

Game Manipulators - the Strategic Implications of Binding Contracts

María Alejandra Ramírez^{1,2}, Yoav Kolumbus³, Rosemarie Nagel⁴, David
Wolpert^{5,6,7,8}, and Jürgen Jost^{1,5,9}

¹Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

²Max Planck Institute for Evolutionary Biology, Plön, Germany

³Cornell University, USA

⁴ICREA, Department of Economics, Universitat Pompeu Fabra, Barcelona School of
Economics, Barcelona, Spain

⁵Santa Fe Institute for the Sciences of Complexity, Santa Fe, NM, USA

⁶Complexity Science Hub, Vienna, Austria

⁷Arizona State University, Tempe, AZ, USA

⁸International Center for Theoretical Physics, Trieste, Italy

⁹Center for Scalable Data Analytics and Artificial Intelligence, Leipzig University, Germany

Abstract

Commitment devices are powerful tools that can influence and incentivise certain behaviours by linking them to rewards or punishments. These devices are particularly useful in decision-making, as they can steer individuals towards specific choices. In the field of game theory, commitment devices can alter a player's payoff matrix, ultimately changing the game's Nash equilibria. Interestingly, agents, whom we term *game manipulators* and who can be external to the original game, can leverage such devices to extract fees from players by making them contingent offers that modify the payoffs of their actions. This can result in a different Nash equilibrium with potentially lower payoffs for the players compared to the original game. For this scheme to work, it is required that all commitments be binding, meaning that once an offer is made, it cannot be revoked. Consequently, we analyse binding contracts as the commitment mechanism that enables *game manipulation* scenarios. The main focus of this study is to formulate the logic of this setting, expand its scope to encompass more intricate schemes, and analyse the behaviour of regret-minimizing agents in scenarios involving game manipulation.

Keywords: Game manipulation - Commitment devices - Binding contracts - Regret minimization - Online learning - Blockchain contracts - Algorithmic game theory - Evolutionary game theory

1 Introduction

Commitment devices are tools that influence decision-making by linking actions to rewards or punishments. In the realm of game theory, the introduction of commitment devices leads to a distortion of payoff matrices, which in turn might result in the modification of the game’s equilibrium [Schelling (1956), Sobel (1981), Vickers (1985)]. For commitment devices to be effective, the offer made to modify a payoff matrix must be credible. That is, the commitment to deliver punishments or rewards is either strategically self enforced [Aumann, R.J. (1974)], or it is *binding* — meaning that once an offer is made, it cannot be revoked and it must be executed once the established conditions are met [Bono and Wolpert (2014)].

In this paper, we explore the logic of binding commitments in game theory. Thus, in the sequel, all offers made by the original players or additional actors, which we call *game manipulators*, are binding. In contrast to previous literature, discussed methodically in Section 11.2, we concentrate on analysing the strategic implications of such commitment schemes.

For this, we assume that offers and actions are common knowledge to all participants. From a classical game theory perspective, it is also assumed that all players are rational. This implies that each player is driven by their individual self-interest, leading them to seek to maximise their own payoffs. Consequently, players select actions that align with the Nash equilibrium in an attempt to unilaterally maximise their payoffs.¹

In such a framework, an actor, referred to as the *game manipulator*, could enter and obtain a fee. The game manipulator simply needs to present the following binding offer to the players: pay me a fee, or I will distort the payoff matrices by benefiting or punishing a certain action. In other words, the manipulator obtains a profit in exchange for not distorting the game structure, effectively exploiting the game.

In this study, we systematically analyse the game-theoretical logic of obtaining fees from players by offering binding contracts. This provides fundamental insight into the potential actions within the realm of blockchain, where intricate systems have been established to guarantee the enforceability of contracts. Moreover, in such blockchain environments binding contracts, like smart contracts, have become the backbone of numerous decentralised activities.

We analyse diverse *game manipulation* scenarios, including cases where one player makes an offer to its co-player in Section 5, cases where an external manipulator makes the offer in Section 6, scenarios where the co-player responds with a counteroffer in Section 7, and finally, in Section 8, scenarios where a second external manipulator makes an offer to the first one according to the same principle.

In each scenario, we complement the simple theoretical analysis with a learning theory approach [Fudenberg and Levine (1998), Hart and Mas-Colell (2000)] to investigate how agents learn to play games altered by binding contracts. Particularly, we

¹Our analysis focuses on games with a single equilibrium. There are, however, cases where multiple equilibria arise in the manipulated game, in which case rationality does not resolve the equilibrium selection problem. In these cases, we study the convergence of no-regret dynamics via simulations.

use the multiplicative weights update algorithm [Arora et al (2012)] and the replicator dynamics [Hofbauer and Sigmund (1998)] to explore how regret-minimizing agents [Blum and Mansour (2007)] behave under game manipulation. The purpose of this is to understand how agents, presumed to possess hindsight rationality [Morrill et al (2021)], act in such scenarios. This assumption is relevant to machine learning-driven agents [Blum (2005)] and can serve as an initial model for human behaviour.

Overall, we observe that depending on the game structure, the dynamics may or may not converge to the predicted subgame perfect equilibria (SPE). This affects the overall contract incentives to the extent that a failure to reach the SPE suggests that external manipulators would have been better off refraining from offering the contracts initially. Thus, to effectively propose and profit from binding contracts, it is necessary to carefully examine the resulting game dynamics along with its strategic implications.

Moreover, the results show that even though regret-minimizing agents effectively avoid dominated strategies, it is not guaranteed for them to converge to subgame perfect equilibria. Therefore, such refinement of the Nash equilibrium, is not necessarily in agreement with the evolutionary stable strategy of the replicator dynamics [Selten (1965), Smith (1982)]. This implies that backward induction fails once again to predict evolutionary stable strategies [Ramírez et al (2023)].

2 How to manipulate a game?

A *game manipulation* scenario requires the following conditions:

- Players have the potential to engage with one another through a game, called the original game or “base game.”
- An actor, known as the “game manipulator”, has the ability to request a fee for allowing such interaction to happen freely (without constraints on the players’ actions), if the fee is paid.
- Otherwise, the manipulator has the power to distort the payoff matrices by providing an incentive to a player for choosing a specific action.
- Once an offer is made by the game manipulator, it cannot be revoked and it should be followed through. To be more precise, the offers are binding commitments.

For simplicity, we study the case where a binding offer is proposed to only one player in a 2×2 asymmetric game. (For a general formulation, we refer the reader to Section 11.1). As a result, game manipulation gives rise to a sequential game with the following steps:

1. An offer, which is a binding commitment, is proposed to player i .
2. Player i must decide to Accept or Decline to pay the fee. If player i Declines, then the player’s payoff matrix is distorted as stated by the game manipulator.
3. Player i plays the chosen game (given Accept or Decline) with its co-player.

Essentially, player i is faced with the alternative to either pay a fee, or have the game modified in such a way that the Nash equilibrium is shifted.

If the game is distorted to player i 's disadvantage, then player i should pay the fee when the fee is smaller than that disadvantage. In such case, the manipulator is able to obtain “free money” by only posing the threat to distort the payoff matrix by paying the players additional rewards. More precisely, the subgame perfect equilibrium in a game manipulation setting encourages player i to Accept to pay the fee.

2.1 Offer of a binding contract

In this analysis, the offers are formulated according to the following format: *Either you pay a fee, or if you decline, I distort the game.*

For example, the offer could be stated as: *Either you give me K_1 (euros), or if you decline, I give you K_2 whenever you play X .*

In this formulation, X refers to an element of the offeree's action set. For example, in a 2×2 game, if the offer is proposed to the row player, X is either Top or Bottom. Likewise, if the offer is proposed to the column player, X is either Left or Right. It should be noted that the offeror can be either a player or an external actor.

2.1.1 Importance of binding commitments

It is important to note that the offers must be proposed as binding contracts, that is, there must be a mechanism in place to ensure that the offeror hands over K_2 when the request for K_1 is declined.

In traditional legal scenarios, the offeror could deposit K_2 with a notaris, who would return it to the offeror if the offeree's Accepts, but otherwise give it to offeree. However, in blockchain environments, such commitments can be made through smart contracts without the need of a third-party. As a result, game manipulation scenarios could proliferate more easily in blockchain platforms.

All the offers analysed in this study use the following “base game”.

3 The base game

Let us start with the following game

	L	R	
T	4,14	9,13	
B	5,6	10,10	(I)

The players are called Row (Row, he) and Column (Col, she). The unique Nash is (*Bottom, Right*) with the payoff of (10, 10), which is Pareto-efficient. Furthermore, Bottom strictly dominates Top for the row player. Right is an iterated dominant strategy for the column player. Thus, the game is a one-sided dominant game.

3.1 Motivation

For simplicity and tractability, we explore a 2×2 asymmetric game, where the binding offers are proposed to only one player. It should be noted that the game must be asymmetric such that the threat to distort the game is done by modifying only player i 's payoff matrix, as described in Section 2.1. This implies that the game manipulator has agency exclusively on player i and not its co-player. Possible modifications of this formulation are discussed in section Section 11.1 and Section 10.

In particular, we propose game (I) as a base game because it is an asymmetric 2×2 that allows a tractable analysis of the multiple game manipulation scenarios presented in further sections. Game (I) is also suitable to describe the buyer-supplier strategic interaction around demerit goods. This type of good is characterised by having negative effects on its consumers, i.e. demerit goods are considered to be unhealthy, demeaning or to have socially undesirable externalities, like pollution [Robbins (1932), Boadway and Bruce (1984)]. As a result, buyers, and society, would be better off if demerit goods are produced and consumed in low quantities, or not at all. To the contrary, suppliers/sellers are better off when they produce large, unrestricted, amounts of the demerit good in order to make sure they always meet the existent demand.

In particular, we assume that the buyers behave according to the supply and demand microeconomic model, such that the demand aims to match the quantity supplied. Precisely, economic equilibrium is reached once the demand coordinates with the supply. Following the interpretation of game (I), the supply is better off by producing large amounts of the good to coordinate with the demand at high production levels. On the other hand, the demand prefers to coordinate with the supply at low levels of production, due to the negative effects of the transacted good.

Consequently, game (I) represents the buyer-supplier interaction for demerit goods, where Row is the supply and Column the demand, as follows.

	Low consumption	High consumption
Low production	4,14	9,13
High production	5,6	10,10

(II)

The Nash equilibrium is (High production, high consumption). In particular, High production is a strictly dominant strategy for the supply. On the other hand, the demand needs to best reply with high consumption, although it would strictly prefer Low production yielding always higher payoffs.

Given that the Nash equilibrium is a socially undesirable equilibrium for the consumers, it is possible to model government intervention through *game manipulation*. Similarly, a game manipulation scenario can be used to model extortive behaviour by dominant firms controlling the market, over smaller firms that have the potential to engage with their existing costumers.

Government intervention: For a demerit good market, the government (external agent) proposes to the supplier (row player): *Either you pay me K_1 , or if you decline, I give you K_2 whenever you play Low production.* This translates into the negotiation of governments with lobbyist from firms selling demerit goods. The government proposes the supplier to either Accept to pay a tax (*Either you pay me K_1*), or to get a subsidy with the goal of reducing the production of the demerit good (*or if you decline, I give you K_2 whenever you play Low production.*).

For example, the government can propose to the car industry selling vehicles powered by fossil fuels “Either you pay me a tax, or if you decline I give you a subsidy to produce electric vehicles, as an incentive to decrease the production of petrol vehicles.” Therefore, the analysis of game manipulation can provide a better understanding on how to implement more efficient regulations on markets.

Extortion from a dominant firm: In a demerit good market, the dominant firm (external agent) proposes to a smaller firm (row player): *Either you pay me K_1 , or if you decline, I give you K_2 whenever you play Low production.* This translates into an extortive threat from a dominant firm against a smaller one. The dominant firm pressures the smaller one to Accept to pay an extortive fee (*Either you pay me K_1*), or the dominant firm will distort the market such that a Low production of the smaller firm is benefited (*or if you decline, I give you K_2 whenever you play Low production.*).

In other words, the only way for the smaller firm to prevent that the dominant firm distorts the base game is to pay the extortion fee. This example serves to illustrate the wide variety of real-life situations where game manipulation acting as an extortion mechanism could come into place.

Furthermore, in the growing economy over blockchain platforms, any platform participant has the power to create binding contracts and may potentially intervene in a strategic interaction between other participants if the incentives to do so exist. However, unlike government interventions, game manipulation coming from extortionists is a concerning activity that is hard to regulate, especially in blockchain platforms.

Consequently, to provide a better understanding of the strategic implications of binding contracts, in the upcoming sections, we analyse various ways in which they can be introduced in games.

We study the logic and dynamics of the following scenarios:

- Section 5: The column player makes an offer to the row player
- Section 6: The first external manipulator enters
- Section 7: Row player’s counter-offer
- Section 8: The second external manipulator enters

All offers are proposed taking the base game as the reference point. Regarding dynamics analysis, we employ the multiplicative weights and replicator dynamics to examine how regret-minimizing agents based on reinforcement learning behave in each scenario.

4 Methods

We employ online algorithms to explore how agents learn to determine the best action in a game. In particular, we simulate the scenario in which agents cannot directly identify the best action from historical data alone; they must actively engage in the game to learn which action is optimal. For this, agents repeatedly play the game, aiming to reduce their cumulative losses, or more precisely, to minimize their regret [Blum and Mansour (2007)]. In particular, we employ the most-studied no-regret algorithm, known as the multiplicative weights update algorithm (see [Arora et al (2012)] and references therein), and its continuum limit [Kleinberg et al (2009)], the replicator dynamic [Hofbauer and Sigmund (1998)].

4.1 Multiplicative weights update algorithm

The multiplicative weights algorithm (MWUA) is a decision making method based on an iterative rule that multiplicatively updates the weights assigned to each player’s possible actions [Littlestone and Warmuth (1994), Fudenberg and Levine (1995)]. Initially, all actions have equivalent weights. After each time step, each action’s weight is updated according to its payoff, thus the actions with higher payoffs have an increased probability of being chosen in the next round. More precisely, the players’ learning process is modelled as a multiplicative weight update, where the action’s distribution is skewed towards the successful actions providing a higher payoff [Grigoriadis and Khachiyan (1995), Arora et al (2012)]. As a result, the main objective of the multiplicative weights algorithm is to reduce its cumulative losses, or more specifically, to minimize the regret for agents playing a repeated game [Bailey and Piliouras (2018)]. As stated in [Arora et al (2012)], the formulation of the algorithm is the following:

Initialisation: fix $\eta \leq \frac{1}{2}$. Associate a weight $w_i^{(1)} := 1$ to each action i , for each player.

For $t = 1, 2, \dots, T$:

1. Choose action i with probability proportional to its weight $w_i^{(t)}$, i.e. use the distribution $\mathbf{p}^{(t)} = (\frac{w_1^{(t)}}{\Phi^{(t)}}, \dots, \frac{w_n^{(t)}}{\Phi^{(t)}})$, where $\Phi^{(t)} = \sum_i w_i^{(t)}$, to choose an action on each round.
2. Observe the payoffs of the actions $\mathbf{m}^{(t)}$.
3. Update the weights for every action as $w_i^{(t+1)} = w_i^{(t)}(1 + \eta m_i^{(t)})$.

4.2 Replicator dynamics

The replicator dynamic is a foundational model for dynamics in evolutionary game theory [Hofbauer and Sigmund (1998), Hofbauer et al (2009), Sandholm (2010)]. In particular, it can be used as a model to analyse how players learn to play an asymmetric game. In this dynamic, at each time step, the row player chooses one of the n possible actions with frequency $\mathbf{x} = (x_1, \dots, x_n)$. Likewise, Column choose one of the m possible actions with frequency $\mathbf{y} = (y_1, \dots, y_m)$. In our setting, Row has the same amount of

possible actions as Column, $n = m$ (see base game (I)). Over time successful actions are positively reinforced, which allows for players to eventually converge to an optimum equilibrium. Thus, the replicator dynamic is represented by the following set of coupled differential equations.

$$\dot{x}_i = x_i[(Ay)_i - x^T Ay] \quad i = 1, \dots, n \quad (1)$$

$$\dot{y}_j = y_j[(Bx)_j - y^T Bx] \quad j = 1, \dots, m. \quad (2)$$

In this equation, A and B are the payoff matrices of the row and column player, respectively. It should be noted that to obtain B , it is required to take the transpose of the game for the column entries, such that the definition of symmetric games holds: $A = B^T$. For equation 1, $(Ay)_i$ is Row's average performance of strategy i over Column's possible actions. On the other hand, $x^T Ay$ is the averaged performance over all of Row's strategies. Thus, the replicator equation describes how the frequency of strategy i changes over time according to its average performance relative to the mean. Equation 2 has an equivalent interpretation for the column player [Börgers and Sarin (1997), Hofbauer and Sigmund (1998), Sato et al (2002)].

Dynamical systems such as the replicator dynamic arise as the continuum limit of discrete no-regret algorithms. More precisely, it has been shown that a multiplicative update process can be approximated to a continuous-time process displaying replicator dynamics [Kleinberg et al (2009), Kleinberg et al (2011), Bailey and Piliouras (2018)].

5 Column makes an offer to Row

To introduce the underlying logic of the binding commitments, we first allow for Col to make the following offer to Row: *Either you give me 3 (Euros), or if you decline, I give you 2 whenever you play T.*

If the row player Accepts, the payoff is as follows.

	L	R	
T	1,17	6,16	
B	2,9	7,13	(III)

The unique Nash equilibrium in this new game does not change compared to the base game (I), as it remains at (B, R) , given that Bottom continues to be a strictly dominant strategy over Top. The main difference with respect to the base game is that now Col is able to extract a profit from Row, which in this case is equivalent to 3.

If the row player Declines, the payoffs are

	L	R	
T	6,12	11,11	
B	5,6	10,10	(IV)

Now the unique Nash in this new game is (T, L) , resulting in $(6, 12)$. The Nash equilibrium changes because Top now becomes the strictly dominant strategy for the row player.

In particular, if the offer is Declined, Row's payoff matrix becomes that of a temptation game, that is, Row is tempted to play T and obtain the fee offered by Col, but the resulting equilibrium leaves him worse off than if he had accepted the offer. Therefore, it is better for Row to Accept the offer and allow for Col to extract the profit. Likewise Col is better off if the Accept Nash equilibrium is reached. Most importantly, the column player is better off in either scenario, Accept or Decline, compared to the base game, which is the main motivation for Col to make the offer to Row.

Overall, for the full game, the subgame perfect equilibrium is the unique Nash equilibrium of the subgame (III), corresponding to (B, R) once Row Accepts given the NE in the subgame (IV) with (T, L) when Declining the offer.

5.1 Formulation conditions

The binding offer to the row player can be generalised as: *Either you give me c_1 (Euros), or if you decline, I give you c_2 whenever you play T .* Therefore, an abstract representation of the offer can be formulated using the payoff matrix of the base game (I)

	L	R	
T	a_{11}, b_{11}	a_{12}, b_{12}	
B	a_{21}, b_{21}	a_{22}, b_{22}	(3)

the Accept fee c_1 and the Decline fee c_2 . Particularly, we find that the crucial conditions are

$$a_{12} + c_2 > a_{22} > a_{12} \tag{4}$$

$$a_{22} - c_1 > a_{11} + c_2 \tag{5}$$

$$b_{11} > b_{12} \tag{6}$$

$$a_{11} + c_2 \geq a_{21}. \tag{7}$$

5.2 Reduced-Normal form game

The sequential game composed of (III) and (IV) can be written in its reduced-normal form, where Row has four possible actions (Accept and Top, Accept and Bottom, Decline and Top, Decline and Bottom). Similarly, Column has four possible actions but that depend on Row's initial decision between (Accept, Decline). For example, the action L, R means that Column chooses Left if Row Accepts and Right if Row Declines.

	L,L	L,R	R,L	R,R
AT	1,17	1,17	6,16	6,16
AB	2,9	2,9	7,13	7,13
DT	6,12	11,11	6,12	11,11
DB	5,6	10,10	5,6	10,10

(V)

The pure Nash equilibria are (DT, L, L) and (AB, R, L) , which coincide with the equilibria of the individual games. There is also a mixed Nash equilibrium which we do not consider in our study. For this representation, it can also be noted that AT and DB are dominated strategies for Row.

5.3 Results

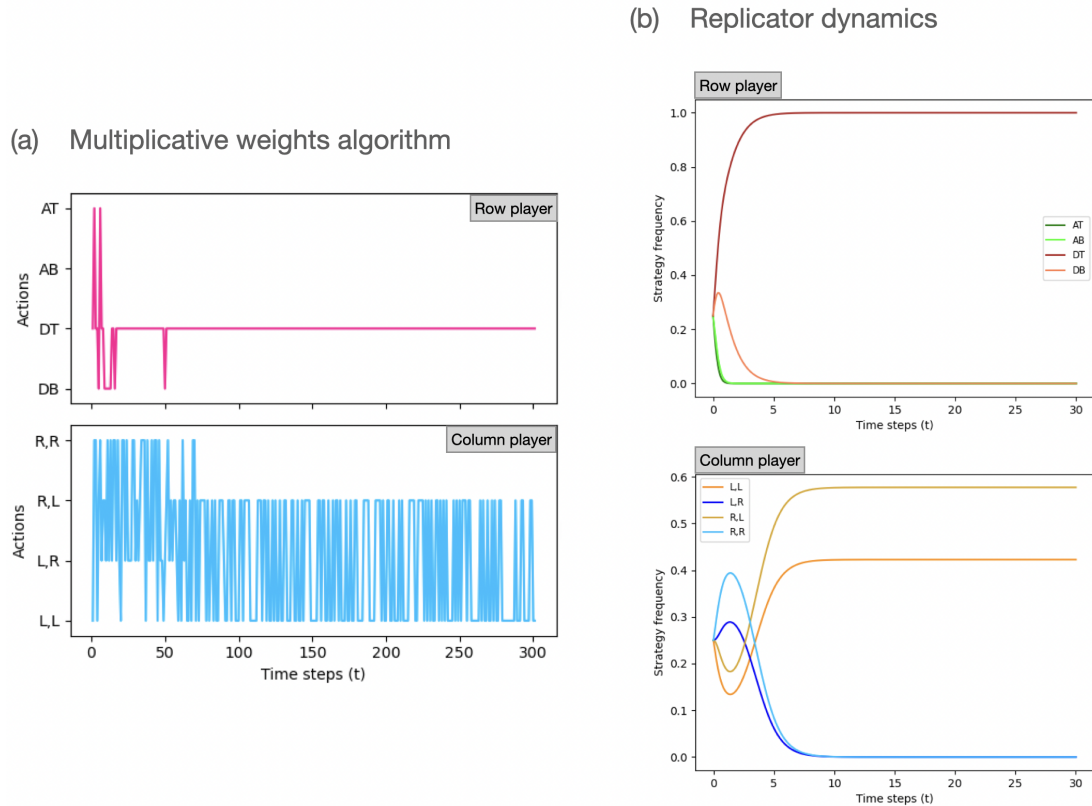


Figure 1: (a) A simulation of multiplicative weights dynamics ($\eta = 0.5$ for both players) and (b) numerical solution to the replicator dynamics equations for Section 5, where Column makes an offer to Row. For both methods, the row player converges to DT , which means to Decline the offer and then playing Top. The column player converges to L, L and R, L , corresponding to Col playing Left once Row has declined the offer. Thus, the system converges to the Nash equilibrium of the Decline game (IV). Interestingly, this equilibrium does not coincide with the subgame perfect equilibrium of the full game.

Figure 1 shows the results for the scenario where Column makes an offer to Row, as considered in the normal form game (VI). It can be seen that the multiplicative weights algorithm converges to the same equilibrium as the replicator dynamics, namely (T, L) , the Nash equilibrium of the Decline game. It should be noted that the MWUA plot shows the actions chosen by each player for every time step, while the replicator dynamics displays how the action frequency evolve in time. In general, both methods identify DT as the best possible action for Row taking into account Column’s available actions. It is surprising that the replicator dynamics and the MWUA equilibrium are consistently not in agreement with the subgame perfect equilibrium of the full game.

Interestingly, this suggests that Row prefers to Decline Column’s offer, such that he is able to receive the $c_2 = 2$ euro offered to him. In this way Row exploits Column back by extracting the decline fee (c_2) from her. The reason for such behaviour is that DT performs better on average with respect to all possible strategies of the column player. Unfortunately, if the row player Declines the drawback is that the game structure leaves him worse off compared to the Accept equilibrium.

6 The first external manipulator enters

The observation that players can change a game to their advantage by making binding commitments dates back to [Schelling (1956)]. Thus, we now introduce a novel scenario where an external “game manipulator” (M_1), makes an offer, in the form of a binding contract, to extort a fee from players. More precisely, the game manipulator makes the offer to Row: *Either you give me 3 (Euros), or if you decline, I give you 2 whenever you play T.*

It should be noted that the main differences with respect to the first scenario is that an external game manipulator is making the offer, instead of a player. Therefore, payoffs injected or retrieved by the manipulator are not transfers between the two players anymore. The equilibrium strategies for Row and Column remain, however, the same.

If the row player Accepts, the payoff matrix is as follows

	L	R
T	1,14,{3}	6,13,{3}
B	2,6,{3}	7,10,{3}

(VI)

The third entry in a cell specifies the payoffs to the external manipulator M_1 . It should be noted that after the offer is proposed, it cannot be revoked. Thus, the payoff of M_1 is solely an outcome of the players’ (Row and Col) decisions, such that the Nash equilibrium in this subgame is (B, R) with payoffs $(7, 10, \{3\})$. For Row, Bottom is a strictly dominant strategy while Column has an iterated dominant strategy. If Row Accepts, the game manipulator extracts a profit of $c_1 = 3$.

If the row player Declines, the payoffs are

	L	R	
T	6,14,{-2}	11,13,{-2}	(VII)
B	5,6,{0}	10,10,{0}	

The Nash equilibrium in this subgame is (T, L) with payoffs $(6, 14, \{-2\})$. In this game, Top is a strictly dominant strategy. Left is an iterated dominant strategy. Therefore, if Row Declines, the game manipulator pays a fee of $c_2 = 2$ in this subgame.

The external game manipulator does not change the strategic situation for Row compared to the previous case in Section 5, where Col made the offer. Again, it is better for Row to Accept the offer, and pay $c_1 = 3$, to reach the equilibrium in the Accept-subgame, $(B, R) = (7, 10, \{3\})$, instead of the equilibrium in the Decline-subgame $(DT, L) = (6, 14, \{-2\})$. In other words, the subgame perfect equilibrium outcome corresponds to the Nash equilibrium outcome of game (VI), namely (B, R) once Row Accepts.

It should be noted that for the logic of the external manipulator scenario, it is not necessary for the fee the manipulator demands ($c_1 = 3$) to be greater than the fee he threatens to pay ($c_2 = 2$). As long as it is better for Row to Accept the offer, any positive fee c_2 would be sufficient. c_2 just cannot be too large, according to (VI), or else the subgame perfect equilibrium would be modified and Row would be better off Declining.

6.1 Reduced-Normal form game

The sequential game described by the games (VI) and (VII) can be described in its reduced-normal form as

	L,L	L,R	R,L	R,R	
AT	1,14,{3}	1,14,{3}	6,13,{3}	6,13,{3}	(VIII)
AB	2,6,{3}	2,6,{3}	7,10,{3}	7,10,{3}	
DT	6,14,{-2}	11,13,{-2}	6,14,{-2}	11,13,{-2}	
DB	5,6,{0}	10,10,{0}	5,6,{0}	10,10,{0}	

The pure Nash equilibria are (DT, L, L) and (AB, R, L) , which coincide with the equilibria of the individual subgames. There is also a mixed Nash equilibrium which we do not consider here. It can be noted that AT and DB are dominated strategies for Row.

6.2 Results

Figure 2 shows that the system converges to Row Declining the offer and then playing Top, and to Column playing Left. More precisely, the system converges to the Nash equilibrium of game (VII). The results are equivalent to the scenario where Column makes an offer to Row. The reason for this is that the general payoff structure is not changed when an external manipulator makes the offer instead of the column

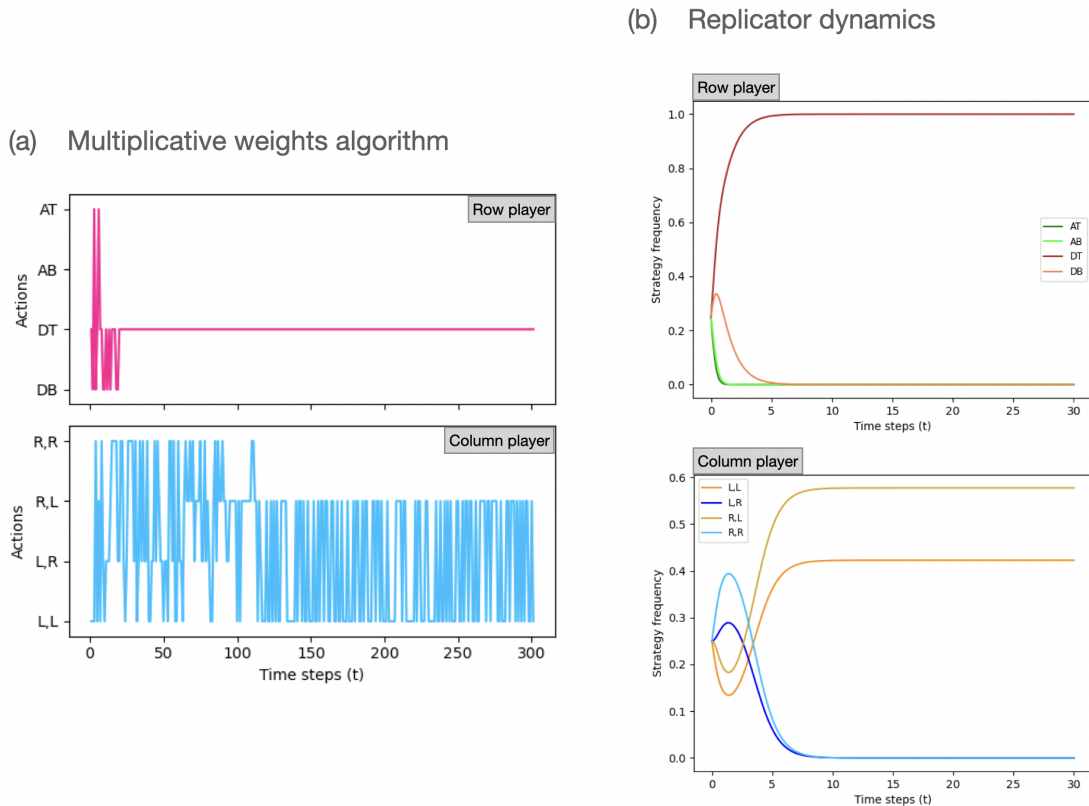


Figure 2: **(a)** A simulation of multiplicative weights dynamics ($\eta = 0.5$ for both players) and **(b)** numerical solution to the replicator dynamics equations for Section 6, where the first external manipulator enters. For both methods, the row player converges to DT , which means to Decline the offer and then playing Top. The column player converges to L, L and R, L , corresponding to Col playing Left once Row has declined the offer. Thus, the system converges to the Nash equilibrium of the Decline game (VII). Interestingly, this equilibrium does not coincide with the subgame perfect equilibrium of the full game.

player. Therefore, once again we reach an equilibrium that does not coincide with the subgame perfect equilibrium of the full game. In this case, the results mean that Row takes advantage of the situation to extract a fee c_2 from the manipulator at the cost of having a lower payoff than the base game, and the Accept equilibrium. This decision indirectly benefits Col who obtains a better payoff for the Decline equilibrium in comparison to the base game and the Accept equilibrium.

It is interesting to note that the results in figure 2 do not correspond to the subgame perfect equilibrium of the full game. The reason for this is the fact that on average DT performs better against all of Column's possible actions. Therefore, the results indicate that despite the theoretical prediction of players reaching the subgame perfect equilibrium, which would enable the external manipulator to earn a profit of c_1 , the manipulator would actually find it more advantageous to abstain from making an offer. This is because players tend to converge towards the Decline equilibrium, which incurs a cost of c_2 for the manipulator.

7 Row's counter-offer

In the scenario of the first external manipulator, Row could propose a counter-offer to Col: *I shall decline the external manipulator's offer and then give you $d_2 = 2$ for you playing Right*. Therefore, if the row player Accepts, the payoff matrix is the same as game (VI), namely

	L	R
T	1,14,{3}	6,13,{3}
B	2,6,{3}	7,10,{3}

(IX)

The unique Nash equilibrium in this subgame is again $(B, R) = (7, 10, \{3\})$, given that Bottom is a strictly dominant strategy and Right is iterated dominant. In this case, the external game manipulator extracts a profit of $c_1 = 3$.

If the row player Declines, Row's counter-offer changes the payoff matrix as follows

	L	R
T	6,14,{-2}	9,15,{-2}
B	5,6,{0}	8,12,{0}

(X)

Therefore, in subgame (XII) $(T, R) = (9, 15, \{-2\})$ is the new Nash equilibrium, where Top and Right are strictly dominant strategies. In this equilibrium, Row and Col obtain the best possible outcome. Even though Row is worse off than in the base game (I), it is better off than in the Accept game and the previous Decline equilibrium. For Column, it is beneficial to accept the offer because she is also better off than in those two cases. In other words, Row has essentially transferred the fee obtained from the external manipulator to Col. But if that is possible, the manipulator should not attempt to pose a threat in the first place, as he risks losing $c_2 = 2$ by his attempt. Of course, this also depends on certain additional inequalities, with d_2 now denoting the payoff increase offered by Row to Col for playing Right.

$$b_{12} < b_{11} < b_{12} + d_2 \quad (8)$$

$$a_{11} + c_2 < a_{12} - d_2 \quad (9)$$

Thus, Row's counter-offer can prevent the game manipulator from extracting a profit.

7.1 Reduced-Normal form game

The sequential game described by the games (IX) and (X) can be represented in its reduced-normal form as

	L,L	L,R	R,L	R,R
AT	1,14,{3}	1,14,{3}	6,13,{3}	6,13,{3}
AB	2,6,{3}	2,6,{3}	7,10,{3}	7,10,{3}
DT	6,14,{-2}	9,15,{-2}	6,14,{-2}	9,15,{-2}
DB	5,6,{0}	8,12,{0}	5,6,{0}	8,12,{0}

(XI)

The pure Nash equilibria are (DT, L, R) , (DT, R, R) and (AB, R, L) , which coincide with the equilibria of the individual subgames games. It can be noted that AT and DB are dominated strategies for Row.

7.2 Results

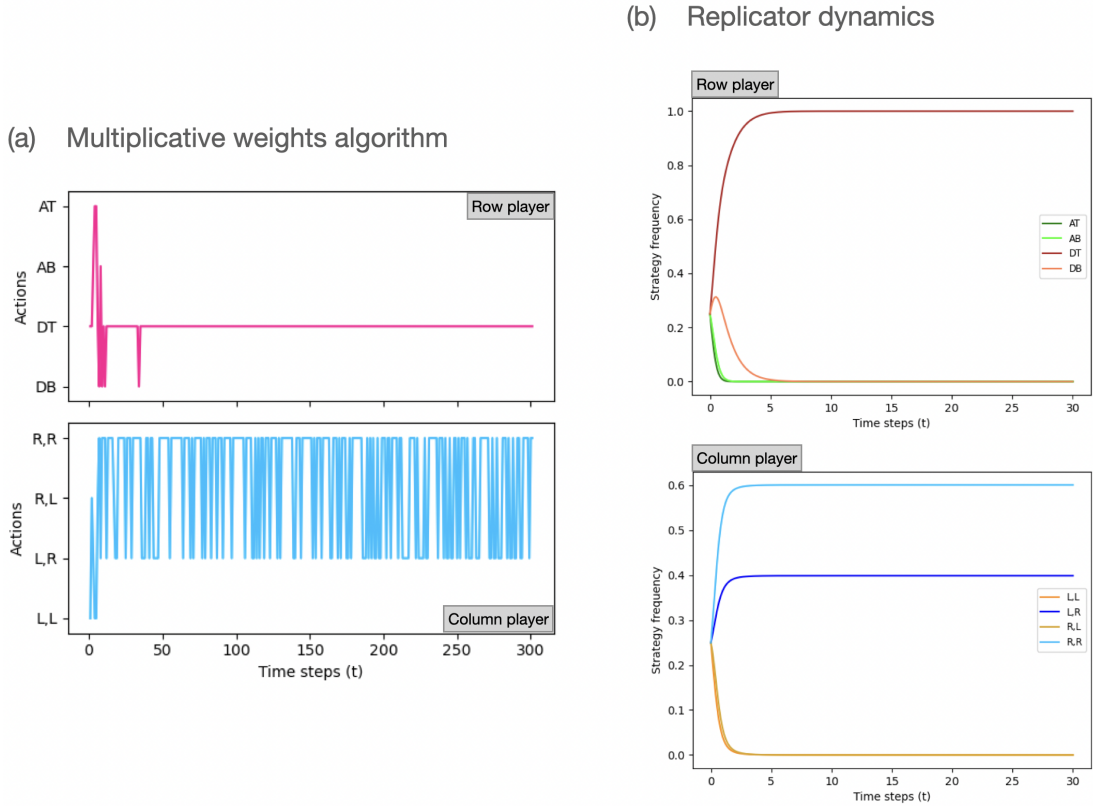


Figure 3: **(a)** A simulation of multiplicative weights dynamics ($\eta = 0.5$ for both players) and **(b)** numerical solution to the replicator dynamics equations for Section 7, which describes Row’s counter-offer. For both methods, the row player converges to DT , which means to Decline the offer and then playing Top. The column player converges to R, R and L, R , corresponding to Col playing Right once Row has declined the offer. Thus, the system converges to the Nash equilibrium of the Decline game (X), which is in agreement with the subgame perfect equilibrium. Such equilibrium allows for both Row and Column to obtain the best possible payoff.

As shown in figure 3, for the multiplicative weights algorithm and the replicator dynamics, the system converges to Row Declining and playing Top, and then for Column to accept Row’s counter-offer to play Right. Therefore, the results are in agreement with the subgame perfect equilibrium. This means that Row’s counter-offer effectively neutralises the external game manipulator to the extent that the manipulator would be better off by not posing a threat on the first place.

8 The second external manipulator enters

In this section, we introduce a second external manipulator into Section 6, where the first external manipulator appears. Thus, the first manipulator's offer is the same as in section Section 6. The second manipulator M_2 (she) offers to the first manipulator M_1 (he): *Either you pay me 2, or else if Row Declines your offer, I will offer Col an additional payoff of 2 for playing Right.*

The chronology of this scenario is as follows:

1. Time t_1 : Row and Column agree to play the original game, at time t_6 .
2. Time t_2 : M_1 proposes the offer described in Section 6 to Row, specifying that they must make their decision precisely at time t_5 .
3. Time t_3 : M_2 proposes the offer described in this section to M_1 , specifying that they must make their decision precisely at time t_4 .
4. Time t_4 : M_1 decides upon the offer made by M_2 .
5. Time t_5 : Row decides upon the offer made by M_1 .
6. Time t_6 : Row and Column play the game determined in previous steps.

We focus on analysing the strategic implications of the offer from M_2 to M_1 (steps t_3 and t_4).

If M_1 accepts M_2 's offer, it implies that the Accept game is equivalent to game (VI), but M_1 's payoff is decreased by 2 amounting now to 1. Additionally, M_2 's payoff of 2 is added to the matrix, such that each entry follows the structure (Row, Col, $\{M_1, M_2\}$).

	L	R
T	1,14,{1,2}	6,13,{1,2}
B	2,6,{1,2}	7,10,{1,2}

(XII)

The Nash equilibrium in this subgame is (B, R) with payoffs $(7, 10, \{1, 2\})$. For Row, Bottom is a strictly dominant strategy and for Col, Right is iterated dominant. The main difference in comparison to Section 6 is that now M_1 only extracts a profit equivalent to 1, instead of 3, and M_2 appears with a payoff of 2.

If M_1 declines M_2 's offer and Row Declines M_1 's offer, the new Decline subgame is as follows:

	L	R
T	6,14,{-2,-2}	11,15,{-2,-2}
B	5,6,{0,0}	10,12,{0,0}

(XIII)

This leads to the Nash equilibrium in this subgame $(T, R) = (11, 15, \{-2, -2\})$ which is the best that Row and Col can hope for. In particular, Top and Right are strictly dominant strategies.

In general, the logic of the scenario is the following: M_1 makes an offer to Row. Either Row accepts M_1 's offer, or the Nash equilibrium is shifted to Row's disadvantage. It is thus better for Row to accept M_1 's offer. But before Row decides, M_2 enters and makes the offer to M_1 . Either M_1 accepts M_2 's offer, or the game is changed in such a way that it now becomes advantageous for Row to decline M_1 's offer. Therefore, it is best for M_1 to accept M_2 's offer, and when Row then decides about M_1 's offer, it is also better to accept.

In fact, the subgame perfect equilibrium states that Row would be better off by Declining M_1 's offer, in case M_1 declines M_2 's offer. Therefore, the threat posed by M_2 to distort the game, by offering an additional payoff to the column player, guarantees that M_1 's offer is Declined by Row. As a result, M_1 is forced to pay M_2 a fee to retain a profit from initially manipulating the Row player.

Interestingly, the results in Section 6 show that regret-minimizing players do not converge initially to Accept M_1 's offer, i.e. M_1 's profit is not guaranteed from the start. As a result, the current situation presents an interesting scenario, where a static analysis of the game contradicts the dynamical outcome obtained from online learning of regret-minimizing agents.

8.1 Reduced-Normal form game

The sequential game described by the games (XII) and (XIII) can be represented in its reduced-normal form as

	L,L	L,R	R,L	R,R	
AT	1,14,{1,2}	1,14,{1,2}	6,13,{1,2}	6,13,{1,2}	
AB	2,6,{1,2}	2,6,{1,2}	7,10,{1,2}	7,10,{1,2}	
DT	6,14,{-2,-2}	11,15,{-2,-2}	6,14,{-2,-2}	11,15,{-2,-2}	
DB	5,6,{0,0}	10,12,{0,0}	5,6,{0,0}	10,12,{0,0}	(XIV)

Note that we do not specify the actions by M_1 and M_2 explicitly. In particular, the pure Nash equilibria are (DT, L, R) , (DT, R, R) and (AB, R, L) , which coincide with the equilibria of the individual subgames. It can be noted that AT and DB are dominated strategies for Row.

8.2 Results

In figure 4, it can be seen that the system converges to Row Declining the offer and playing Top, while Column plays Right. This equilibrium coincides with the Nash equilibrium of game (XIII). Therefore, if M_2 's offer is common knowledge, the players converge to the subgame perfect equilibrium, where both Row and Col reach the best

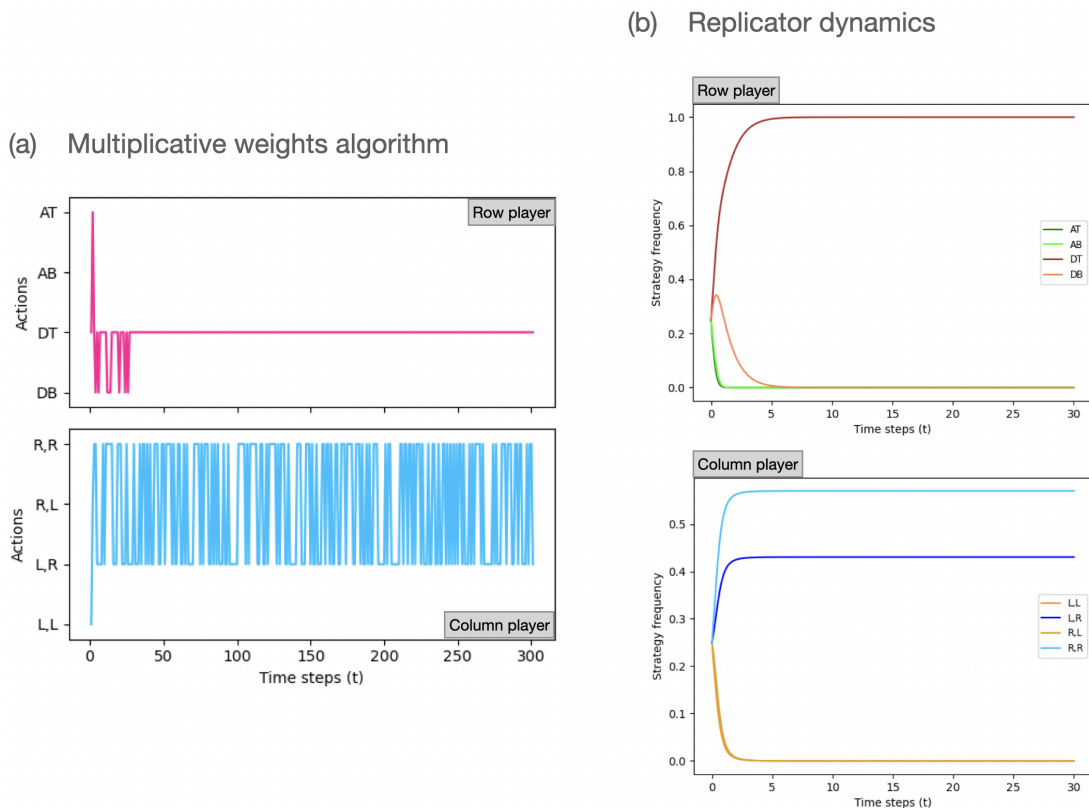


Figure 4: **(a)** A simulation of multiplicative weights dynamics ($\eta = 0.5$ for both players) and **(b)** numerical solution to the replicator dynamics equations for Section 8, where the second external manipulator enters. For both methods, the row player converges to DT , which means to Decline the offer and then playing Top. The column player converges to R, R and L, R , corresponding to Col playing Right once Row has declined the offer. Thus, the system converges to the Nash equilibrium of the Decline game (XIII). If M_2 offer is common knowledge, this equilibrium is equivalent to the subgame perfect equilibrium. This means that M_1 is effectively forced to pay a fee to M_2 to be able to retain a profit from initially manipulating Row. However, this is a paradoxical result because M_1 's profit is not guarantee from the start when regret-minimizing agents are playing the game.

possible outcome. This means that M_1 is effectively forced to pay a fee to M_2 to be able to retain a profit from initially manipulating Row.

The current situation presents an interesting paradox when compared to Section 6, because the results indicate that regret-minimizing players do not converge initially to Accept M_1 's offer. This means that M_1 's profit is not guarantee from the start. As a result, we find ourselves in an intriguing scenario where a static analysis of the game contradicts the dynamical outcome obtained from regret-minimizing agents.

9 Conclusions

In this paper, we explore commitment devices in game theory. More precisely, we study the strategic implications of binding contracts using a tractable 2×2 asymmetric game. Given that game manipulation through binding contracts can take different forms, from government regulations to extortion in blockchain platforms, we systematically analyse the logic of *game manipulation* via diverse setups.

Particularly, we investigate how regret-minimizing agents interact under the influence of game manipulators. For this, we use the multiplicative weights update algorithm (MWUA) and the replicator dynamics. We found that both methods provide consistent results for all analysed scenarios. In particular, the results show that the regret-minimizing agents avoid convergence to dominated strategies, as could be expected.

Furthermore, we observe that depending on how the binding offers distort the base game, regret-minimizing agents might not converge to the subgame perfect equilibrium. For instance, in both Section 5, where Column makes an offer to Row, and Section 6, where the first external manipulator enters, players converge to the Decline equilibrium, (T, L) , while the subgame perfect equilibrium corresponds to the Accept equilibrium, namely (B, R) once Row Accepts. This indicates that regret-minimizing agents do not necessarily behave as backward induction predicts.

In addition, the results reveal that despite the distinction between Section 5, where Column proposes the offer Row, and section Section 6, where an external manipulator (M_1) proposes the offer to Row, both scenarios are dynamically equivalent for MWUA and the replicator dynamics. For both methods, the row player prefers to Decline the offer, such that he is able to extract the fee c_2 offered to him. The reason for such behaviour is that DT performs better on average with respect to all possible strategies of the column player. Unfortunately, the drawback is that if Row Declines, the game structure leaves him worse off compared to the Accept equilibrium.

This suggests a contradicting situation for M_1 , because classical game theory suggests that players would achieve the subgame perfect equilibrium and allow him to extract a profit equivalent to c_1 . However, the equilibrium achieved by regret-minimizing agents indicate that refraining from offering a contract would actually be more beneficial for M_1 . Thus, when offering binding contracts to regret-minimizing agents, or agents only rational at hindsight, it is essential to carefully examine the dynamic implications of the contracts, in order to effectively leverage them. Relying solely on an analysis rooted in classical game theory may not be adequate for this purpose.

On the other hand, in both Section 7, describing Row’s counter-offer, and Section 8, where the second external manipulator enters, MWUA and the replicator dynamics converge to the subgame perfect equilibrium. The reason for this is that both offers alter the game structure in a similar manner, leading to two pure Nash equilibria, in the normal form game, corresponding to Declining the offer. Therefore, despite their differences, both offers have an equivalent impact on the game structure and its dynamical outcome.

In the case of Section 7, the results indicate that Row’s counter-offer to Column effec-

tively neutralises the external game manipulator M_1 intention to profit from posing a threat to distort the game. Particularly, the game manipulator M_1 would be better off by not proposing a binding contract on the first place under Row's counter-offer.

Regarding Section 8, the second external manipulator M_2 proposes the binding contract to M_1 with the intention of extracting a fee from him. To be more precise, when M_1 Declines the offer of M_2 , Row would be better off by also Declining the offer of M_1 . This would inflict a loss of c_2 to M_1 , whereas if M_1 accepts the offer of M_2 , he would retain a profit, because in that case, Row should Accept M_1 's offer according to classical game theory. Moreover, the results obtained from MWUA and the replicator dynamics confirm that M_2 's offer is a credible threat for M_1 .

Interestingly, Section 8, where the second external manipulator enters, poses a paradox in the light of the results of Section 6, where the first external manipulator enters. The reason is that regret-minimizing agents converge to decline M_1 's offer, thus M_1 's profit obtained from initially manipulating Row is not guaranteed from the start. Therefore, M_2 's intervention becomes unreasonable under such consideration.

In conclusion, a static analysis based solely on the concept of Nash equilibrium, an its refinement the subgame perfect equilibrium, is not sufficient to understand the behaviour of regret-minimizing agents in the framework of game manipulation through binding contracts. A dynamic exploration is therefore crucial to effectively leverage binding contracts as commitment devices in game theory.

10 Extensions and open questions

Possible extensions of the presented framework include bargaining over the details of the contract. In that bargaining, the manipulator attempts to maximise the "free money" that a player could give it. This topic is related to mechanism design applied to the offer of the binding contracts. In the paper by [Bono and Wolpert (2014)], the authors derive on the maximal amount a game manipulator can obtain for certain game manipulation scenarios, as one varies over possible payoff matrices of the base game and/or over possible contracts offered by the game manipulator. The analysis in that paper also shows that the results of game manipulation on the resultant joint payoffs of the players in the base game may differ in general from the Coaseian outcome.

In all of these cases, the presence of the game manipulator results in an entirely different strategic environment faced by the players in the base game. Therefore, the manipulator makes considerable profit from recognising a situation in which one party benefits from an output-contingent contract. But if the game manipulator manages to make a profit, the possibility that other game manipulators appear, seeking to pose their own contracts at lower prices, arises. This opens additional strategic implications for the player whose payoff matrix is distorted by such offers.

Another option is for the external game manipulator to offer contracts to both players. As an example, it may be that the manipulator offers contracts to both players with the following properties. First, both players have a strictly dominant strategy to accept the offered contract. Second, when they both inevitably accept, the outcome is that

they are both worse off, and the game manipulator profits considerably. In fact, the game manipulator may even promise to pay the players large sums of money if certain outcomes are reached, knowing well that when all of the players accept the offered contracts, such outcomes will never be obtained. (To do this, the game manipulator effectively creates a prisoner’s dilemma among the players for his own benefit.)

It is also worth noting that in many game manipulation scenarios, “mixed contracts” are needed to guarantee the existence of equilibria. With such contracts, players have uncertainty about their opponent’s payoffs. However, unlike in Bayesian Nash equilibrium, for game manipulation scenarios this uncertainty is resolved before the game is played. Thus, in this setting, the game manipulator’s only role is to make the players indifferent among its own contracts such that players randomise among them according to a probability distribution.

Furthermore, a rich mathematical structure is introduced if the players in the base game are able to offer contracts to the external game manipulator, who can pick only one of them, in which those players “purchase” a move-contingent payoff after the game is played.

Finally, there are important timing issues in game manipulation scenarios, some of which are explored in [Bono and Wolpert (2014)]. When players sequentially sign contracts with game manipulators, there can be a significant first-mover advantage to the first-signing player. This provides the game manipulator with yet another opportunity for profit i.e. they can charge the players to move first. In addition, game manipulation has complicated effects on efficiency. The effects depend on the underlying game as well as the bargaining structure between players and the game manipulator. In general, those effects can either increase equilibrium efficiency or decrease it.

Acknowledgments

David Wolpert thanks the Santa Fe Institute for support. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme (grant agreement no. 740282). Rosemarie Nagel acknowledges the Spanish Ministry of Education grant PRESP04022 - PID2021-125538NB-I00/AEI/10.13039/501100011033 and the support by the Barcelona School of Economics.

11 Appendix

11.1 Generalised game manipulation

For the general formulation, originally proposed in [Bono and Wolpert (2014)] (there termed “game mining”), we first introduce the corresponding notation. Let $T_1 \in \{Row, Col\}$ be the unit for information, or bit, specifying which of the two players, Row or Col, the game manipulator asks for a one-time payment in a single-stage game, before the start of the base game. Let $T_2 \in \{Row, Col\}$ be a bit specifying which of the two players the manipulator will make an output-contingent payment to, if the request

for a one-time payment is refused. Let $W \in \{win, lose\}$ be the bit saying whether the player specified in T_1 has better or worse expected payoff due to the intervention of the game manipulator, compared to what their expected payoff in the base game would be if the manipulator had never intervened. So for example, in the standard game manipulation scenario considered in Section 6, $T_1 = T_2 = Row$ and $W = lose$; the player being exploited by the manipulator is worse off due to the intervention of the manipulator.

Nonetheless, as pointed out by [Bono and Wolpert (2014)], there are base games in which a game manipulator could benefit by intervening for either or both possibilities, $T_1 \neq T_2$ and $T_1 = T_2$. Moreover, there are games where $W = win$ and games in which $W = lose$, i.e., the player in the base game who is being exploited by the manipulator can be made either worse off or better off. Thus, the particular case presented in Section 6 is just one possible configuration in the myriad of all available game manipulation opportunities.

11.2 Game manipulation compared to previous literature

The ideas underlying the game manipulation concept are implicit in a large body of economic literature. As an illustration, in the model presented in [Jackson and Wilkie (2005)] (JW), every player specifies outcome-contingent side-payments that they will make after a non-cooperative strategic form game is played and the payoffs are resolved. These side-payments are binding contracts, so the players are *ex ante* determining their preferences over the game's outcomes. In this regard the game that the players actually play is endogenously determined. JW examine whether a mechanism that allows players to make such outcome-contingent side-payments generally results in efficient outcomes and conclude that it does not. A related idea has been recently explored in the context of mechanism design when considering auctions with aftermarkets [Babaioff et. al. (2023)].

The simplest game-manipulation scenario can be viewed as a special case of JW. In this special case, the only outcome-contingent side-payments are between the players and the game manipulators. The game manipulators would be indifferent over the game outcomes if not for the fact that they will be receiving side-payments dependent on those outcomes. Furthermore, the game manipulators play no part in the game between the players other than to accept contracts for outcome-contingent side-payments and make the contracts public.

In contrast to JW, we do not focus on efficiency issues, and we do not assume that a social planner installs a mechanism for players to make side-payments. Instead, we look at a game without formal mechanisms and ask whether external parties will create contracts for outcome-contingent side-payments in pursuit of profit. In particular, we examine the implications of giving the game manipulator the power to offer contracts, which will in general increase the game manipulator's profits. In addition, we relax the assumption in JW that all side-payments are non-negative. That is, we allow game manipulators to pay players for certain outcomes. This will be important when examining optimal contracts as well as the extent to which a monopolist game manipulator

can extract profits from players. We also consider how things change when there is more than one manipulator, when manipulation contracts are offered in sequence, etc. None of those issues arise in JW.

In another related paper, [Renou (2009)] analyses what happens when players are able to embed the original game in a new two-stage game. In the second stage of that new game the players play the original game. However before they do so, in the first stage, the players each simultaneously commit not to play some subset of their possible moves in that game in the second stage. These *commitment games* can be seen as another special case of JW in which (1) player i 's side-payments are only contingent on i 's action (rather than on the full profile of actions), (2) the side-payments are made to external players and (3) the side-payments are effectively infinite. Renou's analysis does not apply to the full game manipulation scenario. This is because there are many circumstances in which both the player and the game manipulator prefer to make contracts that are fully outcome-contingent and that have non-infinite side-payments. One particular example is when the player and game manipulator find it optimal to agree on a contract that results in an equilibrium where the player uses a mixed strategy with full support (and therefore does not make any commitment in the first stage of Renou's two-stage game).

The idea that there might be pre-game play in which players make choices to affect their own preferences over outcomes is also present in [Wolpert et al (2008)]. The authors analyse the idea that experimentally observed non-rationality is in fact rational, because by committing to play the game with a non-rational "persona," a player may increase her ultimate payoff. This persona has the same effect as a side-payment or commitment, as it is reflected in a temporary change to the player's utility function. The *persona games* model has been successful in explaining non-rational behaviour in non-repeated traveler's dilemma and even in versions of the non-repeated prisoner's dilemma.

A related commitment game between the players has been formalised in [Tennenholtz (2004)] as the notion of program equilibria (see [LaVictoire et. al. (2014), Oosterheld (2019)] for further extensions of this model). The program equilibrium model describes a scenario where each player designs a computer program that will determine their actions in the game. As in the persona game, it is necessary that the commitments will be made public: the programs can read the other players' programs, and their actions are contingent on the content of these programs (a simple example of such an equilibrium in the prisoner's dilemma is "cooperate if the other program is identical to this program or defect otherwise").

In [Kolumbus and Nisan (2022a), Kolumbus and Nisan (2022b)], the authors study scenarios of repeated games where players are assisted by learning algorithms. It is shown that in auctions and other game classes, the long-term outcomes of learning dynamics induce a meta-game between the players in which the players often have incentives to misreport their true preferences to their own learning algorithms. The commitment that players make in this case, namely, to use learning algorithms, is different in nature from notions of punishments or rewards studied in repeated games, since actions are not contingent on the opponents' play. Interestingly, in this model,

the details of the algorithms or the goals provided to them need not be made public, but rather they are revealed through the dynamics.²

There is a subset of the principal-agent literature concerning delegation games that is closely related to game manipulation. In these models, the principal is able to contract with an agent that will engage in a game with the principal’s opponent (or agent of the principal’s opponent). One concern of this literature is detailing the optimal contract for a principal (see [Vickers (1985), Fershtman (1985), Sklivas (1987)]). Another concern is whether a mechanism that allows specific types of contracts can lead to Pareto efficiency (see [Fershtman et al (1991), Katz (1991)]). Game manipulation is closely related to a previously unexplored aspect of principal-agent scenarios: the degree to which the agents can profit from delegation contracts.

Furthermore, our work is related to the general literature on commitment in games because, at its core, game manipulation is about what happens when players benefit by strategically restricting themselves. One well-studied aspect of commitment is the role of timing. Papers such as [Hamilton and Slutsky (1990)], [van Damme and Hurkens (1996)] and [Romano and Yildirim (2005)] concern endogenous timing and Stackelberg-like commitments. Another area of study is the role of commitment in repeated games. In their study of finitely repeated games, [Garcia-Jurado and Gonzalez-Diaz (2006)] introduce a weakening of SPE called virtual subgame perfect equilibrium. [Kalai et al (2007)] also study commitment in finitely repeated games, but do so in a manner similar to JW. That is, they are concerned with the role of commitment in bringing about efficiency.

The contracts offered by an external party in our case are also related to the broad literature on mechanism design, where a platform offers outcome-contingent incentives to the players, for example, in auctions [Milgrom (2004)], capital-raising [Halac et. al. (2022), Babaioff et. al. (2022)], and various social choice scenarios [Brandt et. al. (2016)]. Perhaps most closely related in this literature is [Monderer and Tennenholtz (2013)], who explore the cost needed from an external designer for implementing (i.e., incentivising) a given outcome in a game. This literature mostly focuses on a mechanism designer who aims to implement socially efficient outcomes. By contrast, we explore the potential implications of a self-interested external entity intervening in the interaction between our players. The latter scenario may, in principle, lead to a negative effect on society as long as the external party has profits.

References

[Arora et al (2012)] Arora, Sanjeev, Elad Hazan, and Satyen Kale. ”The multiplicative weights update method: a meta-algorithm and applications.” *Theory of computing* 8.1 (2012): 121-164.

²Our model naturally extends to repeated games with a fixed contract and to players who use no-regret algorithms, as such algorithms (almost) never play dominated strategies. We focus our discussion, however, on the single-shot case for ease of exposition.

- [Aumann, R.J. (1974)] Aumann, R.J., 1974. Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics*, 1(1), pp.67-96.
- [Babaioff et. al. (2022)] Babaioff, M., Kolumbus, Y. and Winter, E., 2022, April. Optimal collaterals in multi-enterprise investment networks. In *Proceedings of the ACM Web Conference 2022* (pp. 79-89).
- [Babaioff et. al. (2023)] Babaioff, M., Immorlica, N., Li, Y. and Lucier, B., 2023. Making Auctions Robust to Aftermarkets. In *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.
- [Bailey and Piliouras (2018)] Bailey, James P., and Georgios Piliouras. "Multiplicative weights update in zero-sum games." *Proceedings of the 2018 ACM Conference on Economics and Computation*. 2018.
- [Blum (2005)] Blum, Avrim. "On-line algorithms in machine learning." *Online algorithms: the state of the art* (2005): 306-325.
- [Blum and Mansour (2007)] A. Blum and Y. Mansour. *Learning, Regret Minimization and Equilibria*, in *Algorithmic Game Theory* N. Nisan, T. Roughgarden, E. Tardos and V. Vazirani (eds), Cambridge University Press, 2007.
- [Boadway and Bruce (1984)] Boadway RW, Bruce N. *Welfare economics*. Oxford: Blackwell, 1984.
- [Bono and Wolpert (2014)] Bono, J.W., Wolpert, D.H. (2014), *Game Mining: How to Make Money from those about to Play a Game*. In "Entangled Political Economy", Horwitz, Steven and Roger Koppl (Eds.), *Advances in Austrian Economics*, Vol. 18
- [Börgers and Sarin (1997)] Börgers, Tilman, and Sarin, Rajiv. "Learning through reinforcement and replicator dynamics." *Journal of economic theory* 77.1 (1997): 1-14.
- [Brandt et. al. (2016)] Brandt, F., Conitzer, V., Endriss, U., Lang, J. and Procaccia, A.D. eds., 2016. *Handbook of computational social choice*. Cambridge University Press.
- [Coase (1960)] Coase R (1960) The problem of social cost. *Journal of Law and Economics* 3(1):1-44
- [van Damme and Hurkens (1996)] van Damme E, Hurkens S (1996) Commitment robust equilibria and endogenous timing. *Games and Economic Behavior* 15:290-311
- [Fershtman (1985)] Fershtman C (1985) Managerial incentives as a strategic variable in a duopolistic environment. *International Journal of Industrial Organization* 3:245-253
- [Fershtman et al (1991)] Fershtman C, Judd KL, Kalai E (1991) Observable contracts: Strategic delegation and cooperation. *International Economic Review* 32(3)

- [Fudenberg and Levine (1995)] Drew Fudenberg and David K Levine. 1995. Consistency and cautious fictitious play. *Journal of Economic Dynamics and Control* 19, 5-7 (1995), 1065–1089.
- [Fudenberg and Levine (1998)] Fudenberg, Drew, and David K. Levine. The theory of learning in games. Vol. 2. MIT press, 1998.
- [Garcia-Jurado and Gonzalez-Diaz (2006)] Garcia-Jurado I, Gonzalez-Diaz J (2006) The role of commitment in repeated games. *Optimization* 55:1–13
- [Grigoriadis and Khachiyan (1995)] Grigoriadis, Michael D., and Khachiyan, Leonid G. "A sublinear-time randomized approximation algorithm for matrix games." *Operations Research Letters* 18.2 (1995): 53-58.
- [Halac et. al. (2022)] Halac, M., Kremer, I. and Winter, E., 2020. Raising capital from heterogeneous investors. *American Economic Review*, 110(3), pp.889-921.
- [Hamilton and Slutsky (1990)] Hamilton JH, Slutsky SM (1990) Endogeneous timing in duopoly games: Stackelberg or cournot equilibria. *Games and Economic Behavior* 2:29–46
- [Hart and Mas-Colell (2000)] Hart, S. and Mas-Colell, A., 2000. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5), pp.1127-1150.
- [Hofbauer and Sigmund (1998)] Hofbauer, Josef, and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge university press, 1998.
- [Hofbauer et al (2009)] J. Hofbauer, S. Sorin, and Y. Viossat. Time Average Replicator and Best-Reply Dynamics. *Mathematics of Operations Research* 34/2 (2009), 263-269.
- [Jackson and Wilkie (2005)] Jackson MO, Wilkie S (2005) Endogenous games and mechanisms: Side payments among players. *The Review of Economic Studies* 72(2):543–566
- [Kalai and Smorodinsky (1975)] Kalai E, Smorodinsky M (1975) Other solutions to Nash's bargaining problem. *Econometrica* 43(3):513-518
- [Kalai et al (2007)] Kalai A, Kalai E, Lehrer E, Samet D (2007) Voluntary commitments lead to efficiency, mimeo, Tel Aviv University
- [Katz (1991)] Katz ML (1991) Game-playing agents: Unobservable contracts as pre-commitments. *The RAND Journal of Economics* 22(3):307–328
- [Kleinberg et al (2009)] Kleinberg, Robert, Georgios Piliouras, and Éva Tardos. "Multiplicative updates outperform generic no-regret learning in congestion games." *Proceedings of the forty-first annual ACM symposium on Theory of computing*. 2009.
- [Kleinberg et al (2011)] R. Kleinberg, K. Ligett, G. Piliouras, and E. Tardos. 2011. Beyond the Nash equilibrium barrier. In *Symposium on Innovations in Computer Science (ICS)*.

- [Kolumbus and Nisan (2022a)] Kolumbus, Yoav, and Noam Nisan. “How and why to manipulate your own agent: On the incentives of users of learning agents.” *Advances in Neural Information Processing Systems* 35 (2022): 28080-28094.
- [Kolumbus and Nisan (2022b)] Kolumbus, Y. and Nisan, N., 2022, April. Auctions between regret-minimizing agents. In *Proceedings of the ACM Web Conference 2022* (pp. 100-111).
- [LaVictoire et al. (2014)] LaVictoire, P., Fallenstein, B., Yudkowsky, E., Barasz, M., Christiano, P. and Herreshoff, M., 2014, July. Program equilibrium in the prisoner’s dilemma via Löb’s theorem. In *AAAI workshop on multiagent interaction without prior coordination*.
- [Littlestone and Warmuth (1994)] Nick Littlestone and Manfred K Warmuth. 1994. The weighted majority algorithm. *Information and computation* 108, 2 (1994), 212–261.
- [Milgrom (2004)] Milgrom, P.R., 2004. Putting auction theory to work. Cambridge University Press.
- [Monderer and Tennenholtz (2013)] Monderer, D. and Tennenholtz, M., 2003, June. k-Implementation. In *Proceedings of the 4th ACM conference on Electronic Commerce* (pp. 19-28).
- [Morrill et al (2021)] Morrill, Dustin, et al. ”Efficient deviation types and learning for hindsight rationality in extensive-form games.” *International Conference on Machine Learning*. PMLR, 2021.
- [Nash (1950)] Nash J (1950) The bargaining problem. *Econometrica* 18(2):155-162
- [Oesterheld (2019)] Oesterheld, C., 2019. Robust program equilibrium. *Theory and Decision*, 86(1), pp.143-159.
- [Ramírez et al (2023)] Ramírez, María Alejandra, et al. ”Diversity enables the jump towards cooperation for the Traveler’s Dilemma.” *Scientific Reports* 13.1 (2023): 1441.
- [Renou (2009)] Renou L (2009) Commitment games. *Games and Economic Behavior* 66(1):488 – 505
- [Robbins (1932)] Robbins, Lionel. *An essay on the nature and significance of economic science*. London: Macmillan, 1932.
- [Romano and Yildirim (2005)] Romano R, Yildirim H (2005) On the endogeneity of cournot-nash and stackelberg equilibria: Games of accumulation. *Journal of Economic Theory* 120:73–107
- [Sandholm (2010)] Sandholm, William H. *Population games and evolutionary dynamics*. MIT press, 2010.

- [Sato et al (2002)] Sato, Yuzuru, Eizo Akiyama, and J. Doyne Farmer. "Chaos in learning a simple two-person game." *Proceedings of the National Academy of Sciences* 99.7 (2002): 4748-4751.
- [Schelling (1956)] Schelling, Thomas C. "An essay on bargaining." *The American Economic Review* 46.3 (1956): 281-306.
- [Selten (1965)] Selten, Reinhard. "Spieltheoretische behandlung eines oligopolmodells mit nachfragerträgeit: Teil i: Bestimmung des dynamischen preisgleichgewichts." *Zeitschrift für die gesamte Staatswissenschaft/Journal of Institutional and Theoretical Economics* H. 2 (1965): 301-324.
- [Sklivas (1987)] Sklivas S (1987) The strategic choice of management incentives. *The Rand Journal of Economics* 18:452–458
- [Smith (1982)] Smith, John Maynard. "Evolution and the Theory of Games." *Did Darwin get it right? Essays on games, sex and evolution.* Boston, MA: Springer US, 1982. 202-215.
- [Sobel (1981)] Sobel J (1981) Distortion of utilities and the bargaining problem. *Econometrica* 49:597–617
- [Tennenholtz (2004)] Tennenholtz, M., 2004. Program equilibrium. *Games and Economic Behavior*, 49(2), pp.363-373.
- [Vickers (1985)] Vickers J (1985) Delegation and the theory of the firm. *Economic Journal* (supplement) 95:138–147
- [Wolpert et al (2008)] Wolpert D, Jamison J, Newth D, Harre M (2008) Schelling revisited: Strategic selection of non-rational personas, available at SSRN: <http://ssrn.com/abstract=1172602>
- [Wolpert and Bono (2010)] Wolpert DH, Bono JW (2010) Distribution-valued solution concepts, SSRN working paper.