Finance and Stochastics (2023) 27:947–984
https://doi.org/10.1007/s00780-023-00519-9

Present-biased lobbyists in linear–quadratic stochastic differential games

Ali Lazrak[1] · **Hanxiao Wang**[2] · **Jiongmin Yong**[3]

This paper is dedicated to the memory of Tomas Björk

Received: 18 April 2023 / Accepted: 18 August 2023

Abstract
We investigate a linear–quadratic stochastic zero-sum game where two players lobby a political representative to invest in a wind farm. Players are time-inconsistent because they discount the utility with a non-constant rate. Our objective is to identify a consistent planning equilibrium in which the players are aware of their inconsistency and cannot commit to a lobbying policy. We analyse equilibrium behaviour in both single-player and two-player cases and compare the behaviours of the game under constant and variable discount rates. The equilibrium behaviour is provided in closed-loop form, either analytically or via numerical approximation. Our numerical analysis of the equilibrium reveals that strategic behaviour leads to more intense lobbying without resulting in overshooting.

Keywords Time inconsistency · Lobbying · Two-player zero-sum dynamic game · Linear–quadratic stochastic differential game

Mathematics Subject Classification (2020) 91A15 · 91B14

JEL Classification C73 · D72

Ali Lazrak is supported in part by SSHRC; Hanxiao Wang is supported in part by NSFC Grant 12201424 and Guangdong Basic and Applied Basic Research Foundation 2023A1515012104; Jiongmin Yong is supported in part by NSF Grant DMS-2305475.

✉ H. Wang
hxwang@szu.edu.cn

A. Lazrak
ali.lazrak@sauder.ubc.ca

J. Yong
jiongmin.yong@ucf.edu

1 Sauder School of Business, University of British Columbia, Vancouver, BC V6T 1Z2, Canada

2 School of Mathematical Sciences, Shenzhen University, Shenzhen, 518060, China

3 Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA

1 Introduction

Time inconsistency refers to a phenomenon in which a decision maker's preferences for different alternatives change over time, even in the absence of new information. This can pose a significant challenge in solving dynamic optimal choice problems as the optimal choice may vary depending on the moment in time from which the decision is being made. This misalignment can create a gap between the optimal policies intended at one point in time and the policies that are implemented later. Standard neoclassical models in macroeconomics and finance assume that decision makers have time-additive preferences that obey the independence axiom and discount utility exponentially. In that context, dynamic programming methods can help break down any dynamic optimisation problem into a sequence of simpler, static problems that can be solved recursively. However, economists are acutely aware that time inconsistency pervades various contexts, rendering the principle of dynamic programming invalid.

Time inconsistency can manifest itself in common economic interactions even if the players' preferences are standard. For instance, in dynamic games, the time inconsistency problem is a common issue, and the players' ability to commit can significantly affect the equilibrium outcome. The game between central banks and the private sector in Kydland and Prescott [19] is an example of this problem. Similarly, in dynamic collective decisions, Jackson and Yariv [17] have shown that time inconsistency is prevalent for almost any collective decision rule, whether it is based on a vote or a utilitarian aggregation rule. Since it emerges from the interaction between players, the time inconsistency problem can provide new insights and allow economists to recommend welfare-improving policies for society (see Kydland and Prescott [19]). Moreover, Calvo and Obstfeld [7] and Bernheim [2] have also highlighted the problem of planners' time inconsistency in intergenerational models. More broadly, time inconsistency and its policy implications have become central concerns of economists over time.

Time inconsistency can also emerge when individual decision makers deviate from the standard assumption of exponential discounting. When a non-exponential discount function is used, the marginal rate of substitution between consumption at two different future dates varies as time passes, unlike in the case of a constant discount rate. A prominent psychological theory proposed by Ainslie [1] argues in favour of varying discount rates. The theory suggests that "hyperbolic discounting"[1] generates bias towards the present and can explain the impulsive behaviour explored in Ainslie [1].

This paper investigates the impact of a non-constant discount rate on continuous-time linear–quadratic game problems. We assume that the decision maker is aware of the time inconsistency and approaches the problem with an intrapersonal game view while lacking any commitment ability. This is the consistent planning solution to the game envisioned by Strotz [21]. To contrast the behaviour of a decision maker with a constant discount rate against one with a non-constant discount rate, we consider

[1] Hyperbolic discounting refers to a phenomenon where people tend to discount the value of future rewards more heavily when they are further away in time, but less so as the reward becomes more immediate.

two variations of the same problem. In the first, we examine the behaviour of a single decision maker. In the second, we analyse a *zero-sum game* between two players. To gain insights into the impact of time inconsistency on decision making in these two situations, we aim to provide closed-form solutions or numerical approximations. This enables us to offer concrete statements about decision-making behaviour and contrast it with the behaviour under constant discounting. It is crucial to underline that our examples and solution methods are exclusively applicable to zero-sum games.

The framework of this paper deliberately maintains simplicity, serving as an entry point to relevant literature and offering insights into how changing discount rates affects decision-making behaviour. Our goal is to present closed-form solutions wherever possible, and when that is impossible, to provide numerical approximation methods. By prioritising simplicity and accessibility, we aim to engage a wider audience in this crucial area of research. Notably, Sect. 5 presents new findings on the zero-sum game with players exhibiting non-constant discount rates which have not yet been explored, to the best of our knowledge.

The topic of time inconsistency has been extensively studied in the literature, and it is not feasible to provide a comprehensive summary in this paper. Therefore, we do not cite every significant paper in this field, and refer the reader to the book by Björk et al. [4] and the recent paper by Hernández and Possamaï [14], in which comprehensive overviews of the literature on time inconsistency in mathematical finance are offered. The seminal paper [21] by Strotz was the first to formalise non-exponential discounting within a dynamic utility maximisation framework. In a simple deterministic "cake eating" problem in continuous time, Strotz formulated the solution of the problem under commitment, representing the optimal choice based on preferences at the beginning of time. He then went on to solve the consistent planning problem that he defined as follows:

"Since precommitment is not always a feasible solution to the problem of intertemporal conflict, the man with insight into his future unreliability may adopt a different strategy and reject any plan which he will not follow through. His problem is then to find the best plan among those that he will actually follow."

The consistent planning solution, also known as the sophisticated solution, provides a useful framework for solving consumption and saving decisions over time when discount rates are non-constant. In a finite-horizon setting with discrete time, the consistent planning solution can be identified by using backward induction. Starting from the last period, the decision maker maximises their utility by selecting a sustainable consumption plan that future selves would also choose. This is done at each step, working backwards to the present time. In each step of the consistent planning solution, lifetime utility is maximised under a consumption sustainability constraint in addition to the standard budget constraint, making the problem non-standard. The sustainability constraints ensure that the needs of future selves are taken into account. Although this method may generate multiple solutions due to the potential non-concavity of intermediate values, the existence of a solution is typically guaranteed.

When the time horizon is infinite, there is no terminal time to start the backward induction and the consistent planning solution can present some mathematical difficulties in discrete time. Nonetheless, researchers have made significant progress in

addressing these difficulties, as seen in Krusell and Smith [18], among others. In the original continuous framework proposed by Strotz, the mathematical formulation of consistent planning was initially established in the context of a deterministic consumption–saving problem by Ekeland and Lazrak [10] (see also Ekeland and Lazrak [11]). The approach involves assuming that the decision maker has control over the consumption of their immediate successors at any given point in time, which enables the formation of a small coalition that isolates the current decision maker from the more distant ones. For a closed-loop consumption strategy that depends on the current value of capital stock to qualify as an equilibrium strategy, it must be the optimal policy for the current planner when the coalition is infinitesimally small. Furthermore, the same strategy must be expected to be employed by the distant planners.

This way of defining the equilibrium is commonly referred to as the "spike variation method". Interestingly, this method is more general than its original framework proposed by Ekeland and Lazrak [10]. The equilibrium is well posed even when a Brownian noise is introduced in different applications. Early research on this topic includes Ekeland and Pirvu [12], Björk and Murgoci [5], Yong [24, 25], Hu et al. [16] and Björk et al. [3].

This paper presents an application of the spike variation method to a two-player, zero-sum linear–quadratic stochastic differential game. Our approach represents a step forward in the understanding of strategic interactions between players when each player has a preference-based time inconsistency. It is worth noting that recent progress has been made in principal–agent models concerning the resolution of problems involving non-constant discounting, as demonstrated in Cetemen et al. [8] and Hernández and Possamaï [15]. Additionally, in the context of optimal control problems with multidimensional states where the control variable appears in the diffusion, Lei and Pun [20] have established the existence of a solution to the equilibrium HJB equation for a small time horizon when dealing with a time-inconsistent single-player situation. Given the focus of our study, we concentrate on specific examples to illustrate the impact of time inconsistency in zero-sum differential games. This approach allows us to define an equilibrium that is global in the time dimension, enabling us to derive valuable economic insights from our findings.

2 The model

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space on which a one-dimensional standard Brownian motion W is defined, whose natural filtration, augmented by all the $\mathbb{P}$-nullsets, is denoted by $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. Let $T > 0$ be a fixed time horizon and for any $0 \leq a \leq b \leq T$, define the set

$$\Delta[a, b] := \{(t, s) \in [a, b]^2 : a \leq t \leq s \leq b\}.$$

2.1 Objectives and policies

We consider a dynamic game where the state is controlled by two players. The state is described by a one-dimensional controlled linear stochastic differential equation

given by

$$dX(s) = \big(u_1(s) + u_2(s)\big)ds + \sigma X(s)dW(s), \quad s \in [t, T], \qquad X(t) = \xi, \quad (2.1)$$

where X is the state process and u_i is the control taken by player i, $i = 1, 2$. We let

$$\mathcal{D} := \{(t, \xi) : t \in [0, T), \xi \in L^2_{\mathcal{F}_t}(\Omega)\},$$

with

$$L^2_{\mathcal{F}_t}(\Omega) := \{\xi : \Omega \to \mathbb{R} : \xi \text{ is } \mathcal{F}_t\text{-measurable}, \mathbb{E}[|\xi|^2] < \infty\}.$$

Any $(t, \xi) \in \mathcal{D}$ is called an *initial pair*. The coefficient σ is a deterministic scalar. For $i = 1, 2$ and $t \in [0, T)$, the set of *admissible (open-loop) controls* of player i on $[t, T]$ is defined by

$$L^2_{\mathbb{F}}(t, T) := \Bigg\{\varphi : [t, T] \times \Omega \to \mathbb{R} : \varphi \text{ is } \mathbb{F}\text{-progressively measurable and } \mathbb{E}\bigg[\int_t^T |\varphi(s)|^2 ds\bigg] < \infty\Bigg\}.$$

To measure the utility for player 1, we introduce the functional

$$J_1(t, \xi; u_1, u_2) = \mathbb{E}_t\bigg[\alpha(T - t)X(T)^2 + \int_t^T \alpha(s - t)\big(-u_1(s)^2 + Ru_2(s)^2\big)ds\bigg], \quad (2.2)$$

where $0 < R \leq 1$ is a deterministic scalar and α is the discount function. Similarly, the utility for player 2 is given by

$$J_2(t, \xi; u_1, u_2) = \mathbb{E}_t\bigg[-\alpha(T - t)X(T)^2 + \int_t^T \alpha(s - t)\big(u_1(s)^2 - Ru_2(s)^2\big)ds\bigg]. \quad (2.3)$$

As we explain below, the condition $0 < R \leq 1$ means that u_1 costs no less than u_2 in the cost functionals. When $R > 1$, the existence of an equilibrium is not guaranteed, and since our objective is to maintain tractability, the assumption $0 < R \leq 1$ is maintained **from now on**.

Starting from the state ξ, player i for $i = 1, 2$ selects her control u_i from the set $L^2_{\mathbb{F}}(t, T)$ to maximise the utility $J_i(t, \xi; u_1, u_2)$. Since

$$J_1(t, \xi; u_1, u_2) + J_2(t, \xi; u_1, u_2) = 0, \quad (2.4)$$

our game is called a two-player zero-sum stochastic linear–quadratic game. The zero-sum feature captures the conflict of interest of the two players due to the presence of externalities. To discuss the implication of the model, it is useful to keep in mind a specific application. Suppose that the quantity $X(T)$ represents the output of a wind farm. The experienced utility of the two players is quadratic in the output. The first player represents a part of the population who value the environmental externality due to green production of electricity and experience a positive utility that is quadratic in

the output. The second player represents a part of the population who experience a disutility that is also quadratic in the output. The aversion to wind turbines of the second player captures opposition to wind turbines due to noise inconvenience for citizens who live near the wind farms. We assume that the level of production of the wind farm is a collective decision that is taken through a political process which is mediated by political representatives. Each player can spend some effort lobbying the politicians in order to have some influence on the output decision of the wind farm. The lobbying can take the form of campaign donations or can be more direct by sending letters to political representatives or engaging in protest campaigns. The lobbying effort is captured by the control. The state equation (2.1) shows that a positive control u_i increases the output and a negative control decreases the level of output. Player 1 (resp. 2) values (dislikes) the output and would apply $u_1 = +\infty$ (resp. $u_2 = -\infty$) in the absence of constraints. However, lobbying generates a disutility that is quadratic for both players. The utility criterion (2.2) shows that player 1 suffers from a disutility from the lobbying effort captured by the term $-u_1(s)^2$. The utility (2.3) shows that player 2 suffers from a disutility from the lobbying effort captured by the term $-Ru_2(s)^2$ with $0 < R \leq 1$. Therefore, for the same level of lobbying, the disutility is weakly larger for player 1 who supports wind farms. This asymmetry captures the idea that opposing an environmental reform requires less lobbying effort because the wind farms represent new technologies that require more subsidy relative to the status quo technology for producing electricity. Finally, in addition to the opposite perception that the players have on wind farms, there is an externality in effort as well. The term $Ru_2(s)^2$ in the functional (2.2) shows that more lobbying efforts from player 2 impact positively the utility of player 1. Similarly, the term $u_1(s)^2$ in the utility (2.2) shows that the utility of player 2 is impacted positively when player 1 exerts more lobbying efforts. This externality in lobbying efforts is tantamount to an assumption of limited supply of political capital. When player 2 increases the lobbying efforts against wind farms, she may exhaust her political capital and player 1 benefits from the resulting depletion. This may reduce player 2's willingness to lobby against other unmodelled political issues, potentially leading to improved outcomes for player 1 on those matters.

2.2 Open-loop versus closed-loop saddle point controls

Due to (2.4), the utilities J_1 and J_2 are not independent. If we set

$$J(t, \xi; u_1, u_2) = J_1(t, \xi; u_1, u_2),$$

then player 1 wants to maximise $J(t, \xi; u_1, u_2)$ by selecting $u_1 \in L^2_{\mathbb{F}}(t, T)$, and player 2 wants to minimise it by selecting $u_2 \in L^2_{\mathbb{F}}(t, T)$. Thus we obtain a two-player zero-sum differential game described by the state equation (2.1) and the functional $J = J_1$ given by (2.2). For convenience, let us name it problem (G). Consequently, in problem (G), player 1 is the maximiser and player 2 the minimiser. We now introduce the following definition of the open-loop saddle point.

Definition 2.1 The control pair $(u_1^*, u_2^*) \in L^2_{\mathbb{F}}(t,T) \times L^2_{\mathbb{F}}(t,T)$ is called an *open-loop saddle point* of problem (G) at the initial pair (t,ξ) if we have

$$J(t,\xi;u_1,u_2^*) \le J(t,\xi;u_1^*,u_2^*) \le J(t,\xi;u_1^*,u_2), \qquad \forall u_1,u_2 \in L^2_{\mathbb{F}}(t,T).$$

An open-loop saddle point, if it exists, is an optimal choice for both players because if player i chooses a control different from u_i^*, her utility index becomes no better or worse. Sun and Yong [23, Theorem 2.2.1] showed that (u_1^*, u_2^*) is an open-loop saddle point of problem (G) on $[t,T]$ with X^* being the corresponding state process if and only if together with another pair (Y^*, Z^*) of adapted process, the so-called optimality system is satisfied (with a stationarity condition). This system is a system of forward-backward stochastic differential equations (FBSDEs, for short), which is actually a two-point boundary value problem of SDEs. Thus the whole information generated by $(X^*(s), u_1^*(s), u_2^*(s))$, $s \in [t,T]$, is needed to determine (Y^*, Z^*). Note that the open-loop saddle point (u_1^*, u_2^*) can be written in terms of (Y^*, Z^*) via the stationarity condition. Hence in seeking open-loop saddle points, the information of the state as well as the opponent's control over the whole time interval $[t,T]$ have to be used (including the initial state ξ). In a game situation, the future information of the state and both controls are not available at the present time to players. Therefore, an open-loop saddle point only has a functional analysis meaning and is not practically feasible. In this paper, we are primarily interested in so-called closed-loop saddle strategies which take the form of linear functions of the state and are non-anticipating, meaning that future information of the state and the two controls is not needed.

We now begin with the definition of closed-loop strategies at a given time t. Let $L^\infty(t,T)$ be the set of all bounded deterministic functions. A *closed-loop strategy* for both players consists of two functions $(\Theta_1, \Theta_2) \in L^\infty(t,T) \times L^\infty(t,T)$. Under such a closed-loop strategy, the state equation reads

$$\begin{aligned} dX(s) &= \big(\Theta_1(s) + \Theta_2(s)\big)X(s)ds + \sigma X(s)dW(s), \qquad s \in [t,T], \\ X(t) &= \xi. \end{aligned} \tag{2.5}$$

We call (2.5) the closed-loop system under (Θ_1, Θ_2) and the corresponding utility functional reads

$$\begin{aligned} &J(t,\xi;\Theta_1 X + v_1, \Theta_2 X + v_2) \\ &= \mathbb{E}_t\bigg[\alpha(T-t)X(T)^2 + \int_t^T \alpha(s-t)\Big(-\big(\Theta_1(s)X(s)\big)^2 + R\big(\Theta_2(s)X(s)\big)^2\Big)ds\bigg]. \end{aligned}$$

To emphasise that the solution X to (2.5) depends on (Θ_1, Θ_2) as well as on the initial pair (t,ξ), we frequently write

$$X(s) = X(s;t,\xi,\Theta_1,\Theta_2), \qquad s \in [t,T].$$

The control pair (u_1, u_2) defined by

$$u_1 = \Theta_1 X, \qquad u_2 = \Theta_2 X \tag{2.6}$$

is called the *outcome* of the closed-loop strategy (Θ_1, Θ_2).

With the above, we are now ready to define the closed-loop saddle strategies of problem (G).

Definition 2.2 For any initial time $t \in [0, T)$, a *closed-loop saddle strategy* for problem (G) consists of a pair $(\Theta_1^*, \Theta_2^*) \in L^\infty(t, T) \times L^\infty(t, T)$ such that for any $\xi \in L^2_{\mathcal{F}_t}(\Omega)$, we have

$$J(t, \xi; u_1, \Theta_2^* X) \le J(t, \xi; \Theta_1^* X^*, \Theta_2^* X^*) \le J(t, \xi; \Theta_1^* X, u_2),$$
$$\forall u_i \in L^2_{\mathbb{F}}(t, T), i = 1, 2. \tag{2.7}$$

One should note that on the most left-hand side of (2.7), we look at the process $X = X(\cdot\,; t, \xi, u_1, \Theta_2^*)$, whereas on the most right-hand side of (2.7), we look at $X = X(\cdot\,; t, \xi, \Theta_1^*, u_2)$, and in the middle of (2.7), we have $X^* = X(\cdot\,; t, \xi, \Theta_1^*, \Theta_2^*)$. Thus the processes X in the most left-hand and the most right-hand side in (2.7) are different in general.

From the above, we see that a closed-loop saddle strategy is determined prior to the state process. In other words, the state process is determined by the state equation under the outcome (2.6) of the closed-loop saddle strategy. Thus the closed-loop saddle strategy works for an arbitrary initial state ξ. It follows from (2.6) that the outcome, as the input of the state equation, is non-anticipating—no future information is used. Hence it is practically feasible. More precisely, one could calculate Θ_i^* off-line first and then apply it via the outcome (2.6) (the current state feedback) in the state equation. It is crucial to underscore that this approach stands in stark contrast to the open-loop saddle points (see Sun and Yong [22]).

In this paper, we concentrate on closed-loop saddle strategies and their outcomes for both the single-player problem and the game problem.

2.3 Discounting

In this paper, we consider first the standard exponential discount function $\alpha(t) = e^{-\rho t}$, where $\rho > 0$ is the constant discount rate defined by $\rho := -\alpha'(t)/\alpha(t)$.

To capture the present bias, we consider a mixture of exponential discount functions of the form

$$\alpha(t) = \lambda e^{-\rho t} + (1 - \lambda) e^{-\gamma t} \qquad \text{for some } \gamma > \rho > 0 \text{ and } \lambda \in (0, 1). \tag{2.8}$$

The implied discount rate is given by

$$-\frac{\alpha'(t)}{\alpha(t)} = \rho + \frac{(1 - \lambda)}{\lambda e^{(\gamma - \rho)t} + 1 - \lambda}(\gamma - \rho),$$

which is monotonically declining from the short term discount rate

$$-\frac{\alpha'(0)}{\alpha(0)} = \rho + (1 - \lambda)(\gamma - \rho)$$

to the long term discount rate $-\frac{\alpha'(\infty)}{\alpha(\infty)} = \rho$. Thus the discount function (2.8) embodies the present bias highlighted in Ainslie [1]. The particular form (2.8) of the discount function appears naturally for social planners in overlapping generations models (see e.g. Blanchard [6] and Calvo and Obstfeld [7]). The discount function (2.8) was also used as a deterministic benchmark in Harris and Laibson [13].[2]

3 Single-player games

In this section, we consider the game problem with a single player. In the first part, we provide the optimal control solution when the player has a constant discount rate. In the second part, we provide the consistent planning solution when the player has a non-constant discount rate. We focus on the behaviour of player 2 because we can identify the closed-loop consistent planning strategies for that players both with constant and non-constant discount rates. By contrast, it can be shown that player 1's optimal strategy may blow up, and hence there is no natural benchmark to study player 1's behaviour in a game setting.

3.1 Single-player games with constant discounting

We consider the state equation (2.1) and the objective given in (2.3) with $u_1 \equiv 0$ and $\alpha(t) = e^{-\rho t}$ for some $\rho > 0$. In other words, player 2 maximises the functional

$$J_2(t, \xi; 0, u_2) = \mathbb{E}_t\left[- e^{-\rho(T-t)} X(T)^2 - \int_t^T e^{-\rho(s-t)} R u_2(s)^2 ds \right] \tag{3.1}$$

by choosing $u_2 \in L^2_{\mathbb{F}}(t, T)$ subject to

$$dX(s) = u_2(s)ds + \sigma X(s) dW(s), \quad s \in [t, T], \qquad X(t) = \xi. \tag{3.2}$$

This is a standard stochastic linear–quadratic optimal control problem which can be solved by the results presented in Yong and Zhou [27, Theorem 6.6.1]. The following proposition describes the single-player optimal control when the discount rate is constant.

Proposition 3.1 *The unique optimal closed-loop optimal strategy of player* 2 *with state equation* (3.2) *and objective* (3.1) *is given by*

$$\hat{\Theta}_2(s) = -\frac{(\rho - \sigma^2) e^{(\rho-\sigma^2)s}}{(1 + R(\rho - \sigma^2)) e^{(\rho-\sigma^2)T} - e^{(\rho-\sigma^2)s}}, \qquad s \in [0, T]. \tag{3.3}$$

[2] In [13, Sect. II.B], utilities at all future periods are discounted exponentially with a discount factor $0 < \rho < 1$. However, distant future periods are additionally discounted with uniform weight $0 < \lambda \leq 1$. As a result, the immediate future receives full weight $e^{-\rho t}$, while more distant future periods are given the lower weight $\lambda e^{-\rho t}$. The immediate future lasts during a length of time that is stochastic and exponentially distributed with an intensity $\gamma - \rho$. Under these assumptions, the expected discount function matches exactly the discount function (2.8).

In other words, for any given initial time $t \in [0, T)$ *and initial state* $\xi \in L^2_{\mathcal{F}_t}(\Omega; \mathbb{R})$, *the unique optimal control is given by*

$$\begin{aligned}\hat{u}_2(s) &= \hat{\Theta}_2(s)\hat{X}(s) \\ &= -\frac{(\rho-\sigma^2)e^{(\rho-\sigma^2)s}}{(1+R(\rho-\sigma^2))e^{(\rho-\sigma^2)T}-e^{(\rho-\sigma^2)s}}\hat{X}(s), \qquad s \in [0,T], \end{aligned} \tag{3.4}$$

with $\hat{X}$ *being the unique solution of the closed-loop system*

$$d\hat{X}(s) = \hat{\Theta}_2(s)\hat{X}(s)ds + \sigma\hat{X}(s)dW(s), \quad s \in [t,T], \qquad \hat{X}(t) = \xi. \tag{3.5}$$

The closed-loop optimal strategy (3.4) as a function of $s \in [0, T]$ for various levels of the model parameters (T, σ, ρ, R) is illustrated in Fig. 1. The figure indicates that the optimal feedback is negative for all parameter values, implying that lobbying against wind turbines is optimal. Figure 1 also implies that the effort against turbines increases with larger output since $|\hat{u}_2|$ increases with $\hat{X}$. The figure also shows that the lobbying effort against turbines increases when the cost of effort R diminishes, the discount rate decreases, and the horizon T decreases. When the discount rate decreases, the decision maker becomes more patient and is willing to exert more effort today to reduce the output in the future. With a shorter horizon, the duration over which the decision maker can lobby is shorter, and as a result, efforts intensify. Moreover, Fig. 1 demonstrates that the lobbying effort intensifies when σ increases. With more risk, it becomes more important to control the state.

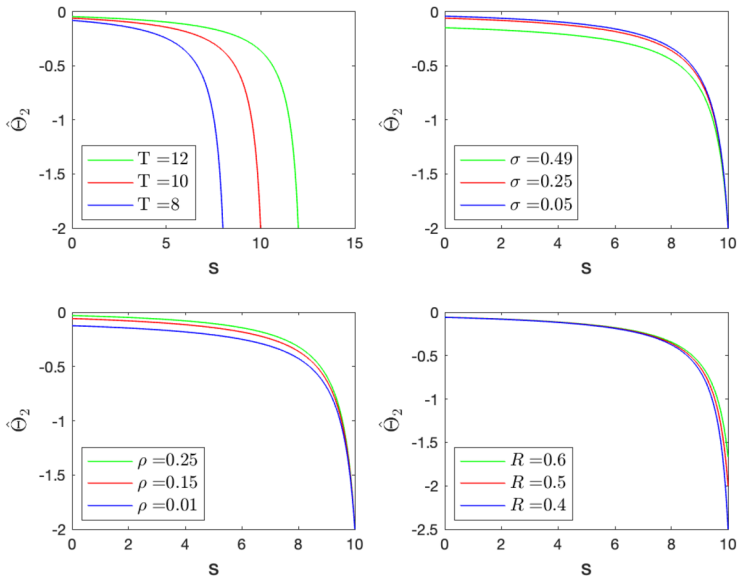

Fig. 1 The panels display the optimal closed-loop strategies $\hat{\Theta}_2(s)$ for $0 \le s \le T$ based on (3.3). The baseline parameter values used are $T = 10$, $\sigma = 0.25$, $\rho = 0.15$ and $R = 0.5$

3.2 Single-player games with non-constant discounting

In this subsection, we consider the state equation (2.1) and the objective (2.3) with $u_1 \equiv 0$ and the discount function $\alpha(t) = \lambda e^{-\rho t} + (1-\lambda)e^{-\gamma t}$ for some $\gamma > \rho > 0$ and $\lambda \in (0,1)$. More precisely, player 2 hopes to maximise the functional

$$J_2(t,\xi;0,u_2) = \mathbb{E}_t\left[-\alpha(T-s)X(T)^2 - \int_t^T \alpha(s-t)Ru_2(s)^2 ds \right] \tag{3.6}$$

by choosing $u_2 \in L^2_{\mathbb{F}}(t,T)$ subject to

$$dX(s) = u_2(s)ds + \sigma X(s)dW(s), \quad s \in [t,T], \qquad X(t) = \xi. \tag{3.7}$$

Since α is a non-exponential function, the above problem is time-inconsistent. We define the consistent planning strategies that we call equilibrium closed-loop strategies as follows.

Definition 3.2 We call a closed-loop strategy $\tilde{\Theta}_2 \in L^\infty(0,T)$ with corresponding state process $\tilde{X}$ an *equilibrium strategy* if

$$\limsup_{\varepsilon \to 0+} \frac{J_2(t,\tilde{X}(t);0,\Theta_2^\varepsilon X^\varepsilon) - J_2(t,\tilde{X}(t);0,\tilde{\Theta}_2\tilde{X})}{\varepsilon} \leq 0 \tag{3.8}$$

for any $t \in [0,T)$ and $u_2 \in L^2_{\mathbb{F}}(t,T)$, where

$$\Theta_2^\varepsilon(s)X^\varepsilon(s) := \begin{cases} \tilde{\Theta}_2(s)X^\varepsilon(s), & s \in [t+\varepsilon, T], \\ u_2(s), & s \in [t, t+\varepsilon), \end{cases} \tag{3.9}$$

with X^ε being the corresponding state process.

The term "equilibrium strategy" refers to the consistent planning approaches envisioned by Strotz [21]. According to the definition in conditions (3.8) and (3.9), an equilibrium strategy exists when no player who can control the system during the time interval $[t, t+\epsilon]$ has an incentive to deviate from the closed-loop policy $\tilde{\Theta}_2$, provided that all players apply the same strategy $\tilde{\Theta}_2$ during the remaining period $[t+\epsilon, T]$. This condition needs to be satisfied when ε is arbitrarily small. In other words, an equilibrium strategy is a closed-loop control that allows all players to maximise their respective objectives without any player having an incentive to deviate from it. The construction (3.8) and (3.9) is sometimes called the spike variation approach, introduced in a deterministic setting by Ekeland and Lazrak [10]. Using the multi-person differential game method given by Yong [26], we get the corresponding equilibrium closed-loop strategy.

Proposition 3.3 *An equilibrium strategy associated with the state equation* (3.6) *and the functional* (3.7) *is given by*

$$\tilde{\Theta}_2(s) = -\frac{P(s,s)}{R}, \qquad s \in [0,T],$$

where for any $0 \le t \le s \le T$, the function P is the unique solution of the Riccati equation

$$\begin{aligned} 0 &= P_s(t,s) + P(t,s)\sigma^2 - 2\frac{P(t,s)P(s,s)}{R} \\ &\quad + \left(\lambda e^{-\rho(s-t)} + (1-\lambda)e^{-\gamma(s-t)}\right)\frac{P(s,s)^2}{R}, \\ P(t,T) &= \lambda e^{-\rho(T-t)} + (1-\lambda)e^{-\gamma(T-t)}. \end{aligned} \tag{3.10}$$

Remark 3.4 The results in Proposition 3.3 can be obtained by simply applying Yong [26, Theorem 6.5] to a special case. Theorem 5.2 presented later in this paper also establishes the solvability of (3.10), but with a method different from that of Yong [26].

Note that (3.10) is a non-local ordinary differential equation (ODE, for short) which does not admit an explicit solution. The non-local feature stems from the fact that (3.10) involves the evaluation of the solution at two different points t and s in time, that is, outside the diagonal of $\Delta[0,T]$.

We now describe an algorithm that provides an explicit approximate sequence for the solution of (3.10). Let Π be a uniform partition of the time interval $[0,T]$ with $t_i = \frac{iT}{N}$ for $i = 0,1,\ldots,N$. Then $\|\Pi\| = \max_{1\le i\le N}(t_i - t_{i-1}) = \frac{1}{N}$. The algorithm begins by constructing an approximation on the final subinterval $\Delta[t_{N-1}, t_N]$ and then proceeds to build the remainder of the approximation through backward recursion.

Step 1: Approximation on $\Delta[t_{N-1}, t_N]$. For $s \in [t_{N-1}, T]$, let

$$P(t_{N-1}; s) = \frac{1}{\frac{1}{e^{\sigma^2(T-s)}\alpha(T-t_{N-1})} + \int_s^T \frac{1}{e^{\sigma^2(r-s)}\alpha(r-t_{N-1})R}dr}. \tag{3.11}$$

Denote

$$\begin{aligned} P^{\Pi}(t,s) &= P(t_{N-1};s), \qquad (t,s) \in \Delta[t_{N-1},T], \\ \Theta_2^{\Pi}(s) &= -\frac{1}{\alpha(s-t_{N-1})R}P(t_{N-1};s), \qquad s \in [t_{N-1},t_N]. \end{aligned}$$

Remaining steps: Approximation on $\Delta[t_k, t_N]$ with $k = 0,1,\ldots,N-2$. Assuming that P^{Π} has been determined on the set $\Delta[t_{k+1},T]$ and that Θ_2^{Π} has been determined on the interval $[t_{k+1},T]$, let

$$\begin{aligned} P(t_k;s) &= \alpha(T-t_k)e^{\int_s^T (2\Theta_2^{\Pi}(\tau)+\sigma^2)d\tau} \\ &\quad + \int_s^T e^{\int_s^r (2\Theta_2^{\Pi}(\tau)+\sigma^2)d\tau}\alpha(r-t_k)R\Theta_2^{\Pi}(r)^2 dr, \qquad s \in [t_{k+1},t_N], \\ P(t_k;s) &= \frac{1}{\frac{1}{e^{\sigma^2(t_{k+1}-s)}P(t_k;t_{k+1})} + \int_s^{t_{k+1}} \frac{1}{e^{\sigma^2(r-s)}\alpha(r-t_k)R}dr}, \qquad s \in [t_k,t_{k+1}]. \end{aligned} \tag{3.12}$$

Since $k = 0, 1, \ldots, N-1$ is an arbitrarily given index, the above approximation on $\Delta[t_k, t_N]$ extends to the interval $[0, T]$. Then for any $N > 0$, P^{Π} and Θ_2^{Π} can be explicitly obtained on $[0, T]$ by induction.

We now state a convergence result in Theorem 3.5. The result is due to Yong [26], but for completeness, we give a shorter and self-contained proof in the Appendix.

Theorem 3.5 *The unique solution P of (3.10) can be obtained as the limit*

$$P(t,s) = \lim_{\|\Pi\| \to 0} P^{\Pi}(t,s), \qquad (t,s) \in \Delta[0,T],$$

where the function P^{Π} is produced by the algorithm (3.11), (3.12). Thus the equilibrium strategy $\tilde{\Theta}_2$ can also be obtained as the limit

$$\tilde{\Theta}_2(s) = \lim_{\|\Pi\| \to 0} \Theta^{\Pi}(s) = \lim_{\|\Pi\| \to 0} -\frac{P^{\Pi}(s,s)}{R}, \qquad s \in [0,T]. \tag{3.13}$$

Figure 2 illustrates the equilibrium strategy $\tilde{\Theta}_2$ obtained through the use of algorithm (3.11) and (3.12), where the step size of the partition Π approaches zero. The results show that the optimal strategy's properties in Fig. 1 hold for the equilibrium strategy $\tilde{\Theta}_2$ as well. Additionally, the absence of "overshooting" is demonstrated in Fig. 3, where the lobbying equilibrium with non-constant discounting is bounded by the optimal lobbying strategy, both when the short-term self is in control and when the long-term self is in control.

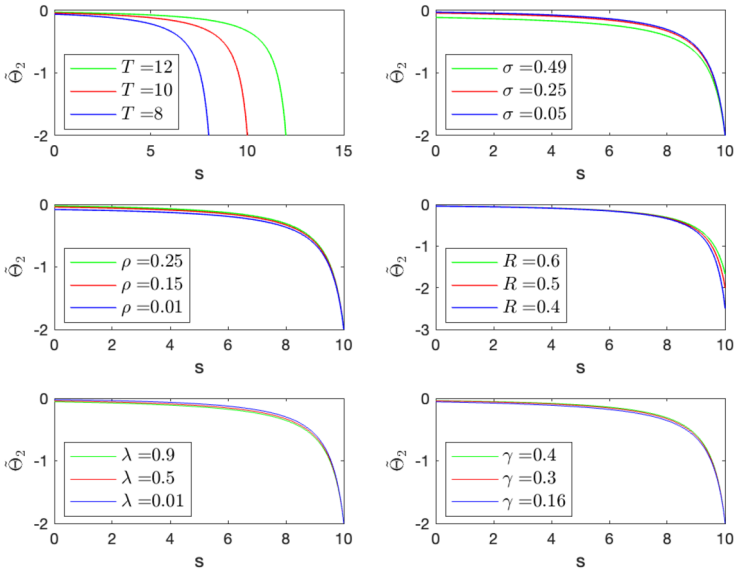

Fig. 2 The panels display the equilibrium closed-loop strategies $\tilde{\Theta}_2(s)$ for $0 \le s \le T$ associated with the single-player game with non-constant discounting, and are based on the approximation given in (3.13). The baseline parameter values used are $T = 10$, $\sigma = 0.25$, $\rho = 0.15$, $R = 0.5$, $\lambda = 0.3$ and $\gamma = 0.3$

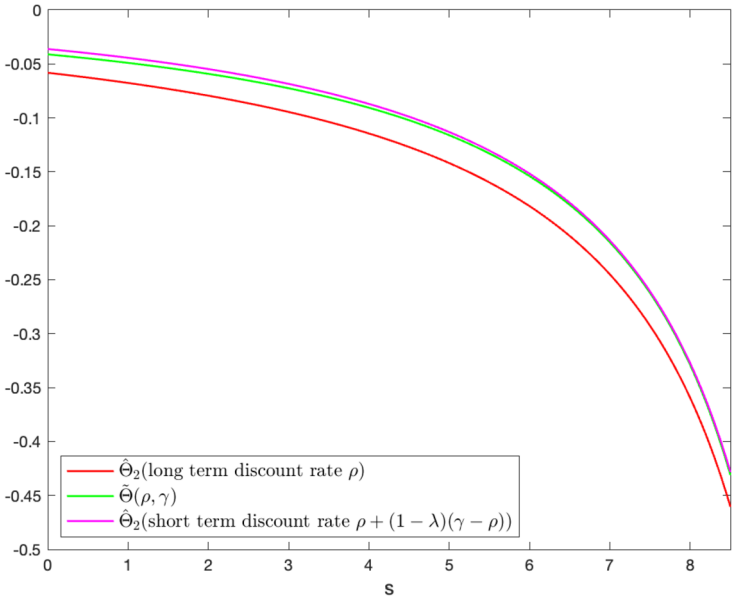

Fig. 3 The figure shows for $0 \le s \le T$ the equilibrium closed-loop strategy $\tilde{\Theta}_2(s)$ in green, and the optimal closed-loop strategy $\hat{\Theta}_2(s)$ when the discount rate is ρ (in pink) or $\rho+(1-\lambda)(\gamma-\rho)$ (in red). The strategy $\tilde{\Theta}_2(s)$ is obtained using the approximation (3.13). The baseline parameter values are $T=10$, $\sigma=0.25$, $\rho=0.15$, $R=0.5$, $\lambda=0.5$ and $\gamma=0.3$

4 Two-player zero-sum games with constant discounting

In this section, we consider the game problem described by (2.1)–(2.3) with the discount function $\alpha(t)=e^{-\rho t}$ for some $\rho>0$. Recall that we denoted the game problem by problem (G). We first define a more general auxiliary game that will be used for solving our game both with constant and non-constant discounting. Second, we provide the closed-form closed-loop saddle strategy for problem (G) with constant discounting.

4.1 An auxiliary game

We introduce the more general game problem with the state equation (2.1) and the objective

$$J(t,\xi;u_1,u_2)=\mathbb{E}_t\Big[e^{-\rho(T-t)}GX(T)^2 + \int_t^T e^{-\rho(s-t)}\big(R_1(s)u_1(s)^2+R_2(s)u_2(s)^2\big)ds\Big] \tag{4.1}$$

for some scalar G and deterministic functions R_1 and R_2. Note that the above reduces to problem (G) with constant discounting when $G=1$, $R_1\equiv -1$ and $R_2\equiv R$. By Sun and Yong [22, Theorem 5.5], we have the following characterisation of the closed-loop saddle strategy of the above game.

Lemma 4.1 *Suppose that $G \geq 0$, $R_1 < 0$ and $R_2 > 0$. The game problem with the state equation (2.1) and the objective (4.1) admits a closed-loop saddle strategy (Θ_1^*, Θ_2^*) if and only if the Riccati equation*

$$0 = \dot{P}(s) + \sigma^2 P(s) - \rho P(s) - \frac{R_1(s) + R_2(s)}{R_1(s)R_2(s)} P(s)^2,$$
$$P(T) = G \tag{4.2}$$

admits a solution P. If (4.2) admits a unique solution P, then the unique closed-loop saddle strategy (Θ_1^*, Θ_2^*) admits the representation

$$\Theta_1^*(s) = -\frac{P(s)}{R_1(s)}, \quad \Theta_2^*(s) = -\frac{P(s)}{R_2(s)}, \qquad s \in [0, T]. \tag{4.3}$$

4.2 Closed-loop saddle strategies

We now turn to problem (G) with constant discounting. By Lemma 4.1, we get the following result.

Proposition 4.2 *Problem (G) with $\alpha(t) = e^{-\rho t}$ admits a unique closed-loop saddle strategy (Θ_1^*, Θ_2^*), given for $s \in [0, T]$ by*

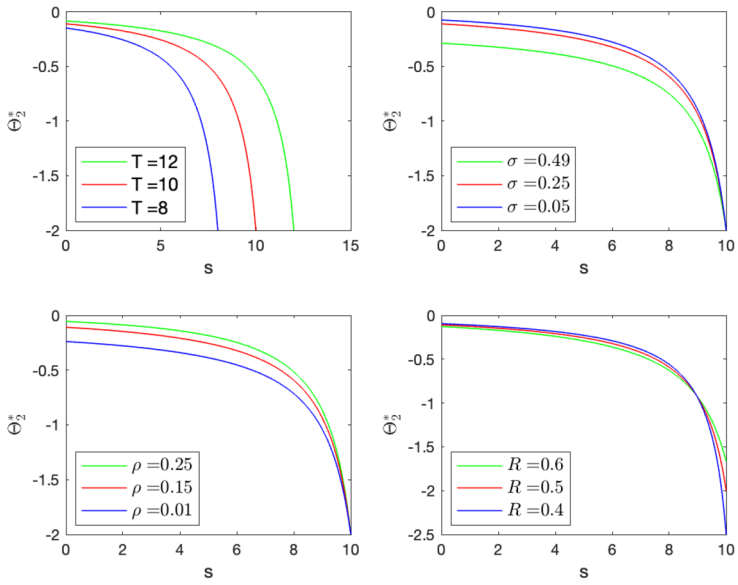

Fig. 4 The figure panels display the closed-loop saddle strategies $\Theta_2^*(s)$ for problem (G) with constant discounting, plotted for $0 \leq s \leq T$ using the closed-form expression given by (4.4). We use the baseline parameter values $T = 10$, $\sigma = 0.25$, $\rho = 0.15$, $R = 0.5$

$$\Theta_1^*(s) = \frac{R(\rho-\sigma^2)e^{(\rho-\sigma^2)s}}{(1-R+R(\rho-\sigma^2))e^{(\rho-\sigma^2)T}-(1-R)e^{(\rho-\sigma^2)s}},$$
$$\Theta_2^*(s) = -\frac{\Theta_1^*(s)}{R}. \tag{4.4}$$

In particular, if $R = 1$ so that the game is symmetric, the unique closed-loop saddle strategy (Θ_1^*, Θ_2^*) admits the representation

$$\Theta_1^*(s) = (\rho-\sigma^2)e^{(\rho-\sigma^2)(s-T)}, \quad \Theta_2^*(s) = -\Theta_1^*(s), \qquad s \in [0,T].$$

To visualise the impact of strategic interaction between the two players, Fig. 4 plots player 2's closed-loop saddle strategies for problem (G) and the optimal strategy for the same players. The strategies are displayed as functions of time, at various parameter levels for player 2. The figure shows that the comparative statics intuitions from the single-player game hold true in the two-player game.

5 Two-player zero-sum games with non-constant discounting

In this section, we explore problem (G) in the case where the discount rate is not a constant, and the discount function is defined by (2.8). First, we define the equilibrium. Second, we introduce a modified version of problem (G) in which we divide the time interval $[0, T]$ into subintervals and assume that players can make commitments during each subinterval. This discretisation enables us to identify a differential equation that has the potential to characterise an equilibrium. Third, we demonstrate that the problem is well posed and that an equilibrium can indeed be characterised by the identified equation. Finally, we present an algorithm that can be used to approximate an equilibrium.

5.1 Equilibrium definition

The introduction of non-constant discounting in problem (G) adds a strategic dimension for each player in two ways. Firstly, each player is strategic in their interaction with the other player. Secondly, each player is strategic in their interaction with their future selves. Consequently, problem (G) effectively becomes a problem with an infinite number of players: the continuum of incarnations of player 1 and the continuum of incarnations of player 2. To account for the time inconsistency generated by non-constant discounting, we define equilibrium closed-loop saddle strategies as follows.

Definition 5.1 We say a closed-loop strategy $(\bar{\Theta}_1, \bar{\Theta}_2) \in L^\infty[0,T] \times L^\infty[0,T]$ with corresponding state process $\bar{X}$ satisfies a *local saddle property* if

$$\limsup_{\varepsilon\to 0+} \frac{J(t,\bar{X}(t);\Theta_1^\varepsilon X^\varepsilon, \bar{\Theta}_2 X^\varepsilon) - J(t,\bar{X}(t);\bar{\Theta}_1\bar{X},\bar{\Theta}_2\bar{X})}{\varepsilon} \le 0,$$
$$\liminf_{\varepsilon\to 0+} \frac{J(t,\bar{X}(t);\bar{\Theta}_1 X^\varepsilon, \Theta_2^\varepsilon X^\varepsilon) - J(t,\bar{X}(t);\bar{\Theta}_1\bar{X},\bar{\Theta}_2\bar{X})}{\varepsilon} \ge 0 \tag{5.1}$$

for any $t \in [0, T)$ and $u_i \in L^2_{\mathbb{F}}(t, T)$, where

$$\Theta_i^{\varepsilon}(s)X^{\varepsilon}(s) := \begin{cases} \bar{\Theta}_i(s)X^{\varepsilon}(s), & s \in [t+\varepsilon, T], \\ u_i(s), & s \in [t, t+\varepsilon), \end{cases} \tag{5.2}$$

with X^{ε} being the corresponding state process. We call a closed-loop strategy satisfying the local saddle property (5.1) an *equilibrium closed-loop saddle strategy*.

The first inequality in (5.1) implies that if player 1 deviates from their strategy during a commitment period of negligible length, their objective function will not improve compared to following the no-deviation strategy. This constraint shares similarities with the definition of a consistent planning strategy for a single player as in Definition 3.2. However, the key difference is that in the game-theoretic setting, each player turns to deviate while keeping unchanged the behaviour of both their future selves during the period $[t+\varepsilon, T]$ and the other player's strategy. Strategies that meet the criteria of the local saddle property (5.2) are denominated as equilibrium closed-loop saddle strategies. Here, "equilibrium" pertains to the game involving multiple selves, while "saddle" refers to the game between player 1 and player 2.

5.2 Time partitions with precommitment in each subinterval

We first let Π be a partition of the time interval $[0, T]$,

$$0 = t_0 < t_1 < t_2 < \cdots < t_{N-1} < t_N = T,$$

with mesh size

$$\|\Pi\| = \max_{1 \le i \le N} (t_i - t_{i-1}).$$

We assume that within each subinterval $[t_k, t_{k+1}]$, the players are allowed to make commitments and adjust their strategies accordingly, before moving on to the next subinterval. As a result, during the time interval $[t_k, t_{k+1}]$, self t_k of each player will play a two-person zero-sum game against self t_k of the other player, with self t_k dictating the strategy choice. To fully specify the problem, we need to define the objective for each player in each subinterval and establish the connection between the game in a given subinterval and the game in the next subintervals. To establish this connection, we assume consistent planning: when solving the game during a given interval $[t_{k-1}, t_k]$, the decision makers internalise how the game will be solved in the subsequent interval $[t_k, T]$. We do this in several steps.

Step 1: The precommitment game on $[t_{N-1}, t_N]$. The precommitment game that we envision is driven by the state equation

$$\begin{aligned} dX^N(s) &= \big(u_1^N(s) + u_2^N(s)\big)ds + \sigma X^N(s)dW(s), \qquad s \in [t_{N-1}, T], \\ X^N(t_{N-1}) &= \xi, \end{aligned} \tag{5.3}$$

and the functionals

$$J_1^N(t_{N-1},\xi;u_1^N,u_2^N)=\mathbb{E}_{t_{N-1}}\Big[\alpha(T-t_{N-1})X(T)^2 + \int_t^T \alpha(s-t_{N-1})\big(-u_1^N(s)^2+Ru_2^N(s)^2\big)ds\Big],$$
$$J_2^N(t_{N-1},\xi;u_1^N,u_2^N)=-J_1^N(t_{N-1},\xi;u_1^N,u_2^N). \tag{5.4}$$

Notice that with precommitment, every player within the interval $[t_{N-1},t_N]$ applies the same discount function that is applied by self t_{N-1}. As a result, the initial time t_{N-1} is fixed in (5.4). With the fixed t_{N-1}, the problem in (5.3) and (5.4) is a standard linear–quadratic stochastic differential game with zero discounting. As such, it can be solved by Lemma 4.1 with $\rho=0$, $G=\alpha(T-t_{N-1})$, $R_1(s)=-\alpha(s-t_{N-1})$ and $R_2(s)=\alpha(s-t_{N-1})R$.

The unique closed-loop saddle strategy $(\bar\Theta_1^N,\bar\Theta_2^N)$ is given by

$$\bar\Theta_1^N(s)=\frac{P(t_{N-1};s)}{\alpha(s-t_{N-1})},\quad \bar\Theta_2^N(s)=-\frac{P(t_{N-1};s)}{R\alpha(s-t_{N-1})},\qquad s\in[t_{N-1},T],$$

with $P(t_{N-1};s)$, $s\in[t_{N-1},T]$, being the unique solution of the Riccati equation

$$0=P_s(t_{N-1};s)+P(t_{N-1};s)\sigma^2-\frac{1-R}{\alpha(s-t_{N-1})R}P(t_{N-1};s)^2$$
$$\text{for } s\in[t_{N-1},T],$$
$$P(t_{N-1};T)=\alpha(T-t_{N-1}). \tag{5.5}$$

Note that in (5.5), t_{N-1} is fixed and only works as a parameter. Denote

$$\bar\Theta_i^\Pi=\bar\Theta_i^N\qquad \text{on } [t_{N-1},T] \text{ for } i=1,2.$$

Steps 2: The precommitment game on $[t_{N-2},t_{N-1}]$. With the strategy $\bar\Theta_i^\Pi$ determined on $[t_{N-1},T]$, the players on $[t_{N-2},T]$ can only control the system on the interval $[t_{N-2},t_{N-1}]$. Consider the state equation

$$dX^{N-1}(s)=\big(u_1^{N-1}(s)+u_2^{N-1}(s)\big)ds+\sigma X^{N-1}(s)dW(s)$$
$$\text{for } s\in[t_{N-2},t_{N-1}],$$
$$dX^{N-1}(s)=\big(\bar\Theta_1^\Pi(s)X^{N-1}(s)+\bar\Theta_2^\Pi(s)X^{N-1}(s)\big)ds+\sigma X^{N-1}(s)dW(s)$$
$$\text{for } s\in[t_{N-1},T],$$
$$X^{N-1}(t_{N-2})=\xi$$

and the functionals $J_i^{N-1}(t_{N-2},\xi;u_1^{N-1},u_2^{N-1})$, $i=1,2$, with

$$
\begin{aligned}
&J_1^{N-1}(t_{N-2},\xi;u_1^{N-1},u_2^{N-1})\\
&=-J_2^{N-1}(t_{N-2},\xi;u_1^{N-1},u_2^{N-1})\\
&=J^{N-1}(t_{N-2},\xi;u_1^{N-1},u_2^{N-1})\\
&=\mathbb{E}_{t_{N-2}}\Bigg[\alpha(T-t_{N-2})X^{N-1}(T)^2\\
&\qquad+\int_{t_{N-1}}^{T}\alpha(s-t_{N-2})\Big(-\big(\bar\Theta_i^{\Pi}(s)X^{N-1}(s)\big)^2+R\big(\bar\Theta_i^{\Pi}(s)X^{N-1}(s)\big)^2\Big)ds\Bigg]\\
&\quad+\mathbb{E}_{t_{N-2}}\Bigg[\int_{t_{N-2}}^{t_{N-1}}\alpha(s-t_{N-2})\big(-u_1^{N-1}(s)^2+Ru_2^{N-1}(s)^2\big)ds\Bigg]\\
&=:\mathrm{(I)}+\mathrm{(II)}.
\end{aligned}
$$

Let $P(t_{N-2};s)$, $s\in[t_{N-1},T]$, be the unique solution of the Lyapunov equation

$$
\begin{aligned}
0&=P_s(t_{N-2};s)+2P(t_{N-2};s)\big(\bar\Theta_1^{\Pi}(s)+\bar\Theta_2^{\Pi}(s)\big)+P(t_{N-2};s)\sigma^2\\
&\quad+\alpha(s-t_{N-2})\big(-\bar\Theta_1^{\Pi}(s)^2+R\bar\Theta_2^{\Pi}(s)^2\big),\qquad s\in[t_{N-1},T],\\
P(t_{N-2};T)&=\alpha(T-t_{N-2}).
\end{aligned}
\tag{5.6}
$$

Then by applying Itô's formula to $s\mapsto P(t_{N-2};s)X^{N-1}(s)^2$ on $[t_{N-1},T]$, we have

$$
\mathrm{(I)}=P(t_{N-2};t_{N-1})X^{N-1}(t_{N-1})^2.
$$

It follows that

$$
\begin{aligned}
&J^{N-1}(t_{N-2},\xi;u_1^{N-1},u_2^{N-1})\\
&=\mathbb{E}\Bigg[P(t_{N-2};t_{N-1})X^{N-1}(t_{N-1})^2\\
&\qquad+\int_{t_{N-2}}^{t_{N-1}}\alpha(s-t_{N-2})\big(-u_1^{N-1}(s)^2+Ru_2^{N-1}(s)^2\big)ds\Bigg].
\end{aligned}
$$

Moreover, from the fact that

$$
-\bar\Theta_1^{\Pi}(s)^2+R\bar\Theta_2^{\Pi}(s)^2=\frac{(1-R)P(t_{N-1};s)^2}{R\alpha(s-t_{N-1})^2}\geq 0,
$$

we have

$$
P(t_{N-2};t_{N-1})\geq 0.
$$

Thus the problem can be solved by Lemma 4.1 again with $\rho=0$, $G=P(t_{N-2};t_{N-1})$, $R_1(s)=-\alpha(s-t_{N-2})$ and $R_2(s)=\alpha(s-t_{N-2})R$.

The unique closed-loop saddle strategy $(\bar\Theta_1^{N-1},\bar\Theta_2^{N-1})$ is given by

$$
\bar\Theta_1^{N-1}(s)=\frac{P(t_{N-2};s)}{\alpha(s-t_{N-2})},\quad \bar\Theta_2^{N-1}(s)=-\frac{P(t_{N-2};s)}{R\alpha(s-t_{N-2})},\qquad s\in[t_{N-2},t_{N-1}],
$$

with $P(t_{N-2};s)$, $s\in[t_{N-2},t_{N-1}]$, being the unique solution of the Riccati equation

$$\begin{aligned}0&=P_s(t_{N-2};s)+P(t_{N-2};s)\sigma^2-\frac{1-R}{\alpha(s-t_{N-2})R}P(t_{N-2};s)^2\\&\qquad\text{for } s\in[t_{N-2},t_{N-1}],\\P(t_{N-2};t_{N-1})&=P(t_{N-2};t_{N-1}).\end{aligned}\tag{5.7}$$

Note that in the above, $P(t_{N-2};t_{N-1})$ has been determined by (5.6). We now extend $\bar\Theta_i^\Pi$ from $[t_{N-1},T]$ to $[t_{N-2},T]$ by

$$\bar\Theta_i^\Pi(s)=\begin{cases}\bar\Theta_i^N(s), & s\in[t_{N-1},T],\\ \bar\Theta_i^{N-1}(s), & s\in[t_{N-2},t_{N-1})\end{cases}$$

for $i=1,2$. With this, we can write (5.6) and (5.7) together as

$$\begin{aligned}0&=P_s(t_{N-2};s)+2P(t_{N-2};s)\big(\bar\Theta_1^\Pi(s)+\bar\Theta_2^\Pi(s)\big)+P(t_{N-2};s)\sigma^2\\&\quad+\alpha(s-t_{N-2})\big(-\bar\Theta_1^\Pi(s)^2+R\bar\Theta_2^\Pi(s)^2\big),\qquad s\in[t_{N-2},T],\\P(t_{N-2};T)&=\alpha(T-t_{N-2}).\end{aligned}$$

Subsequent Steps: The precommitment game on $[t_{k-1},t_k]$ for $k=1,2,\dots,N$. Suppose $\bar\Theta_i^\Pi$, $i=1,2$, has been constructed on $[t_k,T]$ for some $k=1,2,\dots,N-1$. We apply the above procedure to obtain an extension $\bar\Theta_i^\Pi:[t_{k-1},T]\to\mathbb{R}$ of $\bar\Theta_i^\Pi(s)$, $s\in[t_k,T]$, by setting, for $i=1,2$,

$$\bar\Theta_i^\Pi(s)=\begin{cases}\bar\Theta_i^\Pi(s), & s\in[t_k,T],\\ \bar\Theta_i^k(s), & s\in[t_{k-1},t_k),\end{cases}\tag{5.8}$$

where

$$\bar\Theta_1^k(s)=\frac{P(t_{k-1};s)}{\alpha(s-t_{k-1})},\quad \bar\Theta_2^k(s)=-\frac{P(t_{k-1};s)}{R\alpha(s-t_{k-1})},\qquad s\in[t_{k-1},t_k],$$

with $P(t_{k-1};s)$, $s\in[t_{k-1},T]$, being the unique solution of the Riccati equation

$$\begin{aligned}0&=P_s(t_{k-1};s)+2P(t_{k-1};s)\big(\bar\Theta_1^\Pi(s)+\bar\Theta_2^\Pi(s)\big)+P(t_{k-1};s)\sigma^2\\&\quad+\alpha(s-t_{k-1})\big(-\bar\Theta_1^\Pi(s)^2+R\bar\Theta_2^\Pi(s)^2\big)\\&\quad\text{for } s\in[t_k,T],\\0&=P_s(t_{k-1};s)+P(t_{k-1};s)\sigma^2-\frac{1-R}{\alpha(s-t_{k-1})R}P(t_{k-1};s)^2\\&\quad\text{for } s\in[t_{k-1},t_k],\\P(t_{k-1};T)&=\alpha(T-t_{k-1}),\end{aligned}$$

which is equivalent to

$$\begin{aligned}0 &= P_s(t_{k-1};s) + 2P(t_{k-1};s)\big(\bar{\Theta}_1^{\Pi}(s) + \bar{\Theta}_2^{\Pi}(s)\big) + P(t_{k-1};s)\sigma^2 \\ &\quad + \alpha(s-t_{k-1})\big(-\bar{\Theta}_1^{\Pi}(s)^2 + R\bar{\Theta}_2^{\Pi}(s)^2\big) = 0, \qquad s \in [t_{k-1}, T], \\ P(t_{k-1};T) &= \alpha(T-t_{k-1}).\end{aligned}$$

This completes the induction.

The novelty of the construction of a sequence of precommitment zero-sum games is that we can show that under the assumption $0 < R \leq 1$, the zero-sum stochastic linear–quadratic game on every subinterval can be solved. The key step is then to check whether we can connect the Riccati equations associated with each subinterval's two-person zero-sum game with constant discount, and then glue them together. This is what we did in the above steps.

We now describe the global policy constructed from the above steps. Denote the discount function of the precommitment model by

$$\alpha^{\Pi}(t,s) = \begin{cases} \alpha(s-t_{N-1}), & t \in [t_{N-1}, t_N], s \in [t,T], \\ \alpha(s-t_{k-1}), & t \in [t_{k-1}, t_k), s \in [t,T], k = 1, \ldots, N-1. \end{cases}$$

The above discount function formalises the assumption that self t_{k-1} applies her discount function to all selves in the interval $[t_{k-1}, t_k]$. We denote by P^{Π} the function obtained by gluing together the functions $P(t;s)$, $(t,s) \in \Delta[0,T]$, on each subinterval, i.e.,

$$P^{\Pi}(t,s) = P(t_{k-1};s), \qquad t \in [t_{k-1}, t_k), s \in [t,T], k = 1, \ldots, N.$$

Then the closed-loop strategy $(\bar{\Theta}_1^{\Pi}, \bar{\Theta}_2^{\Pi})$ associated with the partition Π can be given by

$$\bar{\Theta}_1^{\Pi}(s) = \frac{P^{\Pi}(s,s)}{\alpha^{\Pi}(s,s)}, \quad \bar{\Theta}_2^{\Pi}(s) = -\frac{P^{\Pi}(s,s)}{R\alpha^{\Pi}(s,s)}, \qquad s \in [0,T], \tag{5.9}$$

with

$$\begin{aligned}0 &= P_s^{\Pi}(t,s) + 2P^{\Pi}(t,s)\big(\bar{\Theta}_1^{\Pi}(s) + \bar{\Theta}_2^{\Pi}(s)\big) + P^{\Pi}(t,s)\sigma^2 \\ &\quad + \alpha^{\Pi}(t,s)\big(-\bar{\Theta}_1^{\Pi}(s)^2 + R\bar{\Theta}_2^{\Pi}(s)^2\big), \qquad (t,s) \in \Delta[0,T], \\ P^{\Pi}(t,T) &= \alpha^{\Pi}(T-t), \qquad t \in [0,T].\end{aligned} \tag{5.10}$$

The purpose of our precommitment strategy based on subintervals constructed in the above steps is to find an equilibrium characterisation in the continuous-time model. To accomplish this, we pretend that P^{Π} converges uniformly to some P. Then formally, P should satisfy the equation

$$\begin{aligned}0 &= P_s(t,s) + 2P(t,s)\big(\bar{\Theta}_1(s) + \bar{\Theta}_2(s)\big) + P(t,s)\sigma^2 \\ &\quad + \alpha(t,s)\big(-\bar{\Theta}_1(s)^2 + R\bar{\Theta}_2(s)^2\big), \qquad (t,s) \in \Delta[0,T], \\ P(t,T) &= \alpha(T-t), \qquad t \in [0,T],\end{aligned} \tag{5.11}$$

where

$$\bar{\Theta}_1(s) = P(s,s), \quad \bar{\Theta}_2(s) = -\frac{P(s,s)}{R}, \qquad s \in [0,T]. \tag{5.12}$$

We call (5.11) an *equilibrium Riccati equation* (ERE, for short).

5.3 Well-posedness and verification theorem

We now show that the solution to the ERE (5.11) exists and is unique.

Theorem 5.2 *The ERE* (5.11) *admits a unique solution* P. *Moreover,* $P(t,s) \geq 0$ *for* $(t,s) \in \Delta[0,T]$. *In particular, if* $R = 1$, *that is, if the game is symmetric, then the unique solution of ERE* (5.11) *can be explicitly given by*

$$P(t,s) = e^{-\sigma^2(s-T)}\alpha(T-t), \qquad (t,s) \in \Delta[t,T]. \tag{5.13}$$

Next we show that the closed-loop strategy $\bar{\Theta}_i$, $i = 1,2$, obtained by (5.11) and (5.12) satisfies the local saddle property and is therefore an equilibrium closed-loop saddle strategy.

Theorem 5.3 *Let* $\bar{\Theta}_i$, $i = 1,2$, *be the closed-loop strategy obtained by* (5.11) *and* (5.12). *Then* $\bar{\Theta}_i$, $i = 1,2$, *satisfies the local saddle property* (5.1). *Thus* $\bar{\Theta}_i$, $i = 1,2$, *is an equilibrium closed-loop saddle strategy for problem* (G) *with non-constant discounting*.

5.4 Convergence and approximation algorithm

We now show that when the partition step is small, the precommitment strategy converges to the equilibrium closed-loop saddle strategy. This is an important result because it builds an equilibrium as a limit of models with precommitment over arbitrarily small periods. The result provides thus a discrete-time foundation of the spike variation method. In a mean–variance model, a similar result was obtained by Czichowsky [9].

Theorem 5.4 *As* $\|\Pi\| \to 0$, *the sequences* $(P^\Pi)_\Pi$ *and* $(\bar{\Theta}_i^\Pi)_\Pi$ *defined by* (5.9) *and* (5.10) *converge uniformly to* P *and* $\bar{\Theta}$, *respectively, where* P *and* $\bar{\Theta}$ *are defined by* (5.11) *and* (5.12).

It should be noted that an explicit solution to the non-local ODE (5.11) cannot be obtained in general. However, for a given partition Π, we can obtain an explicit solution for P^Π. By applying Theorem 5.4, we can use P^Π to approximate the solution of (5.11). This will be our approach going forward.

To start, let Π be an equidistant partition of the time interval $[0,T]$ with $t_i = \frac{iT}{N}$ for $i = 0,1,\ldots,N$. Then $\|\Pi\| = \max_{1\leq i\leq N}(t_i - t_{i-1}) = \frac{1}{N}$. We now describe the algorithm.

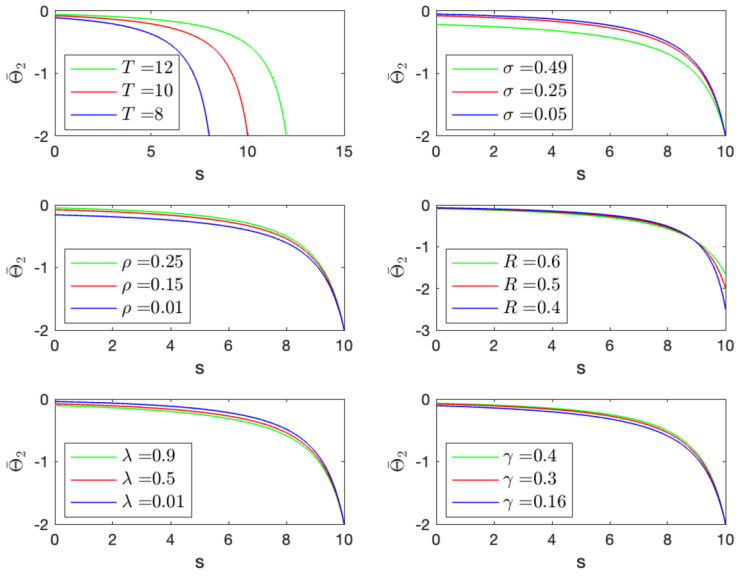

Fig. 5 The figure panels display the equilibrium closed-loop saddle strategies $\bar{\Theta}_2(s)$ for problem (G) with non-constant discounting, plotted for $0 \le s \le T$ using the approximation given by (5.14). We use the baseline parameter values $T = 10$, $\sigma = 0.25$, $\rho = 0.15$, $R = 0.5$, $\lambda = 0.5$ and $\gamma = 0.3$

Step 1: Approximation on $\Delta[t_{N-1}, t_N]$. Let

$$P(t_{N-1}; s) = \frac{1}{\frac{1}{e^{\sigma^2(T-s)}\alpha(T-t_{N-1})} + \int_s^T \frac{1-R}{e^{\sigma^2(r-s)}\alpha(r-t_{N-1})R}dr}, \qquad s \in [t_{N-1}, T].$$

Denote

$$P^{\Pi}(t,s) = P(t_{N-1}; s), \qquad (t,s) \in \Delta[t_{N-1}, T],$$

$$\Theta_1^{\Pi}(s) = \frac{P(t_{N-1}; s)}{\alpha(s - t_{N-1})}, \quad \Theta_2^{\Pi}(s) = -\frac{P(t_{N-1}; s)}{R\alpha(s - t_{N-1})}, \qquad s \in [t_{N-1}, t_N].$$

Step 2: Approximation of general term on $\Delta[t_k, t_N]$ with $k = 0, \ldots, N-2$. Suppose P^{Π} has been determined on $\Delta[t_{k+1}, T]$ and Θ^{Π} on the interval $[t_{k+1}, T]$. Let

$$P(t_k; s) = \alpha(T - t_k)e^{\int_s^T (2\Theta_1^{\Pi}(\tau)+2\Theta_2^{\Pi}(\tau)+\sigma^2)d\tau} + \int_s^T e^{\int_s^r (2\Theta_1^{\Pi}(\tau)+2\Theta_2^{\Pi}(\tau)+\sigma^2)d\tau} \times \alpha(r - t_k)\big(-\Theta_1^{\Pi}(r)^2 + R\Theta_2^{\Pi}(r)^2\big)dr, \qquad s \in [t_{k+1}, t_N],$$

$$P(t_k; s) = \frac{1}{\frac{1}{e^{\sigma^2(t_{k+1}-s)}P(t_k; t_{k+1})} + \int_s^{t_{k+1}} \frac{1-R}{e^{\sigma^2(r-s)}\alpha(r-t_k)R}dr}, \qquad s \in [t_k, t_{k+1}].$$

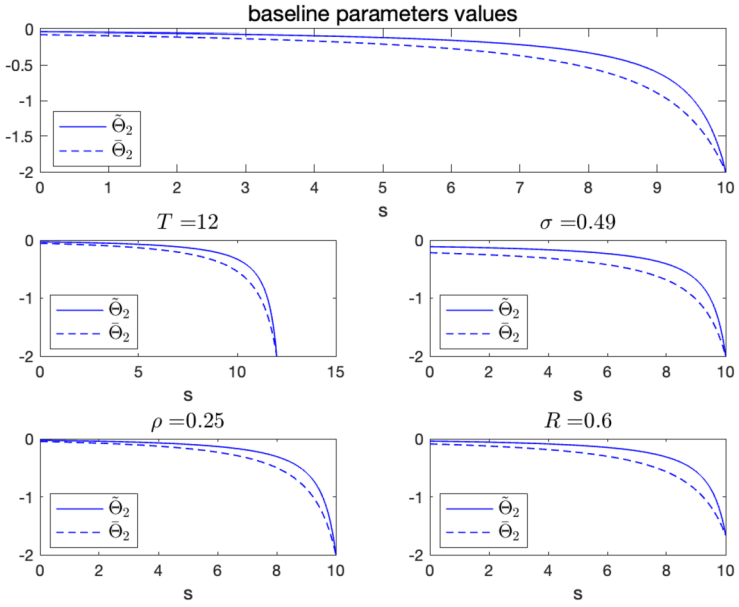

Fig. 6 The figure panels display for $0 \le s \le T$ the equilibrium closed-loop saddle strategies $\bar{\Theta}_2(s)$ for problem (G) with non-constant discounting, as well as the equilibrium closed-loop strategies $\tilde{\Theta}_2(s)$ for the single player with non-constant discounting. The figures are based on the approximations given by (5.14) and (3.13). We use the baseline parameter values $T = 10$, $\sigma = 0.25$, $\rho = 0.15$, $R = 0.5$, $\lambda = 0.5$ and $\gamma = 0.3$

Denote

$$P^{\Pi}(t,s) = P(t_k;s), \qquad (t,s) \in \Delta[t_k,T] \setminus \Delta[t_{k+1},T],$$

$$\Theta_1^{\Pi}(s) = \frac{P(t_k;s)}{\alpha(s-t_k)}, \quad \Theta_2^{\Pi}(s) = -\frac{P(t_k;s)}{R\alpha(s-t_k)}, \qquad s \in [t_k,t_{k+1}].$$

For any $N > 0$, P^{Π} and Θ^{Π} can be explicitly obtained on $[0,T]$ by induction. Then the unique solution P of (3.10) can be obtained by

$$P(t,s) = \lim_{\|\Pi\|\to 0} P^{\Pi}(t,s), \qquad (t,s) \in \Delta[0,T],$$

and the equilibrium strategy $\bar{\Theta}_i$ can be obtained by

$$\bar{\Theta}_1(s) = \lim_{\|\Pi\|\to 0} \Theta_1^{\Pi}(s) = \lim_{\|\Pi\|\to 0} P^{\Pi}(s,s), \qquad s \in [0,T],$$

$$\bar{\Theta}_2(s) = \lim_{\|\Pi\|\to 0} \Theta_2^{\Pi}(s) = -\lim_{\|\Pi\|\to 0} \frac{P^{\Pi}(s,s)}{R}, \qquad s \in [0,T]. \tag{5.14}$$

Figure 5 displays the approximate strategies $(\bar{\Theta}_2^{\Pi})$ obtained by using the above algorithm to visualise the equilibrium closed-loop saddle strategy $\bar{\Theta}_2$. By comparing Figs. 5 and 2, we see that the qualitative characteristics of the lobbying strategies are

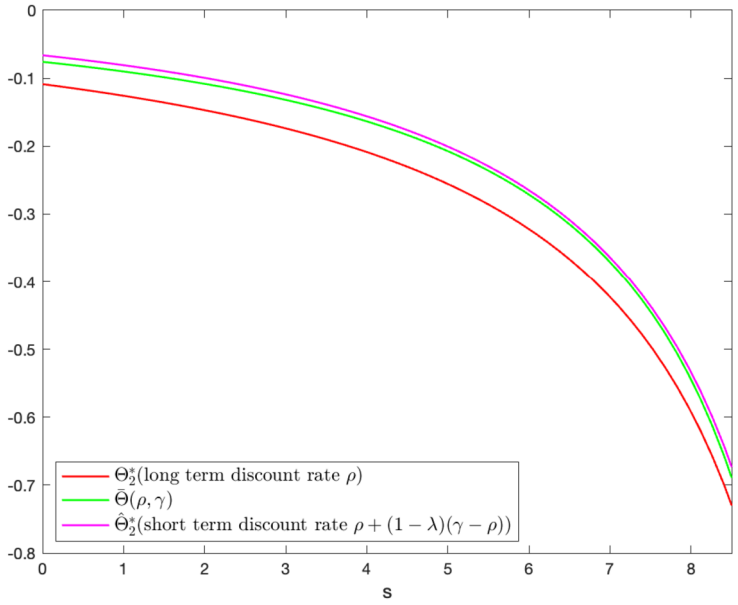

Fig. 7 The figure panels display for $0 \leq s \leq T$ the equilibrium saddle strategies $\bar{\Theta}_2(s)$ for problem (G) with non-constant discounting, and the closed-loop saddle strategies $\Theta_2^*(s)$ for problem (G) with constant discounting. The strategies $\bar{\Theta}_2(s)$ are approximated using the expression (5.14), and the strategies $\Theta_2^*(s)$ are given in the closed-form expression (4.4). We use the baseline parameter values $T = 10$, $\sigma = 0.25$, $\rho = 0.15$, $R = 0.5$, $\lambda = 0.5$ and $\gamma = 0.3$. To visualise the wedge between curves, we truncated the time axis at $s = 8.5$. We have verified numerically that there is no overshooting in the interval $s \in [8.5, 10]$

broadly consistent with those of a single-player game with non-constant discounting. We compare the equilibrium closed-loop saddle lobbying strategy $\bar{\Theta}_2$ to the equilibrium lobbying strategy $\tilde{\Theta}_2$ for a single player with non-constant discounting by plotting them together in Fig. 6. The figure clearly demonstrates that lobbying intensifies in a two-player game setting relative to a single-player setting, indicating that strategic interaction between the two players increases lobbying efforts, even when the discount rate is non-constant. This general observation holds true across all panels in Fig. 6, producing results similar to those obtained when the discount rate is constant and summarised in Fig. 4. Figure 7 further confirms that the equilibrium behaviour effectively prevents the phenomenon of "overshooting". Specifically, the closed-loop saddle lobbying strategy with non-constant discount rates is bounded from above (resp. from below) by the closed-loop saddle lobbying strategies obtained when the players apply the constant short term (resp. long term) discount rate.

5.5 Conclusions

We investigate a zero-sum linear–quadratic stochastic differential game in which two players lobby a political representative to invest in a wind farm. As the players have time-inconsistent preferences, they discount short-term utility with a large discount rate and long-term utility with a low discount rate. Our aim is to determine the equi-

librium lobbying behaviour of the players in both single-player and two-player frameworks.

We find that an equilibrium in the class of linear closed-loop controls exists and identify it in closed form or in a form that can be approximated. Our analysis reveals that non-constant discounting does not significantly alter the comparative statics with respect to most model parameters.

Our study has revealed that despite the aforementioned findings, strategic behaviour consistently leads to an increased lobbying intensity in all situations. Additionally, we have shown that there is no overshooting phenomenon. Specifically, we have found that the equilibrium behaviour under non-constant discounting is constrained from below by the lobbying behaviour when players prioritise long-term utility at a high discount rate, and from above by the lobbying behaviour when players prioritise short-term utility at a low discount rate. These results suggest that strategic behaviour consistently leads to a heightened lobbying intensity, while the degree of discounting impacts the upper and lower bounds of the lobbying behaviour.

We conducted an initial study on the interplay between strategic behaviour of two players and Stackelberg behaviour of multiple selves induced by non-constant discounting in a zero-sum game. To facilitate a broader understanding of the important topic of dynamic games and time inconsistency, we intentionally simplified the model and focused on specific cases for our analysis. Additionally, we discussed the questions in the context of lobbying and employed a basic reduced form political economy model to maintain tractability and the possibility of closed-form solutions. However, there are numerous opportunities for expansion, including more extensive modelling of political economy elements. To the best of our knowledge, this paper is the first to apply dynamic stochastic differential game methods in political analysis, and future extensions could provide further insights into the dynamics of lobbying and other key interest concepts that shape the landscape of political economy.

It is worth noting that our framework's zero-sum property reduces the equilibrium description to one Riccati equation. However, in the interesting case where the players have heterogeneous discount functions, the zero-sum assumption falls apart. Similarly, from the vantage point of the lobbying model, this assumption also curtails the full spectrum of externalities one lobbyist might exert on another. Thus future studies can explore ways to overcome this limitation and expand the applicability of our approach to a broader range of nonzero-sum differential games. If possible, such a result would be a significant advance because it would allow to analyse strategic interactions between players with various degrees of behavioural biases.

Appendix

A.1 Proof of Proposition 3.1

By Yong and Zhou [27, Theorem 6.6.1], the Riccati equation associated with the problem of maximising the objective (3.1) under the state equation (3.2) reads

$$\dot{P}(s) + \sigma^2 P(s) - \rho P(s) - \frac{P(s)^2}{R} = 0, \qquad P(T) = 1. \tag{A.1}$$

Solving (A.1) gives

$$P(s) = \frac{R(\rho-\sigma^2)e^{(\rho-\sigma^2)s}}{(1+R(\rho-\sigma^2))e^{(\rho-\sigma^2)T} - e^{(\rho-\sigma^2)s}}, \qquad s \in [0,T].$$

By [27, Theorem 6.6.1] again, the unique optimal closed-loop strategy is given by

$$\hat{\Theta}_2(s) = -\frac{P(s)}{R} = -\frac{(\rho-\sigma^2)e^{(\rho-\sigma^2)s}}{(1+R(\rho-\sigma^2))e^{(\rho-\sigma^2)T} - e^{(\rho-\sigma^2)s}}, \qquad s \in [0,T].$$

The above expression of $\hat{\Theta}_2$ aligns with (3.3) and implies (3.4) and (3.5). □

A.2 Proof of Theorem 3.5

Denote the discount function of the precommitment model by

$$\alpha^{\Pi}(t,s) = \begin{cases} \alpha(s-t_{N-1}), & t \in [t_{N-1}, t_N], s \in [t,T], \\ \alpha(s-t_{k-1}), & t \in [t_{k-1}, t_k), s \in [t,T], k = 1,\dots,N-1. \end{cases}$$

Note that the function P^{Π} defined by (3.12) satisfies the equation

$$\begin{aligned} 0 &= P^{\Pi}_s(t,s) + 2P^{\Pi}(t,s)\Theta^{\Pi}_2(s) + P^{\Pi}(t,s)\sigma^2 + \alpha^{\Pi}(t,s)R\Theta^{\Pi}_2(s)^2 \\ &\quad \text{for } (t,s) \in \Delta[0,T], \\ P^{\Pi}(t,T) &= \alpha^{\Pi}(T-t), \qquad t \in [0,T]. \end{aligned} \tag{A.2}$$

From the fact that $\alpha(s-t) \le \alpha(s-t')$ for $0 \le t \le t' \le T$ and $s \in [t',T]$, we get

$$P^{\Pi}(t,s) \le P^{\Pi}(s,s), \qquad 0 \le t \le s \le T. \tag{A.3}$$

Let Ξ be the unique solution to the Lyapunov equation

$$\dot{\Xi}(s) + \Xi(s)\sigma^2 = 0, \qquad \Xi(T) = 1.$$

From (3.11), it is easily checked that

$$P^{\Pi}(t,s) = P(t_{N-1};s) \le \Xi(s) = e^{\sigma^2(T-s)}, \qquad t_{N-1} \le t \le s \le T.$$

Then by (A.3), we have

$$P^{\Pi}(t,s) \le \Xi(s), \qquad 0 \le t \le T, t \vee t_{N-1} \le s \le T.$$

In particular,

$$P^{\Pi}(t,t_{N-1}) \le \Xi(t_{N-1}), \qquad 0 \le t \le t_{N-1}.$$

From (3.12), we have

$$P^{\Pi}(t,s) = P(t_{N-2};s) \le e^{\sigma^2(t_{N-1}-s)} P(t_{N-2};t_{N-1})$$
$$\le e^{\sigma^2(t_{N-1}-s)} \Xi(t_{N-1}) = \Xi(s), \qquad t_{N-2} \le t < t_{N-1}, t \le s \le t_{N-1}.$$

It then follows from (A.3) that

$$P^{\Pi}(t,s) \le \Xi(s), \qquad 0 \le t < t_{N-1}, t \vee t_{N-2} \le s \le t_{N-1}.$$

Thus by induction, we have

$$P^{\Pi}(t,s) \le \Xi(s), \qquad 0 \le t \le T, t \vee t_{N-2} \le s \le T.$$

Continuing the above yields

$$P^{\Pi}(t,s) \le \Xi(s), \qquad (t,s) \in \Delta[0,T].$$

Thus, noting that $P^{\Pi} \ge 0$, we have that P^{Π} is uniformly bounded. Then from (3.10) and (A.2), we have for $(t,s) \in \Delta[0,T]$ that

$$|P(t,s) - P^{\Pi}(t,s)| \le K\|\Pi\|$$
$$+ \int_s^T \big(|P(t,r) - P^{\Pi}(t,r)| + |P(r,r) - P^{\Pi}(r,r)|\big) dr,$$

which implies that

$$|P(t,s) - P^{\Pi}(t,s)| \le K\|\Pi\|, \qquad (t,s) \in \Delta[0,T]. \qquad \square$$

A.3 Proof of Proposition 4.2

Using the fact that $G = 1$, $R_1 = -1$, and $R_2 = R$, we can rewrite (4.2) as

$$0 = \dot{P}(s) + \sigma^2 P(s) - \rho P(s) - \frac{1-R}{R} P(s)^2, \qquad P(T) = 1. \tag{A.4}$$

Note that $\frac{1-R}{R} \ge 0$. Then by Yong and Zhou [27, Theorem 6.7.2], the ODE (A.4) admits a unique solution P. Indeed, the solution is explicitly given by

$$P(s) = \frac{R(\rho - \sigma^2) e^{(\rho-\sigma^2)s}}{(1 - R + R(\rho - \sigma^2)) e^{(\rho-\sigma^2)T} - (1-R) e^{(\rho-\sigma^2)s}}, \qquad s \in [0,T].$$

Substituting the above into (4.3), we get the desired results immediately. $\square$

A.4 Proof of Theorem 5.2

By the variation of constants formula, we can see that the ERE (5.11) is essentially a Volterra integral equation. Notice that when the existence of a solution to the ERE

(5.11) holds, the map $t \mapsto P(t,s)$ is differentiable. To prove Theorem 5.2, we introduce the Volterra integro-differential equation (VIDE, for short)

$$\begin{aligned}
0 &= \dot{\Gamma}(t) + \Gamma(t)\sigma^2 - \frac{1-R}{R}\Gamma(t)^2 \\
&\quad - \int_t^T e^{\int_t^s (\sigma^2 + 2\frac{R-1}{R})\Gamma(r)dr} \partial_t \alpha(s-t) \frac{1-R}{R}\Gamma(s)^2 ds \\
&\quad - e^{\int_t^T (\sigma^2 + 2\frac{R-1}{R})\Gamma(r)dr} \partial_t \alpha(T-t), \\
\Gamma(T) &= 1.
\end{aligned} \tag{A.5}$$

If (A.5) has a solution, it is easily checked that

$$\begin{aligned}
&\frac{1-R}{R}\Gamma(t)^2 + \int_t^T e^{\int_t^s (\sigma^2 + 2\frac{R-1}{R})\Gamma(r)dr} \partial_t \alpha(s-t) \frac{1-R}{R}\Gamma(s)^2 ds \\
&+ e^{\int_t^T (\sigma^2 + 2\frac{R-1}{R})\Gamma(r)dr} \partial_t \alpha(T-t) \geq 0.
\end{aligned}$$

Thus by the comparison theorem of ODEs, we have the a priori estimate

$$\Gamma(s) \leq \Xi(s), \qquad s \in [0,T], \tag{A.6}$$

where Ξ is the unique solution to the Lyapunov equation

$$\dot{\Xi}(s) + \Xi(s)\sigma^2 = 0, \qquad \Xi(T) = 1. \tag{A.7}$$

Next, we give an a priori lower bound estimate for Γ. First, (A.5) can be rewritten as

$$\begin{aligned}
0 &= \dot{\Gamma}(t) + \left(\sigma^2 - 2\frac{1-R}{R}\Gamma(t)\right)\Gamma(t) + \frac{1-R}{R}\Gamma(t)^2 \\
&\quad - \int_t^T e^{\int_t^s (\sigma^2 - 2\frac{1-R}{R}\Gamma(r))dr} \partial_t \alpha(s-t) \frac{1-R}{R}\Gamma(s)^2 ds \\
&\quad - e^{\int_t^T (\sigma^2 - 2\frac{1-R}{R}\Gamma(r))dr} \partial_t \alpha(T-t), \\
\Gamma(T) &= 1.
\end{aligned}$$

Denote

$$\begin{aligned}
\widehat{\Gamma}(t) &= \int_t^T e^{\int_t^s (\sigma^2 - 2\frac{1-R}{R}\Gamma(r))dr} \alpha(s-t) \frac{1-R}{R}\Gamma(s)^2 ds \\
&\quad + e^{\int_t^T (\sigma^2 - 2\frac{1-R}{R}\Gamma(r))dr} \alpha(T-t), \qquad t \in [0,T].
\end{aligned}$$

Clearly,

$$\widehat{\Gamma}(t) \geq 0, \qquad t \in [0,T].$$

Note that

$$
\begin{aligned}
0 = \dot{\widehat{\Gamma}}(t) + \left(\sigma^2 - 2\frac{1-R}{R}\Gamma(t)\right)\widehat{\Gamma}(t) + \frac{1-R}{R}\Gamma(t)^2 \\
- \int_t^T e^{\int_t^s (\sigma^2 - 2\frac{1-R}{R}\Gamma(r))dr}\partial_t\alpha(s-t)\frac{1-R}{R}\Gamma(s)^2 ds \\
- e^{\int_t^T (\sigma^2 - 2\frac{1-R}{R}\Gamma(r))dr}\partial_t\alpha(T-t), \\
\widehat{\Gamma}(T) = 1.
\end{aligned}
$$

Then by the uniqueness of the solution to the above linear equation for $\widehat{\Gamma}$, we have

$$\Gamma(t) = \widehat{\Gamma}(t), \qquad t \in [0, T].$$

It follows that

$$\Gamma(t) \geq 0, \qquad t \in [0, T].$$

Combining this with (A.6), we get the a priori estimate

$$0 \leq \Gamma(s) \leq \Xi(s), \qquad s \in [0, T]. \tag{A.8}$$

With the above, we can prove the well-posedness of (A.5).

Lemma A.1 *The VIDE* (A.5) *admits a unique solution* Γ. *Moreover,* $\Gamma(t) \geq 0$ *for* $t \in [0, T]$.

Proof The uniqueness of solutions to (A.5) can be obtained by a standard method, and we only prove the existence of a solution. Let $M := \sup_{t\in[0,T]} |\Xi(t)|$. Let ρ_M be a smooth truncation function with

$$\rho_M(x) = \begin{cases} x, & |x| \leq M+1, \\ 0, & |x| \geq M+2. \end{cases}$$

Let Γ^M be the unique solution to the VIDE

$$
\begin{aligned}
0 = \dot{\Gamma}^M(t) + \Gamma^M(t)\sigma^2 - \frac{1-R}{R}\rho_M\left(\Gamma^M(t)\right)^2 \\
- \int_t^T e^{\int_t^s (\sigma^2 + 2\frac{R-1}{R})\rho_M(\Gamma^M(r))dr}\partial_t\alpha(s-t)\frac{1-R}{R}\rho_M\left(\Gamma^M(s)\right)^2 ds \\
- e^{\int_t^T (\sigma^2 + 2\frac{R-1}{R})\rho_M(\Gamma(r))dr}\partial_t\alpha(T-t), \\
\Gamma^M(T) = 1.
\end{aligned}
$$

Define

$$\tau = \max\{t \in [0, T] : |\Gamma^M(t)| \geq M+1\},$$

and $\tau = 0$ if $\{t \in [0, T] : |\Gamma^M(t)| \geq M+1\} = \emptyset$. Note that

$$|\Gamma^M(T)| = 1 \leq \sup_{t\in[0,T]} |\Xi(t)| < M+1.$$

Thus $\tau < T$. If $\tau = 0$, we have $\rho_M(\Gamma^M) = \Gamma^M$ which implies that Γ^M is a solution to (A.5). If $\tau > 0$, then $|\Gamma^M(\tau)| = M+1$ and Γ^M is a solution to (A.5) on $[\tau, T]$. However, by the a priori estimate (A.8), we have $|\Gamma^M(t)| \leq M$, $t \in [\tau, T]$, which contradicts the fact that $|\Gamma^M(\tau)| = M+1$. Thus $\tau = 0$ and the proof is complete. □

To complete the proof of Theorem 5.2, we now prove the existence of a solution to the ERE (5.11). Denote

$$\begin{aligned} P(t,s) = &\int_s^T e^{\int_s^\tau (\sigma^2 - 2\frac{1-R}{R}\Gamma(r))dr}\alpha(\tau - t)\frac{1-R}{R}\Gamma(\tau)^2 d\tau \\ &+ e^{\int_s^T (\sigma^2 - 2\frac{1-R}{R}\Gamma(r))dr}\alpha(T-t), \qquad 0 \leq t \leq s \leq T. \end{aligned} \tag{A.9}$$

Then

$$P(t,t) = \Gamma(t), \qquad t \in [0,T], \tag{A.10}$$

and

$$\begin{aligned} 0 = P_s(t,s) + P(t,s)\left(\sigma^2 - 2\frac{1-R}{R}\Gamma(s)\right) + \alpha(s-t)\frac{1-R}{R}\Gamma(s)^2 \\ \text{for } (t,s) \in \Delta[0,T], \\ P(t,T) = \alpha(T-t), \qquad t \in [0,T]. \end{aligned} \tag{A.11}$$

Combining (A.10) and (A.11), we can see that the function P defined by (A.9) is a solution to (5.11). If $R = 1$, then (5.11) reads

$$\begin{aligned} 0 = P_s(t,s) + P(t,s)\sigma^2, \qquad (t,s) \in \Delta[0,T], \\ P(t,T) = \alpha(T-t), \qquad t \in [0,T]. \end{aligned}$$

Taking t as a parameter, we get (5.13) immediately by the variation of constants formula. This completes the proof. □

A.5 Proof of Theorem 5.3

We consider a closed-loop strategy $\bar{\Theta}_i$, $i = 1, 2$ satisfying (5.11) and (5.12). We only prove the first inequality in (5.1) since the second can be obtained similarly. We fix $t \in [0, T)$, a constant $\varepsilon > 0$ and controls u_1 and u_2 from $L^2_{\mathbb{F}}(t, T)$. We consider the controls $\Theta_1^\varepsilon X^\varepsilon$ and $\Theta_2^\varepsilon X^\varepsilon$, where $\Theta_i^\varepsilon X^\varepsilon$ is defined by (5.2) for $i = 1, 2$. The corresponding state equation (2.1) becomes

$$\begin{aligned} dX^\varepsilon(s) &= \big(u_1(s) + u_2(s)\big)ds + \sigma X^\varepsilon(s)dW(s), \qquad s \in [t, t+\varepsilon), \\ dX^\varepsilon(s) &= \big(\bar{\Theta}_1(s)X^\varepsilon(s) + \bar{\Theta}_2(s)X^\varepsilon(s)\big)ds + \sigma X^\varepsilon(s)dW(s), \qquad s \in [t+\varepsilon, T], \\ X^\varepsilon(t) &= \bar{X}(t), \end{aligned} \tag{A.12}$$

and the corresponding utility functional

$$J\big(t,\bar{X}(t);\Theta_1^\varepsilon X^\varepsilon,\Theta_2^\varepsilon X^\varepsilon\big):=J_1\big(t,\bar{X}(t);\Theta_1^\varepsilon X^\varepsilon,\Theta_2^\varepsilon X^\varepsilon\big),$$

with J_1 defined by (2.2), is given by

$$\begin{aligned}
&J\big(t,\bar{X}(t);\Theta_1^\varepsilon X^\varepsilon,\Theta_2^\varepsilon X^\varepsilon\big)\\
&=\mathbb{E}_t\bigg[\int_t^{t+\varepsilon}\alpha(s-t)\big(-u_1(s)^2+Ru_2(s)^2\big)ds\bigg]\\
&\quad+\mathbb{E}_t\bigg[\alpha(T-t)X^\varepsilon(T)^2\\
&\qquad+\int_{t+\varepsilon}^T\alpha(s-t)\Big(-\big(\bar{\Theta}_1(s)X^\varepsilon(s)\big)^2+R\big(\bar{\Theta}_2(s)X^\varepsilon(s)\big)^2\Big)ds\bigg].
\end{aligned}\tag{A.13}$$

Applying the Itô formula to the mapping $s\mapsto P(t,s)X^\varepsilon(s)^2$ over $[t+\varepsilon,T]$ gives

$$\begin{aligned}
&P(t,T)X^\varepsilon(T)^2\\
&=P(t,t+\varepsilon)X^\varepsilon(t+\varepsilon)^2\\
&\quad+\int_{t+\varepsilon}^T\Big(P_s(t,r)X^\varepsilon(r)^2+2P(t,r)X^\varepsilon(r)\big(\bar{\Theta}_1(r)X^\varepsilon(r)+\bar{\Theta}_2(r)X^\varepsilon(r)\big)\\
&\qquad\qquad+P(t,r)\big(\sigma X^\varepsilon(r)\big)^2\Big)dr\\
&\quad+\int_{t+\varepsilon}^T2P(t,r)X^\varepsilon(r)\sigma X^\varepsilon(r)dW(r)\\
&=P(t,t+\varepsilon)X^\varepsilon(t+\varepsilon)^2\\
&\quad+\int_{t+\varepsilon}^T\Big(P_s(t,r)+2P(t,r)\big(\bar{\Theta}_1(r)+\bar{\Theta}_2(r)\big)+P(t,r)\sigma^2\Big)X^\varepsilon(r)^2dr\\
&\quad+\int_{t+\varepsilon}^T2P(t,r)X^\varepsilon(r)\sigma X^\varepsilon(r)dW(r).
\end{aligned}$$

Since the function P satisfies (5.11), the above can be rewritten as

$$\begin{aligned}
P(t,T)X^\varepsilon(T)^2&=P(t,t+\varepsilon)X^\varepsilon(t+\varepsilon)^2\\
&\quad-\int_{t+\varepsilon}^T\alpha(t,r)\big(-\bar{\Theta}_1(r)^2+R\bar{\Theta}_2(r)^2\big)X^\varepsilon(r)^2dr\\
&\quad+\int_{t+\varepsilon}^T2P(t,r)X^\varepsilon(r)\sigma X^\varepsilon(r)dW(r),
\end{aligned}$$

which yields

$$\mathbb{E}_t\Big[P(t,T)X^\varepsilon(T)^2+\int_{t+\varepsilon}^{T}\alpha(t,r)\big(-\bar\Theta_1(r)^2+R\bar\Theta_2(r)^2\big)X^\varepsilon(r)^2dr\Big]$$
$$=\mathbb{E}_t[P(t,t+\varepsilon)X^\varepsilon(t+\varepsilon)^2].$$

Substituting the above into (A.13), we get

$$J\big(t,\bar X(t);\Theta_1^\varepsilon X^\varepsilon,\Theta_2^\varepsilon X^\varepsilon\big)$$
$$=\mathbb{E}_t\Big[\int_t^{t+\varepsilon}\alpha(s-t)\big(-u_1(s)^2+Ru_2(s)^2\big)ds+P(t,t+\varepsilon)X^\varepsilon(t+\varepsilon)^2\Big]$$
$$=:J\big(t,\bar X(t);u_1,u_2\big)\big|_{[t,t+\varepsilon]}. \tag{A.14}$$

Thus for fixed t, the game with state equation (A.12) and utility functional (A.13) over the time interval $[t,T]$ can be thought of as a game problem over $[t,t+\varepsilon]$. In that game, the state is given by the first line in (A.12), the utility functional is given by the representation $J(t,\bar X(t);u_1,u_2)|_{[t,t+\varepsilon]}$ in (A.14) and the control u_i is taken from $L^2_{\mathbb{F}}(t,t+\varepsilon)$. Now we introduce over $[\tau,t+\varepsilon]\subseteq[t,t+\varepsilon]$ the dynamic game problem with the state

$$dX^\varepsilon(s)=\big(u_1(s)+u_2(s)\big)ds+\sigma X^\varepsilon(s)dW(s),\qquad s\in[\tau,t+\varepsilon),$$
$$X^\varepsilon(\tau)=\xi, \tag{A.15}$$

and utility functional

$$J(\tau,\xi;u_1,u_2)|_{[t,t+\varepsilon]}$$
$$=\mathbb{E}_\tau\Big[\int_\tau^{t+\varepsilon}\alpha(s-t)\big(-u_1(s)^2+Ru_2(s)^2\big)ds+P(t,t+\varepsilon)X^\varepsilon(t+\varepsilon)^2\Big], \tag{A.16}$$

where $(\tau,\xi)\in[t,t+\varepsilon]\times L^2_{\mathcal{F}_\tau}(\Omega)$ is the initial pair. We remark that in (A.16), the time t has been fixed and only plays the role of a parameter. Thus by Lemma 4.1 with $T=t+\varepsilon$, $\rho=0$, $G=P(t,t+\varepsilon)$, $R_1(s)=\alpha(s-t)$ and $R_2(s)=\alpha(s-t)R$, the unique closed-loop saddle strategy associated with (A.15) and (A.16) can be found. Note that if we take the initial pair $(\tau,\xi)=(t,\bar X(t))$, (A.15) and (A.16) exactly coincide with the first line in (A.12) and the functional $J(t,\bar X(t);u_1,u_2)|_{[t,t+\varepsilon]}$ in (A.14), respectively. Thus the unique closed-loop saddle strategy associated with (A.12) and (A.14) can be found, and it is given by

$$\Theta_1^{*,\varepsilon}(s)=P^t(s),\quad \Theta_2^{*,\varepsilon}(s)=-\frac{P^t(s)}{R},\qquad s\in[t,t+\varepsilon],$$

where P^t is the unique solution to the Riccati equation

$$0=\dot P^t(s)+P^t(s)\sigma^2-\frac{1-R}{\alpha(s-t)R}P^t(s)^2=0,\qquad s\in[t,t+\varepsilon],$$
$$P^t(t+\varepsilon)=P(t,t+\varepsilon). \tag{A.17}$$

In other words, for any $(u_1, u_2) \in L^2_{\mathbb{F}}(t, t+\varepsilon) \times L^2_{\mathbb{F}}(t, t+\varepsilon)$, we have

$$\begin{aligned} J\big(t, \bar{X}(t); \Theta_1^{*,\varepsilon} X^{*,\varepsilon}, \Theta_2^{*,\varepsilon} X^{*,\varepsilon}\big)\big|_{[t,t+\varepsilon]} &\geq J\big(t, \bar{X}(t); u_1, \Theta_2^{*,\varepsilon} X^{1,\varepsilon}\big)\big|_{[t,t+\varepsilon]}, \\ J\big(t, \bar{X}(t); \Theta_1^{*,\varepsilon} X^{*,\varepsilon}, \Theta_2^{*,\varepsilon} X^{*,\varepsilon}\big)\big|_{[t,t+\varepsilon]} &\leq J\big(t, \bar{X}(t); \Theta_1^{*,\varepsilon} X^{2,\varepsilon}, u_2\big)\big|_{[t,t+\varepsilon]}, \end{aligned} \tag{A.18}$$

where

$$\begin{aligned} dX^{*,\varepsilon}(s) &= \big(P^t(s)X^{*,\varepsilon}(s) - R^{-1}P^t(s)X^{*,\varepsilon}(s)\big)ds + \sigma X^{*,\varepsilon}(s)dW(s), \\ X^{*,\varepsilon}(t) &= \bar{X}(t), \\ dX^{1,\varepsilon}(s) &= \big(u_1(s) - R^{-1}P^t(s)X^{1,\varepsilon}(s)\big)ds + \sigma X^{1,\varepsilon}(s)dW(s), \\ X^{1,\varepsilon}(t) &= \bar{X}(t), \\ dX^{2,\varepsilon}(s) &= \big(P^t(s)X^{2,\varepsilon}(s) + u^2(s)\big)ds + \sigma X^{2,\varepsilon}(s)dW(s), \\ X^{2,\varepsilon}(t) &= \bar{X}(t). \end{aligned} \tag{A.19}$$

By applying the Itô formula to the mapping $s \mapsto P^t(s)X^{*,\varepsilon}(s)^2$ over $[t, t+\varepsilon]$ and to the mapping $s \mapsto P(t,s)\bar{X}(s)^2$ over $[t, T]$, we have

$$J\big(t, \bar{X}(t); \Theta_1^{*,\varepsilon} X^{*,\varepsilon}, \Theta_2^{*,\varepsilon} X^{*,\varepsilon}\big)\big|_{[t,t+\varepsilon]} = P^t(t)\bar{X}(t)^2 \tag{A.20}$$

and

$$J\big(t, \bar{X}(t); \bar{\Theta}_1\bar{X}, \bar{\Theta}_2\bar{X}\big) = P(t,t)\bar{X}(t)^2. \tag{A.21}$$

Since $P^t(t+\varepsilon) = P(t, t+\varepsilon)$ and the mappings $s \mapsto P^t(s)$ and $(t,s) \mapsto P(t,s)$ are differentiable with uniformly bounded derivatives, we have

$$\begin{aligned} |P^t(s) - P(t,s)| &\leq |P^t(s) - P(t,t+\varepsilon)| + |P(t,t+\varepsilon) - P(t,s)| \leq K\varepsilon, \\ |P^t(s) - P(s,s)| &\leq |P^t(s) - P(t,s)| + |P(t,s) - P(s,s)| \leq K\varepsilon. \end{aligned} \tag{A.22}$$

Now recall that P^t and P satisfy (A.17) and (5.11), respectively. Noting that we have $P^t(t+\varepsilon) = P(t, t+\varepsilon)$, we get by applying the stability estimate of ODEs to (A.17) and (5.11) over $[t, t+\varepsilon]$ that

$$|P^t(t) - P(t,t)| \leq K\int_t^{t+\varepsilon} \big(|P^t(s) - P(t,s)| + |P^t(s) - P(s,s)|\big)ds \leq K\varepsilon^2.$$

Thus, comparing (A.20) and (A.21), we have

$$\big|J\big(t, \bar{X}(t); \Theta_1^{*,\varepsilon} X^{*,\varepsilon}, \Theta_2^{*,\varepsilon} X^{*,\varepsilon}\big)\big|_{[t,t+\varepsilon]} - J\big(t, \bar{X}(t); \bar{\Theta}_1\bar{X}, \bar{\Theta}_2\bar{X}\big)\big| \leq K\varepsilon^2\bar{X}(t)^2.$$

It follows from (A.18) that

$$J\big(t, \bar{X}(t); u_1, \Theta_2^{*,\varepsilon} X^{1,\varepsilon}\big)\big|_{[t,t+\varepsilon]} \leq J\big(t, \bar{X}(t); \bar{\Theta}_1\bar{X}, \bar{\Theta}_2\bar{X}\big) + K\varepsilon^2\bar{X}(t)^2. \tag{A.23}$$

Thus to prove that the first inequality in (5.1) holds, it suffices to prove that

$$\begin{aligned}&\Big|J\big(t,\bar{X}(t);u_1,\Theta_2^{*,\varepsilon}X^{1,\varepsilon}\big)\big|_{[t,t+\varepsilon]}-J\big(t,\bar{X}(t);\Theta_1^{\varepsilon}X^{\varepsilon},\bar{\Theta}_2X^{\varepsilon}\big)\Big|\\&\leq K\varepsilon^2\mathbb{E}_t\bigg[\bar{X}(t)^2+\int_t^{t+\varepsilon}|u_1(s)|^2ds\bigg]^{\frac{1}{2}}\end{aligned}\tag{A.24}$$

for some $K>0$. Note that

$$\begin{aligned}&J\big(t,\bar{X}(t);u_1,\Theta_2^{*,\varepsilon}X^{1,\varepsilon}\big)\big|_{[t,t+\varepsilon]}\\&=\mathbb{E}_t\bigg[P(t,t+\varepsilon)X^{1,\varepsilon}(t+\varepsilon)^2\\&\qquad+\int_t^{t+\varepsilon}\alpha(s-t)\big(-u_1(s)^2+R^{-1}P^t(s)^2X^{1,\varepsilon}(s)^2\big)ds\bigg],\end{aligned}\tag{A.25}$$

where $X^{1,\varepsilon}$ is the unique solution to (A.19), and

$$\begin{aligned}&J\big(t,\bar{X}(t);\Theta_1^{\varepsilon}X^{\varepsilon},\bar{\Theta}_2X^{\varepsilon}\big)\\&=\mathbb{E}_t\bigg[P(t,t+\varepsilon)X^{\varepsilon}(t+\varepsilon)^2\\&\qquad+\int_t^{t+\varepsilon}\alpha(s-t)\big(-u_1(s)^2+R^{-1}P(s,s)^2X^{\varepsilon}(s)^2\big)ds\bigg],\end{aligned}\tag{A.26}$$

where X^{ε} is uniquely determined by

$$\begin{aligned}&dX^{\varepsilon}(s)=\big(u_1(s)-R^{-1}P(s,s)X^{\varepsilon}(s)\big)ds+\sigma X^{\varepsilon}(s)dW(s),\qquad s\in[t,t+\varepsilon],\\&X^{\varepsilon}(t)=\bar{X}(t).\end{aligned}\tag{A.27}$$

Recall that $X^{1,\varepsilon}$ and X^{ε} are the unique solutions of (A.19) and (A.27), respectively. By applying the stability estimates for SDEs to (A.19) and (A.27), we have

$$\mathbb{E}_t\Big[\sup_{s\in[t,t+\varepsilon]}|X^{\varepsilon}(s)-X^{1,\varepsilon}(s)|^2\Big]\leq K\mathbb{E}_t\bigg[\int_t^{t+\varepsilon}|P^t(s)-P(s,s)||X^{\varepsilon}(s)|ds\bigg]^2.$$

Substituting (A.22) into the above, we get

$$\begin{aligned}\mathbb{E}_t\Big[\sup_{s\in[t,t+\varepsilon]}|X^{\varepsilon}(s)-X^{1,\varepsilon}(s)|^2\Big]&\leq K\varepsilon^2\mathbb{E}_t\bigg[\int_t^{t+\varepsilon}|X^{\varepsilon}(s)|ds\bigg]^2\\&\leq K\varepsilon^4\mathbb{E}_t\Big[\sup_{s\in[t,t+\varepsilon]}|X^{\varepsilon}(s)|^2\Big]\\&\leq K\varepsilon^4\mathbb{E}_t\bigg[\bar{X}(t)^2+\int_t^{t+\varepsilon}|u_1(s)|^2ds\bigg],\end{aligned}\tag{A.28}$$

where the last inequality is due to a standard estimate of SDEs for X^ε. Using (A.28), we obtain

$$\begin{aligned}
&\mathbb{E}_t\Big[\sup_{s\in[t,t+\varepsilon]}|X^\varepsilon(s)^2-X^{1,\varepsilon}(s)^2|\Big]\\
&\le \mathbb{E}_t\Big[\sup_{s\in[t,t+\varepsilon]}|X^\varepsilon(s)+X^{1,\varepsilon}(s)|^2\Big]^{\frac{1}{2}}\mathbb{E}_t\Big[\sup_{s\in[t,t+\varepsilon]}|X^\varepsilon(s)-X^{1,\varepsilon}(s)|^2\Big]^{\frac{1}{2}}\\
&\le K\varepsilon^2\mathbb{E}_t\Big[\sup_{s\in[t,t+\varepsilon]}|X^\varepsilon(s)+X^{1,\varepsilon}(s)|^2\Big]^{\frac{1}{2}}\mathbb{E}_t\Big[\bar{X}(t)^2+\int_t^{t+\varepsilon}|u_1(s)|^2ds\Big]^{\frac{1}{2}}\\
&\le K\varepsilon^2\mathbb{E}_t\Big[\sup_{s\in[t,t+\varepsilon]}|X^\varepsilon(s)|^2+\sup_{s\in[t,t+\varepsilon]}|X^{1,\varepsilon}(s)|^2\Big]^{\frac{1}{2}}\\
&\quad\times\mathbb{E}_t\Big[\bar{X}(t)^2+\int_t^{t+\varepsilon}|u_1(s)|^2ds\Big]^{\frac{1}{2}}\\
&\le K\varepsilon^2\mathbb{E}_t\Big[\bar{X}(t)^2+\int_t^{t+\varepsilon}|u_1(s)|^2ds\Big],
\end{aligned}\tag{A.29}$$

where the last inequality is obtained by standard estimates of SDEs for X^ε and $X^{1,\varepsilon}$. Comparing (A.25) and (A.26), by the estimates (A.22) and (A.29), we get (A.24). Combining this with (A.23), we have

$$\begin{aligned}
J\big(t,\bar{X}(t);\Theta_1^\varepsilon X^\varepsilon,\bar{\Theta}_2X^\varepsilon\big)\le J\big(t,\bar{X}(t);\bar{\Theta}_1\bar{X},\bar{\Theta}_2\bar{X}\big)\\
+K\varepsilon^2\mathbb{E}_t\Big[\bar{X}(t)^2+\int_t^{t+\varepsilon}|u_1(s)|^2ds\Big].
\end{aligned}$$

This implies the first inequality in (5.1), as desired. □

A.6 Proof of Theorem 5.4

This proof is very similar to that of Theorem 3.5. Note that $\alpha(s-t)\le\alpha(s-t')$ for any $0\le t\le t'\le T$ and $s\in[t',T]$. Together with the fact that $-\bar{\Theta}_1(s)^2+R\bar{\Theta}_2(s)^2\ge 0$, we get from (5.10) that

$$P^\Pi(t,s)\le P^\Pi(s,s),\qquad 0\le t\le s\le T.\tag{A.30}$$

Recall (5.5). By the comparison theorem of ODEs, we have

$$P^\Pi(t,s)=P(t_{N-1};s)\le\Xi(s),\qquad t_{N-1}\le t\le s\le T,$$

where Ξ is the unique solution to (A.7). Then by (A.30), we have

$$P^\Pi(t,s)\le\Xi(s),\qquad 0\le t\le T,t\vee t_{N-1}\le s\le T.\tag{A.31}$$

In particular,

$$P^\Pi(t,t_{N-1})\le\Xi(t_{N-1}),\qquad 0\le t\le t_{N-1}.$$

Recall (5.7). By the comparison theorem again, we have

$$P^{\Pi}(t,s) = P(t_{N-2};s) \le \Xi(s), \qquad t_{N-2} \le t < t_{N-1}, t \le s \le t_{N-1}.$$

By (A.30), we have

$$P^{\Pi}(t,s) \le \Xi(s), \qquad 0 \le t < t_{N-1}, t \vee t_{N-2} \le s \le t_{N-1}. \tag{A.32}$$

Combining (A.31) and (A.32), we get

$$P^{\Pi}(t,s) \le \Xi(s), \qquad 0 \le t \le T, t \vee t_{N-2} \le s \le T.$$

By continuing the above, we obtain

$$P^{\Pi}(t,s) \le \Xi(s), \qquad (t,s) \in \Delta[0,T].$$

Thus, noting that $P^{\Pi}(t,s) \ge 0$, we see that P^{Π} is uniformly bounded. Then from (5.10) and (5.11), we have for all $(t,s) \in \Delta[0,T]$ that

$$|P(t,s) - P^{\Pi}(t,s)| \le K\|\Pi\| + \int_s^T \big(|P(t,r) - P^{\Pi}(t,r)| + |P(r,r) - P^{\Pi}(r,r)|\big)dr,$$

which implies that

$$|P(t,s) - P^{\Pi}(t,s)| \le K\|\Pi\|, \qquad (t,s) \in \Delta[0,T].$$

The desired results can now be directly obtained. □

Acknowledgements We extend our gratitude to the Editor Martin Schweizer, an associate editor and an anonymous referee for their valuable feedback which significantly enhanced the quality of this paper. Additionally, we express our appreciation to Dongliang Lu for providing exceptional research assistance.

Declarations

Competing Interests The authors declare no competing interests.

References

1. Ainslie, G.: Specious reward: a behavioral theory of impulsiveness and impulse control. Psychol. Bull. **82**, 463–496 (1975)
2. Bernheim, B.: Intergenerational altruism, dynastic equilibria and social welfare. Rev. Econ. Stud. **56**, 119–128 (1989)
3. Björk, T., Khapko, M., Murgoci, A.: On time-inconsistent stochastic control in continuous time. Finance Stoch. **21**, 331–360 (2017)
4. Björk, T., Khapko, M., Murgoci, A.: Time-Inconsistent Control Theory with Finance Applications. Springer, Berlin (2021)
5. Björk, T., Murgoci, A.: A theory of Markovian time-inconsistent stochastic control in discrete time. Finance Stoch. **18**, 545–592 (2014)
6. Blanchard, O.J.: Debt, deficits, and finite horizons. J. Polit. Econ. **93**, 223–247 (1985)
7. Calvo, G.A., Obstfeld, M.: Optimal time-consistent fiscal policy with finite lifetimes. Econometrica **56**, 411–432 (1988)

8. Cetemen, D., Feng, F.Z., Urgun, C.: Renegotiation and dynamic inconsistency: contracting with non-exponential discounting. J. Econ. Theory **208**, 105606, 1–49 (2023)
9. Czichowsky, C.: Time-consistent mean–variance portfolio selection in discrete and continuous time. Finance Stoch. **17**, 227–271 (2013)
10. Ekeland, I., Lazrak, A.: Being serious about non-commitment: subgame perfect equilibrium in continuous time (2006). math/0604264
11. Ekeland, I., Lazrak, A.: The golden rule when preferences are time inconsistent. Math. Financ. Econ. **4**, 29–55 (2010)
12. Ekeland, I., Pirvu, T.A.: Investment and consumption without commitment. Math. Financ. Econ. **2**, 57–86 (2008)
13. Harris, C., Laibson, D.: Instantaneous gratification. Q. J. Econ. **128**, 205–248 (2013)
14. Hernández, C., Possamaï, D.: Me, myself and I: a general theory of non-Markovian time-inconsistent stochastic control for sophisticated agents. Ann. Appl. Probab. **33**, 1–47 (2023)
15. Hernández, C., Possamaï, D.: Time-inconsistent contract theory (2023). 2303.01601
16. Hu, Y., Jin, H., Zhou, X.Y.: Time-inconsistent stochastic linear–quadratic control. SIAM J. Control Optim. **50**, 1548–1572 (2012)
17. Jackson, M.O., Yariv, L.: Collective dynamic choice: the necessity of time inconsistency. Am. Econ. J. Microecon. **7**, 150–178 (2015)
18. Krusell, P., Smith, A.A.: Consumption-savings decisions with quasigeometric discounting. Econometrica **71**, 365–375 (2003)
19. Kydland, F.E., Prescott, E.C.: Rules rather than discretion: the inconsistency of optimal plans. J. Polit. Econ. **85**, 473–491 (1977)
20. Lei, Q., Pun, C.S.: Nonlocal fully nonlinear parabolic differential equations arising in time-inconsistent problems. J. Differ. Equ. **358**, 339–385 (2023)
21. Strotz, R.H.: Myopia and inconsistency in dynamic utility maximization. Rev. Econ. Stud. **23**, 165–180 (1956)
22. Sun, J., Yong, J.: Linear quadratic stochastic differential games: open-loop and closed-loop saddle points. SIAM J. Control Optim. **52**, 4082–4121 (2014)
23. Sun, J., Yong, J.: Stochastic Linear-Quadratic Optimal Control Theory: Differential Games and Mean-Field Problems. Springer, Berlin (2020)
24. Yong, J.: Time-inconsistent optimal control problems and the equilibrium HJB equation. Math. Control Relat. Fields **2**, 271–329 (2012)
25. Yong, J.: Time-inconsistent optimal control problems. In: Proceedings of 2014 ICM, Sect. 16, Control Theory and Optimization, pp. 947–969 (2014). Version available online at https://dornsife.usc.edu/assets/sites/350/docs/Prof._Jiongmin_Yong_slides.pdf
26. Yong, J.: Linear–quadratic optimal control problems for mean-field stochastic differential equations – time-consistent solutions. Trans. Am. Math. Soc. **369**, 5467–5523 (2017)
27. Yong, J., Zhou, X.Y.: Stochastic Controls: Hamiltonian Systems and HJB Equations. Springer, Berlin (1999)