

# On Regulation with Zero–Determinant Strategies\*

Mehmet Barlo<sup>†</sup>      Sura İmren<sup>‡</sup>

August, 2016

## Abstract

The regulation of two players is modeled as an iterated game with no discounting where first two players (two countries providing carbon emissions) also interact with a regulator (an independent regulator responsible for controlling the carbon emission levels by imposing punishments). In our setting, employing a zero-determinant (ZD) strategy, the regulator gains a unilateral advantage to enforce a linear relation between the expected payoffs of players. We identify two conditions and prove that the first guarantees the existence of a ZD strategy while the second ensures the existence of an optimal one. Furthermore, we propose an intuitive and simple cost structure in order to enable the regulator to employ an uncomplicated ZD strategy and attain a maximal ZD payoff.

*Journal of Economic Literature* Classification Numbers: C72

*Keywords:* Iterated Prisoners' Dilemma, Zero–Determinant Strategies, Regulation

---

\*This is a revised version of the MA Thesis of İmren (2016). We thank Mustafa Oğuz Afacan and Saadettin Haluk Çiftçi for helpful comments and suggestions. All remaining errors are ours.

<sup>†</sup>**Corresponding Author:** Faculty of Arts and Social Sciences, Sabanci University, Orhanli, Tuzla, 34956, Istanbul, Turkey; Phone: +90 216 483 9284; Fax: +90 216 483 9250 (CC. M. Barlo); email: barlo@sabanciuniv.edu.

<sup>‡</sup>Faculty of Arts and Social Sciences, Sabanci University, Orhanli, Tuzla, 34956, Istanbul, Turkey; email: suraimren@sabanciuniv.edu

## 1. Introduction

Repeated games, extensive form games with iterations of a given stage game in every period, present a general structure in which strategic interactions taking place repeatedly over time are analyzed. They are the milestone to understand how dynamic strategies correlate with each other. Thus, they have long been analyzed in economics, evolutionary biology, political science and in many other areas.

That players are able to condition their strategies on the past behavior in each round in repeated games is the reason for an extensive multiplicity of equilibria. As asserted in the Folk Theorem of repeated games by Aumann and Shapley (1994) and Fudenberg and Maskin (1986), if players are sufficiently patient then any individually rational payoff can be sustained as a subgame perfect equilibrium (SPE). Aumann (1981) proposes that the reduction of a multiplicity of equilibria may be provided by imposing boundaries on the players' memory strategies. In contrary to that approach, Barlo et al. (2009) demonstrate that the even if the action spaces of players is sufficiently rich, the Folk Theorem for SPE continues to hold with 1-memory strategies.

However, it is important to point out that the recent progress shows that a different point of view takes place in the world of game theory. The idea that whether or not to set the co-players's payoffs to a fixed value is investigated instead of examining the equilibrium behavior. In the article entitled "Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent", Press and Dyson (2012) find a new kind of strategy which guarantees one player a higher payoff than the opponent.

Press and Dyson discover a considerable mathematical feature of the two players iterated prisoner's dilemma. They demonstrate the existence of Zero-Determinant (ZD) strategies, a new class of memory-one strategies for the iterated prisoner's dilemma. A ZD strategy player is able to enforce a linear relation between his/her payoff and the opponent's payoff, regardless of the opponent's behavior. Thus, the opponent's expected payoff is set to a fixed value by a ZD strategy player. In consequence of ZD strategies, a player is claimed to have a unilateral power in games.

From a distinct approach used in Press and Dyson (2012), being able to pin the opponent's payoff by using memory-one strategies is also derived in Boerlijst et al. (1997)

and Sigmund (2010). The linear relation between the payoffs is provided by Hilbe et al. (2013) through using a different method which does not contain any determinants.

It is noteworthy that Press and Dyson come up with the results under no discounting for the iterated prisoner's dilemma. However, Akin (2013) extend the theory produced by Hilbe et al. (2013) to develop a general equation for the distribution of Markov strategies when discounting factor is 1 ( $\delta = 1$ ). It is a subsequent study that Hilbe et al. (2015) follow the study of Akin to expand the approach in which expected payoffs are discounted for the case  $\delta < 1$ .

Several studies extend the theory of ZD strategies to other two players repeated games and multi-player repeated games. Roemheld (2013) makes a generalization of procedure and implications of ZD strategies for all symmetric two players two action games and also for Battle of the Sexes. Pan et al. (2015) study ZD strategies for an iterated public goods game. In this multi-player game, each player chooses whether or not to contribute a unit of cost into a public pot in every round. The total contribution in the pot is multiplied by a factor greater than one and less than the number of players and then equally divided among all players. They find that ZD strategies still occur in multi-player games. Their results show that even though a player is able to pin the expected total payoff of all other players, the increasing number of players restrains the ability of the ZD player to fix the total payoff. Hilbe et al. (2014) produce a theory for ZD strategies for multi-player social dilemmas. They explore that ZD strategy players constituting alliances can enforce a linear relation between the average payoff in the alliance and the payoffs of all other players. The impact of a ZD strategy alliance relies on the size of the alliance, the type of social dilemma and lastly the distinct strategies.

Press and Dyson work on the iterated prisoner's dilemma in which there are just two strategies for both players. Guo (2014) moves the research a step further by introducing a theory for the two player multi-strategy games. The results mostly seem similar with the main article. The subsequent research is done by He et al. (2016). They generalize the framework of ZD strategies to multi-player multi-action iterated games.

Furthermore, Chen and Zinger (2014) study the robustness of ZD strategies against evolutionary players. They show that having any knowledge about the opponent's evolution a ZD strategy player obtains the maximum payoff provided that he employs a linear relation. Another extension of ZD strategies is presented in the article titled "Extortion

under Uncertainty: Zero-Determinant Strategies in Noisy Games” by Hao et al. (2014). A comprehensive model introducing the performance of ZD strategies for noisy repeated games is proposed. They display that in an environment with uncertainty caused by errors, a ZD strategy player can still set the co-player’s payoff to a fixed value; however, as the noise level increases the possibility of the ZD player to pin the opponent’s payoff decreases.

ZD strategies have raised considerable attention. Even though there are still untouched points in the theory of ZD strategies, applications of the subject into real world problems can be observed. Sharing wireless resources, one of the widely studied topic in communication networks, is formulated as an iterated prisoner’s dilemma. Al Daoud et al. (2014) present a framework for spectrum sharing problem through designing ZD strategies for service providers. In each stage game, service providers supplying downlink services choose transmission power levels and eventually get some downlink rates depending on other providers’ interference. They show that service providers are able to fix their long-run payoffs by taking actions either to transmit at the maximum level or not to transmit. Service providers use power control strategies, i.e. ZD strategies, which permit them to share spectrum and maintain average rates regardless of the other providers’ power control strategies.

In the current study, we analyze an iterated game of reputation (of carbon emissions) with ZD strategies. The critical distinction of our analysis concerns the fact that the regulator, the ZD player, is a social planner who derives payoffs from the payoffs of the regulated players (countries).

The existing literature on carbon emission contains a variety of scientific analysis on this global issue. Since it is a widespread and significant research topic, there is a bulk of literature on this subject. It includes not only the trade of carbon emission rights but also the design of an allocation mechanism associating carbon emissions with GDP. In the article of MacKenzie et al. (2008), an efficient allocation mechanism is designed for a tradable pollution market. They also find the symmetric equilibrium strategy of each firm and the choice of the regulator’s to minimize emission levels. Another leading article written by Tang and Song (2014) investigates a dynamic game of incomplete information where regulators and enterprises are game participants. According to the type of production, there are two kinds of enterprises, environmental friendly, and environ-

mental pollution. Due to the information asymmetry between regulators and enterprises, regulators can not observe the enterprises' type of production. Regulators choose to supervise enterprises or not after observing the signals sent by enterprises. Consequently, four kinds of refining Bias equilibrium is analyzed in this behavior selection model of enterprise based on a signal game. They find that the regulators cannot attain the optimal control provided that there is asymmetric information about the cost of carbon emissions reduction.

In our context, there are two countries and a regulator/social planner that is aimed to govern carbon emission levels for the interests of the environment. In this study, we consider the repeated games approach in which the possible behavior of a regulator in reply to countries to alleviate the global emission is modeled. We incorporate a framework of ZD strategies into our study in order to investigate the impact of the use of such a ZD player. Countries produce carbon emissions as long as the production and the consumption exist. In our game, countries do not just decide about whether to mitigate or not. They also need to determine the amount of emission. For purposes of simplification, we discretize the amount of per capita carbon emissions into three categories:  $h$  for high emission level,  $m$  for medium emission level and  $\ell$  for low emission level.

In every state, each country decides how much pollution to emit and the regulator announces the target aggregate level of emissions. That is to say, two countries choose an emission level from their respective action set which includes the three discretized emission level. The joint actions of two countries bring about the "publicly observed" level of emission. The publicly observed level can be high, medium or low according to the following formulation: (1) the publicly observed level is high if both countries choose  $h$  or only one of them chooses  $h$  while the other chooses  $m$  or only one of them chooses  $h$  while the other chooses  $\ell$  (2) the publicly observed level is medium if both countries choose  $m$  or only one of them chooses  $m$  while the other chooses  $\ell$ , and (3) the publicly observed level is low if both countries choose  $\ell$ . Simultaneously, the regulator determines the aggregate target level by choosing an action from his action set which is the same as that countries. The resulting state space with which this study concentrates on comprises of the publicly observed actions (resulting from the joint actions of the two countries) and the regulator's action.

As long as the gradually increasing global emission problem exists, the social wel-

fare of a country relies on its emissions due to the production and on the emissions of the other countries since the accumulation of emissions creates a negative effect on all countries. Thus, controlling the emission level induces the regulator to employ a punishment system. The regulator sets the cost of emissions separately for each country. In each instance, the costs of emissions are given relative to the previous round's publicly observed level and the regulator's action. Therefore, the difference between the utility that countries attain by choosing an emission level and the cost due to that emission level gives the payoff structure of countries. However, since we consider behavioral strategies involving randomization over pure actions, the resulting utilities of countries under the behavioral strategy profile is given by the probability weighted average of their pure strategy payoffs. When choosing its emission level, a country does not pay attention to the negative externality its emission accumulation imposes on the other country. However, as a global social planner, the regulator needs to control behavior of countries by implementing a cost structure for them and obtains the negative of the joint utility of both countries.

We identify conditions implying the existence of ZD strategies. The regulator is able to pin the linear combination of other players' payoffs to a fixed value. Also, the regulator can fix his expected payoff and ensure the highest returns. Moreover, we propose a simple and intuitive cost structure to which the regulator can restrict attention in order to achieve maximal ZD payoffs.

In the next chapter, we introduce our model. In Chapter 3, we present the ZD strategies for our model and the optimal ZD payoff for the regulator. Also, we employ a special cost structure to obtain a simple ZD strategy for the regulator. Finally, Chapter 4 provides the conclusion.

## **2. The Model**

Our iterated game consists of three players. There are two countries and an independent regulator. In this context, we denote player 1 and player 2 as country 1 and country 2, respectively and player 0 as the regulator. Countries produce carbon emissions as long as the production and the consumption exist. The carbon emission game resembles a

prisoner's dilemma; however, countries do not just decide about whether to mitigate or not, they also need to determine the amount of emissions. For purposes of simplification, we discretize the amount of per capita carbon emissions into three categories:  $h$  for high emission level,  $m$  for medium emission level and  $\ell$  for low emission level.

In every state, which is to be defined in the current paragraph, each country decides how much pollution to emit and the regulator announces the target aggregate level of emissions. That is to say, in every state, two countries take an action from their respective action set  $a_i \in A_i = \{h, m, \ell\}$  with  $i = 1, 2$ .  $\theta : A_1 \times A_2 \rightarrow A$  denotes the “publicly observed” level of emissions given that player 1 and player 2 take actions in  $A$  with the property that  $\theta(h, h) = \theta(h, m) = \theta(h, \ell) = \theta(m, h) = \theta(\ell, h) = h$ ,  $\theta(m, m) = \theta(m, \ell) = \theta(\ell, m) = m$  and  $\theta(\ell, \ell) = \ell$ . For example,  $\theta(h, m) = \theta(m, h) = h$  means when player 1 chooses  $h$  and player 2 chooses  $m$  or vice versa, the publicly observed level of emission is  $h$ . Simultaneously, the regulator determines the aggregate target level by choosing an action  $a_0 \in A_0 = \{h, m, \ell\}$ . The strategies of the regulator depends on the publicly observed level of emission. The resulting state space with which this study concentrates on consists of publicly observed level of emissions and the regulator's actions, i.e.  $S = \{hh, hm, h\ell, mh, mm, m\ell, \ell h, \ell m, \ell\ell\}$ .

Carbon emissions not only have positive effects on social welfare of the countries (since then countries are provided with production hence, welfare) but also lead to harmful results on the environment. Thus, there is a cost of this detrimental consequence. The regulator sets the cost of emissions. The structure of the cost function, denoted by  $c_i(s)$ , depending on the state  $s$  will be discussed later. We define  $\pi_{a_i}$  as the payoff for player  $i$  choosing action  $a_i$  while others choose  $a_{-i}$ . It should be noted that  $\pi_{a_i}$  is constant in  $a_{-i}$ .

For any given  $s \in S$ ,  $G_s = \langle N, (A_i, u_i^s)_{i \in N} \rangle$  is a normal form game (of state  $s \in S$ ) defined by  $N = \{0, 1, 2\}$  the set of players,  $a_i$  actions of player  $i$  with  $A_i = \{h, m, \ell\}$  and  $u_i^s : A \rightarrow \mathbb{R}$  where  $A \equiv \prod_{i=0,1,2} A_i$  the utility function of player  $i$  in state  $s \in S$  and it is defined as follows for  $i = 1, 2$ :

$$u_i^s(a) = \pi_{a_i} - c_i(s)$$

and for the regulator

$$u_0^s(a) = -(\alpha u_1^s(a) + (1 - \alpha)u_2^s(a)).$$

The repeated game  $G^\infty$  given an initial state  $s^0 \in S$  consists of infinite iterations of

$(G_s)_{s \in S}$  which is played in discrete time:  $t \in \mathbb{N}_0 \equiv \{0, 1, 2, \dots\}$  without loss of generality we let  $s^0 = (\ell\ell)$ . The action of player  $i$  in the iterated game at any stage  $t$  is denoted as  $a_i^t \in A_i$ . Let  $a^t = (a_0^t, a_1^t, a_2^t)$  be the action profile at round  $t$ . For any given  $s \in S$  associated with period  $t$ , the period  $t + 1$  state  $\tilde{s}$  consists of  $\tilde{s} = (\theta(a_1^{t-1}, a_2^{t-1}), a_0^{t-1})$  where  $a_i^t \in A_i$  for every  $i \in N$ . Note that  $\theta(a_1^t, a_2^t) \in \{h, m, \ell\}$  and  $a_0^t \in \{h, m, \ell\}$ , therefore  $\tilde{s} \in S$ .

A stage  $t$  history is a vector  $h^t = (s^0, (s^0, a^0), (s^1, a^1), \dots, (s^t, a^t))$  where  $s^0 = (\ell\ell)$  and  $s^1 = (\theta(a_1^0, a_2^0), a_0^0)$  and for any given  $s^{t-1}$ ,  $s^t = (\theta(a_1^t, a_2^t), a_0^t)$ . We symbolize the initial history by  $h^0$  being defined as the initial state  $s^0 = (\ell\ell)$ . The space of all the stage  $t$  histories is  $H_t$ . The set of all histories is the union of stage  $t$  histories  $\mathcal{H} = \cup_{t=0}^{\infty} H_t$  where  $H_0 = s^0$ . Moreover, let  $H_\infty$  be defined by  $H_\infty = \{s^0, (s^0, a^0), (s^1, a^1), \dots, (s^t, a^t), \dots\}$  with for any given  $s^{t-1}$ ,  $s^t = (\theta(a_1^{t-1}, a_2^{t-1}), a_0^{t-1})$ . We often refer to  $H_\infty$  as the set of outcomes.

For any  $\omega \in H_\infty$ , the payoff player  $i$  obtained from  $\omega$  (in period 0) is given by

$$U_i(\omega) = \liminf_{T \rightarrow \infty} \frac{1}{T} \left( \sum_{t=0}^T u_i^{s^t}(a^t) \right).$$

Clearly the use of no-discounting utility implies that for any  $t \in \mathbb{N}_0$  history  $h \in H_t$  associated with  $s^t$ , the continuation utility of player  $i$  at history  $h \in H_t$  from outcome path  $\omega \in H_\infty$  is

$$V_i^t(\omega) = \liminf_{T \rightarrow \infty} \frac{1}{T} \left( \sum_{\tau=t}^{t+T} u_i^{s^\tau}(a^\tau) \right).$$

A pure strategy for player  $i$  is a mapping  $f_i : \mathcal{H} \rightarrow A_i$  for all  $i \in N$ . The set of player  $i$ 's strategies is denoted by  $F_i$  and  $F = \prod_{i \in N} F_i$  is the joint strategy space with  $f \in F$ . We denote the *strategy induced at  $h$*  by  $f_i | h$  given a strategy  $f_i \in F_i$  and a history  $h \in \mathcal{H}$ . This strategy is defined pointwise on  $\mathcal{H} : (f_i | h)(h') = f_i(h \cdot h')$  for every  $h' \in \mathcal{H}$ . We denote  $f | h$  by  $(f_1 | h, \dots, f_n | h)$  for every  $f \in F$  and  $h \in \mathcal{H}$ .

Any strategy  $f \in F$  induces an outcome  $\omega_f \in H_\infty$  as follows:  $\omega_f^0 = (s^0, f(s^0))$  and  $\omega_f^t = (s_f^t, f(\omega_f^0, \omega_f^1, \omega_f^2, \dots, \omega_f^{t-1}))$  for any  $t \in \mathbb{N}_0$  where

$$s_f^1 = (\theta(f_1(\omega_f^0), f_2(\omega_f^0)), f_0(\omega_f^0)) \text{ and}$$

$$s_f^t = (\theta(f_1(\omega_f^0, \omega_f^1, \omega_f^2, \dots, \omega_f^{t-1}), f_2(\omega_f^0, \omega_f^1, \omega_f^2, \dots, \omega_f^{t-1})), f_0(\omega_f^0, \omega_f^1, \omega_f^2, \dots, \omega_f^{t-1})).$$

Note that we will use  $H_\infty$  to denote the set of outcome paths and define a function  $\omega : F \rightarrow H_\infty$  which gives the outcome path induced by any strategy  $f \in F$ . This defines the utility of a (pure) strategy by  $U_i(f) = U_i(\omega_f)$ .

As discussed in the introduction, we restrict attention to 1-memory public strate-



gies of the form  $f_i(h^t) = g_i(s^t)$ , for any  $t$  and  $h^t$  with  $h^t = (s^0, (s^0, a^0), (s^1, a^1), \dots, (s^t, a^t))$ . Moreover, for the purposes of this paper we consider behavioral strategies allowing individual randomization at every stage of the play. Thus, the set of (behavioral) strategies we consider are  $\sigma_i(s) : S \rightarrow \Delta(A_i)$ , while the resulting (von Neumann Morgenstern) utilities are given by the usual linear convex combination of pure strategy payoffs. Thus, for any  $\sigma(s^{t-1}) = (\sigma_i(s^{t-1}))_{i \in N}$  and for any  $s \in S$ ,

$$u_i(\sigma | s) = \sum_{a_i \in \{h, m, \ell\}} \sigma_i(s)(a_i) \cdot (\pi_{a_i} - c_i(s)) \quad i = 1, 2.$$

$u_i(\sigma | s)$  is the expected utility of player  $i = 1, 2$  under the behavioral strategy profile  $\sigma$  in the normal form game associated with state  $s$ . With a slight abuse of notation  $u_i(\sigma | s) = u_i(s)$ . Player 1's expected utility vector constituted under different states  $\mathbf{u}_1(s)$  equals to

$$(u_1(\sigma | hh), u_1(\sigma | hm), u_1(\sigma | hl), u_1(\sigma | mh), u_1(\sigma | mm), u_1(\sigma | ml), u_1(\sigma | lh), u_1(\sigma | lm), u_1(\sigma | ll))$$

and player 2's expected utility vector for every state  $\mathbf{u}_2(s)$  equals to

$$(u_2(\sigma | hh), u_2(\sigma | hm), u_2(\sigma | hl), u_2(\sigma | mh), u_2(\sigma | mm), u_2(\sigma | ml), u_2(\sigma | lh), u_2(\sigma | lm), u_2(\sigma | ll)).$$

Moreover, the expected utility of the regulator under the behavioral strategy profile  $\sigma$  for any state  $s \in S$  is equal to

$$u_0(\sigma | s) = -(\alpha u_1(\sigma | s) + (1 - \alpha)u_2(\sigma | s))$$

Given  $\sigma(s)$  and  $s \in S$ ,

$$\sigma_0(s)(h) = (\sigma_0(hh)(h), \sigma_0(hm)(h), \dots, \sigma_0(mh)(h), \dots, \sigma_0(lh)(h), \dots, \sigma_0(ll)(h))$$

denotes the conditional probabilities to announce  $a_0^t = h$  in the current round  $t$  given the state of the previous round  $s$ . For player 1 and player 2, the conditional probabilities are

$$\sigma_i(s)(a_i) = (\sigma_i(hh)(a_i), \sigma_i(hm)(a_i), \dots, \sigma_i(mh)(a_i), \dots, \sigma_i(ll)(a_i)) \quad \text{for } i = 1, 2.$$

The transition rule among the states is specified by the probabilistic strategies of all players and the resulting state space structure. The Markov transition matrix of the

repeated game which is denoted by  $\mathbf{M}$  is figured out by computing the 81 transition probabilities of players. For example, if the previous state is  $hh$ , the probability that the state transits to a new state  $hm$  will be:

$$\left( \begin{array}{c} \sigma_1(hh)(h)\sigma_2(hh)(h) + \sigma_1(hh)(h)\sigma_2(hh)(m) + \sigma_1(hh)(h)\sigma_2(hh)(\ell) + \\ \sigma_1(hh)(m)\sigma_2(hh)(h) + \sigma_1(hh)(\ell)\sigma_2(hh)(h) \end{array} \right) \sigma_0(hh)(m)$$

The first part of the multiplication denotes the probability of the observed state being  $h$ . The second part of the multiplication  $\sigma_0(hh)(m)$  indicates the probability that the regulator announces  $m$  when the given state is  $hh$ . Notice that the publicly observed state is  $h$  when either (1) both countries choose  $h$  or (2) only one of them chooses  $h$  while the other chooses  $m$  or (3) only one of them chooses  $h$  while the other chooses  $\ell$ . The derivation of the other transition probabilities from one state to another state is presented in the state transition matrix of the Markov chain  $\mathbf{M}$ .

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{21} & \cdots & M_{91} \\ M_{12} & M_{22} & \cdots & M_{92} \\ \vdots & \vdots & \cdots & \vdots \\ M_{19} & M_{29} & \cdots & M_{99} \end{bmatrix}$$

Due to space considerations, the entries of Markov chain matrix is shown explicitly in the Appendix.

#### Zero-Determinant Strategies

Let  $v_s(t)$  be the probability that the outcome of round  $t$  is  $s \in S$ . The following vector notation is:

$$\mathbf{v}(t) = (v_{hh}(t), v_{hm}(t), v_{hl}(t), v_{mh}(t), v_{mm}(t), v_{h\ell}(t), v_{\ell h}(t), v_{\ell m}(t), v_{\ell\ell}(t)).$$

We can define the limit of means distribution as

$$\mathbf{v} = (v_{hh}, v_{hm}, v_{hl}, v_{mh}, v_{mm}, v_{m\ell}, v_{\ell h}, v_{\ell m}, v_{\ell\ell})$$

where  $\mathbf{v} = \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \mathbf{v}(t)$ .

**Definition 1** Let  $\mathbf{M}$  be the Markov transition matrix. The row vector  $\mathbf{v} \in \Delta(S)$  is called a stationary probability distribution if it satisfies

$$\mathbf{v}\mathbf{M} = \mathbf{v}.$$

Let  $\mathbf{M}' \equiv \mathbf{M} - \mathbf{I}$  where  $\mathbf{I}_9$  is the  $9 \times 9$  identity matrix. Then  $\mathbf{vM} = \mathbf{v}$  becomes  $\mathbf{vM}' = \mathbf{0}$ .

**Remark 1** Every stochastic matrix has an eigenvalue equal to 1.<sup>1</sup>

This follows from  $\mathbf{M}$  being a stochastic matrix, i.e for any given  $s \in S$ , the summation of the row items deliver 1, implying that the last column can be obtained using the first 8. Since  $\mathbf{M}$  has a unit eigenvalue, the matrix  $\mathbf{M}'$  is not invertible. That is, the determinant of  $\mathbf{M}'$  must be zero.

We will employ the following Definitions and Proposition in the rest of the thesis.

**Definition 2** Let  $\mathbf{A}$  be an  $n \times n$  matrix. The determinant of the submatrix  $(n-1) \times (n-1)$  obtained by eliminating the  $i$ th row and  $j$ th column of  $\mathbf{A}$  is called the  $(i, j)$ -minor of  $\mathbf{A}$  and denoted by  $\text{minor}\mathbf{A}_{ij}$ . The scalar  $(-1)^{i+j}\text{minor}\mathbf{A}_{ij}$  is called the  $(i, j)$ -cofactor of  $\mathbf{A}$  and denoted by  $\text{cof}\mathbf{A}_{ij}$ .

**Definition 3** The adjoint  $\text{Adj}(\mathbf{A})$  of an  $n \times n$  matrix  $\mathbf{A}$  is the transpose of the cofactor matrix of  $\mathbf{A}$ ,

$$\text{Adj}(\mathbf{A}) = \begin{bmatrix} \text{cof}\mathbf{A}_{11} & \text{cof}\mathbf{A}_{12} & \cdots & \text{cof}\mathbf{A}_{1n} \\ \text{cof}\mathbf{A}_{21} & \text{cof}\mathbf{A}_{22} & \cdots & \text{cof}\mathbf{A}_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \text{cof}\mathbf{A}_{n1} & \text{cof}\mathbf{A}_{n2} & \cdots & \text{cof}\mathbf{A}_{nn} \end{bmatrix}^T.$$

For example, let

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{bmatrix}$$

be the  $3 \times 3$  matrix with nine minors and nine cofactors.

$$\text{The } \text{minor}\mathbf{A}_{12} = \begin{vmatrix} 2 & 6 \\ 1 & 8 \end{vmatrix} = 2 \times 8 - 6 \times 1 = 10.$$

$$\text{The } \text{minor}\mathbf{A}_{23} = \begin{vmatrix} 3 & 1 \\ 1 & 4 \end{vmatrix} = 3 \times 4 - 1 \times 1 = 11.$$

The corresponding cofactors are

---

<sup>1</sup>Please see Stewart (2009)

The  $\text{cof}\mathbf{A}_{12} = (-1)^{1+2}\text{minor}\mathbf{A}_{12} = (-1) \times 10 = -10$ .

The  $\text{cof}\mathbf{A}_{23} = (-1)^{2+3}\text{minor}\mathbf{A}_{23} = (-1) \times 11 = -11$ .

The adjoint matrix for  $\mathbf{A}$  will be

$$\text{Adj}(\mathbf{A}) = \begin{bmatrix} \text{cof}\mathbf{A}_{11} & \text{cof}\mathbf{A}_{12} & \text{cof}\mathbf{A}_{13} \\ \text{cof}\mathbf{A}_{21} & \text{cof}\mathbf{A}_{22} & \text{cof}\mathbf{A}_{23} \\ \text{cof}\mathbf{A}_{31} & \text{cof}\mathbf{A}_{32} & \text{cof}\mathbf{A}_{33} \end{bmatrix}^T = \begin{bmatrix} 16 & -24 & 26 \\ -10 & 28 & -26 \\ 3 & -11 & 13 \end{bmatrix}.$$

**Proposition 1** Let  $\mathbf{A}$  be an  $n \times n$  matrix and  $\mathbf{I}$  be the  $n \times n$  identity matrix. Let  $\text{Adj}(\mathbf{A})$  denote the adjoint matrix of  $\mathbf{A}$ . Then  $\text{Adj}(\mathbf{A}) \mathbf{A} = \det(\mathbf{A})\mathbf{I}$ .

The adjoint matrix of  $\mathbf{M}'$  which is formed by  $M'_{ij}$  indicating the transpose of the  $(i, j)$  cofactor of  $\mathbf{M}'$  is shown as:

$$\text{Adj}(\mathbf{M}') = \begin{bmatrix} M'_{11} & M'_{21} & \cdots & M'_{91} \\ M'_{12} & M'_{22} & \cdots & M'_{92} \\ \vdots & \vdots & \cdots & \vdots \\ M'_{19} & M'_{29} & \cdots & M'_{99} \end{bmatrix}$$

By applying Cramer's rule to  $\mathbf{M}'$  and  $\text{Adj}(\mathbf{M}')$ , we get

$$\text{Adj}(\mathbf{M}')\mathbf{M}' = \det(\mathbf{M}')\mathbf{I}_9 = \mathbf{0}.$$

As stated in Li (2014), having the unique stationary probability distribution there is a unique solution to the  $\mathbf{v}\mathbf{M}' = \mathbf{0}$  up to a scalar factor. Rank of  $\mathbf{M}'$  is 8 since the last column can be written as a linear combination of the other columns. Thus,  $\text{Adj}(\mathbf{M}')$  is a nonzero matrix, so consider that the last row  $(M'_{19}, M'_{29}, \dots, M'_{99})$  of  $\text{Adj}(\mathbf{M}')$  which is guaranteed to be nonzero vector.

Notice that every row of  $\text{Adj}(\mathbf{M}')$  is proportional to the stationary distribution vector  $\mathbf{v}$  since we have  $\mathbf{v}\mathbf{M}' = \mathbf{0}$  and  $\text{Adj}(\mathbf{M}')\mathbf{M}' = \mathbf{0}$ . Hence,  $\mathbf{v} = \mu(M'_{19}, M'_{29}, \dots, M'_{99})$  for some scalar  $\mu \neq 0$ .

Next, we can end up with a formula which is the dot product of an arbitrary vector  $\mathbf{x} = (x_1, x_2, \dots, x_9)$  with the stationary distribution vector  $\mathbf{v}$  of the Markov matrix:

$$\mathbf{v} \cdot \mathbf{x} = \mu(M'_{19}x_1 + M'_{29}x_2 + \dots + M'_{99}x_9).$$

By some elementary column operations on matrix  $\mathbf{M}'$  such as adding the first and fourth columns into the seventh column, we get matrix  $\mathbf{M}''$ . The first column represents the probability that the state transits to  $hh$  given the previous stage's state for any  $s$ . The fourth column shows that the probability that the state transits to  $mh$  and as well as the seventh column gives the probability that the state moves  $\ell h$  given the previous round's state. At the end, we obtain the first column is constituted from the probabilities that the player 0 announces  $h$  emission level in the current round given the previous round's state.

$$\mathbf{M}'' = \begin{bmatrix} \sigma_0(hh)(h) - 1 & [\sigma_1(hh)(h) + \sigma_2(hh)(h)(1 - \sigma_2(hh)(h))] \sigma_0(hh)(m) & \vdots & \sigma_1(hh)(\ell) \sigma_2(hh)(\ell) \sigma_0(hh)(\ell) \\ \sigma_0(hm)(h) & [\sigma_1(hm)(h) + \sigma_2(hm)(h)(1 - \sigma_2(hm)(h))] \sigma_0(hm)(m) - 1 & \vdots & \sigma_1(hm)(\ell) \sigma_2(hm)(\ell) \sigma_0(hm)(\ell) \\ \sigma_0(h\ell)(h) & [\sigma_1(h\ell)(h) + \sigma_2(h\ell)(h)(1 - \sigma_2(h\ell)(h))] \sigma_0(h\ell)(m) & \vdots & \sigma_1(h\ell)(\ell) \sigma_2(h\ell)(\ell) \sigma_0(h\ell)(\ell) \\ \sigma_0(mh)(h) - 1 & \vdots & \vdots & \vdots \\ \sigma_0(mm)(h) & \vdots & \vdots & \vdots \\ \sigma_0(m\ell)(h) & \vdots & \vdots & \vdots \\ \sigma_0(\ell h)(h) - 1 & \vdots & \vdots & \vdots \\ \sigma_0(\ell m)(h) & \vdots & \vdots & \vdots \\ \sigma_0(\ell\ell)(h) & [\sigma_1(\ell\ell)(h) + \sigma_2(\ell\ell)(h)(1 - \sigma_2(\ell\ell)(h))] \sigma_0(\ell\ell)(m) & \vdots & \sigma_1(\ell\ell)(\ell) \sigma_2(\ell\ell)(\ell) \sigma_0(\ell\ell)(\ell) \end{bmatrix}$$

Let  $M''_{ij}$  indicating the transpose of the  $(i, j)$  cofactor of  $\mathbf{M}''$ . Since we conduct elementary column operations, the determinant of the matrix do not change. Also, notice that  $M'_{i9} = M''_{i9}$  for  $i = 1, 2, \dots, 9$  because the last column is not manipulated. If we substitute the last column of  $\mathbf{M}''$  with the transpose of an arbitrary  $\mathbf{x}$  vector and then compute the determinant of the corresponding matrix by extending along the ninth column, we obtain the relation between the determinant and the value of  $\mathbf{v} \cdot \mathbf{x}$ :

$$\det \begin{bmatrix} \sigma_0(hh)(h) - 1 & [\sigma_1(hh)(h) + \sigma_2(hh)(h)(1 - \sigma_2(hh)(h))] \sigma_0(hh)(m) & \vdots & x_1 \\ \sigma_0(hm)(h) & [\sigma_1(hm)(h) + \sigma_2(hm)(h)(1 - \sigma_2(hm)(h))] \sigma_0(hm)(m) - 1 & \vdots & x_2 \\ \sigma_0(h\ell)(h) & [\sigma_1(h\ell)(h) + \sigma_2(h\ell)(h)(1 - \sigma_2(h\ell)(h))] \sigma_0(h\ell)(m) & \vdots & x_3 \\ \sigma_0(mh)(h) - 1 & \vdots & \vdots & x_4 \\ \sigma_0(mm)(h) & \vdots & \vdots & x_5 \\ \sigma_0(m\ell)(h) & \vdots & \vdots & x_6 \\ \sigma_0(\ell h)(h) - 1 & \vdots & \vdots & x_7 \\ \sigma_0(\ell m)(h) & \vdots & \vdots & x_8 \\ \sigma_0(\ell\ell)(h) & [\sigma_1(\ell\ell)(h) + \sigma_2(\ell\ell)(h)(1 - \sigma_2(\ell\ell)(h))] \sigma_0(\ell\ell)(m) & \vdots & x_9 \end{bmatrix}$$

is equal to  $x_1 M''_{19} + x_2 M''_{29} + \dots + x_9 M''_{99} = x_1 M'_{19} + x_2 M'_{29} + \dots + x_9 M'_{99} = \frac{1}{\mu} (\mathbf{v} \cdot \mathbf{x})$ . Thus, we get the result  $\mathbf{v} \cdot \mathbf{x} \equiv D(\sigma_0, \sigma_1, \sigma_2, \mathbf{x})$ . The significant point in this determinant is that

the first column  $\tilde{\sigma}_0(s)(h)$  is equal to

$$\begin{aligned} & (\sigma_0(hh)(h) - 1, \sigma_0(hm)(h), \sigma_0(h\ell)(h), \sigma_0(mh)(h) - 1, \\ & \sigma_0(mm)(h), \sigma_0(m\ell)(h), \sigma_0(\ell h)(h) - 1, \sigma_0(\ell m)(h), \sigma_0(\ell\ell)(h)) \end{aligned}$$

and is only controlled by player 0 while the last column is directly equal to  $\mathbf{x}$ . Resulting from the formula  $\mathbf{v} \cdot \mathbf{x} \equiv D(\sigma_0, \sigma_1, \sigma_2, \mathbf{x})$ , in the stationary state player 1's normalized payoff is obtained as:

$$\begin{aligned} P_1 &= \frac{\mathbf{v} \cdot \mathbf{u}_1}{\mathbf{v} \cdot \mathbf{1}} = \frac{D(\sigma_1, \sigma_2, \sigma_0, \mathbf{u}_1)}{D(\sigma_1, \sigma_2, \sigma_0, \mathbf{1})} \\ &= \frac{(v_{hh}, v_{hm}, \dots, v_{mm}, \dots, v_{\ell\ell}) \cdot (u_1(hh), \dots, u_1(mm), \dots, u_1(\ell\ell))}{(v_{hh}, v_{hm}, \dots, v_{mm}, \dots, v_{\ell\ell}) \cdot (1, 1, \dots, 1)} \\ &= v_{hh}u_1(hh) + v_{hm}u_1(hm) + \dots + v_{\ell\ell}u_1(\ell\ell) \end{aligned}$$

where  $\mathbf{u}_1$  is the player 1's expected payoff vector and  $\mathbf{1}$  is the vector having all components 1. Similarly, player 2's normalized payoff is

$$P_2 = \frac{\mathbf{v} \cdot \mathbf{u}_2}{\mathbf{v} \cdot \mathbf{1}} = \frac{D(\sigma_1, \sigma_2, \sigma_0, \mathbf{u}_2)}{D(\sigma_1, \sigma_2, \sigma_0, \mathbf{1})}.$$

If we replace the arbitrary  $\mathbf{x}$  vector with any linear combination of player 1's and player 2's expected payoff vector  $\alpha\mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2 + \gamma\mathbf{1}$ , we acquire the following matrix

$$\begin{bmatrix} \sigma_0(hh)(h) - 1 & [\sigma_1(hh)(h) + \sigma_2(hh)(h)(1 - \sigma_2(hh)(h))] \sigma_0(hh)(m) & \vdots & \alpha u_1(hh) + (1 - \alpha)u_2(hh) + \gamma \\ \sigma_0(hm)(h) & [\sigma_1(hm)(h) + \sigma_2(hm)(h)(1 - \sigma_2(hm)(h))] \sigma_0(hm)(m) - 1 & \vdots & \alpha u_1(hm) + (1 - \alpha)u_2(hm) + \gamma \\ \sigma_0(h\ell)(h) & [\sigma_1(h\ell)(h) + \sigma_2(h\ell)(h)(1 - \sigma_2(h\ell)(h))] \sigma_0(h\ell)(m) & \vdots & \alpha u_1(h\ell) + (1 - \alpha)u_2(h\ell) + \gamma \\ \sigma_0(mh)(h) - 1 & \vdots & \vdots & \alpha u_1(mh) + (1 - \alpha)u_2(mh) + \gamma \\ \sigma_0(mm)(h) & \vdots & \vdots & \alpha u_1(mm) + (1 - \alpha)u_2(mm) + \gamma \\ \sigma_0(m\ell)(h) & \vdots & \vdots & \alpha u_1(m\ell) + (1 - \alpha)u_2(m\ell) + \gamma \\ \sigma_0(\ell h)(h) - 1 & \vdots & \vdots & \alpha u_1(\ell h) + (1 - \alpha)u_2(\ell h) + \gamma \\ \sigma_0(\ell m)(h) & \vdots & \vdots & \alpha u_1(\ell m) + (1 - \alpha)u_2(\ell m) + \gamma \\ \sigma_0(\ell\ell)(h) & [\sigma_1(\ell\ell)(h) + \sigma_2(\ell\ell)(h)(1 - \sigma_2(\ell\ell)(h))] \sigma_0(\ell\ell)(m) & \vdots & \alpha u_1(\ell\ell) + (1 - \alpha)u_2(\ell\ell) + \gamma \end{bmatrix}$$

Since the normalized payoffs of players' are linearly contingent on their own expected stage payoff vectors, any linear combination of these normalized payoffs of the

two players with coefficients  $\alpha$  and  $\gamma$  is derived as:

$$\begin{aligned}\alpha P_1 + (1 - \alpha)P_2 + \gamma &= \frac{\mathbf{v} \cdot (\alpha \mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2 + \gamma \mathbf{1})}{\mathbf{v} \cdot \mathbf{1}} \\ &= \frac{D(\sigma_0, \sigma_1, \sigma_2, \alpha \mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2 + \gamma \mathbf{1})}{D(\sigma_0, \sigma_1, \sigma_2, \mathbf{1})} \\ &= \sum_s v_s [\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma].\end{aligned}$$

For some values of  $0 < \alpha < 1$ ,  $\gamma$  and  $\rho$  if the regulator can set his strategy  $\sigma_0(s)(h)$  satisfying  $\tilde{\sigma}_0(s^{t-1})(h) = \rho(\alpha \mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2 + \gamma \mathbf{1})$ , then regardless of the two players' strategies, a linear relation between player 1's and player 2's payoff scores will be established :

$$\alpha P_1 + (1 - \alpha)P_2 + \gamma = 0.$$

$D(\sigma_0, \sigma_1, \sigma_2, (\alpha \mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2 + \gamma \mathbf{1}))$  has its first column fully controlled by player 0. When player 0 chooses a satisfying strategy  $\tilde{\sigma}_0(s^{t-1})(h) = \rho(\alpha \mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2 + \gamma \mathbf{1})$ , then the first and the last column will be proportional to constant  $\rho$ . If a matrix has two proportional columns or rows, its determinant is zero. Thus  $D(\sigma_0, \sigma_1, \sigma_2, (\alpha \mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2 + \gamma \mathbf{1})) = 0$ , irrespective of the values of the other columns which gives that

$$(\alpha P_1 + (1 - \alpha)P_2 + \gamma) = \frac{D(\sigma_0, \sigma_1, \sigma_2, (\alpha \mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2 + \gamma \mathbf{1}))}{D(\sigma_0, \sigma_1, \sigma_2, \mathbf{1})} = 0.$$

Therefore, for any given the other players' Markovian behavior, it is possible that player 0 can come up with a stationary behavioral strategy (1-memory public strategy) such that all players' long-run payoffs are fixed to some number. Such strategies of player 0 are called **Zero-Determinant (ZD)** strategies.

The following construction will be used in the existence result:

Given  $\alpha$ ,  $((\pi_{a_i})_{a_i \in A})_{i \in N}$ ,  $((c_i(s))_{s \in S})_{i \in N}$  and  $\rho < 0$  define

$$\gamma_s = \begin{cases} -(\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \in \{hh, mh, lh\} \\ \frac{1}{\rho} - (\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \notin \{hh, mh, lh\} \end{cases}$$

and

$$\bar{\gamma}_s = \begin{cases} \frac{-1}{\rho} - (\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \in \{hh, mh, lh\} \\ -(\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \notin \{hh, mh, lh\} \end{cases}$$

**Condition E** Given  $\alpha$ ,  $((\pi_{a_i})_{a_i \in A})_{i \in N}$ ,  $((c_i(s))_{s \in S})_{i \in N}$ ,  $(\sigma_i(s))_{i=1,2}$  and  $\rho < 0$ . Condition E holds if

1.  $\max_{s \in S} \gamma_s \leq \min_{s \in S} \bar{\gamma}_s$  and
2.  $\gamma \in \left[ \max_{s \in S} \gamma_s, \min_{s \in S} \bar{\gamma}_s \right]$ .

Below we show an existence result for ZD strategies.

**Proposition 2** Let  $(\alpha, \rho, \gamma, ((\pi_{a_i})_{a_i \in A})_{i \in N}, ((c_i(s))_{s \in S})_{i \in N}, (\sigma_i(s))_{i=1,2})$  satisfy condition E. Then player 0 possesses a ZD strategy.

**Proof.** Let  $\tilde{\sigma}_0(s)(h)$  be the strategy that the regulator chooses such that  $\tilde{\sigma}_0(s)(h) = \rho(\alpha \mathbf{u}_1 + (1 - \alpha) \mathbf{u}_2 + \gamma \mathbf{1})$  where  $\rho \neq 0$  leading to the following system of linear equations:

$$\begin{aligned}
\sigma_0(hh)(h) - 1 &= \rho(\alpha u_1(hh) + (1 - \alpha)u_2(hh) + \gamma) \\
\sigma_0(hm)(h) &= \rho(\alpha u_1(hm) + (1 - \alpha)u_2(hm) + \gamma) \\
\sigma_0(h\ell)(h) &= \rho(\alpha u_1(h\ell) + (1 - \alpha)u_2(h\ell) + \gamma) \\
\sigma_0(mh)(h) - 1 &= \rho(\alpha u_1(mh) + (1 - \alpha)u_2(mh) + \gamma) \\
\sigma_0(mm)(h) &= \rho(\alpha u_1(mm) + (1 - \alpha)u_2(mm) + \gamma) \\
\sigma_0(m\ell)(h) &= \rho(\alpha u_1(m\ell) + (1 - \alpha)u_2(m\ell) + \gamma) \\
\sigma_0(\ell h)(h) - 1 &= \rho(\alpha u_1(\ell h) + (1 - \alpha)u_2(\ell h) + \gamma) \\
\sigma_0(\ell m)(h) &= \rho(\alpha u_1(\ell m) + (1 - \alpha)u_2(\ell m) + \gamma) \\
\sigma_0(\ell \ell)(h) &= \rho(\alpha u_1(\ell \ell) + (1 - \alpha)u_2(\ell \ell) + \gamma)
\end{aligned} \tag{1}$$

It suffices to show that  $(\sigma_0(s)(h))_{s \in S}$  satisfies the feasibility condition;  $0 \leq \sigma_0(s)(h) \leq 1$  in order to be a ZD strategy given  $\alpha$ ,  $((\pi_{a_i})_{a_i \in A})_{i \in N}$ ,  $((c_i(s))_{s \in S})_{i \in N}$  and  $(\sigma_i(s))_{i=1,2}$ .

That is, we need to show that

$$\sigma_0(s)(h) = \begin{cases} \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) + 1 & \text{if } s \in \{hh, mh, \ell h\} \\ \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) & \text{if } s \notin \{hh, mh, \ell h\} \end{cases} \text{ is in } [0, 1]$$

for all  $s \in S$ . This leads to



$$\begin{aligned}
0 &\leq \rho(\alpha u_1(hh) + (1 - \alpha)u_2(hh) + \gamma) + 1 \leq 1 \\
0 &\leq \rho(\alpha u_1(hm) + (1 - \alpha)u_2(hm) + \gamma) \leq 1 \\
0 &\leq \rho(\alpha u_1(h\ell) + (1 - \alpha)u_2(h\ell) + \gamma) \leq 1 \\
0 &\leq \rho(\alpha u_1(mh) + (1 - \alpha)u_2(mh) + \gamma) + 1 \leq 1 \\
0 &\leq \rho(\alpha u_1(mm) + (1 - \alpha)u_2(mm) + \gamma) \leq 1 \\
0 &\leq \rho(\alpha u_1(m\ell) + (1 - \alpha)u_2(m\ell) + \gamma) \leq 1 \\
0 &\leq \rho(\alpha u_1(\ell h) + (1 - \alpha)u_2(\ell h) + \gamma) + 1 \leq 1 \\
0 &\leq \rho(\alpha u_1(\ell m) + (1 - \alpha)u_2(\ell m) + \gamma) \leq 1 \\
0 &\leq \rho(\alpha u_1(\ell\ell) + (1 - \alpha)u_2(\ell\ell) + \gamma) \leq 1
\end{aligned} \tag{2}$$

Following from the condition E, given  $\alpha$ ,  $((\pi_{a_i})_{a_i \in A})_{i \in N}$ ,  $((c_i(s))_{s \in S})_{i \in N}$ ,  $(\sigma_i(s))_{i=1,2}$ , and  $\rho < 0$ ,  $\gamma \in \left[ \max_{s \in S} \underline{\gamma}_s, \min_{s \in S} \bar{\gamma}_s \right]$  and  $\max_{s \in S} \underline{\gamma}_s \leq \min_{s \in S} \bar{\gamma}_s$  where

$$\underline{\gamma}_s = \begin{cases} -(\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \in \{hh, mh, \ell h\} \\ \frac{1}{\rho} - (\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \notin \{hh, mh, \ell h\} \end{cases}$$

and

$$\bar{\gamma}_s = \begin{cases} \frac{-1}{\rho} - (\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \in \{hh, mh, \ell h\} \\ -(\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \notin \{hh, mh, \ell h\} \end{cases}$$

all terms are led to be between 0 and 1. This follows from:

Let  $s \in \{hh, mh, \ell h\}$ . Then,  $\rho < 0$  from Condition E and  $\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma \geq 0$  since otherwise  $\rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) + 1 > 1$  will be a contradiction.

As  $\gamma \leq \bar{\gamma}_s$  we have

$$\begin{aligned}
\rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) + 1 &\geq \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \bar{\gamma}_s) + 1 \\
&= \rho[\alpha u_1(s) + (1 - \alpha)u_2(s) - \frac{1}{\rho} - (\alpha u_1(s) + (1 - \alpha)u_2(s))] + 1 \\
&= 0
\end{aligned}$$

Therefore, we get  $\rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) + 1 \geq 0$ .

As  $\gamma \geq \underline{\gamma}_s$  we have

$$\begin{aligned}
\rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) + 1 &\leq \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \underline{\gamma}_s) + 1 \\
&= \rho[\alpha u_1(s) + (1 - \alpha)u_2(s) - (\alpha u_1(s) + (1 - \alpha)u_2(s))] + 1 \\
&= 1
\end{aligned}$$

Thus,  $\rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) + 1 \leq 1$ .

Hence,  $0 \leq \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) + 1 \leq 1$ .

Let  $s \notin \{hh, mh, \ell h\}$  and recall that  $\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma \geq 0$ .

Since  $\gamma \leq \bar{\gamma}_s$  we have

$$\begin{aligned} \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) &\geq \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \bar{\gamma}_s) \\ &= \rho[\alpha u_1(s) + (1 - \alpha)u_2(s) - (\alpha u_1(s) + (1 - \alpha)u_2(s))] \\ &= 0 \end{aligned}$$

Therefore,  $\rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) \geq 0$ .

As  $\gamma \geq \underline{\gamma}_s$  we have

$$\begin{aligned} \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) &\leq \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \underline{\gamma}_s) \\ &= \rho[\alpha u_1(s) + (1 - \alpha)u_2(s) + \frac{1}{\rho} - (\alpha u_1(s) + (1 - \alpha)u_2(s))] \\ &= 1 \end{aligned}$$

Thus,  $\rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) \leq 1$ .

Hence, we obtain  $0 \leq \rho(\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma) \leq 1$ .

So,  $\sigma_0(s)(h)$  satisfying (2) will be a zero-determinant strategy. ■

We show that this condition is critical in the question about the existence of ZD strategies. All components of strategy  $\sigma_0(s)(h)$  of the regulator have to be between 0 and 1 to be feasible. Put differently, for some  $\alpha, \rho$  and  $\gamma$  there is no feasible zero-determinant strategy for the regulator. For example, for  $\alpha \in (0, 1)$  and  $\gamma$ , if we take the difference of the first and the last equations, we get

$$\sigma_0(hh)(h) - \sigma_0(\ell\ell)(h) - 1 = \rho(\alpha(u_1(hh) - u_1(\ell\ell)) + (1 - \alpha)(u_2(hh) - u_2(\ell\ell))).$$

In the case that  $u_i(hh) - u_i(\ell\ell) < 0$  and  $\rho < 0$  for  $i = 1, 2$  the right hand side of the equation is positive but the left hand side can be less than and equal to 0 since the difference of  $\sigma_0(hh)(h) - \sigma_0(\ell\ell)(h)$  can be 1 at maximum. Thus, there is no zero-determinant strategies when  $u_i(hh) - u_i(\ell\ell) < 0$  and  $\rho < 0$  for  $i = 1, 2$ .

By enforcing unilaterally a linear relation between player 1 and player 2 through its zero-determinant strategy  $\sigma_0(s)(h)$ , the regulator obtains the long run average payoff  $P_0 = -(\alpha P_1 + (1 - \alpha)P_2)$  which is equal to  $\gamma$ .

The following is needed for further analysis. Define

$$\Lambda = \{(\sigma_1, \sigma_2) \in \mathbb{R}_+^{27 \times 27} : \exists \gamma \text{ with } (\alpha, (\pi_{a_i}), (c_i(s)), \gamma, \sigma_1, \sigma_2) \text{ sustains (2)}\}.$$

Moreover, let

$$\Lambda(\gamma) = \{(\sigma_1, \sigma_2) \in \mathbb{R}_+^{27 \times 27} : (\alpha, (\pi_{a_i}), (c_i(s)), \gamma, \sigma_1, \sigma_2) \text{ sustains (2)}\}.$$

**Condition EE**  $\Lambda \neq \emptyset$ .

Notice that Condition EE implies Condition E.

## 2.1. The Optimal Zero-Determinant Strategy

Naturally, considering the optimal ZD payoff for player 0 emerges as a research question.

**Claim 1**  $\Lambda$  is compact.

**Proof.** Clearly  $\Lambda$  is bounded by  $[0, 1]^{54}$ . So it suffices to prove that  $\Lambda$  is closed.

Suppose  $(\sigma_1^n, \sigma_2^n)$  is a sequence in  $\Lambda$  obtained with  $(\alpha, (\pi_{a_i}), (c_i(s)))$  and  $\forall n \in \mathbb{N}$ . Therefore, for all  $(\sigma_1^n, \sigma_2^n) \in \Lambda$  there exists  $\gamma^n$  such that  $(\alpha, \pi_{a_i}, c_i(s), \gamma^n, \sigma_1^n, \sigma_2^n)$  satisfies (1), which consists of 9 equalities. While  $(\sigma_1^n, \sigma_2^n) \rightarrow (\sigma_1, \sigma_2)$  and  $\gamma^n \rightarrow \gamma$ , we need to show  $(\alpha, \pi_{a_i}, c_i(s), \gamma, \sigma_1, \sigma_2)$  satisfies (1). This holds trivially because (1) only involves continuous linear equalities. Therefore,  $\Lambda$  is compact. ■

By solving the following maximization problem for  $\gamma_{\sigma_1, \sigma_2}^*$  which is the regulator's highest ever possible payoff under ZD strategy  $\sigma_0^*$ , the regulator determines the optimal ZD strategy for himself:

$$\begin{aligned} & \text{Max } \gamma_{\sigma_1, \sigma_2}^* \\ & \text{s.t. } (\sigma_1, \sigma_2) \in \Lambda \end{aligned}$$

Note that  $\gamma_{\sigma_1, \sigma_2}^* = \min_{s \in S} \bar{\gamma}_s(\sigma_1, \sigma_2)$  where

$$\bar{\gamma}_s = \begin{cases} \frac{-1}{\rho} - (\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \in \{hh, mh, lh\} \\ -(\alpha u_1(s) + (1 - \alpha)u_2(s)) & \text{if } s \notin \{hh, mh, lh\} \end{cases}$$

Therefore,  $\gamma_{\sigma_1, \sigma_2}^*$  is a continuous function on  $\Lambda$  and since  $\Lambda$  is compact, there exists a maximizer to above problem.

In what follows we will show that Press and Dyson's result can be extended to this setting under a mild assumption needed to guarantee the existence of ZD strategies. By using associated ZD strategies  $\sigma_0$ , player 0 can fix all other players to the linear combination of payoffs  $\alpha P_1 + (1 - \alpha)P_2$  and  $\gamma$  (for himself) no matter which strategy  $\sigma_1, \sigma_2$  players 1 and 2 choose in  $\Lambda(\gamma)$ . So, we have a counterpart of Press and Dyson's result for our setting.

Given  $\alpha$ ,  $(\pi_{a_i})$ , and  $(c_i(s))$  there exists such  $\sigma_1, \sigma_2$  and  $\gamma$  satisfying Condition E and therefore, there exists a ZD strategy of player 0 for given  $\sigma_1, \sigma_2$ ,  $\gamma$  and  $\alpha$ ,  $(\pi_{a_i})$ ,  $(c_i(s))$ . In other words,  $(\sigma_1, \sigma_2) \in \Lambda(\gamma)$ . For this strategy, player 0's payoff is fixed to  $\gamma$  and the linear combination of other players' payoffs  $\alpha P_1 + (1 - \alpha)P_2$  is fixed to  $-\gamma$ . If just  $\tilde{\sigma}_1, \tilde{\sigma}_2 \in \Lambda(\gamma)$  changes and with same  $\gamma$  satisfying the (2), Condition E holds with given  $\tilde{\sigma}_1, \tilde{\sigma}_2, \gamma$  and same  $\alpha$ ,  $(\pi_{a_i})$ ,  $(c_i(s))$ . Thus, there exists a ZD strategy  $\tilde{\sigma}_0$  giving payoff  $-(\alpha P_1 + (1 - \alpha)P_2)$  and  $\gamma$ . Therefore, no matter what  $\sigma_1, \sigma_2 \in \Lambda(\gamma)$  is player 0 obtains  $\gamma$  while player 1 and player 2's convex combination of payoffs is given by  $(\alpha P_1 + (1 - \alpha)P_2)$ .

In what follows, we propose an intuitive cost structure which will bring about the emergence of a "simple" and "intuitive" ZD strategy for the regulator.

## 2.2. An Intuitive Cost Structure

We now move forward to our analysis by presenting a special cost structure in order to elaborate on a particular simple ZD strategy for the regulator. Naturally, costs  $(c_i(s))_{i=1,2}$  for all  $s \in S$  are determined by the regulator.

In this section, assume that Condition EE holds and player 1 and 2 are restricted to choose mixed 1-memory public behavior in  $\Lambda$ .

For any  $s \in S$  with  $s, s' \notin \{hh, mh, lh\}$ , we may set  $(c_i(s))_{i=1,2}$  such that  $u_i(s) = u_i(s') = \tilde{u}_i$   $i = 1, 2$ . This follows from

$$1. u_i(hm) = u_i(hl) \Rightarrow \sum_a \sigma_i(hm)(a) \cdot (\pi_{a_i} - c_i(hm)) = \sum_a \sigma_i(hl)(a) \cdot (\pi_{a_i} - c_i(hl))$$

2.  $u_i(hm) = u_i(mm) \Rightarrow \sum_a \sigma_i(hm)(a) \cdot (\pi_{a_i} - c_i(hm)) = \sum_a \sigma_i(mm)(a) \cdot (\pi_{a_i} - c_i(mm))$
3.  $u_i(hm) = u_i(m\ell) \Rightarrow \sum_a \sigma_i(hm)(a) \cdot (\pi_{a_i} - c_i(hm)) = \sum_a \sigma_i(m\ell)(a) \cdot (\pi_{a_i} - c_i(m\ell))$
4.  $u_i(hm) = u_i(\ell m) \Rightarrow \sum_a \sigma_i(hm)(a) \cdot (\pi_{a_i} - c_i(hm)) = \sum_a \sigma_i(\ell m)(a) \cdot (\pi_{a_i} - c_i(\ell m))$
5.  $u_i(hm) = u_i(\ell\ell) \Rightarrow \sum_a \sigma_i(hm)(a) \cdot (\pi_{a_i} - c_i(hm)) = \sum_a \sigma_i(\ell\ell)(a) \cdot (\pi_{a_i} - c_i(\ell\ell))$

From these equations, it is understood that  $c_i(hm)$ ,  $c_i(h\ell)$ ,  $c_i(mm)$ ,  $c_i(m\ell)$ ,  $c_i(\ell m)$ , and  $c_i(\ell\ell)$  are unknowns for given the strategies of player  $i$ 's  $\sigma_i(s)(a)$ . Now, there are six unknowns and five equations so there can be infinitely many solutions. For simplicity, we take  $c_i(\ell\ell) = 0$  because the regulator intends to encourage the countries to produce less emission. By putting  $c_i(\ell\ell) = 0$  into equation 5,

$$u_i(hm) = u_i(\ell\ell) \Rightarrow \sum_a \sigma_i(hm)(a) \cdot (\pi_{a_i} - c_i(hm)) = \sum_a \sigma_i(\ell\ell)(a) \cdot \pi_{a_i}$$

Then, we find that

$$c_i(hm) = \frac{\sum_a [\sigma_i(hm)(a) - \sigma_i(\ell\ell)(a)] \cdot \pi_{a_i}}{\sum_a \sigma_i(hm)(a)}.$$

By putting  $c_i(hm)$  into the other four equation one by one, we acquire the costs for the states  $\ell m, m\ell, mm$  and  $h\ell$ .

$$c_i(\ell m) = \frac{\sum_a [\sigma_i(\ell m)(a) - \sigma_i(\ell\ell)(a)] \cdot \pi_{a_i}}{\sum_a \sigma_i(\ell m)(a)}$$

$$c_i(m\ell) = \frac{\sum_a [\sigma_i(m\ell)(a) - \sigma_i(\ell\ell)(a)] \cdot \pi_{a_i}}{\sum_a \sigma_i(m\ell)(a)}$$

$$c_i(mm) = \frac{\sum_a [\sigma_i(mm)(a) - \sigma_i(\ell\ell)(a)] \cdot \pi_{a_i}}{\sum_a \sigma_i(mm)(a)}$$

$$c_i(h\ell) = \frac{\sum_a [\sigma_i(h\ell)(a) - \sigma_i(\ell\ell)(a)] \cdot \pi_{a_i}}{\sum_a \sigma_i(h\ell)(a)}$$

Therefore,  $c_i(s) = \frac{\sum_a [\sigma_i(s)(a) - \sigma_i(\ell\ell)(a)] \cdot \pi_{a_i}}{\sum_a \sigma_i(s)(a)}$  for  $s \in \{hm, h\ell, mm, m\ell, \ell m\}$  and  $c_i(\ell\ell) = 0$ .

After obtaining costs we can result that  $u_i(s) = u_i(s') = \tilde{u}_i = \sum_a \sigma_i(\ell\ell)(a) \cdot \pi_{a_i}$  for any  $s, s' \notin \{hh, mh, \ell h\}$ .

Similarly, for any  $s, s' \in \{hh, mh, \ell h\}$  we may fix  $(c_i(s))_{i=1,2}$  such that  $u_i(s) = u_i(s') = \bar{u}_i$   $i = 1, 2$ . This follows from:

1.  $u_i(hh) = u_i(mh) \Rightarrow \sum_a \sigma_i(hh)(a) \cdot (\pi_{a_i} - c_i(hh)) = \sum_a \sigma_i(mh)(a) \cdot (\pi_{a_i} - c_i(mh))$
2.  $u_i(hh) = u_i(\ell h) \Rightarrow \sum_a \sigma_i(hh)(a) \cdot (\pi_{a_i} - c_i(hh)) = \sum_a \sigma_i(\ell h)(a) \cdot (\pi_{a_i} - c_i(\ell h))$

For this case,  $c_i(hh)$ ,  $c_i(mh)$  and  $c_i(\ell h)$  are unknowns given the strategies of player  $i$ 's  $\sigma_i(s)(a)$ . Once again, infinitely many solutions can be found as there are three unknowns and two equations. To make the problem easily solvable, we can minimize the cost of the state which takes place while the publicly observed level is low and the announced target level is high,  $c_i(\ell h)$ . The regulator prefers to impose lesser costs for the countries in the case that the publicly observed level is less than the announced target level while the states are  $hh$ ,  $mh$  and  $\ell h$ . By decreasing the cost of this state, the regulator aims that the countries have an incentive to release less carbon emissions. The regulator solves the following minimization problem:

$$\begin{aligned} & \text{Min } c_i(\ell h) \\ & \text{subject to } \begin{cases} \sigma_0(s)(h) - 1 = \rho [\alpha U_1(s) + (1 - \alpha)U_2(s) + \gamma] \quad \forall s \in \{hh, mh, \ell h\} \\ \sigma_0(s)(h) = \rho [\alpha U_1(s) + (1 - \alpha)U_2(s) + \gamma] \quad \forall s \notin \{hh, mh, \ell h\} \\ \gamma \text{ such that } \sum_s v_s [\alpha U_1(s) + (1 - \alpha)U_2(s) + \gamma] = 0 \end{cases} \end{aligned}$$

With the minimum value of  $c_i(\ell h)$ , we can find the other costs, namely  $c_i(hh)$  and  $c_i(mh)$ . Then, we can construct the utility functions via defined cost structure.

$$\text{Thus, } c_i(s) \text{ is identified such that } \begin{cases} u_i(s) = \bar{u}_i \text{ for all } s \in \{hh, mh, \ell h\} \\ u_i(s) = \tilde{u}_i \text{ for all } s \notin \{hh, mh, \ell h\} \end{cases}$$

So by utilizing this cost structure (without off-setting Condition EE) we guarantee that in the stage game player  $i$  obtains the same expected utility level  $\bar{u}_i$  for any state  $s$  involving the announcement of  $h$  by the regulator and separately for another utility level  $\tilde{u}_i$  for any state  $s$  not involving the announcement of  $h$  by the regulator.

Next, we can further construct the payoff of the regulator by means of the recently formed cost structure. Since the uniquely determined  $v_s$  depends on  $\sigma_0(s)(h)$ , we can

write the general equation as

$$\begin{aligned} \sum_s v_s^{\sigma_0} [\alpha u_1(s) + (1 - \alpha)u_2(s) + \gamma] &= 0 \\ \sum_s v_s^{\sigma_0} (\alpha u_1(s) + (1 - \alpha)u_2(s)) + \underbrace{\gamma \sum_s v_s^{\sigma_0}}_1 &= 0 \\ \sum_s v_s^{\sigma_0} (\alpha u_1(s) + (1 - \alpha)u_2(s)) &= -\gamma. \end{aligned}$$

Let  $s \in \{hh, mh, lh\}$  and  $s' \notin \{hh, mh, lh\}$ . Thus, we replace  $u_i(s)$  and  $u_i(s')$  with  $\bar{u}_i$  and  $\tilde{u}_i$ , respectively and we obtain the regulator's long-run payoff as

$$\sum_{s \in \{hh, mh, lh\}} v_s^{\sigma_0} (\alpha \bar{u}_1 + (1 - \alpha) \bar{u}_2) + \sum_{s' \notin \{hh, mh, lh\}} v_s^{\sigma_0} (\alpha \tilde{u}_1 + (1 - \alpha) \tilde{u}_2) = -\gamma.$$

### 3. Concluding Remarks

Press and Dyson have uncovered a significant mathematical feature of iterated prisoner's dilemma and given a different direction to iterated games. Even though the study of ZD strategies brings a new perspective, there are more points waiting to be discovered.

In this thesis, we consider a repeated game of regulation of carbon emissions with ZD strategies utilizing 1-memory strategies. By imposing existence or feasibility conditions of ZD strategies, the regulator is able to unilaterally derive a linear relation between the countries' payoffs. Therefore, the regulator can pin down the probability weighted average of their pure strategy payoffs to a fixed value or guarantee that his long-run payoff is negative of that fixed value. However, since the return of the regulator is bounded from below and above due to the existence condition, the regulator cannot set his payoff to any number.

We continue our analysis by searching the optimal ZD strategy for the regulator which also gives the optimal ZD payoff for him. Then, we propose a method to define the maximum of the optimal payoff. The regulator needs to solve the maximization problem so as to derive the highest possible payoff under ZD strategies.

As a further step, we put forward our study by providing a special and simple cost structure which states that countries obtain the same expected utilities when the regulator announces the target emission level as high and they also obtain the same expected utilities being different from the former when the announcement is not high. Depending on that intuitive cost structure the regulator can easily employ an uncomplicated ZD strategy and attain a maximal ZD payoff. As a final remark, we can advance the analysis by determining how far the optimal payoff is the upper limit, which deserves further researches.



## References

- Akin, E. (2013). Good strategies for the iterated prisoner's dilemma. *arXiv preprint arXiv, v2:1211.0969*.
- Al Daoud, A., Kesidis, G., and Liebeherr, J. (2014, November). Zero-determinant strategies: A game-theoretic approach for sharing licensed spectrum bands. *IEEE Journal on Selected Areas in Communications*, 32(11), 2297-2308. doi: 10.1109/JSAC.2014.141126
- Aumann, R. (1981). Survey of repeated games. *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, 4, 11-42.
- Aumann, R., and Shapley, L. S. (1994). Long-term competition – a game-theoretic analysis. *Essays in Game Theory in Honor of Michael Maschler*.
- Barlo, M., Carmona, G., and Sabourian, H. (2009, January). Repeated games with one-memory. *Journal of Economic Theory*, 144(1), 312-336.
- Boerlijst, M. C., Nowak, M. A., and Sigmund, K. (1997). Equal pay for all prisoners. *The American Mathematical Monthly*, 104(4), 303-305. Retrieved from <http://www.jstor.org/stable/2974578>
- Chen, J., and Zinger, A. (2014). The robustness of zero-determinant strategies in iterated prisoner's dilemma games. *Journal of Theoretical Biology*, 357, 46 - 54.
- Fudenberg, D., and Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3), 533-554.
- Guo, J. (2014). Zero-determinant strategies in iterated multi-strategy games. *CoRR, abs/1409.1786*. Retrieved from <http://arxiv.org/abs/1409.1786>
- Hao, D., Rong, Z., and Zhou, T. (2014). Zero-determinant strategies in noisy repeated games. *CoRR, abs/1408.5208*. Retrieved from <http://arxiv.org/abs/1408.5208>
- He, X., Dai, H., Ning, P., and Dutta, R. (2016, March). Zero-determinant strategies for multi-player multi-action iterated games. *IEEE Signal Processing Letters*, 23(3), 311-315.
- Hilbe, C., Nowak, M. A., and Sigmund, K. (2013). Evolution of extortion in iterated prisoner's dilemma games. *Proceedings of the National Academy of Sciences*, 110(17),

6913-6918.

- Hilbe, C., Traulsen, A., and Sigmund, K. (2015). Partners or rivals? Strategies for the iterated prisoner's dilemma. *Games and Economic Behavior*, 92(C), 41-52.
- Hilbe, C., Wu, B., Traulsen, A., and Nowak, M. A. (2014). Cooperation and control in multiplayer social dilemmas. *Proceedings of the National Academy of Sciences*, 111(46), 16425-16430.
- Li, S. (2014). Strategies in the stochastic iterated prisoner's dilemma. *REU Papers*.
- MacKenzie, I. A., Hanley, N., and Kornienko, T. (2008). A permit allocation contest for a tradable pollution permit market. *Working Paper*, 08(82).
- Pan, L., Hao, D., Rong, Z., and Zhou, T. (2015). Zero-determinant strategies in the iterated public goods game. *Scientific Reports*, 5.
- Press, W. H., and Dyson, F. J. (2012). Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26), 10409-10413.
- Roemheld, L. (2013). Evolutionary extortion and mischief: Zero determinant strategies in iterated 2x2 games. *CoRR*, abs/1308.2576.
- Sigmund, K. (2010). *The calculus of selfishness*. Princeton: Princeton University Press.
- Stewart, W. J. (2009). *Probability, markov chains, queues, and simulation: the mathematical basis of performance modeling*. Princeton: Princeton University Press.
- Tang, H., and Song, G. (2014). Research on behavior of regulators and enterprises about carbon emissions based on game theory. *Open Automation and Control Systems Journal*, 6, 56-61.

# Appendices

## A. Entries of the Markov Chain Matrix

The entries of the Markov Chain Matrix  $\mathbf{M}$  are:

$$\begin{aligned}
M_{11} &= \sigma_0(hh)(h)[\sigma_1(hh)(h)\sigma_2(hh)(h) + \sigma_1(hh)(h)\sigma_2(hh)(m) \\
&\quad + \sigma_1(hh)(h)\sigma_2(hh)(\ell) + \sigma_1(hh)(m)\sigma_2(hh)(h) + \sigma_1(hh)(\ell)\sigma_2(hh)(h)] \\
M_{12} &= \sigma_0(hm)(h)[\sigma_1(hm)(h)\sigma_2(hm)(h) + \sigma_1(hm)(h)\sigma_2(hm)(m) \\
&\quad + \sigma_1(hm)(h)\sigma_2(hm)(\ell) + \sigma_1(hm)(m)\sigma_2(hm)(h) + \sigma_1(hm)(\ell)\sigma_2(hm)(h)] \\
M_{13} &= \sigma_0(h\ell)(h)[\sigma_1(h\ell)(h)\sigma_2(h\ell)(h) + \sigma_1(h\ell)(h)\sigma_2(h\ell)(m) \\
&\quad + \sigma_1(h\ell)(h)\sigma_2(h\ell)(\ell) + \sigma_1(h\ell)(m)\sigma_2(h\ell)(h) + \sigma_1(h\ell)(\ell)\sigma_2(h\ell)(h)] \\
M_{14} &= \sigma_0(mh)(h)[\sigma_1(mh)(h)\sigma_2(mh)(h) + \sigma_1(mh)(h)\sigma_2(mh)(m) \\
&\quad + \sigma_1(mh)(h)\sigma_2(mh)(\ell) + \sigma_1(mh)(m)\sigma_2(mh)(h) + \sigma_1(mh)(\ell)\sigma_2(mh)(h)] \\
M_{15} &= \sigma_0(mm)(h)[\sigma_1(mm)(h)\sigma_2(mm)(h) + \sigma_1(mm)(h)\sigma_2(mm)(m) \\
&\quad + \sigma_1(mm)(h)\sigma_2(mm)(\ell) + \sigma_1(mm)(m)\sigma_2(mm)(h) + \sigma_1(mm)(\ell)\sigma_2(mm)(h)] \\
M_{16} &= \sigma_0(m\ell)(h)[\sigma_1(m\ell)(h)\sigma_2(m\ell)(h) + \sigma_1(m\ell)(h)\sigma_2(m\ell)(m) \\
&\quad + \sigma_1(m\ell)(h)\sigma_2(m\ell)(\ell) + \sigma_1(m\ell)(m)\sigma_2(m\ell)(h) + \sigma_1(m\ell)(\ell)\sigma_2(m\ell)(h)] \\
M_{17} &= \sigma_0(\ell h)(h)[\sigma_1(\ell h)(h)\sigma_2(\ell h)(h) + \sigma_1(\ell h)(h)\sigma_2(\ell h)(m) \\
&\quad + \sigma_1(\ell h)(h)\sigma_2(\ell h)(\ell) + \sigma_1(\ell h)(m)\sigma_2(\ell h)(h) + \sigma_1(\ell h)(\ell)\sigma_2(\ell h)(h)] \\
M_{18} &= \sigma_0(\ell m)(h)[\sigma_1(\ell m)(h)\sigma_2(\ell m)(h) + \sigma_1(\ell m)(h)\sigma_2(\ell m)(m) \\
&\quad + \sigma_1(\ell m)(h)\sigma_2(\ell m)(\ell) + \sigma_1(\ell m)(m)\sigma_2(\ell m)(h) + \sigma_1(\ell m)(\ell)\sigma_2(\ell m)(h)] \\
M_{19} &= \sigma_0(\ell\ell)(h)[\sigma_1(\ell\ell)(h)\sigma_2(\ell\ell)(h) + \sigma_1(\ell\ell)(h)\sigma_2(\ell\ell)(m) \\
&\quad + \sigma_1(\ell\ell)(h)\sigma_2(\ell\ell)(\ell) + \sigma_1(\ell\ell)(m)\sigma_2(\ell\ell)(h) + \sigma_1(\ell\ell)(\ell)\sigma_2(\ell\ell)(h)] \\
M_{21} &= \sigma_0(hh)(m)[\sigma_1(hh)(h)\sigma_2(hh)(h) + \sigma_1(hh)(h)\sigma_2(hh)(m) \\
&\quad + \sigma_1(hh)(h)\sigma_2(hh)(\ell) + \sigma_1(hh)(m)\sigma_2(hh)(h) + \sigma_1(hh)(\ell)\sigma_2(hh)(h)] \\
M_{22} &= \sigma_0(hm)(m)[\sigma_1(hm)(h)\sigma_2(hm)(h) + \sigma_1(hm)(h)\sigma_2(hm)(m) \\
&\quad + \sigma_1(hm)(h)\sigma_2(hm)(\ell) + \sigma_1(hm)(m)\sigma_2(hm)(h) + \sigma_1(hm)(\ell)\sigma_2(hm)(h)]
\end{aligned}$$

$$\begin{aligned}
M_{23} &= \sigma_0(h\ell)(m)[\sigma_1(h\ell)(h)\sigma_2(h\ell)(h) + \sigma_1(h\ell)(h)\sigma_2(h\ell)(m) \\
&\quad + \sigma_1(h\ell)(h)\sigma_2(h\ell)(\ell) + \sigma_1(h\ell)(m)\sigma_2(h\ell)(h) + \sigma_1(h\ell)(\ell)\sigma_2(h\ell)(h)] \\
M_{24} &= \sigma_0(mh)(m)[\sigma_1(mh)(h)\sigma_2(mh)(h) + \sigma_1(mh)(h)\sigma_2(mh)(m) \\
&\quad + \sigma_1(mh)(h)\sigma_2(mh)(\ell) + \sigma_1(mh)(m)\sigma_2(mh)(h) + \sigma_1(mh)(\ell)\sigma_2(mh)(h)] \\
M_{25} &= \sigma_0(mm)(m)[\sigma_1(mm)(h)\sigma_2(mm)(h) + \sigma_1(mm)(h)\sigma_2(mm)(m) \\
&\quad + \sigma_1(mm)(h)\sigma_2(mm)(\ell) + \sigma_1(mm)(m)\sigma_2(mm)(h) + \sigma_1(mm)(\ell)\sigma_2(mm)(h)] \\
M_{26} &= \sigma_0(m\ell)(m)[\sigma_1(m\ell)(h)\sigma_2(m\ell)(h) + \sigma_1(m\ell)(h)\sigma_2(m\ell)(m) \\
&\quad + \sigma_1(m\ell)(h)\sigma_2(m\ell)(\ell) + \sigma_1(m\ell)(m)\sigma_2(m\ell)(h) + \sigma_1(m\ell)(\ell)\sigma_2(m\ell)(h)] \\
M_{27} &= \sigma_0(\ell h)(m)[\sigma_1(\ell h)(h)\sigma_2(\ell h)(h) + \sigma_1(\ell h)(h)\sigma_2(\ell h)(m) \\
&\quad + \sigma_1(\ell h)(h)\sigma_2(\ell h)(\ell) + \sigma_1(\ell h)(m)\sigma_2(\ell h)(h) + \sigma_1(\ell h)(\ell)\sigma_2(\ell h)(h)] \\
M_{28} &= \sigma_0(\ell m)(m)[\sigma_1(\ell m)(h)\sigma_2(\ell m)(h) + \sigma_1(\ell m)(h)\sigma_2(\ell m)(m) \\
&\quad + \sigma_1(\ell m)(h)\sigma_2(\ell m)(\ell) + \sigma_1(\ell m)(m)\sigma_2(\ell m)(h) + \sigma_1(\ell m)(\ell)\sigma_2(\ell m)(h)] \\
M_{29} &= \sigma_0(\ell\ell)(m)[\sigma_1(\ell\ell)(h)\sigma_2(\ell\ell)(h) + \sigma_1(\ell\ell)(h)\sigma_2(\ell\ell)(m) \\
&\quad + \sigma_1(\ell\ell)(h)\sigma_2(\ell\ell)(\ell) + \sigma_1(\ell\ell)(m)\sigma_2(\ell\ell)(h) + \sigma_1(\ell\ell)(\ell)\sigma_2(\ell\ell)(h)] \\
M_{31} &= \sigma_0(hh)(\ell)[\sigma_1(hh)(h)\sigma_2(hh)(h) + \sigma_1(hh)(h)\sigma_2(hh)(m) \\
&\quad + \sigma_1(hh)(h)\sigma_2(hh)(\ell) + \sigma_1(hh)(m)\sigma_2(hh)(h) + \sigma_1(hh)(\ell)\sigma_2(hh)(h)] \\
M_{32} &= \sigma_0(hm)(\ell)[\sigma_1(hm)(h)\sigma_2(hm)(h) + \sigma_1(hm)(h)\sigma_2(hm)(m) \\
&\quad + \sigma_1(hm)(h)\sigma_2(hm)(\ell) + \sigma_1(hm)(m)\sigma_2(hm)(h) + \sigma_1(hm)(\ell)\sigma_2(hm)(h)] \\
M_{33} &= \sigma_0(h\ell)(\ell)[\sigma_1(h\ell)(h)\sigma_2(h\ell)(h) + \sigma_1(h\ell)(h)\sigma_2(h\ell)(m) \\
&\quad + \sigma_1(h\ell)(h)\sigma_2(h\ell)(\ell) + \sigma_1(h\ell)(m)\sigma_2(h\ell)(h) + \sigma_1(h\ell)(\ell)\sigma_2(h\ell)(h)] \\
M_{34} &= \sigma_0(mh)(\ell)[\sigma_1(mh)(h)\sigma_2(mh)(h) + \sigma_1(mh)(h)\sigma_2(mh)(m) \\
&\quad + \sigma_1(mh)(h)\sigma_2(mh)(\ell) + \sigma_1(mh)(m)\sigma_2(mh)(h) + \sigma_1(mh)(\ell)\sigma_2(mh)(h)] \\
M_{35} &= \sigma_0(mm)(\ell)[\sigma_1(mm)(h)\sigma_2(mm)(h) + \sigma_1(mm)(h)\sigma_2(mm)(m) \\
&\quad + \sigma_1(mm)(h)\sigma_2(mm)(\ell) + \sigma_1(mm)(m)\sigma_2(mm)(h) + \sigma_1(mm)(\ell)\sigma_2(mm)(h)] \\
M_{36} &= \sigma_0(m\ell)(\ell)[\sigma_1(m\ell)(h)\sigma_2(m\ell)(h) + \sigma_1(m\ell)(h)\sigma_2(m\ell)(m) \\
&\quad + \sigma_1(m\ell)(h)\sigma_2(m\ell)(\ell) + \sigma_1(m\ell)(m)\sigma_2(m\ell)(h) + \sigma_1(m\ell)(\ell)\sigma_2(m\ell)(h)] \\
M_{37} &= \sigma_0(\ell h)(\ell)[\sigma_1(\ell h)(h)\sigma_2(\ell h)(h) + \sigma_1(\ell h)(h)\sigma_2(\ell h)(m) \\
&\quad + \sigma_1(\ell h)(h)\sigma_2(\ell h)(\ell) + \sigma_1(\ell h)(m)\sigma_2(\ell h)(h) + \sigma_1(\ell h)(\ell)\sigma_2(\ell h)(h)]
\end{aligned}$$

$$\begin{aligned}
M_{38} &= \sigma_0(\ell m)(\ell) [\sigma_1(\ell m)(h)\sigma_2(\ell m)(h) + \sigma_1(\ell m)(h)\sigma_2(\ell m)(m) \\
&\quad + \sigma_1(\ell m)(h)\sigma_2(\ell m)(\ell) + \sigma_1(\ell m)(m)\sigma_2(\ell m)(h) + \sigma_1(\ell m)(\ell)\sigma_2(\ell m)(h)] \\
M_{39} &= \sigma_0(\ell \ell)(\ell) [\sigma_1(\ell \ell)(h)\sigma_2(\ell \ell)(h) + \sigma_1(\ell \ell)(h)\sigma_2(\ell \ell)(m) \\
&\quad + \sigma_1(\ell \ell)(h)\sigma_2(\ell \ell)(\ell) + \sigma_1(\ell \ell)(m)\sigma_2(\ell \ell)(h) + \sigma_1(\ell \ell)(\ell)\sigma_2(\ell \ell)(h)] \\
M_{41} &= \sigma_0(hh)(h) [\sigma_1(hh)(m)\sigma_2(hh)(m) + \sigma_1(hh)(m)\sigma_2(hh)(\ell) \\
&\quad + \sigma_1(hh)(\ell)\sigma_2(hh)(m)] \\
M_{42} &= \sigma_0(hm)(h) [\sigma_1(hm)(m)\sigma_2(hm)(m) + \sigma_1(hm)(m)\sigma_2(hm)(\ell) \\
&\quad + \sigma_1(hm)(\ell)\sigma_2(hm)(m)] \\
M_{43} &= \sigma_0(h\ell)(h) [\sigma_1(h\ell)(m)\sigma_2(h\ell)(m) + \sigma_1(h\ell)(m)\sigma_2(h\ell)(\ell) \\
&\quad + \sigma_1(h\ell)(\ell)\sigma_2(h\ell)(m)] \\
M_{44} &= \sigma_0(mh)(h) [\sigma_1(mh)(m)\sigma_2(mh)(m) + \sigma_1(mh)(m)\sigma_2(mh)(\ell) \\
&\quad + \sigma_1(mh)(\ell)\sigma_2(mh)(m)] \\
M_{45} &= \sigma_0(mm)(h) [\sigma_1(mm)(m)\sigma_2(mm)(m) + \sigma_1(mm)(m)\sigma_2(mm)(\ell) \\
&\quad + \sigma_1(mm)(\ell)\sigma_2(mm)(m)] \\
M_{46} &= \sigma_0(m\ell)(h) [\sigma_1(m\ell)(m)\sigma_2(m\ell)(m) + \sigma_1(m\ell)(m)\sigma_2(m\ell)(\ell) \\
&\quad + \sigma_1(m\ell)(\ell)\sigma_2(m\ell)(m)] \\
M_{47} &= \sigma_0(\ell h)(h) [\sigma_1(\ell h)(m)\sigma_2(\ell h)(m) + \sigma_1(\ell h)(m)\sigma_2(\ell h)(\ell) \\
&\quad + \sigma_1(\ell h)(\ell)\sigma_2(\ell h)(m)] \\
M_{48} &= \sigma_0(\ell m)(h) [\sigma_1(\ell m)(m)\sigma_2(\ell m)(m) + \sigma_1(\ell m)(m)\sigma_2(\ell m)(\ell) \\
&\quad + \sigma_1(\ell m)(\ell)\sigma_2(\ell m)(m)] \\
M_{49} &= \sigma_0(\ell \ell)(h) [\sigma_1(\ell \ell)(m)\sigma_2(\ell \ell)(m) + \sigma_1(\ell \ell)(m)\sigma_2(\ell \ell)(\ell) \\
&\quad + \sigma_1(\ell \ell)(\ell)\sigma_2(\ell \ell)(m)] \\
M_{51} &= \sigma_0(hh)(m) [\sigma_1(hh)(m)\sigma_2(hh)(m) + \sigma_1(hh)(m)\sigma_2(hh)(\ell) \\
&\quad + \sigma_1(hh)(\ell)\sigma_2(hh)(m)] \\
M_{52} &= \sigma_0(hm)(m) [\sigma_1(hm)(m)\sigma_2(hm)(m) + \sigma_1(hm)(m)\sigma_2(hm)(\ell) \\
&\quad + \sigma_1(hm)(\ell)\sigma_2(hm)(m)] \\
M_{53} &= \sigma_0(h\ell)(m) [\sigma_1(h\ell)(m)\sigma_2(h\ell)(m) + \sigma_1(h\ell)(m)\sigma_2(h\ell)(\ell) \\
&\quad + \sigma_1(h\ell)(\ell)\sigma_2(h\ell)(m)]
\end{aligned}$$

$$\begin{aligned}
M_{54} &= \sigma_0(mh)(m)[\sigma_1(mh)(m)\sigma_2(mh)(m) + \sigma_1(mh)(m)\sigma_2(mh)(\ell) \\
&\quad + \sigma_1(mh)(\ell)\sigma_2(mh)(m)] \\
M_{55} &= \sigma_0(mm)(m)[\sigma_1(mm)(m)\sigma_2(mm)(m) + \sigma_1(mm)(m)\sigma_2(mm)(\ell) \\
&\quad + \sigma_1(mm)(\ell)\sigma_2(mm)(m)] \\
M_{56} &= \sigma_0(m\ell)(m)[\sigma_1(m\ell)(m)\sigma_2(m\ell)(m) + \sigma_1(m\ell)(m)\sigma_2(m\ell)(\ell) \\
&\quad + \sigma_1(m\ell)(\ell)\sigma_2(m\ell)(m)] \\
M_{57} &= \sigma_0(\ell h)(m)[\sigma_1(\ell h)(m)\sigma_2(\ell h)(m) + \sigma_1(\ell h)(m)\sigma_2(\ell h)(\ell) \\
&\quad + \sigma_1(\ell h)(\ell)\sigma_2(\ell h)(m)] \\
M_{58} &= \sigma_0(\ell m)(m)[\sigma_1(\ell m)(m)\sigma_2(\ell m)(m) + \sigma_1(\ell m)(m)\sigma_2(\ell m)(\ell) \\
&\quad + \sigma_1(\ell m)(\ell)\sigma_2(\ell m)(m)] \\
M_{59} &= \sigma_0(\ell\ell)(m)[\sigma_1(\ell\ell)(m)\sigma_2(\ell\ell)(m) + \sigma_1(\ell\ell)(m)\sigma_2(\ell\ell)(\ell) \\
&\quad + \sigma_1(\ell\ell)(\ell)\sigma_2(\ell\ell)(m)] \\
M_{61} &= \sigma_0(hh)(\ell)[\sigma_1(hh)(m)\sigma_2(hh)(m) + \sigma_1(hh)(m)\sigma_2(hh)(\ell) \\
&\quad + \sigma_1(hh)(\ell)\sigma_2(hh)(m)] \\
M_{62} &= \sigma_0(hm)(\ell)[\sigma_1(hm)(m)\sigma_2(hm)(m) + \sigma_1(hm)(m)\sigma_2(hm)(\ell) \\
&\quad + \sigma_1(hm)(\ell)\sigma_2(hm)(m)] \\
M_{63} &= \sigma_0(h\ell)(\ell)[\sigma_1(h\ell)(m)\sigma_2(h\ell)(m) + \sigma_1(h\ell)(m)\sigma_2(h\ell)(\ell) \\
&\quad + \sigma_1(h\ell)(\ell)\sigma_2(h\ell)(m)] \\
M_{64} &= \sigma_0(mh)(\ell)[\sigma_1(mh)(m)\sigma_2(mh)(m) + \sigma_1(mh)(m)\sigma_2(mh)(\ell) \\
&\quad + \sigma_1(mh)(\ell)\sigma_2(mh)(m)] \\
M_{65} &= \sigma_0(mm)(\ell)[\sigma_1(mm)(m)\sigma_2(mm)(m) + \sigma_1(mm)(m)\sigma_2(mm)(\ell) \\
&\quad + \sigma_1(mm)(\ell)\sigma_2(mm)(m)] \\
M_{66} &= \sigma_0(m\ell)(\ell)[\sigma_1(m\ell)(m)\sigma_2(m\ell)(m) + \sigma_1(m\ell)(m)\sigma_2(m\ell)(\ell) \\
&\quad + \sigma_1(m\ell)(\ell)\sigma_2(m\ell)(m)] \\
M_{67} &= \sigma_0(\ell h)(\ell)[\sigma_1(\ell h)(m)\sigma_2(\ell h)(m) + \sigma_1(\ell h)(m)\sigma_2(\ell h)(\ell) \\
&\quad + \sigma_1(\ell h)(\ell)\sigma_2(\ell h)(m)] \\
M_{68} &= \sigma_0(\ell m)(\ell)[\sigma_1(\ell m)(m)\sigma_2(\ell m)(m) + \sigma_1(\ell m)(m)\sigma_2(\ell m)(\ell) \\
&\quad + \sigma_1(\ell m)(\ell)\sigma_2(\ell m)(m)]
\end{aligned}$$

$$\begin{aligned}
M_{69} &= \sigma_0(\ell\ell)(\ell)[\sigma_1(\ell\ell)(m)\sigma_2(\ell\ell)(m) + \sigma_1(\ell\ell)(m)\sigma_2(\ell\ell)(\ell) + \sigma_1(\ell\ell)(\ell)\sigma_2(\ell\ell)(m)] \\
M_{71} &= \sigma_0(hh)(h)[\sigma_1(hh)(\ell)\sigma_2(hh)(\ell)] \\
M_{72} &= \sigma_0(hm)(h)[\sigma_1(hm)(\ell)\sigma_2(hm)(\ell)] \\
M_{73} &= \sigma_0(h\ell)(h)[\sigma_1(h\ell)(\ell)\sigma_2(h\ell)(\ell)] \\
M_{74} &= \sigma_0(mh)(h)[\sigma_1(mh)(\ell)\sigma_2(mh)(\ell)] \\
M_{75} &= \sigma_0(mm)(h)[\sigma_1(mm)(\ell)\sigma_2(mm)(\ell)] \\
M_{76} &= \sigma_0(m\ell)(h)[\sigma_1(m\ell)(\ell)\sigma_2(m\ell)(\ell)] \\
M_{77} &= \sigma_0(\ell h)(h)[\sigma_1(\ell h)(\ell)\sigma_2(\ell h)(\ell)] \\
M_{78} &= \sigma_0(\ell m)(h)[\sigma_1(\ell m)(\ell)\sigma_2(\ell m)(\ell)] \\
M_{79} &= \sigma_0(\ell\ell)(h)[\sigma_1(\ell\ell)(\ell)\sigma_2(\ell\ell)(\ell)] \\
M_{81} &= \sigma_0(hh)(m)[\sigma_1(hh)(\ell)\sigma_2(hh)(\ell)] \\
M_{82} &= \sigma_0(hm)(m)[\sigma_1(hm)(\ell)\sigma_2(hm)(\ell)] \\
M_{83} &= \sigma_0(h\ell)(m)[\sigma_1(h\ell)(\ell)\sigma_2(h\ell)(\ell)] \\
M_{84} &= \sigma_0(mh)(m)[\sigma_1(mh)(\ell)\sigma_2(mh)(\ell)] \\
M_{85} &= \sigma_0(mm)(m)[\sigma_1(mm)(\ell)\sigma_2(mm)(\ell)] \\
M_{86} &= \sigma_0(m\ell)(m)[\sigma_1(m\ell)(\ell)\sigma_2(m\ell)(\ell)] \\
M_{87} &= \sigma_0(\ell h)(m)[\sigma_1(\ell h)(\ell)\sigma_2(\ell h)(\ell)] \\
M_{88} &= \sigma_0(\ell m)(m)[\sigma_1(\ell m)(\ell)\sigma_2(\ell m)(\ell)] \\
M_{89} &= \sigma_0(\ell\ell)(m)[\sigma_1(\ell\ell)(\ell)\sigma_2(\ell\ell)(\ell)] \\
M_{91} &= \sigma_0(hh)(\ell)[\sigma_1(hh)(\ell)\sigma_2(hh)(\ell)] \\
M_{92} &= \sigma_0(hm)(\ell)[\sigma_1(hm)(\ell)\sigma_2(hm)(\ell)] \\
M_{93} &= \sigma_0(h\ell)(\ell)[\sigma_1(h\ell)(\ell)\sigma_2(h\ell)(\ell)] \\
M_{94} &= \sigma_0(mh)(\ell)[\sigma_1(mh)(\ell)\sigma_2(mh)(\ell)] \\
M_{95} &= \sigma_0(mm)(\ell)[\sigma_1(mm)(\ell)\sigma_2(mm)(\ell)] \\
M_{96} &= \sigma_0(m\ell)(\ell)[\sigma_1(m\ell)(\ell)\sigma_2(m\ell)(\ell)] \\
M_{97} &= \sigma_0(\ell h)(\ell)[\sigma_1(\ell h)(\ell)\sigma_2(\ell h)(\ell)] \\
M_{98} &= \sigma_0(\ell m)(\ell)[\sigma_1(\ell m)(\ell)\sigma_2(\ell m)(\ell)] \\
M_{99} &= \sigma_0(\ell\ell)(\ell)[\sigma_1(\ell\ell)(\ell)\sigma_2(\ell\ell)(\ell)]
\end{aligned}$$