mathematics

MDPI

Article

Characterizing Manipulation via Machiavellianism

Jacqueline Sanchez-Rabaza, Jose Maria Rocha-Martinez * and Julio B. Clempner

Escuela Superior de Física y Matemáticas, Instituto Politécnico Nacional, School of Physics and Mathematics, National Polytechnic Institute, Edificio 9 U.P. Adolfo Lopez Mateos, Col. San Pedro Zacatenco, Mexico City 07730, Mexico; jsanchezr1217@alumno.ipn.mx (J.S.-R.); jclempnerk@ipn.mx (J.B.C.)
* Correspondence: jrocham@ipn.mx

Abstract: Machiavellianism refers to the propensity of taking advantage of people within a society. Machiavellians have reputations for being cunning and competitive. They are also skilled long-term strategists and planners. Other than their "victories," there are no other successful conclusions for them. The belief component of Machiavellianism includes cynical views of human nature (e.g., manipulated and manipulating individuals), interpersonal exploitation as a technique (e.g., strategic thinking), and a lack of traditional morality that would forbid their behaviors (e.g., immoral behaviors). This paper focuses on a game that involves manipulation. The game was conceptualized using the best and worst Nash equilibrium points as part of our contribution. We constrained the problem to homogeneous, finite, ergodic, and controllable Bayesian–Markov games. Machiavellian players pretended to be in one state when they were actually in another. Moreover, they pretended to perform one action while actually playing another. All Machiavellian individuals engaged in some form of interpersonal manipulation. Manipulating players exhibited a higher preference compared to manipulated participants. The Pareto frontier is defined as the line where manipulating players play the best Nash equilibrium and manipulated players play the worst Nash equilibrium. It is also considered a sequential Bayesian–Markov manipulation game involving multiple manipulating players and manipulated players. Finally, a tractable characterization of the manipulation equilibrium results is provided. To guarantee that the game's solution converged into a singular solution, we used Tikhonov's penalty regularization method. A numerical example describes the results of our model.

Keywords: Machiavellianism; manipulation; Markov chain; game theory

MSC: 91A10; 91A40; 91A80; 91E30; 62C10

Citation: Sanchez-Rabaza, J.; Rocha-Martinez, J.M.; Clempner, J.B. Characterizing Manipulation via Machiavellianism. *Mathematics* **2023**, *11*, 4143. https://doi.org/10.3390/math11194143

Academic Editors: Mahendra Piraveenan, Samit Bhattacharrya and Chunqiao Tan

Received: 25 April 2023
Revised: 4 June 2023
Accepted: 13 June 2023
Published: 30 September 2023

1. Introduction

1.1. Brief Review

The advantages and disadvantages of adding more believable behavioral assumptions to a traditional economic model have been extensively discussed in the literature. The focus on manipulation is one of the most obvious issues. Game-theoretic studies have discovered considerable individual variances in situations where the game allows for unstable behavior, such as manipulation. In this paper, we concentrate on the literature on decision-making and personality linked to Machiavellianism, a personality trait connected to manipulativeness, callousness, and indifference to morality in the study of personality psychology [1].

Christie and Geis [2] employed modified and abridged remarks derived from Niccolo Machiavelli's works to explore variances in human behavior. Niccolo Machiavelli, in his conception of the world, classifies individuals as manipulated and manipulating. Machiavellian individuals often exploit others, hold pessimistic views on human nature, and lack the moral conscience that would make them accountable for their deeds. Machiavellianism captures the lack of coordination in games where individuals are selfish and may have conflicted interests.

The Mach IV exam that Christie and Geis created is now the accepted self-evaluation instrument and scale used for the Machiavellianism concept [3,4]. High-Mach individuals are more likely to exhibit dishonest behavior and possess a cynical, unfeeling disposition compared to low-Mach individuals. Christie established that a Machiavellian individual would have the following qualities: (a) Their *worldview* is determined by manipulators and manipulated individuals (i.e., views component), where manipulators do not feel sympathy for their prey, and manipulated individuals also manipulate to some degree. (b) Machiavellians have a lack of regard for *conventional morality* (e.g., morality component) and do not care about actions, such as lying and cheating. (c) Machiavellians frequently place more emphasis on pragmatic problem-solving than on ideological commitments and they are more likely to employ *power strategies* (e.g., tactical component) to further their personal goals rather than ideological ones.

The *Nash equilibrium* is the most classical method used to characterize the solution of a non-cooperative game involving two or more players in game theory [5]. Each player in a Nash equilibrium is aware of the equilibrium tactics of the other players, and changing one's own strategy would not benefit anyone. The concept of the Nash equilibrium dates back to Cournot, who used it to explain how rival enterprises determined their output levels in 1838 [6]. The current strategy set represents an action plan based on the events that have transpired thus far in the game, and no player can improve their personal expected payoff by altering their strategy while the other players keep theirs unchanged. In optimization problems (with one player), the scenario is the same: if the cost function is strongly convex and we are at the minimal point, then any movement away from this point will result in the greatest number of payout values. From this viewpoint, the Nash equilibrium generalizes the conventional optimal approach for situations involving a single participant. We roughly conceptualize Machiavellianism as a welfare system of the worst-case Nash Equilibrium (manipulated individuals) and the non-cooperative optimal welfare system (manipulating individuals).

1.2. Related Work

Allen and Gorton [7] addressed the idea of a manipulation equilibrium while considering stock price trends and welfare concerns. Stock price manipulation decreases the pre-bid stock price, as demonstrated by Bagnoli and Lipman [8]. They showed that manipulation drives takeover offers and blocks certain effective takeovers when there is limited takeover activity. Clempner [9] provided a game theory method for simulating manipulation that makes use of a reinforcement learning method to incorporate the idea of immorality. Cumming et al. [10] found evidence of the harmful consequences of market manipulation on innovation by using a sample of suspected stock price manipulation incidents based on intra-day data for equities from nine nations over eight years. They demonstrated that these detrimental consequences are more detrimental to innovation in economies with weak intellectual property protections and strong shareholder protections. Clempner and Trejo [11] presented a method based on Nash's bargaining model for modeling manipulation games. Clempner [12] applied the previous approach for repeated Stackelberg security games. Clempner [13] suggested a manipulation game based on a class of ergodic Bayesian–Markov models.

One sender and one receiver are used to illustrate the traditional Bayesian persuasion framework. The sender, who has access to certain private information, creates a signaling strategy in an attempt to influence the receiver, to make a good decision. For models with many senders, see the works by Milgrom [14] and Krishna [15]. Private signal-based persuasion has been examined in a variety of contexts, including two-agent, two-action games [16], and unanimity elections [17]. For a single sender and receiver model, Kamenica and Gentzkow [18] introduced a Bayesian persuasion framework. They demonstrated the necessary and sufficient criteria for the existence of a signal that helps the sender and described the optimum signals. According to Bergemann and Morris [19], the designer of a fixed mechanism in Bayesian persuasion chooses an information structure for the participants in

order to advance their objectives. Gentzkow and Kamenica [20] extended their earlier work [18] by incorporating multiple senders in scenarios where the set of potential signals was extensive, taking into account a lattice structure that enabled intuitive signal comparison and combining. In order to influence a decision-maker's choice, Brocas et al. [21] proposed a framework where two adversaries invest resources in gathering information from the general public regarding an unknown condition of the world. They described the opponents' sampling tactics in the game's equilibrium and demonstrated that they change when one adversary's cost of information acquisition rises. Gul and Pesendorfer [22] presented a model that describes political campaigns involving asymmetric information and signaling about a binary state of the world. The model includes two senders with conflicting objectives. Several authors, such as [23–26], cited frameworks in which the information revealed by the sender was only viewed by the recipients.

1.3. Main Contribution

This paper describes a manipulation game. Our contribution consists of conceptualizing the game by employing the best and the worst Nash equilibrium points. We constrained the problem to an assortment of homogeneous, finite, ergodic, and controllable Bayesian–Markov games, where:

i. Machiavellian players pretended to be in one state when they were actually in another.
ii. Machiavellian players pretended to perform one action while actually playing another.
iii. All Machiavellian individuals engaged in some form of interpersonal manipulation.
iv. Manipulating players exhibited a higher preference compared to manipulated participants.
v. The Pareto frontier is characterized as the boundary where the manipulating players play the best Nash equilibrium and the manipulated players play the worst Nash equilibrium.
vi. It is considered a sequential Bayesian–Markov manipulation game involving multiple manipulating players and manipulated players.
vii. A tractable characterization of the manipulation equilibrium results is provided.
viii. To guarantee that the game's solution converged into a singular solution, we used Tikhonov's penalty regularization method.

1.4. Organization of the Paper

The structure of the paper is as follows. Section 2 describes the manipulation game while Section 3 considers the manipulation equilibrium. In Section 4, considering the Tanaka function, the manipulation equilibrium is computed. Section 5 considers a Bayesian–Markov approach of the model, where the moral hazard is developed. A numerical example is presented in Section 6. Section 7 concludes with some remarks.

2. Manipulation Game

Let $\mathcal{G}$ be a *non-cooperative game* denoted as $\mathcal{G} = (\mathcal{I}, \mathcal{S}, u^l, u^f)$. $\mathcal{I} = \{1, \ldots, n+m\}$ is the set of total Machiavellian players, where $L = \{1, \ldots, n\}$ is the set of *manipulating players* indexed by $l = \overline{1,n}$ and $F = \{1, \ldots, m\}$ is the set of *manipulated players* indexed by $f = \overline{1,m}$. $\mathcal{S} = \mathcal{S}^l \times \mathcal{S}^f$ is a *strategy set*, such that $\mathcal{S}^l$ is the strategy set for the manipulating players ($l \in L$) and $\mathcal{S}^f$ is a strategy set for the manipulated players ($f \in F$). Element $s^l \in \mathcal{S}^l$ is a strategy for player $l \in L$, and element $s^f \in \mathcal{S}^f$ is a strategy for player $f \in F$. Let us consider that, in general, the index set is $\mathcal{I}$, and $s^{-i} = (s^j)_{j \in \mathcal{I}/i}$ denotes the *joint strategy profile* of all players, except player i , i.e., $(s^1, \ldots, s^{i-1}, s^{i+1}, \ldots, s^{n+m})$.

A *strategy profile* $s = (s^1, \ldots, s^n, s^{n+1}, \ldots, s^{n+m})$ is a vector of strategies, one for each player, and the set of all possible strategy profiles is denoted by $\mathcal{S} = \mathcal{S}^1 \times \cdots \times \mathcal{S}^n \times \ldots \mathcal{S}^m$, i.e., $\mathcal{S} = \times_{k=1}^{n+m} \mathcal{S}^k$. For any *admissible strategy set* $s^l \in \mathcal{S}^l_{adm}$ for the manipulating players:

$$\mathcal{S}^l_{adm} := \left\{ s^l : \sum_{l \in L} g^l(s^l) = 0, \sum_{l \in L} h^l(s^l) \leq 0 \right\},$$

and any *admissible strategy space* $s^f \in \mathcal{S}^f_{adm}$ for the manipulating players:

$$\mathcal{S}^f_{adm} := \left\{ s^f : \sum_{f \in F} g^f(s^f) = 0, \sum_{f \in F} h^f(s^f) \leq 0 \right\},$$

such that $\mathcal{S}_{adm} = \mathcal{S}^l_{adm} \times \mathcal{S}^f_{adm}$. The *reward* of a manipulating player $l \in L$ for a strategy $s = (s^1, \ldots, s^n, s^{n+1}, \ldots, s^{n+m})$ is determined by a utility function $u^l : \mathcal{S}_{adm} \to \mathbb{R}$, and the *reward* of a manipulated player $f \in F$ for a strategy $s = (s^1, \ldots, s^n, s^{n+1}, \ldots, s^{n+m})$ is determined by a utility function $u^f : \mathcal{S}_{adm} \to \mathbb{R}$, with the admissible strategy space of the manipulating players being $\mathcal{S}^l_{adm}$ and the admissible strategy space of the manipulating players being $\mathcal{S}^f_{adm}$.

The designations of the manipulating players and the manipulated players indicate the sequential order of play between two types of Machiavellian individuals, meaning that the manipulating players play first and the manipulated players play next. The manipulating and manipulated players are involved in a non-cooperative game. which we will refer to as a *manipulation or Machiavellian game*.

3. Manipulation Equilibrium

Each Machiavellian player must solve an optimization problem, where the admissible set is constrained by convex constraints based on the variables of the other Machiavellian players, as well as integrity constraints with the objective function. The objective function of the optimization problem depends on the variables of the other players.

We assume that the manipulating players obtain a reward u^l, and the manipulated players earn a reward u^f; moreover, we assume that these rewards are continuous and convex in all their arguments. The manipulating players attempt to find a solution to the optimization problem

$$\max_{s^l \in S^l} \left\{ u^l(s^l, s^f) | s^f \in \arg\max_{s'^f \in S^f} u^f(s'^f, s^l) \right\},$$

where the manipulated players attempt to find a solution to the optimization problem

$$\max_{s^f \in S^f} u^f(s^f, s^l).$$

The Machiavellian players obey myopic strategies that move in the direction of improvement with regard to the two optimization problems mentioned above: the former for the manipulating players and the latter for the manipulated players.

Let us first describe the equilibrium concept investigated for simultaneous play games and contrast it with that investigated in the sequential counterpart. A strategy for the Machiavellian players $s^*(\lambda) \in \mathcal{S}_{adm}$, satisfying the following inequality, is known as the *Nash equilibrium* for the Machiavellian players

$$\left.\begin{aligned} F(s, \lambda) := \sum_{i \in I} \lambda^i \left[u^i(s^i, s^{-i}) - u^i(s^*(\lambda)) \right] \leq 0, \\ \lambda \in \Lambda = \left\{ \lambda \in \mathbb{R}^n : \lambda^i \geq 0, \sum_{i \in \mathcal{I}} \lambda^i = 1 \right\}. \end{aligned}\right\} \quad (1)$$

If each Machiavellian player has chosen a strategy s, and no player can benefit from changing strategies while others keep their strategies the same, the present set of strategical choices and payoffs is known as the Nash equilibrium. In other words, for a joint strategy, $s^* \in \mathcal{S}_{adm}$, satisfying for any admissible $s^i \in \mathcal{S}^i_{adm}$, we have $u^i(s^i, s^{-i*}) \leq u^i(s^{i*}, s^{-i*})$ as a Nash equilibrium point. It is important to note that considering the uniform distribution $\lambda^i = n^{-1}$, we obtain the original definition of the Nash equilibrium [5]. We assume that the functions $u^i(s)$ $(i \in \mathcal{I})$ are meant to be multilinear in all of their arguments [27,28].

The condition in Equation (1) implies the *Nash property*:

$$\lambda^i\left[u^i(s^i, s^{-i}) - u^i(s^*(\lambda))\right] \leq 0$$

$$\text{for any } s \in \mathcal{S}_{adm} \text{ and } i \in I.$$

Any equilibrium point $s^*(\lambda) \in \mathcal{S}_{adm}$ belongs to the Pareto set [29,30]

$$\mathcal{P} := \left\{ u^i(s^*(\lambda)) : s^*(\lambda) \in \underset{s\in\mathcal{S}_{adm}}{Arg\max} \sum_{i\in\mathcal{I}} \lambda^i u^i(s) \right\}. \tag{2}$$

Following Tanaka [31,32], a Nash equilibrium $s^*(\lambda)$ that satisfies the conditions in Equation (1) is

$$s^*(\lambda) \in \underset{s\in\mathcal{S}_{adm}}{Arg\,min} F(s,\lambda) = \underset{s\in\mathcal{S}_{adm}}{Arg\,min} \sum_{i\in I} \lambda^i\left[u^i(s^*(\lambda)) - u^i(s)\right] =$$

$$\underset{s\in\mathcal{S}_{adm}}{Arg\,min} \sum_{i\in I} \lambda^i\left[-u^i(s)\right] = \underset{s\in\mathcal{S}_{adm}}{Arg\,max} \sum_{i\in I} \lambda^i u^i(s),$$

which determines the Pareto set given in Equation (2).

Remark 1. *It is important to note that by choosing different $\lambda^i \in \Lambda$, we obtain all feasible Nash equilibria in the proposed game.*

Let us introduce the individual reward for the manipulating players $\varphi^l(s^l, s^f, \lambda^l) = \hat{\lambda}^l u^l(\hat{\mathbf{s}}^l, \hat{\mathbf{s}}^{-l}, s^f)$ for any admissible $s^l = (\hat{\mathbf{s}}^l, \hat{\mathbf{s}}^{-l}) \in S^l_{adm}$, such that

$$\varphi(s^l, s^f, \lambda^l) = \sum_{l\in L} \varphi^l(s^l, s^f, \lambda^l) = \sum_{l\in L} \hat{\lambda}^l u^l(\hat{\mathbf{s}}^l, \hat{\mathbf{s}}^{-l}, s^f) \to \max_{\mathbf{s}^l\in\mathcal{S}^l_{adm}}$$

and the individual reward for the manipulated players $\psi^f(s^f, s^l, \lambda^f) = \hat{\lambda}^f u^f(\hat{\mathbf{s}}^f, \hat{\mathbf{s}}^{-f}, s^l)$ for any admissible $s^f = (\hat{\mathbf{s}}^f, \hat{\mathbf{s}}^{-f}) \in S^f_{adm}$ where

$$\psi(s^f, s^l, \lambda^f) = \sum_{f\in F} \psi^f(s^f, s^l, \lambda^f) = \sum_{f\in F} \hat{\lambda}^f u^f(\hat{\mathbf{s}}^f, \hat{\mathbf{s}}^{-f}, s^l) \to \max_{s^f\in\mathcal{S}^f_{adm}}.$$

The *welfare* of a strategy profile s for the manipulated players is defined as the minimum reward of a Machiavellian player $\min_{\lambda^f\in\Lambda} \psi(s^f, s^l, \lambda^f)$, or the worst-case among all the Nash equilibria. On the other hand, the optimal welfare for the manipulating players is the maximum reward of a Machiavellian player $\max_{\lambda^l\in\Lambda} \varphi(s^l, s^f, \lambda^l)$, or the best-case among all the Nash equilibria.

In a manipulation game, one strategy $s^{l*} \in \mathcal{S}^l_{adm}$ of the manipulating players is called a manipulation (Machiavellian) equilibrium, if

$$\max_{\lambda^l\in\Lambda} \sup_{s^f\in\mathcal{N}(s^l)} \varphi(s^l, s^f, \lambda^l) \leq \max_{\lambda^l\in\Lambda} \sup_{s^f\in\mathcal{N}(s^{l*})} \varphi(s^{l*}, s^f, \lambda^l),$$

where

$$\mathcal{N}(s^l) = \left\{ s^f \in \mathcal{S}^f_{adm} : \psi(s'^f, s^l, \lambda^f) \leq \psi(s^f, s^l, \lambda^f), s'^f \in \mathcal{S}^f_{adm} \right\},$$

represents the set of Nash equilibria $\mathcal{N}(s^l)$, given that the manipulating players are playing strategy s^l and the manipulated players' best reply is the set of Nash equilibria (with s^l fixed).

4. Computing the Manipulation Equilibrium

4.1. The Tanaka Function

We suggest the following representation for the manipulation game

$$\max_{\lambda^l \in \Lambda} \sup_{s^f \in \mathcal{N}(s^l)} \varphi(s^l, s^f, \lambda^l)$$
$$s.t.$$
$$g^l(s^l) = 0, h^l(s^l) \leq 0$$

$$\min_{\lambda^f \in \Lambda} \sup_{s^l \in \mathcal{N}(s^f)} \psi(s^f, s^l, \lambda^f)$$
$$s.t.$$
$$g^f(s^f) = 0, h^f(s^f) \leq 0$$

The complement variables are fixed, for instance, for computing $\max_{\lambda^l \in \Lambda} \sup_{s^f \in \mathcal{N}(s^l)} \varphi(s^l, s^f, \lambda^l)$, variable s^f is fixed.

Let us define the regularized welfare function as follows:

$$\varphi_\delta(s^l, s^f, \lambda^l) = \varphi(s^l, s^f, \lambda^l) - \frac{\delta}{2}\left\|s^l\right\|^2 + \frac{\delta}{2}\left\|\lambda^l\right\|^2, \ \delta > 0, \tag{3}$$

and

$$\psi_\delta(s^f, s^l, \lambda^l) = \psi(s^f, s^l, \lambda^l) - \frac{\delta}{2}\left\|s^f\right\|^2 - \frac{\delta}{2}\left\|\lambda^f\right\|^2, \ \delta > 0. \tag{4}$$

In order to prevent changing the form of the original function, the regularization method additionally focuses on determining the parameter δ for making the original objective functions, $\varphi(s^l, s^f, \lambda)$ and $\psi(s^f, s^l, \lambda^f)$, and the term, $\frac{\delta}{2}$, maximal, which has a unique solution, given that functions (3) and (4) are *strongly convex* if $\delta > 0$.

Applying the penalty approach [27,28] in Equations (3) and (4), we have

$$\Phi_{k,\delta}\left(s^l, s^f, \lambda^l\right) := \varphi_\delta(s^l, s^f, \lambda^l) + k\left[-\tfrac{1}{2}\left\|g^l(s^l)\right\|^2 - \tfrac{1}{2}\left\|h^l(s^l)\right\|^2 - \tfrac{\delta}{2}\left(\left\|s^l\right\|^2 - \left\|\lambda^l\right\|^2\right)\right]. \tag{5}$$

Let $\alpha = k^{-1}$, then we have

$$\Phi_{\alpha,\delta}\left(s^l, s^f, \lambda^l\right) := \alpha\varphi_\delta(s^l, s^f, \lambda^l) - \tfrac{1}{2}\left\|g^l(s^l)\right\|^2 - \tfrac{1}{2}\left\|h^l(s^l)\right\|^2 - \tfrac{\delta}{2}\left(\left\|s^l\right\|^2 - \left\|\lambda^l\right\|^2\right), \tag{6}$$

where

$$s^{l**}_{\Phi_{\alpha,\delta}} = \arg\max_{\lambda^l \in \Lambda} \max_{s^l \in \mathcal{S}^l_{adm}} \Phi_{\alpha,\delta}\left(s^l, s^f, \lambda^l\right),$$

and

$$\Psi_{k,\delta}\left(s^f, s^l, \lambda^f\right) := \psi_\delta(s^f, s^l, \lambda^f) + k\left[-\tfrac{1}{2}\left\|g^f(s^f)\right\|^2 - \tfrac{1}{2}\left\|h^f(s^f)\right\|^2 - \tfrac{\delta}{2}\left(\left\|s^f\right\|^2 + \left\|\lambda^f\right\|^2\right)\right]. \tag{7}$$

Let $\alpha = k^{-1}$, then we have

$$\Psi_{\alpha,\delta}\left(s^f, s^l, \lambda^f\right) := \alpha\psi_\delta(s^f, s^l, \lambda^f) - \tfrac{1}{2}\left\|g^f(s^f)\right\|^2 - \tfrac{1}{2}\left\|h^f(s^f)\right\|^2 - \tfrac{\delta}{2}\left(\left\|s^f\right\|^2 + \left\|\lambda^f\right\|^2\right), \tag{8}$$

where

$$s^{f**}_{\Psi_{\alpha,\delta}} = \arg\min_{\lambda^f \in \Lambda} \max_{s^f \in \mathcal{S}_{adm}} \Psi_{\alpha,\delta}\left(s^f, s^l, \lambda^f\right),$$

such that $s^{**} = (s^{l**}_{\Phi_{\alpha,\delta}}, s^{f**}_{\Psi_{\alpha,\delta}})$.

The players are attempting to reach one of the ε-Nash equilibria [31,32], which involves finding a joint strategy $s \in \mathcal{S}_{adm}$, satisfying for any admissible $s^l \in \mathcal{S}^l_{adm}$ and $s^f \in \mathcal{S}^f_{adm}$ the system of inequalities (*the ε-Nash equilibrium condition*)

$$\begin{gathered} G_{\alpha,\delta}\left(s^l, \tilde{s}^l, s^f, \lambda^l\right) \leq \varepsilon \text{ for any } s^l \in \mathcal{S}^l_{adm} \text{ and } l = \overline{1,n},\ \varepsilon \geq 0, \\ G_{\alpha,\delta}\left(s^l, \tilde{s}^l, s^f, \lambda^l\right) := \sum_{l \in L}\left[\Phi^l_{\alpha,\delta}\left(\hat{\mathbf{s}}^l, \hat{\mathbf{s}}^{-l}, s^f, \hat{\lambda}^l\right) - \Phi^l_{\alpha,\delta}\left(\mathring{\mathbf{s}}^l, \hat{\mathbf{s}}^{-l}, s^f, \hat{\lambda}^l\right)\right] \leq \varepsilon, \end{gathered} \tag{9}$$

where $\tilde{s}^l \in \mathcal{S}^l_{adm}$ is the complement of s^l and s^f is fixed and

$$\mathring{\mathbf{s}}^l := \arg\max_{\hat{s}^l \in \mathcal{S}^l_{adm}} \Phi^l_{\alpha,\delta}\left(\hat{\mathbf{s}}^l, \hat{\mathbf{s}}^{-l}, s^f, \lambda^l\right).$$

Having

$$\begin{gathered} F_{\alpha,\delta}\left(s^f, \tilde{s}^f, s^l, \lambda^f\right) \leq \varepsilon \text{ for any } s^f \in \mathcal{S}^f_{adm} \text{ and } f = \overline{1,m},\ \varepsilon \geq 0, \\ F_{\alpha,\delta}\left(s^f, \tilde{s}^f, s^l, \lambda^f\right) := \sum_{f \in F}\left[\Psi^f_{\alpha,\delta}\left(\hat{\mathbf{s}}^f, \hat{\mathbf{s}}^{-f}, s^l, \hat{\lambda}^f\right) - \Psi^f_{\alpha,\delta}\left(\mathring{\mathbf{s}}^f, \hat{\mathbf{s}}^{-f}, s^l, \hat{\lambda}^f\right)\right] \leq \varepsilon \end{gathered} \tag{10}$$

where $\tilde{s}^f \in \mathcal{S}^f_{adm}$ is the complement of s^f and s^l is fixed and

$$\mathring{\mathbf{s}}^f := \arg\max_{\hat{s}^f \in \mathcal{S}^f_{adm}} \Psi_{\alpha,\delta}\left(\hat{\mathbf{s}}^f, \hat{\mathbf{s}}^{-f}, s^l, \lambda^f\right).$$

Note that the condition $G_{\alpha,\delta}\left(s^l, \tilde{s}^l, s^f, \lambda^l\right) \leq \varepsilon$ and $F_{\alpha,\delta}\left(s^f, \tilde{s}^f, s^l, \lambda^f\right) \leq \varepsilon$ are equivalent to

$$\max_{\lambda^l \in \Lambda} \max_{s^l \in \mathcal{S}^l_{adm}} G_{\alpha,\delta}\left(s^l, \tilde{s}^l, s^f, \lambda^l\right), \tag{11}$$

and

$$\min_{\lambda^f \in \Lambda} \max_{s^f \in \mathcal{S}_{adm}} F_{\alpha,\delta}\left(s^f, \tilde{s}^f, s^l, \lambda^f\right). \tag{12}$$

Then

$$s^{l**}_{G_{\alpha,\delta}} \in Arg\max_{\lambda^l \in \Lambda} \max_{s^l \in \mathcal{S}^l_{adm}} G_{\alpha,\delta}\left(s^l, \tilde{s}^l, s^f, \lambda^l\right),$$

and

$$s^{f**}_{F_{\alpha,\delta}} \in Arg\min_{\lambda^f \in \Lambda} \max_{s^f \in \mathcal{S}_{adm}} F_{\alpha,\delta}\left(s^f, \tilde{s}^f, s^l, \lambda^f\right).$$

Definition 1. *A strategy $s^{l**}_{G_{\alpha,\delta}}$ of the manipulating player, together with the strategy $s^{f**}_{F_{\alpha,\delta}}$, is said to be a manipulating equilibrium (for $\varepsilon = 0$) if they fulfill*

$$\begin{gathered} \left(s^{l**}_{G_{\alpha,\delta}}, s^{f**}_{F_{\alpha,\delta}}\right) \in \arg\max_{\lambda^l \in \Lambda} \min_{\lambda^f \in \Lambda} \max_{s^l \in \mathcal{S}^l_{adm}} \max_{s^f \in \mathcal{S}^f_{adm}} \\ \left\{\Phi_{\alpha,\delta}\left(s^l, s^f, \lambda^l\right) \middle| G_{\alpha,\delta}\left(s^l, \tilde{s}^l, s^f, \lambda^l\right) \leq 0, F_{\alpha,\delta}\left(s^f, \tilde{s}^f, s^l, \lambda^f\right) \leq 0\right\}. \end{gathered} \tag{13}$$

Let us introduce the penalty approach to represent (13) as follows:

$$\mathbb{P}(s^l, \tilde{s}^l, s^f, \tilde{s}^f, \lambda^l, \lambda^f) \to \max_{\lambda^l \in \Lambda} \min_{\lambda^f \in \Lambda} \max_{s^l \in \mathcal{S}^l_{adm}} \max_{s^f \in \mathcal{S}^f_{adm}}$$

where

$$\mathbb{P}_{\alpha,\delta}(s^l,\tilde{s}^l,s^f,\tilde{s}^f,\lambda^l,\lambda^f) = \Phi_{\alpha,\delta}\left(s^l,s^f,\lambda^l\right) - \tfrac{1}{2}\left\|G_{\alpha,\delta}\left(s^l,\tilde{s}^l,s^f,\lambda^l\right)\right\|^2 - \tfrac{1}{2}\left\|F_{\alpha,\delta}\left(s^f,\tilde{s}^f,s^l,\lambda^f\right)\right\|^2. \tag{14}$$

4.2. Proximal Format

The problem (14) can be represented in the following proximal format

$$\begin{aligned}
\lambda_\delta^{l*} &= \arg\max_{\lambda^l\in\Lambda}\left\{-\tfrac{1}{2}\left\|\lambda^l-\lambda_\delta^{l*}\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_\delta^{l*},\tilde{s}_\delta^{l*},s_\delta^{f*},\tilde{s}_\delta^{f*},\lambda^l,\lambda_\delta^{f*})\right\}\\
s_\delta^{l*} &= \arg\max_{s^l\in\mathcal{S}_{adm}^l}\left\{-\tfrac{1}{2}\left\|s^l-s_\delta^{l*}\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s^l,\tilde{s}_\delta^{l*},s_\delta^{f*},\tilde{s}_\delta^{f*},\lambda_\delta^{l*},\lambda_\delta^{f*})\right\}\\
\tilde{s}_\delta^{l*} &= \arg\max_{\tilde{s}^l\in\mathcal{S}_{adm}^l}\left\{-\tfrac{1}{2}\left\|\tilde{s}^l-\tilde{s}_\delta^{l*}\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_\delta^{l*},\tilde{s}^l,s_\delta^{f*},\tilde{s}_\delta^{f*},\lambda_\delta^{l*},\lambda_\delta^{f*})\right\}\\
\lambda_\delta^{f*} &= \arg\min_{\lambda^f\in\Lambda}\left\{\tfrac{1}{2}\left\|\lambda^f-\lambda_\delta^{f*}\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_\delta^{l*},\tilde{s}_\delta^{l*},s_\delta^{f*},\tilde{s}_\delta^{f*},\lambda_\delta^{l*},\lambda^f)\right\}\\
s_\delta^{f*} &= \arg\max_{s^f\in\mathcal{S}_{adm}^f}\left\{-\tfrac{1}{2}\left\|s^f-s_\delta^{f*}\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_\delta^{l*},\tilde{s}_\delta^{l*},s^f,\tilde{s}_\delta^{f*},\lambda_\delta^{l*},\lambda_\delta^{f*})\right\}\\
\tilde{s}_\delta^{f*} &= \arg\max_{\tilde{s}^f\in\mathcal{S}_{adm}^f}\left\{-\tfrac{1}{2}\left\|\tilde{s}^f-\tilde{s}_\delta^{f*}\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_\delta^{l*},\tilde{s}_\delta^{l*},s_\delta^{f*},\tilde{s}^f,\lambda_\delta^{l*},\lambda_\delta^{f*})\right\}
\end{aligned}$$

where the solutions depend of small parameters $\gamma>0$, $\alpha>0$ and $\delta>0$.

The proximal method for calculating the manipulation equilibrium for initial variables $\lambda_0^l, s_0^l, \tilde{s}_0^l, \lambda_0^f, s_0^f, \tilde{s}_0^f$ is given by

$$\left.\begin{aligned}
\lambda_{n+1}^l &= \arg\max_{\lambda^l\in\Lambda}\left\{-\tfrac{1}{2}\left\|\lambda^l-\lambda_n^l\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_n^l,\tilde{s}_n^l,s_n^f,\tilde{s}_n^f,\lambda^l,\lambda_n^f)\right\}\\
s_{n+1}^l &= \arg\max_{s^l\in\mathcal{S}_{adm}^l}\left\{-\tfrac{1}{2}\left\|s^l-s_n^l\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s^l,\tilde{s}_n^l,s_n^f,\tilde{s}_n^f,\lambda_n^l,\lambda_n^f)\right\}\\
\tilde{s}_{n+1}^l &= \arg\max_{\tilde{s}^l\in\mathcal{S}_{adm}^l}\left\{-\tfrac{1}{2}\left\|\tilde{s}^l-\tilde{s}_n^l\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_n^l,\tilde{s}^l,s_n^f,\tilde{s}_n^f,\lambda_n^l,\lambda_n^f)\right\}\\
\lambda_{n+1}^f &= \arg\min_{\lambda^f\in\Lambda}\left\{\tfrac{1}{2}\left\|\lambda^f-\lambda_n^f\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_n^l,\tilde{s}_n^l,s_n^f,\tilde{s}_n^f,\lambda_n^l,\lambda^f)\right\}\\
s_{n+1}^f &= \arg\max_{s^f\in\mathcal{S}_{adm}^f}\left\{-\tfrac{1}{2}\left\|s^f-s_n^f\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_n^l,\tilde{s}_n^l,s^f,\tilde{s}_n^f,\lambda_n^l,\lambda_n^f)\right\}\\
\tilde{s}_{n+1}^f &= \arg\max_{\tilde{s}^f\in\mathcal{S}_{adm}^f}\left\{-\tfrac{1}{2}\left\|\tilde{s}^f-\tilde{s}_n^f\right\|^2+\gamma\mathbb{P}_{\alpha,\delta}(s_n^l,\tilde{s}_n^l,s_n^f,\tilde{s}^f,\lambda_n^l,\lambda_n^f)\right\}
\end{aligned}\right\} \tag{15}$$

To guarantee the convergence of the suggested procedure, let us select the parameters of the algorithm as follows:

$$\begin{gathered}
\delta_n = \begin{cases} \delta_0 & \text{if} \quad n\le n_0 \\ \delta_0\frac{[1+\ln(n-n_0)]}{(1+n-n_0)^\delta} & \text{if} \quad n>n_0 \end{cases},\ \mu_n = \begin{cases} \mu_0 & \text{if} \quad n<n_0 \\ \frac{\mu_0}{(1+n-n_0)^\alpha} & \text{if} \quad n\ge n_0 \end{cases}\\
\gamma_n = \begin{cases} \gamma_0 & \text{if} \quad n<n_0 \\ \frac{\gamma_0}{(1+n-n_0)^\gamma} & \text{if} \quad n\ge n_0 \end{cases},\ \delta,\mu,\gamma>0,\ \delta_0,\mu_0,\gamma_0>0,
\end{gathered} \tag{16}$$

which satisfies $\frac{\mu_n}{\delta_n} \underset{n\to\infty}{\to} 0$ and

$$\delta\le\mu,\ \gamma\ge\delta,\ \gamma+\delta\le 1.$$

5. Markov Games and Moral Hazard

5.1. Markov Model

We consider a discrete-time Markov game played by a set $L = \{1, \dots, n\}$ of *manipulating players* indexed by $l = \overline{1,n}$ and a set $F = \{1, \dots, m\}$ of manipulated players indexed by $f = \overline{1,m}$, where $\mathcal{I} = \{1, \dots, n+m\}$ is the total set of Machiavellian players indexed by $i = \overline{1,n+m}$.

The main results of the paper are as follows:

1. The time is discrete, taking values from $T = \{0, 1, \dots\}$ and the horizon is finite. The Machiavellian players $i \in \mathcal{I}$ are in a sequential game.
2. At each time $t \geq 0$, $t \in T$, the *type* θ_t^l of the manipulating players l is selected from the finite set $\theta_t^l \in \Theta^l$ and revealed to the manipulating players l. In the same manner, the manipulated players f are privately informed about their type $\theta_t^f \in \Theta^f$.
3. For each Machiavellian player $i \in \mathcal{I}$, we use $\Theta = \times_{i \in \mathcal{I}} \Theta^i$ to denote the *set of type profiles* and $\Theta^{-i} = \times_{j \setminus i \in \mathcal{I}} \Theta^j$ to denote the *set of type tuples of all players, except i*.
4. Manipulating players to take an action a_t^l (make a decision) from a finite set $a_t^l \in A^l$, $\mathcal{A}^l = \times_{l \in L} A^l$, and $\mathcal{A}^{-i} = \times_{h \setminus l \in L} A^h$, and manipulated players also take an action $a_t^f \in A^f$, $\mathcal{A}^f = \times_{f \in F} A^f$, and $\mathcal{A}^{-f} = \times_{h \setminus f \in F} A^h$.
5. We use $\Delta(\mathcal{A}^l)(\Delta(\mathcal{A}^f))$ for the set of all probability distributions over $\mathcal{A}^l$ $(\mathcal{A}^f)$ for all l (f).
6. For each Machiavellian player $i \in \mathcal{I}$, the game begins with a belief of an invariant measure $P^i(\theta_0^i) \in \Delta(\Theta^i)$, such that $P^i(\theta_t^i) \in \Delta(\Theta^i)$, where $\Delta(\Theta^i)$ denotes the set of all probability distributions over Θ^i. The beliefs may differ across types, and we do not assume that the Machiavellian players' beliefs are derived from a common prior. However, we assume that all types of players agree on which types have positive or zero probabilities. From now on, we assume that each belief system satisfies $P^i(\theta_t^i) > 0$. We say that type θ_t^i has $P^i(\theta_t^i)$-positive probability if $\theta_t^i \in \Theta^i$ and $P^i(\theta_t^i)$-zero probability otherwise.
7. The transition functions are denoted by

$$p^l(\theta_{t+1}^l | \theta_t^l, a_t^l, \theta_{t-1}^l, a_{t-1}^l, \dots, \theta_0^l, a_0^l) = p^l(\theta_{t+1}^l | \theta_t^l, a_t^l),$$

and

$$p^f(\theta_{t+1}^f | \theta_t^f, a_t^f, \theta_{t-1}^f, a_{t-1}^f, \dots, \theta_0^f, a_0^f) = p^f(\theta_{t+1}^f | \theta_t^f, a_t^f),$$

for the manipulating and manipulated players, respectively.
8. Each chain $(P^l, p^l(\theta_{t+1}^l | \theta_t^l, a_t^l))$ and $(P^f, p^f(\theta_{t+1}^f | \theta_t^f, a_t^f))$ follows controllable, time homogeneous, irreducible, and aperiodic Markov chains. These chains take values in (Θ^l, A^l) and (Θ^f, A^f), respectively.
9. Manipulating players have a known *utility function* $u^l : A^l \times \Theta^l \to \mathbb{R}^+$, which depends on the action $a_t^l \in A^l$ and the privately known type $\theta_t^l \in \Theta^l$. In the same manner, we define the *utility function* $u^f : A^f \times \Theta^f \to \mathbb{R}^+$ for the manipulated players.

Let us relate the message m_t^l for each manipulating player $l \in L$ with its type $m_t^l \in M^l \subseteq \Theta^l$, and the message m_t^f for each manipulated player $f \in F$ with its type $m_t^f \in M^f \subseteq \Theta^f$. Here, $M^i \subseteq \Theta^i$ is the set of messages, such that at time t, a Machiavellian player i sends a message m_t^i to all players that respond with $a_t = (a_t^1, a_t^2, \dots, a_t^{n+m})$. We assume that $M^i = \Theta^i$.

5.2. Strategies and Moral Hazards

The relationship $z^l(\alpha_t^l|a_t^l)$ for the manipulating players is given by $z^l : A^l \to \Delta(A^l)$ and for the manipulated players $z^f(\alpha_t^f|a_t^f)$ is $z^f : A^f \to \Delta(A^f)$. The set of admissible actions is defined as follows:

$$Z_{adm}^l = \left\{ z^l(\alpha_t^l|a_t^l) | \sum_{\alpha_t^l \in A^l} z^l(\alpha_t^l|a_t^l) = 1, a_t^l \in A^l \right\},$$

and

$$Z_{adm}^f = \left\{ z^f(\alpha_t^f|a_t^f) | \sum_{\alpha_t^f \in A^f} z^f(\alpha_t^f|a_t^f) = 1, a_t^f \in A^f \right\}.$$

The relationship $z^i(\alpha_t^i|a_t^i)$ represents the likelihood in which a Machiavellian player i believes that α_t^i is an action a_t^i.

A *strategy* $\sigma^l(m_t^l|\theta_t^l)$ (behavioral strategy) for manipulating players is $\sigma^l : \Theta^l \to \Delta(M^l)$, which represents the likelihood in which a manipulating player l believes that a message m_t^l is of type θ_t^l. For the manipulated players, a strategy $\sigma^f(m_t^f|\theta_t^f)$ is represented by $\sigma^f : \Theta^f \to \Delta(M^f)$. The admissible strategy set is given by

$$\mathfrak{S}_{adm}^l = \left\{ \sigma^l(m_t^l|\theta_t^l) | \sum_{m_t^l \in M^l} \sigma^l(m_t^l|\theta_t^l) = 1, \theta_t^l \in \Theta^l \right\},$$

and

$$\mathfrak{S}_{adm}^f = \left\{ \sigma^f(m_t^f|\theta_t^f) | \sum_{m_t^f \in M^f} \sigma^f(m_t^f|\theta_t^f) = 1, \theta_t^f \in \Theta^f \right\}.$$

A *policy* for the manipulating players is defined as a sequence $\left\{ \pi^l(\alpha_t^l|m_t^l) \right\}$, such that, for each time, t, $\pi^l(\alpha_t^l|m_t^l)$ is a stochastic kernel, and for the manipulated players, we have $\pi^f(\alpha_t^f|m_t^f)$. The set of all admissible policies is denoted as follows:

$$\Pi_{adm}^l = \left\{ \pi^l(\alpha_t^l|m_t^l) | \sum_{\alpha_t^l \in A^l} \pi^l(\alpha_t^l|m_t^l) = 1 \right\},$$

and

$$\Pi_{adm}^f = \left\{ \pi^f(\alpha_t^f|m_t^f) | \sum_{\alpha_t^f \in A^f} \pi^f(\alpha_t^f|m_t^f) = 1 \right\}.$$

We are interested in the average criterion, where the realized average reward function for the Manipulating players is given by

$$U^l(\pi, \sigma, z) = \sum_{\theta_t^l \in \Theta^l} \sum_{m_t^l \in M^l} \sum_{a_t^l \in A^l} \sum_{\alpha_t^l \in A^l} \left(\sum_{\theta_{t+1}^l \in \Theta^l} u^l(a_t^l, \theta_t^l) p^l(\theta_{t+1}^l|\theta_t^l, a_t^l) \right) \cdot$$

$$\prod_{l \in L} \pi^l(\alpha_t^l|m_t^l) \sigma^l(m_t^l|\theta_t^l) z^l(\alpha_t^l|a_t^l) P^l(\theta_t^l) \prod_{f \in F} \pi^f(\alpha_t^f|m_t^f) \sigma^f(m_t^f|\theta_t^f) z^f(\alpha_t^f|a_t^f) P^f(\theta_t^f) = \tag{17}$$

$$\sum_{\theta_t^l \in \Theta^l} \sum_{m_t^l \in M^l} \sum_{a_t^l \in A^l} \sum_{\alpha_t^l \in A^l} \left(\sum_{\theta_{t+1}^l \in \Theta^l} u^l(a_t^l, \theta_t^l) p^l(\theta_{t+1}^l|\theta_t^l, a_t^l) \right) \prod_{i \in I} \pi^i(\alpha_t^i|m_t^i) \sigma^i(m_t^i|\theta_t^i) z^i(\alpha_t^i|a_t^i) P^i(\theta_t^i),$$

and for the manipulated players

$$U^f(\pi,\sigma,z)=\sum_{\theta_t^f\in\Theta^f}\sum_{m_t^f\in M^f}\sum_{a_t^f\in A^f}\sum_{\alpha_t^f\in A^f}\left(\sum_{\theta_{t+1}^f\in\Theta^f}u^f(a_t^f,\theta_t^f)p^f(\theta_{t+1}^f|\theta_t^f,a_t^f)\right)\cdot \prod_{i\in I}\pi^i(\alpha_t^i|m_t^i)\sigma^i(m_t^i|\theta_t^i)z^i(\alpha_t^i|a_t^i)P^i(\theta_t^i). \tag{18}$$

Definition 2. *The interaction between the Machiavellian players induces a manipulation game, given by*

$$G=(I,\Theta^i,A^i,P^i,U^i)_{i\in I}.$$

5.3. Auxiliary Variable

Let us define for the manipulating players:

$$c^l(\theta_t^l,m_t^l,a_t^l,\alpha_t^l)=\pi^l(\alpha_t^l|m_t^l)\sigma^l(m_t^l|\theta_t^l)z^l(\alpha_t^l|a_t^l)P^l(\theta_t^l)$$

and for the manipulated players:

$$c^f(\theta_t^f,m_t^f,a_t^f,\alpha_t^f)=\pi^f(\alpha_t^f|m_t^f)\sigma^f(m_t^f|\theta_t^f)z^f(\alpha_t^f|a_t^f)P^f(\theta_t^f),$$

with

$$C_{adm}^l:=\left\{c^l(\theta_t^l,m_t^l,a_t^l,\alpha_t^l)\Big|\sum_{\theta_t^l\in\Theta^l}\sum_{m_t^l\in M^l}\sum_{a_t^l\in A^l}\sum_{\alpha_t^l\in A^l}c^l(\theta_t^l,m_t^l,a_t^l,\alpha_t^l)=1\right.$$

$$\sum_{m_t^l\in M^l}\sum_{a_t^l\in A^l}\sum_{\alpha_t^l\in A^l}c^l(\theta_t^l,m_t^l,a_t^l,\alpha_t^l)=P^l(\theta_t^l)>0,$$

$$\left.\sum_{\theta_t^l\in\Theta^l}\sum_{m_t^l\in M^l}\sum_{a_t^l\in A^l}\sum_{\alpha_t^l\in A^l}\sum_{\theta_{t+1}^l\in\Theta^l}\left[\delta_{\theta_t^l\theta_{t+1}^l}-p^l(\theta_{t+1}^l|\theta_t^l,a_t^l)\right]c^l(\theta_t^l,m_t^l,a_t^l,\alpha_t^l)=0\right\},$$

where $\delta_{\theta_t^l\theta_{t+1}^l}$ is the Kronecker symbol, such that

$$\Delta^l:=\left\{c^l(\theta_t^l,m_t^l,a_t^l,\alpha_t^l)\Big|\sum_{\theta_t^l\in\Theta^l}\sum_{m_t^l\in M^l}\sum_{a_t^l\in A^l}\sum_{\alpha_t^l\in A^l}c^l(\theta_t^l,m_t^l,a_t^l,\alpha_t^l)=1\right.$$

$$\left.\sum_{m_t^l\in M^l}\sum_{a_t^l\in A^l}\sum_{\alpha_t^l\in A^l}c^l(\theta_t^l,m_t^l,a_t^l,\alpha_t^l)=P^l(\theta_t^l)>0\right\},$$

and

$$C_{adm}^f:=\left\{c^f(\theta_t^f,m_t^f,a_t^f,\alpha_t^f)\Big|\sum_{\theta_t^f\in\Theta^f}\sum_{m_t^f\in M^f}\sum_{a_t^f\in A^f}\sum_{\alpha_t^f\in A^f}c^f(\theta_t^f,m_t^f,a_t^f,\alpha_t^f)=1\right.$$

$$\sum_{m_t^f\in M^f}\sum_{a_t^f\in A^f}\sum_{\alpha_t^f\in A^f}c^f(\theta_t^f,m_t^f,a_t^f,\alpha_t^f)=P^f(\theta_t^f)>0,$$

$$\left.\sum_{\theta_t^f\in\Theta^f}\sum_{m_t^f\in M^f}\sum_{a_t^f\in A^f}\sum_{\alpha_t^f\in A^f}\sum_{\theta_{t+1}^f\in\Theta^f}\left[\delta_{\theta_t^f\theta_{t+1}^f}-p^f(\theta_{t+1}^f|\theta_t^f,a_t^f)\right]c^f(\theta_t^f,m_t^f,a_t^f,\alpha_t^f)=0\right\},$$

where $\delta_{\theta_t^f \theta_{t+1}^f}$ is the Kronecker symbol, such that

$$\Delta^f := \Bigg\{ c^f(\theta_t^f, m_t^f, a_t^f, \alpha_t^f) \Big| \sum_{\theta_t^f \in \Theta^f} \sum_{m_t^f \in M^f} \sum_{a_t^f \in A^f} \sum_{\alpha_t^f \in A^f} c^f(\theta_t^f, m_t^f, a_t^f, \alpha_t^f) = 1$$
$$\sum_{m_t^f \in M^f} \sum_{a_t^f \in A^f} \sum_{\alpha_t^f \in A^f} c^f(\theta_t^f, m_t^f, a_t^f, \alpha_t^f) = P^f(\theta_t^f) > 0 \Bigg\}.$$

The individual aim in the c-variables can be expressed as follows:

$$U^l(\pi, \sigma, z) = \sum_{l \in L} \tilde{U}^l(c) \to \max_{c^l \in C_{adm}^l} \tag{19}$$

where

$$\tilde{U}^l(c) = \sum_{\theta_t^l \in \Theta^l} \sum_{m_t^l \in M^l} \sum_{a_t^l \in A^l} \sum_{\alpha_t^l \in A^l} W(\theta_t^l, a_t^l) \prod_{l \in L} c^l(\theta_t^l, m_t^l, a_t^l, \alpha_t^l) \underbrace{\prod_{f \in F} c^f(\theta_t^f, m_t^f, a_t^f, \alpha_t^f)}_{fixed}$$

$$W(\theta_t^l, a_t^l) = \sum_{\theta_{t+1}^l \in \Theta^l} u^l(a_t^l, \theta_t^l) p^l(\theta_{t+1}^l | \theta_t^l, a_t^l),$$

and

$$U^f(\pi, \sigma, z) = \sum_{f \in F} \tilde{U}^f(c) \to \max_{c^l \in C_{adm}^l} \tag{20}$$

where

$$\tilde{U}^f(c) = \sum_{\theta_t^f \in \Theta^f} \sum_{m_t^f \in M^f} \sum_{a_t^f \in A^f} \sum_{\alpha_t^f \in A^f} W(\theta_t^f, a_t^f) \underbrace{\prod_{l \in L} c^l(\theta_t^l, m_t^l, a_t^l, \alpha_t^l)}_{fixed} \prod_{f \in F} c^f(\theta_t^f, m_t^f, a_t^f, \alpha_t^f),$$

$$W(\theta_t^f, a_t^f) = \sum_{\theta_{t+1}^f \in \Theta^f} u^f(a_t^f, \theta_t^f) p^f(\theta_{t+1}^f | \theta_t^f, a_t^f).$$

Definition 3. *We note that $\tilde{U}^i(c)$ is considered individual rational for the Machiavellian player i if it is the case that*

$$IR = \Big\{ \tilde{U}^i(c) | \lambda^i \tilde{U}^i(c) \geq \tilde{U}^i(c), \lambda \in \Lambda, i \in I \Big\}.$$

Remark 2. *The foundation of rational decision theory, an economic theory that asserts that Machiavellian players always choose options that maximize their own benefit, is rational conduct. Given the available options, these choices offer the Machiavellian players the greatest benefit or satisfaction.*

5.4. Machiavellian Equilibrium

The policies π^*, strategies σ^*, and action kernel z^* that solve the nonlinear programming problem are given by

$$(\pi^*, \sigma^*, z^*) = \arg \max_{\pi^* \in \Pi_{adm}^i} \sum_{i \in I} U^i(\pi, \sigma^*, z^*),$$

where π^*, σ^* and z^* fulfill the *Machiavellian equilibrium,* satisfying

$$U^i(\pi^*, \sigma^*, z^*) \geq U^i(\pi^i, \pi^{-i*}, \sigma^i, \sigma^{-i*}, z^i, z^{-i*}),$$

such that $\pi^{-i*} = (\pi^{1*}, \ldots, \pi^{i-1*}, \pi^{i+1*}, \ldots, \pi^{n+m*})$, $\sigma^{-i*} = (\sigma^{1*}, \ldots, \sigma^{i-1*}, \sigma^{i+1*}, \ldots, \sigma^{n+m*})$, and $z^{-i*} = (z^{1*}, \ldots, z^{i-1*}, z^{i+1*}, \ldots, z^{n+m*})$.

A strategy profile is a *Machiavellian equilibrium* in the manipulation game $G = (I, \Theta^i, A^i, P^i, U^i)_{i\in I}$ if the tuple (π^*, σ^*, z^*) represents the best reply to $(\pi^i, \pi^{-i*}, \sigma^i, \sigma^{-i*}, z^i, z^{-i*})$. Any reward in a Machiavellian equilibrium must belong to the set IR.

A strategy profile is a *sequential Machiavellian equilibrium* in the game $G = (I, \Theta^i, A^i, P^i, U^i)_{i\in I}$ if there exists a tuple (π^*, σ^*, z^*) and history h_t^i, such that, for each player i, the tuple (π^*, σ^*, z^*) is the best reply to $(\pi^i, \pi^{-i*}, \sigma^i, \sigma^{-i*}, z^i, z^{-i*})$.

Remark 3. *The manipulation game $G = (I, \Theta^i, A^i, P^i, U^i)_{i\in I}$ is a sequential game, where the manipulating players play first and the manipulated players play second.*

5.5. Variable Association

Associating these variables with the notations above, let us introduce the following vector:

$$s = s(c) := (c^1(\theta_t^1, m_t^1, a_t^1, \alpha_t^1), \ldots, c^{n+m}(\theta_t^{n+m}, m_t^{n+m}, a_t^{n+m}, \alpha_t^{n+m}))^\intercal \in \mathcal{S}_{adm},\ c^i(\theta_t^i, m_t^i, a_t^i, \alpha_t^i) \in C_{adm}^i$$

and the functions

$$u^i(s(c)) = \tilde{U}^i(c), i \in I.$$

Then we have

$$\left.\begin{array}{c} s^*(\lambda) \in \underset{s\in\mathcal{S}_{adm}}{Arg\min} \sum\limits_{i\in I} \lambda^i u^i(s) \Leftrightarrow c^*(\lambda) \in \underset{c\in C_{adm}^i}{Arg\min} \sum\limits_{i\in I} \lambda^i \tilde{U}^i(c) \\ \\ s^*(\lambda) = s(c^*(\lambda)). \end{array}\right\}$$

Hence,

$$\max_{\lambda^l\in\Lambda} \max_{s^l\in\mathcal{S}_{adm}^l} \varphi(s^l, s^f, \lambda^l) = \max_{\lambda^l\in\Lambda} \max_{s^l\in C_{adm}^l} \phi(c^l, c^f, \lambda^l),$$

$$\min_{\lambda^f\in\Lambda} \max_{s^f\in\mathcal{S}_{adm}^f} \psi(s^f, s^l, \lambda^f) = \min_{\lambda^f\in\Lambda} \max_{c^f\in C_{adm}^f} \varrho(c^f, c^l, \lambda^f),$$

where

$$\phi(c^l, c^f, \lambda^l) = \sum_{l\in L} \lambda^l \tilde{U}^l(c) - \frac{\delta}{2}\left(\left\|c^l\right\|^2 + \left\|\lambda^l\right\|^2\right),\ \delta > 0,$$

$$\varrho(c^f, c^l, \lambda^f) = \sum_{f\in F} \lambda^f \tilde{U}^f(c) - \frac{\delta}{2}\left(\left\|c^f\right\|^2 - \left\|\lambda^f\right\|^2\right),\ \delta > 0.$$

5.6. Recover the Relationship

Lemma 1. *Let us assume that Equations (19) and (20) are solved, then the variables $\pi^{i*}(\alpha_t^i|m_t^i)$, $i \in I$, can be recovered from $c^{i*}(\theta_t^i m_t^i a_t^i \alpha_t^i)$ as follows:*

$$\pi^{i*}(\alpha_t^i|m_t^i) = \frac{1}{|\Theta^i|} \sum_{\theta_t^i\in\Theta^i} \frac{1}{|A^i|} \frac{c^i(\theta_t^i m_t^i a_t^i \alpha_t^i)}{\sum\limits_{\zeta_t^i\in A^i} c^i(\theta_t^i m_t^i a_t^i \zeta_t^i)}. \tag{21}$$

Proof. Let

$$\pi^{i*}(\alpha_t^i|\theta_t^i m_t^i) = \begin{cases} \frac{1}{|A^i|} \dfrac{c^i(\theta_t^i m_t^i a_t^i \alpha_t^i)}{\sum\limits_{\zeta_t^i\in A^i} c^i(\theta_t^i m_t^i a_t^i \zeta_t^i)} & if \quad \sum\limits_{\alpha_t^i\in A^i} c^i(\theta_t^i m_t^i a_t^i \alpha_t^i) > 0, \\ \\ 0 & \sum\limits_{\alpha_t^i\in A^i} c^i(\theta_t^i m_t^i a_t^i \alpha_t^i) = 0. \end{cases}$$

It then follows from

$$\pi^{i*}(\alpha_t^i|m_t^i) = \frac{1}{|\Theta^i|}\sum_{\theta_t^i\in\Theta^i}\frac{1}{|A^i|}\frac{c^i(\theta_t^i m_t^i a_t^i \alpha_t^i)}{\sum\limits_{\zeta_t^i\in A^i} c^i(\theta_t^i m_t^i a_t^i \zeta_t^i)}.$$

that $\sum\limits_{\alpha_t^i\in A^i}\pi^{i*}(\alpha_t^i|m_t^i) = 1$. □

To recover $\sigma^{i*}(m_t^i|\theta_t^i)$, the formula is given by

$$\sigma^{i*}(m_t^i|\theta_t^i) = \frac{\sum\limits_{a_t^i\in A^i}\sum\limits_{\alpha_t^i\in A^i} c^i(\theta_t^i m_t^i a_t^i \alpha_t^i)}{\sum\limits_{\omega_t^i\in M^i}\sum\limits_{a_t^i\in A^i}\sum\limits_{\alpha_t^i\in A^i} c^i(\theta_t^i \omega_t^i a_t^i \alpha_t^i)}. \tag{22}$$

The formula to recover the distribution $P^{i*}(\theta_t^i)$ is

$$P^{i*}(\theta_t^i) = \sum_{m_t^i\in M^i}\sum_{a_t^i\in A^i}\sum_{\alpha_t^i\in A^i} c^i(\theta_t^i m_t^i a_t^i \alpha_t^i) > 0. \tag{23}$$

Finally,

$$z^{i*}(\alpha_t^i|a_t^i) = \frac{\sum\limits_{\theta_t^i\in\Theta^i}\sum\limits_{m_t^i\in M^i} c^{i*}(\theta_t^i m_t^i a_t^i \alpha_t^i)}{\sum\limits_{\theta_t^i\in\Theta^i}\sum\limits_{m_t^l\in M^i}\sum\limits_{\zeta_t^i\in A_t^i} c^{i*}(\theta_t^i m_t^i a_t^i \zeta_t^i)}. \tag{24}$$

6. Numerical Example

We use the notion of manipulation to study non-cooperative games with uncertain information. The forward and futures markets provide a clear illustration of several key problems associated with the economics of uncertainty and information asymmetry. These markets play a significant role in facilitating trade among firms to efficiently distribute risks, taking into account each agent's risk preferences. This topic has long been a subject of study in the literature due to its practical significance. For example, the examination of squeezes or corners, as referred to in institutional terms, brings up the issue of market manipulation. In both scenarios, it is implicitly assumed that certain corporations have access to strategic information and employ this advantage to influence the market in their favor, thereby impacting pricing. In this section, we present an example of how oligopolistic firms behave using manipulation. The game consists of four firms, and the main characteristic of this market is mutual interdependence, i.e., one firm's action influences the profits of its rivals. In the dynamics of the game, firm 1 and firm 2 play first, followed by firms 3 and firm 4. They are tempted to collude to fix a price (produce identical products) or production level (produce different products) in a market in order to maximize their profits. Colluding allows them to act as a monopoly. However, firms pursue their own self-interests, producing greater quantities than the other firms, leading to lower prices (i.e., petroleum). As collusive agreements are illegal, there is a threat that firms may defect by cheating on their associates and manipulating the market. Oligopolists reason that if they are the only ones cheating, it can increase their profits. Then each firm has an incentive to cheat.

Let $\mathcal{G}$ be a *non-cooperative game* determined $\mathcal{G} = (\mathcal{I}, \mathcal{S}, u^l, u^f)$. $\mathcal{I} = \{1, \ldots, 4\}$ is the set of total Machiavellian players where $L = \{1, 2\}$ is the set of *manipulating players* indexed by $l = \overline{1,2}$, and $F = \{3, 4\}$ is the set of *manipulated players* indexed by $f = \overline{3,4}$. In the dynamics of the game, the players take alternate turns. Manipulators commit first to a strategy and then the manipulated players' strategy is played. Let the number of states for each player be $\theta = 4$ and let $a = 2$ be the number of actions for each player.

We intend to reflect the ideas that firms are committed to their manipulation actions for a finite length of time, and that they react to the current manipulation actions of other firms. A firm's instantaneous price competition is given by Equations (17) and (18). The proximal method presented in Equation (15) was applied to calculate the manipulation equilibrium for the initial variables, $\delta_0 = 5.0 \times 10^{-4}$, $\mu_0 = \times 10^{-6}$ and $\gamma_0 = 4.85 \times 10^{-1}$. The resulting values of interest are as follows:

$$\pi^{1*}(a|m) = \begin{bmatrix} 0.5293 & 0.4707 \\ 0.4979 & 0.5021 \\ 0.4798 & 0.5202 \\ 0.5353 & 0.4647 \end{bmatrix}, \quad z^{1*}(\alpha|a) = \begin{bmatrix} 0.5038 & 0.4962 \\ 0.5634 & 0.4366 \end{bmatrix},$$

$$\sigma^{1*}(m|\theta) = \begin{bmatrix} 0.3005 & 0.2316 & 0.2345 & 0.2334 \\ 0.2504 & 0.2458 & 0.2718 & 0.2320 \\ 0.2636 & 0.2637 & 0.2775 & 0.1951 \\ 0.2267 & 0.2242 & 0.2433 & 0.3058 \end{bmatrix}, \quad P^{1*}(\theta) = \begin{bmatrix} 0.2860 \\ 0.2070 \\ 0.2740 \\ 0.2331 \end{bmatrix}.$$

$$\pi^{2*}(a|m) = \begin{bmatrix} 0.5276 & 0.4724 \\ 0.4974 & 0.5026 \\ 0.4820 & 0.5180 \\ 0.6041 & 0.3959 \end{bmatrix}. \quad z^{2*}(\alpha|a) = \begin{bmatrix} 0.5097 & 0.5388 \\ 0.4903 & 0.4612 \end{bmatrix},$$

$$\sigma^{2*}(m|\theta) = \begin{bmatrix} 0.3238 & 0.2240 & 0.2281 & 0.2241 \\ 0.2447 & 0.2451 & 0.2691 & 0.2411 \\ 0.2596 & 0.2668 & 0.2786 & 0.1950 \\ 0.2497 & 0.2491 & 0.2613 & 0.2399 \end{bmatrix}, \quad P^{2*}(\theta) = \begin{bmatrix} 0.2260 \\ 0.1840 \\ 0.3122 \\ 0.2778 \end{bmatrix}.$$

$$\lambda^{l*} = \begin{bmatrix} 0.1802 \\ 0.8198 \end{bmatrix}.$$

$$\pi^{3*}(a|m) = \begin{bmatrix} 0.4375 & 0.5625 \\ 0.5000 & 0.5000 \\ 0.5003 & 0.4997 \\ 0.3487 & 0.6513 \end{bmatrix}, \quad z^{3*}(\alpha|a) = \begin{bmatrix} 0.0124 & 0.9876 \\ 0.2182 & 0.7818 \end{bmatrix},$$

$$\sigma^{3*}(m|\theta) = \begin{bmatrix} 0.9953 & 0.0016 & 0.0016 & 0.0016 \\ 0.1897 & 0.1891 & 0.1879 & 0.4333 \\ 0.0018 & 0.0018 & 0.0018 & 0.9946 \\ 0.0090 & 0.0090 & 0.0089 & 0.9731 \end{bmatrix}, \quad P^{3*}(\theta) = \begin{bmatrix} 0.2534 \\ 0.3275 \\ 0.2239 \\ 0.1953 \end{bmatrix}.$$

$$\pi^{4*}(a|m) = \begin{bmatrix} 0.4375 & 0.5625 \\ 0.5000 & 0.5000 \\ 0.5004 & 0.4996 \\ 0.3285 & 0.6715 \end{bmatrix}, \quad z^{4*}(\alpha|a) = \begin{bmatrix} 0.0251 & 0.9749 \\ 0.1939 & 0.8061 \end{bmatrix},$$

$$\sigma^{4*}(m|\theta) = \begin{bmatrix} 0.9960 & 0.0013 & 0.0013 & 0.0013 \\ 0.1495 & 0.1479 & 0.1472 & 0.5554 \\ 0.0013 & 0.0013 & 0.0013 & 0.9961 \\ 0.0022 & 0.0022 & 0.0022 & 0.9934 \end{bmatrix}, \quad P^{4*}(\theta) = \begin{bmatrix} 0.2981 \\ 0.2092 \\ 0.3112 \\ 0.1815 \end{bmatrix}.$$

$$\lambda^{f*} = \begin{bmatrix} 0.6100 \\ 0.3900 \end{bmatrix}.$$

Depending on the interpretation of the oligopoly, action a can represent the choice of a price, quantity, location, etc. It can also represent a vector of choices. Firms act in discrete time, and the horizon is finite. The convergence of strategies for the manipulating players

is illustrated in Figures 1 and 2; the convergence of strategies for the manipulated players is shown in Figures 3 and 4.

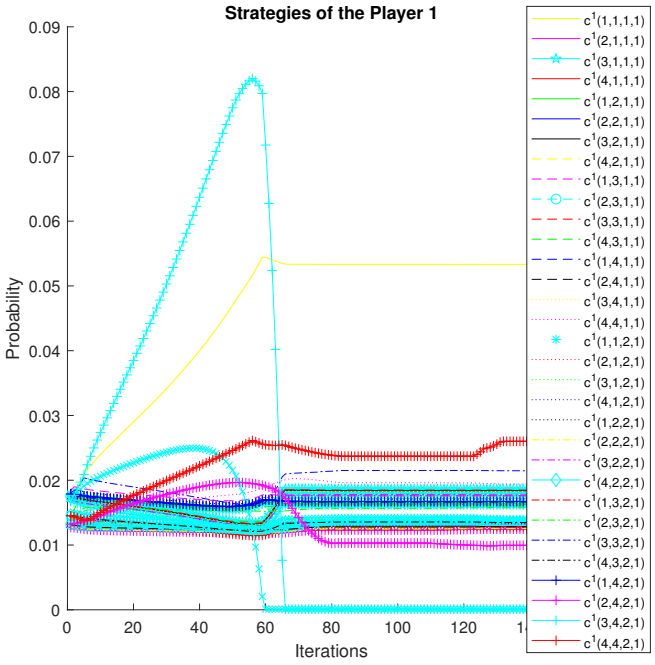

Figure 1. Convergence of strategies c for firm 1.

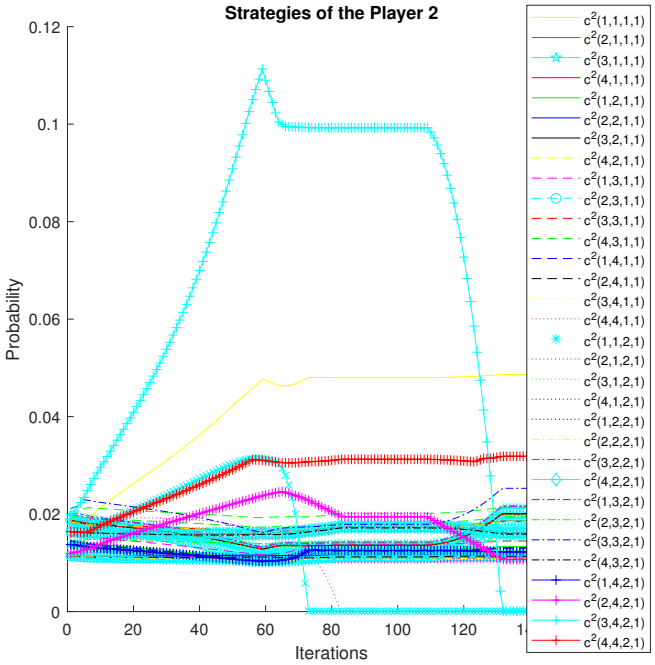

Figure 2. Convergence of strategies c for firm 2.

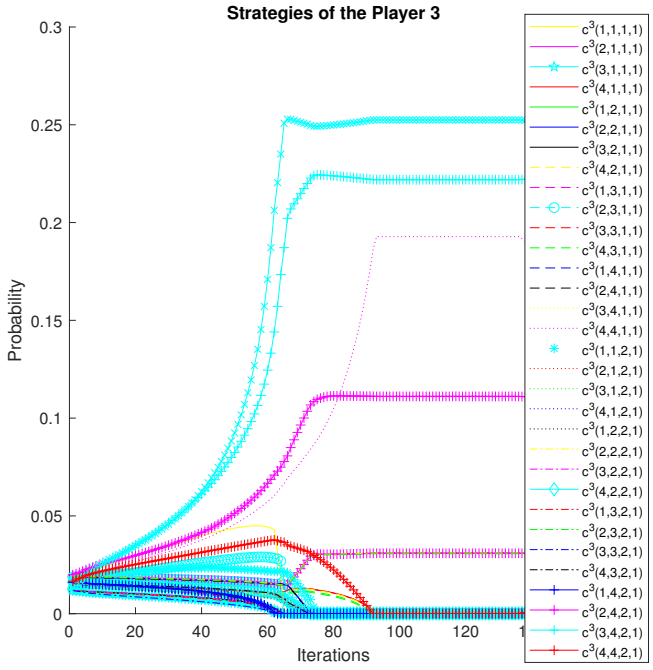

Figure 3. Convergence of strategies c for firm 3.

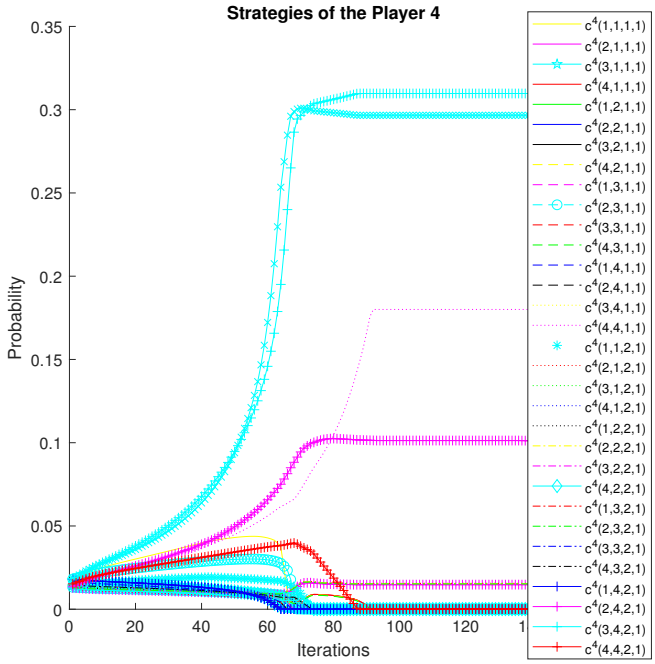

Figure 4. Convergence of strategies c for firm 4.

Based on the understanding of the strategies, the example consists of two key components. According to the manipulation model, firms 1 and 2 accumulate a significant number of positions without causing an increase in the price of these contracts. The manipulated firms 3 and 4 are compelled to maintain prices. A deeper examination of firms 1 and 2 reveals that they possess the necessary market sway to keep futures prices from rising while compelling firms 3 and 4 to maintain lower pricing and significantly strengthen their own positions.

7. Conclusions

This paper proposes a manipulative game. As part of our contribution, the game was modeled using the worst and best Nash equilibrium points. To address the issue, we constrained the problem to homogeneous, finite, ergodic, and controlled Bayesian–Markov games. Players who adhered to Machiavellian principles claimed to be in one state while they were truly in another and pretended to act one way while genuinely acting another. Each participant in the Machiavellian group engaged in some sort of social manipulation. Players who engaged in manipulation had a stronger preference than manipulated participants. The Pareto frontier is defined as the line where manipulating players play the best Nash equilibrium and manipulated players play the worst Nash equilibrium. In the case of multiple manipulators and manipulated players, it is also considered a sequential Bayesian–Markov manipulation game. Lastly, we provided a tractable characterization of the manipulation equilibrium findings. We employed Tikhonov's penalty regularization approach to ensure the convergence of the game's solution to a unique outcome. The outcomes of our concept are illustrated through a numerical example. Without a doubt, there are other issues that should be taken into account in future studies. We are thinking about expanding our method to handle Bayesian–Markov games and a hierarchical Stackelberg model. The paradigm might be expanded to support leaders and followers in a three-level model. Finding a practical economic use for the concept is an intriguing challenge for manipulation games.

Author Contributions: Conceptualization, J.S.-R. and J.B.C.; Investigation, J.S.-R. and J.B.C.; Writing—original draft, J.S.-R., J.M.R.-M. and J.B.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jones, D.N.; Paulhus, D.L. *Handbook of Individual Differences in Social Behavior*; Chapter Machiavellianism; The Guilford Press: New York, NY, USA, 2009; pp. 93–108.
2. Christie, R.; Geis, F. *Studies in Machiavellianism*; Academic Press: Cambridge, MA, USA, 2013.
3. Spielberger, C.D.; Butcher, J. *Advances in Personality Assessment*; Routledge: New York, NY, USA, 2013.
4. Geis, F. *Dimensions of Personality*; London, H., Exner, J.E., Jr., Eds.; Chapter Machiavellianism; Wiley: New York, NY, USA, 1978; pp. 305–363.
5. Nash, J.F. Non-cooperative games. *Ann. Math.* **1951**, *54*, 286–295. [CrossRef]
6. Cournot, A. *Recherches sur les Principes Mathématiques de la Théorie des Richesses*; The Macmillan Company: New York, NY, USA, 1838.
7. Allen, F.; Gorton, G. Stock price manipulation, market microstructure and asymmetric information. *Eur. Econ. Rev.* **1992**, *36*, 624–630. [CrossRef]
8. Bagnoli, M.; Lipman, B.L. Stock Price Manipulationthrough Takeover Bids. *Rand J. Econ.* **1996**, *27*, 124–147. [CrossRef]
9. Clempner, J.B. A game theory model for manipulation based on Machiavellianism: Moral and ethical behavior. *J. Artif. Soc. Soc. Simul.* **2017**, *20*, 12. [CrossRef]
10. Cumming, D.; Ji, S.; Peter, R.; Tarsalewska, M. Market manipulation and innovation. *J. Bank. Financ.* **2020**, *120*, 105957. [CrossRef]
11. Clempner, J.B.; Trejo, K. Manipulation Power in Bargaining Games Using Machiavellianism. *Econ. Comput. Econ. Cybern. Stud. Res.* **2021**, *55*, 299–313.
12. Clempner, J.B. Learning machiavellian strategies for manipulation in Stackelberg security games. *Ann. Math. Artif. Intell.* **2022**, *90*, 373–395. [CrossRef]
13. Clempner, J.B. A Manipulation Game Based on Machiavellian Strategies. *Int. Game Theory Rev.* **2022**, *24*, 2150015. [CrossRef]
14. Milgrom, P.; Roberts, J. Relying on the information of interested parties. *RAND J. Econ.* **1986**, *17*, 18–32. [CrossRef]
15. Krishna, V.; Morgan, J. A model of expertise. *Q. J. Econ.* **2001**, *116*, 747–775. [CrossRef]
16. Taneva, I. Information Design. *Am. Econ. J. Microecon.* **2019**, *11*, 151–185. [CrossRef]
17. Bardhi, A.; Guo, Y. Modes of persuasion toward unanimous consent. *Theor. Econ.* **2018**, *13*, 1111–1149. [CrossRef]
18. Kamenica, E.; Gentzkow, M. Bayesian Persuasion. *Am. Econ. Rev.* **2011**, *101*, 2590–2615. [CrossRef]

19. Bergemann, D.; Morris, S. Information design, Bayesian persuasion, and Bayes correlated equilibrium. *Am. Econ. Rev.* **2016**, *106*, 586–591. [CrossRef]
20. Gentzkow, M.; Kamenica, E. Bayesian persuasion with multiple senders and rich signal spaces. *Games Econ. Behav.* **2017**, *104*, 411–429. [CrossRef]
21. Brocas, I.; Carrillo, J.D.; Palfrey, T.R. Information gatekeepers: Theoryand experimental evidence. *Econ. Theory* **2012**, *51*, 649–676. [CrossRef]
22. Gul, F.; Pesendorfer, W. The war of information. *Rev. Econ. Stud.* **2012**, *79*, 707–734. [CrossRef]
23. Eső, P.; Szentes, B. Optimal Information Disclosure in Auctionsand the Handicap Auction. *Rev. Econ. Stud.* **2007**, *74*, 705–731. [CrossRef]
24. Bergemann, D.; Pesendorfer, M. Information structuresin optimal auctions. *J. Econ. Theory* **2007**, *137*, 580–609. [CrossRef]
25. Rayo, L.; Segal, I. Optimal information disclosure. *J. Political Econ.* **2010**, *118*, 949–987. [CrossRef]
26. Li, H.; Shi, X. Discriminatory Information Disclosure. *Discrim. Inf. Discl.* **2017**, *107*, 3363–3385. [CrossRef]
27. Clempner, J.B.; Poznyak, A.S. A Tikhonov regularized penalty function approach for solving polylinear programming problems. *J. Comput. Appl. Math.* **2018**, *328*, 267–286. [CrossRef]
28. Clempner, J.B.; Poznyak, A.S. A Tikhonov regularization parameter approach for solving Lagrange constrained optimization problems. *Eng. Optim.* **2018**, *50*, 1996–2012. [CrossRef]
29. Clempner, J.B.; Poznyak, A.S. Multiobjective Markov chains optimization problem with strong Pareto frontier: Principles of decision making. *Expert Syst. Appl.* **2017**, *68*, 123–135. [CrossRef]
30. Clempner, J.B.; Poznyak, A.S. Constructing the Pareto front for multi-objective Markov chains handling a strong Pareto policy approach. *Comput. Appl. Math.* **2018**, *3*, 567–591. [CrossRef]
31. Tanaka, K. The Closest Solution To The Shadow Minimum of a Cooperative Dynamic Game. *Comput. Math. Appl.* **1989**, *18*, 181–188. [CrossRef]
32. Tanaka, K.; Yokoyama, K. On ϵ-equilibrium point in a noncooperative n-person game. *J. Math. Anal. Appl.* **1991**, *160*, 413–423. [CrossRef]