Extortion can outperform generosity in the iterated Prisoner's Dilemma

Bin Xu[1,2,*], Yanran Zhou[2,3], Jaimie W. Lien[4,5], Jie Zheng[5], Zhijian Wang[2]

1 College of Economics, Zhejiang Gongshang University, Economics Building, No. 18 Xuezheng Street, Xiasha, Hangzhou 310018, China

2 Experimental Social Science Laboratory, Zhejiang University, Room 312, Dongsan Building, No. 866 Yuhangtang Road, Hangzhou 310058, China

3 College of Economics, Zhejiang University, Room 326, Waijingmao Building, No. 38 Zheda Road, Hangzhou 310027, China

4 Department of Decision Sciences and Managerial Economics, The Chinese University of Hong Kong, Cheng Yu Tung Building, No. 12 Chak Cheung Street, Shatin, Hong Kong 999077, China

5 School of Economics and Management, Tsinghua University, Room 561, Weilun Building, Haidian District, Beijing 100084, China

[*]Correspondence and requests for materials should be addressed to B.X. (email: xubin211@zju.edu.cn).

Abstract

Zero-determinant (ZD) strategies, as discovered by Press and Dyson, can enforce a linear relationship between a pair of players' scores in the iterated Prisoner's Dilemma. Particularly, the Extortionate ZD strategies can enforce and exploit cooperation, providing a player with a score advantage, and consequently higher scores than those from either mutual cooperation or Generous ZD strategies. In laboratory experiments in which human subjects were paired with computer co-players, we demonstrate that both the Generous and the Extortionate ZD strategies indeed enforce a unilateral control of the reward. When the experimental setting is sufficiently long and the computerized nature of the opponent is known to human subjects, the Extortionate strategy outperforms the Generous strategy. Human subjects' cooperation rates when playing against Extortionate and Generous ZD strategies are similar after learning has occurred. More than half of Extortionate strategists finally obtain an average score higher than that from mutual cooperation.

Promoting cooperation under adverse short-term individual incentives is an important social challenge, and the iterated Prisoner's Dilemma (IPD) has been widely studied as the canonical game theoretic framework representing this issue [1–7]. In a one-shot two-person prisoner's dilemma, there are two pure strategies: cooperate (C) and defect (D). Each player receives R if they mutually cooperate; each player receives P if they mutually defect; if one player cooperates and the other defects, the defector receives T and the cooperator receives S, where $T > R > P > S$ guarantees that in this game the commonly used solution concept Nash equilibrium, is mutual defection, while $2R > T + S$ implies that mutual cooperation is actually the socially best outcome.

Ever since the computerized tournaments conducted by Axelrod [3, 4], kindness and fairness [8–10] appeared to yield the best chance to promote and sustain cooperation. A substantial body of literature suggests that reciprocity makes mutual cooperation feasible [11–16], and is a favorable strategy in an evolutionary setting [17–20].

Recently, Press and Dyson discovered a surprising class of so-called zero-determinant (ZD) strategies which allow a player to unilaterally enforce a linear relationship between his score and that of his opponent [21]. A ZD strategy is described by the probabilities of cooperation given the four possible outcomes of the previous round: $p = (p_1, p_2, p_3, p_4)$, where $p_i, i \in [1, 2, 3, 4]$ is the probability of cooperation given the previous outcomes CC, CD, DC and DD, respectively. The premise is that by implementing particular randomizations of actions in a given period which are conditional on the previous period's outcome, the ZD strategist can induce an evolutionary opponent to always choose cooperation as the optimal strategy. A subclass of ZD strategies, namely the Extortionate strategies, can further guarantee that the extortioner's own surplus exceeds the opponent's surplus by a fixed percentage. This means that the Extortionate ZD strategist can maintain cooperation and pursue self-interest at same time. In addition, an extortioner can earn a score which exceeds the best possible score from another subclass of ZD strategies, Generous ZD strategies. This new finding by Press and Dyson has stimulated many researchers to further investigate the performance of ZD strategies in various situations [22–31]. A key insight of the Press-Dyson theory is that if a human being's behavior is developed in an evolutionary manner under an opponent's ZD strategy, he will tend to become more cooperative over time, and if so, the ZD strategist can achieve his maximum possible score by exploiting this cooperative tendency of his opponent [21].

In spite of these significant developments in the literature, the empirical verification of the theory has been nontrivial. Up to now, to our knowledge, there is just one published study which tests the Press-Dyson theory in a laboratory experiment. Hilbe et al [26] provided experimental evidence on the performances of different ZD strategies played by computers against human subjects. They specified the ZD strategists to play an Extortionate strategy or a Generous strategy against human subjects in the context of IPD, in which the Extortionate strategy is predicted to earn a higher score based on the Press and Dyson's theory. However, they find that generosity is in fact the more profitable strategy, and furthermore, that the cooperation rate of human co-players against Extortionate strategies is only half of that against Generous strategies. In other words, in practice, generosity appears to be the winning strategy after all. The experimental results by Hilbe et al [26] appear potentially at odds with the Press-Dyson prediction, and have inspired other related studies [32–34]. This apparent inconsistency between theory and experimental results calls for a closer examination of the source of the discrepancy [35]. In the experiment by Hilbe et al [26], it was not made known to the human subjects that they were actually playing against a strategy executed by a computer program. Thus, in their experiment [26], the effects of two factors, the effectiveness of ZD strategies themselves and human subjects' perception of the other player, may be potentially responsible for the result they obtain. Another factor that may have led to the inconsistency between the Press-Dyson theory and the experimental result in Hilbe et al [26], is insufficient learning opportunities for the human subjects in their 60-round experiments. Until now, there has been no empirical evidence to illustrate the reality of the Press-Dyson theory.

Our study attempts to establish an understanding and empirical validation of the Press-Dyson result. We test the performance of the ZD strategies in a laboratory experiment of IPD which modifies the

design of Hilbe et al [26], by introducing two factors: the knowledge of the opponent's computerized nature and the length of interaction. Each of these factors vary by two conditions, awareness (A) versus unawareness (U) by human subjects of the computerized nature of the opponent for the knowledge of the opponent's nature, and long rounds (500) versus short rounds (60) for length of interaction. Among these conditions, the treatments in which human subjects were both aware of the opponent's computerized nature as well as afforded rich learning experience over long-term interactions, was the ideal condition where the predictions of the Press-Dyson model could be realized. The remaining three settings further confirm the conditions (knowledge, length or both) that are necessary to produce the result. Our experiments show that under long rounds (500) and awareness (A) conditions, the Extortionate ZD strategy indeed outperforms the Generous ZD strategy by inducing a higher cooperation rate among human co-players than under conditions of short rounds (60) or unawareness (U), and thus obtains a higher score by exploiting this tendency. Furthermore, under the awareness condition, after learning has occurred, the number of the Extortionate strategists who obtain an average score higher than that from mutual cooperation is significantly greater than the number who obtain an average score less than or equal to the score from mutual cooperation. To our knowledge, this is the first experimental evidence which supports the predictions of the Press-Dyson theory.

Results

Experiment

The IPD payoff matrix used in our experiment is shown in Fig 1, which is identical to the one used in Refs [21, 26]. Both players receive 3 if they mutually cooperate, both players receive 1 if they mutually defect, the defector receives 5 and the cooperator receives 0 if one player cooperates and the other defects.

We implemented a set of treatments in which human subjects faced the computerized Extortionate ZD strategy (E) or Generous ZD strategy (G), which further varied by the two additional aforementioned features: awareness (A) versus unawareness (U) by human subjects of the computerized nature of the opponent, and length of interaction (500 rounds versus 60 rounds). Table 1 summarizes the experimental design. For E, the four conditional cooperation probabilities are $(p_1, p_2, p_3, p_4) = (0.692, 0.000, 0.538, 0.000)$, and for G the probabilities are $(p_1, p_2, p_3, p_4) = (1.000, 0.182, 1.000, 0.364)$, where $p_i, i \in [1, 2, 3, 4]$ is the probability of cooperation given the previous outcome CC, CD, DC and DD, respectively. In addition, E defects in the first round and G cooperates in the first round. These are the same conditional cooperation probabilities as the Strong Extortionate ZD strategy and Strong Generous ZD strategy in Hilbe et al [26]. Theoretically, as proved by Press and Dyson [21] and specified by Hilbe et al [26], for Extortion, the ZD strategist's score s_z^e and his human co-player's score s_h^e satisfy $\frac{s_z^e - 1}{s_h^e - 1} = 3$, the maximum scores for the ZD strategist and the human co-player are 3.727 and 1.907, respectively. For Generosity, the ZD strategist's score s_z^g and his human co-player's score s_h^g satisfy $\frac{s_z^g - 3}{s_h^g - 3} = 3$, the maximum scores for both the ZD strategist and the human co-player are 3 (see Fig 4). Therefore, if both types of ZD strategies can in fact induce human co-players' cooperative behavior, the Extortionate ZD strategists can achieve higher scores than the Generous ZD strategists.

Altogether 256 graduate and undergraduate students participated in the experiment with each treatment consisting of 32 participants. For further details on the implementation of the experiment, see the **Methods** section.

Score performance of Extortionate and Generous ZD strategies

Fig 2 shows the resulting average scores over each treatment. To begin, we note the higher score achieved by the Extortionate ZD strategist compared to the Generous ZD strategist in the 500A treatments (Mann-Whitney test, $n_E = n_G = 32$, $z = 4.196$, $p = 0.000$), in which human subjects were aware of

the computerized nature of the opponent. On the other hand, the corresponding 60-round awareness treatments (60A) yielded statistically identical average scores between the Extortionate and Generous strategists. Turning to the treatments in which human subjects were unaware of the computerized nature of the opponent, we note that the result of Hilbe et al [26] which demonstrates the favorability of the Generous strategy, is replicated for the 60-round treatments (60U) and is even further enhanced in the 500-round treatments (500U). This indicates that neither awareness alone, nor long-term interaction alone, can deliver the prediction of Press-Dyson. Rather, it is the combination of these features which creates the environment favorable to the Extortionate ZD strategist.

In Treatment 500A, where support for the theory is clear, on average, the Extortionate strategists earn 2.943 ± 0.511 (mean $\pm$ s.d.) scores per round and Generous strategists earn 2.263 ± 0.788 (mean $\pm$ s.d.) scores per round. The average scores for Extortionate strategists are 30% higher than Generous strategists. Furthermore, half of the extortioners even earned scores higher than 3 per round on average, which is the score associated with constant mutual cooperation. By contrast, not a single Generous strategist earned scores exceeding 3.

In the later rounds of the 500AE treatment, the number of the Extortionate strategists obtaining an average score exceeding 3 is significantly greater than those obtaining scores less than or equal to 3 (for further details, see **Supplementary Fig 1 and 3**). In the last 60 rounds of 500AE, the average score of Extortionate strategists rises to 3.127, which is significantly higher than 3 ($z = 2.301$, $p = 0.021$, Wilcoxon signed-rank test), while such a result was not observed in other conditions (for more details, see **Supplementary Fig 2** and **Supplementary Table 5, 6, 7, 8, 9 and 10**). These indicate that Extortion can indeed be a winning strategy under the conditions of awareness and long interaction length.

Effectiveness of promoting cooperation under Extortionate and Generous ZD strategies

On average, in the 500-round awareness treatments (500A), the cooperation rate of human co-players is 0.684 ± 0.198 (mean $\pm$ s.d.) in the Extortionate treatment and 0.645 ± 0.349 (mean $\pm$ s.d.) in the Generous treatment. There is no significant difference between the two treatments (Mann-Whitney test, $n_E = n_G = 32$, $z = 0.537$, $p = 0.591$). By contrast, there is a pronounced and statistically significant difference in human cooperation rates in favor of the Generous strategy for the other treatments. The success of the Generous strategy in the unawareness treatments (500U and 60U) supports the findings in Hilbe et al [26], while the result in the 60-round awareness treatment (60A) may be attributed to the limited learning experiences of human co-players. For details on average cooperation rates, see **Supplementary Table 1**. For statistical results, see **Supplementary Table 11**. For individual level data, see **Supplementary Table 12 and 13**.

Fig 3 shows the dynamic cooperation rate of human subjects for each treatment. In accordance with the similarity in cooperation rates across the 500-round awareness treatments (500AE and 500AG), the two ZD strategies also generate similar upward dynamic patterns in these treatments (for Spearmans rank correlation, see Fig 3). The cooperation rate starts from a relatively lower point then follows an increasing path. For the Extortionate treatment 500AE, compared with the cooperation rate 0.563 ± 0.221 (mean $\pm$ s.d.) in the first 100 rounds, the cooperation rate 0.757 ± 0.220 (mean $\pm$ s.d.) in the last 100 rounds is significantly higher (Wilcoxon signed-rank test, $n = 32$, $z = 4.207$, $p = 0.000$). For the Generous treatment 500AG, compared with the cooperation rate 0.513 ± 0.327 (mean $\pm$ s.d.) in the first 100 rounds, the cooperation rate 0.716 ± 0.386 (mean $\pm$ s.d.) in the last 100 rounds is also significantly higher (Wilcoxon signed-rank test, $n = 32$, $z = 3.314$, $p = 0.000$). By contrast, in the 500U, 60A and 60U treatments, all Generous ZD treatments displayed an upward trend in human cooperation, while the Extortionate ZD treatments did not display this significant upward trend (for Spearmans rank correlation, see each subfigure in Fig 3). The upward trend in human cooperation rates when playing against an Extortionate ZD strategist, observed solely in the 500-round awareness treatment (500AE),

can account for our main finding that Extortion outperforming Generosity in terms of scores, was realized only in the long length awareness treatment (500A) and not in the other treatments.

The score relationship between ZD strategies and human players

The Extortionate ZD strategist earns higher scores than their human co-players. On the contrary, the Generous ZD strategists earn lower scores than their human co-players (as an illustration, see Fig 2, for statistical results, see **Supplementary Table 14**). The relationship of scores between ZD strategists and human co-players follows the linear relationship prediction in Press-Dyson, as shown in Fig 4. These results indicate that ZD strategies can indeed unilaterally enforce a linear relationship between human players' scores and their own scores.

Discussion

The ZD strategies discovered by Press and Dyson allow a player in the IPD to unilaterally invoke a linear relationship between his own payoff and that of his opponent, by fixing a mixed strategy which conditions on the previous round's outcome. A remarkable feature of this class of strategies is that a ZD strategist can create a learning environment in which it is in fact optimal for his opponent to choose the cooperative action in each round. Two types of ZD strategies have attracted particular interest in the literature: so-called Extortionate ZD strategy, in which the ZD strategist sometimes exploits the opponent's cooperation tendencies and never takes the initiative on cooperation, and the so-called Generous ZD strategy, in which the opponent's cooperation tendency is never exploited, and sometimes cooperation is initiated by the ZD strategist. While both of these ZD strategies theoretically yield total cooperation from the opponent, the Extortionate ZD strategy produces a clear score advantage for the ZD strategist [21, 22, 24, 26].

The prediction of the Press-Dyson theory gains support in our experiments. Our experiments showed that the Extortionate strategy earned a higher score than the Generous strategy, while continuously promoting the human co-player's cooperative behavior at the same level of effectiveness as the Generous strategy, when human subjects were aware of the nature of the opponent and had a sufficiently long experimental experience to learn. Under these conditions, after learning, the number of the successful Extortionate strategists obtaining a higher score than mutual cooperation is significantly higher than those who obtain a score less than or equal to the score from mutual cooperation. We also found that the ZD strategies were successful in unilaterally enforcing a linear relationship between human players' scores and their own scores. These findings match the Press-Dyson theory well.

From an economics point of view, human subjects have no reason to reject a strategy which will provide a higher payoff unless there are some additional factors influencing their behaviors, assuming they have correctly found the optimal strategy. The fact that the cooperation rate reaches such a high level in the later stage of the 500AE treatment reflects individuals' pursuit of their own economic interests as well as their cognitive ability. However, human subjects' perception of the other player may be potentially responsible for the experimental result, for example, through fairness [36–39] or conditional cooperation [40–42]. Hilbe et al [26] show that human subjects tended to be more cooperative when their opponent cooperated in the previous round. Human subjects in their study would defect when they played against the Extortionate strategies, because the Extortioner often defected. This conditional behavior may even inhibit the process of discovering the profitable strategy. Therefore, the cooperation rate had difficulty increasing even in 500UE treatment. However, human subjects may be substantially less likely to demand fairness and reciprocity from a machine, compared to when they perceive that they are playing with another human being [36–39]. Through awareness, the performance of the strategy can be independent of potentially more complicated social influences such as human players' attitudes and intentions towards other players. The human subjects' cooperation rate given that the ZD

strategist previously defected, is higher in the awareness condition than in the unawareness condition (see **Supplementary Table 15**). Consequently, the cooperation rates in the 500AE treatment reached a higher level. This indicates that awareness is one of the crucial factors for the ZD theory to hold in our experiment.

Another contributing factor could be the length of the experiment. As in Hilbe et al [26], our results also show that the extortionate ZD strategy cannot outperform the generous ZD strategy under a 60-round setting. This may be due to insufficient learning opportunities for the human subjects. A strategy involving uncertainty (and stochastic actions) takes a longer time to reveal itself than a pure strategy. The 60-round sessions implemented in the experiments may not be sufficient for humans to learn the probabilistic details of a ZD strategy. Lengthier sessions which can accommodate learning are often more ideal when subjects are facing probabilistic environments [43–49]. The ZD strategy is a probabilistic one, human subjects may need ample time in an evolutionary-like setting in order to learn the ZD strategy of his opponent, contemplate and formulate his own optimal strategy (see Chapter 1 in [50]). We observed that the human subjects' cooperation rates and the number of winning Extortionists steadily increased in the 500AE treatment, reflecting the length of interaction effect, while this pattern did not appear in the other treatments.

In this research, we clarify that both factors, the knowledge of the opponent and the length of the interaction, have potential to facilitate the predictions of Press-Dyson. However, we note that there may very well be other conditions in which the prediction of ZD theory can also be observed. Furthermore, our research leaves several related questions open, such as, whether playing against a computer is the same as playing against a human who uses a fixed strategy? Will knowing that the opponent's strategy is a fixed one shorten the time needed to find the profitable strategy? An additional observation from our experiment is that even with a lengthy period of time to learn, some players' cooperative behavior remained unimproved. This may be due to either cognitive limitations of human subjects in realizing the profitability of strategy C, or possible attitudes towards the game such that they deliberately refused to choose C. This points to some additional directions for future work to further check the robustness of ZD strategies' performances in various contexts, such as heterogeneous education levels of human subjects, across genders, or stochastic errors in the game [31]. We note also that in the first 60 rounds, the ZD strategists' scores and the human cooperation rates are different in the long run and short run treatments (see **Supplementary Table 6**). This might be due to the effect of the 'shadow of future' [15, 51] or a 'learning effect' [46, 50], which can be further studied in the future.

In real life, a Prisoner's Dilemma type setting can exist in many situations. For example, the interaction between a firm and its workers. Suppose that payoffs of the game are such that it is a dominant action in a one-shot interaction for workers to shirk instead of work diligently on the job, and it is a dominant action in a one-shot interaction for the firm to keep the worker's salary the same as before instead of giving a raise. The Press-Dyson theory implies that there are conditional randomized policies that the firm can implement which induces the worker to work diligently, some of which exploit the workers cooperative tendencies. Similar scenarios may exist between an online merchant and its consumers, or a developing country government and its citizens. As artificial intelligence products began to enter our daily life (e.g. Siri, Robot, Self-driving cars, etc), as an algorithm, the ZD strategies might also be applied by firms to maximize their profits. The experimental setting of unawareness, as in Hilbe et al [26] and our experiments, provides an ideal situation to test the interaction between individuals. However, the game can also be understood as the strategic interaction between countries, firms, or other institutions, or between institutions and human individuals. For an individual, the prospect of interacting with a machine-like opponent is not uncommon. For example, in the competition between a leader firm and a follower firm [52] in a duopoly market, the leader firm can take a rigid policy towards the opponent, similar to a machine. Between the interaction of a state institution and a private entity, the institution may act like a machine [53]. In these two cases, the machine-like player's stature (leader firm or state institution) informs their opponents (follower firm or private sector). In this way, the ZD strategy also

can be understood as an institutional (machine-like) strategy.

Although in the Prisoner's Dilemma, long-run relationships are typically regarded as a positive influence in preventing opportunistic behavior while supporting cooperative behavior [15, 16, 51, 54], now we can clearly see that this relationship can also be exploited by the Extortionate ZD strategist. Furthermore, since the extortionate strategy can obtain a higher payoff, as a potential mechanism, it will tend to be evolutionarily favored and reproduced. Our experimental results confirm that those policies designed with a ZD theory in mind will have an evolutionary advantage by extorting rather than behaving generously under some conditions (for example, long run interactions and awareness of a computer-like opponent). Against such a strategy, human subjects in this type of strategic setting cannot avoid being extorted unless they hold an outside option to exit the game altogether. Thus, these new strategies discovered by Press-Dyson raise serious challenges for the possibility of generous behavior on the part of payoff maximizing policies towards actual human decision-makers.

Methods

Data source and experimental setting

The data was generated from our laboratory experiments which were conducted at The Experimental Social Science Laboratory of Zhejiang University. The experiments were implemented in accordance with standard social science experiment ethical guidelines and regulations.

Before the formal experiment, the subjects practiced with a matching pennies game against a computer to get acquainted with the laboratory setting. They were then assigned a set of materials including an instruction manual, an informed consent form and a recording chart for their use, and they played the game in a small isolated room with a computer. Oral instructions were also given. Subjects made decisions by clicking the option C or D on the screen. No communication was allowed during the experiment, and the subjects were asked to put their mobile phones in mute and sealed in an envelope until the end of the session. Considering the complexity of the ZD strategies, we provided human subjects with paper and pen to record their decision choices and scores round by round. In the awareness treatments, the human subjects were told that they would play a game with a fixed computer program. On the contrary, in the unawareness treatments, the human subjects were only told that they would play a game with a fixed opponent, which is similar to the implementation by Hilbe et al [26]. The instruction manual of 500A is provided as an example in **Supplementary Note 2**. To obtain the desired number of rounds of data (500 or 60) for each subject while avoiding end-of-game effects [55], human subjects were informed that the game would end with probability 0.1 in each round after the desired number of rounds (500 or 60) is reached.

During the experiments, the player earned scores according to the payoff matrix (see Fig 1) and their choices. After the experiment, the sum of scores were converted to cash according to an exchange rate and paid to the subjects. For more details, see **Supplementary Note 1, 2, and 3**.

Statistical methods

Throughout the paper, we used the Mann-Whitney test for the comparison between treatments, the Wilcoxon signed-rank test for the comparison within a treatment, and the Spearmans rank correlation test for detection of trends. In addition, we used the binomial probability test for comparison of the number of the Extortionate strategists who obtained an average score higher than the score from mutual cooperation, and the number of the Extortionate strategists who obtained an average score less than or equal to the score from mutual cooperation.

References

1. Rapoport, A. & Chammah, A. M *Prisoner's Dilemma: A Study in Conflict and Cooperation* (University of Michigan Press, 1965).

2. Trivers, R. L. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57 (1971).

3. Axelrod, R. M. & Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390–1396 (1981).

4. Axelrod, R. M. *The Evolution of Cooperation* (Basic Books, 1984).

5. Fudenberg, D. & Maskin, E. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* **54**, 533–554 (1986).

6. Boyd, R. & Lorberbaum, J. P. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature* **327**, 58–59 (1987).

7. Kendall, G., Yao, X. & Chong, S. Y. *The Iterated Prisoners' Dilemma: 20 Years On* (World Scientific Publishing Co., Inc., 2007).

8. Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Q. J. Econ.* **114**, 817–868 (1999).

9. Brosnan, S. F. & De Waal, F. B. Monkeys reject unequal pay. *Nature* **425**, 297–299 (2003).

10. Oechssler, J. Finitely repeated games with social preferences. *Exp. Econ.* **16**, 222–231 (2013).

11. Friedman, J. W. A non-cooperative equilibrium for supergames. *Rev. Econ. Stud.* **38**, 1–12 (1971).

12. Binmore, K. G. & Samuelson, L. Evolutionary stability in repeated games played by finite automata. *J. Econ. Theory* **57**, 278–305 (1992).

13. Nowak, M. A., Sigmund, K. *et al.* A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature* **364**, 56–58 (1993).

14. Wedekind, C. & Milinski, M. Human cooperation in the simultaneous and the alternating prisoner's dilemma: Pavlov versus generous tit-for-tat. *Proc. Natl. Acad. Sci. USA* **93**, 2686–2689 (1996).

15. Duffy, J. & Ochs, J. Cooperative behavior and the frequency of social interaction. *Games Econom. Behav.* **66**, 785–812 (2009).

16. Dal Bó, P. & Fréchette, G. R. The evolution of cooperation in infinitely repeated games: Experimental evidence. *Am. Econ. Rev.* **101**, 411–429 (2011).

17. Molander, P. The optimal level of generosity in a selfish, uncertain environment. *J. Conflict Resol.* **29**, 611–618 (1985).

18. Nowak, M. A. & Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **355**, 250–253 (1992).

19. Stephens, D. W., McLinn, C. M. & Stevens, J. R. Discounting and reciprocity in an iterated prisoner's dilemma. *Science* **298**, 2216–2218 (2002).

20. Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).

21. Press, W. H. & Dyson, F. J. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl. Acad. Sci. USA* **109**, 10409–10413 (2012).

22. Stewart, A. J. & Plotkin, J. B. Extortion and cooperation in the prisoners dilemma. *Proc. Natl. Acad. Sci. USA* **109**, 10134–10135 (2012).

23. Hilbe, C., Nowak, M. A. & Sigmund, K. Evolution of extortion in iterated prisoners dilemma games. *Proc. Natl. Acad. Sci. USA* **110**, 6913–6918 (2013).

24. Stewart, A. J. & Plotkin, J. B. From extortion to generosity, evolution in the iterated prisoners dilemma. *Proc. Natl. Acad. Sci. USA* **110**, 15348–15353 (2013).

25. Stewart, A. J. & Plotkin, J. B. Collapse of cooperation in evolving games. *Proc. Natl. Acad. Sci. USA* **111**, 17558–17563 (2014).

26. Hilbe, C., Röhl, T. & Milinski, M. Extortion subdues human players but is finally punished in the prisoner's dilemma. *Nat. Commun.* **5**, 3976 (2014).

27. Hilbe, C., Wu, B., Traulsen, A. & Nowak, M. A. Cooperation and control in multiplayer social dilemmas. *Proc. Natl. Acad. Sci. USA* **111**, 16425–16430 (2014).

28. Chen, J. & Zinger, A. The robustness of zero-determinant strategies in iterated prisoner's dilemma games. *J. Theor. Biol.* **357**, 46–54 (2014).

29. Szolnoki, A. & Perc, M. Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations. *Sci. Rep.* **4**, 5496 (2014).

30. Szolnoki, A. & Perc, M. Evolution of extortion in structured populations. *Phys. Rev. E* **89**, 022804 (2014).

31. Hao, D., Rong, Z. & Zhou, T. Extortion under uncertainty: Zero-determinant strategies in noisy games. *Phys. Rev. E* **91**, 052803 (2015).

32. Bi, Z. & Zhou, H.-J. Optimal cooperation-trap strategies for the iterated rock-paper-scissors game. *PloS One* **9**, e111278 (2014).

33. Bruggeman, J. Deceit and self-deception. In *Sociological Forum*, vol. 30, 243–248 (Wiley Online Library, 2015).

34. Hilbe, C., Traulsen, A. & Sigmund, K. Partners or rivals? strategies for the iterated prisoner's dilemma. *Games Econom. Behav.* **92**, 41–52 (2015).

35. Levine, D. K. & Zheng, J. The relationship between economic theory and experiments. In *Handbook of Experimental Economic Methodology* (Oxford University Press, 2015).

36. Blount, S. When social outcomes aren't fair: The effect of causal attributions on preferences. *Organ. Behav. Hum. Dec.* **63**, 131–144 (1995).

37. Harrison, G. W. & McCabe, K. A. Expectations and fairness in a simple bargaining experiment. *Int. J. Game Theory* **25**, 303–327 (1996).

38. Johnson, E. J., Camerer, C., Sen, S. & Rymon, T. Detecting failures of backward induction: Monitoring information search in sequential bargaining. *J. Econ. Theory* **104**, 16–47 (2002).

39. Gintis, H. Solving the puzzle of prosociality. *Ration. Soc.* **15**, 155–187 (2003).

40. Fischbacher, U., Gächter, S. & Fehr, E. Are people conditionally cooperative? evidence from a public goods experiment. *Econ. Lett.* **71**, 397–404 (2001).

41. Fischbacher, U. & Gächter, S. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *Am. Econ. Rev.* **100**, 541–556 (2010).

42. Chaudhuri, A. & Paichayontvijit, T. Conditional cooperation and voluntary contributions to a public good. *Econ. Bull.* **3**, 1–14 (2006).

43. Selten, R. & Chmura, T. Stationary concepts for experimental 2×2-games. *Am. Econ. Rev.* **98**, 938–966 (2008).

44. Crawford, V. P. Adaptive dynamics in coordination games. *Econometrica* **63**, 103–143 (1995).

45. Fudenberg, D. & Levine, D. K. Steady state learning and nash equilibrium. *Econometrica* **61**, 547–573 (1993).

46. Fudenberg, D. & Levine, D. K. *The Theory of Learning in Games* (The MIT Press, 1998).

47. Erev, I., Roth, A. E., Slonim, R. L. & Barron, G. Learning and equilibrium as useful approximations: Accuracy of prediction on randomly selected constant sum games. *Econ. Theory* **33**, 29–51 (2007).

48. Xu, B., Zhou, H.-J. & Wang, Z. Cycle frequency in standard rock–paper–scissors games: Evidence from experimental economics. *Phys. A* **392**, 4997–5005 (2013).

49. Wang, Z., Xu, B. & Zhou, H.-J. Social cycling and conditional responses in the rock-paper-scissors game. *Sci. Rep.* **4**, 5830 (2014).

50. Camerer, C. *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton University Press, 2003).

51. Dal Bó, P. Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *Am. Econ. Rev.* **95**, 1591–1604 (2005).

52. Simaan, M. & Cruz Jr, J. B. On the stackelberg strategy in nonzero-sum games. *J. Optim. Theory Appl.* **11**, 533–555 (1973).

53. Winden, F. A. *On the Interaction between State and Private Sector* (North-Holland Publishing Company, 1983).

54. Kandori, M. Repeated games. In *New Palgrave Dictionary of Economics* (Palgrave Macmillan, 2008).

55. Selten, R. & Stoecker, R. End behavior in sequences of finite prisoner's dilemma supergames a learning theory approach. *J. Econ. Behav. Organ.* **7**, 47–70 (1986).

	C	D
C	3 3	0 5
D	5 0	1 1

Figure 1. Payoff matrix. If both players cooperate (C), each player receives 3, if one cooperates and one defects (D), cooperator receives 0 and defector receives 5, if both defect, each player receives 1.
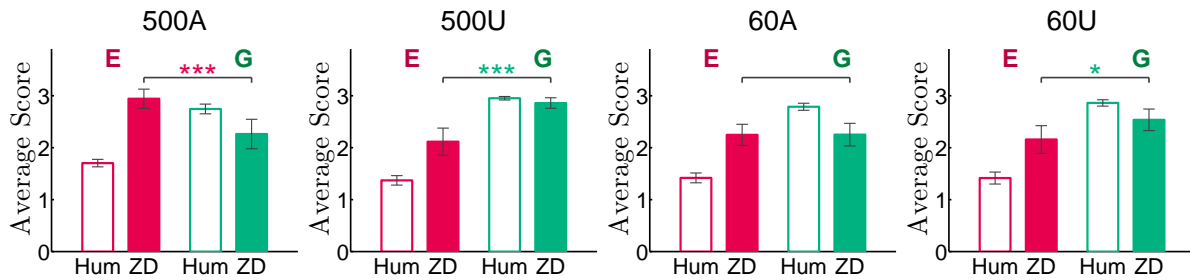
Figure 2. Average scores for human subjects and the ZD strategies for all treatments.
Each graph shows the average score earned by human subjects (unfilled bars) and ZD strategists (filled bars), for the Extortionate (red) and Generous (green) treatments, respectively. The scores for the Extortionate strategy are significantly higher than the Generous strategy in treatment 500A, but not in the other treatments. The scores of the Generous strategy are higher than the Extortionate strategy in both of the unawareness treatments (500U and 60U). There is no significant difference in scores between Extortionate strategy and Generous strategy in the 60A treatment. Three stars indicate $p < 0.001$ (specifically, $p = 0.000$ for the 500A treatment and $p = 0.000$ for the 500U treatment), and one star indicates $p < 0.05$ (specifically, $p = 0.019$ for the 60U treatment). The error bars indicate the 95% confidence interval. For details on average scores and statistical results, see **Supplementary Table 1 and 2**, respectively. For individual level data, see **Supplementary Table 3 and 4**.
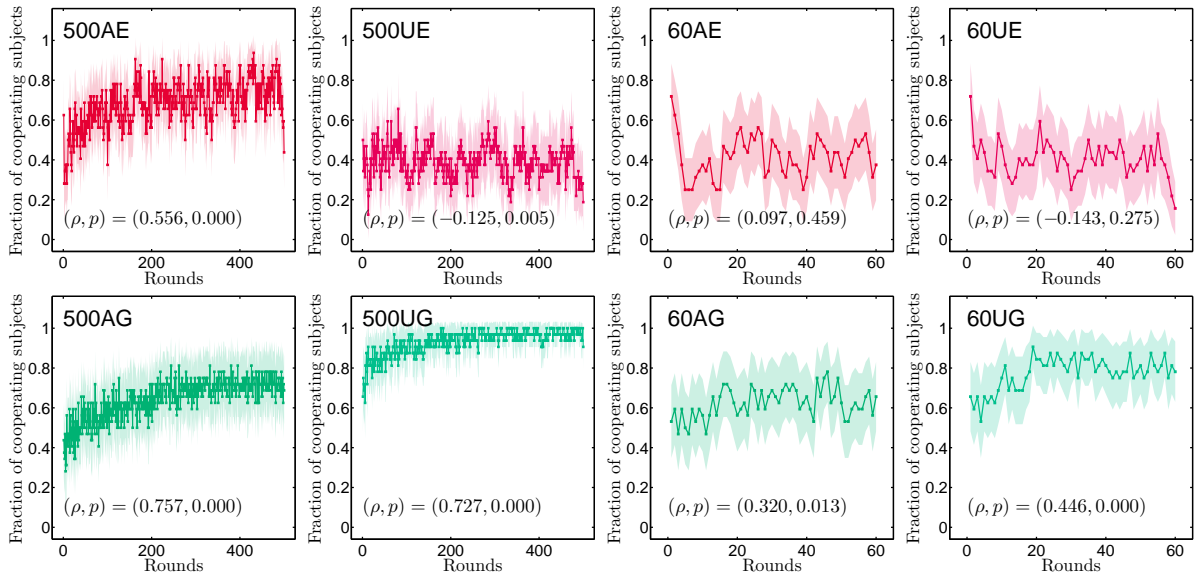
Figure 3. Human cooperation rates over the course of the game. The graph shows the fraction of cooperating human subjects for each round for Extortionate treatment (top row) and Generous treatment (bottom row). Dots represent the average human cooperation rate at a particular round within a treatment, with the shaded areas depicting the 95% confidence interval. The (color) line indicates the time trend. Significant rising trends of cooperation behavior appear in both 500AE and 500AG treatments. ((ρ, p) in each subfigure indicates the Spearman's rank correlation coefficient and the corresponding p-value). Significant positive time trends in cooperation were also found in the 500UG, 60AG, and 60UG treatments, while there was no significant time trend in the 60AE and 60UE treatments. There was a significantly negative time trend in cooperation in the 500UE treatment. **Supplementary Fig 4** also provides the smoothed average time trends. In addition, **Supplementary Fig 5 and 6** provide the smoothed time trend of individual human subjects.
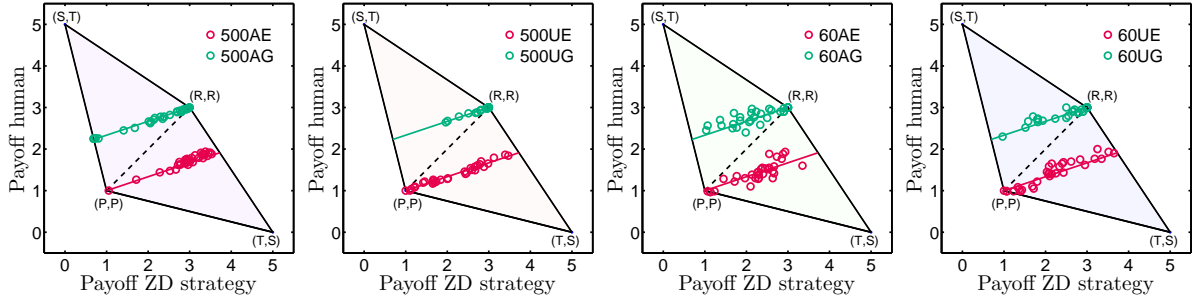
Figure 4. Experimental scores and theoretical prediction. The shaded area depicts the space of possible scores for the ZD strategy implemented by the ZD strategist (x axis) and the human co-player (y axis). The red (green) line corresponds to the theoretical prediction of the expected scores of extortion (generosity), respectively. For Extortion, the ZD strategist's score s_z^e and his human co-player's score s_h^e satisfy $\frac{s_z^e - 1}{s_h^e - 1} = 3$ (red line), and the maximum scores for the ZD strategist and the human co-player are 3.727 and 1.907, respectively, while the minimum scores for both the ZD strategist and the human co-player are 1. For Generosity, the ZD strategist's score s_z^g and his human co-player's score s_h^g satisfy $\frac{s_z^g - 3}{s_h^g - 3} = 3$ (green line), and the maximum scores for both the ZD strategist and the human co-player are 3, while the minimum scores for the ZD strategist and human co-player are 0.692 and 2.232, respectively. The open red (green) circles indicate the outcome of the Extortionate (Generous) treatment, respectively. Each circle indicates a pair of scores, the horizontal axis representing score of the ZD strategist and the vertical axis representing the human co-player's score. Corresponding to each line, there are 32 player pairs (circles) for each treatment (for individual level scores, see **Supplementary Table 3 and 4**).

Table 1. Experimental design

Treatment	Number of Rounds	Information	ZD Strategy	Number of Human Subjects
500AE	500	Awareness	Extortion	32
500AG	500	Awareness	Generosity	32
500UE	500	Unawareness	Extortion	32
500UG	500	Unawareness	Generosity	32
60AE	60	Awareness	Extortion	32
60AG	60	Awareness	Generosity	32
60UE	60	Unawareness	Extortion	32
60UG	60	Unawareness	Generosity	32

Acknowledgments

Author Contributions

Z.W. and B.X. originally designed the main experimental treatments and J.L., J.Z. and Y.Z. joined later to further develop the contribution and experimental design; B.X. and Y.Z. conducted the experiment; B.X. and Z.W. analyzed the data; J.Z., J.L. and B.X. prepared the manuscript; J.L., J.Z., Z.W, Y.Z. and B.X. prepared the supplementary material.

Competing Financial Interests

The authors declare no competing financial interests.