse order) at the cost of computing the Stage 2 value function for every possible initial state, instead of just the value that it will take from the condition $x_2(s,s) = \varphi\big(s, x_1(s), u_1(s)\big)$.

The vintage-structure approach is based on Pontryagin's Maximum Principle (see Dockner et al. [33]): one derives necessary conditions for the solution by calculating the co-state functions for both stages, as solutions of the corresponding adjoint system, and letting the optimal strategies satisfy the resulting maximality conditions. The two stages are necessarily solved together, because x_1 and u_1 enter the initial conditions of Stage 2, and (as we will see) the Stage 2 co-states enter the Stage 1 adjoint equations.

In what follows, we will determine the optimal control of the switching time problem, applying the two approaches and comparing their results.

1.3.1 Backward approach

The backward approach consists in solving the Stage 2 problem with Dynamic Programming, for every possible realization of the switching time and for every initial state. Then, by assuming optimal behavior in Stage 2, the two-stage problem becomes a simple optimal control problem whose solution is the Stage 1 optimal process.

Stage 2 Since, in general, the Stage 2 data may depend on the realization s of the switching time τ, in Stage 2 there is a family of value functions parameterized by s:

$$V_2(s,t,y) := \sup_{u(\theta)\in U_2} \left[\int_t^T g_2\big(s,\theta,x(\theta),u(\theta)\big)\, d\theta + S_2\big(s,x(T)\big) \right]$$

subject to:

$$\begin{cases} \dot{x}(\theta) = f_2\big(s, \theta, x(\theta), u(\theta)\big) \quad \text{for } \theta \in [t, T] \\ x(t) = y \end{cases}$$

For every value of the parameter s, if $(t, x) \mapsto V_2(s, t, x)$ is continuously differentiable, then it is a classical solution of the HJB equation:

$$-\partial_t V_2(s, t, x) = \max_{u \in U_2} \big\{ g_2(s, t, x, u) + \nabla_x V_2(s, t, x) \cdot f_2(s, t, x, u) \big\} \tag{1.9}$$

and, regardless of its regularity, it satisfies the terminal condition:

$$V_2(s, T, x) = S_2(s, x). \tag{1.10}$$

The Verification Theorem holds for every fixed s:

- if $W(s, t, x)$ is a classical solution of the HJB equation (1.9) that satisfies the terminal condition (1.10), and
- if $u_2(s, t)$ is a feasible control, with associated state trajectory $x_2(s, t)$ (for a given initial state), such that $u_2(s, t) \in \Phi_2\big(s, t, x_2(s, t)\big)$ where

$$\Phi_2(s, t, x) = \arg\max_{u \in U_2} \big\{ g_2(s, t, x, u) + \nabla_x W(s, t, x) \cdot f(s, t, x, u) \big\};$$

then $W(s, t, x)$ is the value function, and the process (u_2, x_2) is optimal.

Stage 1 By Bellman's Priciple of Optimality, given a Stage 1 process (u_1, x_1) and assuming (u_2, x_2) is optimal for the Stage 2 problem with initial state $\varphi(s, x_1(s), u_1(s))$, we can write:

$$\int_t^T g_2\big(s, \theta, x_2(s, \theta), u_2(s, \theta)\big)\, d\theta + S_2\big(s, x_2(s, T)\big) = V_2\big(s, t, x_2(s, t)\big)$$

In particular, for $s = t$, we can substitute $x_2(t, t) = \varphi\big(t, x_1(t), u_1(t)\big)$, yielding:

$$V_2\big(t, t, x_2(t, t)\big) = V_2\big(t, t, \varphi(t, x_1(t), u_1(t))\big). \tag{1.11}$$

Assuming optimal behavior in Stage 2, the Stage 2 terms in (1.8) can be substituted by the Stage 2 value function:

$$\int_t^T g_2\big(t, \theta, x_2(t, \theta), u_2(t, \theta)\big)\, d\theta + S_2\big(t, x_2(t, T)\big) = V_2\big(t, t, \varphi(t, x_1(t), u_1(t))\big),$$

obtaining the following objective for Stage 1:

$$\underset{u_1(t)\in U_1}{\text{maximize}} \left[\int_0^T z_1(t)\Big\{ g_1\big(t, x_1(t), u_1(t)\big) + \eta\big(t, x_1(t), u_1(t)\big) V_2\big(t, t, \varphi(t, x_1(t), u_1(t))\big)\Big\}\, dt + z_1(T) S_1\big(x_1(T)\big) \right] \tag{1.12}$$

subject to:

$$\begin{cases} \dot{x}_1(t) = f_1\big(t, x_1(t), u_1(t)\big) \\ \quad x_1(0) = x_0 \\ \dot{z}_1(t) = -\eta\big(t, x_1(t), u_1(t)\big) z_1(t) \\ \quad z_1(0) = 1 \end{cases}$$

This is a single-stage deterministic optimal control problem that can be solved with any preferred method. Here we will tackle it again with Dynamic Programming, and we will show how the particular structure of the problem allows us to simplify the value function by removing the auxiliary state z_1.

It is worth observing that the state variable $z_1(t)$ in this problem behaves in many ways as a discount factor: indeed it is positive and decreasing, and it multiplies the running utility as well as the salvage value. In what follows, by analyzing the value function and the HJB equation, we further validate this claim (subject to limitations on the hazard rate's formulation).

The value function V associated to such a problem is a function of time and the state variables x and z (for all $z > 0$):

$$V(t, y, \omega) := \sup_{u(\theta)\in U_1} \left[\int_t^T z(\theta)\Big\{ g_1\big(\theta, x(\theta), u(\theta)\big) + \eta\big(\theta, x(\theta), u(\theta)\big) V_2\big(\theta, \theta, \varphi(\theta, x(\theta), u(\theta))\big)\Big\}\, d\theta + z(T) S_1\big(x(T)\big) \right]$$

subject to:

$$\begin{cases} \dot{x}(\theta) = f_1\big(\theta, x(\theta), u(\theta)\big) & \text{for } \theta \in [t, T] \\ \quad x(t) = y \\ \dot{z}(\theta) = -\eta\big(\theta, x(\theta), u(\theta)\big) z(\theta) \\ \quad z(t) = \omega \end{cases}$$

However, interpreting state z as a discount factor, we can decompose the

value function V as the product of z and a "current" value function that depends only on time and the original state x.

Proposition 1.7. *The value function V can be decomposed as:*

$$V(t,x,z) = zV^c(t,x)$$

where

$$V^c(t,y) = \sup_{u(\theta)\in U_1} \mathbb{E}\Bigg[\chi_{\tau\le T}\Big\{\int_0^\tau g_1\big(t,x_1(t),u_1(t)\big)\,dt + V_2\big(\tau,\tau,\varphi(\tau,x(\tau),u(\tau))\big)\Big\} + \chi_{\tau>T}\Big\{\int_0^T g_1\big(t,x_1(t),u_1(t)\big)\,dt + S_1\big(x_1(T)\big)\Big\}\;\Big|\;\tau>t\Bigg] \tag{1.13}$$

subject to:

$$\begin{cases}\dot{x}(\theta) = f_1\big(\theta,x(\theta),u(\theta)\big) & \text{for } \theta\in[t,T]\\ x(t) = y\end{cases}$$

Proof. By definition of $V(t,y,\omega)$, the following holds:

$$V(t,y,\omega)/\omega = \sup_{u(\theta)\in U_1}\Bigg[\int_t^T \zeta(\theta)\Big\{g_1\big(\theta,x(\theta),u(\theta)\big) + \eta\big(\theta,x(\theta),u(\theta)\big)V_2\big(\theta,\theta,\varphi(\theta,x(\theta),u(\theta))\big)\Big\}\,d\theta + \zeta(T)S_1\big(x(T)\big)\Bigg] \tag{1.14}$$

subject to:

$$\begin{cases}\dot{x}(\theta) = f_1\big(\theta,x(\theta),u(\theta)\big) & \text{for } \theta\in[t,T]\\ x(t) = y\\ \dot{\zeta}(\theta) = -\eta\big(\theta,x(\theta),u(\theta)\big)\zeta(\theta)\\ \zeta(t) = 1\end{cases}$$

where variable $\zeta(\theta)$ substitutes $z(\theta)$ through the relation

$$\zeta(\theta) = z(\theta)/\omega = z(\theta)/z(t).$$

Because of this relation, and recalling that $\theta > t$, $\zeta(\theta)$ is a conditional probability:

$$\zeta(\theta) = \mathbb{P}(\tau>\theta)/\mathbb{P}(\tau>t) = \mathbb{P}(\tau>\theta\;|\;\tau>t).$$

Retracing backwards the steps that brought us from problem (1.6) to formulation (1.12), while keeping in mind that the initial time is now t and that the probability measure is $\mathbb{P}(\cdot \mid \tau > t)$, we obtain that $V(t,x,z)/z$ equals the RHS in (1.13). Observing that the RHS does not depend on z, we define it as $V^c(t,x)$ obtaining the thesis. □

We have found that V^c is the optimal *expectation* of the 2-stage problem starting from t, *given that we are still in Stage 1 at time* t. From this observation, we gather that V^c carries significant meaning and can be assimilated to a value function, with the advantage that it does not depend on z. Furthermore, we observe that variable z behaves as a discount factor also for what concerns its role in the value function. We will call $V^c(t,x)$ the *current* value function.

Proposition 1.8. *If $V^c(t,x)$ is continuously differentiable, it is the classical solution of the following HJB equation:*

$$\begin{aligned} -\partial_t V^c(t,x) = \max_{u \in U_1} \Big\{ & g_1(t,x,u) + \nabla_x V^c(t,x) \cdot f_1(t,x,u) \\ & + \eta(t,x,u) \Big[V_2\big(t,t,\varphi(t,x,u)\big) - V^c(t,x) \Big] \Big\} \end{aligned} \tag{1.15}$$

Proof. Since $V^c(t,x)$ is C^1, then $V(t,x,z) = zV^c(t,x)$ is also C^1. As such, V is the classical solution of the (canonical) HJB equation

$$\begin{aligned} -\partial_t V(t,x,z) = \max_{u \in U_1} \Big\{ & z \Big[g_1(t,x,u) + \eta(t,x,u) V_2\big(t,t,\varphi(t,x,u)\big) \Big] \\ & + \nabla_x V(t,x,z) \cdot f_1(t,x,u) - \partial_z V(t,x,z) \eta(t,x,u) z \Big\} \end{aligned} \tag{1.16}$$

By substituting $V(t,x,z) = zV^c(t,x)$ and dividing both sides by z (which is positive by assumption and can therefore exit the maximum on the RHS) we obtain

$$\begin{aligned} -\partial_t V^c(t,x) = \max_{u \in U_1} \Big\{ & g_1(t,x,u) + \eta(t,x,u) V_2\big(t,t,\varphi(t,x,u)\big) \\ & + \nabla_x V^c(t,x) \cdot f_1(t,x,u) - V^c(t,x) \eta(t,x,u) \Big\} \end{aligned}$$

and the proposition is proved. □

Remark 1.9. Considering $z(t)$ as a discount factor makes the hazard rate η the corresponding discount rate. Under the assumption that η does not depend on the control variable u, its role as a discount rate is also reflected in the HJB

equation: indeed, we can take the product $\eta(t,x)V^c(t,x)$ out of the maximum and to the LHS, to obtain:

$$-\partial_t V^c(t,x) + \eta(t,x)V^c(t,x) = \max_{u\in U_1}\Big\{g_1(t,x,u) + \eta(t,x)V_2\big(t,t,\varphi(t,x,u)\big) + \nabla_x V^c(t,x)\cdot f_1(t,x,u)\Big\}.$$

Since the canonical value function V satisfies the terminal condition

$$V(T,x,z) = zS_1(x), \tag{1.17}$$

the terminal condition for V^c is straightforward from the fact that $V^c(t,x) = V(t,x,z)/z$:

$$V^c(T,x) = S_1(x). \tag{1.18}$$

The verification theorem still holds:

Theorem 1.10 (Verification Theorem)**.** *Let $W(t,x)$ be a classical solution of the HJB equation* (1.15) *and the terminal condition* (1.18)*. Let $u_1(t)$ be a feasible control with associated state trajectory $x_1(t)$, and suppose that, for almost all t, $u_1(t)\in\Phi_1\big(t,x_1(t)\big)$ where*

$$\Phi_1(t,x) = \arg\max_{u\in U_1}\Big\{g_1(t,x,u) + \nabla_x W(t,x)\cdot f_1(t,x,u) + \eta(t,x,u)\Big[V_2\big(t,t,\varphi(t,x,u)\big) - W(t,x)\Big]\Big\}.$$

Then, $W(t,x)$ is the current value function, and the process (u_1,x_1) is optimal.

Proof. Since $W(t,x)$ is a classical solution of (1.15), then, by retracing back the steps to go from (1.16) to (1.15), $zW(t,x)$ is a classical solution of (1.16). Since W satisfies the terminal condition (1.18), then, by multiplying both sides by z, zW satisfies the terminal condition (1.17).

Since $\Phi_1(t,x)$ maximizes the RHS of (1.15), then, by multiplying the RHS by z (which is positive) and substituting $z\nabla_x W = \nabla_x\big[zW\big]$ and $W = \partial_z\big[zW\big]$, we obtain that $\Phi(t,x,z) := \Phi_1(t,x)$ maximizes the RHS in (1.16).

We have proved that zW and Φ satisfy the hypotheses of the standard verification theorem. Therefore, zW is the (canonical) value function, and hence W is the current value function, and $u_1(t)\in\Phi\big(t,x_1(t),z_1(t)\big) = \Phi_1\big(t,x_1(t)\big)$ is the optimal strategy. □

1.3.2 Vintage-structure approach

In the vintage-structured approach, the optimal solution is characterized by a comprehensive set of PMP-style necessary conditions involving the variables of both stages. The two stages cannot be solved sequentially, because

- the Stage 1 process determines the Stage 2 initial condition;
- the Stage 2 co-states enter the Stage 1 adjoint equations.

To be able to apply this method, the problem formulation in (1.8) needs conversion to a vintage-structure one. Analogously to the more renowned age-structure models (see Anita [5] and Feichtinger et al. [41]), where the variables involved depend on time t and on a heterogeneous parameter a called *age*, in vintage-structure models (see Feichtinger et al. [39]) the heterogeneous parameter is the *vintage* s. As it will be explained below, age and vintage are perfectly interchangeable through a simple variable substitution, yielding that any age-structure problem has an equivalent vintage-structure counterpart and viceversa. Consider, for example, a scenario where the planner wants to implement the optimal age-dependent marketing campaign over time, each age a trivially represents the 'a'-year-old demographic; converting the same problem to a vintage structure, the vintage s represents the demographic that was born at time s. Therefore, the relationship between age and vintage is simply $s + a = t$. This relationship is also manifest in the type of model dynamics of the respective formulations: whereas in age-structured models the state dynamics are PDEs of the kind

$$\partial_a x(a,t) + \partial_t x(a,t) = f\big(a,t,x(a,t),u(a,t)\big),$$

by defining the corresponding vintage-dependent state $\tilde{x}(s,t) = x(t-s,t)$, control $\tilde{u}(s,t) = u(t-s,t)$, and dynamics $\tilde{f}(s,t,x,u) = f(t-s,t,x,u)$, the state dynamics in vintage-structure models take the following form:

$$\partial_t \tilde{x}(s,t) = \tilde{f}\big(s,t,\tilde{x}(s,t),\tilde{u}(s,t)\big).$$

Observe that this is the type of dynamics that we have in the Stage 2 part of our problem (1.6).

Both age-structure and vintage-structure models feature

- *distributed* (also called *heterogeneous*) variables, which depend on both time and age/vintage and are determined by the dynamics described above;

- *aggregated* variables, which are defined as integrals of certain distributed quantities over all ages/vintages, and as such depend only on time.

In our case the Stage 1 variables depend only on time, but they are not defined as integrals of quantities that are distributed over the vintages s (each vintage s representing a possible realization of the switching time): instead, they are determined by the usual ODE-type dynamics. This is the main reason why our problem is not purely vintage-structured, but rather an extension of this framework, where the set of the heterogeneous parameters consists of all the vintages (parameterizing the instances of Stage 2) plus a singleton representing Stage 1.

In what follows, we illustrate the procedure to reformulate (1.8) as a vintage-structure problem. The analogous steps for the infinite horizon problem can be found in Wrzaczek et al. [118]. We are going to separate the payoff into the following additive terms.

- An inter-temporal part, which is an integral over time where the running utility is the sum of two contributions:
  - Stage 1: this needs to be a function of the Stage 1 variables only;
  - Stage 2: this needs to be an integral over all instances of the switching time prior to t, where the integrand is a function of the Stage 2 variables only.

- A terminal part, which is the sum of two contributions:
  - Stage 1: this is a function of the Stage 1 final states only;
  - Stage 2: this needs to be an integral over the switching time where the integrand is a function of the Stage 2 final states only.

The first issue with formulation (1.8) is that $z_1(t)\eta(t, x_1(t), u_1(t))$ – which depends on Stage 1 variables – multiplies both the Stage 2 inter-temporal and terminal parts. We solve this by introducing the auxiliary Stage 2 state variable $z_2(s,t)$ in such a way that $z_2(s,t) = \eta(s, x_1(s), u_1(s))z_1(s)$ for all $t \geq s$:

$$\begin{cases} \dot{z}_2(s,t) = 0 \\ z_2(s,s) = \eta(s, x_1(s), u_1(s))z_1(s) \end{cases}$$

Remark 1.11. By equation (1.7) we have that, for all $t \geq s$, $z_2(s,t)$ is the probability density of τ at time s:

$$z_2(s,t) = f_\tau(s).$$

After conveniently substituting $z_2(t,\theta)$ in place of the $z_1\eta$ that multiplied the inter-temporal part, and $z_2(t,T)$ in place of the $z_1\eta$ that multiplied the terminal part, we obtain:

$$\int_0^T z_1(t)g_1\big(t,x_1(t),u_1(t)\big)\,dt + \int_0^T\int_t^T z_2(t,\theta)g_2\big(t,\theta,x_2(t,\theta),u_2(t,\theta)\big)\,d\theta\,dt$$
$$+ z_1(T)S_1\big(x_1(T)\big) + \int_0^T z_2(t,T)S_2\big(t,x_2(t,T)\big)\,dt$$

To be consistent with the Stage 1 notation, let us change the names of the integration variables for Stage 2. We denote the switching time by s, and the current time by t:

$$\int_0^T z_1(t)g_1\big(t,x_1(t),u_1(t)\big)\,dt + \int_0^T\int_s^T z_2(s,t)g_2\big(s,t,x_2(s,t),u_2(s,t)\big)\,dt\,ds$$
$$+ z_1(T)S_1\big(x_1(T)\big) + \int_0^T z_2(s,T)S_2\big(s,x_2(s,T)\big)\,ds$$

The other issue with this formulation is that, in the Stage 2 inter-temporal part, time is integrated first. Indeed, we are evaluating the Stage 2 inter-temporal payoff for each separate instance of the switching time, and then taking the expectation by integrating all these payoffs *weighted* by their probability of occurrence. Instead, we need time to be the outer integration variable, so let us exchange the order of integration (see figure 1.2 for a graphical representation of this operation). This way, for every time t we take the expectation of the Stage 2

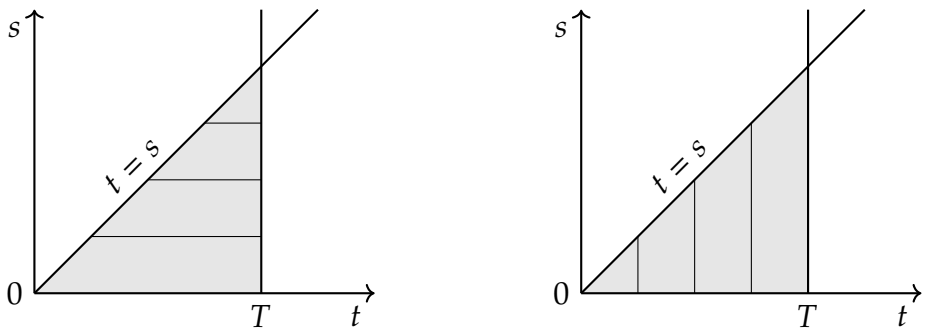


Figure 1.2: Integrating with respect to time first (left); Integrating with respect to the switching time first (right).

contribution to the running utility at t:

$$Q(t) = \int_0^t z_2(s,t) g_2\big(s,t,x_2(s,t),u_2(s,t)\big)\, ds,$$

and then we integrate that over time to obtain the cumulative Stage 2 intertemporal payoff.

The result is the following vintage-structured problem:

$$\begin{aligned}\underset{\substack{u_1(t)\in U_1\\ u_2(s,t)\in U_2}}{\text{maximize}} \Bigg[\int_0^T \Big\{ z_1(t) g_1\big(t,x_1(t),u_1(t)\big)\, dt + \int_0^t z_2(s,t) g_2\big(s,t,x_2(s,t),u_2(s,t)\big)\, ds \Big\}\, dt \\ + \Big\{ z_1(T) S_1\big(x_1(T)\big) + \int_0^T z_2(s,T) S_2\big(s,x_2(s,T)\big)\, ds \Big\} \Bigg]\end{aligned} \tag{1.19}$$

subject to:

$$\begin{cases} \dot{x}_1(t) = f_1\big(t,x_1(t),u_1(t)\big) & x_1(0) = x_0 \\ \dot{z}_1(t) = -\eta\big(t,x_1(t),u_1(t)\big) z_1(t) & z_1(0) = 1 \\ \dot{x}_2(s,t) = f_2\big(s,t,x_2(s,t),u_2(s,t)\big) & x_2(s,s) = \varphi\big(s,x_1(s),u_1(s)\big) \\ \dot{z}_2(s,t) = 0 & z_2(s,s) = \eta(s,x_1(s),u_1(s)) z_1(s) \end{cases} \tag{1.20}$$

Let us define two Hamiltonian functions, one for each stage, as usual:

$$\begin{aligned} H_1(t,x,u,\lambda_x^c) &= g_1(t,x,u) + \lambda_x^c \cdot f_1(t,x,u) \\ H_2(s,t,x,u,\xi_x^c) &= g_2(s,t,x,u) + \xi_x^c \cdot f_2(s,t,x,u) \end{aligned}$$

Theorem 1.12 (Maximum principle)**.** *Suppose that the problem's data satisfy the following assumptions (see Veliov [113]):*

- $U_1 \subseteq \mathbb{R}^{k_1}$, $U_2 \subseteq \mathbb{R}^{k_2}$ *compact.*
- ***Regularity:*** $f_1, g_1, S_1, \eta, \varphi, f_2, g_2, S_2$ *bounded and measurable in* s *and* t*, Lipschitz continuous in the control* u*, differentiable in the state* x*, and the derivatives w.r.t. the states are locally Lipschitz in* (x,u)*. All these properties hold uniformly w.r.t. the other variables.*
- ***Linear growth:*** $|f_1|, |\eta|, |\varphi| < C\big(1+|x|\big)$ *for all* (t,u)*;* $|f_2| < C(1+|x|)$ *for all* (s,t,u)*.*
- ***Admissible controls:*** *bounded and measurable functions* $u_1 : [0,T] \to U_1$ *and* $u_2 : \{(s,t) \,:\, 0 \le s \le t \le T\} \to U_2$

These assumptions imply that the problem (1.19)-(1.20) *is well-posed.*

If $\big((u_1, x_1), (u_2, x_2)\big)$ *is the optimal solution, then there exist co-state functions* $\lambda_x^c(t), \lambda_z(t)$ *for Stage 1 and* $\xi_x^c(s,t), \xi_z(s,t)$ *for Stage 2 such that the following conditions hold:*

Stage 1 First Order Condition: *Suppose that the following additional conditions hold:*

- U_1 *convex.*
- f_1, g_1, η, φ *are differentiable in the control* u *and the derivatives w.r.t. the controls are continuous in* u *uniformly in the rest of the variables.*

Then, for almost all $t \in [0,T]$*, omitting the arguments for simplicity,*

$$\Big[\nabla_u H_1 + \xi_x^c(t,t)\nabla_u\varphi\,\eta + \big[\xi_z(t,t) - \lambda_z(t,t)\big]\nabla_u\eta\Big] \cdot \big(v - u_1(t)\big) \leq 0 \quad \text{(FOC1)}$$

for all $v \in U_1$*.*

If both φ *and* η *do not depend on* u*, condition* (FOC1) *is strengthened into the following:*

Stage 1 Maximum Condition: *for almost all* $t \in [0,T]$*,*

$$u_1(t) \in \arg\max_{v \in U_1} H_1\big(t, x_1(t), v, \lambda_x^c(t)\big) \qquad \text{(MC1)}$$

This does not require the convexity and regularity assumptions w.r.t. the control that were made in the previous point, which were necessary for (FOC1)*.*

Stage 2 Maximum Condition: *for almost all* (s,t) *such that* $0 \leq s \leq t \leq T$*,*

$$u_2(s,t) \in \arg\max_{v \in U_2} H_2\big(s,t,x_2(s,t), v, \xi_x^c(s,t)\big) \qquad \text{(MC2)}$$

Stage 1 Adjoint System: *the following Adjoint Equations (AE) and Transversality Conditions (TC) hold (the arguments are omitted for simplicity):*

$$\begin{cases} \dot\lambda_x^c(t) = -\nabla_x H_1 - \big[\xi_x^c(t,t)\nabla_x\varphi - \lambda_x^c(t)\big]\eta & \\ \qquad\qquad\qquad -\big[\xi_z(t,t) - \lambda_z(t)\big]\nabla_x\eta & \lambda_x^c(T) = \nabla_x S_1\big(x_1(T)\big) \\ \dot\lambda_z(t) = -g_1 - \big[\xi_z(t,t) - \lambda_z(t)\big]\eta & \lambda_z(T) = S_1\big(x_1(T)\big) \end{cases} \qquad \text{(AS1)}$$

Stage 2 Adjoint System: *the following Adjoint Equations (AE2) and Transversality*

Conditions (TC2) hold (the arguments are omitted for simplicity):

$$\begin{cases} \dot{\xi}^c_x(s,t) = -\nabla_x H_2 & \xi^c_x(s,T) = \nabla_x S_2\big(s, x_2(s,T)\big) \\ \dot{\xi}_z(s,t) = -g_2 & \xi_z(s,T) = S_2\big(s, x_2(s,T)\big) \end{cases} \tag{AS2}$$

Remark 1.13. The Stage 1 FOC (FOC1) is a local condition with respect to u_1. The reason is that u_1 can directly influence the initial condition for x_2 (through φ) and z_2 (through η). In fact, the Stage 1 Maximum Condition (MC1), which is a global condition with respect to u_1, is false in general: it only holds if the hazard rate η and the state jump φ do not depend on u_1.

Proof. We apply the Maximum Principle for heterogeneous systems found in Veliov [113] to our vintage-structure problem (1.19). The heterogeneity, in our case, is given by the set of parameters $\{*\} \cup [0,T]$, where the unit-measure singleton $\{*\}$ represents Stage 1, and the interval $[0,T]$, endowed with the Lebesgue measure, represents all possible instances of Stage 2 (each represented by the realization s of the switching time, provided it occurs during the programming interval). We reach the following two-stage adjoint system:

$$\begin{cases} -\dot{\lambda}_x(t) = \big[\lambda_x(t)\nabla_x f_1 - \lambda_z(t)\nabla_x \eta z_1(t)\big] \\ \qquad\qquad + \big[\xi_x(t,t)\nabla_x\varphi + \xi_z(t,t)\nabla_x\eta z_1(t)\big] & \lambda_x(T) = z_1(T)\nabla_x S_1\big(x_1(T)\big) \\ \qquad\qquad + z_1(t)\nabla_x g_1 \\ -\dot{\lambda}_z(t) = -\lambda_z(t)\eta + \xi_z(t,t)\eta + g_1 & \lambda_z(T) = S_1(x_1(T)) \\ -\dot{\xi}_x(s,t) = \xi_x(s,t)\nabla_x f_2 + z_2(s,t)\nabla_x g_2 & \xi_x(s,T) = z_2(s,T)\nabla_x S_2(s, x_2(s,T)) \\ -\dot{\xi}_z(s,t) = g_2 & \xi_z(s,T) = S_2(s, x_2(s,T)) \end{cases} \tag{1.21}$$

Observe that the AE and TC for λ_z and ξ_z are already as in (AS1),(AS2).

Just like in the backward approach, we can view $z_1(t)$ as a discount factor for the Stage 1 problem. Although $z_2(s,t)$ does not behave as a discount factor (it is constant over time), it is merely a weight that multiplies each instance of the Stage 2 running utility. In light of this, it makes sense to define the "current" co-state functions:

$$\lambda^c_x(t) = \lambda_x(t)/z_1(t), \qquad \xi^c_x(s,t) = \xi_x(s,t)/z_2(s,t).$$

Let us differentiate these new co-states with respect to time:

$$\begin{aligned}\dot{\lambda}_x^c(t) &= \dot{\lambda}_x(t)/z_1(t) - \lambda_x^c(t)\dot{z}_1(t)/z_1(t)\\ &= \dot{\lambda}_x(t)/z_1(t) + \lambda_x^c(t)\eta\\ \dot{\xi}_x^c(s,t) &= \dot{\xi}_x(s,t)/z_2(s,t) - \xi_x^c(s,t)\dot{z}_2(s,t)/z_2(s,t)\\ &= \dot{\xi}_x(s,t)/z_2(s,t)\end{aligned}$$

We substitute $\dot{\lambda}_x(t)$ and $\dot{\xi}_x(s,t)$ in each equation, respectively. Upon substitution, in the adjoint equation for λ_x^c, we manipulate the term

$$\xi_x(t,t)/z_1(t) = \xi_x^c(t,t)z_2(t,t)/z_1(t) = \xi_x^c(t,t)\eta$$

to obtain the Adjoint System (AS1)-(AS2).

As for the conditions on the controls, for Stage 1 we obtain the following FOC: for almost all $t \in [0,T]$,

$$\begin{aligned}\Big\{\big[\xi_x(t,t)\nabla_u\varphi + \xi_z(t,t)\nabla_u\eta z_1(t)\big] + \nabla_u\big[z_1(t)g_1 + \lambda_x(t)\cdot f_1 - \lambda_z(t)\eta z_1(t)\big]\Big\}\\ \cdot\big(v - u_1(t)\big) \le 0\end{aligned} \tag{1.22}$$

for all $v \in U_1$. After dividing it by $z_1(t)$ and substituting $\xi_x(t,t)/z_1(t) = \xi_x^c(t,t)\eta$ like before, we obtain (FOC1). If both φ and η do not depend on u, a stronger maximum condition holds for Stage 1: for almost all $t \in [0,T]$,

$$u_1(t) \in \arg\max_{v\in U_1}\Big\{z_1(t)g_1\big(t,x_1(t),v\big) + \lambda_x(t)\cdot f_1\big(t,x_1(t),v\big) - \lambda_z(t)\eta\big(t,x_1(t)\big)z_1(t)\Big\} \tag{1.23}$$

We reach (MC1) by observing that the last term does not depend on u, and dividing the whole expression by $z_1(t)$.

In Stage 2 we obtain that, for almost all (s,t) such that $0 \le s \le t \le T$,

$$u_2(s,t) \in \arg\max_{v\in U_2}\Big\{z_2(s,t)g_2\big(s,t,x_2(s,t),v\big) + \xi_x(s,t)\cdot f_2\big(s,t,x_2(s,t),v\big)\Big\} \tag{1.24}$$

We divide it by $z_2(s,t)$ to obtain (MC2). □

Remark 1.14. With this method, we must necessarily consider all conditions simultaneously:

- In general, the Stage 2 state-costate system depends on the Stage 2 state's initial value, which in turn depends on $x_1(s)$ and $u_1(s)$;

- At the same time, the Stage 1 adjoint system depends on the Stage 2 co-state variables.

These conditions make it impossible to solve the two stages sequentially (unless the data have a very simple structure, such as a linear-state structure, in which case one can solve (AS2) and then plug the Stage 2 co-states into (AS1)).

This section constitutes a formal proof of the validity of the PMP for finite-horizon two-stage optimal control problems with a stochastic switching time. Although an infinite-time version has already been employed in some recent works, including Wrzaczek et al. [118], Kuhn and Wrzaczek [70], and Buratto et al. [20], it has not been formally proved yet. However, its validity carries over by substituting the transversality conditions with the limiting transversality conditions, by the following arguments:

- The Stage 2 timelines (characteristic lines) do not interact, and hence the limiting transversality condition holds for every Stage 2 trajectory.
- The limiting transversality condition holds for the Stage 1 trajectory, as proved by Boukas et al. [11] for the random stopping time problem.

1.3.3 Approach comparison

After presenting the two approaches, let us compare them to see the respective pros and cons, both from the computational point of view and the information that can be deduced from the respective optimality conditions.

The backward approach offers the advantage of deriving the optimal strategy in feedback form, which is convenient if the planner has access to the value of the state variable at every time. However, even if in the two-stage payoff (1.12) V_2 appears only with arguments $\big(t, t, \varphi(t, x_1(t), u_1(t))\big)$, this method still requires computing $V_2(s, t, x)$ for every triplet (s, t, x), which is not an easy task in general. Moreover, the computation of a value function such as V^c and V_2, suffers from the curse of dimensionality (both analytically and numerically): the complexity of the problem increases notably with the dimension of the state space (i.e., the number of state variables).

The vintage-structure approach, on the contrary, allows to derive the strategy only in the open-loop form. On the other hand, it allows for a compact and

unified representation of the model and the necessary conditions, where the interaction between the two stages is made explicit. As such, it enables us to gain analytical insight on the co-states (hence on the optimal solution) by disentangling the contributions that each variable and each term gives to their evolution, even when the problem can only be solved numerically. Moreover, the Maximum Principle for this type of problems is the theoretical basis of a numerical method (see [44]) that is able to withstand high degrees of complexity, such as non-autonomy and high dimensionality.

1.4 Same uncertainty, different behaviors

In this chapter, we have been examining an optimal control problem where the decision-maker is aware of an impending regime shift. They know the hazard rate function governing the switching time, and they understand the potential impacts on the system (or at least make attempts to estimate them). Upon encountering the switch, they adapt their strategy accordingly.

However, this scenario represents the ideal case. In reality, such comprehensive information and adaptability may not be available to the planner. Let us explore alternative types of decision-maker based on these characteristics.

Anticipating vs Myopic. When planning an optimal policy under the threat of a stochastic regime shift, the decision-maker may be more or less informed about the event in question: they may not know its hazard rate, they may not foresee the effects it will have on the system, or they may not expect a regime shift at all. The uninformed planner may try to estimate the unknown aspects of the regime shift and solve the corresponding two-stage problem, or they may act as if nothing were going to happen and solve the simple single-stage problem with the present data. We will call the latter a *myopic* decision-maker, as opposed to an *anticipating* one.

Remark 1.15. In the following chapters, equivalent terms will appear due to the lack of a uniform terminology at the time of the publication of the featured works. The term *anticipative* is equivalent to *anticipating* and is generally used when referring to a behavior or strategy, as opposed to a planner. The term *non-myopic* is also used to describe an anticipating planner. Additionally, *non-anticipative* is synonymous with *myopic*.

Because the anticipating planner maximizes the expected payoff, *on average* they will obtain a higher payoff compared to the myopic planner. However, for late realizations of the switching time, the myopic planner may realize a higher payoff than the anticipating planner (given that they both adapt or do not adapt the strategy after the switch). To explain this effect, let us consider the example of an undesirable regime shift. Indeed, prior to the switch, the anticipating decision-maker compromises on what would be the optimal strategy of the single-stage problem by implementing measures aimed to both fending off the risk of a shift and preparing for Stage 2. For late realizations of the switching time, Stage 1 "weighs more" than Stage 2 in terms of payoff, and it turns out that compromising on the optimal single-stage strategy to prepare for a (relatively) short Stage 2 was not worth it. However, this should not be a reason to willingly behave myopically: in fact, a myopic planner does not take action to delay the shift. This will in turn occur earlier on average, reducing the probability of "being lucky", compared to the anticipating case.

Responsive vs Unresponsive. Upon transition, the decision-maker might not always realize that a regime shift has occurred, or even if they do, they might be bound by contractual obligations to stick with their original plan. In such instances, we describe the planner as being *unresponsive* (to the shift). Conversely, if the planner acknowledges the altered circumstances and adjusts their strategy accordingly to adapt to the new regime, we consider them to be *responsive*.

Of course, assuming that they are both myopic or both anticipating, a responsive planner will always realize a higher payoff compared to an unresponsive planner.

Remark 1.16. Note that the unresponsive scenario is only admissible if the control variables and control sets coincide in both stages, i.e., if $U_1 = U_2$.

Overall, there are four possible combinations of these characteristics: in Stage 1, the planner may be myopic or anticipating; in Stage 2, they may be responsive or unresponsive. We can represent the four decision-makers in a table, as in table 1.2, and refer to them as AR, MR, AU, MU.

1.4.1 Myopic decision-maker

Because they are not expecting a regime shift, a myopic decision-maker will determine their initial strategy by solving the following single stage problem,

	Responsive	**Unresponsive**
Anticipating	AR	AU
Myopic	MR	MU

Table 1.2: Four planner types

with the Stage 1 data:

$$\underset{u_1(t)\in U_1}{\text{maximize}} \left[\int_0^T g_1\big(t, x_1(t), u_1(t)\big)\, dt + S_1\big(x_1(T)\big) \right]$$

subject to:

$$\begin{cases} \dot{x}_1(t) = f_1\big(t, x_1(t), u_1(t)\big) & \text{for } t \in [0, T] \\ x_1(0) = x_0 \end{cases}$$

When the switch occurs, say, at $\tau = s$, if the planner is responsive they will discard strategy u_1 and solve a new single-stage problem with the Stage 2 data. Let us denote the new strategy as $u_2(s,t)$ to stress its dependence on the occurrence of the switching time:

$$\underset{u_2(s,t)\in U_2}{\text{maximize}} \left[\int_s^T g_2\big(s, t, x_2(s,t), u_2(s,t)\big)\, dt + S_2\big(s, x_2(s,T)\big) \right]$$

subject to:

$$\begin{cases} \dot{x}_2(s,t) = f_2\big(s, t, x_2(s,t), u_2(s,t)\big) & \text{for } t \in [s, T] \\ x_2(s,s) = \varphi\big(s, x_1(s), u_1(s)\big) \end{cases}$$

If otherwise they are unresponsive, they will continue with the originally declared strategy $u_1(t)$. However, even in this latter case, the Stage 2 state trajectory in general will not coincide with x_1, as the state dynamics f_2 may be different from f_1 and the state jump φ may not be the identity function.

Given that $u_1(t)$ is computed as above, this is MU's expected payoff:

$$\begin{aligned} \mathbb{E}\Bigg[\chi_{\tau\le T} \Big\{ \int_0^\tau g_1\big(t, x_1(t), u_1(t)\big)\, dt + \int_\tau^T g_2\big(\tau, t, x_2(\tau,t), u_1(t)\big)\, dt + S_2\big(\tau, x_2(\tau,T)\big) \Big\} \\ + \chi_{\tau>T} \Big\{ \int_0^T g_1\big(t, x_1(t), u_1(t)\big)\, dt + S_1\big(x_1(T)\big) \Big\} \Bigg] \end{aligned} \tag{1.25}$$

subject to:

$$\begin{cases} \dot{x}_1(t) = f_1\big(t, x_1(t), u_1(t)\big) & \text{for } t \in [0,T] \\ \qquad x_1(0) = x_0 & \\ \dot{x}_2(s,t) = f_2\big(s,t,x_2(s,t), u_1(t)\big) & \text{for } t \in [s,T] \\ \qquad x_2(s,s) = \varphi\big(s, x_1(s), u_1(s)\big) & \\ \text{Hazard rate of } \tau \text{ at time } t\text{: } \eta\big(t, x_1(t), u_1(t)\big) & \end{cases}$$

Given $u_1(t)$ and $u_2(s,t)$ as computed above, the following is MR's expected payoff (here x_2 is different from the x_2 in (1.25) because it is obtained with u_2 instead of u_1):

$$\mathbb{E}\Bigg[\chi_{\tau\leq T}\Big\{\int_0^\tau g_1\big(t,x_1(t),u_1(t)\big)\,dt + \int_\tau^T g_2\big(\tau,t,x_2(\tau,t),u_2(\tau,t)\big)\,dt + S_2\big(\tau,x_2(\tau,T)\big)\Big\} \\ + \chi_{\tau>T}\Big\{\int_0^T g_1\big(t,x_1(t),u_1(t)\big)\,dt + S_1\big(x_1(T)\big)\Big\}\Bigg] \tag{1.26}$$

subject to:

$$\begin{cases} \dot{x}_1(t) = f_1\big(t, x_1(t), u_1(t)\big) & \text{for } t \in [0,T] \\ \qquad x_1(0) = x_0 & \\ \dot{x}_2(s,t) = f_2\big(s,t,x_2(s,t), u_2(s,t)\big) & \text{for } t \in [s,T] \\ \qquad x_2(s,s) = \varphi\big(s, x_1(s), u_1(s)\big) & \\ \text{Hazard rate of } \tau \text{ at time } t\text{: } \eta\big(t, x_1(t), u_1(t)\big) & \end{cases}$$

1.4.2 Anticipating decision-maker

We already know everything about the anticipating-responsive planner, because it was the focus of the whole chapter: they compute the optimal u_1 and u_2 by maximizing the expected payoff of a two-stage problem with a stochastic switching time (see (1.6)).

The anticipating-unresponsive planner, however, solves an entirely different problem. They possess all the information about the regime shift, but they also know that they will not be able to adjust the strategy once the switch occurs. So they only compute a single strategy $u_1(t)$ by maximizing the following expected

payoff and stick with it in both stages.

$$\underset{u_1(t)\in U_1}{\text{maximize}}\ \mathbb{E}\Bigg[\chi_{\tau\le T}\Big\{\int_0^{\tau} g_1\big(t,x_1(t),u_1(t)\big)\,dt+\int_{\tau}^{T} g_2\big(\tau,t,x_2(\tau,t),u_1(t)\big)\,dt+S_2\big(\tau,x_2(\tau,T)\big)\Big\}$$
$$+\chi_{\tau>T}\Big\{\int_0^{T} g_1\big(t,x_1(t),u_1(t)\big)\,dt+S_1\big(x_1(T)\big)\Big\}\Bigg] \tag{1.27}$$

subject to:

$$\begin{cases}\dot{x}_1(t)=f_1\big(t,x_1(t),u_1(t)\big) & \text{for } t\in[0,T]\\ \quad x_1(0)=x_0 & \\ \dot{x}_2(s,t)=f_2\big(s,t,x_2(s,t),u_1(t)\big) & \text{for } t\in[s,T]\\ \quad x_2(s,s)=\varphi\big(s,x_1(s),u_1(s)\big) & \\ \text{Hazard rate of } \tau \text{ at time } t\text{: } \eta\big(t,x_1(t),u_1(t)\big) & \end{cases}$$

The solution of this problem is beyond the scope of this thesis.

1.5 Switch desirability and the effect of anticipation

In this section we aim to explore the relationship between the co-states arising in the vintage-structure approach and the value functions in the backward approach, and we use it to better understand the difference between the myopic and the anticipating strategy.

By writing the Cauchy problem for ξ_z in integral form, we obtain that

$$\xi_z(s,t)=\int_t^T g_2\big(s,\theta,x_2(s,\theta),u_2(s,\theta)\big)\,d\theta+S_2\big(s,x_2(s,T)\big).$$

Observe that $\xi_z(s,t)$ is the optimal value for the Stage 2 problem starting from time t, given that the switch occurred at time s, i.e.,

$$\xi_z(s,t)=V_2\big(s,t,x_2(s,t)\big). \tag{1.28}$$

By solving the adjoint equation for λ_z, we obtain that

$$\begin{aligned}\lambda_z(t) &= \frac{1}{z_1(t)}\Big[\int_t^T z_1(\theta)[g_1\big(\theta, x_1(\theta), u_1(\theta)\big) + \eta\big(\theta, x_1(\theta), u_1(\theta)\big)\xi_z(\theta,\theta)]\,d\theta \\ &\qquad + z_1(T)S_1\big(x_1(T)\big)\Big] \\ &= \mathbb{E}\Big[\chi_{\tau<T}\Big\{\int_t^\tau g_1\big(\theta, x_1(\theta), u_1(\theta)\big)\,d\theta \\ &\qquad + \int_\tau^T g_2\big(\tau, \theta, x_2(\tau,\theta), u_2(\tau,\theta)\big)\,d\theta + S_2\big(\tau, x_2(\tau,T)\big)\Big\} \\ &\quad + \chi_{\tau\geq T}\Big\{\int_t^T g_1\big(\theta, x_1(\theta), u_1(\theta)\big)\,d\theta + S_1\big(x_1(T)\big)\Big\}\;\Big|\;\tau > t\Big]\end{aligned}$$

where we substituted $\xi_z(\theta,\theta)$ and we reversed the steps that took us from formulation (1.6) to formulation (1.12). We obtain that $\lambda_z(t)$ is the optimal expected value for the 2-stage problem starting from time t (given that we are still in Stage 1 at time t), i.e.,

$$\lambda_z(t) = V^c\big(t, x_1(t)\big). \tag{1.29}$$

Switch desirability. Let us define the *switch desirability* $\Delta V(t,x,u)$ as the expected payoff gain from switching at time t (given that we are still in Stage 1 at time t), if the Stage 1 state and control at t are equal to x and u, respectively. Mathematically:

$$\Delta V(t,x,u) := V_2\big(t,t,\varphi(t,x,u)\big) - V^c(t,x).$$

Written in terms of co-states, in light of equations (1.28) and (1.29), the switch desirability is

$$\Delta V\big(t, x_1(t), u_1(t)\big) = \xi_z(t,t) - \lambda_z(t).$$

As for the co-states corresponding to the original state variables x_1 and x_2, it is clear that

$$\xi_x^c(s,t) = \nabla_x V_2(s,t,x_2(s,t)).$$

It is indeed a well-known fact that the co-state trajectory at t equals the derivative of the value function with respect to the corresponding state variable, evaluated at time t and at the optimal state trajectory at t. In this case, s simply acts as a parameter that identifies the instance of Stage 2 problem.

It is also easy to see that

$$\lambda_x^c(t) = \nabla_x V^c\big(t, x_1(t)\big), \tag{1.30}$$

as $\lambda_x(t)$ is the co-state trajectory corresponding to the state variable x when maximizing the expected payoff by reformulating the two-stage problem as a single-stage problem as in (1.12). Therefore, by the same argument as above, $\lambda_x(t) = \nabla_x V\big(t, x_1(t), z_1(t)\big)$. We obtain (1.30) by dividing both sides by $z_1(t)$.

Effect of anticipation. This bridge between the two approaches enables us to better understand the role that anticipation plays in shaping the optimal strategy, as opposed to a myopic one.

If both φ and η do not depend on u, then the Stage 1 Maximum Condition for the myopic and the anticipating planner looks exactly the same:

$$u_1(t) \in \arg\max_{v \in U_1} H_1\big(t, x_1(t), v, \lambda_x^c(t)\big)$$

where $H_1(t, x, u, \lambda) = g_1(t, x, u) + \lambda \cdot f_1(t, x, u)$. Therefore, the difference in the optimality conditions arises from the adjoint system for λ_x^c. Nevertheless, regardless of the dependence of the data on u, in general the optimal strategy is directly linked to the co-state trajectories, so it makes sense to understand how the co-states are influenced by the anticipation of the regime shift.

The transversality condition is the same for both decision-makers:

$$\lambda_x^c(T) = \nabla_x S_1\big(x_1(T)\big).$$

The adjoint equation for the myopic planner is:

$$\dot{\lambda}_x^c(t) = -\nabla_x H_1,$$

whereas, for the anticipating planner, it is (see (AS1)):

$$\dot{\lambda}_x^c(t) = -\nabla_x H_1 - \big[\xi_x^c(t,t)\nabla_x\varphi - \lambda_x^c(t)\big]\eta - \big[\xi_z(t,t) - \lambda_z(t)\big]\nabla_x\eta.$$

This equation, as it is, allows us to disentangle the partial effects that contribute to the evolution of λ_x^c, such as the influence of each Stage 1 state on the state jump φ and on the hazard rate η. This kind of insight is a strength of the vintage-

structure approach which is not granted when using the backward approach.

However, we want to unify those extra terms in order to identify also the global effect that anticipation yields on the Stage 1 co-states. We begin by observing the following relation:

$$\begin{aligned}\nabla_x\left[\eta\Delta V\right] &= \eta\nabla_x(\Delta V) + \Delta V\nabla_x\eta \\ &= \eta\left[\nabla_x V_2\nabla_x\varphi - \nabla_x V^c\right] + \Delta V\nabla_x\eta.\end{aligned}$$

By evaluating the expression above at $(t, x_1(t), u_1(t))$, we can substitute $\nabla_x V_2$ with ξ_x^c, $\nabla_x V^c$ with λ_x^c, and ΔV with $\xi_z - \lambda_z$. We obtain:

$$\eta\left[\xi_x^c(t,t)\nabla_x\varphi - \lambda_x^c(t)\right] + \left[\xi_z(t,t) - \lambda_z(t)\right]\nabla_x\eta,$$

which are the extra terms that appear, with a negative sign, in the adjoint equation for the anticipative λ_x^c. We can then write:

$$\dot{\lambda}_x^c(t) = -\nabla_x H_1 - \nabla_x\left[\eta\Delta V\right]$$

The (rather intuitive) takeaway is that, compared to the myopic scenario, the anticipation of the switch contributes to increase or decrease λ_i^c, depending on how state i affects the switch desirability weighted by the hazard rate:

- if $x_1^{(i)}$ increases $\eta\Delta V$, then λ_i^c increases in anticipation;
- if $x_1^{(i)}$ decreases $\eta\Delta V$, then λ_i^c decreases in anticipation.

1.6 Appendix

This section features the explicit calculation of the expected payoff from formulation (1.6) to formulation (1.8). Analogous steps are reported in Wrzaczek et al. [118] for the infinite-horizon problem.

Let f_τ denote the probability density of τ. If we compute explicitly the payoff

expectation in (1.6), we obtain the following:

$$\begin{aligned}
&\int_0^T f_\tau(s)\Bigg[\int_0^s g_1\big(s, x_1(t), u_1(t)\big)\,dt + G\big(s, x_1(s), u_1(s)\big) \\
&\qquad\qquad + \int_s^T g_2\big(s, t, x_2(s,t), u_2(s,t)\big)\,dt + S_2\big(s, x_2(s,T)\big)\Bigg]\,ds \\
&+ \mathbb{P}(\tau > T)\Bigg[\int_0^T g_1\big(t, x_1(t), u_1(t)\big)\,dt + S_1\big(x_1(T)\big)\Bigg]
\end{aligned} \tag{1.31}$$

The first integral is then splitted into the Stage 1 part and the remaining part. In the Stage 1 integral, we substitute $f_\tau(s) = -\frac{d}{ds}\mathbb{P}(\tau > s)$:

$$\begin{aligned}
&\int_0^T -\frac{d}{ds}\mathbb{P}(\tau > s)\int_0^s g_1\big(t, x_1(t), u_1(t)\big)\,dt \\
&+ \int_0^T f_\tau(s)\Bigg[G\big(s, x_1(s), u_1(s)\big) + \int_s^T g_2\big(s, t, x_2(s,t), u_2(s,t)\big)\,dt + S_2\big(s, x_2(s,T)\big)\Bigg]\,ds \\
&+ \mathbb{P}(\tau > T)\Bigg[\int_0^T g_1\big(t, x_1(t), u_1(t)\big)\,dt + S_1\big(x_1(T)\big)\Bigg]
\end{aligned} \tag{1.32}$$

The first integral is integrated by parts:

$$\begin{aligned}
&\Bigg[-\mathbb{P}(\tau > s)\int_0^s g_1\big(t, x_1(t), u_1(t)\big)\,dt\Bigg]_{s=0}^{s=T} + \int_0^T \mathbb{P}(\tau > s)\, g_1\big(s, x_1(s), u_1(s)\big)\,ds \\
&+ \int_0^T f_\tau(s)\Bigg[G\big(s, x_1(s), u_1(s)\big) + \int_s^T g_2\big(s, t, x_2(s,t), u_2(s,t)\big)\,dt + S_2\big(s, x_2(s,T)\big)\Bigg]\,ds \\
&+ \mathbb{P}(\tau > T)\Bigg[\int_0^T g_1\big(t, x_1(t), u_1(t)\big)\,dt + S_1\big(x_1(T)\big)\Bigg]
\end{aligned} \tag{1.33}$$

The two terms $\pm\mathbb{P}(\tau > T)\int_0^T g_1\big(t, x_1(t), u_1(t)\big)\,dt$ cancel out, resulting in:

$$\begin{aligned}
&\int_0^T \mathbb{P}(\tau > s)\, g_1\big(s, x_1(s), u_1(s)\big)\,ds \\
&+ \int_0^T f_\tau(s)\Bigg[G\big(s, x_1(s), u_1(s)\big) + \int_s^T g_2\big(s, t, x_2(s,t), u_2(s,t)\big)\,dt + S_2\big(s, x_2(s,T)\big)\Bigg]\,ds \\
&+ \mathbb{P}(\tau > T) S_1\big(x_1(T)\big)
\end{aligned} \tag{1.34}$$

Recalling the definition of $z_1(s) = \mathbb{P}(\tau > s)$, and the fact that $f_\tau = -\dot{z}_1 = \eta z_1$ (1.7), we substitute $\mathbb{P}(\tau > s) = z_1(s)$ and $f_\tau(s) = \eta\big(s, x_1(s), u_1(s)\big)z_1(s)$. After joining

the two integrals and collecting $z_1(s)$ we obtain:

$$\int_0^T z_1(s)\Big\{ g_1\big(s, x_1(s), u_1(s)\big) \\ + \eta\big(s, x_1(s), u_1(s)\big)\Big[G\big(s, x_1(s), u_1(s)\big) + \int_s^T g_2\big(s, t, x_2(s,t), u_2(s,t)\big)\, dt + S_2\big(s, x_2(s,T)\big)\Big]\Big\}\, ds \\ + z_1(T) S_1\big(x_1(T)\big) \tag{1.35}$$

To obtain the formulation in (1.8) we rename the integration variables: s becomes t, and t becomes θ.

2

Health Economics

This second chapter is dedicated to problems related to health economics. In what follows, we consider an application of dynamic optimization methodologies to contrast the spread of diseases with the aid of lockdown and vaccination.

The content of Section 2.1 is a published article titled "Should the COVID-19 lockdown be relaxed or intensified in case a vaccine becomes available?". The article is jointly written by A. Buratto,[1] M. Muttoni,[1] S. Wrzaczek,[2] and M. Freiberger.[2] It was published in PLOS ONE in 2022.

In this paper, we analyze the long-term optimal lockdown strategy for governments to contrast the spread of the COVID-19 pandemic. Our novel approach incorporates the potential development of an effective vaccine at a stochastic time, by directing additional financial resources towards research and development in the field. The theoretical setting of this work is that of an infinite-horizon, two-stage optimal control problem with a stochastic switching time τ, reformulated with a vintage structure and addressed with Pontryagin's Maximum Principle. The hazard rate of τ (also called *switching* rate, because the shift is a desirable event in this case) is assumed to be directly controllable, in that it depends partly on time and partly on the control. Referring to the notation introduced in Classification 1.2.1, in this model the switch has the following effects:

(E1) **Controls:** the planner has two controls in Stage 1 (lockdown and research)

[1]Dipartimento di Matematica "Tullio Levi-Civita", Università degli Studi di Padova
[2]International Institute for Applied Systems Analysis (IIASA)

and one control in Stage 2 (lockdown only);

(E2) **States:** there are three compartments in Stage 1 (S-I-R) and four compartments in Stage 2 (S-I-R-V);

(E3) **Dynamics:** in Stage 2 flows from the S and R compartments to the V compartment are added;

(E4) **Running utility:** the research cost, which is present in Stage 1, disappears in Stage 2.

The first finding is that the lockdown should be intensified after the vaccine approval if the pace of the vaccination campaign is rather slow. Secondly, by comparing the strategies of an anticipating and a myopic planner respectively, we find that the anticipation of the vaccination arrival also leads to a stricter lockdown in the period without vaccination. For both findings, an intuitive explanation is offered.

The content of Section 2.2 is an article titled "Communication strategies to contrast anti-vax action: a differential game approach". The article is jointly written by A. Buratto,[3] R. Cesaretto,[3] and M. Muttoni,[3] and it is published in the Central European Journal of Operations Research in 2024.

Unlike the rest of the applications studied in this thesis, this work presents an infinite-horizon differential game without regime switches, where the players are the healthcare system and a pharmaceutical firm that produces and sells a certain vaccine. Since the spread of misleading information has caused a decline in vaccination coverage, which in turn has lead to the reappearance of the related disease, both parties implement pro-vaccine communication campaigns to contrast anti-vax action. While the healthcare system aims to minimize the healthcare costs that unvaccinated people would entail, the pharmaceutical firm wants to minimize the missed profits from unsold vaccines.

We solve the respective HJB equations to compute the Markovian Nash equilibrium strategies, we then perform a sensitivity analysis of the strategies and steady state with respect to the parameters.

The main findings are that, at the Markovian Nash equilibrium, the strength of the anti-vax word of mouth increases the healthcare system's investment in vaccine communication, while not affecting the firms pro-vaccine campaign. Moreover, upon variation of most parameters, we find that the two players act

[3]Dipartimento di Matematica "Tullio Levi-Civita", Università degli Studi di Padova

as strategic substitutes: the smaller the healthcare system's campaign, the higher the firms one.

2.1 Should the COVID-19 lockdown be relaxed or intensified in case a vaccine becomes available?

Article written by A. Buratto, M. Muttoni, S. Wrzaczek, and M. Freiberger; published in PLOS ONE (2022) [20].

2.1.1 Introduction

Since the beginning of 2020, the COVID-19 pandemic has kept the world in suspense. In addition to distance rules and hygiene measures, the immediate reaction of most countries was a lockdown of all nonessential parts of the economy. Although these measures were quite successful in terms of reducing infections and saving lives (e.g., Flaxman et al. [42] estimates that 3.1 million deaths have been averted by lockdown measures in 11 countries in Europe), the economy suffered enormously (according to Mandel and Veetil [76] the decrease of the world output fell up to 23% in spring 2020), and it became clear quite soon that an efficient vaccination will be the only viable option to end the pandemic in the long term.

Since then, all over the world additional research effort was put into exploring COVID-19 with significant resources allocated towards the development of vaccines and medications. Initially, it was speculated that vaccines could become available about 1.5 years after the beginning of the pandemic. However, due to the large efforts and unprecedented international collaboration, several vaccinations with different technologies (vector vaccine, mRNA, dead vaccine, etc.) have been developed in record breaking time (the vector vaccine "Convidicea", for instance, has already been approved by the end of June 2020 in China, followed by "Sputnik V" in August 2020 in Russia) that no-one could expect in spring 2020.

Naturally, in 2020 governments found themselves in a position where they could only react to the dynamic development of the pandemic. I.e., lockdowns (with different intensities) and travel restrictions have been implemented regarding certain identification numbers (e.g., number of new infections, number

of total infections, number of hospitalized people, number of people in intensive care units) and the dynamics without knowing any details concerning the *time horizon*. This is reflected by two problematic points in the relevant literature on COVID-19.

First, governments intend to take measures to control the pandemic in the best way. The term *best* in this context means that the performance of interventions is valued with respect to some objective function, which might include an economic (e.g., lost GDP per capita) as well as a health part (e.g., monetary valuation of lost lives). Due to the dynamic nature of a pandemic, the application of dynamic optimization or optimal control theory is suitable (see, e.g., Hinderer et al. [62] and Grass et al. [52]).

Second, the end of the pandemic in many other contributions (see below for a review of the relevant literature) is often identified with the availability of an effective vaccine. Neglecting regional differences (industrialized countries, emerging economies, and developing countries), the current situation illustrates that COVID-19 related problems, such as high infection numbers and congested intensive care (IC) units, are not resolved immediately with vaccination availability. Vaccination does not assure an instantaneous coverage of the population due to administration time and management issues that may arise. It may be a long process, and it is not necessarily organized at a constant rate. Lockdown still remains the most effective tool to oppose the virus, and it needs to be considered (probably with certain adaptations) at least for some time.

The latter argument also becomes apparent in the data. Fig 2.1 plots the "intensity of the lockdown" (solid lines, values correspond to the left vertical axis) and the proportion of vaccinated people (dashed lines, values correspond to the right vertical axis) for US (black lines), Israel (red lines) and Germany (blue lines). The data has been taken from `https://ourworldindata.org` (last access May 15, 2021; see Mathieu et al. [79]). The intensity of the lockdown (referred to as "stringency index" in the data set) is measured from 0 to 100, where 0 means no lockdown at all and 100 lockdown of all nonessential parts of the economy. These three countries are highly developed and have the financial power to afford the necessary vaccine doses, but are heterogeneous concerning their vaccination strategies. While the US campaign has been quite quick and efficient, Germany faced a couple of problems at the beginning of the campaign much like numerous other European countries. Israel's campaign has probably been the fastest in the world, resulting in an early relaxation of lockdown measures. Despite these

differences, in all three countries, the lockdown was not relaxed immediately at the vaccine approval (which can also be observed for many other countries). Actually, in Israel and Germany it was even intensified. Relaxation in all three countries occurred at the time when a reasonable proportion of the population had been vaccinated (US and Germany: $15-20\%$, Israel: $\approx 40\%$) which, of course, goes along with a decrease of infection numbers and reduced pressure on IC units.

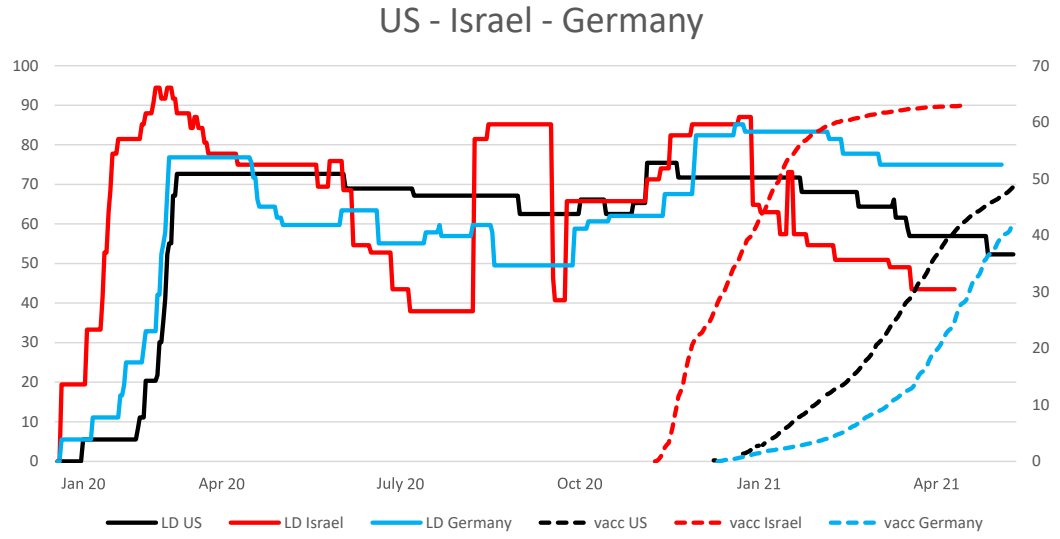


Figure 2.1: **US, Israel and Germany.** Lockdown intensity and proportion of vaccinated people

To overcome these crucial issues, we use a standard SIR model originated in the seminal paper Kermack and McKendrick [67], as already previously done in several mathematical economic papers on the COVID-19 pandemic (see the following paragraphs for a brief review), and include lockdown intensity as a policy measure (that has to be chosen optimally) as well as a stochastic arrival time of an efficient vaccination, that subdivides the time horizon into two stages.

In the stage before an efficient vaccination is available (from now on referred to as Stage 1) the lockdown is the only instrument of the government to push back the virus, and an additional research effort is employed towards the discovery of an effective vaccine. In the second stage, when a vaccine has been discovered (from now on referred to as Stage 2), the administration of the vaccine speeds up the immunization of the population. Lockdown measures can still be undertaken, but additional research effort is not necessary any more.

With this framework we aim to address the following *research questions*:

1. How does the optimal lockdown intensity change during Stage 1, when the vaccination approval is expected (compared to a situation without the expectation of a vaccine development)?

2. How should the optimal lockdown intensity be adapted at the time and after the vaccination is approved?

3. How are the pandemic costs composed w.r.t. different approval times of a vaccination?

These questions are systematically addressed in the section "Numerical results" for different model scenarios. The thorough discussion gives important insights and extends the results of previous papers.

For the model setup, we adapt the baseline model proposed in Alvarez et al. [4]. From a mathematical point of view, any other optimal control model for the current pandemic (see the list of references) could be used. However, adding further compartments (i.e., state variables) or additional policy instruments (i.e., control variables) would neither change the qualitative behaviour nor improve the intuitive understanding. Additional channels would cover up the effect of the stochastic arrival rate of a vaccination.

The literature on mathematical models of the COVID-19 pandemic is already quite rich. At this point, we focus the discussion on papers that are most related to our model. Most papers consider a macroeconomic objective function with lockdown as a control variable but address different aspects of optimal policy interventions during the pandemic. Alvarez et al. [4] and Gonzalez and Niepelt [51] consider optimal lockdown policies over time. Aspri et al. [6] extends the epidemiological model and considers a sophisticated SEIARD model. Ali et al. [3], on the other hand, analyzes the effect of quarantine. El Ouardighi et al. [35] includes the effect of social fatigue and endogenous treatment capabilities. Caulkins et al. [25, 26, 27] concentrate on the qualitative behaviour of the solution and identify parameter regions where the optimal solution is qualitative stable. The parameter regions are separated by bifurcation curves where Skiba solutions or other interesting mathematical phenomena arise. Federico and Ferrari [38] considers a transmission rate originating from a stochastic diffusion process, which can be influenced by confinement policies. Acemoglu et al. [2] is the only paper that includes a stochastic time horizon within a SIR model with different age groups (deriving optimal interventions that may differ between the groups). However, in contrast to our contribution, the period after the vaccination arrival is not considered. Fu et al. [48] embeds work intensity, which

is chosen on an individual basis, into a SIR model and derives the optimal lockdown policy for a given vaccination rate. The main finding is that the lockdown is relaxed when the vaccination effort is intensified (i.e., lockdown and vaccination are substitutes). This is contrasted by Caulkins et al. [24], where a systematic sensitivity analysis of the optimal lockdown intensity with respect to the vaccination rate and the valuation of a lost life is provided. In addition to the substitution property (as in Fu et al. [48]) also scenarios with lockdown and vaccination being complements are identified. Up to the best of our knowledge, Garriga et al. [49] is the only paper that considers an (exogenous) stochastic arrival of a vaccine and the adaptation of the optimal lockdown policy afterwards. However, the paper uses a different model setup and solution method, as well as a considerably different focus in their numerical examples.

Except Acemoglu et al. [2] and Garriga et al. [49], the works mentioned above consider a fixed time horizon assuming that this immediate end of the pandemic is known a priori. As previously mentioned, this assumption is relaxed in our contribution within a model being truly simple in terms of the epidemiological dynamics. However, it allows to analyze the effect of a stochastic arrival rate of a vaccine and to work out the impact of different arrival times, which is not dealt with in Acemoglu et al. [2] and Garriga et al. [49]. This is possible by adopting a novel approach for multi-stage optimal control models with random switching time as presented in Wrzaczek et al. [118]. Based on the deterministic reformulation of the objective function (see Boukas et al. [11]), the problem is transformed into an age-structured optimal control problem. This approach allows treating both stages simultaneously (in contrast to the *traditional* method), thus implying a detailed characterization of the link between the two stages.

The rest of the paper is organized as follows. The following section introduces the model together with some analytic results. Numerical solutions for different cases concerning the speed of the vaccination campaign are presented in the section "Numerical results". The last section concludes.

2.1.2 The model

Within this section, we extend the well-known SIR model proposed in the seminal article Kermack and McKendrick [67] and improved in many other papers. We introduce the control variable *lockdown intensity* to fight the spread of the virus and a stochastic arrival time (i.e., a switch of the dynamics) of an

efficient vaccine that separates the planning horizon into two stages. First, the adapted SIR dynamics are described; then the government's objective function (i.e., costs) is defined. Finally, the full model is formulated.

Dynamics

Consider an infinite planning period $[0, +\infty)$ divided into two stages by the entrance of an effective vaccine at time τ. Assuming that such an entrance is a certain event (sooner or later), the arrival rate of τ is discussed below by the mean of a cumulative distribution function (see Eq (2.5)).

In Stage 1, i.e., in the interval $[0, \tau)$, no effective vaccine is available, therefore the two possible measures to fight the disease are the lockdown intensity, denoted by $\ell(t)$, and the additional research effort (research boost) towards the discovery of an effective vaccine, denoted by $r(t)$. Both measures (entering nonlinearly in the model) are continuous control variables and optimally set by the government. We assume that it is possible to close only nonessential parts of the economy. Essential parts (such as energy supply, health services, and basic food production) have to be kept open. This implies $\ell(t) \in [0, \bar{L}]$, where $\ell(t) = 0$ corresponds to no lockdown and $\ell(t) = \bar{L} > 0$ to full lockdown.

The total population $N(t)$ is divided into the three common compartments, i.e., susceptibles $S(t)$, infected $I(t)$ and recovered $R(t)$. In Stage 2 an additional compartment $V(t)$ for vaccinated individuals is introduced. Due to the negligible effect on the pandemic, the non-COVID-19 mortality rate and the birth rate are ignored. As a result, the total population equals $N(t) = S(t) + I(t) + R(t) + \mathbb{I}_{t\geq\tau}V(t)$, where $\mathbb{I}_{t\geq\tau}$ denotes the usual indicator function (equal to 1 for $t \geq \tau$, zero otherwise).

At any lockdown intensity a proportion of ℓ, $(1 - \theta\ell)$ is active and can transmit the virus, where $\theta \in [0, 1]$ is an exogenous measure of the lockdown effectiveness. Following Alvarez et al. [4] the transmission rate is defined as $\beta(\ell) = \beta_0(1-\theta\ell)^2$, where $\beta_0 > 0$ denotes the transmission rate without lockdown. Thus, during Stage 1 (i.e., $t \in [0, \tau)$) the number of susceptibles evolves according to

$$\dot{S}(t) = -\beta(\ell(t))\frac{S(t)I(t)}{N(t)}, \qquad S(0) = S_0, \qquad \text{for } t < \tau, \tag{2.1}$$

where S_0 is the initial value of the compartment. The increase of the infected duee to new infections equals the decrease of the susceptibles. On the other

hand, the infected decrease because of COVID-19 related deaths and recoveries. Let $1/\gamma$ denote the average dwell time of individuals in $I(t)$, then γ equals the percentage of infected leaving the compartment at any point in time, i.e., γ = (recovered + dead)/infected. For the COVID-19 related death rate, we follow Alvarez et al. [4] and assume the following linear increasing form

$$\varphi(I(t)) = \gamma \cdot (\bar{\varphi} + \kappa I(t)), \tag{2.2}$$

where the assumption $(\bar{\varphi} + \kappa I(t)) < 1$ guarantees a positive recovery rate equal to $\gamma\left(1 - (\bar{\varphi} + \kappa I(t))\right)$. The death rate is increasing in I due to the congestion effects of the health sector (for dramatic examples we refer to Lombardia in Italy or New York City both in spring 2020, or Brazil in spring 2021). Several contributions use a similar form, see e.g., Caulkins et al. [25, 26, 27] who assume an excess death rate if the ICU capacity is exceeded.

Therefore, the infected dynamics during both stages is

$$\dot{I}(t) = \beta(\ell(t))\frac{S(t)I(t)}{N(t)} - \gamma I(t), \qquad I(0) = I_0, \tag{2.3}$$

where I_0 contains the initial number of infected individuals.

The size of the total population at any t diminishes by COVID-19 deaths. This implies the following dynamics for both stages

$$\dot{N}(t) = -\varphi(I(t))I(t), \qquad N(0) = N_0, \tag{2.4}$$

with an initial population size of N_0. The number of recovered people $R(t)$ in both stages can be directly obtained by $R(t) = N(t) - S(t) - I(t) - \mathbb{I}_{t\geq\tau}V(t)$.

The time instant τ, when an efficient vaccination becomes *available*, is assumed to be an absolute continuous random variable. Let us define by $Z(t)$ the probability of discovering an effective vaccine after t. In other words, the probability of remaining in Stage 1 until t. We formalize this probability throughout the complementary cumulative distribution function (tail distribution)

$$Z(t) = Prob\{\tau > t\}. \tag{2.5}$$

As a result, the corresponding switching rate is obtained by

$$\frac{-\dot{Z}(t)}{Z(t)} = \eta(r(t), t). \tag{2.6}$$

Since $Z(0) = Prob\{\tau > 0\} = 1$ holds naturally, this distribution function satisfies the following Cauchy problem

$$\dot{Z}(t) = -\eta(r(t), t)Z(t), \qquad Z(0) = 1. \tag{2.7}$$

We assume $\eta(\cdot, t)$ to be positive, continuous, and to depend on the additional research effort $r(t)$, which is a control variable in Stage 1. Note that $r(t) \in [0, 1]$, i.e., $r(t) = 0$ and $r(t) = 1$ correspond to the cases without and with maximal available additional research effort respectively. Note that research effort cannot exceed a certain level at least on the medium term due to constraints on the availability of experienced researchers, research facilities, etc. Consequently, $\eta(0, t)$ captures the switching rate resulting from base research effort (spent in R&D on a regular basis, without any additional research effort, i.e., with $r = 0$).

We assume that the research boost $r(t)$, aimed at finding an effective vaccine, accelerates its development with efficacy $\eta_1 > 0$. Thus, the switching rate depends on $r(t)$ as follows,

$$\eta(r(t), t) = p(t)\left(\eta_0 + \eta_1 r(t)\right), \tag{2.8}$$

with $\eta_0 > 0$ and where $p(t)$ is an increasing time dependent function such that $p(0) \in (0, 1)$ and $\lim_{t\to\infty} p(t) = 1$. For the numerical calculations in Section "Numerical results", we are using a Gompertz sigmoid function, i.e., $p(t) = e^{-p_1 e^{-p_2 t}}$ with $p_1, p_2 > 0$ (see Table 2.2 for the parameter values). This functional specification guarantees not only a switching rate that increases in $r(t)$, but also a time-dependent learning effect (non-autonomous $p(t)$) resulting in a more efficient use of research efforts.

At the time instant τ a vaccine is developed and becomes available, thus marking the start of the vaccination of the population. We assume that the cost for production and administration of the vaccine (research costs are already covered in Stage 1) are minuscule compared to costs resulting from lockdown measures and lost lives. This implies that at every t as many people as possible (with respect to production and administration capacities) get vaccinated. Hence

the number of vaccinations per unit of time $\alpha(\hat{t})$ (where $\hat{t}$ corresponds to the time after the switch, i.e., $\hat{t} := t - \tau$) follows the exogenously given availability of vaccination after discovery. We assume that $\alpha(\cdot)$ is an increasing and concave function over time, i.e., $\alpha'(\hat{t}) > 0$ and $\alpha''(\hat{t}) < 0$. For the functional specification, we propose the following form

$$\alpha(\hat{t}) = \frac{\alpha_1 \hat{t} + \alpha_2}{\hat{t} + \alpha_3}, \tag{2.9}$$

with positive parameters $\alpha_i > 0$ $(i = 1, 2, 3)$; $\alpha_1 \alpha_3 > \alpha_2$ guarantees that $\alpha(t)$ is increasing and concave. At the switch $\alpha(\tau - \tau) = \alpha(0) = \alpha_2/\alpha_3$ vaccination doses are available. This number increases up to the maximum level of α_1, which is the limit of the above expression.

Although people are vaccinated as quickly as possible starting at τ, as discussed in the introduction (and as observed in lots of countries in the first half of 2021), the lockdown is still an indispensable tool to control the pandemic since people cannot be vaccinated fast enough. We assume that susceptible and recovered people are vaccinated without any prioritization. After getting the vaccine, people enter the compartment of vaccinated people $V(t)$, which is the absorbing state in our model. As a result, the dynamics of $S(t)$ and $R(t)$ in Stage 2 read

$$\begin{aligned} \dot{S}(t) &= -\beta(\ell(t)) \frac{S(t)I(t)}{N(t)} - \alpha(t - \tau) \frac{S}{S + R} \mathbb{I}_{S+R>0} \\ \dot{I}(t) &= \beta(\ell(t)) \frac{S(t)I(t)}{N(t)} - \gamma I(t). \end{aligned} \tag{2.10}$$

where the indicator function guarantees that vaccination ends if everybody is vaccinated. $V(t)$ just collects all vaccinated people in Stage 2

$$\dot{V}(t) = \alpha(t - \tau)\mathbb{I}_{S+R>0}, \qquad V(\tau) = 0. \tag{2.11}$$

Costs – objective function

The decision maker in our model is the government balancing (i) the costs of lost lives and (ii) the costs of lockdown and subsidies in the R&D sector (aiming at accelerating the vaccine development). Therefore, the objective function consists of an economic (lockdown, research subsidies) and a health economic part (lost lives), both measured in GDP (gross domestic product) per day. All objectives

in Stage 1 are represented by the function

$$g_1(I(t), \ell(t), r(t)) := c_h(I(t)) + c_\ell(\ell(t)) + c_r(r(t)). \tag{2.12}$$

Here $c_h(I(t)) := \psi \cdot \varphi(I(t)) \cdot I(t)$ denotes health economic costs, i.e., the costs of lost lives weighted with the value of a statistical life ψ. The range of ψ in the literature ranges from 20 in Alvarez et al. [4] to 150 in Kniesner et al. [69]. In Caulkins et al. [25, 26] this value is used as a bifurcation parameter in a sensitivity analysis in a different model setting with a different research focus.

Lockdown costs are assumed to be quadratic, i.e.,

$$c_\ell(\ell(t)) := w\ell^2(t) \tag{2.13}$$

with a positive parameter w. This results from higher marginal costs as lockdown becomes more intense, due to the interconnection of the economy. Note that the definition of lockdown costs varies in the corresponding literature. While e.g., Alvarez et al. [4] assumes linear costs, e.g., Caulkins et al. [25, 26] assume non-linear costs that also depend on the available workforce (i.e., $N(t)$ diminished by $I(t)$). Moreover, we are relaxing this assumption in the section "Robustness check" and allow for convex-concave costs.

Similar to the lockdown costs, the research costs are assumed to be quadratic $c_r(r(t)) = c_0 r^2(t)$ with a positive parameter c_0. This appropriately reflects reality since the research effort enters linearly in the switching rate and research projects are funded in order with their priority and probability of success.

The costs in Stage 2 are analogous to the ones in Stage 1, they only differ for the lack of additional research efforts (which are assumed to be zero after the arrival of an efficient vaccine), i.e.,

$$g_2(I(t), \ell(t)) := c_h(I(t)) + c_\ell(\ell(t)). \tag{2.14}$$

Note that we do not consider vaccination costs explicitly, while they could be argued to be included. However, as discussed before, vaccinations are given at the maximum pace at any t (due to the fact that the corresponding costs are negligible) implying that the optimal results would not change.

As the decision maker aims for the minimization of the expected aggregated discounted costs over an infinite time horizon (taking the stochasticity of the

switching time τ into account), the objective function can be written as

$$J^*(X(0)) := \min_{\ell(t),r(t)} \mathbb{E}_\tau \left[\int_0^\tau e^{-\rho t} g_1(I(t), \ell(t), r(t))\ dt + \int_\tau^{+\infty} e^{-\rho t} g_2(I(t), \ell(t))\ dt \right] \quad (2.15)$$

Thereby ρ is the discount rate and $X(t)$ is a vector of the compartments $S(t)$, $I(t)$ and $N(t)$.

Following Wrzaczek et al. [118], the objective function (2.15) can be reformulated as

$$J^*(X(0)) = \min_{\ell(t),r(t)} \int_0^\infty e^{-\rho t} Z(t) \left[g_1(I(t), \ell(t), r(t)) + \eta(r(t), t) J_2^*(X(t), t) \right]\ dt \quad (2.16)$$

where $J_2^*(X(t), t)$ denotes the optimal value (in mathematical terms referred to as value function, see Grass et al. [52]) of Stage 2 given a vaccine approval at t (with state variables $X(t)$ at t).

Full model

Using the subscripts 1 and 2 to refer to the (state and control) variables corresponding to Stage 1 or 2 respectively, the problem in Stage 1 can be written as

$$\begin{aligned}
J^*(X(0)) &= \min_{\ell_1(t),r_1(t)} \int_0^\infty e^{-\rho t} Z_1(t) \left[g_1(I_1(t), \ell_1(t), r_1(t)) + \eta(r_1(t), t) J_2^*(X_1(t), t) \right]\ dt \\
\text{s.t. } \dot{S}_1(t) &= -\beta(\ell_1(t)) \frac{S_1(t) I_1(t)}{N_1(t)} \\
\dot{I}_1(t) &= \beta(\ell_1(t)) \frac{S_1(t) I_1(t)}{N_1(t)} - \gamma I_1(t) \\
\dot{N}_1(t) &= -\varphi(I_1(t)) I_1(t) \\
\dot{Z}_1(t) &= -\eta(r_1(t), t) Z_1(t)
\end{aligned} \quad (2.17)$$

with initial conditions

$$S_1(0) = S_0, \quad I_1(0) = I_0, \quad N_1(0) = N_0, \quad Z_1(0) = 1. \quad (2.18)$$

The problem in Stage 2 can be written analogously, and defines $J_2^*(X(t), t)$ which occurs in the objective function of Stage 1. Since for every possible realization

of τ the value function is derived from an optimal control problem, $x_2(t,\tau)$ indicates a (state and control) variable of Stage 2 at t, given a vaccine approval at τ ($\tau \leq t$). This yields

$$\begin{aligned}
J_2^*(X(\tau),\tau) &= \min_{\ell_2(t,\tau)} \int_\tau^\infty e^{-\rho t} g_2(I_2(t,\tau),\ell_2(t,\tau))\, dt \\
\text{s.t. } \dot{S}_2(t,\tau) &= -\beta(\ell_2(t,\tau))\frac{S_2(t,\tau)I_2(t,\tau)}{N_2(t,\tau)} - \alpha(t-\tau)\frac{S_2(t,\tau)}{S_2(t,\tau)+R_2(t,\tau)}\mathbb{I}_{S_2(t,\tau)+R_2(t,\tau)>0} \\
\dot{I}_2(t,\tau) &= \beta(\ell_2(t,\tau))\frac{S_2(t,\tau)I_2(t,\tau)}{N_2(t,\tau)} - \gamma I_2(t,\tau) \\
\dot{N}_2(t,\tau) &= -\varphi(I_2(t,\tau))I_2(t,\tau) \\
\dot{V}_2(t,\tau) &= \alpha(t-\tau)\mathbb{I}_{S_2(t,\tau)+R_2(t,\tau)>0}
\end{aligned} \tag{2.19}$$

with initial conditions

$$S_2(\tau,\tau) = S_1(\tau), \quad I_2(\tau,\tau) = I_1(\tau), \quad N_2(\tau,\tau) = N_1(\tau), \quad V_2(\tau,\tau) = 0. \tag{2.20}$$

The dot-notation in Stage 2 refers to the derivative with respect to time, i.e., $\dot{x}(t,\tau) = dx(t,\tau)/dt$, while τ remains constant.

Clearly, the number of recovered people in both stages can be derived by the identity $N = S+I+R(+V)$. Note that in the problem of Stage 2 the vaccine arrival time τ is only a parameter (therefore entering as parameter in the value function $J_2^*(X(\tau),\tau)$), whereas in the objective function of Stage 1 the value function of Stage 2, $J_2^*(X_1(t),t)$ is evaluated for every t, as a possible arrival time. Table 2.1 summarizes all variables and functions of the model for both stages, at a glance.

Solution and economic intuition

For the derivation of the first-order conditions, we first transform the model into an age-structured optimal control model (see Wrzaczek et al. [118]). Then the age-structured Maximum Principle (see Feichtinger et al. [41]) is applied. This representation offers some advantages compared to the standard approach presented in Boukas et al. [11], as discussed in Wrzaczek et al. [118].

The transformation of Eqs (2.17)–(2.19) into the age-structured form as well as the derivation of the optimality conditions, is quite technical and deferred to the S1 Appendix.

The optimal values of the control variables are summarized in the following Theorem.

Notation	Description	Stage 1	Stage 2
Controls:			
ℓ_i	lockdown intensity	X	X
r_1	research effort (development of vaccination)	X	-
States:			
S_i	susceptibles	X	X
I_i	infected	X	X
R_i	recovered	X	X
V_2	vaccinated	-	X
N_i	total population	X	X
Z_1	probability of vaccine approval after t	X	-
Functions:			
$\eta(r_1, t)$	switching rate	X	-
$\alpha(t-\tau)$	number of vaccinations per unit of time	-	X
$\beta(\ell_i)$	transmission (infection) rate	X	X
$\varphi(I_i)$	fatality rate	X	X
$\gamma - \varphi(I_i)$	recovery rate	X	X

Table 2.1: **Overview of Notation ($i = 1, 2$).**

Theorem 2.1. *Consider the above multi-stage optimal control problem with random switching time* (2.17)–(2.19)*. Then, assuming the existence of an optimal solution, the optimal research efforts in Stage 1 and the optimal lockdown levels in both stages are given by*

$$\begin{aligned}
r_1^*(t) &= \min\left\{\max\left\{\frac{\eta_1 p(t)\left(\xi_Z(t,t)-\lambda_Z(t)\right)}{2c_0}, 0\right\}, 1\right\},\\
\ell_1^*(t) &= \min\left\{\max\left\{\frac{1}{\theta}\frac{\left(\lambda_I(t)-\lambda_S(t)\right)\beta_0\theta\frac{S_1(t)I_1(t)}{N_1(t)}}{\left(\lambda_I(t)-\lambda_S(t)\right)\beta_0\theta\frac{S_1(t)I_1(t)}{N_1(t)} - Z_1(t)\frac{w}{\theta}}, 0\right\}, \bar{L}\right\},\\
\ell_2^*(t,\tau) &= \min\left\{\max\left\{\frac{1}{\theta}\frac{\left(\xi_I(t,\tau)-\xi_S(t,\tau)\right)\beta_0\theta\frac{S_2(t,\tau)I_2(t,\tau)}{N_2(t,\tau)}}{\left(\xi_I(t,\tau)-\xi_S(t,\tau)\right)\beta_0\theta\frac{S_2(t,\tau)I_2(t,\tau)}{N_2(t,\tau)} - Z_1(\tau)\eta(\tau)\frac{w}{\theta}}, 0\right\}, \bar{L}\right\},
\end{aligned} \tag{2.21}$$

where λ_y (ξ_y) denotes the adjoint variable of Stage 1 (Stage 2) of a given state variable y (for the corresponding adjoint equations we again refer to the S1 Appendix).

Giving a more detailed discussion on $\lambda_Z(t)$ and $\xi_Z(t,t)$: $\lambda_Z(t)$ denotes the adjoint variable of $Z_1(t)$, which is the probability that the vaccine is approved after t. $\xi_Z(t,t)$ denotes the adjoint variable of $Z_2(t,t)$, which is the probability

density that the vaccine is approved at t.

Proof. Manipulating the general first order optimality conditions (see S1 Appendix), directly leads to the presented expression for the optimal controls. The max and min operators inside the expressions guarantee that the control variables stay within their admissible regions. □

The formal expression of Theorem 2.1 can be interpreted as follows:

- Optimal research effort (Stage 1): In case the fraction inside the brackets is positive and below one (i.e., $r_1^*(t) \in (0,1)$), the max and min operators can be omitted. Then the optimal value equals the marginal effect of research subsidies over their marginal costs. The marginal effect comprises of the product of the marginal switching rate (w.r.t. research effort) and the shadow price of a switch from Stage 1 to Stage 2 at t (i.e., the difference of the corresponding shadow prices $\xi_Z(t,t)$ and $\lambda_Z(t)$).
  Therefore $r_1^*(t)$ increases if the marginal effect of the switching rate increases, if the marginal value of the vaccine approval increases, or if the marginal cost decreases.

- Optimal lockdown intensity without vaccination (Stage 1): The second-order optimality conditions (see S1 Appendix) imply that the denominator of the fraction inside the brackets is always negative. In the case of an inner solution (i.e., $\ell_1^*(t) \in (0, \bar{L})$) the term $\lambda_I - \lambda_S$, which can be interpreted as the shadow price of one individual getting infected, is strictly negative. Thus the optimal value equals the marginal effect on the epidemiological dynamics (numerator) over the total effect (denominator) divided by the marginal effectiveness of the lockdown (θ). The marginal effect on the dynamics (numerator) comprises of the shadow price of a person getting infected (difference of the shadow prices of the corresponding compartments) weighted by the corresponding probability and the marginal effect of the lockdown intensity. To get the total effect (denominator), the marginal relative cost of lockdown efficiency (weighted by the probability that the vaccine has not been approved at t) is added to the effect on the epidemiological dynamics. This fraction is related to the effectiveness of the lockdown (i.e., division by θ). As a result, the lockdown increases in the effect on the dynamics (either by the shadow price of a person getting infected or by the corresponding probability) and decreases in the costs.

- Optimal lockdown intensity with vaccination (Stage 2): The interpretation of $\ell_2^*(t,\tau)$ is analogous to that of $\ell_1^*(t)$.

In addition to this interpretation Theorem 2.1 can be used for analyzing the lockdown intensity at the vaccination approval, which is not a-priori clear. Intuitively, one would expect that the lockdown will be relaxed or at most remain at the same level. However, as already discussed in the introduction, in the beginning of 2021 most countries in Europe reacted differently. These decision

can potentially be supported by our framework. In particular, the following theorem formalizes conditions for the adaptation of the lockdown intensity at the vaccine approval.

Theorem 2.2. *Consider the multi-stage optimal control problem with random switching time Eqs* (2.17)–(2.19) *and the optimal lockdown intensities given in Theorem 2.1. At the (stochastic) time of a vaccine approval τ the lockdown intensity is in general non-continuous (disruptive) and*

$$\begin{aligned} \textit{(disruptively) increasing} &\iff \frac{\xi_I(\tau,\tau)-\xi_S(\tau,\tau)}{\lambda_I(\tau)-\lambda_S(\tau)} > \eta(r_1(\tau),\tau) \\ \textit{continuous} &\iff \frac{\xi_I(\tau,\tau)-\xi_S(\tau,\tau)}{\lambda_I(\tau)-\lambda_S(\tau)} = \eta(r_1(\tau),\tau) \\ \textit{(disruptively) decreasing} &\iff \frac{\xi_I(\tau,\tau)-\xi_S(\tau,\tau)}{\lambda_I(\tau)-\lambda_S(\tau)} < \eta(r_1(\tau),\tau) \end{aligned} \tag{2.22}$$

given that the total effect of the lockdown intensity of Stage 1 and Stage 2 at τ has the same sign (a different sign just reverses the inequalities).

Proof. From Theorem 2.1 we obtain

$$\begin{aligned} \ell_2^*(\tau,\tau)-\ell_1^*(\tau) = \beta_0\frac{S_1(\tau)I_1(\tau)}{N_1(\tau)}\Big(&\frac{\xi_I(\tau,\tau)-\xi_S(\tau,\tau)}{(\xi_I(\tau,\tau)-\xi_S(\tau,\tau))\,\beta_0\theta\frac{S_1(\tau)I_1(\tau)}{N_1(\tau)}-Z_1(\tau)\eta(\tau)\frac{w}{\theta}} \\ &-\frac{\lambda_I(\tau)-\lambda_S(\tau)}{(\lambda_I(\tau)-\lambda_S(\tau))\,\beta_0\theta\frac{S_1(\tau)I_1(\tau)}{N_1(\tau)}-Z_1(\tau)\frac{w}{\theta}}\Big), \end{aligned} \tag{2.23}$$

for $\ell_1^*(\tau), \ell_2^*(\tau,\tau) \in (0,\bar{L})$ and accordingly for solutions on the boundary. Manipulation of this expression proves the assertion of the theorem. □

According to Theorem 2.2 the adaptation of the lockdown intensity at the time of the vaccine approval depends on the relation of the shadow prices of a new infected person in Stage 2 and Stage 1 at τ. If the ratio exceeds the approval rate $\eta(\cdot)$ at τ, new infections have a stronger effect with the availability of a vaccine than they would have without. Consequently the lockdown intensity has to be intensified disruptively at τ, which seems to be counter intuitive at first glance. However, the reason is hidden in the definition of the shadow price of new infections, which includes the net value of costs (i.e., lockdown, lost lives) and the effect on the dynamics of the epidemic. Therefore, in Stage 1 it might be more costly to keep a lot of people in the S compartment since everyone

has to pass the $S - I - R$-route (while keeping $\varphi(I(t))$ on an acceptable level) to achieve immunity. On the other hand, if the vaccine is available, susceptibles can pass directly from the S to the V compartment by getting vaccinated. Thus, with a more intensive lockdown it pays off to keep more people in the S state (i.e., keeping them healthy), pass them directly to V and therefore save lives (lowering costs). In other words, a vaccination works as a kind of shortcut in the SIR dynamics. A more intense lockdown enables more people to make use of it.

The interpretation of a continuous and disruptive decrease of the lockdown intensity (second and third line of Eq (2.22)) is analogous.

2.1.3 Numerical results

In this section, we present some numerical results obtained by applying the Pontryagin Maximum Principle (PMP) (see S1 Appendix for the analytical expressions). Most of the parameters' values for our baseline calibration are taken from Alvarez et al. [4] and other models of the related literature. We report the complete list of parameters in Table 2.2, with time being measured in years. Note within the plots that the time axis is scaled in months for enhanced readability.

Unlike Alvarez et al. [4], where the proportion S_0 of susceptibles in the population is set equal to 0.97, we fix it equal to 0.99, so that S_0 and I_0 sum up to 1 (and no initial recovered are considered, $R_0 = 0$) as in Caulkins et al. [25].

The parameters concerning research and vaccination must be explained in more detail. Up to the best of our knowledge the effect of research on the discovery of a vaccine has not been considered by other papers. As already explained after the definition of the switching rate Eq (2.8), we use a Gompertz sigmoid function for $p(t)$ with $p_1 = e$ and $p_2 = 3$, such that we end up with

$$p(t) = e^{-e^{1-3t}}. \tag{2.24}$$

This essentially means that research efforts are rather ineffective at the start of the pandemic (i.e., at $t = 0$), and develop their effectiveness soon (e.g., $p(0.5) = 54.5\%$ after half a year, $p(1) = 87.3\%$ after one year). For the base research efficiency η_0 and the efficacy of the research boost η_1 no data is available. We have chosen $\eta_0 = \frac{1}{1.5}$ and $\eta_1 = 4 - \eta_0$. This reflects a situation in which the probability of a vaccine being approved within the first year without an additional research

Parameter	Value	Description
S_0	0.99	initial proportion of susceptibles in population
I_0	0.01	initial proportion of infected in population
N_0	1	initial population
$\bar{L}$	0.7	maximal lockdown intensity
θ	0.5	lockdown efficacy
β_0	$0.13 * 365$	transmission rate without a lockdown
γ	$\frac{1}{18} * 365$	recovery rate (including COVID-19 deaths)
$\bar{\varphi}$	0.0068	COVID-19 fatality rate
κ	0.034	coefficient for additional fatality rate (congestion of health care system)
$\alpha_1, \alpha_2, \alpha_3$	4, 0.1, 1	parameters concerning the availability of the vaccines
ψ	40	value of a statistical life (in terms of GDP)
w	1	lockdown cost parameters (in terms of GDP)
η_0	$\frac{1}{1.5}$	parameter of the switching rate
η_1	$4 - \eta_0$	effectiveness of additional research effort
p_1, p_2	$e, 3$	parameters concerning the switching rate (Gompertz sigmoid)
c_0	0.013	research effort cost parameters (in terms of GDP)
ρ	0.05	discount rate

Table 2.2: **Summary of parameter values (per year).**

boost (i.e., $r_1(t) = 0$ for all t) is only about 30%. On the other hand, for a maximal possible research boost (i.e., $r(t) = 1$ for all t) the probability for the approval within one year increases to more than 80%. We are aware that these parameters can hardly be validated. However, several numerical runs with different values implied that the results are qualitatively very robust with respect to these parameters.

Fig 2.2 plots the benchmark availability of the number of vaccinations per unit of time (modeled according to Eq (2.9)), where we assume that the whole population can be vaccinated in a quarter of a year (i.e., $\alpha_1 = 4$) at the maximum intensity, i.e., if t tends to $+\infty$. Initially the campaign starts with 10% intensity ($\alpha_2 = 0.1$ and $\alpha_3 = 1$). That means that the whole population can in fact be vaccinated within about 10 months. This specification (used as benchmark scenario in the first subsection) reflects a typical situation in Europe, however, these numbers can vary across countries. Hence, a rather hypothetical example is studied in the second subsection, in which $\alpha(t)$ is increased proportionally such that the whole population can be vaccinated within one month.

For the numerical solution, the age-structured formulation of problem Eq (2.17)–

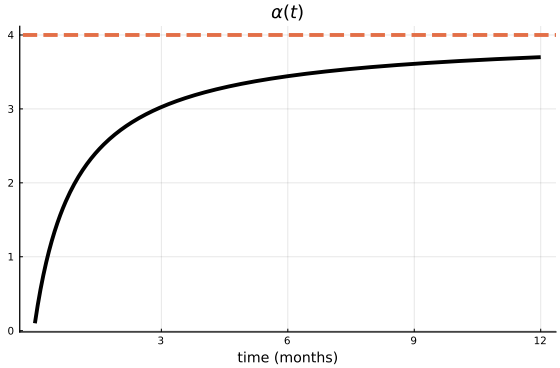


Figure 2.2: **Number of vaccinations per unit of time.**

(2.19) was used (see S1 Appendix). As already mentioned, this includes the advantage to solve both stages simultaneously, as one single problem instead of two sequential ones, i.e., first deriving Stage 2 for all possible values of the state variables at any τ; followed by Stage 1. An age-structured optimal control model implies the solution of a system of partial differential equations, which makes it difficult to use a standard boundary value solver. Thus, we use the established gradient optimization algorithm presented in Veliov [112]. After choosing initial guesses for the control variables, the algorithm uses the gradient to update the control variables iteratively until they converge to an optimum.

Before we start to discuss the specific results, we observe that with the parameters' values declared above, the SOCs and the optimality conditions (see S1 Appendix) are fulfilled. Fig 2.3 introduces the way the results are presented in the following subsections. The horizontal and vertical axes denote time (in months) and lockdown intensity respectively. The black line (subdivided into a solid and a dashed part) shows the optimal lockdown intensity in case the vaccination was not approved until t (i.e., lockdown intensity of Stage 1). A vaccination may be approved at any time. In the current figure the optimal lockdown intensity of Stage 2 is only plotted for the case that the approval happens after approximately $\bar{\tau} \approx 4$ months (green line). Therefore, the optimal lockdown intensity given the approval at $\bar{\tau}$ (which is a specific realization of the random variable τ) follows the solid black line until $\bar{\tau}$ (i.e., $\ell_1(t)$ for $t \in [0, \bar{\tau})$) and the green line afterwards (i.e., $\ell_2(t, \bar{\tau})$ for $t \in [\bar{\tau}, \infty)$). Thus at the approval of the vaccine at $\bar{\tau}$ the lockdown intensity jumps upwards. Before and after $\bar{\tau}$ the lockdown intensity changes continuously (as implied by the Maximum Principle). In all figures of the following subsections, when presenting results, the black line will represent trajectories of Stage 1, whereas the colored lines will

correspond to the trajectories of Stage 2 for different realizations of τ.

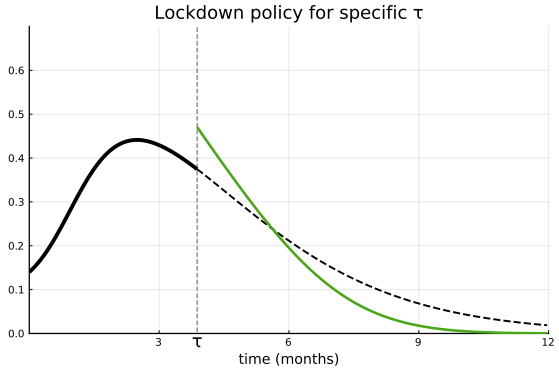


Figure 2.3: **Lockdown intensity for approval at $\tau = 4$.** Lockdown intensity without vaccination (black line, Stage 1) and with vaccination (green line).

In the first subsection, the benchmark scenario is presented with two different assumptions concerning the speed of the vaccination campaign. In Scenario 1 we assume the population to be vaccinated within 10 months, which seems to be pretty realistic considering many European countries. Meanwhile in Scenario 2 the population will be vaccinated within about one month.

In the second subsection, we compare the anticipative behavior of our model (i.e., lockdown measure during Stage 1) with the behavior of a myopic government, which does not include the arrival of a vaccination in the decision process.

Optimal lockdown intensity before and after vaccine approval

Scenario 1: vaccination within 10 months. Fig 2.4 illustrates the most important variables along the course of the pandemic. Black lines in all subplots denote variables that correspond to the period before the vaccine has been approved (Stage 1). Colored lines and dots to variables thereafter. The grey line in panel (d) corresponds to probability that a vaccine gets approved after t without an additional research boost at any time, i.e., $r_1(t) = 0$ for all t. All variables and values are plotted for the first year of the pandemic. Thereafter the solutions follow trajectories, which can be intuitively expected from the development within the first year (which means the pandemic ends, i.e., infections converge to zero, all people are getting the vaccine, lockdown is relaxed).

We start the discussion with the lockdown intensity plotted in panel (a). If a vaccine has not been approved, the intensity follows the course of the epidemic,

i.e., the number of infected (plotted in panel (d)). This is reasonable in our modeling context since the number of deaths directly depends on the number of infected at t (see the definition of COVID-19 related death rate Eq (2.2)), and goes along with all recent papers on COVID-19 which assume lockdown intensity as a control variable. At the time the vaccine is approved, however, the lockdown is intensified if the approval happens in the first seven months of the pandemic. If the approval happens later on, the lockdown intensity remains approximately at the same level. After the upward jump, the lockdown intensity is adapted again continuously and relaxed earlier compared to the case without a vaccine (since the vaccine enables susceptibles to surpass the infected compartment). It is also obvious that for an early vaccine approval the consequent period with more intense lockdown measures is quite long. The number of infected is rather high and it takes time for the vaccination campaign (which is relatively slow at the beginning) to unfold its effect. For a late vaccine approval (after more than seven months) the lockdown jumps only marginally, but also ends earlier. Referring back to Theorem 2.2 an early approval corresponds to the first case of the theorem, while a late approval corresponds to the second or third case.

A different illustration of the lockdown intensity only after the vaccine approval is plotted in panel (b). The three colored lines represent the lockdown intensity over the time that has passed since the vaccine approval (i.e., duration of Stage 2, $d = t - \tau$). For instance, the blue line shows the lockdown intensity that is implemented at three months across all possible approvals since the beginning of the pandemic ($d \in [0, 3]$). Thus, the lockdown value of the blue line with duration zero means the lockdown value of Stage 2 at $t = 3$ for $\tau = 3$. The red and the green lines show the same for six and nine months respectively. From that figure, it becomes evident that (for the majority of points in time) the lockdown intensity is decreasing both in duration and in time. This can be followed from the strict decrease of the colored lines and from the fact that the lines which correspond to higher t start lower and never intersect with one another.

Research boost during Stage 1 is shown in panel (c) and will be discussed together with panel (d), which illustrates the probability that the vaccine will be approved after t (black line: optimal research efforts, grey line: zero research boost). The colored dotted line in panel (d) represents the probability density that the switch happens at t, technically that is the initial value of the auxiliary state variable $Z_2(t, t)$ (see S1 Appendix for details). Starting with the grey line

in panel (d), we see that without additional research effort the probability that a vaccine is available on the market is only about 30% after one year. With the optimal additional research effort this chance increases and is twice as high (about 60%). For almost the whole first three months these additional efforts are at the maximum level in order to increase the probability density of the approval (colored dotted line in panel (d)). After that, additional efforts are decreasing, which is due to the fact that the peak of the pandemic has been passed (see number of infected in panel (f)) and the pandemic starts to decelerate. With a small time lag (which is due to increasing $p(t)$) the probability density for the vaccine approval (colored dotted line) also starts to decrease. For t approaching $+\infty$ the research boost, as well as the probability that the vaccine is approved after t and the probability density of the approval are converging towards zero. Please note again, that the approval rate and the corresponding probabilities cannot be calibrated at all and are furthermore very specific to the country of interest. As already mentioned in the introduction, China and Russia were the first countries approving a (vector) vaccination already in summer of 2020. The US and the EU approved two mRNA vaccines in December of 2020. Moreover, the date of approval does not directly indicate the extent of the availability of vaccine shots (represented by $\alpha(\hat{t})$ in our model), which can be quite diverse across countries too.

The left panel of Fig 2.5 strengthens the intuition of the qualitative shape of the lockdown intensity. It shows the effective reproduction number before and after vaccine approval over time, which can be derived by (see e.g., Driessche and Watmough [109])

$$R_{t,i}(t) = \frac{\beta(\ell_i) S_i / N_i}{\gamma}, \qquad i = 1, 2. \tag{2.25}$$

R_t equals the rate that one infected transmits the virus to one susceptible over the average duration of infectivity. In other words, it can also be interpreted as the average number of susceptibles that is infected by one infected individual. Therefore 1.0 denotes the important threshold, above which an epidemic shows dreaded exponential growth. The black line of Fig 2.5 (corresponding to Stage 1) starts from a considerably high value about 2.0, decreases steadily and drops below 1.0 shortly after the peak of the infection numbers. A stronger reduction is simply too expensive in terms of balancing lockdown and health economic costs (i.e., costs of lost lives). The situation changes, if the vaccine has already been

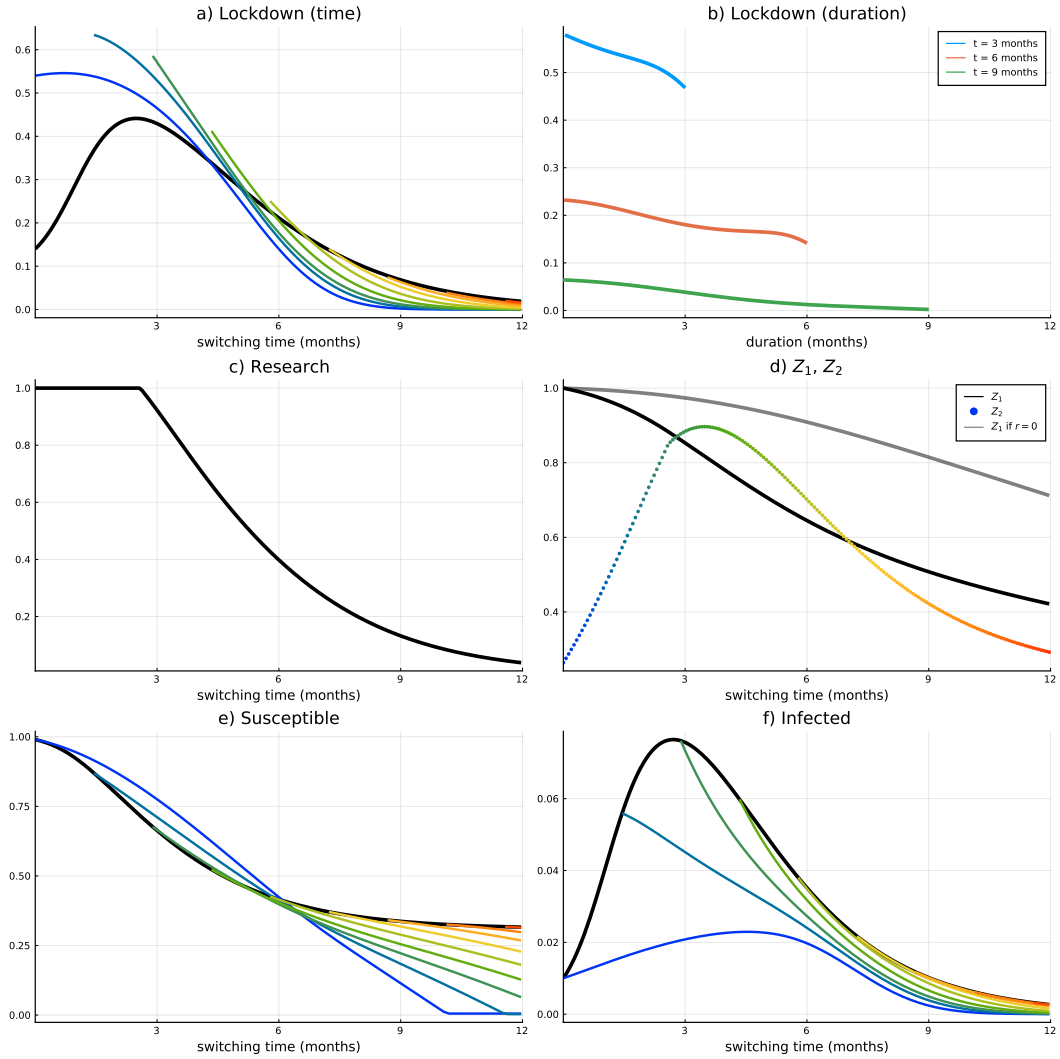


Figure 2.4: **Scenario 1 (Course of the pandemic).** (a) Lockdown intensity over time, (b) Lockdown intensity along duration, (c) Research effort over time, (d) Probability that τ has not set in yet, (e) Susceptibles, (f) Infected.

approved (colored lines), i.e., during Stage 2 of our model. Due to the disruptive upward jump of the lockdown intensity, also the effective reproduction number jumps to a value around the critical threshold 1.0. In most cases, when the vaccine is approved after the first month of the pandemic, R_t directly drops down below 1.0. Again, this results from the more intensive lockdown paying off in these cases, since susceptibles can move directly to the V compartment. The consequent reduction of the S compartment is very fast, which also implies a decrease in the effective reproduction rate (faster decrease of the colored lines in Fig 2.5) and enables the government to relax the lockdown earlier.

The right panel of Fig 2.5 plots the total aggregated objective value (and a

decomposition into its sub-parts), which realizes for different approval dates of the vaccine, i.e.,

$$\int_0^t e^{-\rho s} g_1(I^*(s), \ell^*(s), r^*(s))\ ds + \int_t^{+\infty} e^{-\rho s} g_2(I^*(s), \ell^*(s))\ ds \tag{2.26}$$

where the * denotes the optimality of control and state variables. Consider an approval at $\tau = 3$ months. The corresponding objective value can be read at the value of the dotted black line at $t = 3$. The solid black line, on the other hand, denotes the expected value of the objective values described above, which is the objective function of our original problem Eqs (2.17)–(2.19), i.e.,

$$\mathbb{E}_\tau \left[\int_0^\tau e^{-\rho t} g_1(I^*(t), \ell^*(t), r^*(t))\ dt + \int_\tau^{+\infty} e^{-\rho t} g_2(I^*(t), \ell^*(t))\ dt \right]. \tag{2.27}$$

Trivially this expected value is constant for over t. The blue, red, and green lines, dotted and solid, can be interpreted analogously and represent the decomposition of the objective value into lockdown costs, health economic costs, and research costs (the three corresponding curves and dots sum up to the black ones). The total and health economics costs increase over time, i.e., an early vaccine reduces deaths resulting in lower costs. For early vaccine approvals, the corresponding costs are considerably lower than the expected value, but exceed it shortly before $t = 5$ months. On the other hand the economic costs resulting from the lockdown (red) show a different picture. The upward jump in lockdown intensity at the time of approval (for early vaccine approvals) implies decreasing total economic costs for later approvals. For an early vaccine approval the lockdown costs are higher than on average. However, for a vaccine approval after about four months the relation reverses, since the upward jump at the time of approval gets smaller and the period where the lockdown in Stage 2 is more intense than it would have been in Stage 1 gets shorter.

Research efforts are considerably cheap compared to all other costs. This goes along with empirical evidence. Basically all countries devoted as many financial resources as possible to support research. Qualitatively, they are analogous to total and health economic costs. All in all, health economic costs (blue curve) dominate lockdown (red curve) and research costs (green curve), both qualitatively and in absolute terms. This seems to be realistic considering reality. Being aware of the enormous costs due to lockdown measures, countries all over the world did everything to keep deaths on a low level and agreed on huge financial

supports of the economy. It is at hand that these results are sensitive concerning the (monetary) value of lost lives (measured in GDP per capita, see e.g., Schelling [96], Shepard and Zeckhauser [98], Murphy and Topel [82], Hall and Jones [58], and others). However, governments (and also researchers) face the same dilemma whenever health economic questions are considered (e.g., support research on medications, decisions on the payment of expensive treatment, decisions concerning the extension or reduction of the health system).

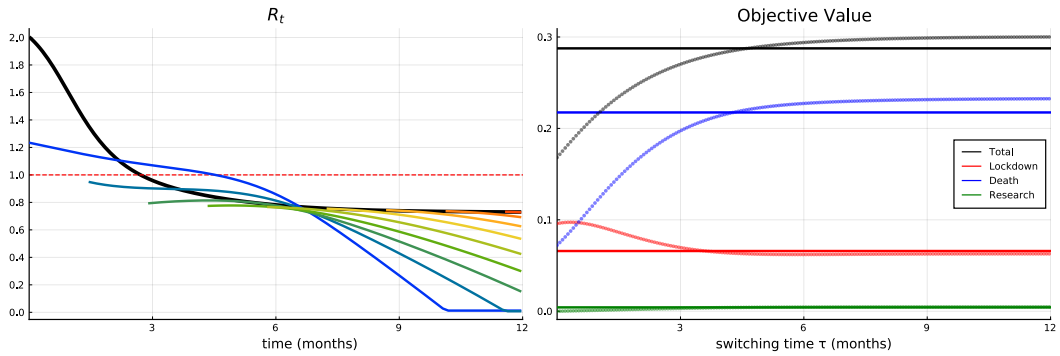


Figure 2.5: **Scenario 1.** Effective reproduction number over time (left) and decomposition of the objective value (right).

Numerical exercises for a broad range of parameters showed the qualitative robustness of the results.

Scenario 2: vaccination within one month Within this scenario, the same parameters as in scenario 1 are used, except the parameters concerning the availability of the vaccines (see Eq (2.9)). We assume that all involved parameters (i.e., $\alpha_1, \alpha_2, \alpha_3$) are increased proportionally (this implies that the qualitative shape of $\alpha(t)$ remains unchanged) such that the whole population can be vaccinated within about one month (instead of 10 as in scenario 1) after the approval. This is unrealistic and artificial. Israel was able to vaccinate $60 - 70\%$ of the inoculable population within three months (see Fig 2.1) and is therefore still slower if compared to this scenario. However, it is useful in order to intensify the understanding of the model.

Fig 2.6 provides the same set of plots for Scenario 2 as Fig 2.4 for Scenario 1. I.e., lockdown intensity over time and along duration (panels (a) and (b)), additional research effort over time (panel (c)), probability that the approval happens after t and probability density of vaccine approval (panel (d)), susceptibles and infected (panels (e) and (f)).

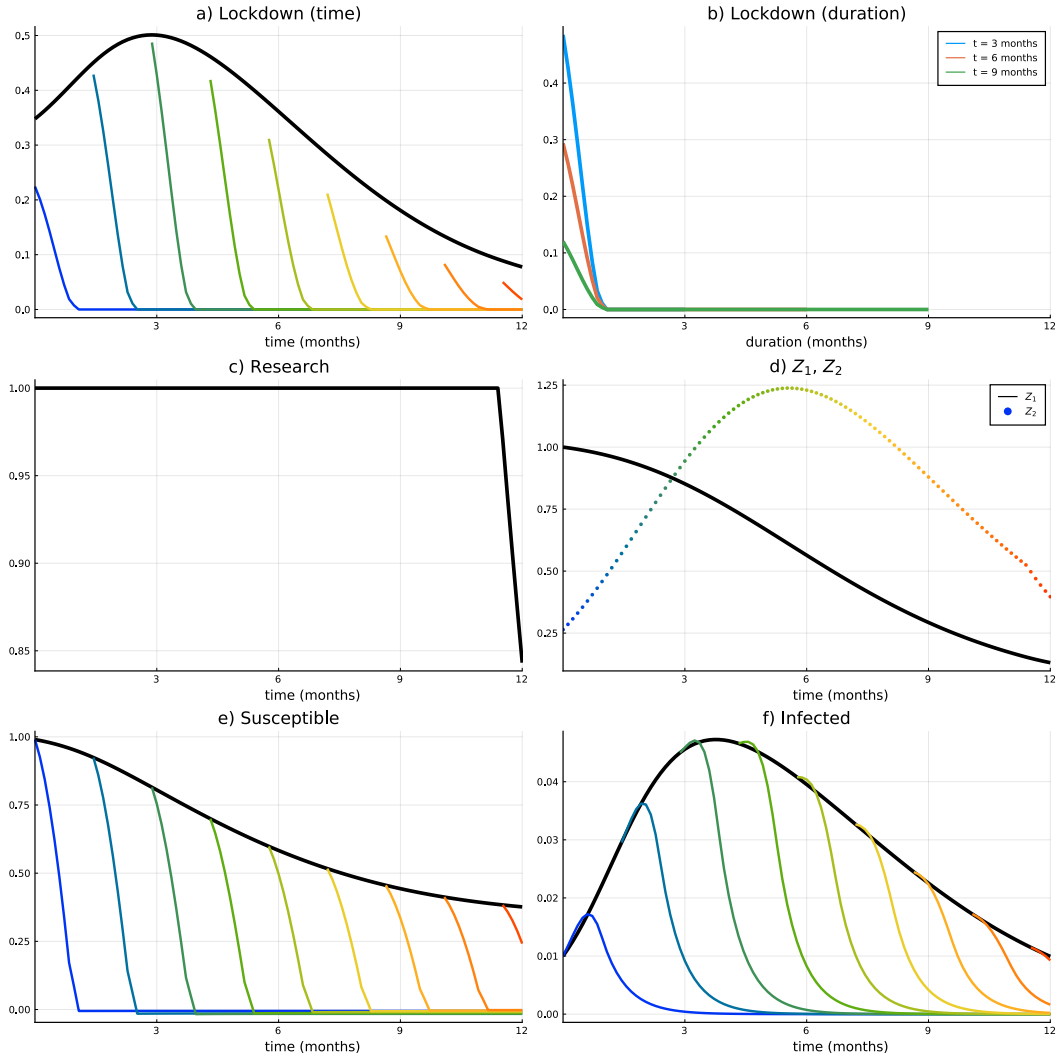


Figure 2.6: **Scenario 2 (Course of the pandemic).** (a) Lockdown intensity over time, (b) Lockdown intensity along duration, (c) Research effort over time, (d) Probability that τ has not set in yet, (e) Susceptibles, (f) Infected.

The high speed of the vaccination campaign has a strong effect on the lockdown intensity at the time of the vaccination approval, as at that very instant it is optimal to immediately (and disruptively) reduce the lockdown measures. In Scenario 1 it was optimal to intensify the lockdown at the vaccine approval, to keep more people in the susceptible compartment and to move them directly to the V compartment instead of letting them undergo the COVID-19 infection process. This effect is undermined by the vaccination speed, since vaccinations work quicker than the infections in this scenario. Of course, the lockdown is not ended immediately, but decreases (continuously) to zero rapidly after the jump at the switch. This corresponds to the third case of Theorem 2.2, which means

that the shadow price of a person getting infected at the switch (Stage 2) is lower than that without the switch (Stage 1) times the density for it.

The drastic continuous decrease of the lockdown after the vaccine approval in Stage 2 gets obvious in panel (b), showing the lockdown intensity along the duration. As in Scenario 1, the blue line corresponds to the lockdown over time if the vaccine is approved within three months, the red one for approval within six months, and the green one for approval within nine months. In all three cases, we see that the lockdown is ended about one month after the vaccine approval, irrespective of the lockdown intensity at the switch. That means that the lockdown should only be ended completely when almost all people have received the vaccination.

Panel (c) illustrates the research boost in Stage 1. Compared to scenario 1, it is now at the maximum for almost one year (until the susceptibles are about 40% of the population). This is a clear consequence of a cost-benefit analysis of the corresponding effects. In Scenario 1 the population is vaccinated within about 10 months, which means that also after the R&D breakthrough considerably high lockdown costs arise. In the current scenario the vaccination administration is so successful that the lockdown can be relaxed shortly after the approval and completely dismissed soon after. Therefore, this implies much lower lockdown costs, and thus it is optimal to allocate as many financial resources as possible to increase the R&D success rate and consequently decrease the expected costs of the pandemic.

The probability to remain in Stage 1 and the probability density for the vaccine approval (both plotted in panel (d)) are changed according to the optimal path of the research efforts of panel (c). The black line (probability to stay in Stage 1) is lower than that of Scenario 1, while the colored one (probability density for approval) is higher.

The susceptibles (panel (e)) and infected (panel (f)) complete the picture for this scenario. Unsurprisingly, the susceptible compartment goes to zero within one month. The speed of the vaccination campaign implies that the pandemic will be ended soon after the vaccine approval. Therefore the government can afford higher infection numbers (compared to Stage 1) for a short time interval, which are due to the downward jump and the sharp decrease of the lockdown intensity.

Anticipative versus non-anticipative behaviour

Within this subsection, we compare the optimal solution of scenario 1 (i.e., assuming a realistic speed of the vaccination campaign) to the case of a myopic government who does not anticipate a possible research breakthrough and the implied vaccine approval. This government just behaves as only Stage 1 would exist. If the vaccine is approved, the government (being surprised) immediately updates its decision and behaves according to the usual Stage 2 (as defined by Eq (2.19)).

Fig 2.7 compares some key results of Scenario 1 with those of a myopic government. Panel (a) plots the lockdown intensity, panel (b) the effective reproduction number, panel (c) and (d) susceptibles and infected over time. All panels focus on Stage 1 since the behaviour in Stage 2 is based on the same model and would overload the figure with too many lines.

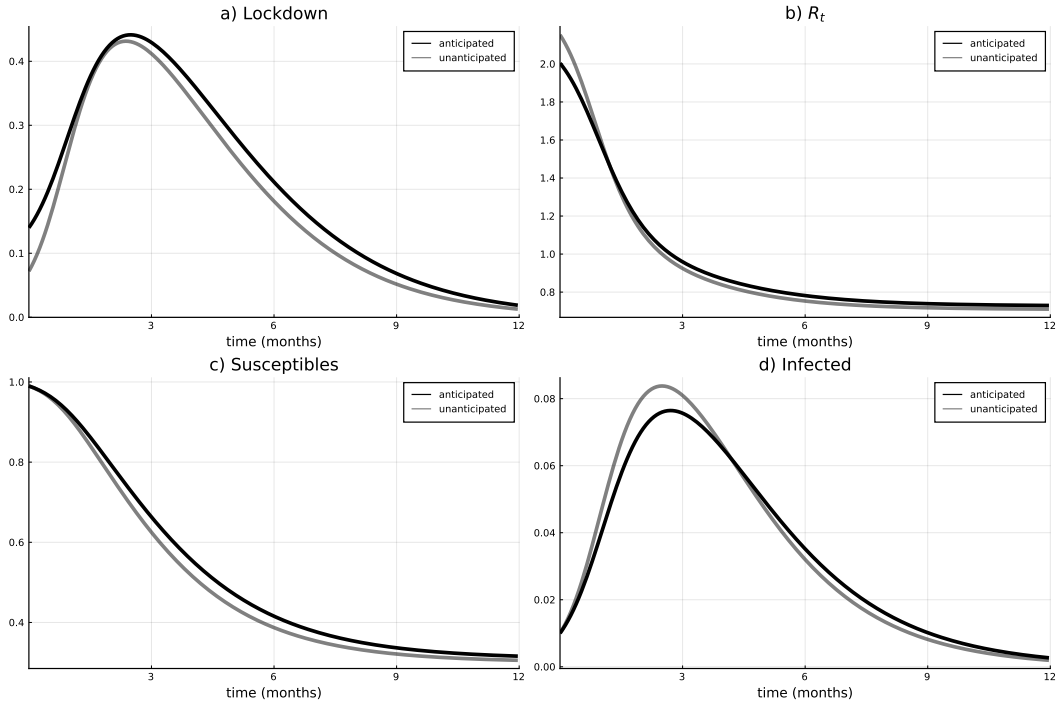


Figure 2.7: **Anticipation versus non-anticipation.** (a) Lockdown intensity in the anticipated and non-anticipated case, (b) Effective reproduction number in the anticipated and non-anticipated case, (c) Susceptibles, (d) Infected.

The optimal lockdown intensity (panel (a)) turns out to be more restrictive if the arrival of a vaccine is anticipated. This results from the analogous effect, which lets the lockdown jump upwards at the time of vaccine approval in Sce-

nario 1. I.e., if a vaccination can be expected, it is better to keep more people in the susceptible compartment so they can avoid the disease by getting the vaccine (if it becomes available). At the beginning of the pandemic, the lockdown intensity of the myopic case is only half of that of Scenario 1. Afterwards the difference gets close to zero at the sharp increase before the peak of the pandemic (around $t \approx 1.5$ months), finally, the difference increases again after the peak. Optimal lockdown intensities become more similar during the sharp increase since within this period the main driver of the lockdown intensity is not the current number of infected, but the dynamic nature of the pandemic (i.e., the SI-term in the dynamics). This implies that the lockdown is intensified in order to diminish this snowball effect, i.e., reducing high infection numbers (and therefore high costs due to lost lives) in the future. After that peak, the lockdown is again more intense in case of anticipation, but the difference is not as big as at the start of the pandemic, i.e., about 10%. Naturally, the difference tends to zero as the pandemic reaches its end.

The effective reproduction rate (panel (b)) is lower at the beginning of the pandemic for the anticipated case due to a more restrictive lockdown. This is due to the definition of $R_{t,i}(t)$ (see Eq (2.25)), where a strict lockdown (entering the numerator) decreases the fraction. However, already before the end of the first 1.5 months the relation switches and the unanticipated case has a lower effective reproduction number (together with a higher level of currently infected). This is due to the closeness of the lockdown intensities and due to the difference in the number of susceptibles. The latter effect dominates the effect of $\varphi(I_i)$ in the denominator (which increases the fraction).

Panels (c) and (d) mirror the trajectories of the lockdown intensity and effective reproduction rate. Initially, infections are higher in the unanticipated case due to the intensity of the lockdown, but they decrease after the peak of the pandemic. The number of susceptibles in the non-anticipated case does not catch up with that of the anticipated case, since the effective reproduction number gets lower when the number of infections already decreases. The intuition behind this is similar to the one given in the subsection presenting Scenario 1. Expecting a vaccine leads to the government trying to keep people healthy and to enable them to take the vaccine before they suffer the disease. The non-anticipative case causes more lost lives at the peak of the pandemic. If no vaccine becomes available, that might be the optimal solution. However, if a vaccine arrives, these additional lives are lost unnecessarily.

Several model runs with different parameters concerning the switching rate show qualitatively analogous results. However, the scale of the difference turns out to be sensitive. Especially, when a vaccine can be expected to be approved earlier (i.e., downward shift of the $Z_1(t)$ trajectory) the difference is remarkably bigger.

Robustness check

In the definition of the model and in the previous section we proposed some simple functional specifications. As a result, we were able to reveal the causes of the qualitative behavior without any disturbances implied by involved (though more realistic) functions. Certainly, one may argue different assumptions and more complex functional forms, which raise questions on the validity of our key messages.

For a robustness check, we therefore analyzed the model with (i) convex-concave lockdown costs, (ii) time-dependent lockdown costs, and (iii) a finite time horizon (with different lengths). We report the obtained results here below.

(i) Convex-concave lockdown costs. In the sections "The model" and "Numerical results", (c.f., Eq (2.13)), we propose convex lockdown costs due to the interconnectedness of the economy and the assumption that lockdown interventions are undertaken in order of cost-effectiveness. Although we believe that this assumption is reasonably argued, there might exist circumstances where initially increasing marginal costs of lockdown stringency $\ell(t)$ (i.e., convex part of the cost function) are followed by decreasing ones, when going beyond a certain threshold (i.e., concave part of the cost function). The intuition is that for a low $\ell(t)$ it becomes more and more difficult to find substitute inputs. Beyond a certain point, however, the economy is so constrained by low production levels, that further closures would have a decreasing effect on the implied costs.

To verify the validity of our core results also under an S-shaped/convex-concave formulation, we assumed the following alternative form of the lockdown costs

$$c_\ell(\ell(t)) = \frac{w}{3}(1 - \cos(\pi\ell(t))). \tag{2.28}$$

Function $c_\ell(\cdot)$ is convex for $\ell(t) \in [0, 0.5)$ and concave for $\ell(t) \in (0.5, \bar{L}]$. In Fig 2.8 we are plotting the optimal lockdown intensities and the research efforts over time for the vaccine availability of Scenario 1. In general, the qualitative nature

of the optimal solution remains unchanged. At the time of vaccine approval, the lockdown has to be intensified for a short period, followed by an earlier relaxation afterward. However, due to the convex part of the cost function, the lockdown curve is not continuous: it jumps around the turning point of the cost function since it is not optimal to remain in the interior of the concave part of the cost function. Consider e.g., the blue curve. It is better to stay a little bit longer on a lower lockdown level and then immediately jump to $\bar{L}$. Also, the relaxation of the lockdown is initiated by a downward jump followed by a continuous change in the following (i.e., on the convex part of the cost function). Optimal research efforts are adapted accordingly.

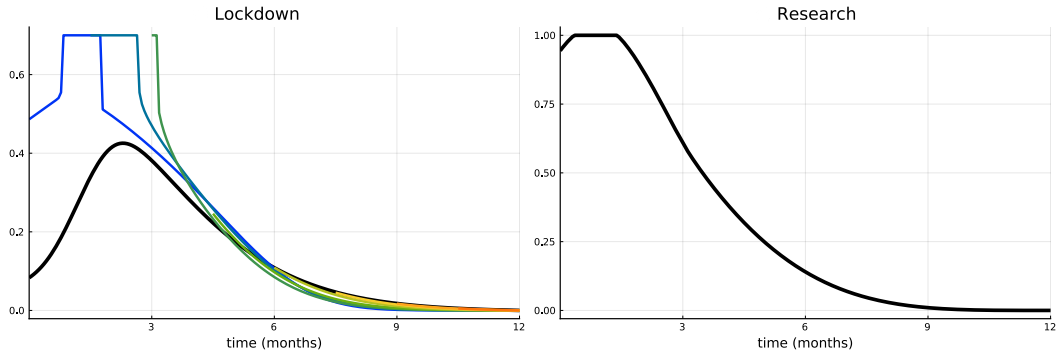


Figure 2.8: **Convex-concave lockdown costs.** Lockdown intensity over time (left) and research effort over time (right).

Note that in general concave costs in an optimal control model frequently imply several problems like, e.g., chattering control (see e.g., Grass et al. [52]). However, in this example, a solution is possible since the cost function is convex only on the highest part and a direct jump to the maximum lockdown level solves the issue.

(ii) Time-dependent lockdown costs. For the second part of the robustness check we relax the assumption of time-independent lockdown costs. In the beginning of the pandemic, firms could use reserves or inventories of inputs to partially absorb the supply problems. After some time, as reserves or inventories are exhausted, economic constraints become binding. Therefore, we introduce an additional factor to the cost function, which increases over time, i.e.,

$$c_\ell(\ell(t)) = e^{3t} w \ell^2(t). \tag{2.29}$$

For $t = 0$ costs are identical to our benchmark and after e.g., six months they are multiplied by a factor $e^{3 \cdot 0.5} \approx 4.48$ (the factor e^{3t} has been chosen for illustration purposes only and does not reflect realistic data). Increasing t makes lockdown more costly. Fig 2.9 plots the optimal lockdown intensity over time and research effort (with the rest of the parameters being equal to Scenario 1). Our core results still hold, but are slightly adapted. Increasing lockdown costs over time naturally implies that the lockdown will tend to be lower (compared to the benchmark). This, in particular, holds for the upward jump at the approval of a vaccine. While at early approvals the jump is evident, later on (about four months after) the discontinuity almost disappears, since during this time the lockdown has already become very expensive. The optimal research efforts also decreased considerably, which is a result of more people getting immunized through infection. Lower lockdown intensity implies more infections at the beginning, but a lower number of susceptibles and a higher number of recovered thereafter.

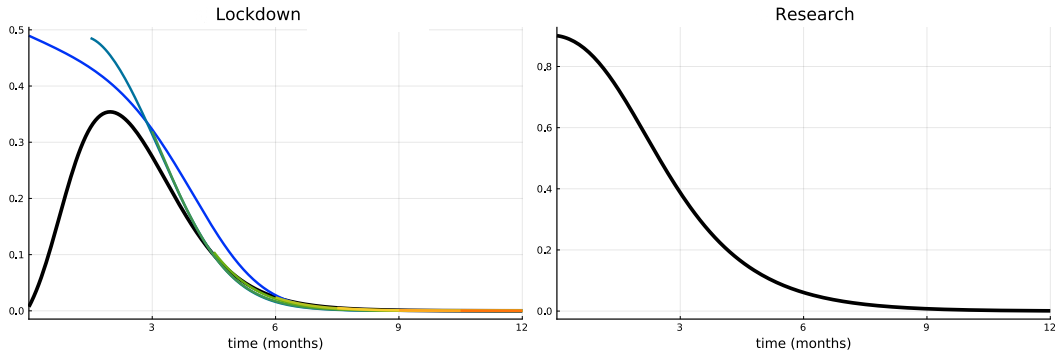


Figure 2.9: **Time-dependent lockdown costs.** Lockdown intensity over time (left) and research effort over time (right).

(iii) Effect of time horizon. As a third robustness check, we vary the time horizon. In our basic setup and in the "Numerical results" we have chosen an infinite time horizon (as used in e.g., Acemoglu et al. [2]), which corresponds to socially optimal decisions (the costs of the entire duration of the pandemic are evaluated). However, in reality, democratically elected governments are in place until the next election. While several papers, therefore, assume a finite time horizon with a zero salvage value function (see e.g., Alvarez et al. [4], or Aspri et al. [6]), we believe that a salvage value function is important. The government has the responsibility for the epidemic situation when elections

take place (see e.g., Caulkins et al. [25, 27], or Freiberger et al. [45]).

For illustration, we assume that the decision-maker evaluates a proxy of the expected health costs of the infected individuals at the end of the time horizon T. This is governed by the following salvage value function

$$\mathcal{S}(I(T), T) = e^{-\rho T} c_h(\bar{I}(T)) \qquad \text{where} \qquad \bar{I}(T) := I(T) \int_T^\infty e^{-\gamma t}\, dt = \frac{e^{-\gamma T}}{\gamma} I(T). \tag{2.30}$$

Individuals are leaving the I-compartment with rate γ. To account for the future time spent in the I-state by the final $I(T)$ infected individuals, $I(T)$ captures the approximate duration-adjusted infected. Deriving the health economic cost of this number is a proxy for the implied costs.

Fig 2.10 plots the optimal lockdown intensity for time horizons of six months (left panel), nine months (middle panel), and one year (right panel). In the case of a time horizon of one year (or longer) the results are practically identical to those of the infinite time horizon (see Scenario 1) since after one year the number of susceptibles is very low in our setting. Reducing the time horizon to nine months is enough to see a slight change in the solution. Decreasing the time horizon further finally reveals a drastic change of the optimal policy. The left panel shows the situation for six months. The lockdown does not abruptly increase but instead drops down at the vaccine approval, because it is not possible to profit from the following earlier relaxation during the remaining short time horizon. Moreover, it is not possible to significantly reduce the pandemic burden to a low level within this time frame. Therefore it is more efficient to save costs in terms of the lockdown.

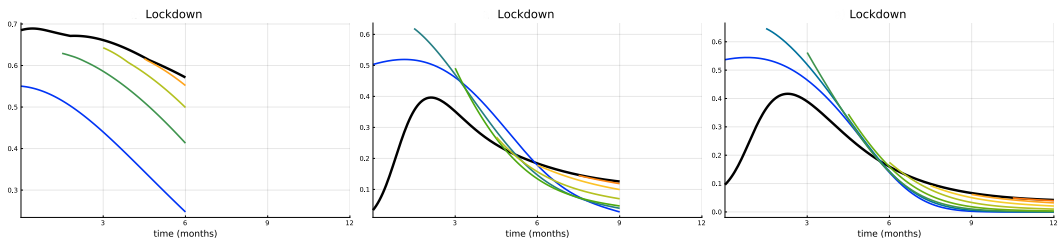


Figure 2.10: **Finite time horizon.** Lockdown intensity with time horizon of six months (left), nine months (middle) and one year (right).

In a nutshell, the difference between the results of the finite time and the infinite time horizon is more pronounced the shorter the time horizon. I.e., policy-makers tend to deviate from the socially optimal solution the sooner

their legislation period ends.

2.1.4 Conclusions

Overview

The current model extends the existing epidemiological models by specifying how a vaccine and its arrival are included in the optimization process. Using our approach we are able to overcome the following two crucial limitations: the *(un)known arrival/approval time* of the vaccine and the *continuation of the COVID-19 pandemic* during a vaccination campaign until enough people are immunized.

The formal analysis (see Theorem 2.2) and the numerical results show that the lockdown intensity (which is assumed to vary continuously in our model) jumps at the approval of the vaccine. In a realistic scenario, where the whole population will be vaccinated within 10 months, the lockdown is even intensified for a short period followed by an earlier relaxation due to the reduction of the susceptibles. This is consistent with the result of Garriga et al. [49], which is obtained using a different model setup and focus. However, it contrasts with other papers (e.g., Fu et al. [48]), which claim that lockdown and vaccination are always substitutes. The main driver behind this upward jump is that more lives can be saved, if more people are kept susceptible (by intensifying the lockdown) and can move directly to the vaccinated compartment. In an artificial scenario, where the vaccination campaign works very fast, i.e., vaccination of the population within one month, the adaptation of the lockdown at the approval date is reversed. The effect of the vaccine dominates the effect discussed above, which saves both lockdown costs, by relaxing the lockdown intensity, and lives, by the rapid reduction of the susceptibles.

A comparison of the optimal solution with the myopic case, where the government does not include the potential approval of an effective vaccine at all, shows that the expectation of a vaccine leads to a stricter lockdown policy. The explanation is similar to the jump at the vaccine approval date: A vaccine enables susceptibles to surpass the infection. Thus a "well-ordered" flow from the susceptibles to the infected, keeping the number of deaths at an acceptable level, becomes unnecessary at the time the vaccine becomes available.

The section "Robustness check" illustrates variations of the assumptions concerning the lockdown costs (convex-concave, time-dependent) and the time horizon. Our key findings turn out to be robust w.r.t. changes in the lockdown

costs and for long finite time horizon. For a short time horizon, the qualitative solution structure changes. It is no longer optimal to increase but to decrease the lockdown at the vaccine approval.

By additional numerical calculations (not included in this paper) we presume that the qualitative results of our analysis are robust against other parameter changes. In all variations, the key message remains that the vaccine availability shapes the structural characteristics of the solution. This is shown by two different scenarios.

Limitations

The paper suffer from a number of limitations, which propose further extensions of the model.

In Stage 2 of our model, it is assumed that the vaccination is 100% effective. While the mRNA vaccines of Biontech/Pfizer and Moderna are relatively close with more than 94−95% effectiveness, other vaccines are less effective. Moreover, this effectiveness has been tested by large studies a couple of months before the corresponding approval date. In the meantime the virus has mutated, and this might have also changed the vaccine's effectiveness. For the "α" and "δ"-variants (earlier called "British variant" or B.1.1.7 and "Indian variant" or B.1.617), which have replaced the original strain of the virus in Europe, the currently approved vaccines seem to remain quite effective (at least in preventing a heavy course of the disease). However, virologists expect also "escape mutations" to arise in the near future (the Omicron-variant took over in late 2021/early 2022). As a consequence, it seems to be realistic that a COVID-19 vaccine booster (adapted to new variants) will be necessary on a yearly basis, as it is common for the influenza virus. Including this into our model means adding more compartments and flows between them. This will certainly add new effects to the optimal solution. To assess their impact in further steps, it is crucial to understand the base effect within the current model first.

Another critical assumption of the current model is that the effectiveness of the lockdown remains constant over time, independently of its duration and intensity. Observing that the resentment in the population is growing since fall 2020 in many countries this assumption seems to be problematic. This so-called "lockdown fatigue" effect is thoroughly addressed in Caulkins et al. [26, 27]. The analysis shows that the shape of the optimal lockdown over time need not

be hump-shaped, but can have several peaks. Even several lockdown periods are possible. However, these papers assume a finite time horizon without vaccination, which corresponds to the myopic scenario in our paper. Therefore, the models are not directly comparable to our setup, and our model should be enriched with the lockdown fatigue effect to investigate this important question.

Finally, note that our paper abstracts from the possibility of testing and tracing (including quarantine) strategies. This is, of course, another important modeling option that might make the results more realistic and comparable to the real world. And indeed, the adaptation could straightforwardly be done by adding additional compartments (for identified people) and a parameter or a control variable (for testing efforts). This route has been followed in Freiberger et al. [45] again without vaccination. The analysis shows that testing has a strong effect on the pandemic if tracing is efficient. In the case of inefficient tracing, the result is basically equal to a solution without any testing. We have chosen to remain in our setup, since the focus of this paper is clearly different and allows a straightforward identification of the additional channel by the stochastic arrival of the vaccine.

2.1.5 S1 Appendix

Age-structured formulation of the model (2.17)–(2.19)

As presented in Wrzaczek et al. [118] the optimal control problem (2.17)–(2.19) with random switching time can be transformed into the following deter-

ministic age-structured form: Let us recall that $R = N - (S + I + V)$.

$$
\begin{aligned}
&\min_{\ell_1(t),\ell_2(t,\tau),r_1(t)} \int_0^\infty e^{-\rho t}\left[Z_1(t)g_1(I_1(t),\ell_1(t),r_1(t)) + Q(t)\right]\ dt \\
&\text{s.t. } \dot{S}_1(t) = -\beta(\ell_1(t))\frac{S_1(t)I_1(t)}{N_1(t)} \\
&\qquad \dot{I}_1(t) = \beta(\ell_1(t))\frac{S_1(t)I_1(t)}{N_1(t)} - \gamma I_1(t) \\
&\qquad \dot{N}_1(t) = -\varphi(I_1(t))I_1(t) \\
&\qquad \dot{Z}_1(t) = -\eta(r_1(t),t)Z_1(t) \\
&\partial_t S_2(t,\tau) = -\beta(\ell_2(t,\tau))\frac{S_2(t,\tau)I_2(t,\tau)}{N_2(t,\tau)} - \alpha(t-\tau)\frac{S_2(t,\tau)}{S_2(t,\tau)+R_2(t,\tau)}\mathbb{I}_{S_2(t,\tau)+R_2(t,\tau)>0} \\
&\partial_t I_2(t,\tau) = \beta(\ell_2(t,\tau))\frac{S_2(t,\tau)I_2(t,\tau)}{N_2(t,\tau)} - \gamma I_2(t,\tau) \\
&\partial_t N_2(t,\tau) = -\varphi(I_2(t,\tau))I_2(t,\tau) \\
&\partial_t V_2(t,\tau) = \alpha(t-\tau)\mathbb{I}_{S_2(t,\tau)+R_2(t,\tau)>0} \\
&\partial_t Z_2(t,\tau) = 0 \\
&\qquad Q(t) = \int_0^t Z_2(t,s)g_2(I_2(t,s),\ell_2(t,s))\ ds
\end{aligned}
\tag{2.31}
$$

with the following initial conditions:

$$
\begin{aligned}
&S_1(0) = S_0, \quad I_1(0) = I_0, \quad N_1(0) = N_0, \quad Z_1(0) = 1, \\
&S_2(\tau,\tau) = S_1(\tau), \quad I_2(\tau,\tau) = I_1(\tau), \quad N_2(\tau,\tau) = N_1(\tau), \quad V_2(\tau,\tau) = 0, \\
&Z_2(\tau,\tau) = Z_1(\tau)\eta(r_1(\tau),\tau).
\end{aligned}
\tag{2.32}
$$

In this formulation, $Z_2(t,\tau)$ is an auxiliary state variable denoting the probability density of the switching time evaluated at τ.

$Q(t)$ denotes the sum of all costs for all possible vaccine approval times up to time t weighted by the probability of their realization. With these two additional state variables the standard age-structured Maximum Principle, as presented e.g., in Feichtinger et al. [41], can be applied, with the advantage of treating Stage 1 and 2 simultaneously. We follow the following four standard steps: (1) formulating the Hamiltonian (for the corresponding maximization problem), (2) deriving the first-order conditions, (3) deriving the adjoint equations, (4) analyzing the transversality conditions.

(1) Hamiltonian:

Let us define the state variables $x(t) = (S_1(t), I_1(t), N_1(t), Z_1(t))$ in Stage 1 and $\tilde{y}(t,s) = (S_2(t,s), I_2(t,s), N_2(t,s), V_2(t,s), Z_2(t,s))$ in Stage 2.

The controls in the two stages are $v(t) = (\ell_1(t), r_1(t))$ and $u(t,s) = \ell_2(t,s)$ respectively.

In order to apply the PMP necessary conditions in Feichtinger et al. [41], we make the change of variables $a = t - s$. Note that these conditions in general hold only for finite time horizon problems. However, since the characteristic lines of the age-structured optimal control problem do not interact, the PMP necessary conditions also hold for infinite time.

Let us define the age-dependent state variable

$$y(t,a) = (x(t), \tilde{y}(t,a)),$$

the aggregated variable

$$q(t) = (\tilde{x}(t), Q(t)) = \int_0^\omega h(t,y,u)\, da$$

where, given $G_2 = Z_2 g_2$,

$$Q(t) = \int_0^t G_2\, ds = \int_0^t G_2 \boxed{da} = \int_0^\omega G_2\, \mathbb{I}_{a<t}\, da$$

and

$$\tilde{x}(t) = \int_0^\omega \frac{1}{\omega} x(t)\, da.$$

From now on, let $h = (G_2\, \mathbb{I}_{a<t},\ \frac{1}{\omega}x(t))$.

The dynamics function of the state variable $y = (x, \tilde{y})$ is

$$f(q,y,t,a) = (f_1(\tilde{x}, v, t), f_2(\tilde{y}, u, t, a)).$$

The border conditions (BC) at $a = 0$ which corresponds to the diagonal $s = t$ (where φ is the BC of $\tilde{y}$)

$$\phi(q,v,t) = (\tilde{x},\ \varphi(\tilde{x}, v, t)).$$

The payoff function (where $G_1 = Z_1 g_1$) is

$$L(q, v, t) = \frac{1}{\omega} e^{-\rho t} \left(G_1(\tilde{x}, v, t) + Q(t) \right).$$

The adjoint functions (ξ, ζ) associated to (y, q) satisfy the following system:

$$\begin{cases} -\left(\partial_t + \partial_a\right) \xi(t,a) = \nabla_y L(t,a) + \xi(t,a) \nabla_y f(t,a) + \zeta(t) \nabla_y h(t,a) \\ \xi(T,a) = 0,\ \xi(t,\omega) = 0 \\ \zeta(t) = \xi(t,0) \nabla_q \phi(t) + \int_0^\omega \left[\nabla_q L(t,a) + \xi(t,a) \nabla_q f(t,a) + \zeta(t) \nabla_q h(t,a) \right] da \end{cases}$$

Recalling that $y = (x, \tilde{y})$, we decompose the adjoint variable $\xi = (\xi_x, \xi_{\tilde{y}})$ from which we get our adjoint variables:

$$\begin{cases} \lambda(t) = \int_0^\omega \xi_x(t,a)\, da \\ \xi = \xi_{\tilde{y}}. \end{cases}$$

Combining the two systems above, we obtain

$$\begin{cases} -\dot{\lambda}(t) = \nabla_{\tilde{x}} \mathcal{H}_1 \\ -\left(\partial_t + \partial_a\right) \xi(t,a) = \nabla_{\tilde{y}} \mathcal{H}_2 \end{cases}$$

or, equivalently, letting $a = t - s$,

$$\begin{cases} -\dot{\lambda}(t) = \nabla_{\tilde{x}} \mathcal{H}_1 \\ -\partial_t \xi(t,s) = \nabla_{\tilde{y}} \mathcal{H}_2 \end{cases}$$

where the present Hamiltonian functions are

$$\begin{aligned} \mathcal{H}_1 &:= e^{-\rho t} G_1(t, \tilde{x}, v) + \lambda f_1(t, \tilde{x}, v) + \xi(t,0) \varphi(t, \tilde{x}, v) \\ \mathcal{H}_2 &:= e^{-\rho t} G_2(t, a, \tilde{y}, u) + \xi f_2(t, a, \tilde{y}, u). \end{aligned}$$

Reformulating the problem as a discounted maximum problem these con-

ditions turn out to be equivalent to the following system

$$\begin{cases} \dot{\lambda}(t) = \rho\lambda(t) - \nabla_{\tilde{x}}\mathcal{H}_1^C \\ \partial_t\xi(t,s) = \rho\xi(t,s) - \nabla_{\tilde{y}}\mathcal{H}_2^C \end{cases}$$

with current-value Hamiltonian functions

$$\begin{aligned} \mathcal{H}_1^C = & - Z_1\left(c_\ell(\ell_1) + \psi\varphi(I_1)I_1 + c_r(r_1)\right) \\ & - \lambda_S\beta(\ell_1)\frac{S_1I_1}{N_1} + \lambda_I\left(\beta(\ell_1)\frac{S_1I_1}{N_1} - \gamma I_1\right) - \lambda_N\varphi(I_1)I_1 - \lambda_Z\eta(r_1,t)Z_1 \\ & + \xi_S(t,t)S_1 + \xi_I(t,t)I_1 + \xi_N(t,t)N_1 + \xi_V(t,t)\cdot 0 + \xi_Z(t,t)\eta(r_1(t),t)Z_1 \end{aligned} \tag{2.33}$$

$$\begin{aligned} \mathcal{H}_2^C = & - Z_2\left(c_\ell(\ell_2) + \psi\varphi(I_2)I_2\right) + \xi_S\left(-\beta(\ell_2)\frac{S_2I_2}{N_2} - \alpha\frac{S_2}{S_2+R_2}\mathbb{I}_{S_2+R_2>0}\right) \\ & + \xi_I\left(\beta(\ell_2)\frac{S_2I_2}{N_2} - \gamma I_2\right) - \xi_N\varphi(I_2)I_2 + \xi_V\left(\alpha\mathbb{I}_{S_2+R_2>0}\right) + \xi_Z\cdot 0 \end{aligned} \tag{2.34}$$

(2) Adjoint equations:

$$\begin{aligned} \dot{\lambda}_S &= \rho\lambda_S + (\lambda_S-\lambda_I)\,\beta(\ell_1)\frac{I_1}{N_1} - \xi_S(t,t) \\ \dot{\lambda}_I &= (\rho+\gamma)\lambda_I + \left(\lambda_N + Z_1\psi\right)\left(\varphi'(I_1)I_1 + \varphi(I_1)\right) + (\lambda_S-\lambda_I)\,\beta(\ell_1)\frac{S_1}{N_1} - \xi_I(t,t) \\ \dot{\lambda}_N &= \rho\lambda_N - (\lambda_S-\lambda_I)\,\beta(\ell_1)\frac{S_1I_1}{N_1^2} - \xi_N(t,t) \\ \dot{\lambda}_Z &= \rho\lambda_Z + c_\ell(\ell_1) + \psi\varphi(I_1)I_1 + c_r(r_1) + (\lambda_Z - \xi_Z(t,t))\,\eta(r_1,t) \\ \partial_t\xi_S &= \rho\xi_S + (\xi_S-\xi_I)\,\beta(\ell_2)\frac{I_2}{N_2} + \xi_S\alpha\frac{R_2}{(S_2+R_2)^2}\mathbb{I}_{S_2+R_2>0} \\ \partial_t\xi_I &= (\rho+\gamma)\xi_I + (\xi_S-\xi_I)\,\beta(\ell_2)\frac{S_2}{N_2} + (\xi_N - Z_2)\,\psi\left(\varphi'(I_2)I_2 + \varphi(I_2)\right) \\ \partial_t\xi_N &= \rho\xi_N - (\xi_S-\xi_I)\,\beta(\ell_2)\frac{S_2I_2}{N_2^2} \\ \partial_t\xi_V &= \rho\xi_V \\ \partial_t\xi_Z &= \rho\xi_Z + c_\ell(\ell_2) + \psi\varphi(I_2)I_2 \end{aligned}$$

(3) First-order and second-order conditions: Taking the derivatives of the Hamil-

tonian function $\mathcal{H}_1^C$ w.r.t. the Stage 1 controls ℓ_1, r_1 and of the Hamiltonian function $\mathcal{H}_2^C$ w.r.t. the Stage 2 control ℓ_2, the first-order conditions (FOC) for positive controls are

$$\begin{aligned} 0 &= -Z_1 c_\ell'(\ell_1) + (\lambda_I - \lambda_S)\,\beta'(\ell_1)\frac{S_1 I_1}{N_1} \\ 0 &= -c_r'(r_1) + \eta'(r_1, t)\,(\xi_Z(t,t) - \lambda_Z) \\ 0 &= -Z_2 c_\ell'(\ell_2) + (\xi_I - \xi_S)\,\beta'(\ell_2)\frac{S_2 I_2}{N_2} \end{aligned} \tag{2.35}$$

The second-order conditions (SOCS), which guarantee that the first-order conditions lead to a maximum of the Hamiltonian function, are obtained by imposing that the Hessian matrix of the Hamiltonian function is negative definite. This Hessian matrix is diagonal, since the mixed derivatives are zero, so the SOCS are

$$\begin{aligned} &-Z_1 c_\ell''(\ell_1) + (\lambda_I - \lambda_S)\,\beta''(\ell_1)\frac{S_1 I_1}{N_1} < 0 \Longleftrightarrow -2Z_1 w + (\lambda_I - \lambda_S)\,2\beta_0\theta^2\frac{S_1 I_1}{N_1} < 0 \\ &-c_r''(r_1) + \eta''(r_1, t)\,(\xi_Z(t,t) - \lambda_Z) < 0 \Longleftrightarrow -2c_0 < 0 \\ &-Z_2 c_\ell''(\ell_2) + (\xi_I - \xi_S)\,\beta''(\ell_2)\frac{S_2 I_2}{N_2} < 0 \Longleftrightarrow -2Z_2 w + (\xi_I - \xi_S)\,2\beta_0\theta^2\frac{S_2 I_2}{N_2} < 0 \end{aligned} \tag{2.36}$$

If $(\lambda_I - \lambda_S) < 0$ and $(\xi_I - \xi_S) < 0$, then the SOCS conditions (2.36) trivially hold.

(4) Transversality conditions: Catching up optimality, i.e., $\liminf_{T\to+\infty} J_T(u^*) - J_T(u) \geq 0$, implies $\liminf_{T\to+\infty} e^{-\rho T}\lambda(T)(x^*(T) - x(T)) \geq 0$ in a standard (time-dependent) optimal control model (using $x(t)$ and $\lambda(t)$ for the state and adjoint variable respectively). Although limiting transversality conditions for an infinite time horizon of an age-structured optimal control model can not be shown in general, they hold in the current model since the characteristic lines are isolated (i.e., do not interact) and can be viewed as trajectories of independent optimal control models.[4] Therefore we have

[4]This holds for all age-structured optimal control models obtained by a transformation of a multi-stage optimal control model with random switching time obtained by the transformation presented in Wrzaczek et al. [118].

$(\forall \tau > 0)$

$$
\begin{aligned}
&\liminf_{T\to+\infty} e^{-\rho T}\lambda_S(T)(S_1^*(T) - S_1(T)) \geq 0\\
&\liminf_{T\to+\infty} e^{-\rho T}\lambda_I(T)(I_1^*(T) - I_1(T)) \geq 0\\
&\liminf_{T\to+\infty} e^{-\rho T}\lambda_N(T)(N_1^*(T) - N_1(T)) \geq 0\\
&\liminf_{T\to+\infty} e^{-\rho T}\lambda_Z(T)(Z^*(T) - Z(T)) \geq 0\\
&\liminf_{T\to+\infty} e^{-\rho T}\xi_S(T,\tau)(S_2^*(T,\tau) - S_2(T,\tau)) \geq 0\\
&\liminf_{T\to+\infty} e^{-\rho T}\xi_I(T,\tau)(I_2^*(T,\tau) - I_2(T,\tau)) \geq 0\\
&\liminf_{T\to+\infty} e^{-\rho T}\xi_N(T,\tau)(N_2^*(T,\tau) - N_2(T,\tau)) \geq 0\\
&\liminf_{T\to+\infty} e^{-\rho T}\xi_V(T,\tau)(V^*(T,\tau) - V(T,\tau)) \geq 0\\
&\liminf_{T\to+\infty} e^{-\rho T}\xi_Z(T,\tau)(Z_2^*(T,\tau) - Z_2(T,\tau)) \geq 0\\
&\liminf_{T\to+\infty} e^{-\rho T}\zeta(T)(Q^*(T) - Q(T)) \geq 0.
\end{aligned}
\tag{2.37}
$$

2.2 Communication strategies to contrast anti-vax action: a differential game approach

Article written by A. Buratto, R. Cesaretto, and M. Muttoni; published in the Central European Journal of Operations Research (2024) [17].

2.2.1 Introduction

Vaccination is one of the greatest discoveries of modern medicine. Thanks to vaccines, diseases such as poliomyelitis, tetanus, smallpox, diphtheria, and rubella have been eradicated in many countries. Furthermore, vaccination has recently been shown to reduce the incidence of some other diseases, for example, human papillomavirus (HPV) infection (Takla et al. [102]), meningitis (Buonomo et al. [15]), and some forms of cancer. The public health benefits brought by such discoveries are so important that several countries have even devised mandatory childhood vaccination policies.[5,6] Committees on Vaccination have been constituted in many countries to plan communication campaigns in favor of vaccination, such as the (STIKO) in Germany by Robert Koch-Institut[7] and the European Centre for Disease Prevention and Control[8] have done.

However, people often do not perceive the importance of vaccines, either because some of the diseases that were eradicated are no longer visible or because their effects may show up only after long periods. In case a given vaccination is not mandatory, without memory of the damage the related disease can cause, the perceived risks of vaccination among some people have begun to outweigh their perceived benefits (Omer et al. [87], Buonomo et al. [16]). Some people focus only on the risk of side effects, which for them appears to be extremely high compared to the risk associated with contracting the disease (Salmon et al. [94]). Furthermore, vaccine efficacy has recently been debated by skeptics who try to spread the idea that vaccines are ineffective and even dangerous (see, e.g., Shim et al. [99], Carrillo and Lopalco [23], and Hotez [63]).

In turn, media, such as magazines, television, the Internet, and social net-

[5] `https://www.immunize.org/laws/`, retrieved on 2023/06/19
[6] `https://ijponline.biomedcentral.com`, retrieved on 2023/06/19
[7] `https://www.rki.de/EN/Content/infections`, retrieved on 2023/06/19
[8] `https://www.ecdc.europa.eu/en`, retrieved on 2023/06/19

works, often present news related to vaccines without submitting them to strict verification by scientific and health authorities. Sometimes, they spread alarming news without any foundation, in the worst cases even claiming an association between vaccines and serious diseases. Just to mention, even though Andrew Wakefield's claim about a causal relation between vaccines and autism was refuted by the scientific community (see Taylor et al. [104]), such a conjecture has caused a decline in vaccination coverage, especially in certain countries.

Due to these fake news, some diseases were taken too lightly and gave rise to the vaccine hesitancy effect (see, e.g., Bozzola et al. [12] and White et al. [117]). Shim et al., in [99], assert that if the great benefits to society of measles vaccination are to be maintained, the public must be educated about these benefits in order to increase public confidence. When organizing vaccine administration, national health systems must consider many aspects, and in the last decade the issue of correct communication about the vaccination campaign has become crucial. The authors observe that the effectiveness of vaccination programs can be jeopardized by public misperceptions of vaccine risk. They illustrate "the importance of public education on vaccine safety and infection risk to achieve vaccination levels that are sufficient to maintain herd immunity."

In this paper, we study the effect of a vaccination advertising campaign in supporting countries to increase the coverage of measles, rubella, and other vaccines. The idea arises from a recent report by the World Health Organization (WHO), which underscores the impact of the COVID-19 pandemic on surveillance and immunization efforts. The report highlights that the suspension of immunization services has decreased the estimates of immunization coverage for many infectious diseases. In fact, during the last three years, the COVID-19 pandemic stopped surveillance and immunization efforts, putting many children at risk for preventable diseases. "Approximately 25 million infants missed at least one dose of the measles vaccine through routine immunization in 2021."[9] Globally low immunization rates increased the chances of outbreaks and endanger unvaccinated children. The World Health Organization just decided to work towards regional measles elimination by strengthening immunization programs (e.g., Measles and rubella strategic framework: 2021-2030) and implementing effective surveillance systems.

[9] `https://www.who.int/news-room/`, retrieved on 2023-06-19

Perceiving the same need for a focused public educational system, this paper aims to formalize in a mathematical context the problem of planning a pro-vax communication campaign to convince people to get vaccinated. In what follows, we assume that the vaccination we are dealing with is not mandatory.

Our research questions are as follows.

- How does negative word of mouth affect the evolution toward herd immunity?
- How can the pro-vax communication campaigns of the national health-care system and of pharmaceutical firms speed up such an achievement?
- How can a proper pro-vaccination campaign sustain the vaccination efficiency of the national health-care system?

The remainder of the paper is organized as follows. In Sect. 2.2.2 we briefly review the literature and position our contribution. In Sect. 2.2.3 we introduce our model for a controlled evolution of the unvaccinated population, together with the cost functionals associated with such an attempt. We formalize the problem in a differential game framework. In Sect. 2.2.4 we determine the optimal communication strategies that constitute the associated steady-state feedback Nash equilibrium. Sect. 2.2.5 presents some sensitivity analysis of the solution together with interpretations of the results. Sect. 2.2.6 concludes.

2.2.2 Brief Literature background

A consistent stream of literature tackles the issue of controlling infectious diseases based on epidemic models, in particular, on behavioral Susceptible-Infected-Recovered (SIR) models. In the book by Manfredi and d'Onofrio [77] a vast literature on vaccination and other influences of human behavior on the spread of infectious diseases is presented. A detailed report reviewing models that account for behavioral feedback and/or the spatial/social structure of the population can be found in Wang et al. [116]. More recent publications among the same stream of literature are, e.g., d'Onofrio and Manfredi [34] and Buonomo et al. [15]. An interesting approach to the vaccination problem in the context of game theory is tackled in Shim et al. [99], where a game-theoretic model of disease transmission and vaccination is formalized as a population game. Here, the game-theoretic epidemiological analysis performed can yield insights into the interplay between anti-vaccine behavior, vaccine coverage, and disease dynamics. More recently in Matusik and Nowakowski [80] a game-theoretical

approach has been used to model the control of COVID-19 transmission, always in the context of SIR dynamics.

The so-called word-of-mouth effect plays a crucial role in the spread of antivax beliefs. Bauch in [8] studies the strategic interaction between individuals when they decide whether or not to vaccinate, using an imitation dynamic game. The role of word of mouth in voluntary vaccination planning is presented in Bhattacharyya et al. [9], where the synergetic feedback between word of mouth and the epidemic dynamics controlled by voluntary vaccination is analyzed. The authors present an epidemiological model with a social learning component that incorporates the reciprocal influence of population groups, as well as the feedback that can occur from the incidence of diseases. They model this social interaction through a game-theoretical framework using the concept of payoff, adapted from applications of Game Theory to Economics.

Just because anti-vax behavior is often associated with word of mouth, rather than with scientific data and information, it seems interesting to tackle the problem not necessarily from an epidemiological point of view but from a sociological point of view. Therefore, the communication approach derived from the theory of dynamic advertising models, described in Huang et al. [64], suggests a correct communication policy to defeat the spread of the anti-vax movement. In a communication framework, El Ouardighi et al. in [35] consider two different types of word of mouth: negative (or adverse) and positive (or favorable) according to the different reactions of satisfied and dissatisfied customers. The authors stress how negative word of mouth is more influential than positive, especially for brands with which potential customers are not familiar, and this could be the case for the vaccine issue.

Within the literature stream based on dynamic advertising models, Grosset and Viscolani in [55] formulate and solve an optimal control problem to determine a provaccination communication campaign that contrasts the effect of antivaccination word of mouth, with the objective of minimizing the care cost induced by unvaccinated people and the cost resulting from the communication campaign. The authors propose an upper bound to the final number of unvaccinated people to guarantee herd immunity. In such a model the programming interval is fixed, while in Grosset and Viscolani [54] the authors study a variable-final-time optimal control problem to stress the importance of reaching the given herd immunity threshold, rather than reducing costs in a fixed-length programming period.

In Buratto et al. [19], a model for the aforementioned national health problem is proposed using a linear-quadratic differential game in a finite time horizon. The model aspires to understand how an optimal interaction between the communication campaigns of the healthcare system and of a pharmaceutical firm that produces a given vaccine can help increase vaccination coverage. The healthcare system wants to minimize the number of unvaccinated people at a minimum cost. The pharmaceutical firm aims to maximize its profits while reducing the number of unvaccinated people.

This paper belongs to the last group of models presented above, named *Communication models*, where the focus is on the communicative approach and considers the "educational plan to vaccination" as a possible communication strategy that can be planned by both the national healthcare system and pharmaceutical firms. The references cited above (Grosset and Viscolani [55, 54], and Buratto et al. [19]) assume that the unvaccinated population, affected by negative word of mouth, diverges to infinity if it is not supported by a pro-vaccine campaign. Here, we adopt a more realistic dynamics where the *à-régime* number of unvaccinated people is increased by the adverse action of anti-vax negative word of mouth and converges to a finite value, which can be deduced from the annual report published by the World Health Organization (WHO).[10]

Moreover, the literature cited above considers only finite-time horizon optimal control problems, while in this paper, being interested in a long-term plan for the provaccination campaign, we consider the plan over an infinite horizon.

We study the interaction between the communication campaigns of the healthcare system and of a pharmaceutical firm that produces a vaccine for a given disease, formulating a differential game played à la Nash. We use the Hamilton-Jacobi-Bellman approach to determine a Markovian Nash equilibrium. In recent years, with the introduction of big data solutions, healthcare management can rely on accurate and up-to-date reports. Mobile contact track&trace apps, recently urged by the COVID-19 emergency, provide constant information on the number of vaccinated (and consequently of unvaccinated) individuals.[11] Therefore, feedback strategies based on the number of unvacci-

[10]To be precise, WHO reports annually the report on global vaccination coverage for each infectious disease in `https://www.who.int/news-room`. The number of unvaccinated people for a given disease can be deduced by subtracting the WHO declaration from the size of the population interested in the related vaccination.

[11]see, e.g., `https://thevaccineapp.com/`, retrieved on 2023/06/19

nated are not only credible but also more reliable than Open-Loop ones, as the former are time-consistent. It is commonly known, as reported also in Della Marca and d'Onofrio [31], that in health economics, it is preferable, whenever possible, to control a system by means of a feedback-based strategy, which aims to minimize the economic and human burden of some health-related phenomena. On the basis of these considerations, we look for communication strategies that constitute a Markovian Nash equilibrium.

2.2.3 The Model

We assume that the population to which vaccination is devoted is stationary. The dynamics we analyze describes the evolution of the unvaccinated population, as is done in the related literature cited above (see Grosset and Viscolani [55, 54], and Buratto et al. [19]).

For this purpose, let $x(t)$ be the number of unvaccinated individuals at time t, and let x_0 be its value at the initial time.

The World Health Organization annually publishes data related to global trends and the total number of reported cases of vaccine-preventable diseases (VPD).[12] In an idealistic vaccination system, the number of unvaccinated should decrease and tend to zero (neglecting people who cannot be vaccinated due to more serious immunodeficences). However, it can be observed from WHO's data that there is always a significantly high average number of unvaccinated. This residual number of unvaccinated can be attributed to the presence of antivax action. On the other hand, in each time unit, a fixed percentage of the unvaccinated population voluntary gets the vaccine, convinced by peers and networks, or even by the spread of the disease. We can formalize this type of scenario with the following dynamics of the unvaccinated people

$$\dot{x}(t) = w - rx(t). \tag{2.38}$$

The term $w > 0$ represents the constant number of people that at a given time may be considered as new-unvaccinated, because they've just entered into the subset of population ought to be vaccinated and did not get the vaccine. The reasons for such a decision can be varied, including the negative information

[12] `https://www.who.int/news-room/fact-sheets/detail/immunization-coverage`, retrieved on 2023/06/19

about vaccines and vaccination that *no-vax* movement spreads throughout a word-of-mouth mechanism. For simplicity, in what follows we will call the parameter w a word-of-mouth parameter.

The parameter $r > 0$ represents the instantaneous vaccination rate. It measures the intensity of vaccination and in some way it describes the efficiency of the national vaccination system: the faster the national healthcare system in the vaccination issue, the higher r. From the dynamics we can observe that the number of unvaccinated increases with w, however, it decreases due to the effect of spontaneous vaccinations. Unlike in Grosset and Viscolani [55], where the number of unvaccinated tends to explode in the absence of a communication action, due to negative word of mouth, here we assume, in a more realistic formalization, that the unvaccinated group converges to a finite positive value. Indeed, as t goes to infinity in dynamics (2.38), the number of unvaccinated tends to the following *à-régime* level

$$\hat{x}_{SS} = \frac{w}{r}. \tag{2.39}$$

This value, which increases in w and decreases in r, is in any case finite, and it corresponds to the nonnegative residual of unvaccinated people that can be deduced by the annual World Health Organization report.

However, as Wang et al. stress in [116], there are scenarios in which voluntary vaccination is not sufficient to provide herd immunity. If word of mouth w is too high or the effectiveness of vaccination r is too low, then $\hat{x}_{SS}$ may be too large; far from the level requested for herd immunity. This issue may become too expensive for the national healthcare system to sustain; therefore, a pro-vaccination communication campaign is needed.

In this paper, we formulate a dynamic model assuming to be exactly in the situation in which the negative word of mouth is so high that a pro-vaccination campaign is necessary. Since we are interested in the scenario where communication is necessary, in what follows we will assume that the parameter w is greater than a given threshold, which guarantees nontrivial equilibrium strategies. The value of such a threshold will be specified later in the next section.

We assume that both the national healthcare system (S) and the pharmaceutical firm (F) are independently planning their own communication campaign to promote vaccination. Let $\phi_S(t) \in U_S$ and $\phi_F(t) \in U_F$ be the pro-vax advertising efforts of the national healthcare system and the pharmaceutical company, re-

spectively. We assume that feasible communication strategies are non-negative and reasonably bounded. In what follows, we assume that the upper bounds for such strategy functions are sufficiently high. This will permit us to concentrate on the characteristics of non-trivial inner solutions. Let $\delta_S, \delta_F > 0$ be the effectiveness of the two communication intensities, as in Buratto et al. [19]. The evolution of the unvaccinated is affected by these controls according to the following dynamics

$$\begin{cases} \dot{x}(t) = w - rx(t) - \delta_S \phi_S(t) - \delta_F \phi_F(t), \\ x(0) = x_0. \end{cases} \tag{2.40}$$

At first glance, we can observe that if the number of unvaccinated x is zero at a given time, then a positive communication effort may drive the state function below zero, and this would not be meaningful. However, it will be proved that under optimal conditions, given our assumption of a significantly high value for the variable w, the positivity of the state function is ensured.

After describing the dynamics, let us focus on the payoff functions of the two players. The national healthcare system and the pharmaceutical firm both seek to minimize their respective costs associated with the number of unvaccinated individuals, along with reducing their communication costs. For any $t \geq 0$, we assume the following cost flows for the two players respectively

$$\mathrm{C_S(t)} = \frac{\beta}{2} x^2(t) + \frac{\kappa_S}{2} \phi_S^2(t), \qquad \mathrm{C_F}(t) = \theta x(t) + \frac{\kappa_F}{2} \phi_F^2(t) \tag{2.41}$$

For what concerns the costs due to the number of unvaccinated individuals, from the national healthcare system point of view, the number of unvaccinated individuals affects national healthcare costs because a relevant percentage of them need medications and sometimes even hospitalization. While in Grosset and Viscolani [55, 54] these costs are assumed to be proportional to the number of unvaccinated people and therefore linear in the variable x, in our model we assume that national healthcare costs are quadratic and convex, formalized by the term $\frac{\beta}{2} x^2(t)$ (with $\beta > 0$), to underscore the significant expenses associated with hospital therapy. Furthermore, this assumption highlights the crucial point that vaccine refusal not only puts those who decline vaccination at risk but also increases the chances of disease transmission for people who interact with unvaccinated individuals. A similar quadratic assumption for national

health system costs due to the number of unvaccinated has also been considered in Buratto et al. [19], where a linear-quadratic differential game is formulated and solved, although in a finite time horizon.

From the pharmaceutical firm point of view, the number of unvaccinated affects the revenue of the pharmaceutical firm, although in a different way. Assuming $\theta > 0$ to be the unit profit of a given vaccine, the consequent missed revenue for each unsold vaccine is here formalized as the linear cost $\theta x(t)$.

Finally, both players sustain the communication costs associated to their pro-vax advertising efforts, here assumed to have the following quadratic and convex formulation $\frac{k_S}{2}\phi_S^2(t)$ and $\frac{k_F}{2}\phi_F^2(t)$ (with $k_S > 0$ and $k_F > 0$) for the healthcare system and the firm respectively.

All the models cited above neglect the cost of vaccination, as we also do; nevertheless, few papers in the related literature evaluate vaccination costs; among them, we mention the real option approach used in Favato et al. [37].

We consider the problem over an infinite horizon. In such a setting, taking into account the dynamics (2.40), and discounting the costs $C_S(t)$ and $C_F(t)$ in (2.41), we formulate the following differential game played à la Nash[13]

healthcare System (S)	pharmaceutical Firm (F)
$\min_{\phi_S \in U_S} \int_0^{+\infty} e^{-\rho t}\left(\frac{\beta}{2}x^2(t) + \frac{\kappa_S}{2}\phi_S^2(t)\right) dt$	$\min_{\phi_F \in U_F} \int_0^{+\infty} e^{-\rho t}\left(\theta x(t) + \frac{\kappa_F}{2}\phi_F^2(t)\right) dt$

$$\begin{cases} \dot{x}(t) = w - rx(t) - \delta_S\phi_S(t) - \delta_F\phi_F(t), \\ x(0) = x_0. \end{cases} \tag{2.42}$$

The table below collects the meaning of the parameters:

[13]Observe that in this formulation in infinite horizon we are discounting the costs that appear in Buratto et al. [19], obviously neglecting the "residual value functions".

ρ	discount rate ($\rho > 0$)
β	health care and social cost ($\beta > 0$)
θ	missed profit due to each unsold vaccine ($\theta > 0$)
κ_S, κ_F	communication cost parameters ($\kappa_S, \kappa_F > 0$)
w	word-of-mouth coefficient ($w > 0$)
r	instantaneous constant vaccination rate,
δ_S, δ_F	pro-vax communication effectiveness ($\delta_S, \delta_F > 0$)

It is interesting to observe the asymmetry of the game due to the different types of costs associated with the unvaccinated group. This asymmetry will emerge in the different forms of the value functions of the two players and, consequently, in the equilibrium strategies of (S) and (F), respectively. The problem above is formulated over an infinite horizon and is autonomous (because there is no other explicit time dependence in its formulation, apart from the discount factors). Therefore, we are interested in looking for a stationary solution. Stationarity means that each player's strategy is determined as a function of the state variable only: $\phi_j^*(x)$ $j \in \{S, F\}$ as stated in Dockner et al. [33, p.210].

2.2.4 The solution (Markovian Nash equilibrium)

We are interested in feedback advertising strategies based on the number of unvaccinated individuals, so we look for a Markovian Nash equilibrium that, being the game autonomous, turns out to be subgame perfect (see Dockner et al. [33, p.105].

Proposition 2.3. *The Markovian Nash Equilibrium Feedback strategies that represent the pro-vax communication strategies are*

$$\phi_S^*(x) = \frac{1}{\delta_S}(\eta - (\rho + r)) \left[x + \left(w - \frac{\delta_F^2 \theta}{\kappa_F \eta} \right) / \eta \right], \tag{2.43}$$

$$\phi_F^*(x) = \frac{\delta_F \theta}{\kappa_F \eta}, \tag{2.44}$$

where

$$\eta = \sqrt{\left[\frac{\rho}{2} + r\right]^2 + \beta\frac{\delta_S^2}{\kappa_S}} + \frac{\rho}{2}. \tag{2.45}$$

Proof. See Appendix 1.

It can be immediately observed that $\phi_F^*(\cdot)$ is strictly positive (being $\delta_F, \theta, k_F > 0$, and $\eta > \rho + r > \rho > 0$). Furthermore, $\phi_F^*(\cdot)$ does not depend on x, therefore it constitutes a degenerate feedback strategy. It represents the constant positive communication contribution of the pharmaceutical firm to contrast the antivax action.[14] From now on, we will denote it by ϕ_F^*. As an economic interpretation, since the firm is interested in minimizing its missed profits due to unvaccinated, its optimal communication strategy is adopted at a positive constant rate, independently of the number of unvaccinated individuals.

Remark 2.4. For what concerns the communication strategy of the national healthcare system $\phi_S^*(\cdot)$, it is easy to prove that $\phi_S^*(x) > 0$ for all feasible x, since $\eta > \rho + r > 0$. Observing that if $w > \frac{\delta_F^2\theta}{\kappa_F\eta}$, then $\phi_S^*(x) > 0$ for all $x \geq 0$, we can conclude that if the negative word of mouth is so high that it compromises herd immunity, as assumed in the problem formulation, then the national healthcare system needs to support the firm's pro-vaccination campaign.

Steady State

These kinds of problems are known in the literature as "Discounted Autonomous Infinite Horizon Models" (DAM) (see Grass et al. [52, p.159]). The main computational effort to solve optimal control models of this type is calculating the stable manifolds of the occurring saddles. Let us substitute the optimal strategies (2.43) and (2.44) in the dynamics of the unvaccinated individuals, then we obtain

$$\dot{x}(t) = w - rx(t) - \delta_S\phi_S^*(x) - \delta_F\phi_F^* = -(\eta - \rho)x(t) + \frac{(\rho + r)}{\eta}\left(w - \frac{\delta_F^2\theta}{\kappa_F\eta}\right). \tag{2.46}$$

[14] Mathematically, such a constant structure could be predicted from the linear form of the value function of player F.

Proposition 2.5. *The steady-state number of unvaccinated people, while adopting the optimal pro-vax communications turns out to be*

$$x_{SS} = \frac{(\rho + r)\left(w - \frac{\delta_F^2 \theta}{\kappa_F \eta}\right)}{\eta(\eta - \rho)}. \tag{2.47}$$

Proof. Differentiating the dynamics in (2.46), we obtain the optimal state trajectory

$$x^*(t) = x_{SS} + (x_0 - x_{SS})e^{-(\eta-\rho)t} = x_{SS}(1 - e^{-(\eta-\rho)t}) + x_0 e^{-(\eta-\rho)t}. \tag{2.48}$$

It is easy to verify that $x^*(t) \geq 0$ for any starting point $x_0 \geq 0$ and, since $\eta > \rho$, the state converges to the stable positive steady state x_{SS} in (2.47). □

As expected, the steady-state level of unvaccinated individuals under the actions of pro-vax communication campaigns of the two involved players increases with negative word of mouth w, and it is lower than the *à-régime* unvaccinated level observed in (2.39) without pro-vax communications. In fact, for any set of values of the problem parameters, it holds

$$x_{SS} \leq \hat{x}_{SS} = \frac{w}{r}. \tag{2.49}$$

As a result, in case the word of mouth is too high, then the joint effects of the pro-vax communication campaigns (2.43) and (2.44) may contribute to lower the number of unvaccinated and to bring it down to the stationary level (2.47). This will help to reach herd immunity.

2.2.5 Sensitivity analysis and simulations

In the previous section, we obtained the analytical form for the strategies that constitute the stationary Markovian Nash equilibrium of the pro-vax communication game and the associated steady-state level of unvaccinated people. The simple dependence of the solution on some of the parameters that characterize the problem allows us to perform a sensitivity analysis.

In Table 2.3, we report the existing monotonicity properties, when analytically computable, of the equilibrium strategies at the steady state, of the steady state level of the unvaccinated, and of the optimal costs for the two players. For the reader's convenience, since the optimal solution contains the constant

η, in the first column of the table, the dependencies of the constant η on all the analyzed parameters are also included. The arrows indicate either increasing monotonicity ($\nearrow$) or decreasing monotonicity ($\searrow$), while the lines "—" mean constant behavior.

	η	ϕ_S^*	ϕ_F^*	x_{SS}	$Cost_S$	$Cost_F$
w	—	$\nearrow$	—	$\nearrow$	$\nearrow$	$\nearrow$
θ	—	$\searrow$	$\nearrow$	$\searrow$	$\searrow$	$\nearrow$
δ_F	—	$\searrow$	$\nearrow$	$\searrow$	$\searrow$	$\searrow$
κ_F	—	$\nearrow$	$\searrow$	$\nearrow$	$\nearrow$	$\nearrow$

Table 2.3: Sensitivity analysis with respect to main parameters

Impact of Word-of-Mouth Effectiveness w (Table 2.3)
Parameter w takes into account the word-of-mouth effects on the evolution of the unvaccinated people. Here we assume, as stated in the previous section, that the negative word of mouth is so high that it requires nonzero pro-vax communications, more precisely $w > \frac{\delta_F^2 \theta}{\kappa_F \eta}$.

The effect of the negative word of mouth w drives the national healthcare system to increase its pro-vax communication strategy (ϕ_S^*), with increased associated costs. Moreover, as expected, the steady state number of unvaccinated (x_{SS}) increases in w. These results are in line with the ones obtained in Buratto et al. [19], where the final level of unvaccinated people $x(T)$ is increasing in the word-of-mouth parameter. Another expected result is that both players have increased costs as anti-vax word of mouth increases. However, it becomes interesting to note that the optimal pharmaceutical strategy ϕ_F^* is not dependent on w; the firm conducts its communication campaign independently of word of mouth because its cost is not an effective loss, but a missed income.

Impact of vaccine unit profit θ (Table 2.3)
Parameter θ represents the unit profit of a vaccine and is considered as the virtual cost of each unsold vaccine. From Table 2.3 we can observe that the higher the unit profit, the higher the strategy communication of the firm (ϕ_F^*) and consequently its cost. With a higher pro-vax campaign, the steady state level of unvaccinated individuals (x_{SS}) decreases. At the same time, the national healthcare system can count on the firm's action, so, as strategic substitutes do, it can reduce its communication campaign (ϕ_S^*) and, therefore, its total cost.

Impact of communication effectiveness and marginal costs (δ_F, κ_F) of the firm (Table 2.3)
It is straightforward to prove that ϕ_F^* increases in δ_F and decreases in κ_F: The more effective (or less costly) the firm pro-vax campaign, the more intensive it will be. This result is typical for dynamic advertising models (see Huang et al. [64]): each optimal communication strategy increases in its corresponding communication efficacy and decreases in its corresponding marginal cost. Moreover, since both strategies act jointly to decrease the number of unvaccinated people, they are strategic substitutes, so that each optimal communication strategy decreases with the communication efficacy of the other player and increases with the communication marginal cost of the other player. In particular, ϕ_S^* is decreasing in δ_F: The more effective the firm's pro-vax campaign, the less the national healthcare system needs to implement its own pro-vax communication. Finally, the steady state (x_{SS}) computed in (2.47) decreases in δ_F, while it increases in the marginal cost κ_F.

Impact of communication strategies to communication effectiveness and marginal costs (δ_S, κ_S) of national health-care system (Table 2.4)
Let us observe in Table (2.4) that η is increasing in δ_S and decreasing in κ_S, therefore ϕ_F^*, which is decreasing in η (see (2.44)), turns out to be decreasing in δ_S and increasing in κ_S. Furthermore, ϕ_S^* turns out to be decreasing in κ_S, while no kind of monotonicity nor regularity can be verified with respect to δ_S, (a "no" appears in the corresponding cell of Table 2.4).

	η	ϕ_S^*	ϕ_F^*
δ_S	$\nearrow$	no	$\searrow$
κ_S	$\searrow$	$\searrow$	$\nearrow$

Table 2.4: Sensitivity analysis with respect to δ_S and κ_S

Sensitivity of communication strategies to social cost β and vaccination rate r (Table 2.5)
The parameter β represents the social cost for the national healthcare system. Analytical results in Table 2.5 highlight how a severe social cost motivates the healthcare system to increase its pro-vaccination campaign, thus allowing the pharmaceutical firm to reduce its own.

	η	ϕ_S^*	ϕ_F^*
β	$\nearrow$	$\nearrow$	$\searrow$
r	$\nearrow$	no	$\searrow$

Table 2.5: Sensitivity analysis with respect to β and r

The parameter r represents the vaccination rate, in other words the efficiency of the national vaccination system. Analytical results highlight how an increase in the efficiency of vaccination in the national healthcare system allows the pharmaceutical firm to lighten its own communication campaign. On the other hand, no kind of monotonicity can be analytically proved in the national health-care system communication effort with respect to r.

Numerical analysis of the steady state

The dependence of the steady state x_{SS} with respect to the parameters δ_S, k_S, β, and r cannot be proved analytically, therefore, we performed some numerical simulations fixing, for each analysis, $n-1$ parameters and letting one parameter vary within its feasible values (such as to guarantee non-zero advertising strategies).[15] In the following figures 2.11–2.14, it is evident, as analytically proven in (2.49), that in the long run, the steady-state number of unvaccinated people while adopting pro-vaccination communication (x_{SS} - solid lines) is smaller compared to the values *à-régime* without such communication policies ($\hat{x}_{SS}$ - dotdashed lines).

All the simulations conducted demonstrate a consistent pattern, specifically a quasi-concave shape characterized by an initial convex-concave increase followed by a subsequent decrease. We can interpret these results by noting the presence of a *threshold effect*: Only sufficiently high values of the parameter with respect to which the analysis is carried out make it possible for the national health-care system to reduce the number of unvaccinated individuals to the extent necessary to achieve herd immunity.

It should be noted in Figure 2.14 that the dotdashed line, representing the à-régime unvaccinated level $\hat{x}_{ss}$ without the pro-vax communications, is not constant (as the corresponding dotdashed lines in the previous figures) because $\hat{x}_{ss}$, defined in (2.39), decreases in r. Furthermore, the gap between the two lines is large in correspondence to small values of r. This permits to conclude that if

[15]Most of parameter values come from Buratto et al. [19], as a reference

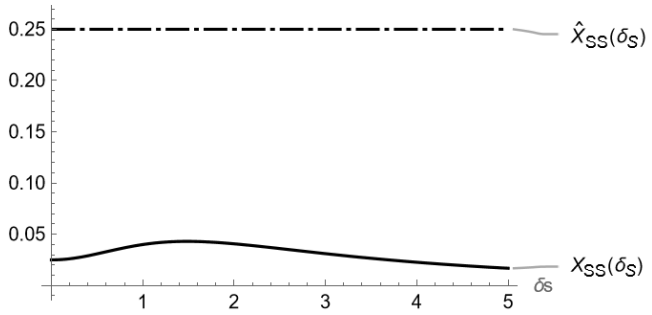


Figure 2.11: Steady state x_{SS} w.r.t. δ_S
($w = 0.1$, $\theta = 0.1$, $\kappa_S = 4$, $\delta_F = 0.6$, $\kappa_F = 0.8$, $\rho = 0.05$, $\beta = 0.3$, $r = 0.1$)

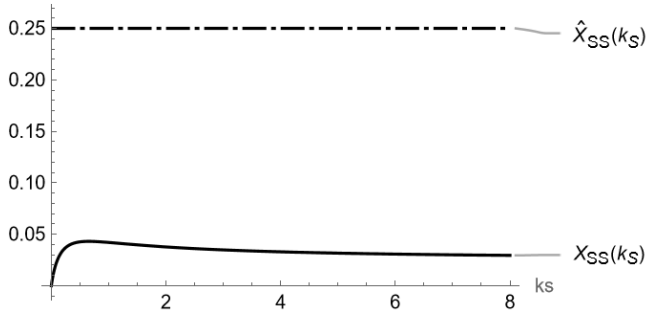


Figure 2.12: Steady state x_{SS} w.r.t. κ_S
($w = 0.1$, $\theta = 0.1$, $\delta_S = 0.6$, $\delta_F = 0.6$, $\kappa_F = 0.8$, $\rho = 0.05$, $\beta = 0.3$, $r = 0.1$)

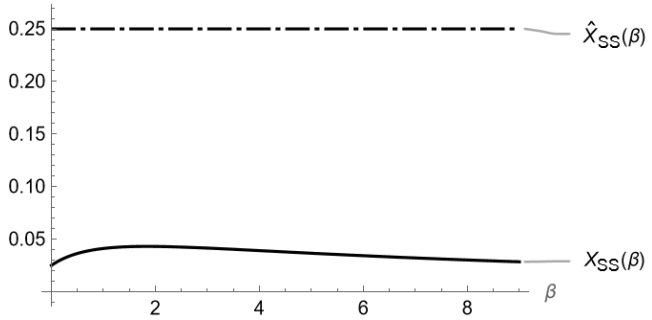


Figure 2.13: Steady state x_{SS} w.r.t. β
($w = 0.1$, $\theta = 0.1$, $\delta_S = 0.6$, $\kappa_S = 4$, $\delta_F = 0.6$, $\kappa_F = 0.8$, $\rho = 0.05$, $r = 0.1$)

the national healthcare system is characterized by a low vaccination rate, then investing in effective and targeted communication strategies becomes crucial in order to significantly slow down the number of unvaccinated individuals and mitigate the potential consequences.

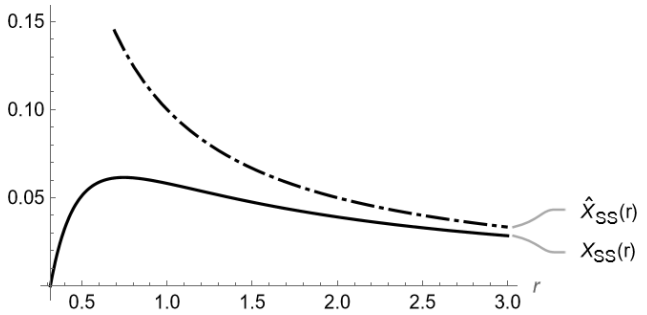


Figure 2.14: Steady state x_{SS} w.r.t. r
($w = 0.1$, $\theta = 0.1$, $\delta_S = 0.6$, $\kappa_S = 4$, $\delta_F = 0.6$, $\kappa_F = 0.8$, $\rho = 0.05$, $\beta = 0.3$)

2.2.6 Conclusion

Vaccination can reduce the incidence of many diseases, but unfortunately, a negative word of mouth based on fake news recently caused a decline in vaccination coverage in many countries. This phenomenon has increased the residual level of unvaccinated people reported each year by the World Health Organisation. In this paper, we tackle the problem of planning a pro-vaccination communication campaign to convince hesitant individuals to get vaccinated, so as to reduce the residual level of unvaccinated people.

We formulate and solve an asymmetric differential game model over an infinite horizon. Two players are involved in the provaccination communication campaign: the national healthcare system and a pharmaceutical firm. The Hamilton-Jacobi-Bellman approach is used to determine the optimal communication strategies that constitute a Markovian Nash equilibrium for the game. The two players act as strategic substitutes; the smaller the healthcare system campaign, the higher the firm's one. Sensitivity analyzes are performed with respect to the parameters of the problem to study their impact on the equilibrium strategies and steady-state solutions.

Let us answer the research questions declared in the Introduction: Our model confirms that the national healthcare system needs to increase its investment in vaccine communication to contrast the effect of negative antivax word of mouth and to aim at herd immunity. On the other hand, the firm's pro-vaccine communication campaign is not affected by negative word of mouth.

The proposed model also includes the vaccination rate as an index that characterizes the efficiency of the national healthcare system. As far as we are aware, such a parameter is not taken into account in the related literature. Our

results show that immunization can be obtained either by increasing vaccination efficiency, i.e., vaccinating at a high rate, or implementing a well-planned provaccination communication campaign. These results emphasize the importance of the "Immunization Agenda 2030 Measles & Rubella Partnership" (M&RP)[16] led by the American Red Cross, United Nations Foundation, Centers for Disease Control and Prevention (CDC), Gavi, the Vaccines Alliance, the Bill and Melinda French Gates Foundation, UNICEF and WHO, to achieve the IA2030 measles and rubella specific goals.

These results also confirm the importance of national healthcare management in maintaining a sufficiently high vaccination rate or, at least, in designing an efficient provaccination campaign.

Different extensions of this work can be envisioned. The first considers the game played à la Stackelberg, where the leader may be the national healthcare system and the follower the pharmaceutical company. A specific analysis of the cases where the optimal communication strategy can turn out to be zero is an idea that deserves investigation in order to take into account the various types of costs that the national healthcare system must bear.

The important role that parameter r plays in the results may suggest considering the vaccination rate as a decision variable to be optimally set by the national healthcare system together with the communication campaign for vaccination.

It can be interesting to analyze the stochastic evolution of the number of unvaccinated people. This goal could be obtained by introducing a stochastic effect in communication campaigns and changing the ordinary differential equation (2) into a stochastic one (see Wang et al. [116]).

2.2.7 Appendix

Let $V_S(x)$ and $V_F(x)$ be the stationary value functions associated with the healthcare system and the pharmaceutical firm, respectively. Let us assume that these functions are differentiable, and let us denote by $V_S'(x)$ and $V_F'(x)$ their derivatives w.r.t. x. They must solve the following Hamilton Jacobi Bellman equations associated to the problems of the two players

[16] `https://measlesrubellainitiative.org/learn/about-us`, retrieved on 2023/07/06

$$\rho V_S(x) = \max_{\phi_S \geq 0} \left\{ \left[w - rx - \delta_F \phi_F(x) - \delta_S \phi_S \right] V_S'(x) - \left[\frac{\beta}{2} x^2 + \frac{\kappa_S}{2} \phi_S^2 \right] \right\} \tag{2.50}$$

$$\rho V_F(x) = \max_{\phi_F \geq 0} \left\{ \left[w - px - \delta_F \phi_F - \delta_S \phi_S(x) \right] V_F'(x) - \left[\theta x + \frac{\kappa_F}{2} \phi_F^2 \right] \right\} \tag{2.51}$$

Maximising the r.h.s. of (2.50) and (2.51) with respect to ϕ_S and ϕ_F respectively, we obtain

$$\phi_S(x) = \max \left\{ 0, -\frac{\delta_S}{\kappa_S} V_S'(x) \right\}, \qquad \phi_F(x) = \max \left\{ 0, -\frac{\delta_F}{\kappa_F} V_F'(x) \right\}. \tag{2.52}$$

Being interested in a non-trivial Nash equilibrium, we look for the best response strategies in the region $\{x \in \mathbb{R} : V_S'(x), V_F'(x) \leq 0\}$, in such a case

$$\phi_i(x) = -\frac{\delta_i}{\kappa_i} V_i'(x), \quad i \in \{S, F\}. \tag{2.53}$$

After substituting (2.53) in (2.50) and (2.51), the HJB equations for the two players can be rewritten as the following system

$$\begin{cases} \rho V_S(x) = \left[w - rx + \dfrac{\delta_F^2}{\kappa_F} V_F'(x) \right] V_S'(x) - \dfrac{\beta}{2} x^2 + \dfrac{\delta_S^2}{2\kappa_S} (V_S'(x))^2, \\ \rho V_F(x) = \left[w - rx + \dfrac{\delta_S^2}{\kappa_S} V_S'(x) \right] V_F'(x) - \theta x + \dfrac{\delta_F^2}{2\kappa_F} (V_F'(x))^2. \end{cases} \tag{2.54}$$

Notice the asymmetry in the resulting equations with respect to the state variable x: This entails the corresponding value functions featuring the same asymmetry; therefore, we assume the following forms for V_S and V_F:

$$V_S(x) = \frac{1}{2} A x^2 + Bx + C, \qquad V_F(x) = Dx + E, \tag{2.55}$$

where the coefficients A, B, C, D, and E are assumed constant, as we focus on stationary strategies.

After substituting the value functions (2.55) into the HJB equations (2.54) and comparing the coefficients of the resulting polynomials in the variable x, the real coefficients A, B, C, D, and E satisfy the following system of algebraic

equations (known as algebraic Riccati equations)

$$\begin{cases} \frac{\delta_S^2}{2\kappa_S}A^2 - \left[\frac{\rho}{2} + r\right]A - \frac{\beta}{2} = 0 \\ \left[\rho + r - \frac{\delta_S^2}{\kappa_S}A\right]D + \theta = 0 \\ \left[\rho + r - \frac{\delta_S^2}{\kappa_S}A\right]B - \left(w + \frac{\delta_F^2}{\kappa_F}D\right)A = 0 \\ \rho E - \left[w + \left(\frac{\delta_S^2}{\kappa_S}B + \frac{\delta_F^2}{2\kappa_F}D\right)\right]D = 0 \\ \rho C - \left[w + \left(\frac{\delta_F^2}{\kappa_F}D + \frac{\delta_S^2}{2\kappa_S}B\right)\right]B = 0 \end{cases} \tag{2.56}$$

From the first equation we get the two solutions

$$A_\pm = \frac{\kappa_S}{\delta_S^2}\left\{\left[\frac{\rho}{2} + r\right] \pm \sqrt{\left[\frac{\rho}{2} + r\right]^2 + \beta\frac{\delta_S^2}{\kappa_S}}\right\},$$

where $A_+ > 0$ and $A_- < 0$. We discard the positive root because once substituted into the second equation in (2.56) to obtain the coefficient D, it would be incompatible with assumption $V_F'(x) \leq 0$. From now on, let us denote $A = A_- < 0$. Observe that if we define

$$\eta = \left[\rho + r - \frac{\delta_S^2}{\kappa_S}A\right] = \sqrt{\left[\frac{\rho}{2} + r\right]^2 + \beta\frac{\delta_S^2}{\kappa_S}} + \frac{\rho}{2}, \tag{2.57}$$

it is easy to prove that $\eta > \rho + r > \rho > 0$, and that the system (2.56) admits the following unique solution

$$\begin{cases} A = \frac{\kappa_S}{\delta_S^2}\left\{\rho + r - \eta\right\} < 0, \\ B = \left(w + \frac{\delta_F^2}{\kappa_F}D\right)A/\eta = \left(w - \frac{\delta_F^2\theta}{\kappa_F\eta}\right)A/\eta, \\ C = \left(w + \frac{\delta_F^2}{\kappa_F}D + \frac{\delta_S^2}{2\kappa_S}B\right)B/\rho, \\ D = -\theta/\eta < 0, \\ E = \left(w + \frac{\delta_F^2}{2\kappa_F}D - \frac{\delta_S^2}{\kappa_S}B\right)D/\rho. \end{cases} \tag{2.58}$$

Substituting these values into (2.55) we obtain the continuously differentiable value functions, whose derivatives are

$$V_S'(x) = Ax + B = A\left[x + \frac{1}{\eta}\left(w - \frac{\delta_F^2\theta}{\kappa_F\eta}\right)\right], \qquad V_F'(x) = D < 0. \tag{2.59}$$

Observe that indeed the constant A is negative; moreover, from the second equation, we can infer that if the word-of-mouth level w is high, $\left(w > \frac{\delta_F^2 \theta}{\kappa_F \eta}\right)$, then the positivity of the term inside the round brackets also guarantees the negativity of the constant B. This implies that if we restrict the domain to nonnegative values of x, which are those that have a physical meaning, then $V_S'(x) < 0$ for all $x \in \mathbb{R}^+$.

The Markovian Nash equilibrium strategies representing the optimal pro-vax communication efforts can be obtained in feedback form by substituting coefficients (2.58) into (2.59), and in turn into (2.53), obtaining

$$\phi_S^*(x) = -\frac{\delta_S}{\kappa_S}(A\ x + B) \qquad \phi_F^*(x) = -\frac{\delta_F}{\kappa_F} D.$$

Once substitutes the values of parameters A, B, C, and D, we obtain (2.43) and (2.44).

3

Climate Change Economics

This chapter contains a research report in the field of climate change economics, titled "How to prepare for and adapt to a climate tipping point" [83]. The monograph was written by M. Muttoni[1] under the supervision of S. Wrzaczek,[2] as the final report of a research project conducted over the three-month Young Scientists Summer Program at IIASA. It was published in the IIASA repository in 2023.

This work focuses on a continuous-time version of the renown DICE (Dynamic Integrated Climate-Economy) model by Nordhaus [86], modified to include the possibility of a climate tipping point. Climate tipping points are abrupt regime shifts in the Earths climate systems that are mainly driven by the anthropogenic increase in global mean temperature. Given the concrete possibility that increasingly frequent and severe climate-related catastrophes will negatively impact the world economy, how should we model our policies to optimally prepare for and adapt to such events?

We formulate a finite-horizon, two-stage optimal control problem with a stochastic switching time τ (representing the tipping point), we convert it to a vintage structure and address it with Pontryagin's Maximum Principle. The hazard rate of τ is assumed to be exclusively state-dependent, more specifically, increasing in the global mean temperature. Among our results, we find ana-

[1]Dipartimento di Matematica "Tullio Levi-Civita", Università degli Studi di Padova
International Institute for Applied Systems Analysis (IIASA)
[2]International Institute for Applied Systems Analysis (IIASA)

lytical conditions for the optimal savings and emission abatement policy, and a decomposition of the growth rate of the social cost of carbon.

The complexity of the model, determined by its relatively high dimensionality and involved dyamics, entails the use of numerical simulations to compute the optimal solution of the problem. We simulate four different tipping point scenarios (denoted A, B, C, and D), each characterized by different effects of the switch on the dynamics, the running utility, and the state jump. Referring to the notation introduced in Classification 1.2.1, the switch has the following effects:

(E3) **Dynamics:** scenarios A and C feature a change in the geophysical dynamics governing the atmospheric temperature and the carbon cycle, respectively; scenario D features a change in the economic production function, which influences the capital dynamics.

(E4) **Running utility:** as the economic production also enters the utility function, the change in production that characterizes scenario D also constitutes a change in the running utility function.

(E7) **State jump:** scenario B features the instantaneous destruction of a fraction of capital (which is a state variable), due to an environmental catastrophe occurring at the switching time.

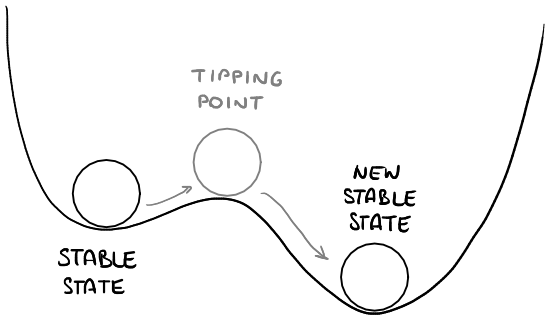
3.1 How to prepare for and adapt to a climate tipping point

3.1.1 Introduction

Tipping points

Consider a dynamical system in a stable state. By definition, a *tipping point* occurs when, by the influence of an external agent possibly triggering feedback mechanisms, the system shifts and crosses a threshold beyond which it is drawn away from its original state and into a different one.

Although the concept of tipping point can apply to any type of system shifting from one state to another as described, the scientific community is increasingly interested in *climate* tipping points, i.e., tipping points that occur in the Earth's climate system. This interest towards climate tipping points is fueled by the tangible possibility that increasingly frequent and severe climate-related



catastrophes will negatively impact global economy in the best case scenario, or even lead to human extinction in the worst. To assess resilience – i.e., the ability to recover from a shock in the long run – in any socio-economic and/or climate model, it is therefore key to take into consideration the occurrence of one, or more, tipping points.

The main driver behind climate tipping points is the anthropogenic increase in global mean temperature triggered by GHGs emissions and, according to recent IPCC Special Reports (2018, 2019), their thresholds could lie between 1 and 2řC of global warming above pre-industrial levels, instead of 5řC as it was previously thought (Lenton et al. [73]). It is therefore plausible that multiple tipping elements will occur within this century under anthropogenic climate change (Lenton et al. [72]).

To give some concrete examples, the most critical climate tipping points we are likely to face are, according to Lenton et al. [73], the melting of the Arctic sea ice and of the Greenland and Antarctic ice sheets, the slowdown of the Atlantic circulation, and several transformations of carbon sinks into carbon sources (e.g., destruction of the Amazon rainforest and of the Boreal forest, die-off of coral reefs, thawing of the Permafrost).

Addressing the matter only in the attempt to prevent the tipping point would be too demanding for the economy, and there is a strong possibility that actual tentative implementations of preventive policies may come as too little, too late. This is why we also need to consider building resilience to tipping points into the global system: how can we (optimally) change our policies and behaviors in preparation for and in adaptation to the occurrence of a climate tipping point?

We implement the problem in the Optimal Control environment: we suppose that one single policymaker controls the world's emission abatement and economic investments – which in turn influence the evolution of the Earth's cli-

mate system and of the population's goods and services – in order to maximize a social welfare function. The occurrence of a climate-related tipping point is modeled as a random instant whose hazard rate increases with global mean temperature, after which the system's dynamics are irreversibly changed. We then apply the Maximum Principle for heterogeneous systems to derive necessary conditions for the optimal savings and abatement policies.

The main advantage of treating this problem with Optimal Control Theory is being able to learn about the interplay of the different variables, actions, and factors, thus gaining more in-depth knowledge behind what constitutes an optimal solution. This precise disentanglement of the various effects, which cannot be achieved through other optimization techniques (e.g., classic nonlinear optimization), is important in order to correctly implement decisions and adaptations of decisions.

In our fragmented world, different countries have different priorities when it comes to climate and economy, especially considering the issue of climate justice: poor countries suffer the most from the climate change that has been mainly caused by past emissions of the rich countries. Because of this disparity, modeling one single policymaker may be unrealistic, but it results in the best possible *aggregated* behavior. Indeed, in decentralized models, this 'single planner' solution acts as a benchmark to tend to by setting the correct carbon taxes.

An objective limitation of our current work is the inability to introduce more than one tipping point into the system: in fact, due to the strong interconnections between different tipping points in the form of feedback mechanisms and domino effects, crossing even only one of them is likely to trigger a tipping point *cascade*. This phenomenon will be investigated in future research, as an application of a new theoretical framework to study multi-stage OC problems with a sequence of random switching times.

Research question

Suppose that the policymaker is aware of the risk of a tipping point, that they know how this risk grows with global temperature, and what effect it will have (it could be either on the Earth's climate system, or on economic production, or both).

Given this context, the main objectives of this study are the following. First:

we aim to derive the optimal emission abatement policy that should be implemented under the uncertainty of a pending climate catastrophe. Such a policy is expressed as the portion of reduced emissions as a function of time, and it can be adjusted upon occurrence of the tipping point. Second: we aim to analyze the evolution of Social Cost of Carbon over time and the different factors that contribute to its composition.

We will show how the expectation of different effects produces different behaviors in anticipation of the event, by considering four different types of tipping point: A) the melting of the Arctic sea ice, which compromises the self-cooling ability of the atmosphere, B) the destruction of capital due to an environmental catastrophe, which instantaneously decreases the available capital without bringing any long-term changes, C) the transformation of a carbon sink into a carbon source, which alters the rates of carbon uptake and release between the atmosphere and the biosphere/shallow ocean, and D) an increased impact of temperature on production, where the rate of economic production becomes less efficient.

Literature review

A few articles feature tipping points in theoretical dynamic approaches, and they can be divided into two main streams of literature: continuous-time consumption/pollution models and discrete-time integrated-assessment models based on DICE.

Among the stream of continuous-time consumption/pollution models, Cropper [29], Clarke and Reed [28], and Gjerde et al. [50] feature an unforeseen and sudden drop in society's level of consumption due to the environment's vulnerability to pollution-related catastrophic collapse. The drop in consumption may be complete or partial, reversible or irreversible. Tsur and Zemel [105, 106, 107] feature an unpredictable and detrimental event, associated with the greenhouse effect, that inflicts damage in the form of an instantaneous penalty upon occurrence. In particular, Tsur and Zemel [106] also features irreversible events, i.e., those which result in a truncation of the time horizon, and multiple-occurrence reversible events. In Cropper [29] and Tsur and Zemel [105] the uncertainty concerning the time of occurrence refers to our ignorance about the exact pollution level required to trigger the event, whereas in Clarke and Reed [28], Gjerde et al. [50], and Tsur and Zemel [106, 107] it stems from the intrinsic stochastic nature

of the environmental processes that control occurrence. In the latter case, the hazard rate may increase with the GHG stock (Clarke and Reed [28], Tsur and Zemel [106, 107]) or with global mean temperature change (Gjerde et al. [50]). Differently from all of the above, Ploeg and Zeeuw [110] features a productivity shock, whose hazard rate increases with the pollution stock. It may consist in a sudden drop in technological stock or in the production's response to temperature. Also recoverable catastrophes are considered, like a sudden rise in the carbon stock or a partial destruction of capital.

Among the stream of discrete-time integrated-assessment models based on DICE, Keller et al. [66] features a DICE model with the addition of central uncertainties and stochastic damages, caused by an uncertain environmental threshold imposed by an ocean circulation change, giving rise to a probabilistic optimization problem. Cai et al. [21] features an extension of the DICE model by Nordhaus featuring a stochastic economy and with the addition of uncertain economic impact of possible (multiple) climate tipping events. The latter enters the climate damage function as a discrete Markov chain, starting from zero in the pre-tipping stage, with nondecreasing fractional values over time. Lemoine and Traeger [71] includes a climate tipping point in Nordhaus's DICE model, which is then solved in a dynamic programming setting where the policymaker learns about a threshold's location by observing the system's response in each period. Lontzek et al. [75] features a stochastic IAM based on DICE where the likelihood of tipping points increases with global warming, but it is uncertain.

Our work aims to bridge the existing gap between these two streams by adapting a continuous-time formulation of DICE to include a tipping point.

3.1.2 Methods and Discussion

We implement a finite-horizon, two-stage Optimal Control problem with a random switching time: we suppose that one single policymaker controls the world's emission abatement and economic investments in order to maximize a social welfare function. These control variables enter the dynamical system that governs the evolution of the state variables describing the Earth's climate – such as temperature and carbon stock – and the population's goods and services.

The occurrence of a climate-related tipping point is modeled as a random instant whose hazard rate increases with global mean temperature. Before the switch, the system is a continuous-time version of the DICE model by W. Nord-

haus; after the switch, the system enters a new regime, that is characterized – depending on the nature of the tipping point – either by different dynamics or running payoff, or by a jump discontinuity in the state variables.

We then apply the Maximum Principle for heterogeneous systems developed by V. M. Veliov (see Veliov [113] and Wrzaczek et al. [118]) and derive necessary conditions for the optimal savings and abatement policies.

A feature of the Maximum Principle is that it involves the computation of the co-state trajectories: each one of these functions represents the evolution of the *shadow-value* of the corresponding state variable over time, i.e., the marginal increase in the optimal payoff for an instantaneous increase in the state variable. This plays a crucial role in the analysis of the Social Cost of Carbon: due to its relation with the shadow-value functions, we are able to isolate the contributions of the elementary factors that add up to compose the Social Cost of Carbon.

Continuous-time DICE model

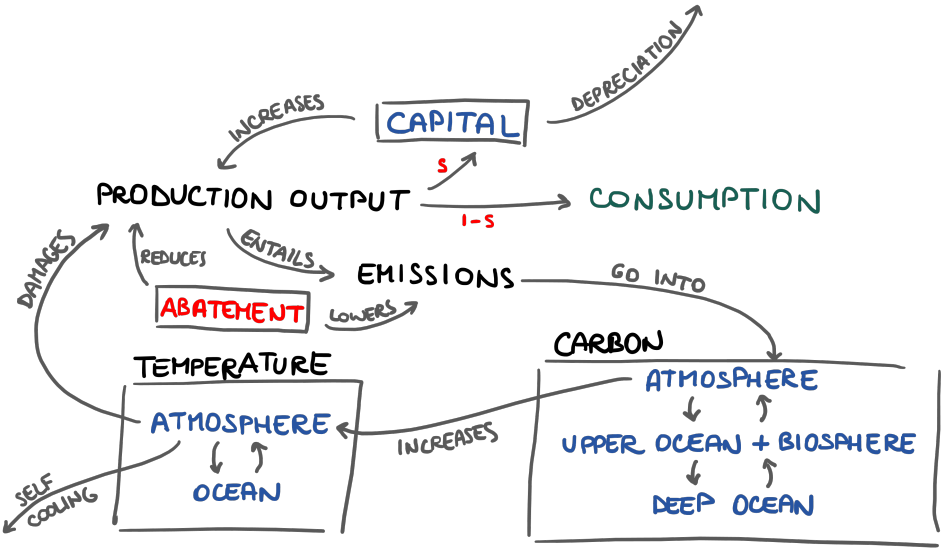
Before the occurrence of the tipping point, the model is a continuous-time version of the DICE that is described in Nordhaus [86]. We use the continuous-time version of the functional forms and parameters in Freiberger et al. [47], that were extracted from the official GAMS code DICE-2016R.

Control and State Variables. At every time, the state of the system is described by the following state variables: the *capital* stock K; the *carbon* stock vector $M = (M^{AT}, M^{UP}, M^{LO})$ has three components which measure the amount of carbon that is accumulated in the atmosphere, in the biosphere and shallow ocean, and in the deep ocean, respectively; the *temperature* vector $T = (T^{AT}, T^{OC})$ has two components which measure the global mean temperature change in the atmosphere and in the deep ocean, respectively.

There is a single policymaker who, at every time, sets the values the following control variables: the *savings* rate $s \in [0, 1]$ represents the fraction of the net economic output that is re-invested in capital; the *abatement* rate $\mu \in [0, 1]$ represents the fraction of industrial carbon emissions that is cut.

Controls:		**States:**	
$s \in [0,1]$	savings	K	capital
$\mu \in [0,1]$	abatement	M^{AT}	carbon (atmosphere)
		M^{UP}	carbon (shallow ocean + biosphere)
		M^{LO}	carbon (deep ocean)
		T^{AT}	temperature change (atmosphere)
		T^{OC}	temperature change (ocean)

The controlled dynamical system that describes the state variables' evolution, which is quite complex and involved, is presented in detail in the following pages. However, here is a graphic (and qualitative) representation of the mechanisms behind the mathematical expression of the dynamics.



Objective Function. The policymaker adjusts their policy (the savings rate s and the abatement rate μ) in order to optimize a social welfare function. This function is the flow over time of the "generalized" consumption, i.e., not only of traditional market goods and services, like food and shelter, but also of nonmarket items such as leisure, health status, and environmental services. This is obtained by maximizing the discounted integral of the population-weighted utility of per-capita consumption:

$$\underset{s,\mu\in[0,1]}{\text{maximize}} \int_0^T e^{-\rho t}\, L(t)\, u\big(c(t)\big)\, dt$$

where $\rho > 0$ is the rate of time preference (a discount is applied on the economic well-being of future generations), $L(t)$ is the exogenous population/labor, $u(c)$ is the utility of consumption, and $c(t)$ is the per-capita consumption $C(t)/L(t)$. The total consumption $C(t) = \big(1 - s(t)\big)Q(t)$ amounts to the fraction of the net production output $Q(t)$ that is not reinvested in capital, and as such it depends on the control and state variables (as explained in the next paragraph). The utility of consumption is isoelastic:

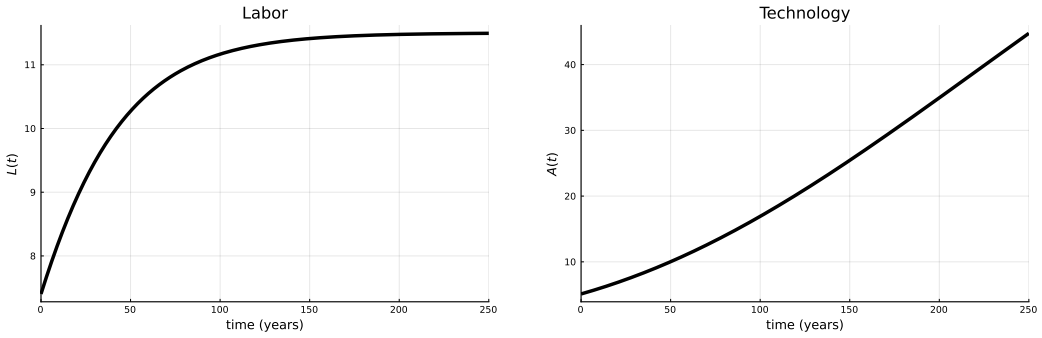
$$u(c) = k\,\frac{c^{1-\upsilon} - 1}{1 - \upsilon} \text{ with } \upsilon \in (0, 1) \quad \text{or} \quad u(c) = k \log(c)$$

Economic Variables. The exogenous function $L(t)$ counts the world population in billions. Its growth rate declines over time so that the total population approaches a limit of 11.5 billion.

The global production function $Y(t, K)$ is assumed to be a constant-returns-to-scale Cobb-Douglas production function in capital K, exogenous labor $L(t)$, and Hicks-neutral exogenous technological change $A(t)$ (a coefficient whose growth represents technological advancement):

$$Y(t, K) = A(t)K^{\gamma}L(t)^{1-\gamma}$$

where $\gamma \in (0, 1)$ is the capital elasticity.



The rising atmospheric temperature entails damages to production. Such damages include estimated damages to major sectors (such as agriculture), the cost of sea-level rise, adverse impacts on health, and nonmarket damages. They are assumed to be proportional to world output and are polynomial functions of global mean temperature change. Taking the temperature damage into con-

sideration, production is thus resized by a factor $(1 + \Omega(T^{AT}))^{-1}$ where

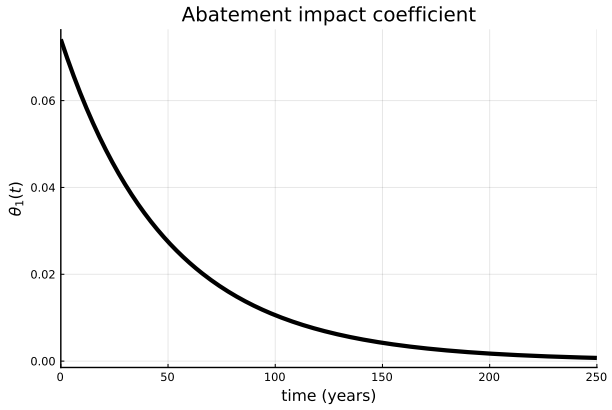
$$\Omega(T^{AT}) = \psi_1 (T^{AT})^{\psi_2},$$

with ψ_1, ψ_2 parameters.

Abatement of carbon emissions also constitutes costs in terms of production. Abatement costs are assumed to be proportional to global output and to a strictly convex polynomial function of μ. A backstop technology, i.e., an all-purpose environmentally benign zero-carbon energy technology that can replace all fossil fuels, is introduced into the model by setting the exogenously time-dependent coefficient in the abatement-cost to be equal to the backstop price for each year. Such a price is assumed to be initially high and to decline over time with carbon-saving technological change. Considering abatement costs, production is again resized by a factor $(1 - \Lambda(t, \mu))$ with

$$\Lambda(t, \mu) = \theta_1(t) \mu^{\theta_2},$$

where the coefficient $\theta_1(t)$ is the price of the backstop technology, in units of capital, and $\theta_2 > 1$ is a parameter.



After applying temperature-related damages and emission abatement costs to the total production, the net economic output amounts to:

$$Q(t, \mu, K, T^{AT}) = \frac{1 - \Lambda(t, \mu)}{1 + \Omega(T^{AT})} Y(t, K)$$

Assuming that the capital depreciates at a rate δ_K, and that a fraction s of the net output Q is reinvested, the capital K evolves according to the following

dynamics:

$$\dot{K} = sQ(t, \mu, K, T^{AT}) - \delta_K K$$

The rest of the net output is consumed:

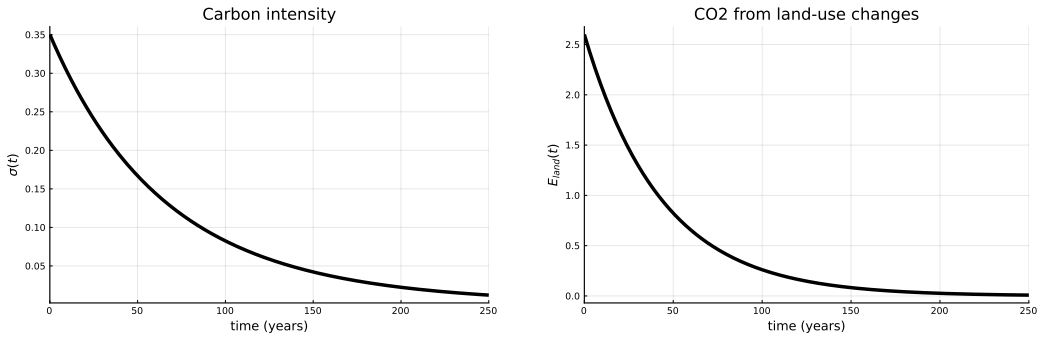
$$C(t) = \big(1 - s(t)\big)Q\big(t, \mu(t), K(t), T^{AT}(t)\big)$$

In absence of emission abatement, the industrial CO_2 emissions are given by the carbon intensity, i.e., the emission per unit of production output, times the total production output. The carbon intensity $\sigma(t)$ is assumed to be an exogenous function of time that is linked to technological advancement. Actual industrial emissions are then reduced by the abatement rate μ, yielding:

$$E_{ind}(t, \mu, K) = (1 - \mu)\sigma(t)Y(t, K)$$

Whereas industrial CO_2 emissions are endogenous, the CO_2 arising from land-use changes is assumed to be exogenous (other GHGs are also accounted for, in the exogenous part of the radiative forcing term in the temperature dynamics, as we explain later). The total flow of CO_2 emissions is thus

$$E(t, \mu, K) = E_{ind}(t, \mu, K) + E_{land}(t)$$



Geophysical Equations. The carbon cycle is represented by a three-reservoir model: we consider the stock of carbon in the atmosphere M^{AT}, a quickly mixing reservoir in the shallow ocean and the biosphere M^{UP}, and the carbon stored in the deep ocean M^{LO}. CO_2 emissions are released into the atmosphere (coefficient $\alpha = 3/11$ converts the mass of CO_2 to the mass of carbon), and there

is a constant exchange of carbon between adjacent layers: the biosphere takes up atmospheric carbon at a rate δ_1 and releases it back into the atmosphere at a rate δ_2, whereas the deep ocean absorbs carbon from the shallow ocean at a rate δ_3 and releases it back at a rate δ_4.

$$\begin{cases} \dot{M}^{AT} = -\delta_1 M^{AT} + \delta_2 M^{UP} + \alpha E(t, \mu, K) \\ \dot{M}^{UP} = \delta_1 M^{AT} - (\delta_2 + \delta_3) M^{UP} + \delta_4 M^{LO} \\ \dot{M}^{LO} = \delta_3 M^{UP} - \delta_4 M^{LO} \end{cases}$$

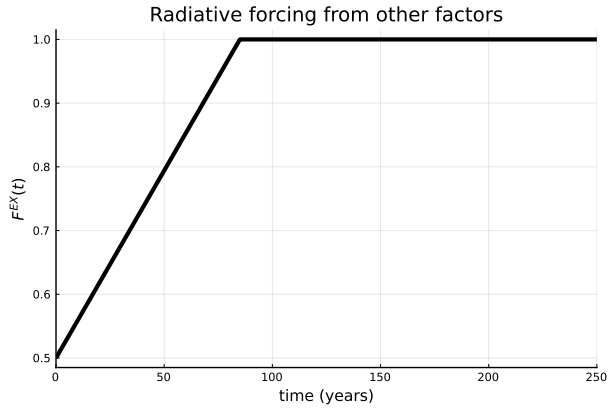
By introducing the following vector notation,

$$M = \begin{pmatrix} M^{AT} \\ M^{UP} \\ M^{LO} \end{pmatrix}, \qquad \Phi = \begin{bmatrix} -\delta_1 & \delta_2 & 0 \\ \delta_1 & -(\delta_2 + \delta_3) & \delta_4 \\ 0 & \delta_3 & -\delta_4 \end{bmatrix}, \qquad e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

we can reformulate the carbon dynamics as

$$\dot{M} = \Phi M + \alpha E(t, \mu, K) e_1$$

Radiative forcing, i.e., the greenhouse effect, is the mechanism that links the accumulation of GHGs in the atmosphere to warming at the earth's surface. Although the main driver of such phenomenon is the atmospheric CO_2, exogenous forcing from other GHGs (e.g., aerosols and ozone) is also present. The increase



in atmospheric temperature due to radiative forcing is given by:

$$F(t, M^{AT}) = \eta \log_2 \left(\frac{M^{AT}}{\tilde{M}} \right) + F^{EX}(t)$$

where $\tilde{M}$ is the pre-industrial level of atmospheric carbon, and η is a parameter. Furthermore, the atmosphere has a self-cooling ability and it also constantly exchanges heat with the ocean through diffusion.

$$\begin{cases} \dot{T}^{AT} = \xi_1 \left[-\xi_2 T^{AT} + \xi_3 (T^{OC} - T^{AT}) + F(t, M^{AT}) \right] \\ \dot{T}^{OC} = \xi_4 (T^{AT} - T^{OC}) \end{cases} \tag{3.1}$$

By introducing the following vector notation,

$$T = \begin{pmatrix} T^{AT} \\ T^{OC} \end{pmatrix}, \qquad \zeta = \begin{bmatrix} -\xi_1(\xi_2 + \xi_3) & \xi_1 \xi_3 \\ \xi_4 & -\xi_4 \end{bmatrix}, \qquad e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

we can reformulate the temperature dynamics as

$$\dot{T} = \zeta T + \xi_1 F(t, M^{AT}) e_1$$

Unlike the original DICE model, we do not include the following features: the resource constraint on carbon fuels

$$\int_0^T E_{ind}(t, \mu(t), K(t))\, dt \leq CCum,$$

nor the possibility of negative emissions after some time

$$\mu(t) \in [0, 1.2] \quad \text{for } t \geq t_0,$$

nor the constraint on atmospheric temperature

$$T^{AT} \leq \mathcal{T}.$$

There are two reasons behind this choice: because the theory we apply to solve our problem does not feature this kind of constraints or changing control sets, and because these specifications would not provide further insight for our purposes anyway.

Full model. Putting it all together, the DICE model in continuous time results in the following optimal control problem:

$$\begin{aligned}
&\underset{s,\mu\in[0,1]}{\text{maximize}} \int_0^T e^{-\rho t}\, L(t)\, u\big(c(t)\big)\, dt \\
&\text{subject to} \quad \begin{cases} \dot{K}(t) = sQ(t,\mu,K,T^{AT}) - \delta_K K & K(0) = K_0 \\ \dot{M}(t) = \Phi M + \alpha E(t,\mu,K)e_1 & M(0) = M_0 \\ \dot{T}(t) = \zeta T + \xi_1 F(t,M^{AT})e_1 & T(0) = T_0 \end{cases} \\
&\text{where} \quad c(t) = \big(1 - s(t)\big)\, Q\big(t,\mu(t),K(t),T^{AT}(t)\big) \,/\, L(t)
\end{aligned}$$

This is the problem that one should solve in a scenario without any tipping point. However, it is also the problem that a policymaker solves if they are not *expecting* a tipping point, regardless of it then occurring or not. For this reason, in a scenario where a tipping point is indeed under way, we can call this the *myopic* problem: the agent plans their policy as if no event is pending, and is caught by surprise by its occurrence.

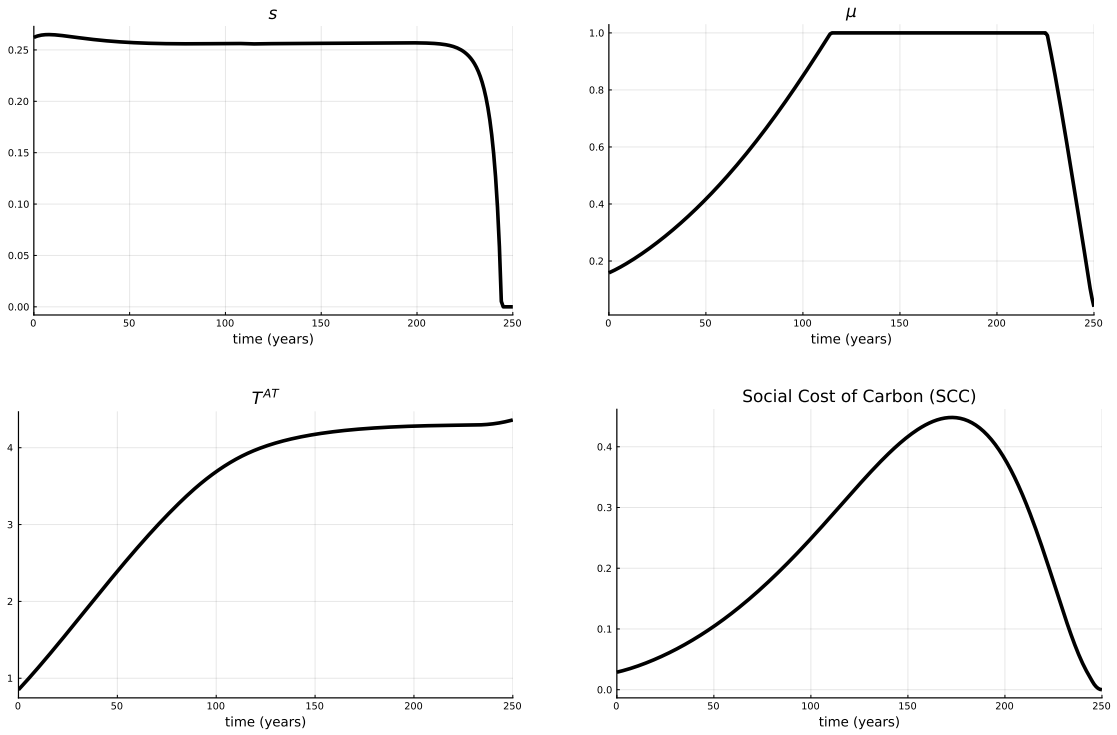
Below we show part of the output that was generated by the numerical simulation. The GAMS calibration was used for the parameter values.

The drop in savings rate and emission abatement toward the end of the time horizon is due to the absence of a scrap value function: indeed, if one is indifferent to the system's conditions at the final time, it makes sense to stop abating (to maximize production) and to stop saving (to maximize consumption). Usually, in this type of models, the plots are truncated to get rid of this trivial effect.

As for the controls, here we see that the savings rate s is quite stable slightly above 25%, whereas the abatement rate μ starts slightly below 20% and gradually increases towards net zero over the course of about 115 years.

The global atmospheric temperature change T^{AT} increases linearly from 0.85řC to about 3.5řC in the first 100 years, then its growth rate declines and the value settles at about 4.3řC.

The Social Cost of Carbon grows from about 0.03 to about 0.45 in 173 years, then it sharply declines at the end as a result of the end-time effect mentioned earlier; however, since the SCC is the fraction of adjoints (see Appendix 3.1.5), it starts decreasing about 50 years earlier than savings and abatement, which also include direct effects.



Integrating a Climate Tipping Point

Modelling a Stochastic Tipping Point. We add to the continuous-time DICE model described in section 3.1.2 the possibility of a stochastic climate tipping point, i.e., an unpredictable climate-related event that changes the system abruptly and irreversibly.

Due to current limitations in the mathematical tools that we use to tackle multi-stage optimal control problems with random switching times, we assume the tipping point to be unique, in that one and only one tipping point will eventually occur, either during the planning horizon or afterwards. This is insightful to study the separate effects of tipping points of different natures one by one; however, once an adequate theoretical framework is developed, it will be fundamental to consider in future research the possibility of a sequence of multiple interconnected tipping points, as feedback mechanisms and domino effects may realistically trigger a tipping point cascade.

Let us denote by τ the instant at which the tipping point occurs, also called *switching time* because it constitutes a shift from one regime to another. We will call *Stage 1* the time interval $[0, \tau \wedge T)$[3], and *Stage 2* the time interval $[\tau, T]$ if $\tau < T$.

[3]We denote $\tau \wedge T = \min(\tau, T)$

We model τ as a random variable, taking values in the positive half-line $[0, \infty)$, whose hazard rate is assumed to increase with the global mean temperature change:

$$\lim_{dt \to 0^+} \frac{\mathbb{P}\big(\tau \leq t + dt \mid \tau > t\big)}{dt} = h\big(T(t)\big)$$

with $h_{T^{AT}}, h_{T^{OC}} > 0$.[4]

The effects of the tipping point may be a change in the state dynamics, or a jump discontinuity in the state trajectory at the switching time. We only consider regime shifts that concern the carbon cycle dynamics (in Stage 2 the carbon cycle matrix $\Phi_2 \neq \Phi$), the temperature dynamics (in Stage 2 the temperature dynamics matrix $\zeta_2 \neq \zeta$), the impact of temperature on production (in Stage 2 the function $\Omega_2(T^{AT}) \neq \Omega(T^{AT})$, yielding $Q_2 \neq Q$), and instantaneous change in capital stock ($K(\tau^+) \neq K(\tau^-)$). This last case is not technically a tipping point, in that it does not change the nature of the system irreversibly, but since the concept of climate catastrophe entailing substantial economic damage *una tantum* is widely featured in the literature, we chose to include it in our analysis.

Although we are aware that we are considering only a limited range of regime shifts, this is only to lighten the notation. The same theory is applicable to any change in the system's dynamics, any change in the instant payoff, any jump discontinuity in the state trajectory of the kind $X(\tau^+) = \varphi\big(\tau, X(\tau^-)\big)$.

Information and Strategy. In Stage 1, the policymaker does not know when τ will occur, they know that it could occur at any time, and they are aware of its effects on the system. So, when planning a Stage 1 strategy in advance, one needs to plan for the whole time horizon (even if the switch then ends up happening during the time horizon): for all $t \in [0, T]$ we denote by

$$s_1(t), \qquad \mu_1(t)$$

the Stage 1 savings and abatement rate respectively, at time t.

In Stage 2, the policymaker knows that τ has occurred and when, and can therefore adjust their strategy accordingly. So, when planning a Stage 2 strategy in advance, one needs to plan for all possible occurrences of the switching time:

[4]We denote f_x the partial derivative of a function f with respect to the variable x.

for all $\tau \in [0, T]$ and for all $t \in [\tau, T]$ we denote by

$$s_2(\tau, t), \qquad \mu_2(\tau, t)$$

the Stage 2 savings and abatement rate respectively, at time t, given that the regime shift occurred at time τ. The same holds for the corresponding state trajectories: we will have $K_1(t), M_1(t), T_1(t)$ in Stage 1 and $K_2(\tau, t), M_2(\tau, t), T_2(\tau, t)$ in Stage 2.

In light of these considerations, and due to the stochastic nature of the switching time τ, the policymaker's objective is to maximize the *expectation* of the integral payoff over all possible realizations of τ:

$$\underset{s_1,\mu_1,s_2,\mu_2\in[0,1]}{\text{maximize}} \quad \mathbb{E}_\tau\left[\int_0^{\tau\wedge T} e^{-\rho t}\, L(t)\, u\big(c_1(t)\big)\, dt + \int_{\tau\wedge T}^{T} e^{-\rho t}\, L(t)\, u\big(c_2(\tau,t)\big)\, dt\right]$$

$$\text{subject to} \quad \begin{cases} \dot K_1(t) = s_1 Q(t, \mu_1, K_1, T_1^{AT}) - \delta_K K_1 & K_1(0) = K_0 \\ \dot M_1(t) = \Phi M_1 + \alpha E(t, \mu_1, K_1) e_1 & M_1(0) = M_0 \\ \dot T_1(t) = \zeta T_1 + \xi_1 F(t, M_1^{AT}) e_1 & T_1(0) = T_0 \\ \frac{d}{dt}\mathbb{P}(\tau > t) = -h\big(T_1(t)\big)\mathbb{P}(\tau > t) & \mathbb{P}(\tau > 0) = 1 \\ \dot K_2(\tau, t) = s_2 Q_2(t, \mu_2, K_2, T_2^{AT}) - \delta_K K_2 & K_2(\tau,\tau) = \varepsilon\big(T_1^{AT}(\tau)\big) K_1(\tau) \\ \dot M_2(\tau, t) = \Phi_2 M_2 + \alpha E(t, \mu_2, K_2) e_1 & M_2(\tau,\tau) = M_1(\tau) \\ \dot T_2(\tau, t) = \zeta_2 T_2 + \xi_1 F(t, M_2^{AT}) e_1 & T_2(\tau,\tau) = T_1(\tau) \end{cases}$$

$$\text{where} \quad c_1(t) = \big(1 - s_1(t)\big)\, Q_1\big(t, \mu_1(t), K_1(t), T_1^{AT}(t)\big) \,/\, L(t)$$

$$\text{and} \quad c_2(\tau, t) = \big(1 - s_2(\tau, t)\big)\, Q_2\big(t, \mu_2(\tau, t), K_2(\tau, t), T_2^{AT}(\tau, t)\big) \,/\, L(t)$$

In the next sections, we will sometimes use the following collective notation to denote the Stage 2 initial condition:

$$X_2(\tau, \tau) = \varphi\big(X_1(\tau)\big).$$

Optimality conditions

We address the problem with a new approach, that was recently developed by Wrzaczek et al. [118]. This is the first time that this methodology is employed in a climate/economy model with random switching time: the existing literature has only been featuring the backward approach so far.

The method consists in introducing two auxiliary state variables, Z_1, Z_2 in Stage 1 and 2 respectively, that are linked to the distribution of the random variable τ, then reformulating the objective value using a vintage structure, and finally applying the Maximum Principle for heterogeneous systems by Veliov [113].

The auxiliary state variables are

$$Z_1(t) = \mathbb{P}(\tau > t), \qquad Z_2(\tau, t) = h\big(T_1(\tau)\big)Z_1(\tau)$$

and as such they are subject to the following dynamics:

$$\begin{cases} \dot{Z}_1(t) = -h\big(T_1(t)\big)Z_1(t) \\ Z_1(0) = 1 \end{cases} \qquad \begin{cases} \dot{Z}_2(\tau, t) = 0 \\ Z_2(\tau, \tau) = h\big(T_1(\tau)\big)Z_1(\tau) \end{cases}$$

Observe that Z_1 is the tail of τ's distribution, whereas $Z_2(s,t) = -\dot{Z}_1(s) = -\frac{d}{ds}\mathbb{P}(\tau > s) = f_\tau(s)$, which is τ's probability distribution.

By explicitly computing the expectation using Z_1 and Z_2, and after some basic integral manipulations (e.g., integrating by parts and applying Fubini's theorem), we obtain the following vintage structure formulation for the problem:

$$\underset{s_1,\mu_1,s_2,\mu_2\in[0,1]}{\text{maximize}} \int_0^T e^{-\rho t}\left[Z_1(t)\,L(t)\,u\big(c_1(t)\big) + \int_0^t Z_2(\tau,t)\,L(t)\,u\big(c_2(\tau,t)\big)\,d\tau\right]dt$$

$$\text{subject to} \quad \begin{cases} \dot{K}_1(t) = s_1 Q(t,\mu_1,K_1,T_1^{AT}) - \delta_K K_1 & K_1(0) = K_0 \\ \dot{M}_1(t) = \Phi M_1 + \alpha E(t,\mu_1,K_1)e_1 & M_1(0) = M_0 \\ \dot{T}_1(t) = \zeta T_1 + \xi_1 F(t, M_1^{AT})e_1 & T_1(0) = T_0 \\ \dot{Z}_1(t) = -h\big(T_1(t)\big)Z_1(t) & Z_1(0) = 1 \\ \dot{K}_2(\tau,t) = s_2 Q_2(t,\mu_2,K_2,T_2^{AT}) - \delta_K K_2 & K_2(\tau,\tau) = \varepsilon\big(T_1^{AT}(\tau)\big)K_1(\tau) \\ \dot{M}_2(\tau,t) = \Phi_2 M_2 + \alpha E(t,\mu_2,K_2)e_1 & M_2(\tau,\tau) = M_1(\tau) \\ \dot{T}_2(\tau,t) = \zeta_2 T_2 + \xi_1 F(t, M_2^{AT})e_1 & T_2(\tau,\tau) = T_1(\tau) \\ \dot{Z}_2(\tau,t) = 0 & Z_2(\tau,\tau) = h(T_1(\tau))Z_1(\tau) \end{cases}$$

Maximum Principle. A Hamiltonian function is defined for each stage:

$$H_1 = H_1(t, s_1, \mu_1, K_1, M_1, T_1, \lambda),$$
$$H_2 = H_2(t, s_2, \mu_2, K_2, M_2, T_2, \xi)$$

where λ, ξ denote the co-state vectors of Stage 1 and 2 respectively. In particular, we will denote by λ_X, ξ_X the adjoint variables of the canonical state variables, and by λ_Z, ξ_Z the adjoint variables of the auxiliary state variables. Of course, the co-state trajectories depend on the same variables as the corresponding state trajectories:

$$\lambda = \lambda(t), \qquad \xi = \xi(\tau, t)$$

From now on we may omit the arguments of H_1, H_2 to lighten the notation.

$$H_1 = Lu(c_1) + \lambda_K [s_1 Q - \delta_K K_1] + \lambda_M \cdot \Phi M_1 + \lambda_{M^{AT}} \alpha E + \lambda_T \cdot \zeta T_1 + \lambda_{T^{AT}} \xi_1 F$$
$$H_2 = Lu(c_2) + \xi_K [s_2 Q_2 - \delta_K K_2] + \xi_M \cdot \Phi_2 M_2 + \xi_{M^{AT}} \alpha E + \xi_T \cdot \zeta_2 T_2 + \xi_{T^{AT}} \xi_1 F$$

Let us denote $(s_i^*, \mu_i^*, K_i^*, M_i^*, T_i^*)$, $i = 1, 2$, the optimal process. The maximality condition from the Maximum Principle states that, for almost all t, τ,

$$\big(s_1^*(t), \mu_1^*(t)\big) \in \arg\max_{s,\mu} H_1\big(t, s, \mu, K_1^*(t), M_1^*(t), T_1^*(t), \lambda(t)\big)$$
$$\big(s_2^*(\tau, t), \mu_2^*(\tau, t)\big) \in \arg\max_{s,\mu} H_2\big(t, s, \mu, K_2^*(\tau, t), M_2^*(\tau, t), T_2^*(\tau, t), \xi(\tau, t)\big)$$

where the co-state trajectories satisfy

$$\begin{cases} \dot\lambda_X(t) = \rho\lambda_X(t) - \nabla_X H_1(t) - \big[\xi_X(t,t) \cdot \nabla_X \varphi\big(X_1(t)\big) - \lambda_X(t)\big] h\big(T_1(t)\big) \\ \qquad\qquad - \big[\xi_Z(t,t) - \lambda_Z(t)\big] \nabla_X h\big(T_1(t)\big) \\ \lambda_X(T) = 0 \\ \dot\lambda_Z(t) = \rho\lambda_Z(t) - L(t)u\big(c_1(t)\big) - \big[\xi_Z(t,t) - \lambda_Z(t)\big] h\big(T_1(t)\big) \\ \lambda_Z(T) = 0 \\ \dot\xi_X(\tau, t) = \rho\xi_X(\tau, t) - \nabla_X H_2(\tau, t) \\ \xi_X(\tau, T) = 0 \\ \dot\xi_Z(\tau, t) = \rho\xi_Z(\tau, t) - L(t)u\big(c_2(\tau, t)\big) \\ \xi_Z(\tau, T) = 0 \end{cases}$$

Observe that the Stage 2 co-states enter the Stage 1 adjoint equation: through this, the anticipating effect of the tipping point influences the Stage 1 behavior.

First Order Conditions. The maximum condition yields that, whenever s_1^* or s_2^* are inner controls, i.e., they belong to the inner part $(0,1)$ of the control set $[0,1]$, they are stationary points for the respective Hamiltonians:

$$0 = \partial_s H_1 = [-u_c(c_1) + \lambda_K]\, Q$$
$$0 = \partial_s H_2 = [-u_c(c_2) + \xi_K]\, Q_2$$

So, for all t such that the optimal savings rate $s_1^*(t)$ is an inner solution (arguments are omitted to lighten the notation),

$$u_c(c_1) = \lambda_K \tag{3.2}$$

and for all τ, t such that the optimal savings rate $s_2^*(\tau, t)$ is an inner solution,

$$u_c(c_2) = \xi_K \tag{3.3}$$

Analogously, whenever μ_1^* or μ_2^* are inner controls,

$$0 = -\partial_\mu H_1 = \left[(1 - s_1^*)u_c(c_1) + s_1^* \lambda_K\right] Q_\mu + \lambda_{M^{AT}} \alpha E_\mu$$
$$0 = -\partial_\mu H_2 = \left[(1 - s_2^*)u_c(c_2) + s_2^* \xi_K\right] (Q_2)_\mu + \xi_{M^{AT}} \alpha E_\mu$$

If, furthermore, s_1^* or s_2^* respectively are inner controls,

$$0 = -\partial_\mu H_1 = \lambda_K Q_\mu + \lambda_{M^{AT}} \alpha E_\mu$$
$$0 = -\partial_\mu H_2 = \xi_K (Q_2)_\mu + \xi_{M^{AT}} \alpha E_\mu$$

yielding

$$\alpha \frac{-\lambda_{M^{AT}}}{\lambda_K} = \frac{Q_\mu}{E_\mu} \tag{3.4}$$

$$\alpha \frac{-\xi_{M^{AT}}}{\xi_K} = \frac{(Q_2)_\mu}{E_\mu} \tag{3.5}$$

This is quite meaningful because, since s_i^* is inner, the term on the LHS coincides with the Social Cost of Carbon (see Appendix 3.1.5).

Adjoint Equations. Although we can of course compute the adjoint equations for all the variables, due to their relevance in the calculation of the Social Cost of Carbon (see Appendix 3.1.5) we are more interested in the co-state trajectories of capital and atmospheric carbon, and particularly in their growth rates.

Here we are omitting the dependence on time and switching time. Recall that in the adjoint equation for Stage 1 at time t, the Stage 2 co-states are evaluated on the diagonal, i.e., $\xi = \xi(t,t)$.

Observe that in both cases the terms of the two equations are symmetric up until the final one: in general, the Stage 1 shadow values also depend on the Stage 2 shadow values and on the 'switching conditions' of the states (i.e., condition $X_2(t,t) = \varphi(X_1(t))$). This dependence encloses the policymaker's awareness in regard to the possibility of a tipping point.

We start with the Stage 1 and 2 adjoint equations for capital:

$$\dot{\lambda}_K = (\rho + \delta_K)\lambda_K - \left\{\left[(1 - s_1^*)u_c(c_1) + s_1^*\lambda_K\right]Q_K + \lambda_{M^{AT}}\alpha E_K\right\} - \left[\varepsilon\xi_K - \lambda_K\right]h$$
$$\dot{\xi}_K = (\rho + \delta_K)\xi_K - \left\{\left[(1 - s_2^*)u_c(c_2) + s_2^*\xi_K\right](Q_2)_K + \xi_{M^{AT}}\alpha E_K\right\}$$

In this case, the final term in the Stage 1 equation is $-\left[\varepsilon((T_1^{AT})^*(t))\xi_K(t,t) - \lambda_K(t)\right]h(T_1^*(t))$. Recalling that $\varepsilon(T^{AT})$ is the fraction of capital that remains after the tipping point, and by educated-guessing that both $\lambda_K, \xi_K > 0$ (which is confirmed by the numerical simulations), we can observe that the preservation of capital upon the regime shift has a frontloading effect on λ_K. This finds an intuitive explanation in the fact that if you know that part of the capital is going to be destroyed at the switch, then, before the switch occurs, it will hold less value than it would if it were to be completely preserved.

Suppose that s_1^*, s_2^* are inner controls and that $\mu_1^*, \mu_2^* \neq 0$.[5] Then, by the FOCs

[5]This assumption is reasonable, since from the numerical simulations we obtain that s_i^* is inner up to the last 5 years in the planning horizon, and that μ_i^* is always nonzero except at the final time. Abatement $\mu^* = 1$ is acceptable even if we are making the substitution $-\alpha\lambda_{M^{AT}}/\lambda_K = Q_\mu/E_\mu$ (which only holds if $\mu^* \in (0,1)$), because the substituted terms are multiplied by $E_K = (1 - \mu^*)\sigma Y_K = 0$.

(3.2-3.3) and (3.4-3.5), the growth rates of λ_k, ξ_K are:

$$\frac{\dot{\lambda}_K}{\lambda_K} = \rho + \delta_K - \left[Q_K - \frac{Q_\mu}{E_\mu}E_K\right] - \left[\varepsilon\frac{\xi_K}{\lambda_K} - 1\right]h$$
$$\frac{\dot{\xi}_K}{\xi_K} = \rho + \delta_K - \left[(Q_2)_K - \frac{(Q_2)_\mu}{E_\mu}E_K\right]$$

Next, we compute the adjoint equations for the atmospheric carbon:

$$\dot{\lambda}_{M^{AT}} = \rho\lambda_{M^{AT}} - \left[\delta_1(\lambda_{M^{UP}} - \lambda_{M^{AT}}) + \xi_1 F_{M^{AT}}\lambda_{T^{AT}}\right] - [\xi_{M^{AT}} - \lambda_{M^{AT}}]h$$
$$\dot{\xi}_{M^{AT}} = \rho\xi_{M^{AT}} - \left[\delta_1^{(2)}(\xi_{M^{UP}} - \xi_{M^{AT}}) + \xi_1 F_{M^{AT}}\xi_{T^{AT}}\right]$$

Here the final term in the Stage 1 equation is $-\left[\xi_{M^{AT}}(t,t) - \lambda_{M^{AT}}(t)\right]h(T_1^*(t))$. By guessing that both $\lambda_{M^{AT}}, \xi_{M^{AT}} < 0$ (confirmed by the numerical simulations), we observe that the continuity of the atmospheric carbon upon the regime shift constitutes a frontloading factor for (the negative value of) $\lambda_{M^{AT}}$: indeed, knowing that carbon will stay around after the shift makes it worse to accumulate it in Stage 1.

Their growth rates are:

$$\frac{\dot{\lambda}_{M^{AT}}}{\lambda_{M^{AT}}} = \rho - \left[\delta_1\left(\frac{\lambda_{M^{UP}}}{\lambda_{M^{AT}}} - 1\right) + \xi_1 F_{M^{AT}}\frac{\lambda_{T^{AT}}}{\lambda_{M^{AT}}}\right] - \left[\frac{\xi_{M^{AT}}}{\lambda_{M^{AT}}} - 1\right]h$$
$$\frac{\dot{\xi}_{M^{AT}}}{\xi_{M^{AT}}} = \rho - \left[\delta_1^{(2)}\left(\frac{\xi_{M^{UP}}}{\xi_{M^{AT}}} - 1\right) + \xi_1 F_{M^{AT}}\frac{\xi_{T^{AT}}}{\xi_{M^{AT}}}\right]$$

Social Cost of Carbon. The Social Cost of Carbon (SCC) is defined as the Marginal Rate of Substitution (MRS) of consumption for emissions: it measures how much consumption one is willing to give up to reduce CO_2 emissions by 1 unit. Whenever s_1^* or s_2^* are inner controls, then

$$SCC_1 = \alpha\frac{-\lambda_{M^{AT}}}{\lambda_K}, \qquad SCC_2 = \alpha\frac{-\xi_{M^{AT}}}{\xi_K}$$

respectively (proved in Appendix 3.1.5). As we will see from the numerical simulations, the savings rate (both in Stage 1 and 2) is inner up to the last ~5 years in the planning horizon, so we can safely use these formulae for the SCC.

Hence

$$\frac{\dot{SCC}_1}{SCC_1} = \frac{\dot{\lambda}_{M^{AT}}}{\lambda_{M^{AT}}} - \frac{\dot{\lambda}_K}{\lambda_K}, \qquad \frac{\dot{SCC}_2}{SCC_2} = \frac{\dot{\xi}_{M^{AT}}}{\xi_{M^{AT}}} - \frac{\dot{\xi}_K}{\xi_K}$$

Substituting the growth rates of the co-states, which were computed in the previous subsection, under the further hypothesis that $\mu_1^*, \mu_2^* \neq 0$ (which is confirmed by the numerical simulations, except in the final time),

$$\frac{\dot{SCC}_1}{SCC_1} = \left[Q_K - \delta_K - \frac{Q_\mu}{E_\mu} E_K\right] - \left[\delta_1\left(\frac{\lambda_{M^{UP}}}{\lambda_{M^{AT}}} - 1\right) + \xi_1 F_{M^{AT}} \frac{\lambda_{T^{AT}}}{\lambda_{M^{AT}}}\right] + \left[\varepsilon \frac{\xi_K}{\lambda_K} - \frac{\xi_{M^{AT}}}{\lambda_{M^{AT}}}\right] h \tag{3.6}$$

$$\frac{\dot{SCC}_2}{SCC_2} = \left[(Q_2)_K - \delta_K - \frac{(Q_2)_\mu}{E_\mu} E_K\right] - \left[\delta_1^{(2)}\left(\frac{\xi_{M^{UP}}}{\xi_{M^{AT}}} - 1\right) + \xi_1 F_{M^{AT}} \frac{\xi_{T^{AT}}}{\xi_{M^{AT}}}\right] \tag{3.7}$$

where, in both stages, the first group of terms comes from the growth rate of the co-state of capital and the second group comes from the growth rate of the co-state of the atmospheric carbon; in Stage 1 an additional third group is present, which reflects the anticipation of the tipping point.

One can prove that:

$$\begin{aligned} \xi_X(\tau, t) &= \nabla_X V_2\big(\tau, t, X_2^*(\tau, t)\big) \\ \lambda_X(t) &= \nabla_X V\big(t, X_1^*(t)\big) \end{aligned}$$

so that, through further calculations, we can attach a meaning to the extra terms in the Stage 1 SCC growth rate:

$$\begin{aligned} \varepsilon \frac{\xi_K}{\lambda_K} - \frac{\xi_{M^{AT}}}{\lambda_{M^{AT}}} &= \frac{\varepsilon \cdot \partial_K V_2\big(t, t, \varphi(X_1^*(t))\big)}{\partial_K V\big(t, X_1^*(t)\big)} - \frac{\partial_{M^{AT}} V_2\big(t, t, \varphi(X_1^*(t))\big)}{\partial_{M^{AT}} V\big(t, X_1^*(t)\big)} \\ &= \left[\frac{\partial_K \left[V_2\big(t, \varphi(X)\big)\right]}{\partial_K V\big(t, X\big)} - \frac{\partial_{M^{AT}} \left[V_2\big(t, \varphi(X)\big)\right]}{\partial_{M^{AT}} V\big(t, X\big)}\right]\Bigg|_{X = X_1^*(t)} \end{aligned}$$

where the first (resp. second) ratio is the marginal gain in the Stage 2 value from a unit increment in K_1 (resp. M_1^{AT}) if switching now, compared to the marginal gain in the expected value of the 2-stage problem from a unit increment in K_1 (resp. M_1^{AT}). All that is weighted by the hazard rate of τ at time t.

Guessing that the numerator and denominator are both positive in the first fraction, and both negative in the second fraction, we could say that the more is gained from switching *now* with a unit increment in K_1, the more the SCC is backloaded, and the more is lost from switching *now* with one unit increment in M_1^{AT}, the more the SCC is frontloaded.

3.1.3 Numerical results: four tipping point scenarios

In this section we present the numerical results in four different tipping point scenarios: A) Melting of sea ice or ice sheets, B) Destruction of capital due to an environmental catastrophe, C) Transformation of a carbon sink into a carbon source, and D) Increase in the impact of temperature on production.

We set a time horizon of 250 years. For the myopic scenario, where the planner does not expect a tipping point, and for the parametrization of Stage 1 in the problem with the tipping point, we use the continuous-time version of the functional forms and parameters in Freiberger et al. [47], that were extracted from the official GAMS code DICE-2016R.

In all the simulations we formulate a linear hazard rate with respect to the temperature: $h(T) = (T^{AT} + T^{OC})/60$. As for the Stage 2 parametrization, we purposely exaggerated the effects of the tipping points on the system, leading to exaggerated (and hopefully unrealistic!) outcomes in terms of the Stage 2 temperature and/or carbon. Keep in mind that this was done in order to better highlight the *qualitative* adjustments to the optimal behavior in preparation for and adaptation to the tipping point, and not to accurately quantify the tipping point's effect or the relative optimal strategy. A sensitivity analysis will be performed in future research for this purpose.

The numerical results were obtained through a gradient descent method based on the Maximum Principle for heterogeneous systems: starting from an initial guess for the controls, at each step the corresponding states and co-states are computed through forward and backward integration respectively, then a line search in a descent direction for $-H_1$ and $-H_2$ with respect to the controls is performed by comparing objective values, and finally the controls are updated.

How to read the figures

We include essentially two types of figures.

In the first kind we compare a variable's behavior in the *myopic* scenario, in gray, with its behavior in *anticipation* of the tipping point, in black (see figure 3.1).

The second kind features the anticipating *Stage 1* behavior (in black) and the *Stage 2* behavior after the tipping point (in color). A Stage 2 line will live in the time interval between the tipping point τ and the final time T, and since τ could occur at any time, in principle there should be a colored line starting

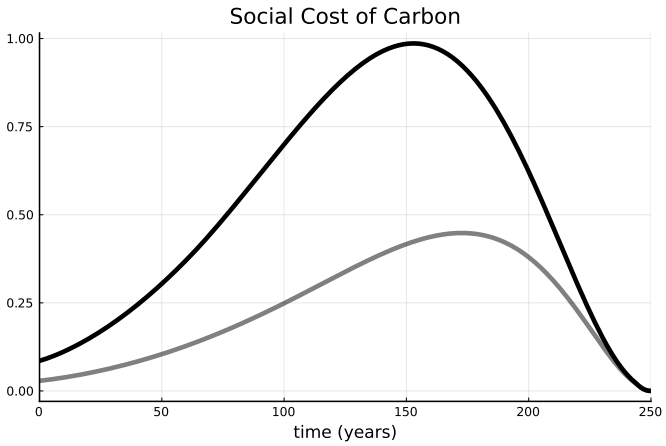


Figure 3.1: Anticipating (black) vs myopic (gray)

from every instant. Of course this would be impossible to represent, so we only display ten Stage 2 lines (for one realization of τ every 25 years). To consult the variable's global behavior for a specific occurrence $\tau = s$, one should follow the black Stage 1 line from $t = 0$ to $t = s$, and then jump to the colored Stage 2 line which starts from $t = s$ (see figure 3.2).

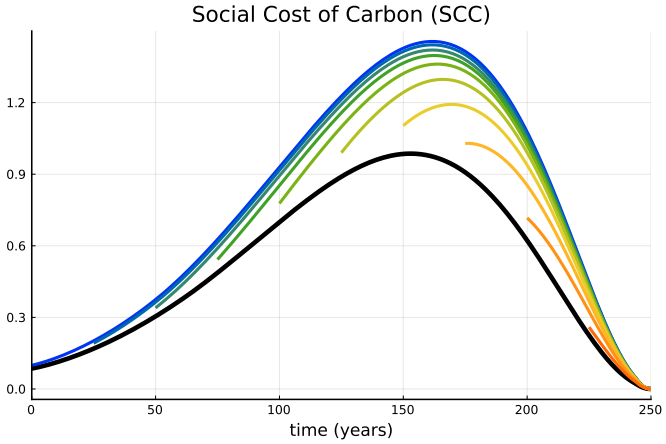


Figure 3.2: Stage 1 (black) vs Stage 2 (color)

A) Melting of sea ice or ice sheets

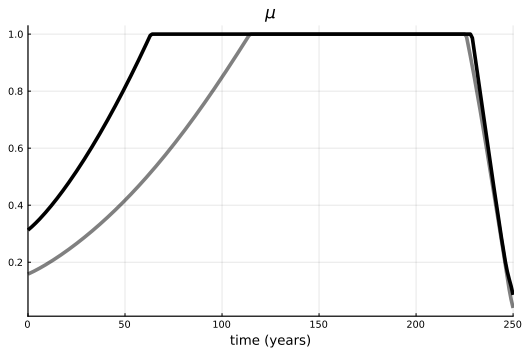
The ice reservoirs on the Earth's surface are important allies in keeping the atmosphere cool: this is due to the ice-albedo effect, which consists in the ice's ability to reflect part of the sun's radiation back to space. In the DICE model, this effect is enclosed in the self-cooling term $-\xi_2 T^{AT}$ in the atmospheric temperature's dynamics (3.1).

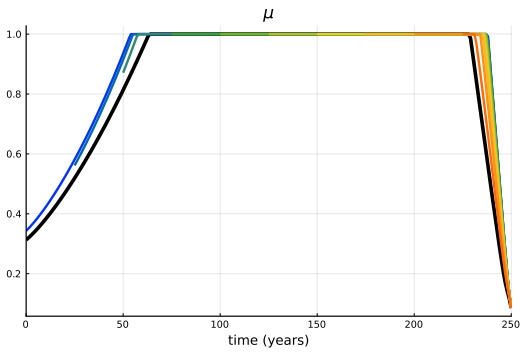
However, global warming is triggering a feedback mechanism involving the ice-albedo effect: higher temperatures reduce the extension of the ice sheets, which in turn reflect less of the radiation away from our atmosphere, allowing for the temperature to rise even more.

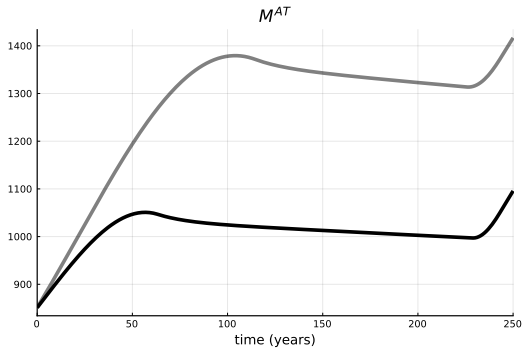
Supposing that a tipping point may consist in the complete melting of an important ice reservoir, such as the Arctic sea-ice or the Greenland ice sheet, we model this effect as a drastic lowering of the atmosphere's cooling rate, upon the regime shift:

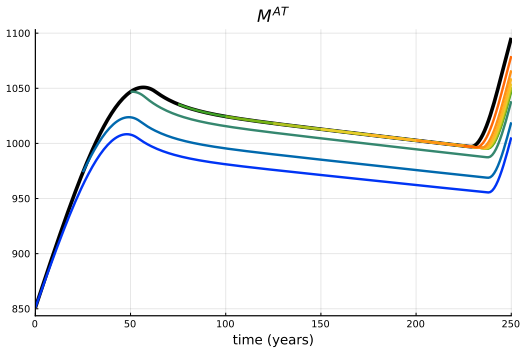
$$\xi_2^{(2)} < \xi_2$$

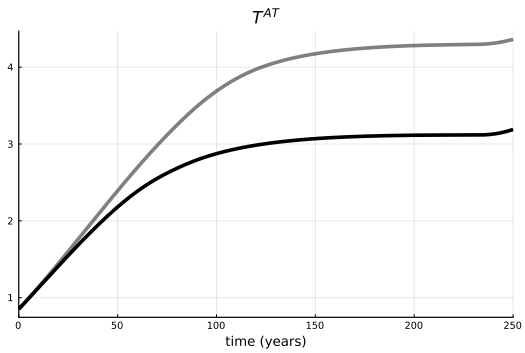
In the following simulation, $\xi_2^{(2)} = \xi_2/5$.

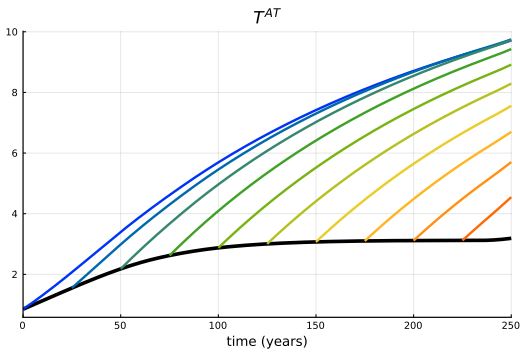


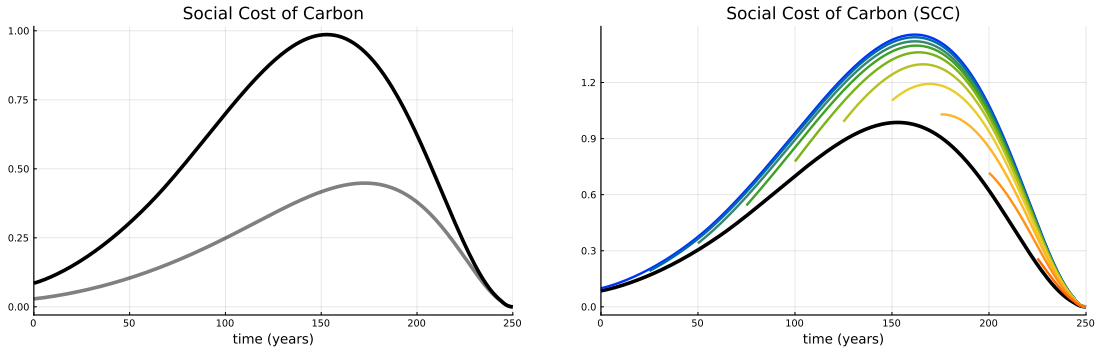












The tipping point has a catastrophic effect on atmospheric temperature: in Stage 2 the mean temperature change rises up to an exaggerated 10řC. Compared to the myopic scenario, in anticipation of the tipping point we observe a precautionary increase in the emission abatement μ. The purpose is to postpone the tipping point by keeping the atmospheric temperature (which is the only risk factor) as low as possible: the higher the abatement, the lower the carbon in the atmosphere, the lower the temperature. The SCC peaks about 20 years earlier than in the myopic scenario, and at more than double the value: indeed emissions are way more costly when trying to fend off such a catastrophic event.

B) Destruction of capital due to an environmental catastrophe

A realization of this scenario may consist in a serious climate event, such as severe wind or precipitation, flooding, wildfire, extreme heatwave, ecosystem disruption, etc. Even though these catastrophes have a regional characterization, there may be an impact on the global economy, which could be due to side-effects of the event such as mass migrations, pandemic outbreaks, disruptions in global supply chains, etc.

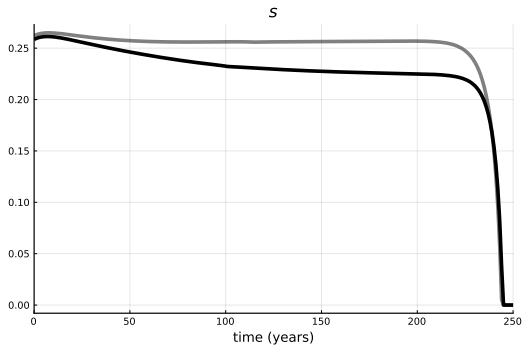
This is not technically a tipping point, because even though there is an instantaneous damage, the system is not permanently changed. This means that, once the only switch has occurred, we are back to the original system where the planner is not expecting any tipping point, i.e., the myopic scenario. This fact will be evident from the simulation figures below.

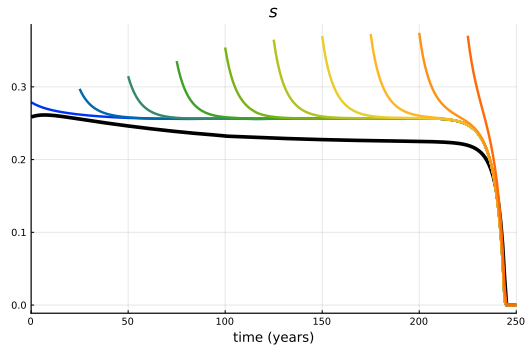
We assume the damage to be proportional to the capital itself, and to the global mean temperature change in the atmosphere:

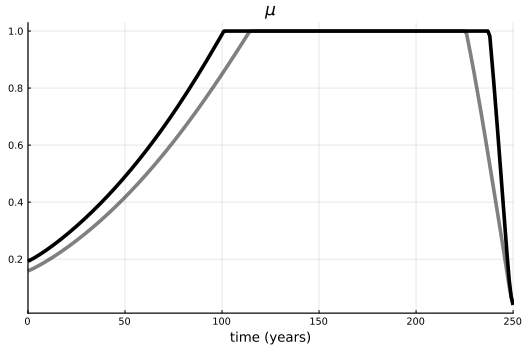
$$K_2(\tau, \tau) = K_1(\tau) - \varepsilon_T T_1^{AT}(\tau) K_1(\tau)$$

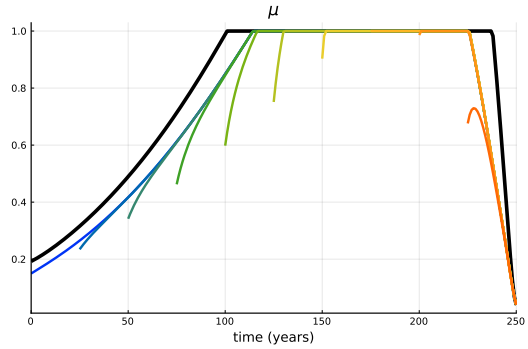
$$\varepsilon(T^{AT}) = 1 - \varepsilon_T T_1^{AT}$$

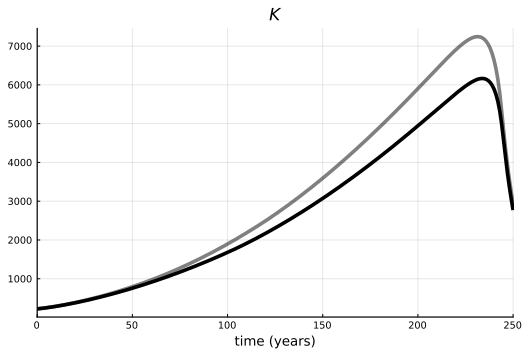
In the following simulation, $\varepsilon_T = 0.15$.

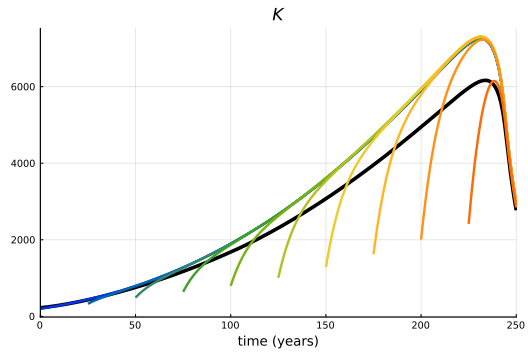


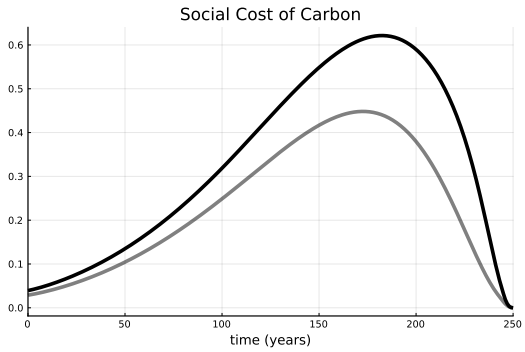


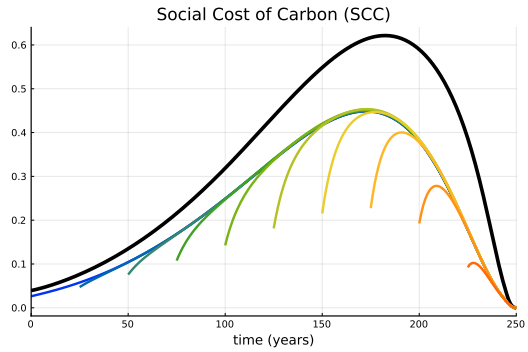












Compared to the myopic scenario, in anticipation of the tipping point the planner is saving less, thus accumulating less capital: indeed, since part of

the capital is going to be destroyed at the switch, it is less valuable in Stage 1. The abatement is higher in order to postpone the switch and to reduce the damage (both depend on atmospheric temperature). The SCC is higher because of the negative effect of emissions, but peaks slightly later: indeed capital – and therefore, damage – is increasing with time, so a later switch yields greater damage, and therefore emissions are more costly later on.

Observe that, as we anticipated, all the variables' behavior in Stage 2 retrace the myopic case, after an adjustment period of about 25 years, in which we restore capital through higher savings and lower abatement.

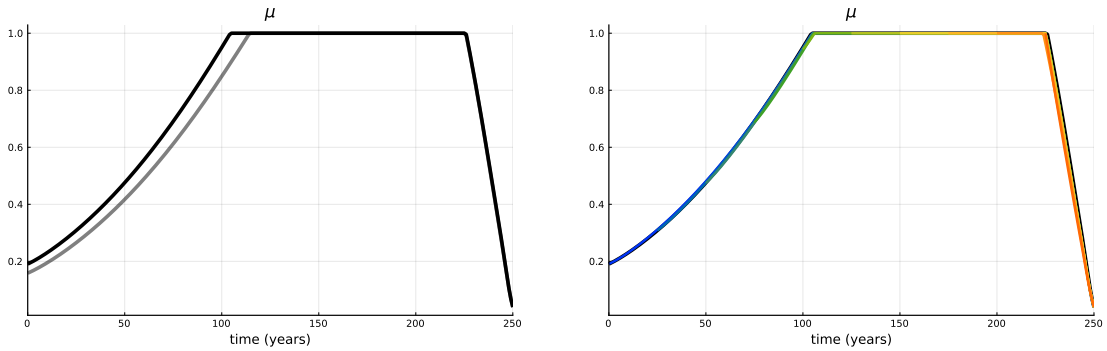
C) Transformation of a carbon sink into a carbon source

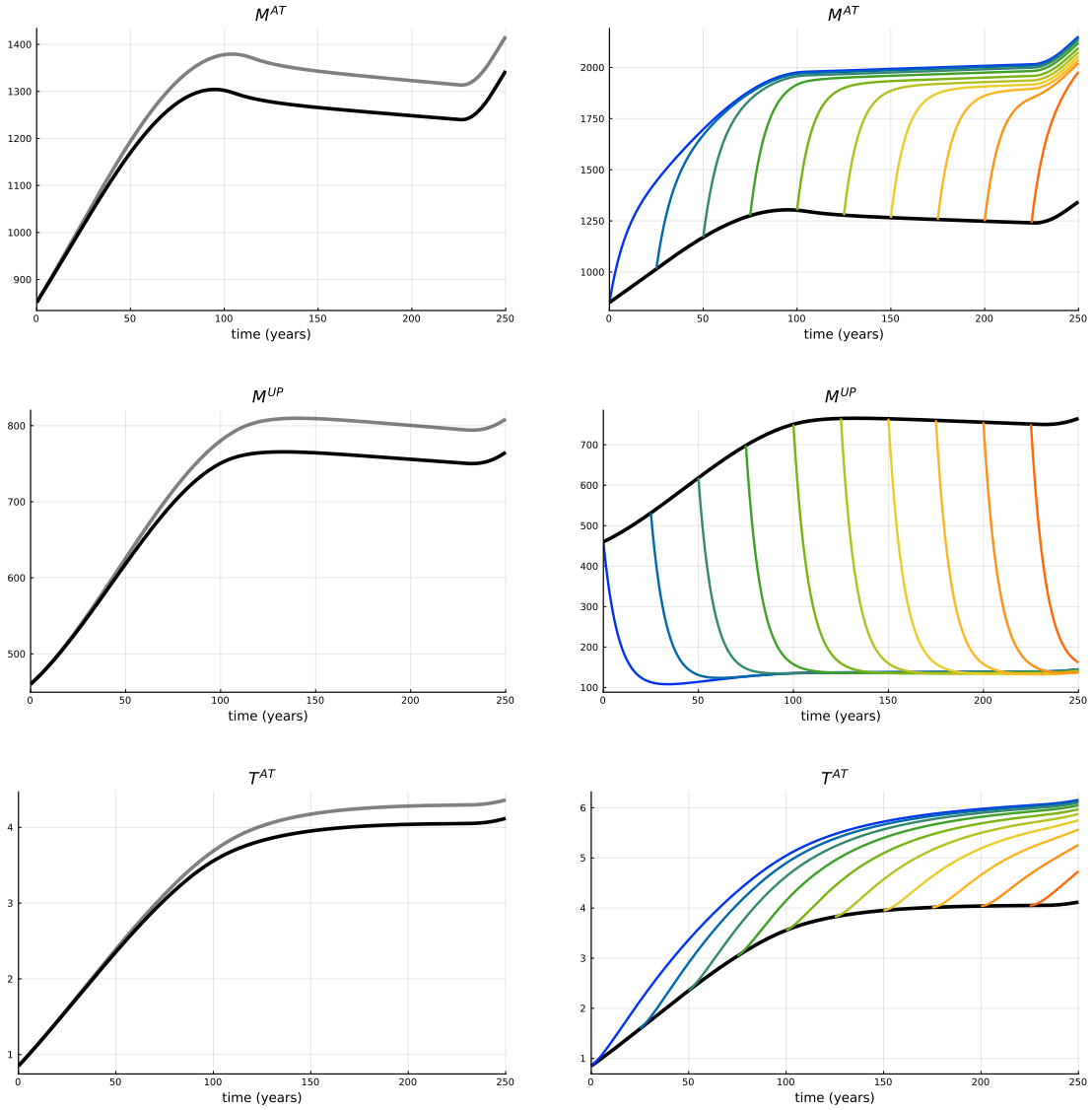
The Earth's biosphere is an important ally in taking up the atmospheric carbon, however, anthropogenic factors are endangering and compromising its activity. Some examples of this phenomenon are: the destruction of the Amazon rainforest and of the Boreal forest, the die-off of coral reefs, and the thawing of the Permafrost. When such an ecosystem is destroyed, the CO_2 that was absorbed and stored is released back into the atmosphere.

This could be modeled in Stage 2 by reducing the carbon uptake rate from the atmosphere to the middle layer, and increasing the release rate:

$$\delta_1^{(2)} < \delta_1, \qquad \delta_2^{(2)} > \delta_2$$

In the following simulation, $\delta_1^{(2)} = \delta_1/3$ and $\delta_2^{(2)} = 3 \cdot \delta_2$.





Very little is done to fend off the tipping point: compared to the myopic scenario, the abatement in the anticipating case is only slightly higher (hence the carbon and temperature are slightly lower).

Interestingly enough, nothing changes in the planner's policy from Stage 1 to Stage 2: indeed, the carbon stocks and temperature reach stable values that are different from the Stage 1 stable values (higher for the atmosphere, lower for the biosphere).

D) Increased impact of temperature on production

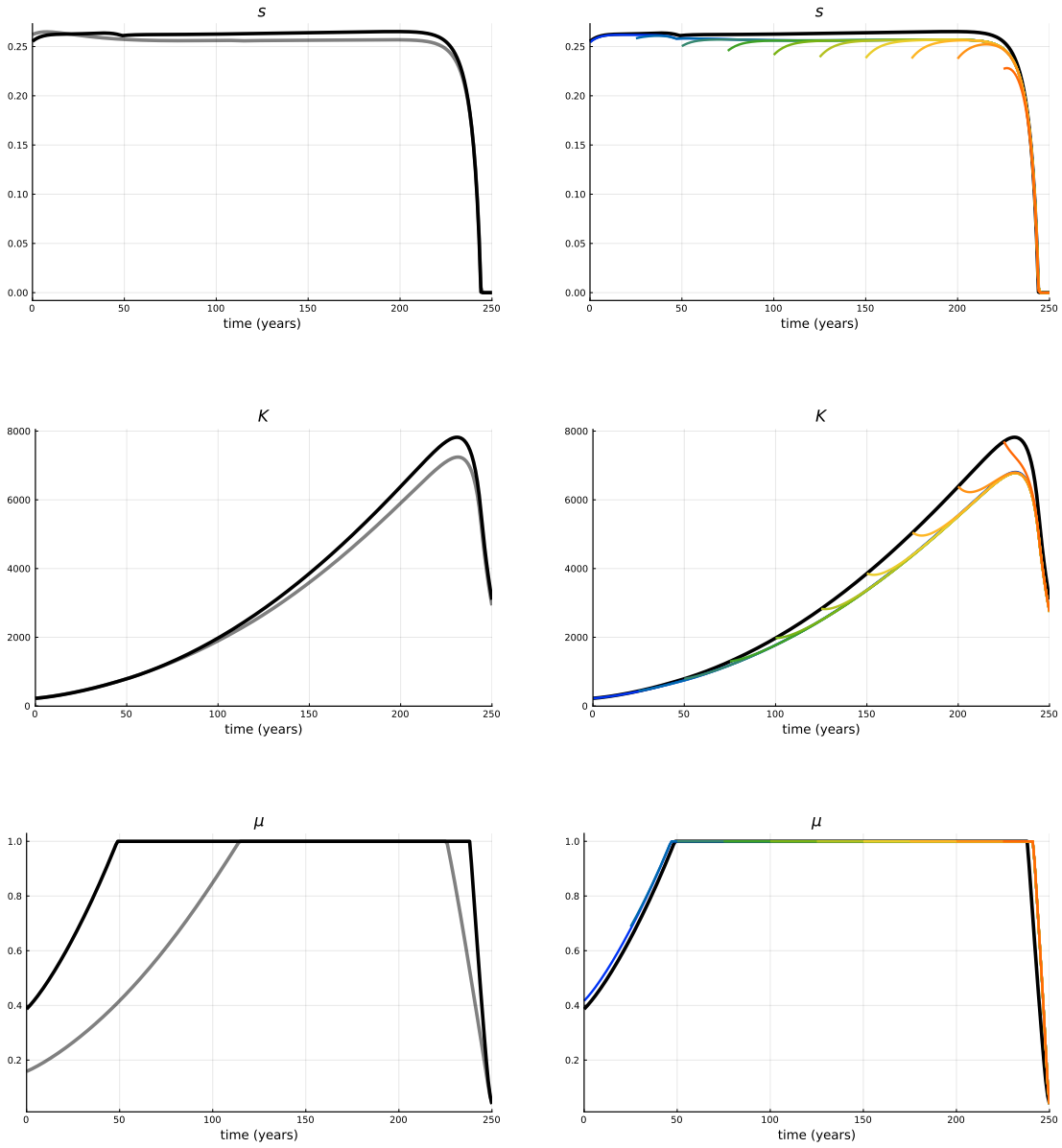
Some production sectors, like agriculture, highly depend on climate conditions that are not represented in this model, such as humidity or rainfall. In this

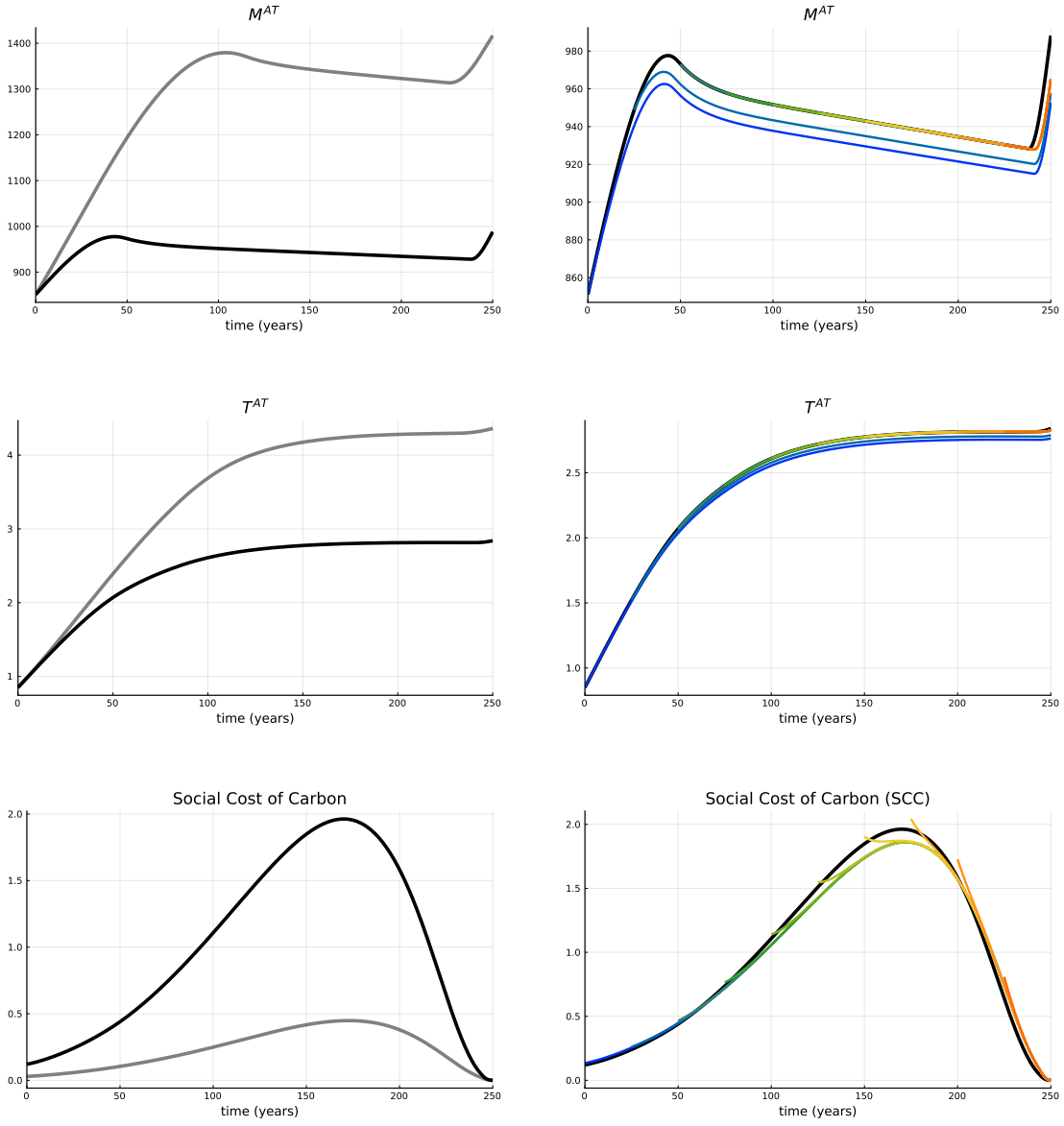
scenario, we suppose that the climate change brought on by the tipping point alters some of these climate variables, thus making the economic production more sensitive to global warming.

This is modeled by increasing the temperature damage function upon the regime shift:

$$Q_2 = \frac{1 - \Lambda(t, \mu)}{1 + \Omega_2(T^{AT})} Y(t, K) \quad \text{with} \quad \Omega_2(T^{AT}) > \Omega(T^{AT})$$

In the following simulation, $\Omega_2 = 5 \cdot \Omega$.





Compared to the myopic scenario, in the anticipating case much more abatement is employed to fend off the tipping point (high abatement implies low carbon, hence low temperature). The importance of preventing the tipping point rather than adjusting to it in this case is also evident in the anticipating vs myopic SCC: the anticipating peak is more than quadruple the myopic peak, reflecting the extremely costly emissions in anticipation of the event.

As for the Stage 1 to Stage 2 adaptation, there is only a slight change in policy. One can observe that the savings rate in Stage 2 retraces the Stage 1 policy after an adjustment period, but the capital stays lower due to the decreased production output.

Decomposition of the SCC

In the section "Social Cost of Carbon", we decomposed the growth rate of the Social Cost of Carbon into the contribution of different terms (see equations 3.6 and 3.7). In both stages we can separate a backloading term,

$$Q_K - \delta_K - \frac{Q_\mu}{E_\mu} E_K$$

representing the effect of capital and the production dynamics, and a frontloading term,

$$-\left[\delta_1 \left(\frac{\lambda_{M^{UP}}}{\lambda_{M^{AT}}} - 1\right) + \xi_1 F_{M^{AT}} \frac{\lambda_{T^{AT}}}{\lambda_{M^{AT}}}\right]$$

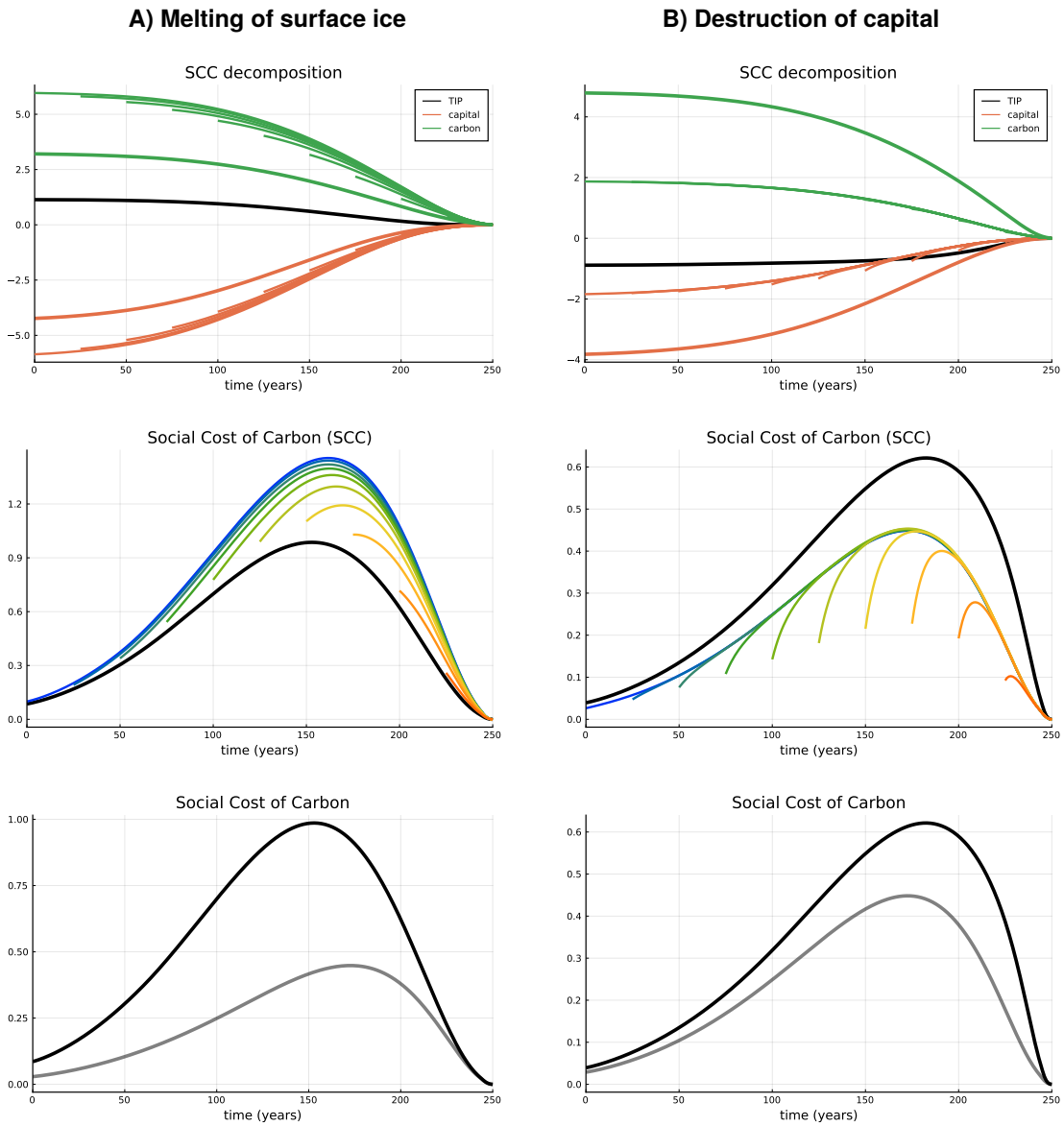
representing the effect of carbon and the geophysical dynamics.

In Stage 1 there is also a term representing the effect of the anticipation of the tipping point:

$$\varepsilon \frac{\xi_K}{\lambda_K} - \frac{\xi_{M^{AT}}}{\lambda_{M^{AT}}}$$

which is in turn made up of a backloading term related to capital and a frontloading term related to carbon. Depending on the scenario, one of the two may prevail, thus making the anticipation of the tipping point either a backloading or a frontloading factor for the SCC.

Here below we show two figures with the SCC decomposition in scenarios A and B, respectively, and the resulting SCCs for reference. The red lines (one for Stage 1 and the other ten for Stage 2) represent the backloading contribution of the capital dynamics, the green lines (again, one for Stage 1 and the other ten for Stage 2) represent the frontloading contribution of the carbon dynamics. The black line (Stage 1 only) represents the anticipating effect of the tipping point: it is frontloading in scenario A, and backloading in scenario B.



We can observe how in scenario A the tipping point's frontloading effect shifts the SCC's Stage 1 peak sooner than it would occur in the myopic scenario and in Stage 2, whereas in scenario B the tipping point's backloading effect shifts the SCC's peak later.

3.1.4 Conclusion

We began our analysis by formulating the continuous-time version of the DICE model; we then included the possibility of a climate tipping point, modeled as a random instant whose hazard rate increases with the global temperature, and whose effect is either a permanent or an instantaneous change in

the system. The resulting problem is an optimal control problem where the planner controls the savings rate and the abatement rate, aiming to maximize the expectation of a 2-stage payoff. Explicitly computing the expectation, we obtained a heterogeneous system where the Stage 1 variables depend on time alone, whereas the Stage 2 variables depend on both time and switching time.

We wrote the necessary conditions for the optimal controls and the co-states, through a version of the Maximum Principle that was derived as a special case of [113]; we then employed such conditions to gain insight about the composition of the Social Cost of Carbon.

We made numerical simulations in four different tipping point scenarios, reporting the relevant figures that highlight the changes to the optimal policies in anticipation of the tipping point (compared to a myopic scenario) and in the new regime in response to the tipping point. Thanks to the analytical decomposition of the SCC we were also able to identify – through backward numerical integration – which factors have a frontloading/backloading effect on the SCC.

The planner's takeaway from our simulations is the importance of prevention. Whereas in some scenarios there is no significant adjustment in the optimal policy upon the regime shift, in all of them we see some degree of prevention effort to fend off the tipping point through higher abatement rates (compared to the myopic scenario). Ideally we want to postpone the regime shift towards the end of the planning horizon or – even better – after the end.

Our results at the moment are purely qualitative. To complete the work we will perform the Stage 2 parameter calibration and the robustness analysis of the effects of the tipping point, in order to make a meaningful comparison of different scenarios.

In future developments we are going to update the current Stage 1 parameters to the most recent DICE values, and also integrate the possibility of negative emissions from a certain time onward.

With the upcoming advancements in the theoretical equipment, we intend to include multiple tipping points and represent the interconnections between them, that is given by feedback mechanisms and domino effects.

3.1.5 Appendix: SCC derivation

Let $V(C, E)$ be the Value function depending on total consumption and emissions. Suppose that the level set of V given by $\{V = V(C^*, E^*)\}$ can be locally parametrized as the graph of a function $C(E)$

$$\{(C(E), E) : E \in (E^* - \varepsilon, E^* + \varepsilon)\}$$

(which is true in a neighborhood of (C^*, E^*) if $V_C(C^*, E^*) \neq 0$).

The Social Cost of Carbon (SCC) is defined as the Marginal Rate of Substitution (MRS) of consumption for emissions:

$$SCC = \frac{dC}{dE}.$$

Since V is constant on the level set,

$$\begin{aligned} 0 &= \frac{dV}{dE}(C(E), E) \\ &= V_C \frac{dC}{dE} + V_E \\ &= V_C \cdot SCC + V_E \end{aligned}$$

which yields

$$SCC = -\frac{V_E}{V_C} = -\frac{\Lambda_E}{\Lambda_C}.$$

Since E and C are not state variables and V does not depend on them explicitly, we identify which quantities that are directly impacting V they influence, compute the derivatives of such quantities with respect to E and C, and then apply the chain rule to V_E and V_C.

E enters the dynamics of M^{AT}:

$$\dot{M}^{AT} = -\delta_1 M^{AT} + \delta_2 M^{UP} + \alpha E$$

therefore the derivative $\frac{dM^{AT}}{dE}(t)$ is to be intended in the sense of Fréchet, i.e. the variation of $M^{AT}(t)$ that is produced by $\Delta E = 1$ in the small interval preceding

t and 0 otherwise ("blip")

$$\begin{aligned}\frac{d}{dt}(M^{AT} + \Delta M^{AT}) &= -\delta_1(M^{AT} + \Delta M^{AT}) + \delta_2 M^{UP} + \alpha(E+1) \\ &= -\delta_1 M^{AT} + \delta_2 M^{UP} + \alpha E - \delta_1 \Delta M^{AT} + \alpha \\ &= \dot{M}^{AT} - \delta_1 \Delta M^{AT} + \alpha\end{aligned}$$

yielding the Cauchy problem

$$\begin{cases} \frac{d}{dt}\Delta M^{AT} = -\delta_1 \Delta M^{AT} + \alpha \\ \Delta M^{AT}(t-\varepsilon) = 0 \end{cases}$$

$$\Delta M^{AT}(t) = \frac{\alpha}{\delta_1}\left(1 - e^{-\delta_1 \varepsilon}\right)$$

$$\begin{aligned}\frac{dM^{AT}}{dE}(t) &= \lim_{\varepsilon \to 0^+} \frac{\Delta M^{AT}(t)}{\varepsilon} \\ &= \frac{\alpha}{\delta_1} \lim_{\varepsilon \to 0^+} \frac{1 - e^{-\delta_1 \varepsilon}}{\varepsilon} \\ &= \alpha\end{aligned}$$

So $V_E = V_{M^{AT}} \frac{dM^{AT}}{dE} = \lambda_{M^{AT}} \alpha$.

C **enters the utility function** $u(c) = u\big(C/L\big)$**:**

$$\frac{du}{dC} = u_c(c)/L$$

The derivative $\frac{dV}{du}$ is again a Fréchet derivative, this time backwards: it is the variation in $V(t)$ that is produced by $\Delta u = 1$ in the small interval following t.

$$\Delta V(t) = \frac{1}{Z_1(t)} \int_t^{t+\varepsilon} e^{-\rho(s-t)} Z_1(s) L(s)\, ds$$

$$\begin{aligned}\frac{dV}{du}(t) &= \lim_{\varepsilon \to 0^+} \frac{\Delta V(t)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \frac{1}{Z_1(t)} \int_t^{t+\varepsilon} e^{-\rho(s-t)} Z_1(s) L(s)\, ds = L(t)\end{aligned}$$

by continuity of the integrand function. So $V_C = \frac{dV}{du}\frac{du}{dC} = u_c(c)$, which, **if** s **is an inner control,** equals λ_K.

Putting everything together,

$$SCC = -\frac{V_E}{V_C} = \alpha\frac{-\lambda_{M^{AT}}}{u_c} = \alpha\frac{-\lambda_{M^{AT}}}{\lambda_K}.$$

Analogous calculations can be carried out for Stage 2, yielding

$$SCC_2 = \alpha\frac{-\xi_{M^{AT}}}{\xi_K}.$$

4

Marketing & Production

This fourth chapter is dedicated to problems related to marketing strategies and production management. In what follows, we consider an application of the theoretical framework of two-stage optimal control problems with a stochastic switching time to a Nerlove&Arrow-type goodwill model and to an offshoring/reshoring model.

The content of Section 4.1 is a paper titled "The Cost of Myopia with Respect to a Switching Time in an Advertising Model". The article is jointly written by A. Buratto,[1] L. Grosset,[1] M. Muttoni,[1] and B. Viscolani.[1] It will be published as a chapter of the book "The unaffordable price of myopia in economics and management" by F. El Ouardighi and G. Feichtinger.

This paper aims to assess the cost of adopting a myopic approach toward system changes. We consider a marketing problem where the demand for a product is influenced by the goodwill of the firm that produces, advertises and sells it. Moreover, we consider the possibility of a sudden and unpredictable increase in production costs. We compare the expected profit of a planner who anticipates such a change against that of a myopic planner who overlooks this possibility. The theoretical setting of this work is that of a finite-horizon, two-stage optimal control problem with a stochastic switching time τ, reformulated with a vintage structure and addressed with Pontryagins Maximum Principle. The hazard rate of τ is assumed to be state-dependent, specifically increasing in the demand (which in turn is an increasing function of the state variable

[1]Dipartimento di Matematica "Tullio Levi-Civita", Università degli Studi di Padova

representing the goodwill). Referring to the notation introduced in Classification 1.2.1, in this model the switch has the following effect:

(E4) **Running cost:** the unit production cost rises in Stage 2.

The content of Section 4.2 is a work-in-progress joint paper titled "Offshoring and reshoring under social and economic uncertainty".

This study examines a firm that is adopting offshoring, the strategic relocation of business processes abroad, and that needs to reconsider its production plans, possibly reshoring its business operations back home. In fact, we assume that unforeseen circumstances, such as wars, pandemics, or social unrests, can disrupt operations abroad, prompting the company to reassess its production strategy. Among the various factors, we also take into account the so-called *Made-in* effect (the preference/aversion to products based on their country of origin) along with rising ethical concerns about labor exploitation and environmental sustainability that can challenge the efficacy of offshoring. The idea of this work is to investigate how the firm should adapt its offshoring policy to these disrupting events.

The theoretical setting of this work is that of an infinite-horizon, two-stage optimal control problem with a stochastic switching time τ, addressed with the backward approach and solved with Dynamic Programming. The hazard rate of τ is assumed to be constant. Referring to the notation introduced in Classification 1.2.1, in this model the switch has the following effect:

(E3) **Dynamics:** the parameter representing the intensity of the *Made-in* effect increases in Stage 2;

(E4) **Running cost:** the unit cost associated with offshoring rises in Stage 2;

(E6) **Switching cost:** the firm incurs in instantaneous costs at τ, that are increasing in the offshore production.

4.1 The Cost of Myopia with Respect to a Switching Time in an Advertising Model

Article written by A. Buratto, L. Grosset, M. Muttoni, and B. Viscolani; to be published as a chapter of the forthcoming book "The unaffordable price of myopia in economics and management" by F. El Ouardighi and G. Feichtinger.

4.1.1 Introduction

Advertising is one of the most effective marketing tools that can influence consumer behavior, leading to changes in demand for a product or a service. There are two primary models in the literature on optimal control applications to advertising, both proposed around the same time. The milestone model in 1957, by Vidale and Wolfe [114], describes the response of sales to advertising and aims to represent typical behaviors observed in real data. The second fundamental model in 1962, by Nerlove and Arrow [85], assumes that the demand and sale intensity of a product depend on a state variable called goodwill, which represents the effects of a firm's investment in advertising. The Nerlove-Arrow model has become an essential reference for advertising and marketing research, as seen in the review articles Sethi [97], Feichtinger et al. [40], and Huang et al. [64]. From this seminal model, we analyze how the optimal advertising campaign adapts in anticipation of a regime shift. In more detail, we consider an advertising problem for a firm assuming that its production costs can disruptively change during the programming interval and affect its (marginal) profits. We assume that the time at which the switch occurs is affected by the demand due to the concept of economies or diseconomies of scale. While it's commonly expected that production costs will decrease as production volume increases (economies of scale), there are situations where the opposite occurs due to various challenges that arise with growth and increased demand. For example, increased demand can lead to raw material shortages, and therefore suppliers might struggle to meet the higher demand. Additionally, a growing demand could lead to labor shortages, particularly if specialized skills are required. This can result in companies needing to offer higher wages or overtime pay to attract workers. In both cases, as a result, production costs can increase abruptly.

The stochastic time corresponding to an increase in production costs can be modeled as a random variable, named the switching time, whose distribution is influenced by the state variable of the system.

Reacting to sudden changes is an important skill for decision makers. Strategic planning and a farsighted perspective are crucial to managing the potential risks associated with irreversible changes in production costs.

In this paper, we compare two different types of behavior, assuming that in any case, the entire advertising campaign must be planned at the beginning of the programming interval. In the first case, we assume that the firm has

complete information about the switching time and can plan how to adjust its advertising campaign for any occurrence of the switching time if such a switch occurs during the programming interval.

In the second case, we analyze a firm that has no information on the time of the switch and plans its advertising strategy as if nothing would change at all.

We denote the latter firm as *myopic with respect to the switching time*, (in short, *myopic*). In the literature, the former type of behavior is called *anticipative* (see Buratto et al. [20]); nevertheless, for a more immediate distinction between the two types of planners, in this paper we shall refer to it as *non-myopic*.

Within the attitude of the myopic firm's with respect to the switching time, we further distinguish two scenarios. First, the decision maker is unable to update its control if the switching time is realized. Due to initial agreements, its advertising campaign will remain as fixed at the initial time. Alternatively, in a second scenario, the decision maker, although myopic, can adapt their strategy to the situation that arises after the switching time.

With our model, we want to analyze the following three research questions:

- How do optimal advertising policies and expected profits vary for the two types of decision maker?
- What is the cost of myopia? It can be quantified by examining the decrease in expected profit, which is directly related to the decrease in the level of knowledge of the decision maker about the time of the switch.
- While the expected profit of the myopic planner is lower than that of the non-myopic planner, there are certain instances where, due to the specific realization of the switching time (e.g., in our model, if the change occurs later on), the actual profit of the myopic planner may be higher than that of the farsighted planner. Therefore, our third research question is: With what probability does the myopic decision maker achieve a higher profit?

This work is organized as follows. In Section 4.1.2, we present a marketing scenario based on the Nerlove-Arrow framework and model its fundamental features, particularly focusing on switching time, to describe a disruptive change in the firm's production costs. In Section 4.1.3, using the necessary conditions for optimal control of heterogeneous systems (see Veliov [113]), we present the necessary conditions for a non-myopic decision maker and find the optimal advertising campaign up to integration of the state-adjoint system of ODEs that is fully nonlinear. In Section 4.1.4, we present the necessary conditions for a myopic decision maker and we find the optimal advertising campaign. In Section 4.1.5, we numerically compare the optimal advertising campaign and the

optimal expected profits. In the Conclusions, we describe some open questions connected with this advertising model.

4.1.2 Model

We consider a finite-time marketing problem in which a company invests in advertising at a rate $a(t) \geq 0$ to increase the demand for its product. The time horizon $[0, T]$ is finite (with $T > 0$), allowing us to set a constant selling price for the product. In fact, we are assuming that during the programming interval, the price remains constant and cannot be modified by the firm. Following the Nerlove and Arrow model, we assume that the *goodwill* summarises the effect of advertising investment; hence $G(t)$ is a state variable which increases with the firm's advertising intensity $a(t)$, and it decays exponentially at a constant rate $\delta > 0$ if not sustained by the advertising:

$$\begin{cases} \dot{G}(t) = a(t) - \delta G(t) \quad \text{for } t \in [0, T] \\ G(0) = G_0 > 0 \end{cases} \tag{4.1}$$

The firm's objective is to maximize its payoff, which is composed of an intertemporal term and a salvage value. The intertemporal term captures the trade-off between the profit from selling the product and the cost of promoting it, whereas the salvage value captures the interest of the firm in sustaining the brand value. The unit production cost $c > 0$ is constant. As mentioned above, we also assume that the unit selling price $p > c$ is constant throughout the programming interval. Let us assume that instantaneous demand depends linearly on the goodwill's value, according to the following formula ($\alpha, \beta > 0$):

$$D(G) = \alpha + \beta G \tag{4.2}$$

Hence, the instantaneous firm's profit from selling the product is $(p-c)D(G(t)) = (p - c)(\alpha + \beta G(t))$ The advertising cost is assumed to be quadratic, which is a standard hypothesis in the related literature, and we denote by $\kappa > 0$ the advertising cost parameter. The salvage value is assumed to be proportional to the final goodwill $G(T)$, with weight $\sigma > 0$. This parameter allows the firm to maximize its brand value even at the end of the programming interval.

Summarizing, the firm wants to solve the following optimal control problem

$$\underset{a(t)\geq 0}{\text{maximize}} \int_0^T \left[(p-c)(\alpha+\beta G(t)) - \frac{\kappa}{2}a^2(t)\right] dt + \sigma G(T) \tag{4.3}$$

subject to (4.1):

$$\begin{cases} \dot{G}(t) = a(t) - \delta G(t) \\ G(0) = G_0 \end{cases}$$

Table 4.1: Variables and parameters

$a(t) \geq 0$	advertising investment at time t (control function)
$G(t)$	goodwill level at time t (state function)
α, β	demand parameters, $\alpha, \beta > 0$
$c > 0$	unit production cost
$p > c$	unit selling price
$\kappa > 0$	parameter for quadratic advertising cost
$\sigma > 0$	marginal weight of the final goodwill
$\delta > 0$	goodwill's depreciation rate

Stochastic switching time: rise in production costs

A stochastic switching time in a dynamic system is an unpredictable event that occurs at a stochastic time τ that abruptly changes the nature of the problem. The instant τ can be modeled as an absolutely continuous random variable with support $[0, +\infty)$. Denoting by Stage 1 the period before the switch, (i.e. all $t \leq \tau$) and Stage 2 the period after the switch, (i.e. all $t > \tau$), we have $c_1 < p$ as a unit production cost in Stage 1, while $c_2 < p$ in Stage 2, with the assumption that $c_1 < c_2$.

In our model, we assume such a change to be also irreversible and the instant τ represents a sudden rise in the production cost, which in turn can be formalized as follows

$$c = \begin{cases} c_1 & \text{in Stage 1} \\ c_2 & \text{in Stage 2} \end{cases} \tag{4.4}$$

The literature on problems related to switching time is wide and many articles tackle the problem of reacting to abrupt changes. A well-established and widely used method to solve this kind of optimal control problems, is given by the *backward approach*, a particular case of the more general theory of piecewise

deterministic optimal control problems (see Dockner et al. [33, ch.8.1]), where the system switches between "modes" at stochastic times, but the dynamics and running payoff in each mode are deterministic. In the backward approach (see, e.g., Boukas et al. [11]), the Stage 2 value function acts as a salvage value for the Stage 1 problem, and the random switching time constitutes a random endpoint for Stage 1.

A different example on an infinite horizon, in Tsur and Zemel [108], assumes the existence of multiple (although with identical hazard rates and effects) catastrophic threats with state-dependent hazards. The authors are very prolific in the stream of literature on two-stage optimal control with stochastic switching time, mostly with an infinite horizon, state-dependent hazard rate, and with a backward approach.

A parallel literature substream, shared by this paper, concerns the same type of problem, tackled with a new solution method, here called *heterogeneous approach*. This method entails formulating an equivalent deterministic optimal control problem, distributed along an additional variable that represents the occurrence of the switching time. Wang [115] and Brogan [14] employ a dynamic programming approach to derive the MP, enabling the interpretation of adjoint variables as shadow prices. Feichtinger et al. in [41] obtain a global maximum principle to tackle this kind of problem, while Veliov [113] provides the necessary conditions for the solution of a more general heterogeneous optimal control problem, which can be applied to this particular case.

More recently Wrzaczek et al. [118] describes the transformation of the two-stage problem into a heterogeneous one, discusses the advantages of this approach compared to the standard backward approach, and provides a simple example on a macroeconomic shock with state-dependent hazard rate in infinite horizon. The same approach is used in Kuhn and Wrzaczek [70], for an infinite-time, two-stage rational addiction model that explicitly incorporates a pre-addiction phase and a stochastic transition into addiction with a state-dependent hazard. In the same sub-stream, Buratto et al. [20] features an infinite-time, two-stage SIR model with lockdown measures and R&D, where the stochastic switch represents the discovery of an effective vaccine. To this day, this is the only published work in which the hazard rate is directly controllable (depending on time and R&D effort). A recent comprehensive overview of dynamic economic problems with regime switches can be found in Haunschmied

et al. [60]. Finally, Freiberger et al. [46][2] present necessary optimality conditions for age-structured optimal control models and apply them to a model on the health impacts of air pollution.

Our work fits into the latter heterogeneous approach substream; however, differently from all the works discussed above, we consider a model in a finite time horizon.

A classical way to tackle optimal control problems with stochastic switching time is to introduce a function called *hazard rate*, or *switching rate*, which describes the probability of the occurrence of such a switch. The hazard rate may be exogenous or endogenous and in the latter case it may depend both on the state variable and the control variable (as, for example, in Sorger [100] and in Dawid et al. [30]) . In this model, we assume the hazard rate to be endogenous and dependent on the demand of the product $D(G)$, in (4.2), therefore, it only depends on the state $G(\cdot)$ of the system.

We assume that the absolutely continuous random variable τ is defined by the following equation:

$$\lim_{h\to 0^+} \frac{\mathbb{P}(\tau \le t+h \,\big|\, \tau > t)}{h} = \eta\big(D(G(t))\big) \tag{4.5}$$

where $\eta : (0,+\infty) \to (0,+\infty)$ is the *hazard rate* function. The distribution of τ can be derived from the definition of η, however, since the goodwill function $G(t)$ is not defined after T, we can determine the distribution of τ only within the programming interval. Nevertheless, we are not interested in the distribution of τ after T: It suffices to know that it can be extended in any way so that the total integral of τ's probability density equals 1 in $[0,+\infty)$.

In the following, we assume that the hazard rate is a linear and increasing function of the demand

$$\eta(D) = \varepsilon D = \varepsilon(\alpha + \beta G), \qquad \varepsilon > 0. \tag{4.6}$$

where ε represents the marginal hazard with respect to the demand.

Since we consider a finite time horizon, τ could occur during the programming interval or later. If it occurs before the final time T, it splits the planning

[2]which will appear as a Chapter of the same book "The unaffordable price of myopia in economics and management" by F. El Ouardighi and G. Feichtinger.

horizon into a Stage 1 and a Stage 2, respectively before and after τ; if it occurs after T, the entire planning horizon is covered by Stage 1. The latter case implies that, with positive probability, the unit production price can be equal to c_1 throughout the whole programming interval.

Switching time: Information and adaptability

In this work, we study the key role that information plays in planning the optimal strategy under the uncertainty of a pending switching time. Our aim is to analyze how the ability to adapt strategies based on the available level of information allows to increase profits, emphasizing the importance of dynamic methods for accurate managerial prescriptions. We do so by considering two types of decision makers.

The first type of decision maker is familiar with both the goodwill's dynamics and the influence of their control on the evolution of the probability distribution of the switching time. In addition, they anticipate the effects that the switch will have on the system. This leads them to define, for Stage 1, a control that covers the entire programming interval, because they do not know, a priori, when the switch will occur. This control will be truncated as soon as the switch occurs and Stage 2 starts. Hereafter, we will denote by *Non-myopic* a decision maker belonging to this first type.

Concerning Stage 2, non-myopic decision makers plan, at the beginning of the programming interval, a strategy that adapts to the realization of the random variable, so that they determine a family of controls, parameterized by the realization of the switching time, each of them defined in the time interval after the switch. In other words, the Stage 2 control for $\tau = s$ is defined in the Stage 2 interval $[s, T]$.

Let us delve into the details, trying to establish the mathematical framework that allows us to describe this problem.

- Stage 1: the planner is expecting τ to occur at any time, knows its hazard rate function $\eta(D(G(t)))$ and the effect it will have on the system, that is, the future production cost c_2. Therefore, they will have to balance the increase in goodwill and consequently in demand with the probability of sudden increases in production costs. Since the planner does not know exactly when τ will occur, their Stage 1 advertising strategy needs to cover the whole programming interval $[0, T]$: the Stage 1 process will be described by the following couple of functions

$$(a_1(t),\ G_1(t)) \ \text{ for } t \in [0, T]$$

- Stage 2: the planner notices when τ occurs and can update their strategy according to the new regime in the interval $[\tau, T]$. We assume that the decision maker establishes a parametric control function at the beginning of the process. Since different realizations of τ may lead to different optimal Stage 2 strategies, the Stage 2 process will actually be parametrized by the realization s of the switching time τ during the planning horizon:

$$(a_2(s,t),\ G_2(s,t)),\ \text{ for } (s,t) \in \Delta := \left\{(s,t) \,\middle|\, s \in [0,T],\ t \in [s,T]\right\}$$

The firm plans their strategy ahead for both stages and for every possible occurrence of τ: before the programming interval starts they will have decided both $a_1(t)$ and $a_2(s,t)$. If τ occurs at time s, the firm will implement the strategy $a_1(t)$ for $t \in [0,s]$, and then the strategy $a_2(s,\cdot) : [s,T] \to [0,+\infty)$.

Remark 4.1. We emphasize that throughout the remainder of this paper we will use this notation for the state and control functions of Stage 2: the first variable s represents the realization of the switching time, while the second variable $t \in [s,T]$ represents the time.

A second type of decision maker is familiar with the goodwill's dynamics but ignores the possibility of a switching time. They choose an advertising strategy that covers the entire programming interval $[0,T]$. This strategy is fixed at the beginning of the programming interval. The control and state functions are then represented by

$$(a(t), G(t)),\ \text{ for } t \in [0,T].$$

In what follows, we will denote *Myopic* this second type of decision maker.

It is worth observing that even though myopic decision makers do not consider the switching time, such an event can indeed occur. Therefore, recalling that our model is based on the assumption that the hazard rate depends on the demand for the product, myopic decision makers determine their strategies without knowing that their controls influence the distribution of the random variable τ. To properly compare the two types of decision makers, we will calculate the expected profit both for non-myopic and myopic decision makers, although taking into account that the latter are unaware of the randomness of the system. Indeed, the non-myopic planners have all the information about the system's stochasticity, so they are able to compute the optimal strategy to maximize the expected payoff over all possible realizations of the switching time. On the other hand, myopic planners maximize their profit as if no switch were to occur; however, even if they do not consider it, their actions do influence the switch's hazard rate, and their production cost will indeed increase at

some point. Knowing this, we can evaluate the *actual* expected profit of myopic planners, which will necessarily differ from their objective value.

For simplicity, it is convenient to perform the computation of the profit first for the non-myopic decision maker. We will see that it is possible to treat the myopic case as a formal instance of the non-myopic case.

4.1.3 Non-myopic decision maker

Starting for the benchmark model (4.3), in order to formalize the switching time optimal control problem for the non-myopic decision maker, we need to introduce some notation. Planners who are aware that a change in marginal costs can occur aim to maximize their expected profit, in the set of feasible control paths $a(\cdot)$; so that the density probability $\mathbb{E}_{a(\cdot)}$ is needed. More precisely, since it is the control used in the Stage 1 that modifies the state variable and, in turn, the distribution function of the random variable τ, this dependence must be indicated on the expected value operator. The formulation of the problem has to take into account both the possibility that the switch occurs before the final time T of the programming period or after it. With such an attempt, let us introduce the indicator functions $\mathbb{1}_{\{\tau<T\}}$ and $\mathbb{1}_{\{\tau\geq T\}}$, respectively. In the first case, there will be a Stage 1, with a payoff equal to (4.3) and a Stage 2 with a payoff with a new marginal cost c_2. On the other hand, in the latter case, i.e., if the switch occurs after the end of the programming period, nothing will change, and the problem essentially remains equal to the one in (4.3) with the original marginal cost c_1. Finally, the problem of the non-myopic decision maker can be stated as follows.

$$\begin{aligned}\underset{a_1(t),a_2(s,t)\geq 0}{\text{maximize}}\ \mathbb{E}_{a_1(t)}\Bigg[&\mathbb{1}_{\{\tau<T\}}\Bigg\{\int_0^{\tau}\left[(p-c_1)D(G_1(t))-\frac{\kappa}{2}a_1(t)^2\right]dt\\&+\int_{\tau}^{T}\left[(p-c_2)D(G_2(\tau,t))-\frac{\kappa}{2}a_2(\tau,t)^2\right]dt+\sigma G_2(\tau,T)\Bigg\}\\&+\mathbb{1}_{\{\tau\geq T\}}\Bigg\{\int_0^{T}\left[(p-c_1)D(G_1(t))-\frac{\kappa}{2}a_1(t)^2\right]dt+\sigma G_1(T)\Bigg\}\Bigg]\end{aligned}\tag{4.7}$$

subject to:

$$\begin{cases} \dot{G}_1(t) = a_1(t) - \delta G_1(t), & \text{for } t \in [0, \tau) \\ G_1(0) = G_0 & \\ \dot{G}_2(\tau, t) = a_2(\tau, t) - \delta G_2(\tau, t), & \text{for } t \in [\tau, T] \\ G_2(\tau, \tau) = G_1(\tau) & \end{cases} \tag{4.8}$$

where the hazard rate of τ is defined in (4.5) and, with a common abuse of notation, see e.g. Wrzaczek et al. [118], we have written $\dot{G}_2(s,t) = \partial_t G_2(s,t)$.

We emphasize that the probability law that governs the process also depends on the chosen control $a(t)$ and henceforth, in accordance with Sorger [100] and Dockner et al. [33, p.204], we write $\mathbb{E}_{a(t)}$ to denote the expectations computed with respect to that law.

To study this problem, we need to compute the expectation through the probability density of τ. For this purpose, it is convenient to introduce an auxiliary state variable $z_1(t) = \mathbb{P}(\tau > t)$ for Stage 1. It represents the probability of still being in Stage 1 at time t. This definition allows us to write the probability density of τ, which is the derivative of $\mathbb{P}(\tau \leq t) = 1 - \mathbb{P}(\tau > t)$, as:

$$f_\tau(t) = -\dot{z}_1(t). \tag{4.9}$$

Following the same computation performed in Wrzaczek et al. [118], we can prove that $z_1(t)$ is the solution of the following Cauchy problem:

$$\begin{cases} \dot{z}_1(t) = -\eta(D(G_1(t))z_1(t), & \text{for } t \in [0, T], \\ z_1(0) = 1. & \end{cases} \tag{4.10}$$

These two results allow us to write explicitly the expected value introduced in the objective functional. After basic integral manipulation, the firm's objective functional becomes:

$$\begin{aligned} &\int_0^T z_1(t)\left[(p - c_1)D(G_1(t)) - \frac{\kappa}{2}a_1(t)^2\right] dt + z_1(T)\sigma G_1(T) \\ &+ \int_0^T \eta\big(G_1(s)\big)z_1(s)\Big\{ \int_s^T \left[(p - c_2)D(G_2(s,t)) - \frac{\kappa}{2}a_2(s,t)^2\right] dt + \sigma G_2(s,T)\Big\} ds \end{aligned}$$

It is worth observing how the auxiliary state variable $z_1(t)$ acts as a discount factor for the Stage 1 payoff. In order to be able to treat this maximization

problem with the theory provided in Veliov [113], we first need to separate the payoff into two additive terms containing Stage 1 and Stage 2 variables. The problem with the above formulation is that the Stage 2 payoff (starting from s) is multiplied by $\eta\big(G_1(s)\big)z_1(s)$, which depends on the Stage 1 variables G_1 and z_1. We work around this as in Wrzaczek et al. [118], by introducing the auxiliary Stage 2 variable $z_2(s,t) = \eta\big(G_1(s)\big)z_1(s)$, i.e.,

$$\begin{cases} \dot{z}_2(s,t) = 0 \\ z_2(s,s) = \eta\big(D(G_1(s))\big)z_1(s) \end{cases} \tag{4.11}$$

The variable $z_2(s,t)$ depends on the switching time s and it is constant in time t. It represents the probability density of τ at time s. After substituting z_2 in the objective functional, we obtain

$$\begin{aligned} &\int_0^T z_1(t)\big[(p-c_1)D(G_1(t)) - \frac{\kappa}{2}a_1(t)^2\big]\,dt + z_1(T)\sigma G_1(T) \\ &+ \int_0^T \Big\{ \int_s^T z_2(s,t)\big[(p-c_2)D(G_2(s,t)) - \frac{\kappa}{2}a_2(s,t)^2\big]\,dt + z_2(s,T)\sigma G_2(s,T)\Big\}\,ds. \end{aligned}$$

Problem (4.7) can be reformulated as a deterministic, heterogeneous one:

$$\begin{aligned} \underset{a_1(t),a_2(s,t)\geq 0}{\text{maximize}} \Bigg[&\int_0^T z_1(t)\big[(p-c_1)D(G_1(t)) - \frac{\kappa}{2}a_1(t)^2\big]\,dt + z_1(T)\sigma G_1(T) \\ &+ \int_0^T \Big\{ \int_s^T z_2(s,t)\big[(p-c_2)D(G_2(s,t)) - \frac{\kappa}{2}a_2(s,t)^2\big]\,dt \\ &\qquad + z_2(s,T)\sigma G_2(s,T)\Big\}\,ds \Bigg] \end{aligned} \tag{4.12}$$

subject to:

$$\begin{cases} \dot{G}_1(t) = a_1(t) - \delta G_1(t), & G_1(0) = G_0, \\ \dot{z}_1(t) = -\eta\big(D(G_1(t))\big)z_1(t), & z_1(0) = 1, \\ \dot{G}_2(s,t) = a_2(s,t) - \delta G_2(s,t), & G_2(s,s) = G_1(s), \\ \dot{z}_2(s,t) = 0, & z_2(s,s) = \eta\big(D(G_1(s))\big)z_1(s). \end{cases} \tag{4.13}$$

We have transformed the optimal control problem with stochastic switching time described at the beginning of this section into a heterogeneous deterministic

optimal control problem. The idea now is to characterize its optimal solutions with necessary conditions.

Theorem 4.2. *Let* $\big(a_1^*(t), G_1^*(t), z_1^*(t), a_2^*(s,t), G_2^*(s,t), z_2^*(s,t)\big)$ *be the optimal solution of the heterogeneus problem* (4.12) *and* (4.13) *for the non-myopic decision maker, then the optimal advertising efforts* $a_1^*(t), a_2^*(s,t)$ *in Stage 1 and stage 2 respectively are:*

$$a_1^*(t) = \big[\lambda_G(t)/\kappa\big]^+, \qquad a_2^*(s,t) = \xi_G(s,t)/\kappa, \tag{4.14}$$

where

$$\xi_G(s,t) = \frac{(p-c_2)\beta}{\delta}(1-e^{-\delta(T-t)}) + \sigma e^{-\delta(T-t)}, \quad t \geq s \tag{4.15}$$

and the following co-state system holds:

$$\begin{cases} \dot{\lambda}_G(t) = -(p-c_1)\beta + \delta\lambda_G(t) - \varepsilon(\alpha+\beta G_1^*(t))\big[\xi_G(t,t) - \lambda_G(t)\big] - \\ \qquad\qquad -\varepsilon\beta\big[\xi_z(t,t) - \lambda_z(t)\big] \\ \lambda_G(T) = \sigma \\ \dot{\lambda}_z(t) = -\big((p-c_1)D(G_1^*(t)) - \frac{\kappa}{2}a_1^*(t)^2\big) - \varepsilon(\alpha+\beta G_1^*(t))\big[\xi_z(t,t) - \lambda_z(t)\big] \\ \lambda_z(T) = \sigma G_1^*(T) \\ \dot{\xi}_z(s,t) = -\big((p-c_2)D(G_2^*(s,t)) - \frac{\kappa}{2}a_2^*(s,t)^2\big) \\ \xi_z(s,T) = \sigma G_2^*(s,T) \end{cases} \tag{4.16}$$

Proof. Let us denote by λ_G and λ_z the Stage 1 co-state variables, and by ξ_G and ξ_z the Stage 2 ones, as in Buratto et al. [18].[3] By Veliov [113], the maximality condition for the Stage 2 control is:

$$a_2^*(s,t) \in \arg\max_{\mathsf{a}\geq 0}\left\{(p-c_2)D\big(G_2^*(s,t)\big) - \frac{\kappa}{2}\mathsf{a}^2 + \xi_G(s,t)\big[\mathsf{a} - \delta G_2^*(s,t)\big]\right\},$$

yielding

$$a_2^*(s,t) = \big[\xi_G(s,t)/\kappa\big]^+.$$

Concerning Stage 1, since the initial condition of the Stage 2 state variables and the hazard rate function do not depend on the control, the necessary condition

[3]For simplicity, we omit the superscript "c" for the current-value co-state functions that correspond to the state variable G.

for the control is a maximality condition of the following form (see also Buratto et al. [18]):

$$a_1^*(t) \in \arg\max_{\mathsf{a}\geq 0} \left\{(p-c_1)D\big(G_1^*(t)\big) - \frac{\kappa}{2}\mathsf{a}^2 + \lambda_G(t)\big[\mathsf{a} - \delta G_1^*(t)\big]\right\},$$

yielding

$$a_1^*(t) = \big[\lambda_G(t)/\kappa\big]^+.$$

We obtain the co-state system from the more general formulation in Buratto et al. [18], with $\phi(t,G,a) = G$, due to the continuity of the goodwill upon the switch, and hence $\partial_G\phi(t,G,a) = 1$. The co-state functions, recalling that the motion equation is the same in the two stages, satisfy the following system.

$$\begin{cases} \dot\lambda_G(t) = -(p-c_1)D'(G_1^*(t)) + \delta\lambda_G(t) - \eta\big(D(G_1^*(t))\big)\big[\xi_G(t,t) - \lambda_G(t)\big] - \\ \qquad\qquad -\frac{d}{dG}\big[\eta(D(G_1^*(t)))\big]\big[\xi_z(t,t) - \lambda_z(t)\big] \\ \lambda_G(T) = \sigma \\ \dot\lambda_z(t) = -\big((p-c_1)D(G_1^*(t)) - \frac{\kappa}{2}a_1^*(t)^2\big) - \eta\big(D(G_1^*(t))\big)\big[\xi_z(t,t) - \lambda_z(t)\big] \\ \lambda_z(T) = \sigma G_1^*(T) \\ \dot\xi_G(s,t) = -(p-c_2)D'(G_2^*(s,t)) + \delta\xi_G(s,t) \\ \xi_G(s,T) = \sigma \\ \dot\xi_z(s,t) = -\big((p-c_2)D(G_2^*(s,t)) - \frac{\kappa}{2}a_2^*(s,t)^2\big) \\ \xi_z(s,T) = \sigma G_2^*(s,T) \end{cases}$$

Observe that the Cauchy problem for ξ_G can be solved independently from the other ones, obtaining equation (4.15). The remaining equations, recalling from (4.2) and (4.6) that $D'(G_1^*) = \beta$ and $\frac{d}{dG}\big[\eta(D(G_1^*))\big] = \varepsilon\beta$, constitute the co-state system (4.16).

It can be easily proved that $\xi_G(s,t) > 0$ being $\sigma > 0$, thus guaranteeing the positivity of the Stage 2 advertising effort $a_2^*(s,t)$ in (4.14). □

Since, in both stages, the optimal control depends solely on the co-state corresponding to the goodwill, it is of interest to analyze the evolution of such co-states λ_G and ξ_G (see system (4.1.3)), and the role played by the anticipation of the switch in shaping such evolution.

Regarding ξ_G, we observe that its adjoint equation (for any fixed s) is the same as it would be in the case of a single-stage problem with production cost equal to c_2. Indeed, after the switch has occurred, there is no uncertainty about future disruptions, and therefore the planner may equivalently be facing a new simple single-stage optimal control problem on the interval $[s, T]$.

As for λ_G, if we compare its adjoint equation with that in the case of a single stage problem with production cost equal to c_1, we notice the presence of two additional terms:

$$-\eta\big(D(G_1^*(t))\big)\big[\xi_G(t,t) - \lambda_G(t)\big] \qquad \text{and} \qquad -\varepsilon\beta\big[\xi_z(t,t) - \lambda_z(t)\big]. \tag{4.17}$$

These terms represent the anticipating effect on the Stage 1 shadow value of the goodwill. Let us illustrate their meaning, starting from the latter term.

By integrating the equation for ξ_z, we obtain that $\xi_z(s,t)$ equals the optimal value of the Stage 2 problem starting from t (with $t \geq s$), given that the switch occurred at time s:

$$\xi_z(s,t) = \int_t^T \Big[(p - c_2)D(G_2^*(s,\theta)) - \frac{\kappa}{2}a_2^*(s,\theta)^2\Big]\, d\theta + \sigma G_2^*(s,T) \tag{4.18}$$

By integrating the equation for λ_z, we obtain that $\lambda_z(t)$ equals the optimal expected value of the two-stage problem starting from t, given that the switch

has not occurred yet at time t:

$$\begin{aligned}
\lambda_z(t) &= \frac{1}{z_1^*(t)}\Big\{ \int_t^T z_1^*(\theta)\big[(p-c_1)D(G_1^*(\theta)) - \frac{\kappa}{2}a_1^*(\theta)^2 \\
&\qquad + \eta\big(D(G_1^*(\theta))\big)\xi_z(\theta,\theta)\big]\, d\theta + z_1^*(T)\sigma G_1^*(T)\Big\} \\
&= \frac{1}{z_1^*(t)}\Big\{ \int_t^T z_1^*(\theta)\big[(p-c_1)D(G_1^*(\theta)) - \frac{\kappa}{2}a_1^*(\theta)^2 \\
&\quad + \eta\big(D(G_1^*(\theta))\big)\Big(\int_\theta^T \big[(p-c_2)D(G_2^*(\theta,u)) - \frac{\kappa}{2}a_2^*(\theta,u)^2\big]\, du + \sigma G_2^*(\theta,T)\Big)\big]\, d\theta \\
&\quad + z_1^*(T)\sigma G_1^*(T)\Big\} \\
&= \mathbb{E}\Big[\chi_{\tau<T}\Big(\int_t^\tau \big[(p-c_1)D(G_1^*(\theta)) - \frac{\kappa}{2}a_1^*(\theta)^2\big]\, d\theta \\
&\qquad + \int_\tau^T \big[(p-c_2)D(G_2^*(\tau,\theta)) - \frac{\kappa}{2}a_2^*(\tau,\theta)^2\big]\, d\theta + \sigma G_2^*(\tau,T)\Big) \\
&\quad + \chi_{\tau\geq T}\Big(\int_t^T \big[(p-c_1)D(G_1^*(\theta)) - \frac{\kappa}{2}a_1^*(\theta)^2\big]\, dt + \sigma G_1^*(T)\Big)\,\Big|\, \tau > t\Big]
\end{aligned}$$

To write this in terms of value functions, we denote (as in Buratto et al. [18]) $V_2(s,t,G)$ for Stage 2 and $V(t,G)$ for Stage 1[4] and obtain

$$\xi_z(s,t) = V_2(s,t,G_2^*(s,t)), \qquad \lambda_z(t) = V(t,G_1^*(t)). \tag{4.19}$$

Having observed this, we can interpret the difference $\big[\xi_z(t,t) - \lambda_z(t)\big]$, which occurs frequently in the co-state system, as the "desirability" of the switch at time t. Let us denote it by $\lambda_\tau(t)$ (recall that $G_2^*(t,t) = G_1^*(t)$):

$$\begin{aligned}
\lambda_\tau(t) :&= \xi_z(t,t) - \lambda_z(t) \\
&= V_2(t,t,G_1^*(t)) - V(t,G_1^*(t))
\end{aligned} \tag{4.20}$$

The reasoning behind this interpretation is that $\lambda_\tau(t)$ measures the expected gain in profit if the switch were to occur at time t, given that it has not occurred

[4] In the cited paper, the Stage 1 problem which originates from plugging V_2 into the Stage 1 objective value (backward approach) is not solved, as it is a simple single-stage problem that the reader can solve with either dynamic programming or Pontryagin's Maximum Principle. With dynamic programming, one obtains a value function of the form $V(t,G,z) = zV^c(t,G)$. With the same abuse of notation as we employed for the co-state variables, we denote $V(t,G)$ the current value function $V^c(t,G)$.

up to t. So, for example, $\lambda_\tau(t) < 0$ means that the profit that would be realized if the switch were to occur at time t is lower than the expected profit, given that the system is still in Stage 1 at time t. This can be intuitively translated as "at time t, the switch is undesirable". Viceversa, if $\lambda_\tau(t) > 0$, the switch is desirable because the realized profit if $\tau = t$ is higher than the expected profit conditional on $\tau > t$ (i.e., the system is still in Stage 1 at time t).

In light of this, an undesirable (resp., desirable) switch has a backloading (resp., frontloading) effect on the Stage 1 goodwill's shadow value λ_G (and therefore on a_1) that is proportional to the marginal hazard with respect to G and the desirability of the switch. Intuitively, a farsighted planner will postpone (resp., advance) his advertising effort compared to a myopic planner, knowing that the switch will have a negative (resp., positive) effect on their optimal payoff. In this model, where $c_2 > c_1$, the switch turns out to be undesirable,

By comparing the ODEs for ξ_G and λ_G, with the PDEs for the Value functions V_2 and V, one can prove that

$$\xi_G(s,t) = \partial_G V_2(s,t,G_2^*(s,t)), \qquad \lambda_G(t) = \partial_G V(t, G_1^*(t)) \tag{4.21}$$

and therefore (recalling that $G_2^*(t,t) = G_1^*(t)$)

$$\xi_G(t,t) - \lambda_G(t) = \partial_G \left[V_2(t,t,G_1^*(t)) - V(t,G_1^*(t)) \right] \tag{4.22}$$

which is the expected marginal gain in profit from switching at time t (given that the switch has not occurred yet) for a unit increment in the goodwill G at time t. A negative expected marginal gain (that is, $\xi_G(t,t) - \lambda_G(t) < 0$) means that, on average, at time t, a (slightly) higher goodwill than $G_1^*(t)$ would make switching at time t (slightly) less convenient. For example, in the case that switching at time t is undesirable for a given value of $G_1^*(t)$, (see previous paragraph), it may become even more undesirable if the goodwill were greater than $G_1^*(t)$.

From a different perspective, the same conclusion can be reached by observing that $\xi_G(t,t) - \lambda_G(t) = \kappa \left[a_2^*(t,t) - a_1^*(t) \right]$ If $\xi_G(t,t) - \lambda_G(t) < 0$ (resp., $>$), then the anticipation of the switch has a backloading (resp., frontloading) effect on λ_G (and therefore on a_1). Intuitively, a non-myopic planner will postpone (resp., hasten) his advertising compared to a myopic planner, knowing that a higher goodwill would make the switch less (resp., more) convenient.

Remark 4.3. It is worth highlighting how the special structure of this model

simplifies the necessary conditions and, as a consequence, the solution of the switching-time problem. Indeed, the adjoint equation for ξ_G does not depend explicitly on s because the problem is autonomous. Additionally, due to the linear state structure of the model, $G_2^*(s,t)$ (which would introduce an implicit, non-trivial dependence on s) does not factor into the equation either. Consequently, the dependence of the solution $\xi_G(s,t)$ on s is trivial, as is that of the strategy $a_2^*(s,t)$.

In order to characterize the optimal advertising efforts, we need to solve the forward-backward system of the state and co-state dynamics constituted by (4.13) and (4.16). Observe that such a system is nonlinear, due for example to the presence of the multiplicative term between η (depending on G) and λ_G, therefore, we resort to a numerical solution. In Section 4.1.5 we report some graphics with the optimal controls, states, costates, and payoffs obtained with the numerical simulations.

The problem we have discussed so far relates to a planner who has all the information about the upcoming τ (hazard rate function and effects on the system) and has the ability to update his strategy upon the occurrence of τ. In what follows, we will discuss planners with no information about τ or without the ability to update their strategy to the new regime after τ. We will see how each of them has a specific functional objective, leading to different optimal strategies. This, of course, entails different expected payoffs, which, intuitively, are increasing with the planner's knowledge/ability.

4.1.4 Myopic decision makers

A planner is considered myopic if it does not take into account the possibility of a switch. Technically, myopic decision makers consider only the first equation of the system (4.13), and this definition is consistent with the game-theoretical one, by which a myopic player ignores the dynamics of a certain state variable, see e.g. Taboubi and Zaccour [101]. Formally, their problem corresponds to (4.12) and (4.13) in which the hazard rate is identically 0, hence the two auxiliary state variables z_1 and z_2 are constant and equal to 1 and 0, respectively.

Now, the question moves to the potential behavior of a myopic planner after the switch has occurred during the programming period. In light of these considerations, in the following definitions, we introduce two additional features that a myopic planner can have.

Definition 4.4. A decision maker is *myopic with respect to a switching time and is unable to update their strategy* if, at the initial time, they compute their optimal advertising strategy by solving the single-stage problem (4.3) and cannot modify their strategy after the possible realization of the switching time.

In Definition 1, we are assuming that the myopic planner, once realized that τ has occurred and observed a sudden increase in production costs, may not be able to re-evaluate and update the advertising strategy according to the new regime. There may be several reasons for the impossibility to update the strategy to the abrupt event: they may have committed to a long-term advertising campaign or have contractual obligations with advertising agencies. Ultimately, the decision to continue to advertise a product at the same intensity, despite higher production costs, depends on several factors, including market conditions, competitive landscape, brand strategy, and long-term business goals.

Theorem 4.5. *Let* $\big(a^*(t), G^*(t)\big)$ *be the optimal path in the optimal control problem* (4.3) *for decision makers who are myopic regarding switching time and are unable to update their strategy, then the optimal control function is*

$$a^*(t) = \frac{1}{\kappa}\Big(\frac{(p-c_1)\varepsilon\beta}{\delta}(1-e^{-\delta(T-t)}) + \sigma e^{-\delta(T-t)}\Big), \quad t \in [0,T]. \tag{4.23}$$

Proof. Let us solve (4.3) using the necessary standard conditions Grass et al. [52, Th.3.4, p.109]. Let us introduce the Hamiltonian function:

$$H(G, a, \lambda_G) = \big[(p-c_1)\,\varepsilon(\alpha+\beta G) - \kappa a^2/2\big] + \lambda_G(a-\delta G)$$

Maximizing with respect to $a \geq 0$ we obtain

$$a(t) = \big[\lambda_G(t)/\kappa\big]^+$$

with co-state equation satisfying

$$\begin{cases} \dot{\lambda}_G(t) = -(p-c_1)\varepsilon\beta + \delta\lambda_G(t) \\ \lambda_G(T) = \sigma \end{cases}$$

By a direct integration we get

$$\lambda_G(t) = \frac{(p - c_1)\varepsilon\beta}{\delta}(1 - e^{-\delta(T-t)}) + \sigma e^{-\delta(T-t)} > 0.$$

Since $\lambda_G(t) > 0$ for all t, the optimal control turns out to be (4.23). □

Note that while for the non-myopic decision maker the optimal controls are characterized by numerically solving a system of ODEs, the optimal control of the myopic decision maker can be calculated explicitly.

Once the optimal solutions for both non-myopic and myopic planners are obtained, to evaluate the cost of myopia, we need to compare their expected profits. It is worth observing that even though the decision maker is myopic with respect to the switching time, such a random event can still occur, and therefore the profit that we need to consider in the comparison is in any case an expected value. To be precise, we need to use the optimal control $a^*(t)$ to calculate its associated optimal state $G^*(t)$ and then determine the probability distribution of the switching time τ. The differential equation governing the evolution of goodwill is linear, and therefore it is possible to calculate the optimal state function explicitly. Recall that the optimal control for myopic decision makers who do not adapt to the new regime remains the same before and after the switching time. Furthermore, the dynamics does not change upon the swith, so the corresponding objective functional can be deduced by (4.7) where the denomination of the optimal control function $a^*(t)$ and its corresponding optimal state $G^*(t)$ do not change in the two stages.

$$\begin{aligned} J^*_{\text{Myopic}} = \mathbb{E}_{a^*(t)}\Bigg[& \mathbb{1}_{\{\tau<T\}}\Big\{ \int_0^\tau \Big[(p - c_1)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\, dt \\ & + \int_\tau^T \Big[(p - c_2)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\, dt\Big\} \\ & + \mathbb{1}_{\{\tau\geq T\}}\Big\{ \int_0^T \Big[(p - c_1)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\, dt\Big\}\Bigg] + \sigma G^*(T) \end{aligned}$$

Denoting by $f^*_\tau(t)$ the density function of the random variable τ, and by $F^*_\tau(t)$ its cumulative distribution function, we can explicitly calculate the expected value of the profit for the myopic decision maker who cannot adapt after the switch. Using the previous notation, we get

$$J^*_{\text{Myopic}} = \int_0^T f^*_\tau(s)\Big\{ \int_0^s \Big[(p-c_1)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\,dt + \int_s^T \Big[(p-c_2)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\,dt\Big\}\,ds + (1-F^*_\tau(T))\Big\{ \int_0^T \Big[(p-c_1)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\,dt\Big\}\Big] + \sigma G^*(T).$$

After integrating by parts the first line in the expression above we obtain the following:

$$J^*_{\text{Myopic}} = \int_0^T \mathbb{P}(\tau > t)\Big[(p-c_1)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\,dt + \int_0^T \int_s^T f_\tau(s)\Big[(p-c_2)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\,dt\,ds + \sigma G^*(T)$$

So far, we have assumed that the decision maker is unable to adjust its control in response to a change in production cost. Let us now provide a further definition that describes a different feature for a myopic decision maker who can update his control after the occurrence of the switch time.

Definition 4.6. A decision maker is *myopic with respect to a switching time, but is able to update their strategy* if, at the initial time, he computes his optimal advertising strategy by solving the single-stage problem (4.3) and uses this strategy until the possible occurrence of the switching time. After the possible realization of the switching time, he updates his strategy by solving a new optimal control starting from the state level that is achieved at the realization of the switching time.

For this kind of decision maker, we need to provide two controls: one to be used before the realization of the switching time and the other afterwards. It is useful to express these two controls in a way that depends on the random variable τ.

Theorem 4.7. *Let $a^\sharp(t)$ be the optimal strategy for a decision maker who is myopic with respect to a switching time, but is able to update his strategy, then the optimal control is*

$$a^\sharp(t) = \begin{cases} a^*(t) & t \in [0,\tau) \\ \frac{1}{\kappa}\Big(\frac{(p-c_2)\varepsilon\beta}{\delta}(1-e^{-\delta(T-t)}) + \sigma e^{-\delta(T-t)}\Big) & t \in [\tau,T] \end{cases} \tag{4.24}$$

Proof. This results come straightforward from Theorem 2. A crucial point is in the form of the necessary conditions for the problem (4.3). The optimal control depends only on the co-state variable, which in turns is decoupled from the motion equation and can therefore be independently solved backward. □

Using the same notation and the same calculations described in this section, we can find the expected value for the decision maker who is myopic with respect to a switching time but who can update his strategy.

$$\begin{aligned} J^{\sharp}_{\text{Myopic}} = \int_0^T f^*_\tau(s)\Big\{ \int_0^s \Big[(p-c_1)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\, dt \\ + \int_s^T \Big[(p-c_2)D(G^\sharp(s,t)) - \frac{\kappa}{2}a^\sharp(t)^2\Big]\, dt + \sigma G^\sharp_2(s,T)\Big\}\, ds \\ + (1-F^*_\tau(T))\Big\{ \int_0^T \Big[(p-c_1)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\, dt + \sigma G^*(T)\Big\}\Big] \end{aligned}$$

After integrating by parts the first line in the expression above we obtain the following:

$$\begin{aligned} J^{\sharp}_{\text{Myopic}} = \int_0^T \mathbb{P}(\tau > t)\Big[(p-c_1)D(G^*(t)) - \frac{\kappa}{2}a^*(t)^2\Big]\, dt + \mathbb{P}(\tau > T)\sigma G^*(T) \\ + \int_0^T \Big[\int_s^T f_\tau(s)\Big[(p-c_2)D(G^\sharp(s,t)) - \frac{\kappa}{2}a^\sharp(t)^2\Big]\, dt + f_\tau(s)\sigma G^\sharp(s,T)\Big]\, ds \end{aligned}$$

In this section, we have characterized optimal solutions for a myopic decision maker in a closed form. However, to perform the comparison with respect to the non-myopic planner, numerical simulations will be necessary.

It is interesting to observe that in the single-stage optimal control problem solved by the myopic planners in Stage 1, their goodwill co-state functions satisfies the same Cauchy problem (4.16) of the non-myopic planner except for the value of epsilon which is $\varepsilon = 0$. The co-state equations for $\lambda_G(t)$ become

$$\begin{cases} \dot{\lambda}_G(t) = -(p-c_1)\pi + \delta\lambda_G(t) \\ \lambda_G(T) = \sigma \end{cases} \tag{4.25}$$

4.1.5 The cost of Myopia

In this section, our objective is to draw a comparison between a myopic approach and a non-myopic one, addressing the research questions outlined in the Introduction. Specifically, we aim to quantify the cost of adopting a myopic perspective with respect to the switching time. With this attempt, we proceed numerically by assigning fixed values to some parameters and letting the marginal hazard with respect to the demand (ε) take values in $\{0.005, 0.01, 0.04\}$. The higher ε the higher the hazard risk η and therefore the more likely is the switch, which in our model corresponds to an increase in the production cost. The simulations have been produced with an algorithm based on Freiberger [44]. Figure 2 collects the parameters that are kept constant in all numerical simulations.

p	c_1	c_2	α	β	κ	σ	δ
1	.5	.8	1	.5	.3	1	.05

Table 4.2: Parameter values

In the following figures, the horizontal axes denote time (in months).
In the figures presented below, the myopic trajectories are represented by gray lines. The myopic planner who does not adapt will follow these trajectories even in the event of a regime shift. Conversely, the myopic planner who adapts will interrupt such strategies upon occurrence of the switching time. The non-myopic trajectories are depicted by black lines. Just like the myopic planner who adapts to the regime shift, the non-myopic planner will follow these lines throughout Stage 1, until the switching time occurs. Finally, the colored lines represent the Stage 2 behavior for different realizations of τ. In the case of the myopic planner who adapts, these lines are slightly transparent, whereas for the non-myopic planner, they are presented in solid color. It is important to note that the Stage 2 advertising aligns in both scenarios. So, assuming that the switch occurs at a given $\overline{\tau}$, the optimal trajectory can be observed following the black line until the instant $\overline{\tau}$ and then "jumping" to the colored line from $\overline{\tau}$ on.
At a glance, we can observe that in both Figure 4.1 and Figure 4.2, the black lines are lower than the gray lines, regardless of ε. This is because a non-myopic planner knows that higher demand increases the likelihood of a switch, leading to increased costs. Consequently, the optimal advertising efforts of the

non-myopic planner in Stage 1 turn are less intensive compared to those of the myopic planner. A similar pattern is observed in the goodwill trajectories in Figure 4.2.

Moreover, in Figure 4.1 it is worth noting that the myopic trajectories remain constant with changes in ε, while the non-myopic trajectories decrease as ε increases. This indicates that the higher the marginal hazard with respect to the demand, the lower the advertising effort of the non-myopic planner. Again, the goodwill trajectories in Figure 4.2 exhibit the same pattern.

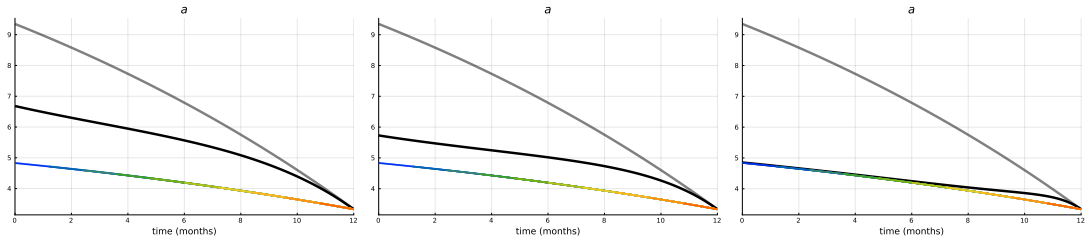


Figure 4.1: Optimal advertising efforts, $\varepsilon = 0.005$, $\varepsilon = 0.01$, $\varepsilon = 0.04$

In Fig. 4.2 each colored line starts at a given switch instant; therefore, the optimal goodwill trajectory can be seen following the black line from the initial time zero until the instant of the switch, while upon that instant, the associated colored line has to be considered. Each figure shows that the later the switch, the higher the optimal goodwill, while comparing the three figures it appears that the higher the marginal risk, the lower the optimal goodwill, as expected.

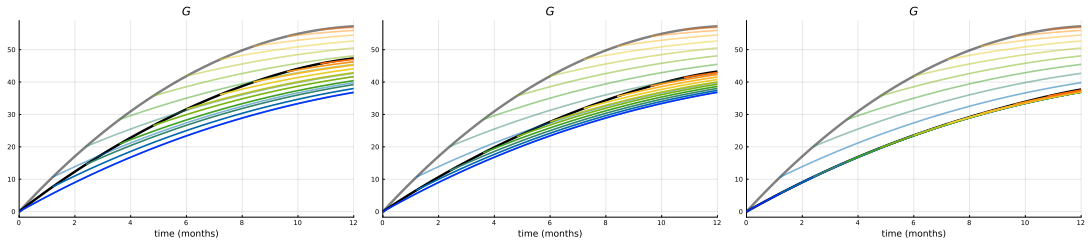


Figure 4.2: Optimal goodwill trajectories, $\varepsilon = 0.005$, $\varepsilon = 0.01$, $\varepsilon = 0.04$

In Fig. 4.3 the objective values related to all the different types of behavior are plotted, with the further specific features the myopic planner can have, referring to Definitions 4.4 and 4.6: Gray=myopic+can update; Light gray=myopic+cannot update. Expected profits (represented by the solid constant lines) and realized profits (represented by the dotted lines) are shown as functions of the realization of the switch time. As before, the black color stands for non-myopic. Naturally,

expected payoffs do not depend on the realization of τ and are therefore represented by constant lines.

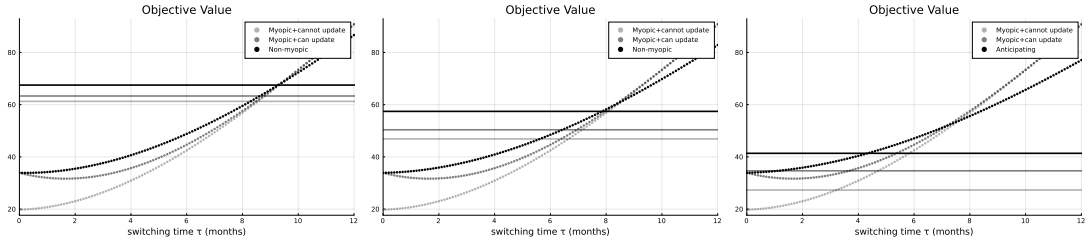


Figure 4.3: Profits; $\varepsilon = 0.005$ ($\tilde{\tau} \approx 9.5$), $\varepsilon = 0.01$ ($\tilde{\tau} \approx 8.5$), $\varepsilon = 0.04$ ($\tilde{\tau} \approx 7.5$)

The task of our analysis is to assess the cost of being myopic with respect to an abrupt switch. Indeed, we found that the expected profit for the non-myopic planner is greater than the ones of the myopic ones. Moreover, among myopic decision planners, the one who is able to update obtains higher profit because they can adapt to the new situation by reacting to the increase in production costs through a decrease in their advertising investment.

However, regarding realized payoffs, for sufficiently small realizations of τ non-myopic payoff lays over the myopic ones, nevertheless, it is a well-known result that there may be late realizations of τ where myopic planners achieve a higher payoff than their non-myopic counterparts. The dotted lines effectively intersect as τ approaches T, and, after such an intersection (let us call it $\tilde{\tau}$), the black dotted lines lay below the gray ones. In any case, this situation occurs with a very low probability, as clearly represented in Fig. 4.4, where the probability (z_1) of still being in Stage 1 is represented as a function of time. In all graphs, this probability decreases with t, consistent with its definition. Moreover, as the hazard risk increases, then the probability of arriving at late realizations of τ that provide a higher profit for the myopic decision planner is very small. For example, from the third graph of Figure 4.3 ($\varepsilon = 0.04$) we can observe that the dark gray dotted line (myopic who can update) intersects the black one in $\tilde{\tau} \approx 7.1$ which in Figure 4.4 (using the myopic gray strategy) leads to $z_1(7.1) \approx 0.02$[5]. In other words, the probability that a myopic decision planner makes a profit greater than the non-myopic one is statistically not significant (less than 0.05). As a final comment on Figure 4.4, we observe that the black lines are higher than

[5]The value of z_1 evaluated in the $\tilde{\tau}$ corresponding to the intersection of the myopic who cannot update (light gray dotted line) is even smaller

the gray because the non-myopic planner takes action to postpone the switch time.

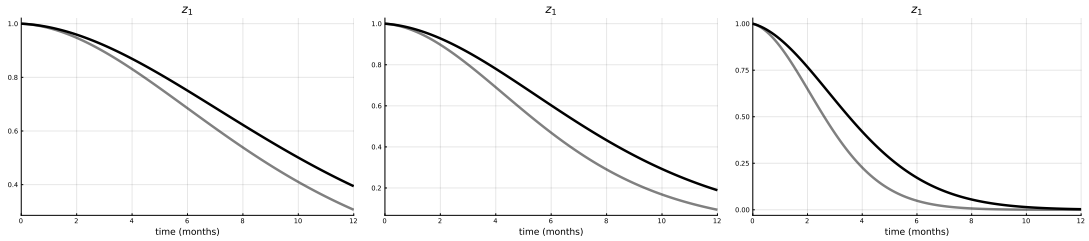


Figure 4.4: $z_1(t)$, $\varepsilon = 0.005$, $\varepsilon = 0.01$, $\varepsilon = 0.04$

4.1.6 Conclusions

This research offers a compelling analysis of the implications of myopia in decision making, particularly in the context of dynamic marketing problems. Our study centers on assessing the cost of a myopic approach when faced with potential abrupt changes in production costs, a scenario increasingly relevant in today's fast-paced and unpredictable market environments.

Through the formulation and solution of an optimal control problem with stochastic switching time, we were able to quantitatively compare the outcomes of a myopic planner with those of a planner who anticipates potential changes.

Referring to the first two research questions, our analysis shows how different decision-making approaches (myopic vs. non-myopic) affect advertising strategies and the resulting profits.

Second, we examine the impact of short-term thinking on expected profits. We evaluated the cost of myopia by assessing how a decrease in knowledge about crucial time-sensitive decisions affects profitability.

Our findings revealed a significant divergence in the profit outcomes between these two approaches. The myopic planner, constrained by a lack of foresight into possible system changes, invariably encountered a reduction in profit when the switch in production costs occurred. This reduction is directly attributable to the planner's inability to adjust strategies preemptively, highlighting the cost of myopia in decision-making.

On the contrary, the non-myopic planner, equipped with the awareness of potential changes, demonstrated a more adaptable and resilient approach. This planner's ability to anticipate and plan for potential disruptions not only min-

imized losses but often led to more optimized use of resources, thereby maximizing profits. These results underscore the importance of strategic foresight in management and planning.

Our third research question:"With what probability does the myopic decision maker achieve a higher profit?" is intriguing, as it challenges the assumption that non-myopic (farsighted) planners always yield higher profits. The first answer can be found by observing that among myopic planners, there are those who can update their strategies after the switch and can indeed aim at achieving higher profits. However, in the cases of late switch realizations, myopic planners can occasionally outperform non-myopic ones. However, this scenario can occur with a low probability. In fact, our simulations prove that when the hazard rate is increased, the probability of myopic planners gaining higher profits in late switch realizations becomes statistically insignificant (less than 0.05 probability). This finding underscores that non-myopic planners, who proactively adapt to delay the switching time, generally yield better outcomes.

In conclusion, this paper highlights the tangible benefits of strategic anticipation and adaptability in dynamic decision-making contexts. It serves as a call to action for planners and managers to cultivate a forward-looking perspective, integrating predictive analytics and scenario planning into their strategic toolkit. By doing so, they can significantly reduce the risks associated with myopia and harness the full potential of their decision-making capabilities in an ever-evolving market landscape.

4.2 Offshoring and reshoring under social and economic uncertainty

4.2.1 Introduction

Over the past decades, many firms have offshored their production to foreign countries to gain high profit margins. Lower labor expenses were among the benefits that offshore locations provided mainly (Jensen and Pedersen [65]), together with other motivations as counter-trade agreements (Nassimbeni et al. [84]). Although transportation costs are involved, production may be cheaper abroad, allowing for higher profit margins. Indeed, the costs issue is one of the main reasons that leads a firm to produce offshore, as observed by Yang et al. [120], who analyze the effect of tariffs and production cost on firms profits under competition and offshoring.

However, in today's dynamic market, the strategic advantage of offshoring can quickly transform into a liability for firms. Recently, in times of global crisis, policy-makers in several western countries have considered bringing back production to their home country (Booth [10]), leading to a phenomenon known as *reshoring* or *backshoring*. Different and more pressing matters, such as political events, the recent global pandemic hit, the start of war conflicts, together with the governments' subsidies for capital expenditure, encouraged firms to reshore their business operations back home, see Ellram et al. [36], Gupta et al. [57], and more recently Henkel et al. [61].

In what follows, we consider some of the factors that may influence the reshoring decision. We discuss, on one hand, those related to firm objectives, such as the costs associated with offshoring. For instance, Radi et al. [91] emphasize how increasing labor costs in a technologically lagging country can reduce its attractiveness and potentially trigger a reshoring process. On the other hand, we consider another aspect, less tangible, which is not included as a cost, but it influences the sales, namely the *Made-in* effect, i.e. the perceived value and quality associated with products that are explicitly labeled as being manufactured in a particular country. In fact, customers' demand may decrease when they realize that the product is made abroad (Di Mauro et al. [32], Sawik [95]).

There may be many drivers that lead companies to repatriate their offshore production, and managers can address the 'location decision dilemma' more

effectively if they properly take into account such factors. For example, reshoring can contribute to mitigating the rising unemployment rates (Tate et al. [103]) and to supporting re-industrialization (Pisano and Shih [89, 88]). As a result of these beneficial impacts, certain governments (such as the United States) are motivated to promote and subsidize repatriations (Guenther [56], Livesey [74]).

Moreover, reshoring production boosts domestic economic growth, enhances supply chain resilience, reduces transportation costs, improves quality control, increases intellectual property protection, fosters innovation, creates local job opportunities, and strengthens the national economy.

Furthermore, some studies have pointed out that geographical and cultural distance can increase the risk of opportunistic behavior by offshore suppliers or proprietary offshore production sites (Kinkel and Maloca [68], McIvor [81], Martínez-Mora and Merino [78]). These threats may cause unsustainable costs to negotiate, manage, and maintain cross-border transactions and, in turn, steer the company towards the reshoring of production activities. Moreover, the firms inability to develop critical assets abroad, or to access and exploit the resources of the host country, can also lead to reshoring (Canham and Hamilton [22], Fratocchi et al. [43]). They also highlight that the decision-making process for the production-location choice is affected by inappropriate selection criteria (e.g. little consideration of total cost analysis) or by the so-called "bandwagon effect" (Abrahamson and Rosenkopf [1]).

Another aspect to consider is the aforementioned *Made-in* effect: if the product's image among the consumers was damaged by offshoring, reshoring can foster the firms ability to create value and maintain competitive advantage through quality and innovation.

Last but not least, among the various factors that may challenge the efficacy of offshoring, we can mention rising ethical concerns about labor exploitation and environmental sustainability.

In short, a firm's decision to reshore may be seen as a simple correction of a previous misjudged decision or as a deliberate response to exogenous or endogenous changes (Tate [103]).

The purpose of this paper is to explore how a firm should consider the risk of a potential regime shift and accordingly adjust its offshoring policy, considering, if necessary, the possibility of reshoring a part or all its production.

In this work, we consider a firm that offshores a fraction of its production to a foreign country. We assume that at a random instant τ, due to social, political or

economic causes, an unpredictable disruptive event occurs, entailing a variety of effects that may potentially hinder the offshoring practice. Taking examples from recent history: the COVID-19 pandemic, the Suez Canal blockage, unforeseen war outbreaks. Therefore, the firm must reconsider the appropriate "shoring" decision, possibly reducing the offshoring policy (partial reshoring) or stopping it entirely (full reshoring).

The problem is formulated as an infinite-horizon, two-stage, discounted optimal control problem with random switching time, see Veliov [113] and Wrzaczek et al. [118], and is solved with the backward approach (Grass et al. [52, p.386]).

To the best of our knowledge, this is the first work in dynamic optimization that explicitly accounts for the *Made-in* effect. While Sawik in [95] considers the *Made-in* effect, his work is formulated within a static context. Conversely, Brambilla et al. [13] features a continuous-time controlled Markov chain with binary states, addressing offshoring and reshoring decisions within a dynamic framework, though without incorporating the *Made-in* effect explicitly.

Among the possible negative effects that may prompt reshoring, together with the amplification of the *Made-in* effect, in the present model we consider the increase in offshoring costs, as well as potential instantaneous damages or penalty costs. This damage/cost, which is inflicted on the firm *una tantum* at the switching time, may capture the physical damage sustained by the firm's offshore facilities, due for example to a war outbreak or a violent manifestation, or the loss of perishable goods due to a disruption in the transportation chain from the offshore country to the home country, or again it could represent assessment costs sustained by the firm in order to analyze the implications that a change in the offshore socio-political scenario may entail on the firm's production. As such, said damage/cost depends on the firm's policy and on the state of things just before the switch happened, and not at all on what happens afterwards.

Research questions. The objective of this research is to investigate how the firm should adjust its offshoring strategy in anticipation of and in response to such a regime switch. The study aims to provide insights into proactive measures that the firm can employ to adapt its offshore operations effectively, mitigating risks, and maintaining competitiveness in a dynamic global environment. We can summarize the research questions as follows:

(i) What is the role played by the anticipation of the switch in shaping the firm's optimal offshoring strategy in the time period before the switch?

(ii) How should the firm adjust its offshoring strategy in response to the regime switch?

The paper is structured as follows. In Section 4.2.2, we formalize the two-stage model in which the firm has some information about the impending event and takes it into account, demonstrating an anticipating behavior. In Section 4.2.3 we solve the problem of the anticipating planner, who aims to maximize the expectation of the total payoff over all possible realizations of the switching time; in Section 4.2.4 we solve the problem of a myopic planner, who on the contrary is unaware of the possibility of a switch and/or does not take it into consideration. This reduces to a simple single stage optimal control problem that can be viewed as a particular case of the anticipating problem where the hazard rate is set to zero. Section 4.2.5 concludes.

4.2.2 Model description and assumptions

Let $u(t) \in [0,1]$ be the fraction of the total output produced offshore and controlled by the firm. We assume that the goods produced are not stocked, therefore, at every time, the total production coincides with sales, which are represented by the state variable $S(t)$. Suppose that, over an infinite time horizon, if no offshoring is implemented, sales converge to the regime level $\sigma > 0$, at rate $\gamma > 0$. In this paper we want to model also the *Made-in* effect by penalizing the sales dynamics with a negative term that is proportional to the offshore production $u(t)S(t)$. The parameter $\lambda > 0$ represents the impact of the *Made-in* effect on sales. The resulting dynamics is

$$\dot{S}(t) = \gamma(\sigma - S(t)) - \lambda u(t)S(t). \tag{4.26}$$

Observe that if $u(t) = 0$, that is if the firm does not produce offshore, then no negative *Made-in* effect can be observed. On the other hand, if $u(t) > 0$ then the firm offshores the amount of production $u(t)S(t)$ abroad, and therefore sales suffer from the *Made-in* effect due to the customers' perception. Nevertheless, due to the linearity of the ODE 4.26, and the positivity of parameters γ and σ, sales are always positive for any given offshoring policy $u(t)$.

As seen above, the cost increase is one of the main issues that the firm must consider in its objectives. An optimal shoring plan requires taking into

account both *domestic* ($\kappa^d > 0$) and *offshoring* ($\kappa^o > 0$) costs. In this model, we assume that the instantaneous comprehensive offshoring cost is linear with respect to the offshore production, and amounts to $\kappa^o u(t)S(t)$. In the same way, the instantaneous production cost in the home country is linear with respect to the domestic production, and amounts to $\kappa^d(1-u(t))S(t)$. We also assume that $\kappa^o < \kappa^d$, otherwise, factoring in the negative *Made-in* effect, it would be trivially optimal to not produce offshore at all.

Goods are sold at a constant unit price p, therefore, assuming $p > \kappa_d$ it follows that the firm revenue is originally $pS(t)$. After subtracting the production costs, split into "d"-omestic and "o"-ffshored, we obtain the profit flow

$$\big[p - \kappa^d(1-u(t)) - \kappa^o u(t)\big]S(t). \tag{4.27}$$

We formulate a discounted problem over an infinite horizon. In this analysis, we also assume that a disruptive external event occurs at a random time τ, with constant hazard rate $\eta > 0$, splitting the programming interval into Stage 1 (before τ) and Stage 2 (after τ). We take into consideration any kind of event of exogenous nature which may entail the following effects on the system:

- An amplification of the *Made-in* effect, represented by an increase in the parameter λ affecting the state dynamics: it equals λ_1 in Stage 1 and $\lambda_2 > \lambda_1$ in Stage 2.
- An increase in the costs associated with offshoring: the unit cost is κ_1^o in Stage 1 and $\kappa_2^o > \kappa_1^o$ in Stage 2.
- Instantaneous damage/cost related to the entity of the offshore production at the time of the event: in this early version of the work we assume this cost to be equal to $\frac{1}{2}\kappa^\tau u^2 S$.

The first two factors will prompt the firm to reconsider its offshoring strategy after the change has occurred. However, the anticipation of each of these three effects, will influence the firm's offshoring strategy right from the start.

The firm plans to maximize the expected discounted profit (see Chapter 1), where the profit flow associated to each stage takes the form as in (4.27):

$$\begin{aligned}\underset{\substack{u_1(t)\in[0,1]\\ u_2(s,t)\in[0,1]}}{\text{maximize}}\,\mathbb{E}\Big[&\int_0^\tau e^{-\rho t}\big[p - \kappa^d(1-u_1(t)) - \kappa_1^o u_1(t)\big]S_1(t)\,dt - \frac{1}{2}\kappa^\tau u_1(\tau)^2 S_1(\tau) \\ &+ \int_\tau^\infty e^{-\rho t}\big[p - \kappa^d(1-u_2(\tau,t)) - \kappa_2^o u_2(\tau,t)\big]S_2(\tau,t)\,dt\Big]\end{aligned} \tag{4.28}$$

subject to

$$\begin{cases} \dot{S}_1(t) = \gamma(\sigma - S_1(t)) - \lambda_1 u_1(t) S_1(t) & S_1(0) = S_0 \\ \dot{S}_2(s,t) = \gamma(\sigma - S_2(s,t)) - \lambda_2 u_2(s,t) S_2(s,t) & S_2(s,s) = S_1(s). \end{cases} \tag{4.29}$$

Knowing that the hazard rate of τ is constant and equal to η, and therefore τ is an exponential random variable with probability density $\eta e^{-\eta t}$, we can compute the expectation explicitly. By assuming optimal behavior in Stage 2, the expected payoff takes the following form:

$$\begin{aligned} \underset{u_1(t)\in[0,1]}{\text{maximize}} \int_0^\infty e^{-(\rho+\eta)t} \Big\{ \Big[p - \kappa^d(1-u_1(t)) - \kappa_1^o u_1(t) - \frac{1}{2}\eta\kappa^\tau u_1(t)^2 \Big] S_1(t) \\ + \eta V_2(S_1(t)) \Big\} \, dt \end{aligned} \tag{4.30}$$

subject to

$$\begin{cases} \dot{S}_1(t) = \gamma(\sigma - S_1(t)) - \lambda_1 u_1(t) S_1(t) \\ S_1(0) = S_0, \end{cases} \tag{4.31}$$

where V_2 denotes the Stage 2 value function. Note that we already know that V_2 depends only on the state. Indeed, whereas V_2 would depend in general also on the switching time realization s and on time t, the infinite horizon together with the independence of the Stage 2 data of t (autonomy) and s ensure that V_2 is independent of those as well.

The Stage 2 problem, to be solved with Dynamic Programming in order to compute the Stage 2 value function V_2, if τ occurs at time s is as follows:

$$\underset{u_2(s,t)\in[0,1]}{\text{maximize}} \int_s^\infty e^{-\rho t} \big[p - \kappa^d(1-u_2(s,t)) - \kappa_2^o u_2(s,t) \big] S_2(s,t) \, dt \tag{4.32}$$

subject to

$$\begin{cases} \dot{S}_2(s,t) = \gamma(\sigma - S_2(s,t)) - \lambda_2 u_2(s,t) S_2(s,t) \\ S_2(s,s) = S_1(s). \end{cases} \tag{4.33}$$

Observe that solving this problem would require knowing $S_1(s)$, unless we opt to solve it for every possible value of $S_2(s,s)$, which requires the application of the theory of Dynamic Programming.

Variables	$(i = 1, 2)$
u_i	portion of production that is offshored (control), $u_i \in [0, 1]$
S_i	total sales, which coincide with total production (state)
Parameters	$(i = 1, 2)$
σ	sales regime with no offshoring
γ	rate of convergence to sales regime
λ_i	impact of offshoring on sales (*Made-in* effect)
κ^d	unit cost associated with domestic production
κ_i^o	unit cost associated with offshore production
κ^τ	marginal switching damage/cost
ρ	discount rate
η	hazard rate

Table 4.3: Variables and parameters

4.2.3 Anticipating-responsive solution

Thanks to the low dimension and simple structure of the problem, we can comfortably address it with the backward approach. As expected, the optimal control turns out to be constant, and both value functions V_1 and V_2 turn out to be affine in the state variable S.

Stage 2

Since the Stage 2 problem is fully autonomous, the Stage 2 value function depends only on the state. The HJB equation for Stage 2 is

$$\rho V_2(S) = \max_{u \in [0,1]} \left\{ [p - \kappa^d(1-u) - \kappa_2^o u]S + V_2'(S)[\gamma(\sigma - S) - \lambda_2 u S] \right\}. \tag{4.34}$$

The control that maximizes the RHS is

$$\begin{cases} u_2 = 1 & \text{if } \kappa^d - \kappa_2^o - \lambda_2 V_2'(S) > 0 \\ u_2 = 0 & \text{if } \kappa^d - \kappa_2^o - \lambda_2 V_2'(S) < 0 \\ u_2 \in [0,1] & \text{otherwise.} \end{cases} \tag{4.35}$$

• Let us suppose first that $\kappa^d - \kappa_2^o - \lambda_2 V_2'(S) < 0$, and hence $u_2 = 0$. The updated HJB turns out to be

$$\rho V_2(S) = (p - \kappa^d)S + V_2'(S)\gamma(\sigma - S), \tag{4.36}$$

whose solution is

$$V_2(S) = \frac{p - \kappa^d}{\rho + \gamma}(S + \gamma\sigma/\rho). \tag{4.37}$$

The condition on $V_2'(S)$ is satisfied if and only if

$$(\rho + \gamma)(\kappa^d - \kappa_2^o) < \lambda_2(p - \kappa^d). \tag{4.38}$$

• Now, let us suppose that $\kappa^d - \kappa_2^o - \lambda_2 V_2'(S) > 0$, and hence $u_2 = 1$. The updated HJB turns out to be

$$\rho V_2(S) = (p - \kappa_2^o)S + V_2'(S)[\gamma(\sigma - S) - \lambda_2 S], \tag{4.39}$$

whose solution is

$$V_2(S) = \frac{p - \kappa_2^o}{\rho + \gamma + \lambda_2}(S + \gamma\sigma/\rho). \tag{4.40}$$

The condition on $V_2'(S)$ is satisfied if and only if

$$(\rho + \gamma)(\kappa^d - \kappa_2^o) > \lambda_2(p - \kappa^d). \tag{4.41}$$

• If $\kappa^d - \kappa_2^o - \lambda_2 V_2'(S) = 0$, then every feasible control maximizes the RHS. The resulting HJB equation is as in (4.36), yielding the same solution as (4.37). The condition on $V_2'(S)$ is satisfied if and only if $(\rho + \gamma)(\kappa^d - \kappa_2^o) = \lambda_2(p - \kappa^d)$.

Putting everything together, the Stage 2 value function is

$$V_2(S) = A(S + \gamma\sigma/\rho), \quad \text{where} \quad A = \max\left\{\frac{p - \kappa^d}{\rho + \gamma}, \frac{p - \kappa_2^o}{\rho + \gamma + \lambda_2}\right\}. \tag{4.42}$$

By the Verification Theorem (see Chapter 1, or Bardi and Capuzzo-Dolcetta [7, ch.1], or Dockner et al. [33, ch.3]), the Stage 2 solution can be summarized by the following proposition:

Proposition 4.8 (Optimal offshoring strategies at Stage 2)**.**

- *If* $(\rho + \gamma)(\kappa^d - \kappa_2^o) < \lambda_2(p - \kappa^d)$*, then the optimal offshoring strategy is no offshoring* ($u_2 \equiv 0$) *with corresponding value function*

$$V_2(S) = \frac{p - \kappa^d}{\rho + \gamma}(S + \gamma\sigma/\rho). \tag{4.43}$$

- *If* $(\rho + \gamma)(\kappa^d - \kappa_2^o) > \lambda_2(p - \kappa^d)$*, then the optimal offshoring strategy is full*

offshoring ($u_2 \equiv 1$), with corresponding value function

$$V_2(S) = \frac{p - \kappa_2^o}{\rho + \gamma + \lambda_2}(S + \gamma\sigma/\rho). \tag{4.44}$$

- *If $(\rho+\gamma)(\kappa^d - \kappa_2^o) = \lambda_2(p - \kappa^d)$, then any feasible offshoring strategy is indifferent, and the resulting value function is equal to both the formulations in the two previous cases.*

The inequalities above can be rewritten in a way that separates the quantities that changed at the switch from those that remained the same, obtaining:

$$\frac{\kappa^d - \kappa_2^o}{\lambda_2} \gtrless \frac{p - \kappa^d}{\rho + \gamma}. \tag{4.45}$$

The responsive planner, who (by definition) intends to reassess the offshoring strategy in adaptation to the new regime, should therefore compare the updated ratio $(\kappa^d - \kappa_2^o)/\lambda_2$ with the domestic profit margin $p - \kappa^d$ divided by the sum of the discount rate and the rate of convergence to the sales regime. Unsurprisingly, the fraction on the LHS can be interpreted as the "convenience of offshoring": it is the difference in the unit costs associated to domestic and offshore production respectively, "weighted" by the negative impact of the *Made-in* effect.

Steady state. Since the time horizon is infinite, any sufficiently large t belongs to Stage 2. Therefore, it makes sense to compute the steady state of S_2. If the optimal offshoring strategy is $u_2(s,t) = 0$, then

$$\begin{cases} \dot{S}_2(s,t) = \gamma(\sigma - S_2(s,t)) \\ S_2(s,s) = S_1(s). \end{cases} \tag{4.46}$$

The solution can be computed explicitly:

$$S_2(s,t) = \sigma + (S_1(s) - \sigma)e^{-\gamma(t-s)} \xrightarrow[t\to\infty]{} \sigma. \tag{4.47}$$

If, otherwise, the optimal offshoring strategy is $u_2(s,t) = 1$, then

$$\begin{cases} \dot{S}_2(s,t) = \gamma(\sigma - S_2(s,t)) - \lambda_2 S_2(s,t) \\ S_2(s,s) = S_1(s), \end{cases} \tag{4.48}$$

and the corresponding trajectory is:

$$S_2(s,t) = \frac{\gamma\sigma}{\gamma+\lambda_2} + \Big(S_1(s) - \frac{\gamma\sigma}{\gamma+\lambda_2}\Big)e^{-(\gamma+\lambda_2)(t-s)} \xrightarrow[t\to\infty]{} \frac{\gamma\sigma}{\gamma+\lambda_2}. \tag{4.49}$$

Observe how the *Made-in* effect (represented by parameter λ_2) impacts the steady state: with no offshoring the steady state is the sales regime σ; with full offshoring the steady state is a fraction $\gamma/(\gamma+\lambda_2) \in (0,1)$ of the sales regime σ. The stronger the *Made-in* effect, the smaller the firm's steady state total sales .

Stage 1

Since the Stage 1 problem and the Stage 2 value function are fully autonomous, the Stage 1 value function depends only on the state S. The HJB equation for Stage 1 is

$$\begin{aligned}(\rho+\eta)V_1(S) = \max_{u\in[0,1]} \big\{&[p-\kappa^d(1-u)-\kappa_1^o u - \eta\kappa^\tau u^2/2]S \\ &+ V_1'(S)[\gamma(\sigma-S)-\lambda_1 uS] + \eta V_2(S)\big\}.\end{aligned} \tag{4.50}$$

The control which maximizes the RHS is

$$\begin{cases} u_1 = 0 & \text{if } \kappa^d - \kappa_1^o - \lambda_1 V_1'(S) < 0 \\ u_1 = 1 & \text{if } \kappa^d - \kappa_1^o - \lambda_1 V_1'(S) > \eta\kappa^\tau \\ u_1 = [\kappa^d - \kappa_1^o - \lambda_1 V_1'(S)]/(\eta\kappa^\tau) & \text{if } \kappa^d - \kappa_1^o - \lambda_1 V_1'(S) \in [0, \eta\kappa^\tau]. \end{cases} \tag{4.51}$$

• Let us suppose first that $\kappa^d - \kappa_1^o - \lambda_1 V_1'(S) < 0$, and hence $u_1 = 0$. The updated HJB turns out to be

$$(\rho+\eta)V_1(S) = (p-\kappa^d)S + V_1'(S)\gamma(\sigma-S) + \eta V_2(S), \tag{4.52}$$

whose solution is

$$V_1(S) = \frac{p-\kappa^d+\eta A}{\rho+\eta+\gamma}[S+\gamma\sigma/(\rho+\eta)] + \frac{\eta\gamma\sigma A}{(\rho+\eta)\rho}. \tag{4.53}$$

The condition on $V_1'(S)$ is satisfied if and only if

$$(\rho+\eta+\gamma)(\kappa^d-\kappa_1^o) < \lambda_1(p-\kappa^d+\eta A). \tag{4.54}$$

• Now, let us suppose that $\kappa^d - \kappa_1^o - \lambda_1 V_1'(S) > \eta\kappa^\tau$, and hence $u_1 = 1$. The updated HJB turns out to be

$$(\rho + \eta)V_1(S) = (p - \kappa_1^o - \eta\kappa^\tau/2)S + V_1'(S)[\gamma(\sigma - S) - \lambda_1 S] + \eta V_2(S), \tag{4.55}$$

whose solution is

$$V_1(S) = \frac{p - \kappa_1^o - \eta\kappa^\tau/2 + \eta A}{\rho + \eta + \gamma + \lambda_1}[S + \gamma\sigma/(\rho + \eta)] + \frac{\eta\gamma\sigma A}{(\rho + \eta)\rho}. \tag{4.56}$$

The condition on $V_1'(S)$ is satisfied if and only if

$$(\rho + \eta + \gamma)[\kappa^d - (\kappa_1^o + \eta\kappa^\tau)] > \lambda_1(p - \kappa^d + \eta\kappa^\tau/2 + \eta A). \tag{4.57}$$

Remark 4.9. Condition (4.57) is clearly stronger than the negation of (4.54), as its RHS is strictly smaller and its LHS is strictly larger. This means that there is a nonempty set of parameter configurations such that neither (4.57) nor (4.54) hold.

• Therefore, let us suppose that $\kappa^d - \kappa_1^o - \lambda_1 V_1'(S) \in [0, \eta\kappa^\tau]$, and hence $u_1 = \bar{u} = [\kappa^d - \kappa_1^o - \lambda_1 V_1'(S)]/(\eta\kappa^\tau)$. The updated HJB is

$$(\rho + \eta)V_1(S) = (p - \kappa^d + \eta\kappa^\tau\bar{u}^2/2)S + V_1'(S)\gamma(\sigma - S) + \eta V_2(S). \tag{4.58}$$

If we look for a solution $V_1(S)$ that is affine in S, then $\bar{u}$ is a constant. Given the linear relationship between $\bar{u}$ and $V_1'(S)$, and since we are more interested in the optimal strategy than in the value function, let us compute $\bar{u}$ instead of $V_1'(S)$. We know that

$$V_1(S) = \frac{\kappa^d - \kappa_1^o - \eta\kappa^\tau\bar{u}}{\lambda_1}[S + \gamma\sigma/(\rho + \eta)] + \frac{\eta\gamma\sigma A}{(\rho + \eta)\rho}. \tag{4.59}$$

After substituting the expression for $V_1(S)$ in the HJB, we solve for $\bar{u}$ and obtain

$$\bar{u} = \frac{\rho + \eta + \gamma}{\lambda_1}\left[\sqrt{\frac{\lambda_1}{(\rho + \eta + \gamma)^2}\frac{2}{\eta\kappa^\tau}\left[(\rho + \eta + \gamma)(\kappa^d - \kappa_1^o) - \lambda_1(p - \kappa^d + \eta A)\right] + 1} - 1\right]. \tag{4.60}$$

The condition on $V_1'(S)$ is satisfied if and only if $\bar{u} \in [0, 1]$. The optimal control $\bar{u}$ is nonnegative if and only if

$$(\rho + \eta + \gamma)(\kappa^d - \kappa_1^o) \geq \lambda_1(p - \kappa^d + \eta A), \tag{4.61}$$

which is the negation of condition (4.54). Moreover, we have that $\bar{u} \leq 1$ if and only if

$$(\rho + \eta + \gamma)[\kappa^d - (\kappa_1^o + \eta\kappa^\tau)] \leq \lambda_1(p - \kappa^d + \eta\kappa^\tau/2 + \eta A), \tag{4.62}$$

which is the negation of (4.57).

Putting everything together, by the Verification Theorem, the Stage 1 solution can be summarized by the following proposition:

Proposition 4.10 (Optimal offshoring strategies at Stage 1)**.**

- *If* $(\rho + \eta + \gamma)(\kappa^d - \kappa_1^o) < \lambda_1(p - \kappa^d + \eta A)$*, then the optimal offshoring strategy is no offshoring* ($u_1 \equiv 0$) *with corresponding value function*

$$V_1(S) = \frac{p - \kappa^d + \eta A}{\rho + \eta + \gamma}[S + \gamma\sigma/(\rho + \eta)] + \frac{\eta\gamma\sigma A}{(\rho + \eta)\rho}. \tag{4.63}$$

- *If* $(\rho + \eta + \gamma)[\kappa^d - (\kappa_1^o + \eta\kappa^\tau)] > \lambda_1(p - \kappa^d + \eta\kappa^\tau/2 + \eta A)$*, then the optimal offshoring strategy is full offshoring* ($u_1 \equiv 1$) *with corresponding value function*

$$V_1(S) = \frac{p - \kappa_1^o - \eta\kappa^\tau/2 + \eta A}{\rho + \eta + \gamma + \lambda_1}[S + \gamma\sigma/(\rho + \eta)] + \frac{\eta\gamma\sigma A}{(\rho + \eta)\rho}. \tag{4.64}$$

- *Otherwise, the optimal offshoring strategy is partial offshoring at a constant intensity* $u_1 \equiv \bar{u}$*, where*

$$\bar{u} = \frac{\rho + \eta + \gamma}{\lambda_1}\left[\sqrt{\frac{\lambda_1}{(\rho + \eta + \gamma)^2}\frac{2}{\eta\kappa^\tau}\left[(\rho + \eta + \gamma)(\kappa^d - \kappa_1^o) - \lambda_1(p - \kappa^d + \eta A)\right] + 1} - 1\right]. \tag{4.65}$$

with corresponding value function

$$V_1(S) = \frac{\kappa^d - \kappa_1^o - \eta\kappa^\tau\bar{u}}{\lambda_1}[S + \gamma\sigma/(\rho + \eta)] + \frac{\eta\gamma\sigma A}{(\rho + \eta)\rho}. \tag{4.66}$$

Remark 4.11. Note that the Stage 1 value function V_1 depends on the Stage 2 data through the constant A (defined in (4.42)), which is present both in the additive term $\frac{\eta\gamma\sigma A}{(\rho+\eta)\rho}$ and in the multiplicative fraction $V_1'(S)$ (explicitly in the first two cases, and through $\bar{u}$ in the third case).

The first inequality in Proposition 4.10 permits to establish whether to implement any offshoring or none at all: specifically, if it holds, the firm should not offshore production. Such inequality can be rewritten in a way that separates the "convenience of offshoring" (on the LHS below) from everything else:

$$\frac{\kappa^d - \kappa_1^o}{\lambda_1} < \frac{p - \kappa^d}{\rho + \gamma} + \left[\frac{\eta}{\rho + \eta + \gamma}\left(A - \frac{p - \kappa^d}{\rho + \gamma}\right)\right]. \tag{4.67}$$

Recalling that $A = \max\left\{\frac{p-\kappa^d}{\rho+\gamma}, \frac{p-\kappa_2^o}{\rho+\gamma+\lambda_2}\right\}$, it follows that the second term (within square brackets) on the RHS is nonnegative. Therefore, the anticipating firm, who plans the offshoring while also considering the future effects of the switch and its hazard rate, should compare the "convenience of offshoring" with the domestic profit margin $p - \kappa^d$ divided by $(\rho + \gamma)$ (like in Stage 2) *plus* a positive quantity that takes into account the hazard rate η and the Stage 2 profits (represented by A).

Note that, whereas in Stage 2 the optimal offshoring is essentially either zero or full offshoring, in Stage 1 the anticipating firm may also implement partial offshoring.

Sales trajectory. We found that the optimal offshoring strategy in Stage 1 is constant. Let us denote it by $u_1(t) = \bar{u} \in [0, 1]$ (here considering also the cases $u_1 \equiv 0$ and $u_1 \equiv 1$). Then, the Cauchy problem for the Stage 1 sales trajectory is

$$\begin{cases} \dot{S}_1(t) = \gamma(\sigma - S_1(t)) - \lambda_1 \bar{u} S_1(t) \\ S_1(0) = S_0, \end{cases} \tag{4.68}$$

with the following solution:

$$S_1(t) = \frac{\gamma\sigma}{\gamma + \lambda_1\bar{u}} + \left(S_0 - \frac{\gamma\sigma}{\gamma + \lambda_1\bar{u}}\right)e^{-(\gamma+\lambda_1\bar{u})t}. \tag{4.69}$$

The Stage 1 sales trajectory starts from S_0 and, if we ignored the regime switch, would converge exponentially to $\gamma\sigma/(\gamma + \lambda_1\bar{u})$. However, the trajectory changes as soon as the switch occurs, settling on the Stage 2 trajectory and converging to the Stage 2 steady state.

4.2.4 Myopic solution

Here we compute the offshoring strategy of a myopic planner, i.e., a firm who is unaware of the impending regime switch and therefore does not take it into consideration.

The myopic decision-maker solves the following single-stage optimal control

problem with the Stage 1 data:

$$\underset{u_1(t)\in[0,1]}{\text{maximize}} \int_0^\infty e^{-(\rho+\eta)t}\left\{\left[p-\kappa^d(1-u_1(t))-\kappa_1^o u_1(t)\right]S_1(t)\right\}dt \tag{4.70}$$

subject to

$$\begin{cases}\dot{S}_1(t)=\gamma(\sigma-S_1(t))-\lambda_1 u_1(t)S_1(t)\\ S_1(0)=S_0.\end{cases} \tag{4.71}$$

Note that the myopic problem above is equivalent to the Stage 1 anticipative problem (4.30)–(4.31) where the hazard rate η is set to zero.

Observe also that the myopic problem has the same form of the Stage 2 problem (4.32)–(4.33), with κ_1^o instead of κ_2^o and λ_1 instead of λ_2. Having already computed the Stage 2 solution, we can exploit the results by simply substituting the parameters accordingly. Therefore, the optimal strategy for the myopic planner is

$$u_1(t)=\begin{cases}0 & \text{if } (\rho+\gamma)(\kappa^d-\kappa_1^o)<\lambda_1(p-\kappa^d)\\ 1 & \text{if } (\rho+\gamma)(\kappa^d-\kappa_1^o)>\lambda_1(p-\kappa^d)\\ \text{any feasible control} & \text{if } (\rho+\gamma)(\kappa^d-\kappa_1^o)=\lambda_1(p-\kappa^d).\end{cases} \tag{4.72}$$

For the purpose of answering our research questions, it is interesting to compare the myopic and the anticipating behaviors: under which conditions do the respective firms implement any offshoring? Can we prove that such conditions are stricter in the anticipative case?

To understand the role played by anticipation in determining whether it is optimal to implement zero or positive offshoring, we compare condition (4.67), under which the anticipating offshoring strategy is $u_1(t)=0$, with the condition in (4.72) for the myopic strategy to be $u_1(t)=0$, i.e.,

$$\frac{\kappa^d-\kappa_1^o}{\lambda_1}<\frac{p-\kappa^d}{\rho+\gamma}. \tag{4.73}$$

Recalling that $A=\frac{p-\kappa^d}{\rho+\gamma}$ if $u_2\equiv 0$ and $A>\frac{p-\kappa^d}{\rho+\gamma}$ if $u_2\equiv 1$, we can write the following proposition:

Proposition 4.12. *If the optimal offshoring in Stage 2 is $u_2\equiv 0$, then condition* (4.67)

is equivalent to (4.73). *If otherwise the optimal offshoring strategy in Stage 2 is* $u_2 \equiv 1$, *then condition* (4.67) *is stronger than* (4.73), *indicating that the anticipating planner offshores more cautiously, compared to the myopic one.*

4.2.5 Conclusions and further developments

This section summarizes the research content, the research contribution, and the research limits of this paper.

Our results provide valuable insights, especially when comparing anticipative and myopic strategies. Anticipating a potential disruptive event that could affect current offshoring conditions leads the planner to adopt a more cautious approach compared to a myopic planner who chooses to either offshore none or all of the production based on immediate convenience. Indeed, the condition required for an anticipating firm to engage in positive offshoring is stricter than that for a myopic firm, making it "less likely" for an anticipating firm to offshore any production. Moreover, even when the optimal anticipative strategy involves offshoring, such offshoring may only be partial. In conclusion, the anticipative firm might offshore only a portion of the production, if offshoring appears to be beneficial at all, to account for possible future disruptions.

However, this study still has limitations: the simplicity of the resulting optimal offshoring strategy, which in the myopic case and in Stage 2 essentially can only take the extreme values of 0 or 1, indicates that our toy model is too simplistic to capture the full complexity of this real-world problem.

For this reason, future research will focus on enhancing and developing this model to better reflect reality. This will likely render an analytical solution unattainable, necessitating the use of numerical simulations for qualitative results. Consequently, we will need to adopt a different solution method. Pontryagin's Maximum Principle, as opposed to Dynamic Programming, offers a promising alternative, as it allows us to gain analytical insights into the co-states and controls without requiring explicit computation.

Conclusions

This thesis has presented a comprehensive exploration of two-stage optimal control problems with a stochastic switching time and their applications to various economic models. The work is structured around the development of theoretical frameworks and the application of these frameworks to practical, real-world problems, demonstrating the utility and flexibility of the proposed methodologies. This final chapter synthesizes the findings and suggests directions for future research.

Summary of Findings

Theory of Two-Stage Optimal Control Problems. The thesis begins by establishing a robust theoretical foundation for two-stage optimal control problems with a stochastic switching time. Two primary solution techniques are explored: the backward approach and the vintage-structure approach. The backward approach solves the second stage first, leveraging information about the future to optimize current decisions. The vintage-structure approach simultaneously solves both stages, providing a more integrated view of the problem dynamics.

Health Economics Applications. In the domain of health economics, the thesis applies the developed theory to model optimal lockdown and vaccination strategies during the COVID-19 pandemic. The analysis reveals that an anticipative approach, which considers the potential discovery of a vaccine, leads to stricter initial lockdown measures compared to a myopic strategy. This anticipative strategy is shown to balance the trade-offs between economic costs and health outcomes more effectively, resulting in better overall outcomes.

In the same field, the thesis addresses vaccine hesitancy by modeling a differential game between the healthcare system and a pharmaceutical firm, both engaging in pro-vaccine communication campaigns. The study finds that

effective communication is crucial to counter anti-vax actions, influencing public perception and vaccination rates. The joined communication effort is effective at mitigating the healthcare costs associated with lower vaccination rates and the economic implications for pharmaceutical firms due to unsold vaccines.

Climate Change Economics. Another significant application is in climate change economics. The thesis examines optimal savings and emission abatement policies in the face of potential climate tipping points. The findings emphasize the importance of preventive measures, where even modest anticipatory abatement efforts can significantly mitigate the risks associated with tipping points. This work highlights the need for proactive environmental policies that account for the stochastic nature of climate-related disruptions.

Marketing and Production Strategies. The research further extends to marketing strategies under the threat of abrupt changes in production costs. By comparing the strategies of planners with different levels of foresight, the study demonstrates that those who anticipate potential disruptions and adjust their marketing intensity accordingly can achieve better outcomes. This section underscores the cost of myopia and the value of incorporating uncertainty into strategic planning.

In the area of production management, the thesis explores strategic decisions regarding offshoring and reshoring under uncertainty, incorporating factors like the *Made-in* effect, ethical concerns, and potential disruptions due to unforeseen circumstances. The study emphasizes the need for adaptive strategies that account for these uncertainties and highlights how rising ethical concerns and preferences for local products influence firms' decisions. The findings stress the importance of flexibility in global production strategies.

Directions for Future Research

While this thesis has made substantial contributions, it also opens several avenues for future research. Some of those are mentioned here below.

Extension to Multi-Stage Problems. Future work could extend the current two-stage models to multi-stage problems, incorporating multiple switching

times and stages. This would provide a more nuanced understanding of decision-making in even more complex and dynamic environments.

Broader Economic Applications. Expanding the applications to other areas of economics, such as financial markets, supply chain management, and labor economics, would further demonstrate the versatility and impact of the developed methodologies.

Conclusion

In conclusion, this thesis has advanced the understanding of two-stage optimal control problems with a stochastic switching time, providing both theoretical formalization and practical applications. The research highlights the critical importance of anticipatory strategies in managing uncertainty and optimizing outcomes across various domains. As the world continues to face complex and unpredictable challenges, the insights and methodologies developed in this thesis will serve as valuable tools for researchers, policy-makers, and practitioners aiming to navigate and mitigate these uncertainties effectively.

Bibliography

[1] E. Abrahamson and L. Rosenkopf. Institutional and competitive bandwagons: Using mathematical modeling as a tool to explore innovation diffusion. *Academy of management review*, 18(3):487–517, 1993.

[2] D. Acemoglu, V. Chernozhukov, I. Werning, M. D. Whinston, et al. *A multi-risk SIR model with optimally targeted lockdown*, volume 2020. National Bureau of Economic Research Cambridge, MA, 2020.

[3] M. Ali, S. T. H. Shah, M. Imran, and A. Khan. The role of asymptomatic class, quarantine and isolation in the transmission of covid-19. *Journal of biological dynamics*, 14(1):389–408, 2020.

[4] F. Alvarez, D. Argente, and F. Lippi. A simple planning problem for covid-19 lock-down, testing, and tracing. *American Economic Review: Insights*, 3(3):367–382, 2021.

[5] S. Anita. *Analysis and control of age-dependent population dynamics*, volume 11. Springer Science & Business Media, 2013.

[6] A. Aspri, E. Beretta, A. Gandolfi, and E. Wasmer. Mortality containment vs. economics opening: optimal policies in a seiard model. *Journal of mathematical economics*, 93:102490, 2021.

[7] M. Bardi, I. Capuzzo-Dolcetta, et al. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, volume 12. Springer, 1997.

[8] C. Bauch. Imitation dynamics predict vaccinating behaviour. *Proceedings of the Royal Society B: Biological Sciences*, 272(1573):1669–1675, 2005.

[9] S. Bhattacharyya, C. Bauch, and R. Breban. Role of word-of-mouth for programs of voluntary vaccination: A game-theoretic approach. *Mathematical Biosciences*, 269:130–134, 2015.

[10] T. Booth. Special report: outsourcing and offshoring: here, there and anywhere. *The Economist*, 2013.

[11] E.-K. Boukas, A. Haurie, and P. Michel. An optimal control problem with a random stopping time. *Journal of optimization theory and applications*, 64(3):471–480, 1990.

[12] E. Bozzola, G. Spina, R. Russo, M. Bozzola, G. Corsello, and A. Villani. Mandatory vaccinations in european countries, undocumented information, false news and the impact on vaccination uptake: the position of the italian pediatric society. *Italian Journal of Pediatrics*, 44(1):67, 2018.

[13] C. Brambilla, L. Grosset, and E. Sartori. A binary-state continuous-time markov chain model for offshoring and reshoring. *Axioms*, 13(5):300, 2024.

[14] W. L. Brogan. Optimal control theory applied to systems described by partial differential equations. In *Advances in Control Systems*, volume 6, pages 221–316. Elsevier, 1968.

[15] B. Buonomo, A. d'Onofrio, S. M. Kassa, and Y. H. Workineh. Modeling the effects of information-dependent vaccination behavior on meningitis transmission. *Mathematical Methods in the Applied Sciences*, 45(2):732–748, 2022.

[16] B. Buonomo, A. dOnofrio, and D. Lacitignola. The geometric approach to global stability in behavioral epidemiology. *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases*, pages 289–308, 2013.

[17] A. Buratto, R. Cesaretto, and M. Muttoni. Communication strategies to contrast anti-vax action: a differential game approach. *Central European Journal of Operations Research*, May 2024.

[18] A. Buratto, L. Grosset, and M. Muttoni. Two different solution techniques for an optimal control problem with a stochastic switching time. *WSEAS TRANSACTIONS ON MATHEMATICS*, 22:730–735, 2023.

[19] A. Buratto, L. Grosset, and B. Viscolani. *A LQ Vaccine Communication Game*, pages 353–367. Springer International Publishing, Cham, 2020.

[20] A. Buratto, M. Muttoni, S. Wrzaczek, and M. Freiberger. Should the COVID-19 lockdown be relaxed or intensified in case a vaccine becomes available? *PLOS ONE*, 17(9):e0273557, Sept. 2022.

[21] Y. Cai, K. L. Judd, and T. S. Lontzek. The social cost of stochastic and irreversible climate change. Working Paper 18704, National Bureau of Economic Research, January 2013.

[22] S. Canham and R. T. Hamilton. Sme internationalisation: off-shoring,backshoring, or staying at home in new zealand. *Strategic Outsourcing: An International Journal*, 6(3):277–291, 2013.

[23] P. Carrillo Santisteve and P. Lopalco. Measles still spreads in europe: who is responsible for the failure to vaccinate? *Clinical Microbiology and Infection*, 18:50–56, 2012.

[24] J. Caulkins, D. Grass, G. Feichtinger, R. Hartl, P. Kort, M. Kuhn, A. Prskawetz, M. Sanchez-Romero, A. Seidl, and S. Wrzaczek. The hammer and the jab: Are covid-19 lockdowns and vaccinations complements or substitutes? *European Journal of Operational Research*, 2023.

[25] J. Caulkins, D. Grass, G. Feichtinger, R. Hartl, P. M. Kort, A. Prskawetz, A. Seidl, and S. Wrzaczek. How long should the covid-19 lockdown continue? *Plos one*, 15(12):e0243413, 2020.

[26] J. P. Caulkins, D. Grass, G. Feichtinger, R. F. Hartl, P. M. Kort, A. Prskawetz, A. Seidl, and S. Wrzaczek. The optimal lockdown intensity for covid-19. *Journal of Mathematical Economics*, 93:102489, 2021.

[27] J. P. Caulkins, D. Grass, G. Feichtinger, R. F. Hartl, P. M. Kort, A. Prskawetz, A. Seidl, and S. Wrzaczek. Covid-19 and optimal lockdown strategies: The effect of new and more virulent strains. In *Pandemics: Insurance and Social Protection*. Springer International Publishing, 2022.

[28] H. R. Clarke and W. J. Reed. Consumption/pollution tradeoffs in an environment vulnerable to pollution-related catastrophic collapse. *Journal of Economic Dynamics and Control*, 18(5):991–1010, 1994.

[29] M. L. Cropper. Regulating activities with catastrophic environmental effects. *Journal of Environmental Economics and Management*, 3(1):1–15, 1976.

[30] H. Dawid, M. Y. Keoula, M. Kopel, and P. M. Kort. Product innovation incentives by an incumbent firm: A dynamic analysis. *Journal of Economic Behavior & Organization*, 117:411–438, Sept. 2015.

[31] R. Della-Marca and A. dOnofrio. Volatile opinions and optimal control of vaccine awareness campaigns: chaotic behaviour of the forward-backward sweep algorithm vs. heuristic direct optimization. *Communications in Nonlinear Science and Numerical Simulation*, 98:105768, 2021.

[32] C. Di Mauro, L. Fratocchi, G. Orzes, and M. Sartor. Offshoring and backshoring: A multiple case study analysis. *Journal of Purchasing and Supply Management*, 24(2):108–134, 2018.

[33] E. Dockner, S. Jørgensen, N. V. Long, and G. Sorger. *Differential Games in Economics and Management Science*. Cambridge University Press, Nov. 2000.

[34] A. d'Onofrio and P. Manfredi. The interplay between voluntary vaccination and reduction of risky behavior: A general behavior-implicit. *Current Trends in Dynamical Systems in Biology and Natural Sciences*, 21:185, 2020.

[35] F. El Ouardighi, G. Feichtinger, D. Grass, R. F. Hartl, and P. M. Kort. Advertising and quality-dependent word-of-mouth in a contagion sales model. *Journal of Optimization Theory and Applications*, 170(1):323–342, 2016.

[36] L. M. Ellram, W. L. Tate, and K. J. Petersen. Offshoring and reshoring: an update on the manufacturing location decision. *Journal of Supply Chain Management*, 49(2):14–22, 2013.

[37] G. Favato, G. Baio, A. Capone, A. Marcellusi, and F. Saverio Mennini. A novel method to value real options in health care: The case of a multicohort human papillomavirus vaccination strategy. *Clinical Therapeutics*, 35(7):904–914, 2013.

[38] S. Federico and G. Ferrari. Taming the spread of an epidemic by lockdown policies. *Journal of mathematical economics*, 93:102453, 2021.

[39] G. Feichtinger, R. F. Hartl, P. M. Kort, and V. M. Veliov. Capital accumulation under technological progress and learning: a vintage capital approach. *European Journal of Operational Research*, 172(1):293–310, 2006.

[40] G. Feichtinger, R. F. Hartl, and S. P. Sethi. Dynamic optimal control models in advertising: Recent developments. *Management Science*, 40(2):195–226, 1994.

[41] G. Feichtinger, G. Tragler, and V. M. Veliov. Optimality conditions for age-structured control systems. *Journal of Mathematical Analysis and Applications*, 288(1):47–68, 2003.

[42] S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, et al. Estimating the effects of non-pharmaceutical interventions on covid-19 in europe. *Nature*, 584(7820):257–261, 2020.

[43] L. Fratocchi, A. Ancarani, P. Barbieri, C. Di Mauro, G. Nassimbeni, M. Sartor, M. Vignoli, and A. Zanoni. Motivations of manufacturing reshoring: an interpretative framework. *International Journal of Physical Distribution & Logistics Management*, 46(2):98–127, 2016.

[44] M. Freiberger. Two-stage-optimal-control, Mar 2023. `https://doi.org/10.5281/zenodo.7772076`.

[45] M. Freiberger, D. Grass, M. Kuhn, A. Seidl, and S. Wrzaczek. Chasing up and locking down the virus: Optimal pandemic interventions within a network. *Journal of Public Economic Theory*, 24(5):1182–1217, 2022.

[46] M. Freiberger, M. Kuhn, A. Fürnkranz-Prskawetz, M. Sanchez-Romero, and S. Wrzaczek. Optimization in age-structured dynamic economic models. 2024.

[47] M. Freiberger, M. Kuhn, and S. Wrzaczek. Opening the dice black-box. In *Systems Analysis for Reducing Footprints and Enhancing Resilience*, November 2022.

[48] Y. Fu, H. Jin, H. Xiang, and N. Wang. Optimal lockdown policy for vaccination during covid-19 pandemic. *Finance research letters*, 45:102123, 2022.

[49] C. Garriga, R. Manuelli, and S. Sanghi. Optimal management of an epidemic: Lockdown, vaccine and value of life. *Journal of Economic Dynamics and Control*, 140:104351, 2022.

[50] J. Gjerde, S. Grepperud, and S. Kverndokk. Optimal climate policy under the possibility of a catastrophe. *Resource and Energy Economics*, 21(3):289–317, 1999.

[51] M. Gonzalez-Eiras and D. Niepelt. On the optimal 'lockdown' during an epidemic. 2020. Swiss National Bank, Study Center Gerzensee, CEPR Discussion Paper 14612.

[52] D. Grass, J. P. Caulkins, G. Feichtinger, G. Tragler, D. A. Behrens, et al. Optimal control of nonlinear processes. *Berlino: Springer*, 2008.

[53] D. Gromov and E. Gromova. Differential games with random duration: a hybrid systems formulation. *Contributions to game theory and management*, 7:104–119, 2014.

[54] L. Grosset and B. Viscolani. Advertising in a vaccination campaign: A variable time control problem. *Journal of Interdisciplinary Mathematics*, 23(6):1197–1211, 2020.

[55] L. Grosset and B. Viscolani. A dynamic advertising model in a vaccination campaign. *Central European Journal of Operations Research*, 29(2):737–751, June 2021.

[56] G. Guenther. Federal tax benefits for manufacturing: current law, legislative proposals, and issues for the 112th congress. 2012.

[57] S. Gupta, Y. Wang, and M. Czinkota. Reshoring and sustainable development goals. *British Journal of Management*, 32(3):E6–9, 2021.

[58] R. E. Hall and C. I. Jones. The value of life and the rise in health spending. *The Quarterly Journal of Economics*, 122(1):39–72, 2007.

[59] P. Hartman. *Ordinary Differential Equations: Second Edition*. SIAM, Jan. 1982.

[60] J. L. Haunschmied, R. M. Kovacevic, W. Semmler, and V. M. Veliov, editors. *Dynamic Economic Problems with Regime Switches*, volume 25 of *Dynamic Modeling and Econometrics in Economics and Finance*. Springer International Publishing, Cham, 2021.

[61] M. Henkel, A. Boffelli, J. Olhager, and M. Kalchschmidt. A case survey of offshoringbackshoring cases: The influence of contingency factors. *International Journal of Production Economics*, 253:108615, 2022.

[62] K. Hinderer, U. Rieder, and M. Stieglitz. *Dynamic optimization*. Springer, 2016.

[63] Hotez, P. J. How the anti-vaxxers are winning. `https://www.nytimes.com/2017/02/08/opinion/how-the-anti-vaxxers-are-winning.html`, 2017. Accessed: 2022-04-11.

[64] J. Huang, M. Leng, and L. Liang. Recent developments in dynamic advertising research. *European Journal of Operational Research*, 220(3):591–609, 2012.

[65] P. D. Ø. Jensen and T. Pedersen. The economic geography of offshoring: the fit between activities and local context. *Journal of Management Studies*, 48(2):352–372, 2011.

[66] K. Keller, B. M. Bolker, and D. F. Bradford. Uncertain climate thresholds and optimal economic growth. *Journal of Environmental Economics and Management*, 48(1):723–741, 2004.

[67] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[68] S. Kinkel and S. Maloca. Drivers and antecedents of manufacturing offshoring and backshoringa german perspective. *Journal of purchasing and Supply Management*, 15(3):154–165, 2009.

[69] T. J. Kniesner, W. K. Viscusi, C. Woock, and J. P. Ziliak. The value of a statistical life: Evidence from panel data. *Review of Economics and Statistics*, 94(1):74–87, 2012.

[70] M. Kuhn and S. Wrzaczek. Rationally Risking Addiction: AăTwo-Stage Approach. In J. L. Haunschmied, R. M. Kovacevic, W. Semmler, and V. M. Veliov, editors, *Dynamic Economic Problems with Regime Switches*, Dynamic Modeling and Econometrics in Economics and Finance, pages 85–110. Springer International Publishing, Cham, 2021.

[71] D. Lemoine and C. Traeger. Watch your step: Optimal policy in a tipping climate. *American Economic Journal: Economic Policy*, 6(1):137–166, 2014.

[72] T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber. Tipping elements in the earth's climate system. *Proceedings of the National Academy of Sciences*, 105(6):1786–1793, 2008. Publisher: Proceedings of the National Academy of Sciences.

[73] T. M. Lenton, J. Rockström, O. Gaffney, S. Rahmstorf, K. Richardson, W. Steffen, and H. J. Schellnhuber. Climate tipping points too risky to bet against. *Nature*, 575(7784):592–595, 2019. Bandiera_abtest: a Cg_type: Comment Number: 7784 Publisher: Nature Publishing Group Subject_term: Climate change, Climate sciences, Environmental sciences, Policy.

[74] F. Livesey. The need for a new understanding of manufacturing and industrial policy in leading economies. *Innovations: Technology, Governance, Globalization*, 7(3):193–202, 2012.

[75] T. S. Lontzek, Y. Cai, K. L. Judd, and T. M. Lenton. Stochastic integrated assessment of climate tipping points indicates the need for strict climate policy. *Nature Climate Change*, 5(5):441–444, 2015. Number: 5 Publisher: Nature Publishing Group.

[76] A. Mandel and V. Veetil. The economic cost of covid lockdowns: an out-of-equilibrium analysis. *Economics of Disasters and Climate Change*, 4:431–451, 2020.

[77] P. Manfredi and A. d'Onofrio. Modeling the interplay between human behavior and the spread of infectious diseases. New York, NY: Springer (ISBN 978-1-4614-5473-1/hbk; 978-1-4614-5474-8/ebook). xiii, 329 p. (2013)., 2013.

[78] C. Martínez-Mora and F. Merino. Offshoring in the spanish footwear industry: a return journey? *Journal of Purchasing and Supply Management*, 20(4):225–237, 2014.

[79] E. Mathieu, H. Ritchie, E. Ortiz-Ospina, M. Roser, J. Hasell, C. Appel, C. Giattino, and L. Rodés-Guirao. A global database of covid-19 vaccinations. *Nature human behaviour*, 5(7):947–953, 2021.

[80] R. Matusik and A. Nowakowski. Control of covid-19 transmission dynamics, a game theoretical approach. *Nonlinear Dynamics*, 110:857–877, 2022.

[81] R. McIvor. Understanding the manufacturing location decision: The case for the transaction cost and capability perspectives. *Journal of Supply Chain Management*, 49(2):23–26, 2013.

[82] K. M. Murphy and R. H. Topel. The value of health and longevity. *Journal of political Economy*, 114(5):871–904, 2006.

[83] M. Muttoni. How to prepare for and adapt to a climate tipping point, Nov. 2022. Place: Laxenburg, Austria Publisher: IIASA.

[84] G. Nassimbeni, M. Sartor, and G. Orzes. Countertrade: compensatory requests to sell abroad. *Journal for Global Business Advancement*, 7(1):69–87, 2014.

[85] M. Nerlove and K. J. Arrow. Optimal advertising policy under dynamic conditions. *Economica*, pages 129–142, 1962.

[86] W. Nordhaus. *A Question of Balance: Weighing the Options on Global Warming Policies*. Yale University Press, 2014. Google-Books-ID: XD_XBQAAQBAJ.

[87] S. B. Omer, D. A. Salmon, W. A. Orenstein, M. P. deHart, and N. Halsey. Vaccine refusal, mandatory immunization, and the risks of vaccine-preventable diseases. *New England Journal of Medicine*, 360(19):1981–1988, 2009. PMID: 19420367.

[88] G. P. Pisano and W. C. Shih. Does america really need manufacturing. *Harvard business review*, 90(3):94–102, 2012.

[89] G. P. Pisano, W. C. Shih, et al. Restoring american competitiveness. *Harvard business review*, 87(7/8):114–125, 2009.

[90] S. Polasky, A. de Zeeuw, and F. Wagener. Optimal management with potential regime shifts. *Journal of Environmental Economics and Management*, 62(2):229–240, Sept. 2011.

[91] D. Radi, F. Lamantia, and G. Italo Bischi. Offshoring, reshoring, unemployment, and wage dynamics in a two-country evolutionary model. *Macroeconomic Dynamics*, 25(3):705732, 2021.

[92] W. J. Reed. Protecting a forest against fire: optimal protection patterns and harvest policies. *Natural Resource Modeling*, 2(1):23–53, 1987.

[93] W. J. Reed. Optimal harvesting of a fishery subject to random catastrophic collapse. *Mathematical Medicine and Biology: A Journal of the IMA*, 5(3):215–235, 1988.

[94] D. Salmon, S. Teret, C. MacIntyre, D. Salisbury, M. Burgess, and N. Halsey. Compulsory vaccination and conscientious or philosophical exemptions: Past, present, and future. *The Lancet*, 367(9508):436–442, Feb. 2006.

[95] T. Sawik. Reshore or not reshore: A stochastic programming approach to supply chain optimization. *Omega*, 118:102863, 2023.

[96] T. C. Schelling. The life you save may be your own. *Problems in public expenditure*, pages 127–162, 1968.

[97] S. P. Sethi. Dynamic optimal control models in advertising: A survey. *SIAM Review*, 19(4):685–725, 1977.

[98] D. S. Shepard and R. J. Zeckhauser. Survival versus consumption. *Management Science*, 30(4):423–439, 1984.

[99] E. Shim, J. J. Grefenstette, S. M. Albert, B. E. Cakouros, L. Bohn, and D. S. Burke. *Impact of Vaccine Behavior on the Resurgence of Measles*, pages 255–266. Springer New York, New York, NY, 2013.

[100] G. Sorger. Maximum principle for control problems with uncertain horizon and variable discount rate. *Journal of Optimization Theory and Applications*, 70(3):607–618, 1991.

[101] S. Taboubi and G. Zaccour. Impact of retailers myopia on channels strategies. In *Optimal Control and Differential Games: Essays in Honor of Steffen Jørgensen*, pages 179–192. Springer, 2002.

[102] A. Takla, M. Wiese-Posselt, T. Harder, J. J. Meerpohl, M. Röbl-Mathieu, M. Terhardt, M. van der Sande, O. Wichmann, F. Zepp, and S. J. Klug. Background paper for the recommendation of hpv vaccination for boys in germany. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 61(9):1170–1186, 2018.

[103] W. L. Tate. Offshoring and reshoring: Us insights and research challenges. *Journal of Purchasing and Supply Management*, 20(1):66–68, 2014.

[104] L. Taylor, A. Swerdfeger, and G. Eslick. Vaccines are not associated with autism: An evidence-based meta-analysis of case-control and cohort studies. *Vaccine*, 32(29):3623–3629, 2014.

[105] Y. Tsur and A. Zemel. Accounting for global warming risks: Resource management under event uncertainty. *Journal of Economic Dynamics and Control*, 20(6):1289–1305, 1996.

[106] Y. Tsur and A. Zemel. Pollution control in an uncertain environment. *Journal of Economic Dynamics and Control*, 22(6):967–975, 1998.

[107] Y. Tsur and A. Zemel. Policy tradeoffs under risk of abrupt climate change. *Journal of Economic Behavior & Organization*, 132:46–55, 2016.

[108] Y. Tsur and A. Zemel. Coping with Multiple Catastrophic Threats. *Environmental and Resource Economics*, 68(1):175–196, Sept. 2017.

[109] P. Van den Driessche and J. Watmough. Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1-2):29–48, 2002.

[110] F. van der Ploeg and A. de Zeeuw. Climate tipping and economic growth: Precautionary capital and the price of carbon. *Journal of the European Economic Association*, 16(5):1577–1617, 2018.

[111] F. van der Ploeg and A. de Zeeuw. Pricing carbon and adjusting capital to fend off climate catastrophes. *Environmental and Resource Economics*, 72(1):29–50, 2019.

[112] V. M. Veliov. Newton's method for problems of optimal control of heterogeneous systems. *Optimization Methods and Software*, 18(6):689–703, 2003.

[113] V. M. Veliov. Optimal control of heterogeneous systems: Basic theory. *Journal of Mathematical Analysis and Applications*, 346(1):227–242, Oct. 2008.

[114] M. L. Vidale and H. B. Wolfe. An Operations-Research Study of Sales Response to Advertising. *Operations Research*, 5(3):370–381, 1957.

[115] P. Wang. Control of distributed parameter systems. *Advances in control systems*, 1:75–172, 1964.

[116] Z. Wang, C. T. Bauch, S. Bhattacharyya, A. d'Onofrio, P. Manfredi, M. Perc, N. Perra, M. Salathé, and D. Zhao. Statistical physics of vaccination. *Physics Reports*, 664:1–113, Dec. 2016.

[117] J. L. White, M. K. Grabowski, A. F. Rositch, P. E. Gravitt, T. C. Quinn, A. A. R. Tobian, and E. U. Patel. Trends in adolescent human papillomavirus vaccination and parental hesitancy in the United States. *The Journal of Infectious Diseases*, 03 2023. jiad055.

[118] S. Wrzaczek, M. Kuhn, and I. Frankovic. Using age structure for a multistage optimal control model with random switching time. *Journal of Optimization Theory and Applications*, 184(3):1065–1082, 2020.

[119] M. E. Yaari. Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies*, 32(2):137–150, 1965.

[120] H. Yang, J. Ou, and X. Chen. Impact of tariffs and production cost on a multinational firm's incentive for backshoring under competition. *Omega*, 105:102500, 2021.

Acknowledgments

Completing this PhD thesis has been a significant journey, and it would not have been possible without the support, guidance, and encouragement of many individuals. I would like to express my sincere appreciation:

To Simone and Brianna, for being a constant reminder of what I deem truly important: my family. You are my world and my reason for being.

To Flavia, Popy, Chicco and Ninò, for your unconditional love and always appreciated support, for backing my life choices, and for giving me the chance to pursue my interest in mathematics at my own pace, free from judgement or pressure.

To Alessandra, Natale, Federica, Mattia, Nicoletta and Tiziano, for the good times, for your welcoming heart, and for your precious assistance in juggling a PhD alongside raising a young child.

To Alessandra, my academic godmother, for going out of your way to make sure I wouldn't miss an opportunity, for teaching me that sometimes it's better to choose diplomacy over winning an argument, and for helping and supporting me in whichever path I chose, even when you'd have done things differently. You've helped me in more ways than it was required of you, and in more ways than I could have asked for.

To Stefan, for your always stimulating research ideas, for believing in my abilities, and for giving me the chance to work alongside great scientists in a nurturing and inspiring environment. My time in Laxenburg is my most cherished PhD experience.

To Michael, my friend and virtuous example, for your endless patience and clever observations, for the shared coffees and music sessions. It's hard to imagine this PhD journey without you in it.

To Stefano and Prof. Meda, for nudging me towards the PhD path. Were it not for you I'd probably be stuck at some corporate job by now.