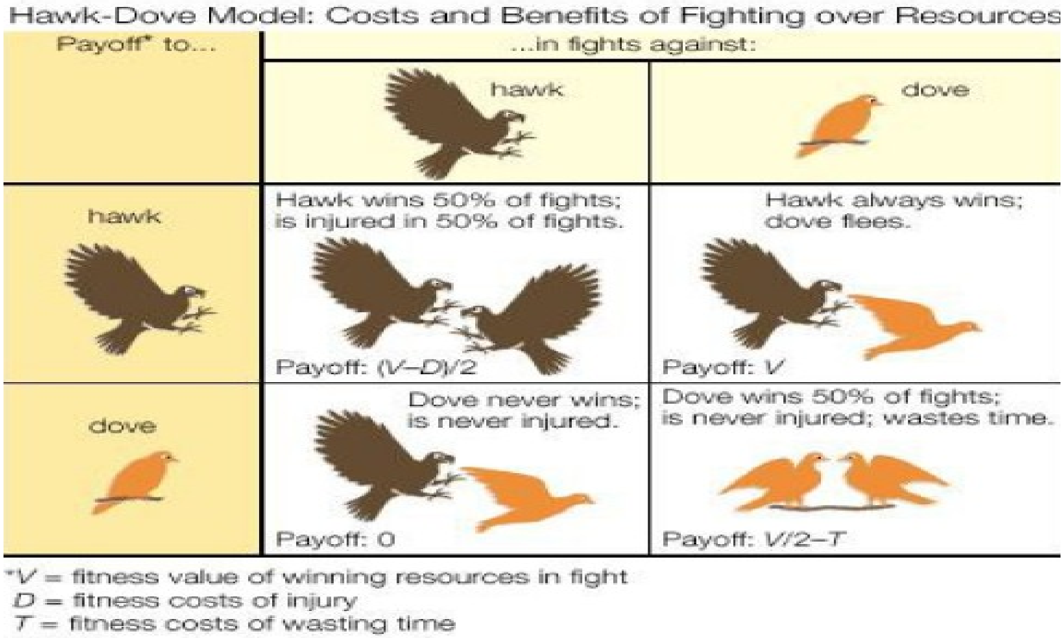
r
/
n

Lecture 4 : Evolutionary Stable Strategies (ESS), Game Theory, and ZD Strategy that rules them all

Evolutionary game theory examines how strategies survive over time subject to invasions of other strategies. Equilibrium solutions are often mixed strategies where the key is how a strategy does against itself.

I) HAWK-DOVE IS THE REPRESENTATIVE ESS GAME



IJ = what I gets from meeting J

HH = reward to H encounter with H; DD = reward to D encounter with D;
DH = what D gets from interaction with H; HD = result to H of HD

With (1-p) Hs and p Ds, profits are :W(H) = (1-p) HH+ p HD and W(D) = (1-p) DH + p DD

Strategy H is an ESS if W(H)>W(D) **when H dominates (p~0)**. This requires:

1.W(H) > W(D) --> HH> DH – ie H does better than D interacting with H: **or**

2. If H and D do same against H, HH= DH then H does better against D than D does against D: **HD > DD**

Strategy D is an ESS if W(D) >W(H) **when D dominates (large p)**. In a population largely of Ds, this requires:

1.D gets more interacting with itself DD >HD or

2.If DD =HD then D does better against H than H does: **DH > HH**

ESS refines Nash equilibrium by adding the 2nd criterion. All ESS are Nash equilibrium but not all Nash are ESS.

Key to HD is that all H or D *invite* invasion since invaders do better against dominant group:**HD > DD > DH > HH.** Compare to PD where DC>CC>DD>CD. Hawk-Dove lowest outcome is when two hawks meet. In PD the two defects do better than CD.[1] PD, max output for C is with CC but in Hawk-Dove, max output for H is with HD.

Here is representative game matrix

	H	D
H	(V-C)/2	V
D	0	V/2

where C= cost of fighting and V the reward you fight over
V/C is the ratio of rewards to cost of fighting

And Matrix when V=4 and C=6.

	H	**D**
H	**-1**	**4**
D	**0**	**2**

Note HH – DH = -1 (better to invade H as Dove while HD – DD = 2 (better to invade D as Hawk)

Another interpretation HD is a "chicken" car crash: if the other guy backs off, you win (HD) , worst is if you both fight and crash (HH), if both chicken you do better than if you are only chicken (DD>DH). Note P/V = 2/3

1There are other game matrices: HD> HH> DD> DH is thuggery. It always pays to be a hawk. Since HH> DH, V> C, it pays to battle against another hawk. But if you alternated H and D you would have bigger total amount since HD+DH = 4 > HH +HH = 2

	H	D
H	1	4
D	0	0.5

The ESS is an equilibrium concept for games of **pairwise competition.** If the population does the ESS, mutants -- alternative strategies -- cannot easily invade. ESS generalizes Nash equilibrium from requiring a strategy do better against itself than against any other strategy X to what happens when Nash and X give equal outcomes and replicator dynamics (aka a "computer tournament") change their population dependent on their profitability.

1) There is an ecology or population of strategies. ECOLOGY IS IMPORTANT
2) Higher scoring strategies grow over time. Do better (worse) than average, you increase (decrease)your share of population. The more you beat average, the more you are into next generation. DYNAMIC RULE IS IMPORTANT.

Some games have pure ESS – you drive on right or left (coordination game). Some have mixed strategy, where you combine strategies: do TFT p% of the time and Pavlov 1-p%.

HD has a **MIXED STRATEGY**,which can reflect:

1) Heterogeneity in the population: r percent of the population always does H or D
2) Individual variation in strategies: everyone does H r% of time (or some folk do H r% and others less)

This difference matters in some problems. If agents meet other agents and there is a reputation effect, reputations need heterogeneity in the population. Individual variation destroys reputations.

Analysis of HD: p% of the population are doves. You choose to enter as D or H

Start with All D world, p = 1 $\quad$ $W(H) = (1-p)\,HH + pHD = HD = 4$
$W(D) = p\,DH + (1-p)\,DD = DH = 2$ So invade as Hawk.

Start with all H world , so p =0 $\quad$ $W(H) = (1-p)HH + p\,HD = HH = -1$
$W(D) = p\,DH + (1-p)\,DD = DD = 0$ So invade as Dove

Since in all H world, D scores more and in all D world H scores more, the ESS might be **a mixed strategy**. So look for a p that equates W(H) and W(D):

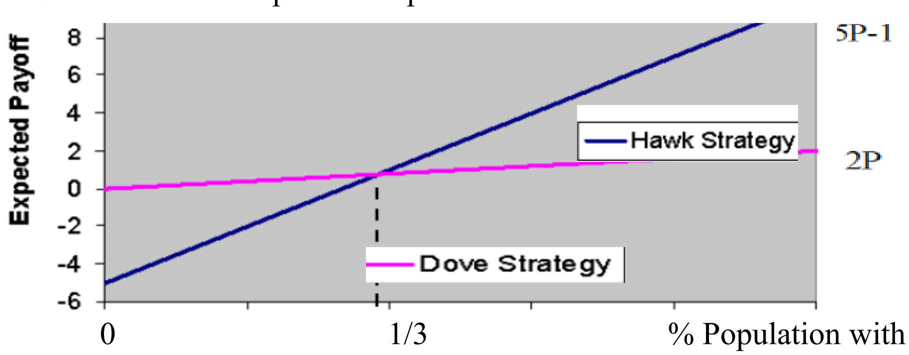
$W(H) = (1-p)HH + p\,HD$ and $W(D) = (1-p)DH + p\,DD$. Set W(H) = W(D) and find p
$HH - pHH + pHD = DH - p\,DH + p\,DD \rightarrow p = HH-DH / [\,HH - DH + DD - HD]$
Plug in the numbers and you get p = -1/(-1 - 2) = 1/3. **Equilibrium is 1/3 dove**

With p = 1/3: $W(H) = (1-p)(-1) + (p)\,4 = -2/3 + 4/3 = 2/3$ and $W(D) = (1-p)\,0 + p\,2 = 2/3$
In this model, the equilibrium payoff is 2/3, which is the ratio of the benefits of winning to the cost of losing, **V/C.**

The two strategies coexist because **each is best as an invader but is a poor defender.**

H as an invader of D gives E(HD)> E(DD) [D as a defender] in example, 4>2
H as a defender against D E(HH)< E(HH). [D as an invader] in example, 0 > -1.

The graph shows derivation of the equilibrium p:



With a mixed strategy M (p% D and 1-p% H), **Bishop-Cannings Theorem says the mixed must play as well against H and D.** HM = DM-- M must defenc equally well against either invasion. If HM > DM then increase the percentage using H, and conversely if D did better as invader. If there is a mixed strategy ESS, you can find it by solving HM = DM for p.

In the model with the 1/3rd D strategy, do H and D play equally well against the mixed strategy? Do the algebra and you will see YES. So could find the mixed strategy by solving $(1-p)\,HH + p\,EHD = (1-p)\,DH + p\,DD$

Plug in numbers from example and you indeed get p = 1/3

1)No guarantee an ESS exists. With 2 strategies, an ESS exists. For 3 or more strategies, an ESS need not.
2) T**here can be more than one ESS.** An example is the right/left game below.

	R	L
R	4	1
L	1	4

p is % who do Left

	p small	p large
W(R) = (1-p) RR + p RL ≈	4	0
W(L) = (1-p) LR + p LL ≈	0	4

3) Some ESSs are better than others – RL coordination game

	R	L
R	40	0
L	0	4

RR is the **GLOBAL MAXIMUM**. LL is **LOCAL MAXIMUM**. The **BASIN OF ATTRACTION** of R is larger because need only a modest 1-p % choosing R for it to be more profitable than choosing L. If 1- p% play R, you gain 40 (1-p) from R compared to 4p from playing L. The p that equates these two returns is 10/11 (44p = 40). If you start at R, you need 10/11 + e changers to make it worthwhile shifting to L. Conversely, start at L you need 1/11 + e changers to make shifting to R worthwhile. Since R needs fewer changes to gain dominance and more to lose its position, it is more powerful attractor.

With 11 people do L, need just 2 switchers to R to make it profitable for all to shift: if two shift, the shifters gain 40 while the remaining nine who do L gain 36 each. If fraction r switch every period, the chance of the best choice changing is proportionate to r^2 .

By contrast, if the 11 do R, and nine switch to L, R still does better. So it takes 10 shifters to make L more profitable, proportionate to r^{10} (Note with just one R, get 0)

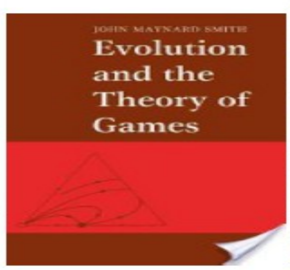
With 110 people you need 20 switchers to move from L to R, which is proportionate to r^{20} so CHANCE OF A SHIFT IS SMALLER the LARGER THE POPULATION.

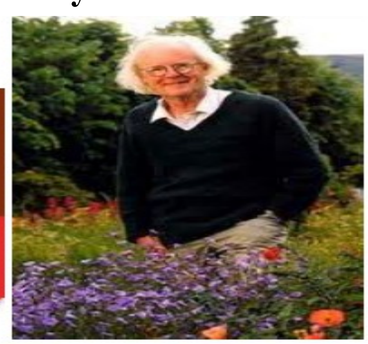
4) The move from LL to RR is discontinuous. Others must move with you for it to be profitable, and then it is profitable for everyone to move.

5) **TFT is not an ESS. "No pure strategy is evolutionary stable in the repeated prisoner's dilemma game"**
TFT type strategies have "large basins of attraction"so THEY do not need large proportion of the population to be an ESS but whether they are ESS depends on distribution of strategies in population. N**o single ESS for the Iterated PD**

6)**ESS does not rule out ALL invasions.** Depends on the type of invasion (whether single or multiple invaders) and dynamic process by which successful strategies expand. Original definition of ESS refers to invasion by a single strategy. With just 2 strategies, H and D, and H scores better than D, the way the superior strategy expands does not matter. But this does not hold for invasions with more strategies.

II) **TFT and ESS in Evolutionary Games**





AN EVOLUTIONARY GAME has three components

1)GAME MATRIX – allowable strategies and what they get when they meet other strategies. Strategies can be:
NONCONTINGENT – ALL C or D;
CONTINGENT – I respond to what you do – TFT (which can cost more since requires observe and act);
AKA Stochastic strategy – usually markov "mem one" – just depends on last turn's play/outcome.
INFERENTIAL – If you move contingently and you play C I don't know if you are TFT or all C or ... another "nice" strategy. So build models to try to INFER the other strategy;

ZD strategy – strategy that imposes relation between it and other strategies regardless of what they do.

2) AN ECOLOGY OF STRATEGIES – All possible strategies? Invasion by a single or by multiple strategies?
A strategy that invades with a servant strategy can readily conquer others. Consider a **NEUTRAL strategy --** one that does as well against the dominant strategy as it does against itself. If neutral strategy invades with a servant strategy that helps it and hurts the dominant one it will win.

Example: TFT, TF2T, and STFT: TF2T – Cooperate until partner Defects twice, then D
STFT – Defect then do what partner did last period

TFT and TF2T are neutral.	TFT does badly against STFT;	TF2T beats TFT against STFT	STFT vs STFT;
TFT C—>	TFT C D C	TF2T C C C —>	STFT D D -->
TF2T C--->	STFT D C D	STFT D C C —>	STFT D D -->

So TF2T+ STFT wins against TFT

Example: take the PD game matrix

	C	D
C	3	0
D	5	1

When TFT meets TFT it generates 3/(1-w) in present value with w discount: $(1+ w + w^2 + \ldots) \Rightarrow 1/(1-w)$ for $w < 1$
When TFT meets TF2T it generates the same 3/(1-w).
When TF2T meets TF2T it generates 3/(1-w) **Nice strategies just keep playing C**
When TF2T meets TFT it generates 3/(1-w)

TFT vs STFT generates 0, 5, 0, 5 which is about 2.5 each period. In PV it sums to $5w/(1-w^2)$
(because PV of alternating 5s is $5 w \ (1 + w^2 + \ldots = 1/(1-w^2))$
TF2T vs STFT generates 0, 3, 3, which is about 3 each period. In PV it sums to 3w/(1-w)
STFT vs TFT generates 5, 0, 5,0 which sums as $5 /(1-w^2)$
STFT vs TF2T generates 5, 3, 3, … 5+ 3w/(1-w)

TF2T gets 3/(1-w) + 3/(1-w) + 3w/(1-w) while TFT gets $3/(1-w) + 3/(1-w) + 5w/(1-w^2)$

Thus TF2T beats TFT when $3w/(1-w) > 5w/(1-w^2)$, which holds if $3 (1-w^2) > 5(1-w)$ or $3(1+w) > 5 \rightarrow w > 2/3$. Less we discount the future better chance for 3,3,3, .. etc to outdo 5,0,5 .. Simple calculation shows with no discount that 3 periods TF2T does worse 9 < 10 but with 4th period , 12>10.

3) A DYNAMIC (EVOLUTIONARY) PROCESS: Different dynamics can lead to different equilibrium:

1. **Proportional Fitness Reproduction (PFR):** You grow proportionate to your score relative to average. Your share of tomorrow/your share of today = Your Score/Average score. If p is current proportion of population and Wi is score and W is average, better performers increase share of future while poorer performers reduce share

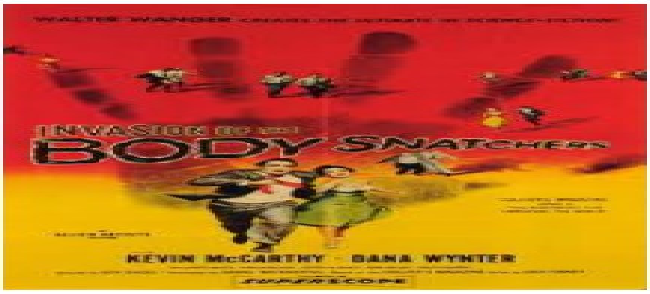
Replicator Dynamics: dp/p = Pk - Pk(-1)]/ Pk(-1)= λ (K's score last period - average score last period)

2. **Imitate the winner** – the highest scoring strategy grows. Often used with a local neighborhood model, where people copy the neighbor that does best if it does better than they do. More likely to be discontinuous.
With IMITATE THE WINNER, A NEUTRAL + SERVANT strategy will immediately win.
With PFR, the NEUTRAL will grow until the SERVANT is wiped out.
Dynamic process can be more subtle in some games with lots of players: a certain number of players randomly choose someone else and copy them if their score exceeds yours – followers along the line-- idea.
A STRATEGY MAY BE STABLE IN SOME ECOLOGY BUT NOT IN OTHERS
A STRATEGY MAY BE STABLE UNDER SOME DYNAMICS BUT NOT OTHERS

BINMORE AND SAMUELSON: "That varying the details of a dynamic game can alter the equilibrium selected shows that the **institutional environment ... can matter for equilibrium selection."**

Maynard Smith definition of an ESS with single mutants implies that dynamics don't matter: if you can beat out the mutant, you win; so the dynamics and ecology are important together.

MINIMAL STABILIZING FREQUENCY: WHAT WE NEED TO SURVIVE MASS INVASIONS .



Bendor-Swistak (political scientists) ask: if you face invasion of multiple strategies, is there a condition under which you keep from being wiped out? Answer is YES, there is minimal proportion of the population with PFR that guarantees you survive. The minimal proportion is lower for "nice strategies" that have evolutionary advantage (usually from interacting with themselves). A small minimal stabilizing frequency <--> a large basin of attraction. You are more likely to survive if you need 50% to survive than if you need 70%. When evolution follows PFR and the native population is playing nice, invaders of small size will fail to drive the natives to extinction but large group can wipe you out.

1) The Minimal Stabilizing frequency > 50% – have to get over 50% in a world with lots of different strategies to assure you are not wiped out.

2) With PFR, nice and **retaliatory** strategies (TFT and its clones) have a minimum stabilizing frequency that ---> 0.5 as the future becomes more important (w --> 1). Just need ½ to keep that proportion. (Nice and retaliatory are CONTINGENT not probabilistic strategies).

3)With PFR, observed behavior will end up with all nice strategies (not necessarily TFT)

4) Under *any* evolutionary dynamic, all strategies that have minimum stabilizing frequency --> .5 as w--> 1 must be "almost" nice and "almost" retaliatory.

5) Under *any* evolutionary dynamic, the minimum stabilizing frequency of a NASTY strategy (NEVER cooperate first) converges to 1 as w converges to 1.

TFT type "nice" strategies (those that never defect first)can survive in world where future matters when they are a majority; **while nasty strategies (those that defect "without cause" need higher proportion to survive).** If w is high, TFT-types have a bigger basin of attraction than Nasties and are "weakly stable" under PFR.

Provisos: http://plato.stanford.edu/entries/prisoner-dilemma/ – Bendor and Swistak's results are not consistent with Nowak/Sigmund simulations with **probabilistic** strategies where defects did better. Pavlov (win-stay; lose-change) has low minimum stabilizing frequency but is overturned by invasions of unconditional defectors exceeding 10% of the population. The **simulations allow for errors** that open door for different results; need broader definition of "nice" or "retaliatory" for strategies that are probabilistic /have errors of communication; "if the number of generations is large compared with the original population that plays min stabilizing strategy, the sequence of invading groups could reduce the original strategy to less than half of the population."

Dal Bó and Fréchette The Evolution of Cooperation … **Experimental Evidence**, AER (February 2011): 411–429: "... cooperation may not prevail even when it is a possible equilibrium action. This provides a word of caution against … assuming that subjects will cooperate whenever it is an equilibrium action... cooperation does prevail under some treatments—namely, when the probability of continuation and the payoff from cooperation are high enough."

III – One strategy to Rule Them All? ZD Determinant

"It would be surprising if any significant mathematical feature of IPD has remained undescribed, but that appears to be the case" (Dyson& Press. 2012). Also surprising 93yr old Dyson added it!

"Axelrod's 1980 tournaments … strategies have been condensed into … Don't be too clever, don't be unfair. Press and Dyson have shown that cleverness and unfairness triumph after all."W. Poundstone

Zero Determinant strategy "controls" outcomes regardless of what an opposing strategy does by playing conditional probability between 0 and 1 depending on last period's play. Press & Dyson Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent" PNAS June 26, 2012, 109 (26) 10409-10413. Paper shows that an opponent who considers earlier encounters does no better playing against a strategy that remembers only last period , so analysis need only consider strategies has "mem 1". Below: ZD and 4 other "mem 1" strategies:

Conditional Probability of Playing C

	ZD Strategy	All Coop	All Defect	TFT	"Pavlov – WSLC"
CC	Pcc	1	0	1	1
CD	Pcd	1	0	0	0
DC	Pdc	1	0	1	0
DD	Pdd	1	0	0	1

ZD will likely set Pcc high; Pcd low; Pdc high; Pdd low but not 0. It's flexibility gives it an "edge:"

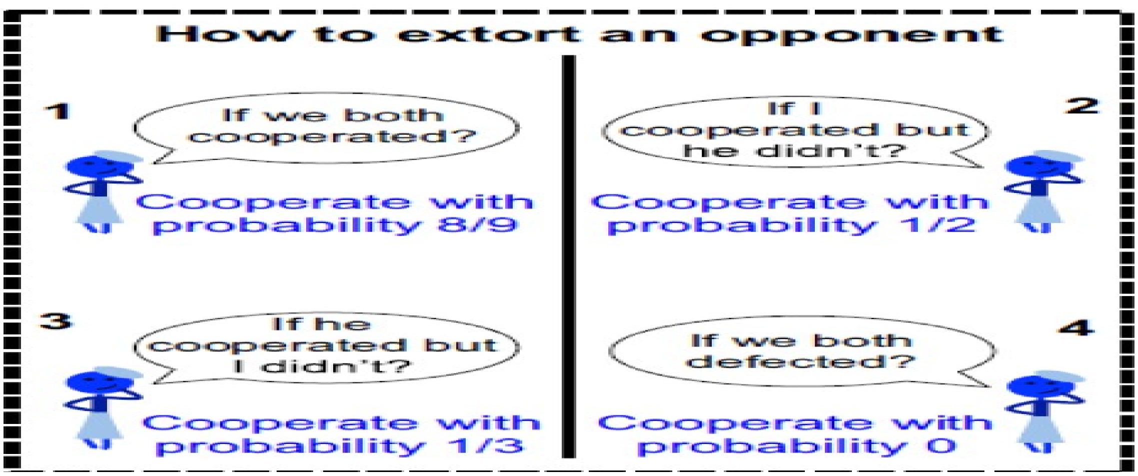
Let Qcc, Qcd, Qdc, Qdd be conditional probabilities of 2nd player. The two conditional probs give a distribution for the outcome of each round, conditional on previous round – a 4 by 4 Markov transition matrix M for the four outcomes in this period to the next

	CC	CD	DC	DD
CC	Pcc Qcc	Pcc (1-Qcc)	(1-Pcc) Qcc	(1-Pcc) (1-Qcc)
CD	Pcd Qdc	Pcd (1-Qdc)	(1-Pcd) Qdc	(1-Pcd) (1-Qdc)
DC	Pdc Qcd	Pdc (1-Qcd)	(1-Pdc) Qcd	(1-Pdc) (1-Qcd)
DD	Pdd Qdd	Pdd (1-Qdd)	(1-Pdd) Qdd	(1-Pdd) (1-Qdd)

Press and Dyson show that the equation can be expressed as a determinant in which one column involves only the four probabilities of one player's strategy and another column that involves the probabilities of the other player's strategy. By setting the determinant to zero, hence the name ZD, and solving resultant linear equation. This allows a player to force a given linear relation between the outcomes of both players **regardless** of the other players' strategy. ZD can force other person's outcomes to be low, high, smaller than own.

http://s3.boskent.com/prisoners-dilemma/fixed.html forces target of 2 on other player. If you cooperated last time, it cooperates with prob 2/3. If you defected while it cooperated, it cooperates with probability 0; If you defected last time and it defected, it cooperates with prob 1/3 → Whatever the non-ZD player does its long term outcome is 2.

ZD can also "Extort" gains by **defecting enough times to win in any one on one contest.** Extort-2 below forces the relationship where ZD gains twice the share of payoffs above a value P compared with opponent.



But ZD can also be generous. Stewart and Plotkin From extortion to generosity, evolution in the Iterated Prisoner's DilemmaPNAS September 17, 2013 110 (38) 15348-15353 define **"generous ZD strategies,"** that reward cooperation but punish defection only mildly and thus score lower payoffs than those of defecting opponents. (but dominate in evolving populations).

What is the ZD innovation? Different information – the ZD player knows ZD and the other does not, giving ZD edge

Responses: "An underway revolution in game theory" that "changes the research paradigm of game theory …with ... dozens of ingenious ideas and untraditional approaches " H. Dong, et al (Chinese Physics B, 2014)

W. Press: "When both players have a theory of mind (that is, are not just evolving to maximize their own score) are all games in some deep way, actually Ultimatum Games.

F. Dyson;"Cooperation loses and defection wins ... My view of the evolution of cooperation is colored by my memories of … Christmas and Guy Fawkes. … Christmas was boring and Guy Fawkes was fun. We were born with an innate reward system that finds joy in punishing cheaters. The system evolved to give cooperative tribes an advantage ... using punishment to give cooperation an evolutionary advantage within the tribe. **This double selection**

of tribes and individuals goes way beyond the Prisoners' Dilemma model."

Chad English (http://blogs.plos.org/neuroanthropology/2012/06/24/ prisoners-dilemma-and-the-evolution-of-inequality-does-unfairness-triumph-after-all/) "provides the demonstrable benefit of unions and governments … companies in capitalist societies make use of ZD strategies to exploit their shorter term acting employees ("employees who do not know ZD strategies"). It is …in the interests of workers to create … ZD strategy organization aka a union).

Key to social/bio importance is how ZD does in evolutionary games.

ZD opens up "a vast set of strategies linking the scores of two players deterministically (as TFT does), but asymmetrically (unlike TFT). This enriches the canvas of individual interactions, but **not necessarily the range of outcomes open to evolving populations."** Indeed, winning one-on-one does NOT mean winning in evolution. Why? Because ZD scores lows against itself. If two ZD extortionary strategies meet, they end up with DD lowest value.

As extortionate ZD beats opponents, its share of population grows and then it meets ZD and gets low score. If 'extortionate' strategy guarantees that two players do twice as well as opponent, both get nothing. Simulations of ZD in evolutionary context shows that ZD does not fare well. (C Hilbe M A. Nowak,A Traulsen PLOS 2013 "Adaptive Dynamics of Extortion and Compliance). **ZD-extortion loses to** Pavlov (WSLC). ZD scores higher than Pavlov in each encounter but Pavlov scores higher because it cooperates with itself. ZD-extortion must cooperate against itself. It needs a marker/communication to win in evolutionary game.

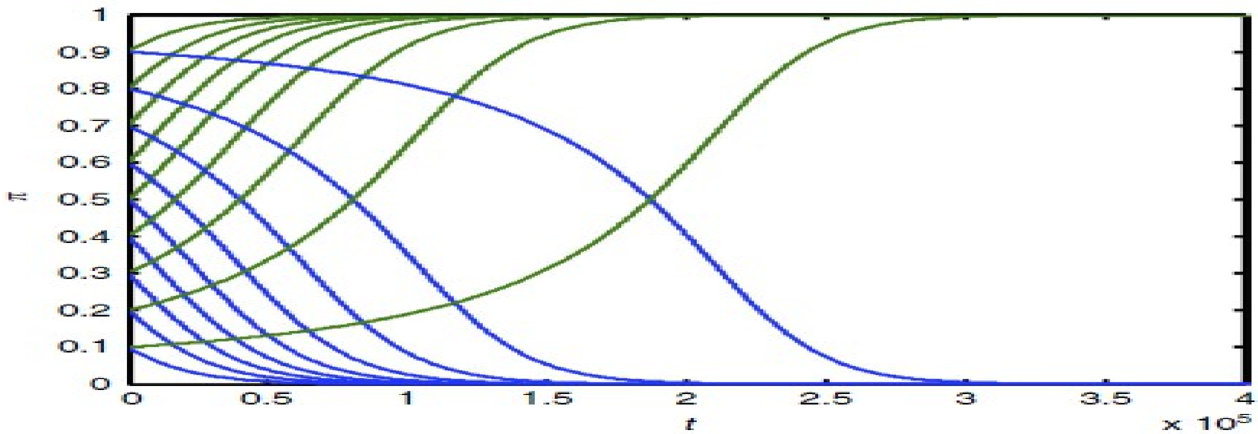
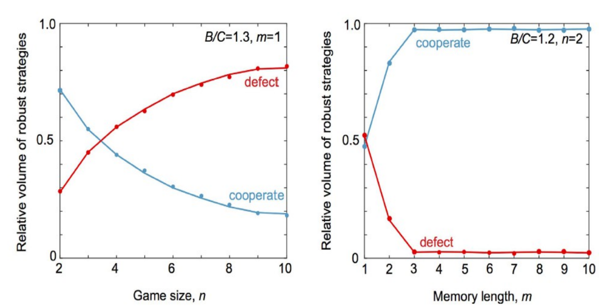


Figure 2 | Population fractions of ZD versus Pavlov over time. Population fractions π_{ZD} (blue) and π_{PAV} (green) as a function of time for initial ZD concentrations $\pi_{ZD}(0)$ between 0.1 and 0.9.

Stewart AJ, Plotkin JB (2013) find that "For all but the smallest population ... **generous** ZD strategies are robust to being replaced … and can selectively" replace any non-cooperative ZD strategy. " … "In some regimes, generous strategies outperform even the most successful ...IPD strategies, including Pavlov" "Whether cooperative strategies are favored in the long run critically depends on the size of the population ...cooperation is most abundant in large populations, in which case average payoffs approach the social optimum. … "

But in **multi-player games**, results differ: Stewart and Plotkin (2016) "Small groups and long memory promote cooperation" http://www.nature.com/articles/srep26889 "Free rider incentives:. In a two-player game, pairwise interactions occur in the population at each generation. If the whole population plays the game each generation all players interact simultaneously. **Memory of past events...** allows for more complex strategies, such …punish rare defection or reward rare cooperation.



Active research area, with shifting views on whether ZD extortion or cooperation dominate

Hilbe, C., Wu, B., Traulsen, A., & Nowak, M. A. (2015). Evolutionary performance of zero-determinant strategies in multiplayer games. Journal of Theoretical Biology, 374, 115–124. Repeated interactions provide an important explanation for the evolution of cooperation: individuals cooperate because they can expect to be rewarded in future reciprocal strategies... Our simulations suggest that the evolutionary success of ZD strategies critically depends on the **size of the group**. Repeated interactions can only help sustaining cooperation when groups are sufficiently small. The downfall of cooperation in large groups can be prevented if large-scale endeavors have an efficiency advantage: if there are additional pairwise incentives to cooperate; they can increase their strategic power by coordinating their actions and by forming alliances or ...implement central institutions that enforce mutual cooperation.

Asymmetric Power Boosts Extortion in an Economic Experiment. PLoS ONE 11(10): Hilbe C, Hagel K, Milinski M (2016) Direct reciprocity is a major mechanism for the evolution of cooperation. but subjects who use ZD extortionate strategies are able to exploit and subdue cooperators... we show with a model and **an economic experiment** that extortionate strategies readily emerge once subjects differ (in power)... . In our main treatment... one randomly chosen group member ... is unilaterally allowed to exchange one of the other group members after every ten rounds of the social dilemma.... asymmetric replacement opportunity generally promotes cooperation, but often the resulting payoff distribution reflects the underlying power structure. **Almost half of the subjects in a better strategic position turn into extortioners**, who … exploit their peers. By adapting their cooperation probabilities consistent with ZD theory, extortioners force their co-players to cooperate without being similarly cooperative themselves. … Our results thus highlight how power asymmetries can endanger mutually beneficial interactions, and transform them into exploitative relationships … and that the extortionate strategies predicted from ZD theory could play a more prominent role in our daily interactions than previously thought.,

Extortion can outperform generosity in the iterated prisoner's dilemma Nat Commun. 2016; 7: 11125.Z Wang, YZhouJ Lien JZheng and Bxua, 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... 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... 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.

Extortion strategies resist disciplining when higher competitiveness is rewarded with extra gain *Nat Commun* **10,** 783 (2019). L Becks & M Milinski. Cooperative strategies are predicted for repeated social interactions....ZD strategies enforce the partner's cooperation because the 'generous' ZD players help their cooperative partners while 'extortionate' ZD players exploit their partners' cooperation. Partners may accede to extortion because it pays them to do so, but the partner can sabotage his own and his extortioner's score by defecting to discipline the extortioner... we show with human volunteers that an additional monetary incentive (bonus) paid to the finally competitively superior player maintains extortion. **Unexpectedly, extortioners refused to become disciplined**, thus forcing partners to accede. Occasional opposition reduced the extortioners' gain so that using extortion paid off only because of the bonus. With no bonus incentive, players used the generous ZD strategy.

Autocratic strategies for iterated games with arbitrary action spaces Proceedings of the National Academy of Sciences Mar 2016, 113 (13) 3573-3578; A.McAvoya and C.Hauert The recent discovery of zero-determinant strategies for the iterated prisoner's dilemma sparked a surge of interest in the surprising fact that a player can exert unilateral control over iterated interactions. These remarkable strategies, however, are known to exist only in\ games in which players choose between two alternative actions such as "cooperate" and "defect." Here we introduce a broader class of autocratic strategies by extending zero-determinant strategies to iterated games with more general action spaces. We use the continuous donation game as an example, which represents an instance of the prisoner's dilemma that intuitively extends to a continuous range of cooperation levels. Surprisingly, despite the fact that the opponent has infinitely many donation levels from which to choose, a player can devise an autocratic strategy to enforce a linear relationship between his or her payoff and that of the opponent even when restricting his or her actions to merely two discrete levels of cooperation. In particular, a player can use such a strategy to extort an unfair share of the payoffs from the opponent. Therefore, although the action space of the continuous donation game dwarfs that of the classic prisoner's dilemma, players can still devise relatively simple autocratic … extortionate strategies.

Strategies that enforce linear payoff relationships under **observation errors** in Repeated Prisoner's Dilemma game A Mamiya, G. Ichinose Journal of Theoretical Biology 477, 63-76 (2019) Errors often happen between players in the real world. Even in the case with observation errors, the only strategy sets that enforce a linear payoff relationship are either ZD strategies or unconditional strategies.

PAPER STUDYING RESPONSE TO Press & Dyson would be extraordinary. My bet would be get wide mix of references from math, comp sci, biology, econ/other social science. Changing/differing views on robust cooperative.