Chapter 1

BASIC CONCEPTS

1.1. What Are Dynamic Games?

Let us begin with a preliminary characterization of dynamic games. A game involves two or more players. A player can be any decision-making entity, such as a government, a political party, a firm, a regulatory agency, an international organization, a household, or an individual. (I will use the pronoun "it" to refer to a generic player, even though in some contexts, the pronouns "he" or "she" may be more appropriate). A dynamic game extends over a time horizon (finite or infinite) and normally displays the following properties:

- the players may receive payoffs in every period (or at every point of time);
- the overall payoff for a player is the sum (or integral) of its discounted payoffs over the time horizon, possibly plus some terminal payments;
- the payoff that a player receives in a period may depend on both the actions taken in that period and the "state of the system" in that period, as represented by one or several "state variables";
- the state of the system changes over time, and the rate of change of the state variables may depend on the actions of the players, as represented by their "control variables"; and
- the rate of change of a state variable is described by a difference equation or a differential equation, often called the "transition equation" or "dynamic equation".

Thus, I will exclude from consideration "repeated games" (such as the repeated prisoners' dilemma etc.) since in such games there are no state variables and no transition equations that describe the changing environment in which the players operate.

Economists have used dynamic games to analyze a variety of problems in various fields, such as dynamic oligopoly, dynamic contributions to a public good, dynamic game of optimal tariffs and retaliation, redistributive taxation in the presence of forward-looking agents, exploitation of common property resources, non-cooperative environmental policies, and the arms race. In this survey, I intend to introduce the readers to the basic equilibrium concepts in dynamic games and present a number of interesting dynamic game models and results in various fields of economics.

This chapter introduces the basic concepts and some ideas about solution techniques. In Sec. 1.2, I introduce two main equilibrium concepts, open-loop Nash equilibrium (OLNE) and Markov-perfect Nash equilibrium (MPNE), and illustrate their difference by means of simple examples. In Sec. 1.3, I introduce the concept of hierarchical dynamic games and two equilibrium concepts for such games, the open-loop Stackelberg equilibrium (OLSE) and the feedback Stackelberg equilibrium (FBSE).

1.2. Open-loop Nash Equilibirum and Markov-perfect Nash Equilibrium

One of the most important distinctions in dynamic games is that between "open-loop strategies" (or pre-commitment strategies) and "Markov-perfect strategies" (or feedback strategies).[1] A player's open-loop strategy is a planned time path of its actions.

An OLNE is a profile of open-loop strategies (one for each player) such that each player's open-loop strategy maximizes its payoff, given the open-loop strategies of other players. Some early articles analyze dynamic games using exclusively this equilibrium concept. See, for example, Clark and Munro (1975), Salant (1976), Dasgupta and Heal (1979, Chap. 12), Spence (1979, 1981), Flaherty (1980a,b), Crawford *et al.* (1980), Reinganum (1981a, 1982a), Fudenberg and Tirole (1983), Lewis and Schmalensee (1980), Chiarella *et al.* (1984) and Loury (1986).[2] More recent articles that focus exclusively on OLNEs include Gaudet and Long (1994, 2003), Lambertini and Rossini (1998), Sorger (2002), Benchekroun *et al.* (2009, 2010), and Fujiwara and Long (2010).

[1] I use the terms feedback strategy and Markov-perfect strategy interchangeably. Similarly, a feedback Nash equlibrium (FBNE) and a Markov-perfect Nash equilibrium (MPNE) are alternative names for the same concept.

[2] It turns out that in the case of Reinganum (1981a, 1982a), the OLNE is also the FBNE. Basically, this is because these models are linear in the state variable (after a suitable transformation of variables).

OLNEs are time consistent: along the equilibrium path, no player has at any stage any incentive to deviate from its original plan.[3] However, if perhaps by error someone has deviated from the equilibrium path, then the state of the system is revealed to be different from what was predicted, and it will be no longer optimal for a player to continue with its originally planned time path of actions. Thus, an OLNE is not robust to perturbations. In contrast, an MPNE, which consists of Markov-perfect strategies that are best replies to each other, is robust to deviations.[4] A Markov-perfect strategy is a rule that conditions action at any date on the observed state of the system at that date, such that the objective function of the player, starting from any (date, state) pair, is maximized, given the Markov-perfect strategies of the other players.[5] This equilibrium concept corresponds to the idea of subgame perfection.[6]

An OLNE is founded on the assumption that each players has the ability to make a credible precommitment. (This will become clearer in the examples that follow.) The emerging consensus is that in analyzing dynamic games, one should try where possible to find an equilibrium in Markov-perfect strategies.

Reinganum and Stokey (1985) offer an intuitive, nontechnical, explanation of the distinction between open-loop strategies and Markov-perfect strategies. They use the expressions "path strategies" and "decision rule strategies".[7] If agents use path strategies, at the initial date each player must make a binding commitment about the actions it will take at all future dates. In contrast, when a player uses a decision rule, the action at any future date t will depend on the observed value of the state variable at t, denoted by $S(t)$. Decision rules that specify action at time t as a function of the observed pair $(t, S(t))$ only are called Markovian decision rules.[8] As Reinganum and Stokey (1985) point out, a Nash equilibrium in decision rules is not necessarily Markov-perfect. To be Markov-perfect, a Nash equilibrium in decision rules must satisfy the additional property that the continuation of the given decision rules constitutes a Nash equilibrium

[3]For a discussion of this point, see Dockner *et al.* (2000).
[4]For a brief discussion of Markov-perfect strategies, see Maskin and Tirole (1988a, p. 553).
[5]For a precise definition, see Dockner *et al.*, 2000.
[6]Of course, in continuous time, the concept of a subgame is problematic.
[7]See also Kydland (1975).
[8]Trigger strategies are examples of decision rules that are not Markovian. See Dockner *et al.* (2000) for a discussion of trigger strategies.

when viewed from *any* future (date, state) pair. Dockner *et al.* (2000, Ex. 4.2) give an example of a Nash equilibrium in decision rules that fails to be Markov-perfect.

Economic models that make use of the equilibrium in feedback strategies are increasingly popular. Beginning with Clemhout *et al.* (1973), Simaan and Cruz (1975), and Levhari and Mirman (1980), more and more dynamic game models focus exclusively on this equilibrium concept.

The economic profession's increasing preference of MPNE over OLNE does not mean that the latter is a useless concept. It is often quite useful to characterize an OLNE and compare it with an MPNE. As Fudenberg and Levine (1988) and Fudenberg and Tirole (2000) point out, an OLNE provides a useful benchmark for understanding the added strategic effect of Markov-perfect strategies.

1.2.1. *A simple transboundary pollution game in discrete time*

To illustrate the difference between OLNE and MPNE, let us consider a simple example: a two-period model of transboundary pollution.

There are two players (say two countries), called SMALL and CAP. Player SMALL chooses its level of CO_2 emissions for periods 1 and 2, denoted by the lowercase symbols x and y. Player CAP's levels of emissions for period 1 and 2 are denoted by the capital letters X and Y. Let S_t be the global stock of CO_2 concentration in the atmosphere at the beginning of period t. The initial stock (at the beginning of period 1) is denoted by S_1. Assume that the stock at the beginning of period 2 is determined by

$$S_2 = S_1 + x + X$$

and that the stock at the end of period 2 (which is the beginning of period 3) is

$$S_3 = S_2 + y + Y.$$

The variable S_t is called the state variable of the system. The control variables in this model are the emission levels.

The period 2 payoff to SMALL is the benefit from consuming the output, $ay - (1/2)y^2$, minus the cost of period 2 environmental damages, $(1/2)S_2^2$:

$$u_2 = ay - \frac{1}{2}y^2 - \frac{1}{2}S_2^2,$$

where a is a positive constant. Similarly, CAP's payoff in period 2 is

$$U_2 = AY - \frac{1}{2}Y^2 - \frac{1}{2}S_2^2,$$

where A is a positive constant.

We assume SMALL's payoff in period 1 is

$$u_1 = ax - \frac{1}{2}x^2 - \frac{1}{2}S_1^2$$

and that of CAP is

$$U_1 = AX - \frac{1}{2}X^2 - \frac{1}{2}S_1^2.$$

The objective of SMALL is to maximize its overall payoff w, defined as the sum of its period 1 and period 2 payoffs, minus a term that reflects its guilt of passing on to the next generation the stock of pollution S_3. We denote this term by $g_3(S_3)$ and assume it is increasing in S_3. For simplicity, let $g_3(S_3)$ take the simple form $(1/2)S_3^2$. Thus, $w \equiv u_1 + u_2 - g_3 = u_1 + u_2 - (1/2)S_3^2$. Similarly, CAP's guilt function is $G_3(S_3) = (1/2)S_3^2$ and it wants to maximize $W \equiv U_1 + U_2 - G_3 = U_1 + U_2 - (1/2)S_3^2$. Let us find the OLNE for this game.

Finding the OLNE

An open-loop strategy of a player is a planned time path of actions over the time horizon. Each player assumes that the other player is going to carry out its planned course of actions.

SMALL, taking as given the time path of actions of CAP denoted by $\{X, Y\}$, chooses its time path $\{x, y\}$ to maximize w:

$$w = ax - \frac{1}{2}x^2 - \frac{1}{2}S_1^2 + ay - \frac{1}{2}y^2 - \frac{1}{2}S_2^2 - \frac{1}{2}S_3^2,$$

where

$$S_2 = S_1 + x + X \quad \text{and} \quad S_3 = S_2 + y + Y = S_1 + x + X + y + Y$$

The first-order conditions are

$$\frac{\partial w}{\partial x} = a - x - (S_1 + x + X) - (S_1 + x + y + X + Y) = 0$$
$$\frac{\partial w}{\partial y} = a - y - (S_1 + x + y + X + Y) = 0$$

CAP, taking $\{x, y\}$ as given, chooses $\{X, Y\}$ to maximize W. The first-order conditions are

$$\frac{\partial W}{\partial X} = A - X - (S_1 + x + X) - (S_1 + x + y + X + Y) = 0$$
$$\frac{\partial W}{\partial Y} = A - Y - (S_1 + x + y + X + Y) = 0$$

Solving these four first-order conditions simultaneously, we obtain the OLNE. Let the superscript OL denotes open-loop strategies. The equilibrium open-loop strategy of SMALL is $\{x^{\mathrm{OL}}, y^{\mathrm{OL}}\}$, where

$$x^{\mathrm{OL}} = \frac{6}{11}a - \frac{4}{11}S_1 - \frac{5}{11}A$$
$$y^{\mathrm{OL}} = \frac{7}{11}a - \frac{1}{11}S_1 - \frac{4}{11}A$$

Similarly, that of CAP is $\{X^{\mathrm{OL}}, Y^{OL}\}$, where

$$X^{\mathrm{OL}} = \frac{6}{11}A - \frac{4}{11}S_1 - \frac{5}{11}a$$
$$Y^{\mathrm{OL}} = \frac{7}{11}A - \frac{1}{11}S_1 - \frac{4}{11}a$$

It follows that the time path of the state variable under the OLNE is $\{S_1, S_2^{\mathrm{OL}}, S_3^{\mathrm{OL}}\}$ where

$$S_2^{\mathrm{OL}} = S_1 + x^{\mathrm{OL}} + X^{\mathrm{OL}} = \frac{1}{11}(3S_1 + a + A)$$
$$S_3^{\mathrm{OL}} = S_2 + y^{\mathrm{OL}} + Y^{\mathrm{OL}} = \frac{1}{11}(S_1 + 4A + 4a)$$

The two-period welfare of SMALL in the OLNE is

$$w = ax^{\mathrm{OL}} - \frac{1}{2}(x^{\mathrm{OL}})^2 - \frac{1}{2}S_1^2 + ay^{\mathrm{OL}} - \frac{1}{2}(y^{\mathrm{OL}})^2$$
$$-\frac{1}{2}\left[\frac{1}{11}(3S_1 + a + A)\right]^2 - \frac{1}{2}\left[\frac{1}{11}(S_1 + 4A + 4a)\right]^2.$$

Similarly, the two-period welfare of CAP in the OLNE can be computed.

The OLNE is *time-consistent* in the following sense. Given that in period 1 both players have carried out their respective actions x^{OL} and X^{OL}, at the beginning of period 2, if both players are given an opportunity to revise their plan to maximize the remaining part of their payoffs,

$u_2 - g_3 = (ay - \frac{1}{2}y^2 - \frac{1}{2}S_2^2) - \frac{1}{2}S_3^2$ and $U_2 - G_3 = (AY - \frac{1}{2}Y^2 - \frac{1}{2}S_2^2) - \frac{1}{2}S_3^2$, they would choose respectively the same y^{OL} and Y^{OL}.

However, suppose for some reason (perhaps by error), CAP did not emit the amount X^{OL} in period 1. Say, for example, CAP's emission was $X^{\text{OL}} + \varepsilon$. Then, at the beginning of period 2, both players observe that the stock is

$$S_2 = S_2^{\text{OL}} + \varepsilon = \frac{1}{11}(3S_1 + a + A) + \varepsilon$$

Will it remain optimal for SMALL to choose $y = 7/11a - 1/11S_1 - 4/11A$? Recall that

$$\begin{aligned} u_2 - g_3 &= \left(ay - \frac{1}{2}y^2 - \frac{1}{2}S_2^2\right) - \frac{1}{2}S_3^2 \\ &= ay - \frac{1}{2}y^2 - \frac{1}{2}\left[\frac{1}{11}(3S_1 + a + A) + \varepsilon\right]^2 \\ &\quad - \frac{(\frac{1}{11}(3S_1 + a + A) + \varepsilon + y + Y)^2}{2} \end{aligned}$$

The first-order condition for SMALL's optimal choice y is

$$\frac{\partial(u_2 - g_3)}{\partial y} = a - y - \left(\frac{1}{11}(3S_1 + a + A) + \varepsilon + y + Y\right) = 0$$

Similarly for CAP

$$\frac{\partial(U_2 - G_3)}{\partial Y} = A - Y - \left(\frac{1}{11}(3S_1 + a + A) + \varepsilon + y + Y\right) = 0$$

Because of the presence of the perturbation ε, these two first-order conditions yield emissions (y, Y) that are not the same as the originally planned quantities $(y^{\text{OL}}, Y^{\text{OL}})$. This shows that the OLNE is not robust to deviation, that is, it is not subgame perfect.

When the second period comes, a player would be foolish to stick to the previously planned emission Y^{OL} if the observed pollution stock S_2 turned out to be different from S_2^{OL}. It would seem, then, that in an environment where perturbations are possible, each player would be wise to think that the other player would act in each period according to the observed level of stock. This gives rise to the idea of a feedback (or Markov-perfect) strategy: optimal current action should depend on currently observed state.

Finding the MPNE

Markov-perfect strategies are found by solving the game backward. This method gives us "feedback decision rules" (e.g., emission in a given period depends on the observed stock at the beginning of that period).

At the beginning of period 2, given the observed stock level S_2, SMALL chooses y to maximize

$$u_2 - g_3 = ay - \frac{1}{2}y^2 - \frac{1}{2}S_2^2 - \frac{1}{2}(S_2 + y + Y)^2$$

The first-order condition is

$$a - y - (S_2 + y + Y) = 0$$

Thus, we obtain SMALL's reaction function for period 2:

$$y = \frac{a - Y - S_2}{2}$$

Similarly, CAP's reaction function in period 2 is

$$Y = \frac{A - y - S_2}{2}$$

The intersection of the two reaction curves determines the period-two Nash equilibrium emissions, expressed as functions of the observed stock S_2. Thus, we get the feedback (FB) decision rules for SMALL and CAP

$$y^{\mathrm{FB}}(S_2) = \frac{2}{3}a - \frac{1}{3}S_2 - \frac{1}{3}A$$
$$Y^{\mathrm{FB}}(S_2) = \frac{2}{3}A - \frac{1}{3}S_2 - \frac{1}{3}a,$$

where the superscript FB indicates that this is a feedback equilibrium. Note that these decision rules indicate that a player will pollute less in period 2 if it observes a greater level of pollution at the beginning of that period.[9]

The resulting feedback equilibrium stock S_3 is then

$$S_3 = S_2 + y^{\mathrm{FB}}(S_2) + Y^{\mathrm{FB}}(S_2) = \frac{1}{3}(S_2 + A + a).$$

Let us work backward to find the equilibrium decision rules for period 1. The feedback-equilibrium payoff to SMALL for period 2 (including the terminal

[9] As we show below, knowing this, each player *has an incentive to pollute a bit more in period* 1 (to increase its first period payoff), as it believes the other player will reduce its own emissions in period 2.

term $g_3(S_3)$) is

$$u_2 - g_3 = ay^{\mathrm{FB}}(S_2) - \frac{1}{2}\left(y^{\mathrm{FB}}(S_2)\right)^2 - \frac{1}{2}S_2^2 - \frac{1}{2}(S_2 + y^{\mathrm{FB}}(S_2) + Y^{\mathrm{FB}}(S_2))^2 \equiv v_2(S_2).$$

Similarly

$$U_2 = AY^{\mathrm{FB}}(S_2) - \frac{1}{2}\left(Y^{\mathrm{FB}}(S_2)\right)^2 - \frac{1}{2}S_2^2 - \frac{1}{2}\left(S_2 + y^{\mathrm{FB}}(S_2) + Y^{\mathrm{FB}}(S_2)\right)^2 \equiv V_2(S_2).$$

Working backward, at the beginning of period 1, SMALL chooses x to maximize

$$u_1 + v_2(S_2) \equiv ax - \frac{1}{2}x^2 - \frac{1}{2}S_1^2 + v_2(S_2),$$

where $S_2 = S_1 + x + X$.

SMALL's first-order condition is

$$\frac{\partial u_1}{\partial x} + \frac{dv_2}{dS_2}\frac{\partial S_2}{\partial x} = 0,$$

where

$$\begin{aligned}\frac{dv_2}{dS_2} &= \frac{\partial(u_2 - g_3)}{\partial y^{\mathrm{FB}}}\frac{dy^{\mathrm{FB}}}{dS_2} + \frac{\partial(u_2 - g_3)}{\partial Y^{\mathrm{FB}}}\frac{dY^{\mathrm{FB}}}{dS_2} + \frac{\partial(u_2 - g_3)}{\partial S_2} \\ &= 0 - S_3\frac{dY^{\mathrm{FB}}}{dS_2} - S_2 - S_3 \\ &= -S_2 - \frac{1}{3}(S_2 + A + a)\left[1 + \frac{dY^{\mathrm{FB}}}{dS_2}\right] = -S_2 - \frac{2}{9}(S_2 + A + a) \\ &= -\frac{(11S_2 + 2A + 2a)}{9}.\end{aligned}$$

Therefore, SMALL's first-order condition is simply

$$a - x - \frac{(11S_2 + 2A + 2a)}{9} = 0$$

That is,

$$a - x - \frac{(11S_1 + 11x + 11X + 2A + 2a)}{9} = 0.$$

This equation yields SMALL's period 1 reaction function:

$$x = \frac{-11S_1 - 11X - 2A + 7a}{20}$$

Similarly, CAP's period 1 reaction function is

$$X = \frac{-11S_1 - 11x - 2a + 7A}{20}$$

The intersection of these two reaction curves gives the equilibrium feedback decision rules in period 1:

$$x^{\text{FB}}(S_1) = \frac{18}{31}a - \frac{11}{31}S_1 - \frac{13}{31}A$$
$$X^{\text{FB}}(S_1) = \frac{18}{31}A - \frac{11}{31}S_1 - \frac{13}{31}a.$$

The two-period welfare of each player can then be computed, and expressed as functions of the parameters of the model (in our example, the parameters are S_1, a, and A).

Comparing OLNE and MPNE of the transboundary pollution game

Let us compare the first-period emissions under OLNE with those under MPNE. For SMALL,

$$x^{\text{OL}} = \frac{6}{11}a - \frac{4}{11}S_1 - \frac{5}{11}A$$
$$x^{\text{FB}}(S_1) = \frac{18}{31}a - \frac{11}{31}S_1 - \frac{13}{31}A$$

Both equations indicate that a higher initial stock level S_1 will cause players to opt for lower emissions in period 1, but since $4/11 > 11/31$, we can see that players are more willing to cut back their emissions under OLNE than under MPNE. On the other hand, an exogenous increase in the preference parameter a will lead to more emissions, but this effect is much more pronounced under the MPNE.

An interesting question is: does the OLNE give rise to a higher overall payoff for each player than the MPNE? If it does, one could argue that in the context of this model, *institutions that facilitate precommitment would be welfare enhancing.*

Let us consider a numerical example. Let $A = a = 20$, and $S_1 = 2$. Then, $x^{\text{OL}} = X^{\text{OL}} = 1.09$, $y^{\text{OL}} = Y^{\text{OL}} = 5.27$, $S_2^{\text{OL}} = 4.18$ and $S_3^{\text{OL}} = 14.72$. Thus, the overall payoff of each player in the OLNE is equal to -6.41.

Compare with the MPNE: $x^{\mathrm{FB}} = X^{\mathrm{FB}} = 2.51 > 1.09$. This results in $S_2 = 7.03$. The second-period emissions are $y^{\mathrm{FB}} = Y^{\mathrm{FB}} = 4.32 < 5.27$. The resulting S_3 is 15.67. The overall payoff of each under the MPNE is -25.34.

The above numerical example indicates that both players are better off if they both use open-loop strategies, under which, given S_1, each player would commit to a time path of actions, regardless of what S_2 turns out to be. However, in the absence of a mechanism to ensure that they honor their commitments, each player would believe that the other player will deviate from the committed action for period 2 if S_2 turns out to be different from the level implied by their committed period 1 emissions. Suppose CAP believes this. Then, it will deviate from the OLNE by increasing its period 1 emissions beyond its OLNE level, so as to increase S_2, knowing that SMALL will then reduce its own period 2 emissions, as dictated by the rule $y^{\mathrm{FB}}(S_2)$ that we discovered above. Thus, CAP's first-period deviation from OLNE will increase its overall payoff (because it manages to pollute more while getting SMALL to pollute less in period 2). If SMALL anticipates this, it will also deviate from first period OLNE emissions.

The above discussion indicates that while an OLNE might achieve higher overall payoff for both players, their mutual suspicion and opportunistic behavior will prevent an OLNE from being realized.

1.2.2. *Choice among equilibrium concepts*

When one formulates and analyzes a dynamic game, should one focus only on MPNE? Is the MPNE likely to be a better prediction of the outcome of the game?

Recall that in an MPNE, each player assumes that the other player will act in each period according to the observed level of the state variable in that period. Each player therefore has an incentive to influence the state variable in period t with the objective of manipulating the other player's action in period $t + 1$. Thus, MPNE is perhaps a better concept if one believes that players are sophisticated and manipulative.

If players realize that, in a given game, the OLNE gives higher welfare to each than the MPNE does, would there be incentives for them to agree on playing open-loop strategies? Such an agreement would require an ability to commit to an initially announced time path of actions.

One may argue that in situations where agents are able to commit, an OLNE may well be a good prediction of the outcome of the game. Another advantage of OLNE is that it is relatively easier to compute than MPNE.

In situations where calculations are extremely costly, there is a plausible presumption that economic agents may opt for the easy-to-compute OLNE. In addition, there are situations where an OLNE is so plausible and intuitive that it is not worthwhile to try to compute MPNE for a small gain in sophistication. We later provide an example of this type.

It is worth noting that OLNE and MPNE can be thought of as based on two alternative, both extreme, assumptions about ability to precommit. In the OLNE, players commit to a whole time path of play. In the MPNE, players cannot precommit at all. Reinganum and Stokey (1985) argue that in some cases, players may be able to commit to actions in the near future (e.g., by forward contracts) but not to actions in the distant future. They develop a simple model where the game begins at time 0 and ends at a fixed time T, and there are k periods of equal length δ, where $k\delta = T$. At the beginning of each period, agents can commit to a path of action during that period. The special case where $k = 1$ corresponds to the open-loop formulation, and the OLNE is then the appropriate equilibrium concept. At the other extreme, where $\delta \to 0$, the appropriate equilibrium concept is MPNE. This issue is discussed further in the chapter on natural resources (Chap. 3).

We now turn to a simple example where there is a very plausible OLNE.

1.2.3. *Another simple dynamic game: non-cooperative cake eating*

Two players must share a cake that has been cut into six identical pieces. They have three days, $t = 1, 2, 3$, to play this game. Let s_t denote the number of pieces of cake that remain at the beginning of day t. Then, $s_1 = 6$. Each day t, after observing the number of pieces of cake that remain, s_t, each player must touch one of the three buttons marked $0, 1$, and 2 on its computer screen. To put it more formally, the control variable for player i, denoted by b_{it}, can take one of the following three values $0, 1$, and 2. According to the rules of the game, if $b_{1t} + b_{2t} \leq s_t$ then player i $(i = 1, 2)$ is given b_{it} pieces and consumes them. Its utility for that day is the square root of the number of pieces it consumes:

$$u_{it} = \sqrt{b_{it}}$$

If $b_{1t} + b_{2t} > s_t$, then both players get nothing, and their utility is 0.

The three-period welfare of player i ($i = 1, 2$) is assumed to be the non-discounted sum of utilities:

$$w_i = u_{i1} + u_{i2} + u_{i3}$$

It is easy to see that the following pair of open-loop strategies constitutes an OLNE: each player chooses 1 each day. Along the equilibrium play, each will consume one piece each day. Its equilibrium three-period welfare is then

$$\sqrt{1} + \sqrt{1} + \sqrt{1} = 3$$

It can be verified that it does not pay any player to deviate from this equilibrium. Suppose player 2 uses the open-loop equilibrium strategy ($b_{2t} = 1$ for $t = 1, 2, 3$) and player 1 deviates, say, by choosing $b_{11} = b_{12} = 2$ and $b_{13} = 2$ (or 1) then player 1's payoff will be

$$\sqrt{2} + \sqrt{2} + \sqrt{0} = 2.8284 < 3$$

Therefore, this deviation is not profitable. It is easy to check that in fact there is no profitable deviation.

It is interesting to observe that the OLNE of this game is Pareto optimal. If the two players were allowed to collude in their choice of strategies, they would not achieve better outcome.

Let us show that the OLNE is not subgame perfect. Suppose both players have decided to play the OLNE strategies ($b_{it} = 1$ for $t = 1, 2, 3$ for $i = 1, 2$) but, on the second day, player 1 observes that there are only three pieces of cake left (i.e., at least one player has made a mistake by touching the wrong button). Would player 1 still want to continue with its original open-loop strategy? If it does (and assuming player 2 will play $b_{2t} = 1$ for $t = 2, 3$), it will get

$$\sqrt{1} + \sqrt{1} + \sqrt{0} = 2$$

If it chooses instead $b_{12} = 2$, it will get a better payoff

$$\sqrt{1} + \sqrt{2} + \sqrt{0} = 2.4142$$

assuming player 2 will play $b_{2t} = 1$ for $t = 2, 3$. This shows that the OLNE is, in general, not subgame perfect: if the time path of the state variable is observed to have deviated from the equilibrium path, at least one player will want to deviate from the original plan.

The reader is invited to find all the MPNEs for this game (recall that an MPNE is a pair of feedback strategies that are best replies to each other) and to show that the outcome of any MPNE of this game is that each player consumes one piece of cake each day, that is, the same outcome as that obtained under OLNE.

1.2.4. *An infinite horizon game of transboundary pollution in continuous time*

Let us consider the following game proposed by Long (1992). There are two countries. Country i's output at date t is denoted by $y_i(t)$. Assume all output is consumed. Emissions are proportional to output, $E_i(t) = y_i(t)$ where the factor of proportionality is normalized at unity. (In what follows, we use y_i and E_i interchangeably.) The stock of pollution, common to both countries, is $S(t)$. The rate of accumulation of the stock is equal to the sum of emissions minus the natural decay:

$$\dot{S}(t) = E_1(t) + E_2(t) - \delta S(t), \tag{1.1}$$

where $\delta > 0$ is the decay rate. The pollution damage suffered by country i at time t is

$$D_i(S(t)) = \frac{c_i}{2}(S(t))^2,$$

where $c_i > 0$ is the damage parameter. The utility of consumption is $U_i(y_i(t)) = A_i y_i(t) - 1/2(y_i(t))^2$, where A_i is a positive constant. The net utility, denoted by $B_i(t)$, is defined as the utility of consumption minus the damage cost:

$$B_i(t) = A_i y_i(t) - \frac{1}{2}(y_i(t))^2 - \frac{c_i}{2}(S(t))^2$$

The government of country i perceives that the country's social welfare is

$$W_i = \int_0^\infty e^{-\rho t} B_i(t) dt,$$

where $\rho > 0$ is the rate of discount. Its objective is to maximize the country's social welfare subject to the transition Eq. (1.1). In doing so, it must know if the government of the other country uses an open-loop or a feedback-emission strategy. These two cases are examined separately below.

OLNE in the infinite horizon model

If country i believes that country j uses an open-loop emission strategy, $E_j(t) = \phi_j^{\mathrm{OL}}(t)$, its optimization problem becomes

$$\max_{E_i(.)} \int_0^{\infty} e^{-\rho t} \left[A_i E_i(t) - \frac{1}{2}(E_i(t))^2 - \frac{c_i}{2}(S(t))^2 \right] dt \tag{1.2}$$

subject to

$$\dot{S}(t) = E_i(t) + \phi_j^{\mathrm{OL}}(t) - \delta S(t), \qquad S(0) = S_0. \tag{1.3}$$

This is a simple optimal control problem, where $\phi_j^{\mathrm{OL}}(t)$ is taken as an exogenously given time path. Let us solve this problem using the Maximum Principle.[10] Let H_i denote the Hamiltonian function for the optimal control problem (1.2), and ψ_i be the co-state variable. Then

$$H_i = A_i E_i - \frac{1}{2}(E_i)^2 - \frac{c_i}{2}(S)^2 + \psi_i(E_i + \phi_j^{\mathrm{OL}} - \delta S)$$

The necessary conditions are

$$\frac{\partial H_i}{\partial E_i} = A_i - E_i + \psi_i = 0 \tag{1.4}$$

$$-(\dot{\psi}_i - \rho\psi_i) = \frac{\partial H_i}{\partial S} = -c_i S - \psi_i \delta \tag{1.5}$$

$$\dot{S} = \frac{\partial H_i}{\partial \psi_i} = E_i + \phi_j^{\mathrm{OL}} - \delta S \tag{1.6}$$

and the transversality condition is

$$\lim_{t\to\infty} e^{-\rho t}\psi_i(t)S(t) = 0 \tag{1.7}$$

The variable ψ_i can be eliminated by using the necessary condition (1.4), and thus we get the following conditions

$$-(\dot{E}_i - \rho(E_i - A_i)) = -c_i S - \delta(E_i - A_i)$$

$$\dot{S} = E_i + \phi_j^{\mathrm{OL}} - \delta S, \qquad S(0) = S_0$$

$$\lim_{t\to\infty} e^{-\rho t}(E_i(t) - A_i)S(t) = 0.$$

[10] See, for example, Kamien and Schwartz (1991) or Léonard and Long (1992).

A similar set of equations applies to country j's optimization problem, if j believes that i uses an open-loop emissions strategy $E_i(t) = \phi_i^{\mathrm{OL}}(t)$:

$$-(\dot{E}_j - \rho(E_j - A_j)) = -c_j S - \delta(E_j - A_j)$$

$$\dot{S} = E_j + \phi_i^{\mathrm{OL}} - \delta S$$

$$\lim_{t\to\infty} e^{-\rho t}(E_j(t) - A_j)S(t) = 0$$

To find an OLNE, we must find a pair of functions $(\phi_1^{\mathrm{OL}}, \phi_2^{\mathrm{OL}})$ such that $\phi_1^{\mathrm{OL}}(t) = E_1^*(t)$ and $\phi_2^{\mathrm{OL}}(t) = E_2^*(t)$ where $(E_1^*(t), E_2^*(t), S^*(t))$ satisfy the three differential equations

$$\dot{E}_1(t) = c_1 S(t) + (\rho + \delta)(E_1(t) - A_1) \tag{1.8}$$

$$\dot{E}_2(t) = c_2 S(t) + (\rho + \delta)(E_2(t) - A_2) \tag{1.9}$$

$$\dot{S}(t) = E_1(t) + E_2(t) - \delta S(t), \qquad S(0) = S_0 \tag{1.10}$$

and, in addition, the following transversality conditions

$$\lim_{t\to\infty} e^{-\rho t}(E_1(t) - A_1)S(t) = 0 \tag{1.11}$$

$$\lim_{t\to\infty} e^{-\rho t}(E_2(t) - A_2)S(t) = 0 \tag{1.12}$$

This is a system of three differential equations with three boundary conditions. Before solving for an OLNE for this system let us consider the special case where the cost and preference parameters of the two countries are identical, that is, $A_1 = A_2 = A$ and $c_1 = c_2 = c$. In this case, it is reasonable to assume that the two countries will adopt identical open-loop strategies, that is, $E_1^*(t) = E_2^*(t) = E^*(t)$. Then, the system reduces to two differential equations

$$\dot{E}(t) = cS(t) + (\rho + \delta)(E(t) - A) \tag{1.13}$$

$$\dot{S}^*(t) = 2E(t) - \delta S(t), \qquad S(0) = S_0 \tag{1.14}$$

with the transversality condition

$$\lim_{t\to\infty} e^{-\rho t}(E(t) - A)S(t) = 0. \tag{1.15}$$

The pair of differential Eqs. (1.13) and (1.14) admits a unique steady-state pair $(\widehat{S}, \widehat{E})$, where

$$\widehat{S} = \frac{2A(\delta + \rho)}{2c + \delta(\delta + \rho)} \tag{1.16}$$

$$\widehat{E} = \frac{A\delta(\delta+\rho)}{2c+\delta(\delta+\rho)} = \frac{\delta\widehat{S}}{2}, \tag{1.17}$$

Remark 1 If the two countries co-operate and maximize the sum of their welfares, the steady-state stock of pollution will be lower than $\widehat{S}$.

We can show that steady-state pair $(\widehat{S}, \widehat{E})$ has the saddle-point property, in the sense that for any given S_0, there exists a *unique* corresponding $E^*(0)$ such that if both countries choose $E(0)$ as their initial emission rate, the pair of time paths $(S^*(t), E^*(t))$ starting with $(S_0, E^*(0))$ at time zero will converge to the steady-state pair $(\widehat{S}, \widehat{E})$. Of course, this pair of time paths satisfies the transversality condition (1.15). We show below how to determine $E^*(0)$, given S_0.

To prove the saddle-point property, we must show that the Jacobian matrix J of the system (1.13)–(1.14) has exactly one negative real root, that is, it satisfies the condition $\det J < 0$ (recall that the product of the roots equals the determinant of J). Now

$$J = \begin{bmatrix} \rho+\delta & c \\ 2 & -\delta \end{bmatrix} \equiv \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$
$$\det J = a_{11}a_{22} - a_{12}a_{21} = -\delta(\rho+\delta) - 2c < 0.$$

Recall that the trace of matrix J is $\text{trace}(J) = a_{11} + a_{12}$. The characteristic equation is

$$\lambda^2 - \text{trace}(J)\lambda + \det J = 0$$

or

$$\lambda^2 - \rho\lambda - \delta(\rho+\delta) - 2c = 0. \tag{1.18}$$

Equation (1.18) has one negative real root, denoted by λ_1, and one positive real root, denoted by λ_2:

$$\lambda_1 = \frac{\text{trace}(J) - \sqrt{(\text{trace}(J))^2 - 4\det J}}{2}$$
$$= \frac{\rho - \sqrt{(\rho+2\delta)^2 + 8c}}{2} < 0 \tag{1.19}$$
$$\lambda_2 = \frac{\rho + \sqrt{(\rho+2\delta)^2 + 8c}}{2} > 0 \tag{1.20}$$

We take the negative root for convergence. The convergent paths of E and S (along the stable branch of the saddle point) are given by

$$\begin{bmatrix} E^*(t) - \widehat{E} \\ S^*(t) - \widehat{S} \end{bmatrix} = \beta \begin{bmatrix} \mathbf{k}_1^{(1)} \\ \mathbf{k}_2^{(1)} \end{bmatrix} e^{\lambda_1 t} \equiv \beta \mathbf{k} e^{\lambda_1 t}, \tag{1.21}$$

where $\mathbf{k}$ is a characteristic vector corresponding to the negative root λ_1, and β is a constant. Take

$$\mathbf{k} = \begin{bmatrix} \mathbf{k}_1^{(1)} \\ \mathbf{k}_2^{(1)} \end{bmatrix} = \begin{bmatrix} 1 \\ (\lambda_1 - a_{11})a_{12}^{-1} \end{bmatrix}.$$

Then, Eq. (1.21) gives

$$\frac{E^*(t) - \widehat{E}}{S^*(t) - \widehat{S}} = \frac{1}{(\lambda_1 - a_{11})a_{12}^{-1}} \tag{1.22}$$

Therefore, the right-hand side (RHS) of Eq. (1.22) is the slope of the stable branch of the saddle point in the space (S, E). At $t = 0$, $S^*(0) = S_0$ (given). It follows that given S_0, the initial emission rate dictated by the OLNE is

$$E^*(0) = \widehat{E} + \frac{S_0 - \widehat{S}}{(\lambda_1 - a_{11})a_{12}^{-1}} = \widehat{E} + \frac{(S_0 - \widehat{S})c}{\lambda_1 - \rho - \delta} \tag{1.23}$$

Remark 2 The matrix Eq. (1.21) gives

$$E^*(t) - \widehat{E} = \beta \mathbf{k}_1^{(1)} e^{\lambda_1 t} \tag{1.24}$$

$$S^*(t) - \widehat{S} = \beta \mathbf{k}_1^{(2)} e^{\lambda_1 t}. \tag{1.25}$$

Differentiating these equations with respect to t, we get, after substitution,

$$\dot{E}^*(t) = \lambda_1 (E^*(t) - \widehat{E}) \tag{1.26}$$

$$\dot{S}^*(t) = \lambda_1 (S^*(t) - \widehat{S}). \tag{1.27}$$

The entire time path of E^* can be obtained using the first-order differential Eq. (1.26) and the initial value $E^*(0)$ as given by Eq. (1.23). Thus

$$E^*(t) = \widehat{E} + \left(E^*(0) - \widehat{E}\right) e^{\lambda_1 t} \tag{1.28}$$

Similarly

$$S^*(t) = \widehat{S} + (S_0 - \widehat{S}) e^{\lambda_1 t} \tag{1.29}$$

Remark 3 If initially the environment is perfectly clean, that is, if $S_0 = 0$, then the (OLNE) initial emission rate is

$$E_0^* = \widehat{E} - \frac{\widehat{S}}{(\lambda_1 - a_{11})a_{12}^{-1}} = \widehat{S}\left[\frac{\delta}{2} - \frac{1}{(\lambda_1 - a_{11})a_{12}^{-1}}\right] \tag{1.30}$$

Remark 4 An alternative representation of the open-loop equilibrium path is a diagram in the $(S, \dot{S})$ space. Equation (1.27) describes the movement of $S^*(t)$ that results from the OLNE. It can be depicted in a simple diagram in the $(S, \dot{S})$ space.

Remark 5 An alternative method of finding $E^*(0)$ is as follows. At time $t = 0$, Eq. (1.27) gives

$$2E^*(0) - \delta S(0) = \lambda_1(S(0) - \widehat{S}), \tag{1.31}$$

which determines $E^*(0)$ uniquely. Note that in the special case where initially the environment is perfectly clean, that is, if $S_0 = 0$, Eq. (1.31) gives

$$2E_0^* = \lambda_1 \widehat{S} \tag{1.32}$$

Equation (1.32) is consistent with Eq. (1.30) because $2a_{12} = (\lambda_1 + \delta)(\lambda_1 - a_{11})$.

Remark 6 From Eq. (1.27), we get

$$2E^*(t) - \delta S^*(t) = \lambda_1\left[S^*(t) - \widehat{S}\right]$$

Hence, along the open-loop equilibrium play, there is a linear relationship between E^* and S^*:

$$E^* = \frac{1}{2}\left[(\lambda_1 + \delta)S^* - \lambda\widehat{S}\right] \tag{1.33}$$

This equation is sometimes called the "feedback representation of the OLNE". It should not be interpreted as a feedback strategy.

Let us return to the general case where the two countries differ, that is, $A_1 \neq A_2$ and $c_1 \neq c_2$. We must solve for the OLNE using the system of three differential Eq. (1.8)–(1.10). We show that there exists a unique solution $(E_1^*(t), E_2^*(t), S^*(t))$ that converges to a unique steady state $(\widehat{E}_1, \widehat{E}_2, \widehat{S}^a)$. Here, the superscript a in $\widehat{S}^a$ indicates that we are dealing with the case of an asymmetric world. The vector $(\widehat{E}_1, \widehat{E}_2, \widehat{S}^a)$ is the solution of the matrix equation

$$\begin{bmatrix} \rho+\delta & 0 & c_1 \\ 0 & \rho+\delta & c_2 \\ 1 & 1 & -\delta \end{bmatrix} \begin{bmatrix} \widehat{E}_1 \\ \widehat{E}_2 \\ \widehat{S}^a \end{bmatrix} = \begin{bmatrix} -(\rho+\delta)A_1 \\ -(\rho+\delta)A_2 \\ 0 \end{bmatrix}$$

The steady state $(\widehat{E}_i, \widehat{E}_j, \widehat{S}^a)$ has the saddle-point property. To verify this, consider the Jacobian matrix

$$J_3 \equiv \begin{bmatrix} \rho+\delta & 0 & c_1 \\ 0 & \rho+\delta & c_2 \\ 1 & 1 & -\delta \end{bmatrix}$$

Then

$$\det J_3 = -\delta(\rho+\delta)^2 - (c_1+c_2)(\rho+\delta) < 0$$

The steady-state values are

$$\begin{bmatrix} \widehat{E}_1 \\ \widehat{E}_2 \\ \widehat{S} \end{bmatrix} = \frac{1}{\det J_3} \begin{bmatrix} -A_1\delta(\rho+\delta)^2 - A_1c_2(\rho+\delta) + A_2c_1(\rho+\delta) \\ -A_2\delta(\rho+\delta)^2 - A_2c_1(\rho+\delta) + A_1c_2(\rho+\delta) \\ -(A_1+A_2)(\rho+\delta)^2 \end{bmatrix}$$

Thus

$$\widehat{S}^a = \frac{(A_1+A_2)(\rho+\delta)}{(c_1+c_2)+\delta(\rho+\delta)} \tag{1.34}$$

The determinant of J_3 is negative and the trace of J_3 is positive. This implies there is a negative root, and two roots with positive real parts. This shows that the system exhibits saddle-point stability: given S_0, one can determine a unique pair $(E_i^*(0), E_j^*(0))$ such that starting from $(E_i^*(0), E_j^*(0), S_0)$, the path $(E_i^*(t), E_j^*(t), S(t))$ converges to the steady state $(\widehat{E}_i, \widehat{E}_j, \widehat{S}^a)$. The stable branch of the saddle-point is a line in the three-dimensional space (E_i, E_j, S). It is a one-dimensional stable manifold.

Remark 7 We can directly compute the roots of the characteristic equation

$$\det[J_3 - \lambda I] = 0$$

This equation is a cubic equation in λ

$$(\lambda-\delta-\rho)^2(\lambda+\delta) - (\lambda-\delta-\rho)(c_1+c_2) = 0$$

and the roots are

$$\lambda_1 = \frac{1}{2}\left[\rho - \sqrt{(\rho+2\delta)^2 + 4(c_1+c_2)}\right] < 0$$
$$\lambda_2 = \frac{1}{2}\left[\rho + \sqrt{(\rho+2\delta)^2 + 4(c_1+c_2)}\right] > 0$$
$$\lambda_3 = \delta + \rho > 0$$

We choose the negative root λ_1 for convergence. Then, we have the solution path for S:

$$S(t) = \widehat{S}^a + (S_0 - \widehat{S}^a)e^{\lambda_1 t}$$

Finding an MPNE in the infinite horizon model

Suppose country i believes that country j uses a feedback emissions strategy, $E_j(t) = \phi_j^{\mathrm{FB}}(S(t))$, that is, the rate of emissions at t is dependent on the currently observed level of the stock, $S(t)$. Then, the optimal control problem for country i is

$$\max_{E_i(.)} \int_0^\infty e^{-\rho t}\left[A_i E_i(t) - \frac{1}{2}(E_i(t))^2 - \frac{c_i}{2}(S(t))^2\right] dt \tag{1.35}$$

subject to

$$\dot{S}(t) = E_i(t) + \phi_j^{\mathrm{FB}}(S(t)) - \delta S(t), \qquad S(0) = S_0. \tag{1.36}$$

Notice that $\phi_j^{\mathrm{FB}}(S)$ is a function of S, and not a function of t. Country i therefore knows that if it influences S, it will indirectly influence the emission rate chosen by country j. This adds a strategic consideration, which was not present in the open-loop case. Let us see how this additional strategic consideration affects the necessary conditions for i's optimal control problem. The Hamiltonian for country i is

$$H_i = A_i E_i - \frac{1}{2}(E_i)^2 - \frac{c_i}{2}(S)^2 + \psi_i(E_i + \phi_j^{\mathrm{FB}}(S) - \delta S).$$

The necessary conditions are

$$\frac{\partial H_i}{\partial E_i} = A_i - E_i + \psi_i = 0 \tag{1.37}$$

$$-(\dot{\psi}_i - \rho\psi_i) = \frac{\partial H_i}{\partial S} = -c_i S + \psi_i \frac{d\phi^{\mathrm{FB}}(S)}{dS} - \psi_i\delta \tag{1.38}$$

$$\dot{S} = E_i + \phi_j^{\mathrm{FB}}(S) - \delta S, \qquad S(0) = S_0 \tag{1.39}$$

and the transversality condition is

$$\lim_{t\to\infty} e^{-\rho t}\psi_i(t)S(t) = 0 \tag{1.40}$$

Comparing Eqs. (1.38) and (1.5), we see that in the feedback case, there is an extra term, $\psi_i d\phi^{\rm FB}(S)/dS$. This term reflects the additional strategic consideration that when i takes an action, it realizes that the action will change the future level of S, which will in turn influence j's emissions.

Let us consider the simplest case, where the two countries have identical preference and cost parameters. We focus on the symmetric equilibrium. Substituting $E - A$ for ψ, we obtain from Eq. (1.38)

$$-(\dot{E} - \rho(E - A)) = -cS + (E - A)\,\frac{dE(S)}{dS} - (E - A)\delta \tag{1.41}$$

and

$$\dot{S} = 2E(S) - \delta S \tag{1.42}$$

Now, recall that we are looking for a feedback strategy $E = E(S) \equiv \phi^{\rm FB}(S)$. Thus,

$$\dot{E} = \frac{dE}{dS}\frac{dS}{dt} = (2E(S) - \delta S)\frac{dE}{dS}.$$

Therefore, Eq. (1.41) can be written as

$$-(2E(S) - \delta S)\frac{dE}{dS} + (\rho + \delta)[E(S) - A] = -cS + (E(S) - A)\frac{dE(S)}{dS}. \tag{1.43}$$

Simplify to get

$$\frac{dE}{dS} = \frac{(\rho+\delta)\,[E(S) - A] + cS}{3E(S) - A - \delta S} \tag{1.44}$$

Equation (1.44) is a first-order differential equation, where S is the independent variable and E is the dependent variable. Does this differential equation, together with the transversality condition (1.40) where $\psi_i(t)$ is replaced by $E(t) - A$, help determine an MPNE? It turns out that a more direct approach, using the Hamilton–Jacobi–Bellman (HJB) equation, would give us a clearer picture.

The HJB equation for country i is

$$\rho V_i(S) = \max_{E_i} \left[AE_i - \frac{1}{2}E_i^2 - \frac{c}{2}S^2 + V_i'(S)(E_i + E_j(S) - \delta S) \right], \quad (1.45)$$

where $E_j(S)$ is country j's feedback strategy, and $V_i(S)$ is country i's value function, to be solved for. We also impose the condition that along the equilibrium path, the value does not grow too fast:[11]

$$\lim_{t \to \infty} e^{-\rho t} V(S(t)) = 0 \quad (1.46)$$

Maximizing the RHS of the HJB equation with respect to E_i gives the FOC

$$A - E_i + V_i'(S) = 0$$

Rearranging terms

$$E_i = A + V_i'(S)$$

This gives $E_i = E_i(S)$, that is, the chosen emission rate at any time t depends only on the observed level of the stock, and is independent of time. Now, since we are assuming that the two countries have identical parameter values, we focus on the symmetric solution, where $V_i'(S) = V_j'(S) = V'(S)$ and $E_i(S) = E_j(S) = E(S)$. Then, substituting E by $A + V'(S)$ into the HJB equation, we obtain

$$\rho V(S) = \frac{1}{2}[A^2 + 4AV' + 3(V')^2] - \delta SV' - \frac{c}{2}S^2 \quad (1.47)$$

This is a first-order differential equation which, together with condition (1.46), helps determine an MPNE.

Remark 8 If we differentiate Eq. (1.47) with respect to S, then re-arrange terms and substitute $V'(S) = E(S) - A$ and $V''(S) = E'(S)$, we obtain an equation identical to Eq. (1.44). This shows that the two approaches are, in fact, equivalent.

[11]See Dockner *et al.* (2000, Sec. 3.6) for weaker conditions.

Let us conjecture that the value function is quadratic[12]

$$V(S) = -\frac{\alpha S^2}{2} - \beta S - \mu \tag{1.48}$$

Then,

$$V'(S) = -\alpha S - \beta \tag{1.49}$$

and we get the linear feedback strategy

$$E(S) = A - \beta - \alpha S \tag{1.50}$$

We expect that $\alpha > 0$, that is, a higher stock will lead countries to reduce emissions, and that $\beta > 0$, that is, if $S = 0$, the marginal effect on welfare of an exogenous increase in S is negative.

Substituting Eqs. (1.48) and (1.49) into Eq. (1.47), we obtain a quadratic equation of the form

$$p_0 + p_1 S + p_2 S^2 = 0,$$

where p_0, p_1, and p_2 are expressions involving the parameters δ, ρ, and c and the coefficients α, β, and μ (to be determined). Since this equation must hold for all S, it follows that the following conditions must be satisfied:

$$p_0 = 0$$

$$p_1 = 0$$

$$p_2 = 0,$$

where the last condition implies that

$$p_2 \equiv \frac{3}{2}\alpha^2 + \left(\delta + \frac{\rho}{2}\right)\alpha - \frac{c}{2} = 0 \tag{1.51}$$

These conditions yield the following values for α, β, and μ:

$$\alpha = \frac{1}{3}\left[-\left(\delta + \frac{\rho}{2}\right) + \sqrt{\left(\delta + \frac{\rho}{2}\right)^2 + 3c}\ \right] \equiv \alpha_m \tag{1.52}$$

[12]The quadratic value function gives rise to linear strategies. Dockner and Long (1993) show that for this model, there exist non-linear strategies as well, in which case the corresponding value function is not quadratic. This possibility was pointed out by Reynold (1987), Tsutsui and Mino (1990), and further discussed in Clemhout and Wan (1994). The non-uniqueness is due to the lack of a natural boundary condition, that is, the assumption of feedback strategies is not restrictive enough to yield a unique steady-state and unique feedback strategies.

(We show below that for convergence to a steady state, we must choose the positive root of α.)

$$\beta = \frac{2A\alpha}{\delta + \rho + 3\alpha} \equiv \beta_m$$
$$\mu = \frac{(A-\beta)}{2\rho}(3\alpha - \delta - \rho) \equiv \mu_m.$$

Note that

$$A - \beta = \frac{A(\delta+\rho+\alpha)}{\delta+\rho+2\alpha} > 0$$

and $\mu > 0$ if c is sufficiently large. The linear feedback strategy is

$$E = \frac{A(\delta+\rho+\alpha)}{\delta+\rho+2\alpha} - \alpha S$$

From these, we obtain

$$\dot{S} = \frac{2A(\delta+\rho+\alpha)}{\delta+\rho+2\alpha} - (2\alpha+\delta)S \tag{1.53}$$

For S to converge to a steady state, it is necessary that $2\alpha+\delta > 0$. It can be verified that this condition is satisfied if and only if the positive root for α is used. Note also that $\alpha > 0$ implies $\beta > 0$. From Eq. (1.53) the steady-state pollution stock under the MPNE with linear feedback strategies is

$$\widehat{S}^M = \frac{2A(\delta+\rho+\alpha)}{(\delta+\rho+2\alpha)(2\alpha+\delta)} \tag{1.54}$$

Comparing steady states and transient emissions of the transboundary game under OLNE and under MPNE

Comparing the OLNE steady-state pollution stock $\widehat{S}$ (see Eq. (1.16)) with the MPNE steady state $\widehat{S}^M$, we see that the former is smaller than the latter if and only if

$$(\delta+\rho+\alpha)(2c+\delta(\rho+\delta)) > (\delta+\rho)[\delta(\rho+\delta) + 6\alpha^2 + (5\delta+2\rho)\alpha] \tag{1.55}$$

This relationship always holds, in view of Eq. (1.51). Therefore, $\widehat{S} < \widehat{S}^M$.

Do emissions respond to pollution stock more under the MPNE? Under the MPNE, we have from Eq. (1.53)

$$\frac{\partial \dot{S}^M}{\partial S^M} = -2\alpha - \delta < 0 \tag{1.56}$$

while under the OLNE, we have

$$\frac{\partial \dot{S}^{\mathrm{OL}}}{\partial S^{\mathrm{OL}}} = \lambda_1 = \frac{\rho - \sqrt{(\rho + 2\delta)^2 + 8c}}{2} < 0 \tag{1.57}$$

The absolute value of the RHS of Eq. (1.56) is smaller than that of Eq. (1.57) if and only if

$$\frac{2}{3}\delta + \frac{\rho}{3} + \sqrt{\frac{4}{9}(\rho + 2\delta)^2 + \frac{16}{3}c} < \sqrt{(\rho + 2\delta)^2 + 8c}$$

that is, if and only if

$$\sqrt{\frac{1}{9}(\rho + 2\delta)^2} + \sqrt{\frac{4}{9}(\rho + 2\delta)^2 + \frac{16}{3}c} < \sqrt{(\rho + 2\delta)^2 + 8c}$$

Clearly, this inequality holds for all positive c, ρ, and δ. It follows from this result and from $\widehat{S} < \widehat{S}^M$ that the initial emission under the MPNE is also higher than that under the OLNE.

1.3. Stackelberg Equilibrium

In a two-player game, if one player can make a commitment on what strategy it will use before the other player can choose its strategy, the former is called the Stackelberg leader, and the latter is the follower. In differential games, we make the distinction between open-loop Stackelberg leadership and feedback Stackelberg leadership.

An open-loop Stackelberg leader knows that for any given time path of its control variables, which it announces at the start of the game, the follower will choose a best reply to maximize its payoff. The leader therefore can compute its payoff that would result from each of its feasible announced path, and choose the optimal one. Its best announced path, together with the best reply of the follower, constitute an open-loop stackelberg equilibrium (OLSE).

It turns out that, unlike OLNEs (which are time-consistent), OLSE is generically not time-consistent: if at some time after the game has started the leader is relieved of its commitment to follow its preannounced path, it will typically find it optimal to deviate from that path. (There are exceptions, as we illustrate later.) The intuition behind this time-inconsistency is as follows. If you promise to pay someone a stream of rewards on the condition that it carries out some investment, then once the investment has been sunk, you will have no incentive to keep your promise (given the implicit assumption that there is no loss of reputation, or no cost arising from a loss of reputation).[13] In contrast, in an OLNE, because of the *simultaneous* choice of time paths, no player is trying to influence the action of any other players.

The time-inconsistency of OLSE does not mean that this equilibrium concept is useless. It is a useful equilibrium concept in situations where the leader can credibly precommit, for example by signing a contract that is perfectly enforceable. For example, university teachers in Canada are often required to announce in advance what topics will be covered in the next 13 weeks, in what order, what articles students must read, in what week a midterm exam will be held, what is the percentage of final grade for each assignment, etc. Deviations will be punished (e.g., in some cases promotion can be denied for serious deviations). Thus, a Canadian university teacher is often an open-loop Stackelberg leader, while the students are the followers.

In the next subsection, we consider an example of an OLSE. This is followed by a discussion of feedback stackelberg equilibrium (FBSE).

1.3.1. *Open-loop Stackelberg Equilibrium*

Recall the infinite horizon game of transboundary pollution that we consider in the preceding section. Assume $c_1 = c_2 = c$ and $A_1 = A_2 = A$ for simplicity. Given S_0, consider the OLNE described by the open-loop strategy $E^*(t)$, which is the unique solution of the differential equation

$$\dot{E}^*(t) = \lambda_1 \left[E^*(t) - \widehat{E}\right] \text{ with } E^*(0) = \widehat{E} + \frac{(S_0 - \widehat{S})c}{\lambda_1 - \rho - \delta}, \tag{1.58}$$

where $\lambda_1, \widehat{E}$, and $\widehat{S}$ are given by Eq. (1.19), (1.17), and (1.16), and $E^*(0)$ is given by Eq. (1.23). By construction, this strategy will be chosen by

[13]The time-inconsistency of open-loop Stackelberg games was pointed out by Simaan and Cruz (1973b, p. 619), and recognized in the macroeconomic literature by Kydland and Prescott (1977). In Kydland and Prescott (1977), the government is the open-loop Stackelberg leader and private individuals are followers. In public economics, the models of optimal redistributive taxation formulated by Chamley (1986) and Judd (1985) are formally open-loop Stackelberg games, and therefore are subject to time inconsistency.

player j if it believes that player i uses the same strategy. Now, suppose the rules of the game change as follows: the players are no longer required to choose their strategies simultaneously. Suppose player 1 is allowed to be the Stackelberg leader and player 2 is the follower. Assume that the leader can credibly commit. Clearly, the leader can achieve the above OLNE outcome, simply by announcing that it will use the open-loop strategy found above, see Eq. (1.58). Therefore, the leader can ensure for itself a payoff at least as great as its OLNE payoff. However, it can do better. It knows that player 2 will not choose the strategy (1.58) if it announces its commitments to a different strategy. In fact the leader knows how the follower would react to any time path $E_1(t)$, by computing its best reply $E_2(t)$ that satisfies the following first-order conditions

$$\begin{aligned} A - E_2 + \psi_2 &= 0 \\ \dot{S} &= E_1 + E_2 - \delta S,\ S(0) = S_0 \\ \dot{\psi}_2 &= cS + (\delta + \rho)\psi_2 \end{aligned} \tag{1.59}$$

as well as the transversality condition

$$\lim_{t\to\infty} e^{-\rho t}\psi_2(t)S(t) = 0 \tag{1.60}$$

Since E_2 can be expressed as a function of ψ_2, namely $E_2 = A + \psi_2$, the constraints that the leader faces are Eqs. (1.59), (1.60), and

$$\dot{S} = E_1 + A + \psi_2 - \delta S,\ S(0) = S_0 \tag{1.61}$$

That is, the leader faces two differential equations and a transversality condition.[14] Therefore, the Hamiltonian function for the leader must contain two co-state variables, which we denote by θ and γ, respectively, and which correspond respectively to the "state variables" S and ψ_2. Note that ψ_2 is the follower's shadow price of S, but since the leader faces the differential Eq. (1.59), the variable ψ_2 must be technically treated as a state variable as far as the leader is concerned. The co-state variable γ is then the leader's shadow price of the follower's shadow price. In this model, ψ_2 is the follower's marginal valuation of the pollution stock. This marginal valuation evolves slowly, that is, it must obey the differential Eq. (1.59) and therefore it cannot jump. The more negative is ψ_2, the smaller will be the follower's emissions E_2. The follower's $\psi_2(t)$ depends

[14] This solution method was suggested by Simaan and Cruz (1973a).

on its expectation of the whole future time path of emissions by the leader.

Let us formulate and solve the leader's problem. Its Hamiltonian is

$$H = AE_1 - \frac{E_1^2}{2} - \frac{c}{2}S^2 + \theta(E_1 + (A + \psi_2) - \delta S) + \gamma(cS + (\delta + \rho)\psi_2).$$

The necessary conditions are

$$\frac{\partial H}{\partial E_1} = A - E_1 + \theta = 0 \tag{1.62}$$

$$-(\dot{\theta} - \rho\theta) = \frac{\partial H}{\partial S} = -cS - \delta\theta + c\gamma \tag{1.63}$$

$$-(\dot{\gamma} - \rho\gamma) = \frac{\partial H}{\partial \psi_2} = \theta + \gamma(\delta + \rho) \tag{1.64}$$

$$\dot{S} = E_1 + A + \psi_2 - \delta S, S(0) = S_0 \text{ given}$$

or, substituting for E_1

$$\dot{S} = 2A + \theta + \psi_2 - \delta S,\ S(0) = S_0 \text{ given} \tag{1.65}$$

$$\dot{\psi}_2 = cS + (\delta + \rho)\psi_2,\ \psi_2(0) \text{ free} \tag{1.66}$$

In addition, since $\psi_2(0)$ is not exogenously given (it depends on the whole time path of the leader's control variable, E_1), it can be normally chosen by the leader.[15] The leader's optimal choice of $\psi_2(0)$ implies that its initial shadow price for the "state variable" ψ_2 is zero, that is, $\gamma(0) = 0$.[16] Let us write this "transversality condition" as a numbered equation:

$$\gamma(0) = 0 \tag{1.67}$$

[15] An important point is that for the Stackelberg leader's problem, one should ask what primitive assumptions of the specific models at hand imply $\psi_2(0)$ can (or cannot) be influenced by the leader. This is a type of "controllability problem", see Xie (1997) and Dockner *et al.* (2000, Chap. 5) for details.

[16] This is technically a transversality condition at the beginning of the time horizon, when the initial state variable is not fixed; see, for example, Léonard and Long (1992, Chap. 7). If you can choose the initial level of a state variable, you should set that variable at the level where its marginal contribution to your objective function is zero. Hence $\gamma(0) = 0$.

In addition, we have the usual transversality conditions at the end of the horizon

$$\lim_{t\to\infty} e^{-\rho t}\theta(t)S(t) = 0$$

$$\lim_{t\to\infty} e^{-\rho t}\gamma(t) = 0$$

as well as the constraint (1.60). These conditions are satisfied as long as θ, γ, S, and ψ_2 converge to finite levels.

Let us focus on the steady state of the system of Eq. (1.63)–(1.66). At the steady state, $\dot{\gamma} = \dot{\theta} = \dot{S} = \psi_2 = 0$, and we have four linear equations:

$$\theta(\delta + \rho) + cS - c\gamma = 0 \tag{1.68}$$

$$\gamma\delta + \theta = 0 \tag{1.69}$$

$$2A + \theta + \psi_2 - \delta S = 0 \tag{1.70}$$

$$cS + (\delta + \rho)\psi_2 = 0. \tag{1.71}$$

Solving these four equations, we get the steady-state pollution stock under the OLSE, which we denote by $\widetilde{S}$:

$$\widetilde{S} = \frac{2A}{\delta + \frac{c}{\rho+\delta} + \frac{c}{\rho+\delta+(c/\delta)}} \tag{1.72}$$

Thus, we have obtained the interesting result that the open-loop Stackelberg game leads to a *higher steady-state pollution stock* than the OLNE $\widehat{S}^a$ in Eq. (1.34). It can be verified that in the OLSE, the leader's level of emissions is greater, and the follower's level is smaller, than in their OLNE counterparts. At the steady state, the leader's shadow price γ of the follower's shadow price is positive:

$$\widetilde{\gamma} = \frac{c\widetilde{S}}{c + \delta(\rho + \delta)} > 0.$$

The leader's steady-state shadow price θ of the pollution stock is

$$\widetilde{\theta} = -\delta\widetilde{\gamma} = -\frac{\delta c\widetilde{S}}{c + \delta(\rho + \delta)} < 0$$

and the follower's steady-state shadow price is

$$\widetilde{\psi}_2 = -\frac{c\widetilde{S}}{\delta + \rho} < 0.$$

Does the steady state have the saddle-point properties? Since there are four differential equations, and two initial values are known (they are $S(0) = S_0$ and $\gamma(0) = 0$ by Eq. (1.67) above), saddle-point stability requires the existence of exactly two negative real roots. This issue was taken up in Long (1992), where an analytical formula for computing the four roots of a more general model is also provided.[17]

Finally, we offer a remark about the lack of time consistency. An OLSE is said to be time inconsistent if, at some time $t_1 > 0$, the leader would want to deviate from the originally announced path once it is no longer required to honor its commitment. Here is a method of showing time inconsistency in our model. The leader's co-state variable $\gamma(t)$ (associated with follower's shadow price ψ_2) is positive at the steady state and therefore is also positive when the steady state is almost reached. If the leader can replan at some time $t_1 > 0$ where $\gamma(t_1) > 0$, it will reset $\gamma(t_1) = 0$, which implies a change in the path for $\theta(t)$, and hence in its emissions. This proves time inconsistency. The intuition is as follows. At the start of the game, the leader would announce that it will emit a lot. This induces the follower to choose a low path of emissions, which is good for the leader. However, once its announced emission plan has "worked", it would no longer want to emit so much. Therefore, at the replanning time t_1, it would choose a lower path of emissions. This reduces the marginal environmental damage to both players. It would be reflected in a less negative value for $\psi_2(t_1)$. The fact that $\gamma(t_1) > 0$ means that at t_1, a new announcement of a lower emission path of the leader would increase its payoffs. In a different context, a similar result on time inconsistency is found in Kemp and Long (1980d), and also Karp (1984), who offers an intuitive explanation of the meaning of the shadow price of the follower's shadow price (p. 80).

1.3.2. *Feedback Stackelberg Equilibrium (FBSE)*

Because of the time inconsistency of OLSE (as noted by Simaan and Cruz, 1973b, p. 619), many authors have attempted to obtain results for leader–follower games using an alternative equilibrium concept: FBSE. This equilibrium requires that the leader must use a feedback strategy: its action (e.g., E_1) at any time t depends on the observed pair $(t, S(t))$, such that

[17]The formula originated from Dockner (1985), and has been used by Dockner and Feichtinger (1991) and Kemp *et al.* (1993) to study the existence of limit cycles in optimal control problems and OLSE, respectively.

starting at any (date, state) pair, the continuation of the optimal strategy remains optimal for the leader. The leader must announce its feedback strategy (its decision rule) to the follower, who would believe it only if latter knows that the former would use the same decision rule for all (date, state) pairs. Decision rules that specify action at time t as a function of the observed pair $(t, S(t))$ only are called Markovian decision rules. Unlike OLSE, there is no established general methodology for finding an FBSE.[18]

There is the concept of "stagewise feedback Stackelberg equilibrium", which is most easily explained using a model in discrete time with a finite horizon.[19] At the last period, T, the leader moves first, and the follower responds. Therefore, they both know their equilibrium leader–follower payoffs for period T, as a function of the opening stock x_T. Working backward, in period $T-1$, again the leader makes the first moves, and the follower responds. This type of dynamic programming approach, based on the assumption that the leader can move first in each period, gives rise to the "stagewise feedback Stackelberg equilibrium".[20] However, in general, the concept of FBSE does not require that in each period the leader takes an action before the follower can move. One can think of the feedback Stackelberg leader as a player who can tell the other player(s): "this is my Markovian decision rule, and you can verify that it is in my interest to use the same decision rule at every point of time."The main difficulty with this "global"Stackelberg leadership is twofold: first, for each Markovian decision rule, we must calculate how the follower will react and second, we must choose from the space of all possible decision rules the one that maximizes the (overall) payoff of the leader. Since this space can be very large, the task of finding a FBSE is formidable.

In some models with very special features, global FBSEs can be found relatively easily. For example, if we can find a strategy for the leader that would achieve the "command and control"outcome (i.e., as if it can control the actions of the other players) then clearly it has no incentive to deviate from it under any circumstances. The optimal pollution tax rule obtained by Benchekroun and Long (1998) for a polluting oligopoly is an instance where the feedback Stackelberg leader (the government) is able to achieve

[18]For recent discussions of this issue, see Shimomura and Xie (2008) and Long and Sorger (2009).
[19]The papers by de Zeeuw and van der Ploeg (1991) and Kydland (1975) contain results for stagewise FBSE in discrete time with a finite horizon.
[20]See Başar and Haurie (1984) or Turnovsky *et al.* (1988) for an illustration.

the command and control outcome. Alternatively, one can restrict the space of decision rules that the leader can choose from. Thus, for linear quadratic games, it may seem natural to restrict the leader to decision rules that are linear affine in the state variables. There are several across examples of global Stackelberg leadership and stagewise Stackelberg leadership in the chapters to follow.