It Pays to Pay in Bi-Matrix Games – a Rational Explanation for Bribery

Anshul Gupta
Department of Computer Science
University of Liverpool, UK
Liverpool, United Kingdom
anshul@liverpool.ac.uk

Sven Schewe
Department of Computer Science
University of Liverpool
Liverpool, United Kingdom
sven.schewe@liverpool.ac.uk

ABSTRACT

We discuss bribery as a rational behaviour in two player bi-matrix games in game settings, where one player can assign the strategies to the other player. This can be observed as leader (or: Stackelberg) equilibria, where a leader assigns the strategy to herself and to the other player, who follows this lead unless he has an incentive not to. We make the rational assumption that a leader can further incentivise decisions of her follower, by bribing him with a small payoff value, and show that she can improve her gain this way. This results in an asymmetric equilibrium for a strategy profile: the incentive equilibrium. By 'asymmetric equilibrium', we refer to the strategy profile where a leader might benefit from deviation, while her follower does not. We observe that this concept is strong enough to obtain social optimum in the classic example of the prisoners' dilemma. We show that computing such incentive equilibria is no more expensive than computing leader equilibria: as opposed to Nash equilibria, they are both tractable. We evaluated our techniques on a large set of benchmarks (100,000 bi-matrix games) and provide the experimental results for incentive equilibrium.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent systems

General Terms

Economics

Keywords

Nash Equilibrium, Bi-matrix Games, Leader Equilibrium

1. INTRODUCTION

Nash Equilibria [11, 12] are a widely accepted way to reflect rational behaviour of players by defining the stable strategies with the intuition that only if no player gains from changing her strategy unilaterally, the strategy will be maintained. Thus, Nash equilibria are defined completely symmetrically.

We study bi-marix games [7] under game settings, where such a symmetry in determining the strategy profile is not natural. We therefore challenge the dominating role of symmetry in bi-matrix games. While symmetry is a common way of generalising rationality from two-player zero-sum games, we argue that it is not the only way, and that symmetry is not a global concept that always holds. To the contrary, the setting of many games is distinctively asymmetric. In some settings, one player has some power over the rules of the game. This includes customer relationship models, where the company has more power over the way they interact with their customers than the customer has. Likewise, in auctions, the auctioneer sets the rules of the auction in addition to being a stakeholder in each individual auction. It is natural in these settings to observe that a player with more power (the *leader* or *dictator* in this paper) would be in a position to select a strategy profile for herself and for the other player (the *follower* or *subject*). Therefore, the leader has the liberty to select a strategy profile where she might benefit upon deviation. This results in an asymmetry in the definition of optimal strategy profiles. A similar asymmetry in the definition of optimal strategy profiles occurs in von Stackelberg's *leader equilibria* [14], where the leader commits to a strategy and advises her follower to follow a particular strategy, which can then be pure.

In some sense, the leader dictates an equilibrium, and we therefore refer to her as the *dictator*. As the dictator assigns the strategies, we give her leeway to select those strategy profiles, she may improve upon. This gives her a wider base of strategies to select from as compared to Nash. The weaker player follows the assigned strategy, though only if it is in his immediate interest in the simpler sense that he cannot improve over the selected strategy: changing his strategy will not yield a better payoff. We therefore refer to him as the *subject* of the dictator. Such equilibria are not necessarily optimal for the leader in the Nash sense: it may be optimal for a leader to fix her strategy in a way that she would benefit from changing it.

In order to define such strategies, the leader needs to be able to communicate her strategy to the follower, and the follower needs to trust her to follow up on the strategy she has committed to. We argue that a leader who satisfies both requirements is also in a position to make further offers to her follower. We therefore consider it a natural assumption that she can do this by *incentivising* her follower to follow the strategy she has advised to, e.g., by paying her follower a small bribe. Incentivising the choice of her follower by paying an incentive $\iota \geq 0$ effectively changes the payoff matrices: it decreases the leader's payoff by the same amount ι, by which the follower's payoff is increased. We refer to these equilibria as *incentive equilibria*. Like with leader equilibria, the way the dictator arrives at an incentive equilibrium is manipulative and reflective. Although the result is an equilibrium, its definition is—like Stackelberg's leader equilibria [14]—by no means symmetric.

Bribery seems to be a very natural concept, and it might be argued to be almost unnatural to disregard it. In leader equilibria, the leader needs to announce what she does, and her follower needs to trust her to do so. In our view, the capacity of offering further incentives seems natural in all such settings, and we failed to think

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May, 4–8, 2015, Istanbul, Turkey.*

of an economic example where the popular leader equilibria sound appropriate while incentive equilibria do not. We therefore incorporated a simple incentive scheme, where the leader can incentivise different actions of the subject by offering different pay-off values.

As a first motivating example for incentive equilibria, we consider a travel agent, who sells flight tickets to a customer. The travel agent receives a fee from the company. These fees differ between different companies. Customers want to travel, but the possible connections are not of equal value to them. E.g., they might prefer arriving in Heathrow over arriving in Stansted, or prefer leaving at 10am over leaving at 5am. They value their benefit from buying the various connections on offer differently, e.g., as follows:

Airline	agent-value	customer-value
Air India:	$100	$170
German Wings:	$50	$100
Lufthansa:	$120	$80
Ryan Air:	$5	$200
US Airways:	$110	$150

The agent-value reflects the fees the agent charges to the company. The customer-value reflects the value the customer considers an offer to have (value of flight to him minus the cost of the ticket).

When the travel agent learns the priorities and valuations of the customer, she can make an offer to suit their needs. The agent cannot increase the offer by the company, but she can discount them on her own expenses. When the agent does not discount any offer, the customer would choose the Ryan Air flight, but the travel agent can simply offer the customer an incentive in form of a $30 discount on the Air India flight, thus increasing her gain from $5 to $70.

Customer relationship models of this type usually lead to only the customer making choices, as in the example above. When both parties have choices, we face the same constraints as for Stackelberg equilibria. In particular, the leader must know ex ante that the follower observes her action. We argue that, accepting that leader can do this, it is natural to assume that she can announce a reward for compliance with a suggestion.

When the leader announces her action, she effectively simplifies the bi-matrix game to a vector, like the one from the travel agent example. For a given selection of strategies, transferring utilities makes sense for the same reason as shown in the travel agent example. Optimising the chosen action and the incentive at the same time would naturally provide the best results in cases, where such a transferral of utilities is possible. It also explains why the transferral of utilities is uni-directional.

If there is insufficient trust in the transfer of utilities, one can involve a trusted third party, e.g., a solicitor/notary. While relying on a solicitor/notary is heavy handed, we would remark that this is only the extreme position that would arise when the follower does not trust the leader to transfer the utilities as promised. We would argue that, in this case, it is far less likely that he would trust her to follow the announced strategy: the prerequisite of Stackelberg's leader equilibria is the critical part, not the transfer of utilities.

Incentive equilibria in bi-matrix games. An important category of bi-matrix games consists of games similar to the prisoners dilemma [13]. This class of games, which is also known as arms-race games, reflects a situation, where two purely 'rational' individuals might not agree to co-operate even if it lies in their best mutual interest. Such situations where two entities might benefit significantly by co-operating—and similarly suffer if they fail to co-operate—are the greatest strength of incentive equilibria: in such cases, incentive equilibria provide socially optimal results.

Incentive equilibria in the prisoners' dilemma. The concept of incentive equilibria is strong enough to bring the socially optimum solution in games similar to the prisoners' dilemma. To exemplify the way incentive equilibria work, we use the concurrent bi-matrix game shown in Table 1, which represents the classic prisoners' dilemma [13], one of the most researched problems in the area of Nash equilibria. It has a famous antinomy that both prisoners do better if they both play co-operate (C), however, both of them have the dominating strategy to defect (D). Consequently, (D, D) is the only Nash equilibrium that is also the socially worst solution with a joint return of -16, while, if both of them had co-operated, it would have resulted in the socially optimal return of -2.

		Prisoner II	
		C	D
Prisoner I	C	$-1, -1$	$-10, 0$
	D	$0, -10$	$-8, -8$

Table 1: Prisoners' Dilemma

The first observation is that leader equilibria are not powerful enough to overcome this antinomy. The only strategy dictator, say Prisoner I, can advise successfully to her subject is the strategy to defect as defect is strictly superior for all dictator strategies here. In turn, her best strategy would be to defect, too, ending up in the same (D, D) as the only leader equilibrium. If the dictator also makes use of her power to *incentivise*, then she can bribe her subject, Prisoner II, into co-operation. An analysis with the techniques we develop in this paper shows that it would be optimal for her to commit to co-operating and to incentivise co-operation of Prisoner II by paying a bribery of value 1. As a result, (C, C) becomes the optimal choice, and we note that this is indeed the social optimum. Interestingly, in this example the power of the dictator benefits the subject more than the dictator: the dictator return after bribery is -2, while the subject return is 0. We will see that this is not always the case, and it is also not empirically the case on the random examples we generated to test our proof-of-concept implementation. As another example from the class of arms-race games, consider two countries, say the United States and Iran, who, because of a mutual threat, have to spend a considerable amount of their resources on the production of arms and their military. The countries can save a considerable amount if they both agree on not entering in an arms race. To achieve this, the more powerful country, say the United States, may pledge not to enter an arms race and, at the same time, offer the less powerful country (follower, here Iran) a trade advantage, or promise support in obtaining public events like Olympic games or any other major event in return for it not entering into an arms-race either. If the pledge is believed and the incentive is sufficiently high that the follower would have no incentive to enter the arms-race, then the dilemma is overcome. This also exemplifies that the incentive is not the main obstacle, but the trust in the pledge not to enter into the arms race.

The improvement the dictator (and the follower) can obtain through incentive equilibria is ε close to 1 compared to leader (and Nash) equilibria, when we norm all entries of the bi-matrix to be in $[0, 1]$. For this, we refer to Table 2, a variant of the prisoner's dilemma with pay-offs between 0 and 1. The social return in the incentive equilibrium is $2 - 2\varepsilon$, while in leader and Nash equilibria the social return is 2ε. Note that, for any $\delta > 0$, we can choose an ε, e.g., $\varepsilon = \min\{0.1, \delta/3\}$, to obtain an improvement greater than $1 - \delta$, for the dictator's return and the subject's return at the same time, using only values in $[0, 1]$ for the pay-offs.

Another well studied class of games are battle of sexes games.

We do not expand on these games as leader equilibria (and even

		Prisoner II	
		C	D
Prisoner I	C	$1-\varepsilon, 1-\varepsilon$	$0, 1$
	D	$1, 0$	ε, ε

Table 2: Prisoners' Dilemma

		Player 2	
		I	II
Player 1	I	$1, 2$	$0, 0$
	II	$0, 0$	$2, 1$

Table 3: A Battle-of-Sexes game

Nash equilibria) are sufficient to obtain any optimal solution in such games. The incentive would therefore play no role in them. Note that the example Battle-of-Sexes game from Table 3 has only one leader/incentive equilibrium, but three Nash equilibria: the pure strategies where both players play I or both play II and a third mixed strategy equilibrium where Player 1 plays I with probability $\frac{1}{3}$, and Player 2 plays I with probability $\frac{2}{3}$. The outcome of these strategy profiles is $(1, 2)$, $(2, 1)$, and $(\frac{2}{3}, \frac{2}{3})$, respectively. Note that, even when one restricts the focus to those strategy profiles only, it needs to be negotiated which of them is taken. The complexity of this negotiation is outside of the complexity to determine Nash equilibria in the first place, but it is yet another level of complexity that is reduced by using incentive (or leader) equilibria.

Friendly equilibria in bi-matrix games. The classic prisoners' dilemma example discussed above explains how incentive equilibria improve over leader equilibria and Nash equilibria. While always beneficial for the dictator, it is interesting to note that it is often beneficial for both players. To get an intuition why incentive equilibria improve over leader and Nash equilibria, note that the dictator can choose from strategies that satisfy the side-constraint that the subject cannot improve over it by unilateral deviation. For incentive equilibria, the dictator can use incentives, whereas leader equilibria would optimise only over strategies without incentive (or: with zero incentives). Thus, incentive equilibria are again an optimum over a large base than leader equilibria. This optimisation may further leave a plateau of jointly optimal strategies, strategies with the same optimal payoff (after bribery) for the dictator. We argue that the dictator can-and should-move a step ahead and choose an equilibrium that is optimal for her subject among these otherwise equivalent solutions. This is another advantage of a clear outcome for the dictator – the option to choose ex-aequo an incentive equilibrium, which is good for the subject. We call an incentive equilibrium **friendly** if it is optimal and assigns, among this class of equilibria, the highest payoff to the subject. Friendly incentive equilibria are therefore in favour of the subject as well: when the dictator assigns strategies to herself and the other player, her primary objective is to maximise her own benefit and her secondary objective is that the subject receives a high gain, too. Thus, they refer to a situation where both, the dictator and her subject, benefit, increasing the social quality of result.

Tractability and purity of incentive equilibria. Another appealing property of general and friendly incentive equilibria is their simplicity: it suffices for the subject to consider pure strategies, or, similarly, it suffices for the dictator to assign pure strategies to her subject. Another strong argument in favour of general and friendly incentive equilibria is that they are *tractable*, whereas the complexity of finding mixed strategy Nash equilibria is known to be PPAD-complete [3].

We give an algorithm for the computation of incentive and friendly incentive equilibria and empirically evaluated the technique presented here on the randomly generated 100,000 data-sets with continuous pay-off values and integer pay-off values for the evaluation of friendly incentive equilibria.

Related Work. Leader equilibria, introduced by von Stackelberg [14] and therefore sometimes referred to as Stackelberg equilibria, have been studied in depth in Oligopoly theory [4]. The main contribution of our work is conceptual, and the closest relation of our equilibrium concept is to Stackelberg's leader equilibria [2, 15]. Ehtamo and Hamalainen in [**?**] considered the construction of optimal incentive strategies in two-player dynamic game problems that are described by integral convex cost criteria. They considered the problem of finding an incentive strategy for the leader by looking at the rational response from the follower, such that in an optimal strategy, cost functional of leader could be minimised. They further considered in [**?**], the analytical methods for constructing memory incentive strategies for continuous time decision problem. The strategies they study are time-consistent, that means the continuation of equilibrium solution remains an equilibrium. Stark in [**?**] studied how introduction of altruism into non cooperative game settings can lead to an improved quality for both the agents. Stark [**?**] further discussed special altruism – 'a mutual altruism' and their role in various contexts. Stengel and Zamir [15] studied a commitment model in a leadership game with mixed extensions of bi-matrix games. They show that the possibility to commit to a strategy profile in bi-matrix games is always beneficial for the committing player. Conitzer and Sandholm [2] gives first insight into the computation of Stackelberg strategies in normal-form games. [1] considered a mechanism designer modelled as a player in the game who has the opportunity to modify the game. In their setting, the players' utility and social welfare is seen as counter intuitive. E.g., social welfare may arbitrarily come worse and they focus completely on pure strategies. Whereas, our solution approach always gives socially optimal outcomes and the equilibrium we study is mixed: it suffices for the leader to play mixed strategies while to assign pure strategy to the subject. Another related work is [10], in which an external party who has no control over the rules of the game can influence the outcome of the game by committing to non-negative monetary transfers for the different strategy profiles that may be selected by the agents in a multi-agent interaction.

2. DEFINITIONS

Formally, a bi-matrix game is defined as $\mathcal{G}(A, B)$, where A and B are the real valued $m \times n$ payoff matrices for the dictator and the subject, respectively. In our settings, the dictator is the row player and the subject, is the column player. In the remainder, we refer to the number of rows by m and to the number of columns by n. We also refer to the entry in row i and column j of A and B by a_{ij} and b_{ij}, respectively. A (mixed) strategy of the dictator is a probability vector $\sigma = (p_1, \ldots, p_m)$, i.e., $\sum_{i=1}^{m} p_i = 1$ (the sum of the weights is 1) and $p_i \geq 0$ for $1 \leq i \leq m$. Likewise, a (mixed) strategy of the subject is a probability vector $\delta^T = (q_1, \ldots, q_n)$, i.e., $\sum_{j=1}^{n} q_j = 1$ and $q_j \geq 0$ for $1 \leq j \leq n$. Where convenient, we read δ^T as a function and refer to the j^{th} column of δ by $\delta^T(j)$. For a given subject strategy $\delta^T = (q_1, \ldots, q_n)$, we define its support as the positions with non-zero probability, and denote it as $\mathsf{support}(\delta^T) = \{j \leq n \mid \delta^T(j) > 0\}$. We also define a bribery vector $\beta^T = (\beta_1, \ldots, \beta_n)$, with non-negative entries $\beta_j \geq 0$ for all $j = 1, \ldots, n$, corresponding to each decision 1 through n of the subject, where the dictator pays her subject the bribery β_j when he plays j. We call strategies with singleton support *pure*. We are particularly interested in pure subject strategies, and abbreviate the strategy with $\delta^T(j) = 1$ by j, or by $\overrightarrow{j}$ to emphasise that j refers to a strategy. We call a bribery vector $\beta^T = (\beta_1, \ldots, \beta_n)$ a j-bribery if it incentivises only playing j, that is, if $\beta_i = 0$ for all $i \neq j$.

A pair of strategies $\langle\sigma, \delta\rangle$ is called a *strategy profile*. For a strategy profile $(\langle\sigma, \delta\rangle, \beta)$ with bribery vector β, we denote the dictator payoff by $\mathsf{dpayoff}(A; \langle\sigma, \delta\rangle, \beta) = \sigma A\delta - \delta\beta$ and subject payoff by $\mathsf{spayoff}(B; \langle\sigma, \delta\rangle, \beta) = \sigma B\delta + \delta\beta$. A strategy profile $(\langle\sigma, \delta\rangle, \beta)$ is called *stable* under bribery or bribery stable, if $\mathsf{spayoff}(B; \langle\sigma, \delta\rangle, \beta) \geq \mathsf{spayoff}(B; \langle\sigma, \delta'\rangle, \beta)$ holds for all subject strategies δ'. A bribery stable strategy profile $(\langle\sigma, \delta\rangle, \beta)$ is called an *incentive equilibrium*, if $\mathsf{dpayoff}(A; \langle\sigma, \delta\rangle, \beta) \geq \mathsf{dpayoff}(A; \langle\sigma', \delta'\rangle, \beta')$ holds for all bribery stable strategy profiles $(\langle\sigma', \delta'\rangle, \beta')$. An incentive equilibrium $(\langle\sigma, \delta\rangle, \beta)$ is called *friendly*, if $\mathsf{spayoff}(B; \langle\sigma, \delta\rangle, \beta) \geq \mathsf{spayoff}(B; \langle\sigma', \delta'\rangle, \beta')$ holds for all incentive equilibria $(\langle\sigma', \delta'\rangle, \beta')$. A bribery stable strategy profile (and an incentive equilibrium) $(\langle\sigma, \delta\rangle, \beta)$ is called simple, if δ is pure.

3. INCENTIVE EQUILIBRIA

In this section, we show that friendly incentive equilibria always exists. Moreover, there is always a *simple* friendly incentive equilibrium.

Existence of bribery stable strategy profiles. The existence of bribery stable strategy profiles is implied by the existence of Nash equilibria [3], as Nash equilibria are special cases of bribery stable strategy profiles with the zero bribery vector.

THEOREM 1. *Every bi-matrix game $\mathcal{G}(A, B)$ has a Nash equilibrium [7].*

COROLLARY 2. *Every bi-matrix game $\mathcal{G}(A, B)$ has a bribery stable strategy profile.*

Optimality of simple bribery stable strategy profiles. Different to Nash equilibria, there are always simple incentive equilibria. In order to show this, we first show that there is always a simple bribery stable strategy profile. This is because, for a given bribery stable strategy profile $(\langle\sigma, \delta\rangle, \beta)$ and a j in the support of δ^T, $(\langle\sigma, j\rangle, \beta)$ is an incentive equilibrium, too.

THEOREM 3. *For every bi-matrix game $\mathcal{G}(A, B)$ and every bribery stable strategy profile $(\langle\sigma, \delta\rangle, \beta)$ and $\mathsf{dpayoff}(A; \langle\sigma, j\rangle, \beta) = v$ for the dictator, there is always a simple bribery stable strategy profile with $\mathsf{dpayoff}(A; \langle\sigma, j\rangle, \beta) \geq v$ for the dictator.*

PROOF. Let $S = \mathsf{support}(\delta^T)$ and let $s = \mathsf{spayoff}(B; \langle\sigma, \delta\rangle, \beta)$ be the payoff for the subject in this simple bribery stable strategy profile. We first argue that, for all $j \in S$, $(\langle\sigma, j\rangle, \beta)$ is a simple bribery stable strategy profile with the same subject payoff as $(\langle\sigma, \delta\rangle, \beta)$. First, the subject payoff cannot be higher, as $(\langle\sigma, j\rangle, \beta)$ would otherwise not be a simple bribery stable strategy profile (the subject could improve his payoff by changing his strategy to j). Assuming for contradiction that there is an $j \in S$ with $\mathsf{spayoff}(B; \langle\sigma, j\rangle, \beta) < s$ implies, together with the previous observation that $\mathsf{spayoff}(B; \langle\sigma, j\rangle, \beta) \leq s$ holds for all $j \in S$, that $\sum_{j\in S} \delta^T(j) \cdot \mathsf{spayoff}(B; \langle\sigma, j\rangle, \beta) < s$, which contradicts $s = \mathsf{spayoff}(B; \langle\sigma, \delta\rangle, \beta)$. Taking into account that the dictator payoff $v = \sum_{j\in S} \delta^T(j) \cdot \mathsf{dpayoff}(A; \langle\sigma, j\rangle, \beta)$ is an affine combination of the dictator payoffs for these simple bribery stable strategy profiles, there is some $j \in S$ with $\mathsf{dpayoff}(A; \langle\sigma, j\rangle, \beta) \geq v$. □

Description of simple bribery stable strategy profile. Theorem 3 allows us to seek incentive equilibria only among simple bribery stable strategy profiles. Note that this is in contrast to general Nash equilibria, cf. the rock-paper-scissors game. Simple bribery stable strategy profiles are defined by a set of linear inequations.

THEOREM 4. *For a bi-matrix game $\mathcal{G}(A, B)$, $(\langle\sigma, j\rangle, \beta)$ is a simple bribery stable strategy profile if, and only if, $\sigma B\overrightarrow{j} + \beta\overrightarrow{j} \geq \sigma B\overrightarrow{i} + \beta\overrightarrow{i}$ holds for all pure strategies i=$1, \ldots, n$ of the subject.*

PROOF. If $(\langle\sigma, j\rangle, \beta)$ is a simple bribery stable strategy profile, then in particular changing the strategy to a different pure strategy i cannot be beneficial for the subject. Consequently, it holds for all $i = 1, \ldots, n$ that $\sigma B\overrightarrow{j} + \beta\overrightarrow{j} \geq \sigma B\overrightarrow{i} + \beta\overrightarrow{i}$. If it holds for all $i = 1, \ldots, n$ that $\sigma B\overrightarrow{j} + \beta\overrightarrow{j} \geq \sigma B\overrightarrow{i} + \beta\overrightarrow{i}$, then we note that, for any subject strategy δ, the payoff under σ is an affine combination $\mathsf{spayoff}(B; \langle\sigma, \delta\rangle, \beta) = \sum_{i\in S} \delta(i) \cdot \sigma B\overrightarrow{i} + \beta\overrightarrow{i}$ of the payoffs for the individual pure strategies. Using $\sigma B\overrightarrow{j} + \beta\overrightarrow{j} \geq \sigma B\overrightarrow{i} + \beta\overrightarrow{i}$, we get $\mathsf{spayoff}(B; \langle\sigma, \delta\rangle) = \sum_{i\in S} \delta(i) \cdot \sigma B\overrightarrow{i} + \beta\overrightarrow{i} \leq \sum_{i\in S} \delta(i) \cdot \sigma B\overrightarrow{j} + \beta\overrightarrow{j} = \sigma B\overrightarrow{j} + \beta\overrightarrow{j} = \mathsf{spayoff}(B; \langle\sigma, j\rangle, \beta)$. □

THEOREM 5. *For a bi-matrix game $\mathcal{G}(A, B)$ with a bribery vector $\beta' = (\beta'_1, \ldots, \beta'_n)^T$, and for a simple bribery stable strategy profile $(\langle\sigma, j\rangle, \beta')$, we can replace β' by a j-bribery vector $\beta = (\beta_1, \ldots, \beta_n)^T$ with $\beta_j = \beta'_j$.*

PROOF. According to Theorem 3, for a bi-matrix game $\mathcal{G}(A, B)$ and a bribery vector $\beta' = (\beta'_1, \ldots, \beta'_n)^T$, there is always a simple bribery stable strategy profile $\delta^T = \overrightarrow{j}$ with optimal pay-off for the dictator. We now choose $\delta^T = \overrightarrow{j}$ and $\beta = (\beta_1, \ldots, \beta_n)^T$, with $\beta_j = \beta'_j$ and $\beta_i = 0$ for all $i \neq j$. We first observe that $(\langle\sigma, j\rangle, \beta)$ provides the same dictator and subject payoff as $(\langle\sigma, j\rangle, \beta')$. Next we observe that if $(\langle\sigma, j\rangle, \beta')$ is bribery stable, then so is $(\langle\sigma, j\rangle, \beta)$ by Theorem 4. □

In our equation systems, we can therefore focus on simple incentive equilibria $(\langle\sigma, j\rangle, \beta)$ with j-bribery β. The value of this j-bribery β can then be described by the value of the incentive $\iota = \beta_j$ that the dictator gives to her subject to solicit him to play j.

Computing Incentive equilibria. This invites the definition of a constraint system $\mathcal{C}_j^{\mathcal{G}(A,B)}$ for each pure subject strategy j, that describes the vectors σ by $\sigma = (p_1, \ldots, p_m)$. Theorem 4 and Theorem 5 invites to reflect the fact that $(\langle\sigma, j\rangle, \beta)$ is a simple bribery stable strategy profile with j-bribery β by using a constraint system. This constraint system $\mathcal{C}_j^{\mathcal{G}(A,B)}$ consists of $m + n + 1$ constraints, where $m + 1$ constraints describe that σ is a strategy,

- $\sum_{i=1}^{m} p_i = 1$ (the sum of the weights is 1) and
- the m non-negativity requirements $p_i \geq 0$ for $1 \leq i \leq m$,

and $n - 1$ constraints reflect the conditions from Theorem 4 on an incentive equilibrium. That is, $\mathcal{C}_j^{\mathcal{G}(A,B)}$ contains $n - 1$ constraints of the form

- $\sum_{k=1}^{m} (b_{jk} - b_{ik})p_k + \iota \geq 0$,

one for each $i \neq j$ with $1 \leq i \leq n$, and a non-negativity constraint on the bribery value ι,

- $\iota \geq 0$,

As this reflects the conditions from Theorem 4, we first get the following corollary.

COROLLARY 6. *The solutions to $\mathcal{C}_j^{\mathcal{G}(A,B)}$ describe the set of dictator strategy and j-bribery vector pairs (σ, β) such that $(\langle\sigma, j\rangle, \beta)$ is bribery stable.*

EXAMPLE 7. *If we consider the first strategy of PrisonerII from the Prisoner's Dilemma from Table 1, then $\mathcal{C}_1$ consists of the constraints*

- $p_1 + p_2 = 1$,
- $\iota, p_1, p_2 \geq 0$, *and*
- $\iota - p_1 - 2p_2 \geq 0$.

Note that the constraints do *not* depend on the payoff matrix of the dictator. What depends on the payoff matrix of the dictator is the formalisation of the objective. We denote with $\mathcal{LP}_j^{\mathcal{G}(A,B)}$ the linear programming problem that consists of the constraint system $\mathcal{C}_j^{\mathcal{G}(A,B)}$ and the objective function

$$\sum_{k=1}^{m} a_{jk} p_k - \iota \mapsto \max,$$

where $\sum_{k=1}^{m} a_{jk} p_k - \iota = \mathsf{dpayoff}(A; \langle\sigma, j\rangle, \beta)$ is the payoff the dictator obtains for such a simple bribery stable strategy profile $(\langle\sigma, j\rangle, \beta)$ with $\sigma = (p_1, \ldots, p_m)$ and $\beta_j = \iota \geq 0$ is the bribery value of the j-bribery vector β.

COROLLARY 8. *The solutions to $\mathcal{LP}_j^{\mathcal{G}(A,B)}$ describe the set of dictator strategy and j-bribery vector pairs (σ, β) such that $(\langle\sigma, j\rangle, \beta)$ is bribery stable and the dictator return $\mathsf{dpayoff}(A; \langle\sigma, j\rangle, \beta)$ is maximal among simple bribery stable strategy profiles with subject strategy j and j-bribery vectors β.*

EXAMPLE 9. *If we consider the first strategy of PrisonerII from the Prisoner's Dilemma from Table 1, then the $\mathcal{LP}_1$ consists of the constraints from $\mathcal{C}_1$ and the objective*

$$-p_1 - \iota \mapsto \max.$$

This provides us with a simple algorithm for determining (simple) incentive equilibria.

COROLLARY 10. *To find an incentive equilibrium for a game $\mathcal{G}(A, B)$, it suffices to solve the linear programming problems $\mathcal{LP}_j^{\mathcal{G}(A,B)}$ for all $1 \leq j \leq n$, to select a i with maximal solution among them, and to use a solution $(\langle\sigma_i, i\rangle, \beta)$, where β is a i-bribery vector with $\beta_i = \iota$ from the solution of $\mathcal{LP}_i^{\mathcal{G}(A,B)}$. This solution $(\langle\sigma_i, i\rangle, \beta)$ is an incentive equilibrium.*

PROOF. Obviously, $\mathcal{LP}_j^{\mathcal{G}(A,B)}$ has *some* solution iff $\mathcal{C}_j^{\mathcal{G}(A,B)}$ is satisfiable, and thus, by Corollary 6, if there is a dictator strategy σ and a j-bribery vector β such that $(\langle\sigma, j\rangle, \beta)$ is a simple bribery stable strategy profile. In this case, a solution to $\mathcal{LP}_j^{\mathcal{G}(A,B)}$ reflects an optimal simple bribery stable strategy profile among the pure subject strategies that always play j; in particular, such an optimum exists. (Note that the dictator return is bounded from above by the highest entry a_{ij} in her payoff matrix). Thus, the maximal return from some $\mathcal{LP}_i^{\mathcal{G}(A,B)}$ is the optimal solution among all simple bribery stable strategy profiles. Assuming that a better non-simple bribery stable strategy profile exists implies by Theorem 3 that a better simple incentive equilibrium exists as well and thus leads to a contradiction. Note that the existence of some simple bribery stable strategy profile is established by Corollary 2 and Theorem 3. □

As linear programming problems can be solved in polynomial time [5, 6], finding an incentive equilibria is tractable.

COROLLARY 11. *An optimal incentive equilibrium can be constructed in polynomial time.*

Friendly incentive equilibria. As discussed in the introduction, the dictator follows a secondary objective of being benign to the subject. We have seen that it is cheap and simple to determine the value $v_{\max}^d$ that the dictator can at most acquire in an incentive equilibrium. It is therefore an interesting follow-up question to determine the highest payoff $v_{\max}^s$ for her subject in an incentive equilibrium with dictator payoff $v_{\max}^d$, i.e., to construct and evaluate *friendly* incentive equilibria. We first observe that friendliness does not come to the cost of simplicity.

THEOREM 12. *For every bi-matrix game $\mathcal{G}(A, B)$ and every incentive equilibrium $(\langle\sigma, \delta\rangle, \beta)$ with payoff v for the subject, it holds for all $j \in \mathsf{support}(\delta^T)$ that $(\langle\sigma, j\rangle, \beta)$ is an incentive equilibrium with payoff v for the subject.*

PROOF. Let $S = \mathsf{support}(\delta^T)$ be the support of δ^T and let $v_{\max}^d = \mathsf{dpayoff}(A; \langle\sigma, \delta\rangle, \beta)$ be the dictator payoff for $(\langle\sigma, \delta\rangle, \beta)$. In the proof of Theorem 3 we have shown that, for all $j \in S$, $(\langle\sigma, j\rangle, \beta)$ is a simple bribery stable strategy profile with the same payoff $\mathsf{spayoff}(B; \langle\sigma, j\rangle, \beta) = \mathsf{spayoff}(B; \langle\sigma, \delta\rangle, \beta)$ for the subject. To establish that $(\langle\sigma, j\rangle, \beta)$ is also an incentive equilibrium, we first note that the dictator payoff cannot be higher than in an incentive equilibrium, such that $\mathsf{dpayoff}(A; \langle\sigma, j\rangle, \beta) \leq v_{\max}^d$ holds for all $j \in S$. Assuming for contradiction that there is an $j \in S$ with $\mathsf{dpayoff}(A; \langle\sigma, j\rangle, \beta) < v_{\max}^d$ would, together with the previous observation that $\mathsf{dpayoff}(A; \langle\sigma, j\rangle, \beta) \leq v_{\max}^d$ holds for all $j \in S$, imply that the affine combination $\sum_{j \in S} \delta(j) \cdot \mathsf{spayoff}(B; \langle\sigma, j\rangle, \beta) < v_{\max}^d$ of these values defined by δ is strictly smaller than $v_{\max}^d$, and would therefore lead to a contradiction. □

THEOREM 13. *For a bi-matrix game $\mathcal{G}(A, B)$ and every incentive equilibrium $(\langle\sigma, \delta\rangle, \beta)$ with payoff v for the subject, there is always a pure subject strategy j with a j-bribery vector β' such that $(\langle\sigma, j\rangle, \beta')$ is an incentive equilibrium with subject payoff $\geq v$.*

PROOF. First, according to Theorem 12, we can observe that there is always a pure subject strategy j, such that $(\langle\sigma, j\rangle, \beta)$ is an incentive equilibrium for the subject as it returns the maximal gain for the subject. Following the same argument as in the proof of Theorem 5, we can amend β by only incentivising the subject to play j, replacing all other entries β_i, $i \neq j$, by $\beta_i' = 0$ and setting $\beta_j' = \beta_j$. The payoff for the dictator and subject are unaffected, and if $(\langle\sigma, j\rangle, \beta)$ is bribery stable, so is $(\langle\sigma, j\rangle, \beta')$: the subject return for playing j is unaffected, while the subject return for all other strategies is not increased. □

Like in the quest for ordinary incentive equilibria, we can therefore focus on pure subject strategies j and the respective j-bribery vectors when seeking friendly incentive equilibria. Recall that each constraint system $\mathcal{C}_j^{\mathcal{G}(A,B)}$ describes the set of dictator strategies σ and gives a j-bribery vector β, such that $(\langle\sigma, j\rangle, \beta)$ is a simple bribery stable strategy profile. In order to be an incentive equilibrium, it also has to satisfy the optimality constraint

$$\sum_{k=1}^{m} a_{jk} p_k - \iota \geq v_{\max}^d.$$

We refer to the extended constraint system by $\mathcal{E}_j^{\mathcal{G}(A,B)}$ By Corollary 6, the set of solutions to this constraint system is non-empty iff there is an incentive equilibrium of the form $(\langle\sigma, j\rangle, \beta)$.

COROLLARY 14. *The solutions to $\mathcal{E}_j^{\mathcal{G}(A,B)}$ describes the set of dictator strategies σ and a bribery value ι, such that $(\langle\sigma, j\rangle, \beta)$ is an incentive equilibrium for $\mathcal{G}(A, B)$.*

EXAMPLE 15. *Considering again the Prisoner's Dilemma from Table 1, $\mathcal{E}_1$ consists of the constraints from $\mathcal{C}_1$ plus the optimality constraint*

$$-p_1 - \iota \geq -2$$

We now extend the constraint system $\mathcal{E}_j^{\mathcal{G}(A,B)}$ to an extended linear programming problem $\mathcal{ELP}_j^{\mathcal{G}(A,B)}$ by adding the objective

$$\sum_{k=1}^{m} b_{jk} p_k + \iota \mapsto \max$$

COROLLARY 16. *The solutions to $\mathcal{LP}_j^{\mathcal{G}(A,B)}$ describe the set of dictator strategy and j-bribery vector pairs (σ, β) such that $(\langle \sigma, j \rangle, \beta)$ is an incentive equilibrium that satisfies, if such a solution exists, that the subject return $\mathsf{spayoff}(B; \langle \sigma, j \rangle, \beta)$ is maximal among these simple bribery stable strategy profiles with subject strategy j and j-bribery vectors β.*

EXAMPLE 17. *If we consider the Prisoner's Dilemma from Table 1, then $\mathcal{ELP}_1$ consists of the constraints from $\mathcal{E}_1$ and the objective*

$$-p_1 - 10p_2 + \iota \mapsto \max$$

Together with the observation of Theorem 12, Corollary 16 provides an algorithm for finding a friendly incentive equilibrium.

COROLLARY 18. *To find a friendly incentive equilibrium for a game $\mathcal{G}(A, B)$, it suffices to solve the linear programming problems $\mathcal{ELP}_j^{\mathcal{G}(A,B)}$ for all $1 \leq j \leq n$, to select a i with maximal solution among them, and to use a solution $(\langle \sigma_i, i \rangle, \beta)$, where β is a i-bribery vector with $\beta_i = \iota$ from the solution of $\mathcal{ELP}_i^{\mathcal{G}(A,B)}$. This solution $(\langle \sigma_i, i \rangle, \beta)$ is a friendly incentive equilibrium.*

PROOF. $\mathcal{ELP}_j^{\mathcal{G}(A,B)}$ has *some* solution iff $\mathcal{E}_j^{\mathcal{G}(A,B)}$ is satisfiable, and thus, by Corollary 12, if there is a dictator strategy σ such that $(\langle \sigma, j \rangle, \beta)$ is an incentive equilibrium. In this case, a solution to $\mathcal{ELP}_j^{\mathcal{G}(A,B)}$ reflects an incentive equilibrium with the maximal subject payoff among the incentive equilibria of the form $(\langle \sigma, j \rangle, \beta)$; in particular, such an optimum exists. Thus, the returned result is the optimal solution among all simple incentive equilibria. Assuming that a better non-simple friendly incentive equilibrium exists implies by Theorem 12 that a better simple friendly incentive equilibrium exists as well, and thus leads to a contradiction. Note that the existence of some simple optimal incentive equilibrium is implied by Corollary 10. $\square$

As linear programming problems can be solved in polynomial time [5, 6], finding friendly incentive equilibria is tractable.

COROLLARY 19. *A simple friendly incentive equilibrium can be constructed in polynomial time.*

Incentive equilibria in zero-sum games. Here, we establish friendliness of incentive equilibria in zero-sum games and show that the dictator can not gain anything by paying a bribery in zero-sum games. Zero-sum games are bi-matrix games where the gain of the dictator is the loss of the subject and vice versa. They satisfy $a_{ij} = -b_{ij}$ for all $1 \leq i \leq m$ and all $1 \leq j \leq n$. Different to the general bi-matrix games, zero-sum games are determined in that they have determined expected payoffs for both the players when both play rational. A rational behaviour for zero-sum games is therefore an acid test for new concepts: they do not have different levels of reasoning, and the opponent is predictable. We would like to make a few simple observations to show that incentive equilibria pass this acid test.

THEOREM 20. *If $\langle \sigma, \delta \rangle$ is a Nash equilibrium in a zero-sum game and $j \in \mathsf{support}(\delta^T)$, then $\langle \sigma, j \rangle$ is a friendly incentive equilibrium with zero bribery vector β_0 and with $\mathsf{dpayoff}(A; \langle \sigma, j \rangle, \beta_0) = \mathsf{dpayoff}(A; \langle \sigma, \delta \rangle, \beta_0)$.*

PROOF. We first establish that $(\langle \sigma, j \rangle, \beta_0)$ is a simple bribery stable strategy profile. To see this, we use that $(\langle \sigma, \delta \rangle, \beta_0)$ is a Nash equilibrium, and therefore $\mathsf{spayoff}(B; \langle \sigma, j \rangle, \beta_0) \leq \mathsf{spayoff}(B; \langle \sigma, \delta \rangle, \beta_0)$ holds for all $j \leq n$. Assuming that this inequation is strict for any $j \in \mathsf{support}(\delta^T)$ violates $\mathsf{spayoff}(B; \langle \sigma, \delta \rangle, \beta_0) = \sum_{j=1}^{n} \delta(j) \cdot \mathsf{spayoff}(B; \langle \sigma, j \rangle, \beta_0))$. Second, we observe that playing δ offers a return $\mathsf{spayoff}(B; \langle \sigma', \delta \rangle, \beta_0) \geq \mathsf{spayoff}(B; \langle \sigma, \delta \rangle, \beta_0)$, as otherwise the dictator could benefit by deviating from $(\langle \sigma, \delta \rangle, \beta_0)$. By a similar "$\leq$, but not $<$ as it would contradict $\geq$ from the affine combination" argument, we can lead the assumption that $\mathsf{spayoff}(B; \langle \sigma', \delta \rangle, \beta_0) > \mathsf{spayoff}(B; \langle \sigma', j \rangle, \beta_0)$ holds to a contradiction for all $j \in \mathsf{support}(\delta^T)$. Consequently there is, for all dictator strategies σ', an $j \in \mathsf{support}(\delta^T)$ such that $\mathsf{spayoff}(B; \langle \sigma', j \rangle, \beta_0) \geq \mathsf{spayoff}(B; \langle \sigma, \delta \rangle, \beta_0)$. The zero-sum property then provides the claim. $\square$

Note that all incentive equilibria are for this reason friendly in zero-sum games. The definition of a simple bribery stable strategy profile now shows that the dictator return $\mathsf{dpayoff}(A; \langle \sigma, j \rangle, \beta_0)$ can only be improved when the subject changes her strategy. (Note that this does not generally hold for non-zero-sum games.) Consequently, her incentive equilibrium provides her with the same guarantee as her rational strategy from zero-sum games. Her subject might be left exploitable, but in the selected strategy profile, he will receive the same payoff as with a rational strategy, but does not have to resort to randomisation. As these games are symmetric, this in particular implies that both players can play dictator strategies in zero-sum games, and thus dictator can not gain anything by paying a bribery (as it is applicable only when there is an asymmetry).

COROLLARY 21. *If $(\langle \sigma, j \rangle, \beta_0)$ is an incentive equilibrium in a zero-sum game $\mathcal{G}(A, B)$ and $(\langle \delta, i \rangle, \beta_0')$ is an incentive equilibrium in $\mathcal{G}(B, A)$, where β_0 and β_0' are the zero vectors, then $\langle \sigma, \delta \rangle$ is a Nash equilibrium in $\mathcal{G}(A, B)$.*

Naturally, this does not extend to general bi-matrix games.

THEOREM 22. *The dictator payoff in incentive equilibria grows monotonously in the payoff matrix of the dictator. If all entries grow strictly, so does the dictator payoff.*

PROOF. As observed earlier, the individual constraint systems do not depend on the payoff matrix of the dictator, and the set of simple bribery stable strategy profiles is not affected by replacing a payoff matrix A by an entry-wise greater payoff matrix A'. An incentive equilibrium for A is thus a bribery stable strategy profile for A', too, and the payoff of this equilibrium for A' has the required properties. Consequently, an incentive equilibrium for A' has them as well. $\square$

THEOREM 23. *If $(\langle \sigma, j \rangle, \beta)$ is a friendly incentive equilibrium with j-bribery vector β, then j is the socially optimal response to σ.*

PROOF. First, there is always a pure socially optimal response. Let us assume for contradiction that there is a socially better pure response i. For i to be socially strictly better than j, it must hold that $\sigma A \overrightarrow{i} + \sigma B \overrightarrow{i} > \sigma A \overrightarrow{j} + \sigma B \overrightarrow{j}$ (Note that bribery is socially neutral). This is equivalent to $\sigma A \overrightarrow{j} + \sigma B \overrightarrow{j} - \sigma A \overrightarrow{i} < \sigma B \overrightarrow{i}$. As $(\langle \sigma, j \rangle, \beta)$ is a friendly equilibrium and $\beta^T = (\beta_1, \ldots, \beta_n)$ is a

j-bribery vector, we know that $\sigma B \overrightarrow{j} + \beta_j \geq \sigma B \overrightarrow{i}$. In order to incentivise the subject to play i, it suffices to choose a i-bribery vector $\beta' = (\beta'_1, \ldots, \beta'_n)^T$ such that $\sigma B \overrightarrow{i} + \beta'_i \geq \sigma B \overrightarrow{j} + \beta_j$, for which we can choose $\beta'_i = \beta_j + \sigma B \overrightarrow{j} - \sigma B \overrightarrow{i}$, which is non-negative as $(\langle \sigma, j \rangle, \beta)$ is bribery stable. Note that this immediately provides all inequalities from Theorem 4, such that $(\langle \sigma, j \rangle, \beta')$ is bribery stable. The dictator payoff, however, would be $\mathsf{dpayoff}(A; \langle \sigma, i \rangle, \beta') = \sigma A \overrightarrow{i} - \beta'_i > (\sigma A \overrightarrow{j} + \sigma B \overrightarrow{j} - \sigma B \overrightarrow{i}) - \beta'_i = \sigma A \overrightarrow{j} - \beta_j = \mathsf{dpayoff}(A; \langle \sigma, j \rangle, \beta)$, which contradicts the optimality requirement of incentive equilibria. □

4. EVALUATION

In this section, we give detail of our proof-of-concept implementation and the results obtained. We randomly generated two set of benchmarks with uniformly distributed entries in the bi-matrices. One set of benchmarks uses continuous pay-off values in the range from 0 to 1 (Table 4), and the other set of benchmarks uses integer pay-off values in the range from -10 to 10 (Table 5). They contain samples of 100,000 games for each matrix form covered.

An incentive equilibrium (*IE*) of a bi-matrix game $\mathcal{G}(A, B)$ can be computed by solving the linear programming problems from Corollaries 8 and 16. The result is a simple strategy profile $(\langle \sigma, j \rangle, \beta)$, where j is a pure strategy of the subject, σ is given as a tuple of probabilities that describe the likelihood the dictator chooses her individual strategies, and β is a j-bribery vector. We implemented Algorithm 1 for computing friendly incentive equilibrium, using the LP solver [8], taken from `http://lpsolve.sourceforge.net/5.5/`. The algorithm returns a friendly *IE* in form of a strategy profile $(\langle \sigma, j \rangle, \beta)$, as well as the payoffs obtained by the subject and dictator in the friendly *IE* under the j-bribery vector returned for the given bi-matrix game. We implemented our algorithm in C and our implementation is available [**?**]. We have used GAMBIT [9] to compute the Nash equilibria(*NE*). The data size is given in terms of number of subject strategies ($\#SSt$) and the number of dictator strategies($\#DSt$).

Experimental results. We analysed the outcome of the random games along the parameters 'average optimal return value of the dictator ($DictR$)', 'average optimal return value of the subject ($SubjR$)', 'average bribery value', and the confidence interval radius for both dictator and subject return for a 95% confidence interval. We also gave the execution time for 100,000 games. The highest average execution time observed, which is obtained for friendly incentive equilibria, is 21.5 milliseconds.

We summarise the results for friendly incentive equilibria (*IE*) and leader equilibria (*LE*), respectively, in Table 4 for continuous variables in the 0 to 1 range and in Table 5 for integer variables in the -10 to 10 range. The results indicate that the bribery value falls with the number of dictator strategies and, less pronounced, with the number of subject strategies. This is not very surprising: the limit value for infinitely many strategies is 0, as there is, with limit probability 1, an entry arbitrarily close[1] to the social return 2 in the continuous case, and with the social return 20 in the integer case. For the same reason, it is not surprising that the dictator benefits from an increase in her own strategies and in the strategies of the subject alike, whereas the subject does not seem to benefit from an increase in the number of dictator strategies.

Table 6 compares incentive, leader, and Nash equilibria for dictator return and subject return, respectively, for uniformly distributed integer entries in the range from -10 to 10. We have used Gambit [9] to compute Nash equilibrium. If there is more than one *NE* in a data-set, we have considered the optimal one with the maximum pay-off for the dictator, using the subject payoff as a tie-breaker. Our results show that, for both dictator and subject, return is always higher in *IE* as compared to *LE* and even more so when compared to *NE*. Table 6 also outputs the gain for the dictator in *IE* as compared to *NE*. Its value is given by $i\_gain = (DictRet_{bribery} - DictRet_{Nash})/(Max_{dictvalue} - DictRet_{Nash})$ to describe the improvement obtained.

The data set used there is tiny, 10 samples each. This is because it is expensive to compute optimal Nash equilibria. The unsurprisingly large differences to the values from Tables 4 and 5 confirm that the values have to be read with caution, but they suffice to give an impression on the advantage obtained over Nash equilibria.

The improvement obtained (numerator) is the difference between the dictator payoff in the *IE* and the best *NE*, while the maximum improvement possible (denominator) is the difference between the

[1] ε close for an arbitrarily small, but fixed, $\varepsilon > 0$

Algorithm 1: The Algorithm outputs a (pure) friendly incentive equilibrium $(\langle \sigma, j \rangle, \beta)$, a j-bribery vector β, dictator payoff $\mathsf{dpayoff}(A; \langle \sigma, j \rangle, \beta)$ and subject payoff $\mathsf{spayoff}(B; \langle \sigma, j \rangle, \beta)$

```
Input: A bi-matrix game G(A, B)
Output: A friendly incentive equilibrium ⟨σ, j⟩, a j-bribery vector β
        (represented by the value ι = β_j), dpayoff(A; ⟨σ, j⟩, β), and
        spayoff(B; ⟨σ, j⟩, β)
max ← min{a_ij | i ≤ m, j ≤ n}
// initialise the dictator payoff to minimal entry of A
opt ← ∅ // initialise optimal pure subject strategies
for j ← 1 to n //for each pure subject strategy do
    write linear programme LP_j^{G(A,B)};
    call LPsolver();
    get objective value obj_val
    if obj_val > max then
        max ← obj_val
        opt ← {j}
    end
    if obj_val = max then
        opt ← obj ∪ {j}
    end
end
max' ← min{b_ij | i ≤ m, j ≤ n} − 1
for j ∈ opt //for pure subject strategies with IE do
    write linear programme ELP_j^{G(A,B)};
    call LPsolver();
    get objective value obj_val and solution σ
    if obj_val > max' then
        max' ← obj_val
        IE ← (⟨σ, j⟩, β)
    end
end
return IE, max', max, and ι
```

		dictator return				subject return		
#SSt	#DSt	Incentive	Leader	Nash	i_gain	Incentive	Leader	Nash
2	2	5.96	4.28	1.34	0.56	5.24	4.05	3.53
2	3	6.13	4.78	2.45	0.54	3.77	2.61	1.9
2	5	5.33	5.33	2.81	0.48	2.87	0.94	0.14
2	10	8.81	8.81	7.1	0.66	5.69	5.26	1.89
3	2	4.93	3.43	1.23	0.54	5.31	5.3	5.3
3	3	5.94	4.71	0.67	0.57	5.56	4.3	4.3
5	3	5.95	4.34	1.12	0.52	6.24	6.12	5.2
10	10	9.13	9.01	6.72	0.78	7.27	6.7	1.96

Table 6: average dictator return and subject return for same datasets with pay-offs -10 to 10 in different equilibria

#SSt	#DSt	Dict R average	Subj R average	Bribery average	conf I dict	conf I subject	ET (minutes)	Dict R average	Subj R average	conf I dict	conf I subject	ET (minutes)
2	2	0.726498	0.621347	0.031252	0.001188	0.001705	4.2175	0.698048	0.612999	0.001332	0.001774	2.7368
2	3	0.798458	0.585959	0.020400	0.000926	0.002014	4.6354	0.785525	0.577285	0.001009	0.002068	3.8177
2	5	0.868985	0.529380	0.010426	0.000642	0.002537	4.5459	0.865601	0.523492	0.000667	0.002555	4.0808
2	10	0.929643	0.434452	0.003310	0.000365	0.003249	6.4351	0.929367	0.432039	0.000367	0.003250	4.5798
3	2	0.752211	0.697569	0.043221	0.001051	0.001526	5.0878	0.710645	0.683939	0.001274	0.001623	4.5323
3	3	0.818569	0.662023	0.027266	0.000809	0.001893	6.5401	0.800711	0.647836	0.000926	0.001970	4.9813
5	3	0.856830	0.558777	0.026226	0.000646	0.003482	11.1242	0.817277	0.725141	0.000843	0.001852	8.0363
10	10	0.967437	0.184819	0.003880	0.000160	0.005209	35.8585	0.954231	0.684151	0.000207	0.003089	34.0453

Table 4: dictator return(avg), subject return(avg), bribery value(avg), confidence interval radius(dictator), confidence interval radius(subject), total execution time for 100,000 samples in incentive equilibria (left) and leader equilibria (right) with pay-offs 0 to 1

#SSt	#DSt	Dict R average	Subj R average	Bribery average	conf I dict	conf I subject	ET (minutes)	Dict R average	Subj R average	conf I dict	conf I subject	ET (minutes)
2	2	4.748595	3.030555	0.661109	0.024802	0.030519	3.5780	4.227947	2.982509	0.027691	0.030597	3.5780
2	3	6.272594	2.880576	0.432742	0.019294	0.030389	3.9091	6.041986	2.797888	0.020913	0.030337	3.3044
2	5	8.723898	2.894013	0.217244	0.013183	0.029920	4.1927	7.666887	2.831469	0.013692	0.029740	3.3426
2	10	8.982101	3.039720	0.068128	0.006878	0.029555	4.9849	8.977945	3.013178	0.006930	0.029482	4.1737
3	2	5.295267	4.623187	0.901498	0.021894	0.025617	4.9897	4.552556	4.473518	0.026218	0.025743	3.6521
3	3	6.680038	4.444846	0.575029	0.016841	0.025512	5.3623	6.356434	4.286942	0.019193	0.025418	3.9804
5	3	7.4977325	6.740712	0.548419	0.013251	0.032891	7.6684	6.728477	5.887561	0.017346	0.019969	5.7481
10	10	9.727398	7.82520	0.065963	0.002783	0.046590	37.3025	9.41755	7.587172	0.003934	0.013824	24.9537

Table 5: dictator return(avg), subject return(avg), bribery value(avg), confidence interval radius(dictator), confidence interval radius(subject), total execution time for 100,000 samples in incentive equilibria (left) and leader equilibria (right) with integer pay-offs from -10 to 10

maximal entry in the payoff matrix of the dictator and her payoff in her best *NE*. The value thus norms the dictator's gain if she pays bribery to the subject. The higher the value, the more is dictator's gain in *IE* by paying bribery as compared to *NE*. For the subject, her gain with bribery is also always higher or equal to her gain without bribery, or in *NE*. Thus, the friendly *IE* guarantees a local social optimum in the form of a socially optimal subject return.

Note that the execution time given for each data-size is the total time required for the complete data-set, i.e., 100,000 games. The execution time rises faster with an increasing number of subject strategies than with an increasing number of dictator strategies. This was to be expected, as the number of subject strategies determines the number of linear programmes need to be solved to find a friendly incentive equilibrium. Our expectation for randomly drawn examples was to find roughly a quadratic growth in the number of subject strategies and a linear growth in the number of dictator strategies. The actual growth seems to be a bit lower, but this may well be due to noise and random effects. The execution time is tiny in all instances.

Symbolic analysis of relevant classes. As discussed in the introduction, prisoners dilemma / arms race games are one standard class of problems, where our technique provides very nice results. We give a short overview on their symbolic solution. A general bi-matrix game of this class are games in the form of Table 7 that satisfy the following constraints:

- $d > b > h > f$ and $e > a > g > c$, and
- $\min\{b-g, a-g\} > \max\{d-b, h-f, e-a, g-c\}$.

		Player II 1	Player II 2
Player I	1	a, b	c, d
Player I	2	e, f	g, h

Table 7: "prisoners dilemma" pay-off matrix

It is easy to see that the only Nash and leader equilibrium is to play (2,2), with dictator return g and subject return h. In an incentive equilibrium, however, the dictator can incentivise her subject to play 1 when she pledges to play 1 herself and promises to pay him a bribery of $d-b$ for playing 1. The subject return then increases to d, while the dictator return increases to $b+a-d$.

As mentioned in the introduction, the class of battle of sexes games—these are the games satisfying $g > a > \max\{c, e\}$ and $b > h > \max\{d, f\}$—is a class of games, for which incentive equilibria provide optimal result for the dictator, namely the strategy (2,2) with bribery 0. This is, however, also a leader equilibrium.

Finally, when we consider the games where the dictator has no choice except for the selection of the bribery value. We note that an incentive equilibrium provides the social optimum, without effecting the outcome for the subject. If, e.g., the individual values for the dictator and subject are $N(0,1)$ normal distributed, the social outcome for each pair is $N(0,\sqrt{2})$ normal distributed, such that the expected dictator return is $(\sqrt{2}-1)$ times the expected subject return for all numbers of subject strategies, whereas it would be 0 for leader and Nash equilibria.

5. CONCLUSION

With incentive equilibria, we have introduced an asymmetric type of equilibria that reflects the asymmetry in the evaluation of bi-matrix games by allowing the more powerful agent to suggest solutions to her follower and by incentivising her follower's actions. It translates into more freedom for the dictator to select the strategy profiles. As observed for the prisoners' dilemma (cf. Table 1), incentivising proves to be individually better for *both* players. Naturally, the observation *incentive equilibria* beat *leader equilibria* beat *Nash equilibria* was confirmed. Our experimental results also suggest that the subject would often benefit from incentive equilibria. Incentive equilibria, therefore, provide a natural explanation for bribery in asymmetric non-zero sum games.

6. REFERENCES

[1] A. Anderson, Y. Shoham, and A. Altman. Internal implementation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 191–198. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[2] V. Conitzer and T. Sandholm. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, EC '06, pages 82–90, New York, NY, USA, 2006. ACM.

[3] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a nash equilibrium. *Commun. ACM*, 52(2):89–97, 2009.

[4] J. Friedman. *Oligopoly and the theory of games*. Advanced textbooks in economics. North-Holland Pub. Co., 1977.

[5] N. Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the 16th Annual ACM Symposium on Theory of Computing (STOC 1984), 30 April – 2 May, Washington, DC, USA*, pages 302–311. ACM Press, 1984.

[6] L. G. Khachian. A polynomial algorithm in linear programming. *Doklady Akademii Nauk SSSR*, 244:1093–1096, 1979.

[7] C. E. Lemke and J. T. Howson, Jr. Equilibrium points of bimatrix games. *Journal of the Society for Industrial & Applied Mathematics*, 12(2):413–423, 1964.

[8] K. E. M. Berkelaar and P. Notebaert. lp_solve, a Mixed Integer Linear Programming (MILP) solver. Website.

[9] R. D. Mckelvey, A. M. Mclennan, and T. L. Turocy. Gambit: Software Tools for Game Theory, http://www.gambit-project.org. Technical report, Version 13.1.0., 2013.

[10] D. Monderer and M. Tennenholtz. K-implementation. *Journal of Artificial Intelligence Research*, 21:37–62, 2004.

[11] J. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.

[12] M. J. Osborne and A. Rubinstein. *A course in game theory*. The MIT Press, Cambridge, USA, 1994. electronic edition.

[13] A. Tucker. *A two-person dilemma*. Stanford University, 1950.

[14] H. von Stackelberg. *Marktform und Gleichgewicht*. J. Springer, 1934.

[15] B. von Stengel and S. Zamir. Leadership games with convex strategy sets. *Games and Economic Behavior*, 69(2):446–457, 2010.