Spelteoretisk jämvikt som ett resultat av inlärning eller evolution

Erik Mohlin

Abstract
One way of justifying the use of equilibrium concepts in game theory is by appeal to learning or evolution. This paper provides a non-technical introduction to evolutionary game theory. A short summary of basic standard (non-evolutionary) game theory sets the stage and introduces a number of simple games. Evolutionary game theory is then introduced via a simple learning rule, according to which individuals occasionally revise their strategy choices by imitating currently successful peers. The resulting dynamic is studied in the previously introduced games, investigating whether behaviour converges to an equilibrium, and if so, which one. Finally, some empirical evidence is discussed.

Spelteori beskrivs ofta som en teori om perfekt rationella aktörer vars beteende är i jämvikt. Var och en fattar ett beslut som är *rationellt* i den mening att det är optimalt utifrån respektive beslutsfattares målsättning och vad de *tror* att andra kommer göra. Teoretiskt fångas beslutsfattarens målsättning i en så kallad nyttofunktion; rationalitet består i att välja det alternativ som maximerar den förväntade nyttan givet beslutsfattarens subjektiva probabilistiska förväntningar om hur andra beter sig. I en *jämviktsanalys* måste vi dessutom förutsätta att vars och ens förväntningar om vad motparten kommer att göra visar sig vara fullständigt korrekta. Därmed fattar var och en ett optimalt (nyttomaximerande) beslut givet sin målsättning och givet vad motparterna *faktiskt* gör.

Det finns definitivt situationer där aktörerna fattar övertänkta beslut som kännetecknas av en mycket hög grad av rationalitet. Det finns till och med sällsynta fall där man kan motivera de korrekta jämviktsförväntningarna med

Erik Mohlin är docent i nationalekonomi vid Lunds universitet och Institutet för Framtidsstudier.
E-post: erik.mohlin@nek.lu.se

hänvisning till rent logiska resonemang om vad gemensam kunskap om rationalitet implicerar för vilka beteenden man kan förvänta sig av sin motpart.[1] I allmänhet verkar det dock rimligare att tänka sig att rationellt beslutsfattande är erfarenhetsbaserat, dvs att beslutsfattaren haft tid och möjlighet att observera, pröva och utvärdera olika handlingsalternativ. Framförallt verkar det svårt att se hur aktörerna skulle kunna skaffa sig korrekta förväntningar om vad andra kommer att göra ifall de inte har erfarenhet av interaktionen – och andras beteende i den.

Lyckligtvis är det inte så att spelteorin behöver börja och sluta med perfekt rationalitet och jämvikt. Det finns en omfattande teoretisk och experimentell forskning om hur inlärning (exempelvis i form av upprepade ”försök och misstag”) och gradvis anpassning av beteende (i ljuset av vad man lär sig om andras beteende och handlingars konsekvenser) kan leda till jämviktstillstånd där alla inblandade beter sig på ett sätt som är rationellt (nyttomaximerande) för var och en, givet de andras beteende.[2] I föreliggande uppsats ger jag en kort introduktion till detta forskningsfält. Jag kommer mestadels att hålla mig på ett teoretiskt plan, men även ge hänvisningar till experimentella studier som testar teorin.[3] Det matematiska inslaget hålls på ett absolut minimum, även om en del resonemang utvecklas i ett matematiskt appendix. Mitt mål är att texten skall kunna läsas även av dem som på sin höjd såg några spelmatriser i en kurs för många år sedan. Därför följer en kort repetition av klassisk spelteori.

Lite klassisk spelteori

SPELARE, STRATEGIER OCH NYTTOR

Spelteorin är ett verktyg för att analysera interaktiva beslutssituationer, det vill säga situationer där konsekvenserna av en persons beslut även beror på andra personers beslut. Namnet till trots är teorin inte begränsad till att handla om spel i vardaglig mening, som poker (även om också sådana spel kan analyseras med hjälp av spelteori). Tvärtom är spelteorins räckvidd mycket stor. Teorin kan i princip användas till att studera alla sociala interaktioner – från

1 En strategi i ett spel sägs vara dominerad om den ger en lägre nytta än en annan strategi, oavsett vad de andra spelarna gör. En rationell spelare kommer naturligtvis aldrig att välja en dominerad strategi. En spelare som vet att motparten är rationell kan sluta sig till att motparten aldrig väljer en dominerad strategi. En spelare som dessutom vet att de andra spelarna vet att han eller hon är rationell kan eventuellt dra ytterligare slutsatser om vilka strategier de andra spelarna kommer undvika. Vi kan fortsätta resonemanget för godtyckligt komplexa hierarkier om kunskap om andras kunskap. För mig framstår detta angreppssätt som en fullständigt orealistisk väg att rättfärdiga jämvikt.

2 Redan i sin doktorsavhandling (Nash 1950) föreslog Nash en evolutionär tolkning av det jämviktsbegrepp som senare skulle komma att bära hans namn.

3 Den experimentella litteraturen om inlärning i spel sammanfattas väl av Camerer (2003), även om en del har hänt sedan den kom ut.

marknadsutbyten och förhandlingar till krig och samarbete, liksom interaktioner mellan djur.[4]

I spelteorin kallas en interaktiv situation för ett spel och deltagarna i interaktionen kallas spelare. Spelarna väljer var sin strategi. Varje spelares utfall beror på vad alla andra spelare gör. Spelarnas drivkrafter representeras av nyttofunktioner som anger hur en spelare värderar vart och ett av de möjliga utfallen. Närmare bestämt antas i allmänhet von Neumann-Morgenstern nyttor, vilken innebär att tilldelningen av nyttor till de olika utfallen är sådan att spelarna strävar efter att maximera förväntad nytta.[5]

Om du exempelvis är ute och kör bil kan strategierna vara dina olika färdvägar och spelarna kan utgöras av alla andra trafikanter som är ute och rör sig samtidigt som du. Din nytta kan tänkas bero på om du kan undvika köer och slippa olyckor. Dina möjligheter att komma fram fort och säkert beror på vilka färdvägar andra trafikanter väljer. På samma sätt beror dina medtrafikanters utfall på hur du beter dig.

Ett vanligt missförstånd är att tolka nytta som det används i vardagsspråket, kanske i termer av materiella nyttigheter. Nytta skall istället ses som en teknisk term som representerar vad helst spelaren värderar och motiveras av. Spelteorin tillåter såväl altruistiska och idealistiska som själviska och materialistiska drivkrafter. Med detta sagt måste det tilläggas att när man försöker testa spelteori i laboratorieexperiment är det ofta bekvämt att motivera deltagarna med monetära belöningar. När man tolkar resultaten från sådana experiment måste man naturligtvis vara medveten om att de subjektiva nyttorna, eller inre belöningarna, kan skilja sig systematiskt från de objektiva materiella belöningarna.

För att förstå eller förutsäga vad som händer i sådana här interaktioner använder spelteorin vanligen olika jämviktsbegrepp. Det vanligaste och mest grundläggande jämviktsbegreppet inom spelteorin är Nashjämvikt. En Nashjämvikt är en uppsättning strategier (en strategi per spelare), sådana att ingen spelare kan uppnå en högre nytta genom att ensidigt byta strategi, det vill säga genom att själv byta strategi samtidigt som de andra spelarna inte byter strategier.

Låt oss nu konkretisera detta abstrakta ramverk genom att studera två enkla (men viktiga) exempel på spel. Det så kallade ”Hök och Duva”-spelet beskriver en konfliktsituation med delvis motstridiga intressen. Spelet ”Hjortjakten” beskriver istället en situation med sammanfallande intressen.

4 Huruvida analysen är fruktbar är en annan fråga som bör bedömas från fall till fall.

5 Man kan visa matematiskt att om en individs preferenser rörande de olika utfallen i spelet uppfyller vissa axiom, så finns det någon nyttofunktion som är sådan att individerna beter sig som om de maximerade förväntad nytta.

KONFLIKTSPEL: HÖK OCH DUVA

Tänk dig två länder som gör anspråk på en obebodd ö. På ön finns naturresurser till ett värde av 6. Varje land kan välja om det ska följa en aggressiv strategi för att lägga beslag på hela ön, något som vi kan kalla en hökstrategi, eller följa en mera generös strategi som syftar till fredlig samexistens på ön, något som vi kallar för en duvstrategi. Om båda länder väljer att bete sig som hökar så utbryter en konflikt som minskar öns värde med 8, och i slutändan får båda länder dela på ön, så att vartdera landet får nytta $-1 = (6 - 8) / 2$. Om ett land väljer hökstrategin och det andra landet väljer duvstrategin, får höklandet hela ön medan duvlandet får ingenting. Om båda länder följer duvstrategin så delar de på ön utan att behöva genomlida konflikten som minskar värdet på ön. Vi kan sammanfatta detta med matrisen i figur 1.

Den första siffran i varje ruta i figur 1 anger nyttan för radspelaren, den spelare som väljer mellan strategierna som representeras av raderna. Den andra siffran i varje ruta anger nyttan för kolumnspelaren, den spelare som väljer mellan strategierna som representeras av kolumnerna.

Figur 1. Konfliktspel: hök och duva.

	Hök	Duva
Hök	-1 , -1	6 , 0
Duva	0 , 6	3 , 3

En alternativ tolkning av matrisen i figur 1 är att det handlar om två spelare som är inbegripna i en förhandling. Det skulle kanske kunna handla om en arbetsgivare och en arbetstagare som förhandlar om hur mycket av vinsten som ska tillfalla respektive part. Om konflikt uppstår (i form av strejk eller lock-out) så förlorar båda sidor på det. Man kan också tolka spelet som en förhandling mellan två företag.[6]

Oavsett vilken tolkning vi gör visar matrisen att om motspelaren spelar hök så är det rationellt att spela duva. Att spela duva när motparten spelar hök ger visserligen bara nyttan 0, men att spela hök när motparten spelar hök ger en ännu lägre nytta, nämligen -1. På motsvarande sätt är det, om motspelaren spelar duva, rationellt att spela hök eftersom man då får 6 medan man bara skulle fått 3 ifall man spelade duva. Det betyder att både kombinationen duva-hök och kombinationen hök-duva är jämvikter.

Man kan fråga sig hur de två kontrahenterna bestämmer sig för vem som ska spela duva och vem som ska spela hök. Båda spelare föredrar ju den jämvikt

6 Rent allmänt är det naturligtvis så att en och samma spelmatris kan tolkas som en beskrivning av många olika situationer.

där de själva spelar hök och motparten spelar duva. Går det att finna en annan jämvikt?

Hittills har vår diskussion utgått från att varje spelare måste bestämma sig för en strategi. Det verkar dock rimligt att tillåta att spelare kan välja mellan sina handlingsalternativ på ett mera slumpmässigt sätt och därmed välja olika alternativ med olika sannolikhet. En spelare som utför flera handlingar med positiv sannolikhet sägs följa en blandad strategi. Motsatsen, att spela en enda handling med sannolikheten ett, kallas för att följa en ren strategi. För att förstå hur blandade strategier fungerar kan det vara bra att tänka på straffsparkar i fotboll. En framgångsrik målvakt slänger sig inte alltid åt samma håll och en framgångsrik straffläggare skjuter inte alltid åt samma håll. Just möjligheten att blanda sin strategi är avgörande i sammanhanget.

Låt oss anta att kolumnspelaren följer en blandad strategi som slumpmässigt alternerar mellan två rena strategier, så att vederbörande väljer hökalternativet med sannolikheten q och duvalternativet med sannolikheten $1 - q$. Radspelarens förväntade nytta av att bete sig som hök blir då

$$U_{Hök} = (-1) \cdot q + 6 \cdot (1 - q) = 6 - 7q$$

och radspelarens förväntade nytta av att bete sig som duva blir

$$U_{Duva} = 0 \cdot q + 3 \cdot (1 - q) = 3 - 3q.$$

Notera att de båda handlingsalternativen ger samma förväntande nytta om och endast om $6 - 7q = 3 - 3q$, vilket är fallet om och endast om $q = 0{,}75$. Det betyder att ifall kolumnspelaren slumpmässigt väljer mellan rena strategier på ett sådant sätt att man väljer hök med sannolikheten $q = 0{,}75$, då kan radspelaren inte förbättra sin förväntade nytta genom att ändra sitt beteende; skälet är att båda strategier ger exakt samma förväntade nytta.

Situationen är likadan för kolumnspelaren. Anta att kolumnspelaren slumpmässigt väljer mellan de två alternativen hök och duva, så att hon agerar som hök med sannolikheten p och som duva med sannolikheten $1 - p$. Om radspelaren väljer hök med sannolikheten $q = 0{,}75$, så kan kolumnspelaren inte förbättra sin förväntade nytta genom att ändra sitt beteende; återigen är skälet att båda strategier ger exakt samma förväntade nytta.

Det betyder att om radspelaren väljer hök med sannolikheten $p = 0{,}75$, samtidigt som kolumnspelaren väljer hök med sannolikheten $q = 0{,}75$, så står båda spelarna inför situationen att båda deras strategier ger samma förväntade nytta. Således kan ingendera spelaren förbättra sin förväntande nytta genom att ensidigt ändra sitt beteende. Detta är ju precis vad definitionen av Nashjämvikt kräver. Alltså har vi identifierat en tredje jämvikt. Vi säger att denna tredje jämvikt är blandad eftersom den innebär att spelarna använder blandade strategier. De två övriga jämvikterna sägs vara rena eftersom spelarna där bara utför en handling var.

Man kan fråga sig hur det går till när spelare väljer ett alternativ med en viss sannolikhet. Kanske kan spelarna låta magkänslan avgöra vilken alternativ de väljer: ett mera precist sätt skulle kunna vara att använda en fyrsidig tärning och sedan spela duva om tärningen landar på siffran 1, 2, eller 3 (eftersom sannolikheten för att en fyrsidig tärning landar på siffran 1, 2 eller 3 är 0,75). I avsnittet om hjortjakten nedan kommer vi se att inlärning ger oss en annan (och ofta rimligare) tolkning av vad en blandad strategi är.

Man kan även fråga sig vilken, eller vilka, av de tre jämvikterna i konfliktspelet som troligast kommer att blir utfallet i en verklig situation. Innan vi undersöker dessa frågor ska vi stifta bekantskap med ytterligare ett spel som innehåller tre jämvikter, men som skiljer sig från konfliktspelet på avgörande sätt.

KOORDINATIONSSPEL: HJORTJAKTEN

Två jägare kan välja mellan att jaga hjort eller hare. För att jaga hjort framgångsrikt krävs det att båda jägare hjälps åt; om en jägare försöker jaga hjort medan den andra jagar hare så får den förre inte något byte. Att jaga hare ger däremot ett visst byte även om den andre spelaren jagar hjort, men bytet blir mindre ifall båda jagar hare eftersom de då konkurrerar om samma djur. Jägarna bryr sig endast om att komma hem med så mycket kött som möjligt. Vi kan sammanfatta dessa preferenser i matrisen i figur 3.[7]

Figur 2. Koordinationsspel: hjortjakten.

	Hjort	**Hare**
Hjort	6 , 6	0 , 3
Hare	3 , 0	2 , 2

Båda spelare föredrar att jaga hjort ifall den andre spelaren jagar hjort. På samma vis föredrar båda att jaga hare ifall den andre spelaren jagar hare. Därför finns det en jämvikt där båda jagar hjort och en annan jämvikt där båda jagar hare.

Spel med flera symmetriska jämvikter kan illustrera hur sociala normer eller konventioner, det vill säga outtalade beteenderegler, fungerar. Många olika konventioner är möjliga men om man förväntar sig att motparten följer en viss konvention så ligger det i ens intresse att också följa just den konventionen. Huruvida man kör på höger eller vänster sida om vägen är ett enkelt men tydligt exempel. I Sverige kör vi på höger sida om vägen för att lagen kräver det men framför allt därför att vi annars riskerar att råka ut för en olycka. Före 1967 körde svenskar istället på vänster sida. Det krävdes ett regeringsbeslut och en

7 Situationen, om än inte spelmatrisen, beskrevs redan av Jean-Jacques Rousseau.

noga planerad övergång för att genomföra bytet av konvention.[8] För det mesta växer dock konventioner fram över lång tid och är resultatet av gradvis förändrat beteende hos en grupp av människor. Evolutionär spelteori är ett sätt att modellera dessa långsamma processer genom vilka konventioner uppstår.[9]

Liksom i konfliktspelet finns det ytterligare en (blandad) jämvikt. Låt oss anta att kolumnspelaren följer en blandad strategi och väljer slumpmässigt mellan sina två rena strategier: hjort med sannolikheten q och hare med sannolikheten $1 - q$. Radspelarens förväntade nytta av att jaga hjort blir då

$$U_{Hjort} = 6 \cdot q + 0 \cdot (1 - q) = 6q$$

och radspelarens förväntade nytta av att jaga hare blir

$$U_{Hare} = 3 \cdot q + 2 \cdot (1 - q) = 2 + q$$

Vi ser att de båda strategierna ger samma förväntande nytta om och endast om $6q = 2 + q$, det vill säga om och endast om $q = 0{,}4$. Situationen är precis likadan för kolumnspelaren. Den tredje, blandade, jämvikten är alltså sådan att båda spelare följer den blandade strategin som spelar hjortalternativet med sannolikhet $0{,}4$.

Evolutionär spelteori

Det är nu dags att bli tydligare med hur vi tänker oss att inlärning går till och hur det sker en gradvis anpassning av beteende som eventuellt kan hamna i jämvikt. Vi behöver inte anta att spelarna är fullt rationella. Det räcker att de följer tumregler som i någon mån anpassas efter om en strategi är framgångsrik eller inte. Det kan exempelvis handla om att man härmar de medspelare som verkar lyckas bättre med sina val av strategier, eller att man prövar nya strategier på måfå när man tycker att ens nuvarande strategi ger ett dåligt utfall.

Den formella teorin som ligger till grund för diskussionen här kallas för *evolutionär spelteori*. Termen evolutionär ska dock inte förstås bokstavligt. Den matematiska teorin är tillräckligt generell för att kunna beskriva både biologiska evolutionära processer och sociala eller psykologiska eller sociala inlärningsprocesser. I både det biologiska och det kulturella fallet handlar det om populationer vars sammansättning och beteenden förändras gradvis på ett sätt som åtminstone delvis beror på hur, i någon mening, framgångsrika olika beteende är. I det följande kommer vi huvudsakligen att använda termen inlärning och inte evolution.

8 Man kan eventuellt invända att anledningen till att vi kör på höger sida är att det är en trafikregel och att man bötfälls om man bryter mot regeln. Jag skulle dock gissa att det inte främst är rädslan för böter utan rädslan för krock som får oss att respektera trafikregeln och den konvention den kodifierar.

9 Young (1998) analyserar många konventioner på detta sätt, inklusive konventionen om vilken sida av vägen man kör på.

EN ENKEL TUMREGEL FÖR INLÄRNING

Vi tänker oss en stor population av individer som varje dag delas in i olika par för att spela ett spel med varandra. *För det mesta följer var och en av individerna samma strategi som de följde igår. Då och då får de dock för sig att undersöka om de kanske ska pröva en annan strategi.*[10] *En individ som funderar på att byta strategi jämför sig då med en slumpvist vald annan individ i populationen. Om den individen följer en strategi som ger en högre förväntad nytta så är den jämförande individen benägen att byta till den mer framgångsrika strategin. Åtminstone är sannolikheten för ett byte till en strategi högre ju större skillnaden i förväntad nytta är.*

Man kan visa att om alla individer beter sig på det beskrivna sättet kommer andelen som spelar en viss strategi att öka om och endast om den strategi de spelar ger en förväntad nytta som ligger över den genomsnittliga, förväntade nyttan i populationen.[11] När det bara finns två strategier betyder det att den strategi som för närvarande ger mest nytta ökar i popularitet och att den som ger minst nytta minskar i popularitet. Det finns många andra tumregler för inlärning än den som beskrivits här, som också leder till att strategier som ger mer än andra växer i antal medan de som ger mindre minskar i antal.[12] I det följande utgår vi helt enkelt från att den strategi som för ögonblicket ger högst förväntad nytta ökar i popularitet i populationen och att den strategi som ger lägst förväntad nytta minskar i popularitet.

Vi kan även definiera en biologisk evolutionär process med i stort sett identiska egenskaper. Återigen tänker vi oss en stor population av individer som varje dag delas in i olika par för att spela ett spel med varandra. Varje individ är programmerad att välja en av handlingarna i spelet. Då och då dör en individ. Vederbörande ersätts då av en nyfödd individ. I och med den darwinistiska principen om att de bäst anpassade överlever (*survival of the fittest*) antar vi att sannolikheten för att den nyfödda individen ska vara programmerad att följa en viss strategi är proportionell mot hur hög förväntad nytta strategin ger. Mer exakt kan man specificera reproduktionen på ett sådant sätt att denna evolutionära process beskrivs med precis samma ekvationer som den tidigare beskrivna

10 Notera likheten med det evolutionsbiologiska perspektivet på beslutsfattande som Jessica Abbott beskriver i sin uppsats i detta nummer.

11 Björnerstedt & Weibull (1996) samt Weibull (1995) visar att om den beskrivna inlärningsprocessen följs av individerna i populationen så kommer populationsandelarna att utvecklas i enlighet med den så kallade replikatordynamiken. Olika former av imitation, med åtföljande dynamiska konsekvenser, undersöks även av Schlag (1998).

12 Till exempel visar Börgers & Sarin (1998) att om individer lär sig genom positiv och negativ förstärkning när de provar olika strategier så kommer populationen att utvecklas enligt replikatordynamiken. Det finns självfallet många inlärningsregler som inte fångas av denna dynamik; se Fudenberg & Levine (1998) och Sandholm (2010).

inlärningsprocessen.[13] I ett appendix går jag igenom mer exakt på vilket sätt matematiken för biologisk evolution och social inlärning kan sammanfalla.

HJORTJAKT

Låt oss nu tillämpa vår enkla inlärningsmodell på hjortjaktsspelet som beskrevs ovan. För att kunna beskriva hur beteenden ändras över tid använder vi oss av en funktion $x(t)$ som för varje tidpunkt t anger hur stor andel av populationen som följer hjortstrategin.

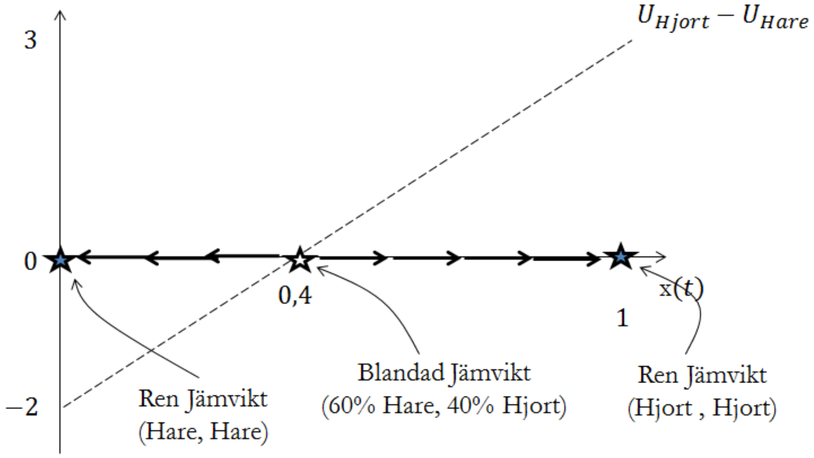
Eftersom alla individer paras ihop slumpvis för att spela mot varandra så är sannolikheten att möta en individ som spelar hjort lika med $x(t)$. Om man själv spelar hjort så kommer man därför att uppnå den förväntade nyttan $U_{Hjort} = 6x(t)$. Om man istället spelar hare är den förväntade nyttan $U_{Hare} = 2 + x(t)$. Andelen hjortspelare ökar om $U_{Hjort} > U_{Hare}$, och minskar om $U_{Hjort} < U_{Hare}$. Om vi jämför dessa nyttor så ser vi att hjortspelarna tjänar mer än harespelarna om och endast om $x(t) > 0{,}4$, med andra ord om och endast om andelen som spelar hjort är större än 40 procent. Med vår inlärningsmodell betyder det att andelen hjortspelare ökar om det redan är fler än 40 procent som spelar hjortstrategin och minskar om det är färre än 40 procent som gör det.

För att tydliggöra vad detta innebär kan vi illustrera med figur 3. Eftersom $x(t)$ står för andelen hjortspelare är $x(t)$ en siffra mellan noll och ett. Eftersom andelen harespelare är $1 - x(t)$ räcker det med att ange siffran $x(t)$ för att fullständigt beskriva i vilket tillstånd populationen befinner sig. Intervallet mellan noll och ett representerar därför de möjliga populationstillstånden. När $x(t) = 0$ är alla individer harespelare, vilket betyder att hare–hare-jämvikten spelas varje gång ett par möts. När $x(t) = 1$ är istället alla individer hjortspelare, vilket betyder att hjort–hjort-jämvikten spelas varje gång ett par möts.

När $x(t) = 0{,}4$ så är 40 % av populationen hjortspelare och 60 % harespelare. Det innebär att om vi slumpmässigt drar två individer från populationen för att spela så är sannolikheterna för de olika utfallen (hjort–hjort, hjort–hare, hare–hjort, och hare–hare) precis desamma som om vi hade betraktat två spelare som var för sig valde att spela hjort med sannolikheten $0{,}4$ och hare med sannolikheten $0{,}6$. Det betyder alltså att en blandad jämvikt inte bara kan användas för att beskriva en situation där spelarna väljer handlingar slumpmässigt utan även en situation där det istället är spelarna som väljs slumpmässigt. Om man tycker att slumpmässigt valda handlingar verkar orimliga i vissa interaktioner så kan man alltså ändå ha nytta av idén om en blandad jämvikt när man betraktar populationen som helhet.

13 I själva verket var det teoretiska biologer som först studerade dessa ekvationer. Replikatordynamiken introducerades av Taylor & Jonker (1978). Hofbauer & Sigmund (1988) beskriver utförligt dess biologiska tillämpning.

Figur 3. Inlärning i koordinationsspel.



I den streckade linjen i figur 3 har skillnaden i nytta mellan hjort- och hare-alternativen ritats in som en funktion av $x(t)$. Denna skillnad är $U_{Hjort} - U_{Hare} = 6x(t) - (2 + x(t)) = 5x(t) - 2$. När denna skillnad är positiv är det bättre att spela hjort än hare, och då ökar andelen som spelar hjort, det vill säga att $x(t)$ ökar. Detta visas i figuren med pilar som pekar åt höger längs den horisontella axeln. När skillnaden $U_{Hjort} - U_{Hare}$ är negativ så minskar istället andelen som spelar hjort, det vill säga att $x(t)$ minskar, vilket indikeras av pilarna som pekar åt vänster längs den horisontella axeln.

Från figur 3 blir det tydligt att om populationen initialt befinner sig i en punkt till höger om $x(t) = 0{,}4$ så ökar $x(t)$ till dess att alla är hjortspelare. Från ett utgångsläge till vänster om $x(t) = 0{,}4$ så minskar istället $x(t)$ till dess att alla är harespelare. Det innebär att den blandade jämvikten är mycket instabil. Om populationen i utgångsläget befinner sig i $x(t) = 0{,}4$ så räcker det med en liten störning som för populationen bort från $x(t) = 0{,}4$ för att inlärningsprocessen obönhörligen ska föra populationen till någon av de rena jämvikterna.[14] Det blir uppenbart hur inlärning implicit koordinerar förväntningarna om andras beteende, vilket i detta spel leder till att beteende koordineras.

En liknelse kan tydliggöra vad som menas med stabilitet i sammanhanget. Föreställ dig en boll som ligger på toppen av en kulle. En liten vindpust flyttar bollen ett stycke från toppen, bollen börjar rulla nedför kullen och hamnar till

14 Det är rimligt att tro att strategier inte alltid uppdateras på så sätt att mer framgångsrika strategier väljs framför mindre framgångsrika. Om det finns en positiv sannolikhet för att val av mindre framgångsrika strategier inträffar så kan det hända, av ren slump, att en population som befunnit sig i en jämvikt under lång tid på ganska kort tid förflyttar sig till den andra jämvikten. Man kan då studera vilken jämvikt som är svårast att lämna och lättast att hamna i. En jämvikt som är svår att på detta vis lämna sägs vara stokastiskt stabil. Teorin om stokastisk stabilitet utvecklades av Young (1993) och Kandori m.fl. (1993).

slut långt från toppen. Vi säger att bollen befann sig i ett instabilt läge eftersom en liten förändring av dess utgångsläge ledde till att den fortsatte att röra sig bort från utgångsläget. Jämför detta med situationen för en boll som ligger i botten av en sänka. Om vi knuffar bollen lite grand åt något håll så kommer sluttningen i sänkan göra att bollen rullar tillbaka till utgångsläget när vi släpper den. Vi säger då att bollen befann sig i ett stabilt läge. Den blandade jämvikten är instabil som bollen på toppen av kullen – en liten förändring fortplantar sig och tar oss långt bort från utgångsläget. De båda rena jämvikterna är däremot stabila som bollen i botten av sänkan – efter en liten förändring kommer vi tillbaka till utgångsläget.

HÖK–DUVA

I hjortjaktsspelet fann vi att den blandade jämvikten är instabil och att de rena jämvikterna är stabila. Vi ska nu se att slutsatsen blir annorlunda i hök–duva-spelet. Låt $x(t)$ beteckna andelen hökspelare vid tidpunkten t, så att $x(t)$ är sannolikheten att möta en individ som spelar hök.

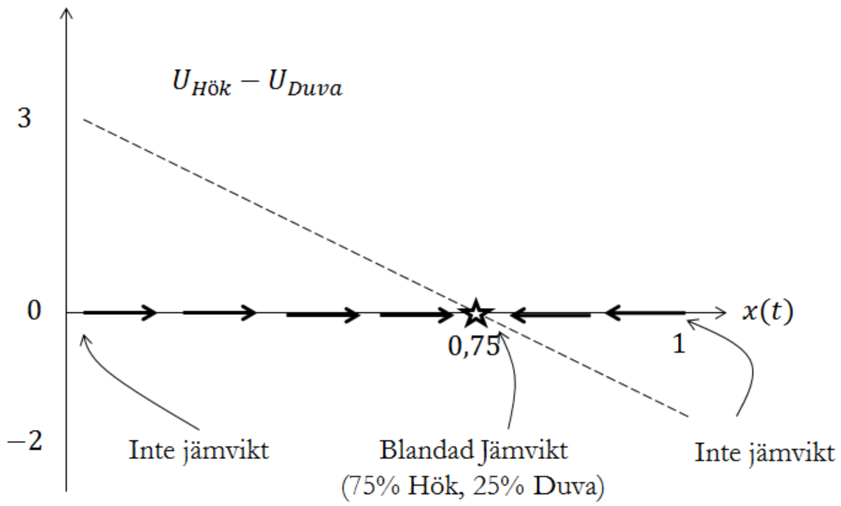
När $x(t) = 0{,}75$ beter sig populationen som i den blandade jämvikten. Om vi slumpmässigt drar två individer från populationen för att spela så är sannolikheterna för de olika utfallen (hök-hök, hök-duva, duva-hök, och duva-duva) samma när två spelare var för sig spelar hök med sannolikheten $0{,}75$ och duva med sannolikheten $0{,}25$. Liksom i koordinationsspelet ovan kan vi här tolka den blandade jämvikten som en situation där det inte är spelarna som väljer handlingar slumpmässigt utan istället spelarna som väljs slumpmässigt.

Det finns däremot inget värde på $x(t)$, alltså ingen blandning av hök- och duvspelare, som svarar mot de båda asymmetriska jämvikterna (duva-hök och hök-duva). Anledningen till detta är att dessa jämvikter förutsätter att man kan välja olika strategier beroende på om man är en kolumn- eller radspelare. Eftersom alla spelare i populationen slumpmässigt tilldelas roller som rad- eller kolumnspelare kan den asymmetriska jämvikten inte spelas. Som vi skall se nedan ändras detta om vi hade två populationer och drog radspelarna från den ena populationen och kolumnspelarna från den andra populationen.

Den förväntade nyttan av att spela hök är $U_{Hök} = 6 - 7x(t)$ och den förväntade nyttan av att spela duva är $U_{Duva} = 3 - 3x(t)$. Hökspelarna tjänar mer än duvspelarna, och ökar därmed som andel av befolkningen, om och endast om $x(t) < 0{,}75$. Det betyder att $x(t)$ växer om $x(t) < 0{,}75$ och minskar om $x(t) > 0{,}75$. Detta illustreras i figur 4. Där har skillnaden $U_{Hök} - U_{Duva}$ ritats in som en funktion av $x(t)$. När skillnaden är positiv så ökar $x(t)$, vilket indikeras av pilarna som pekar åt höger längs den horisontella axeln. När skillnaden istället är negativ så ökar $x(t)$, vilket indikeras av pilarna som pekar åt vänster längs den horisontella axeln. Vi ser att inlärningsprocessen för populationen mot den blandade jämvikten. Till skillnad från koordinationsspelet ovan är den blandade jämvikten här stabil. Om populationen ursprungligen befinner sig i punkten

$x(t) = 0{,}75$ och sedan störs lite grand, så att hamnar över eller under 0,75, kommer inlärningsprocessen så småningom att föra tillbaka populationen till $x(t) = 0{,}75$.[15]

Figur 4. Inlärning i hök och duva-spelet.



HÖK–DUVA MED TVÅ POPULATIONER

En allmän lärdom av analysen ovan är att studiet av inlärning hjälper oss skilja på vilka jämvikter som det är rimligt och orimligt att förvänta sig. Detaljerna i inlärningsmodellen kan dock spela roll för slutsatserna. Hittills har alla individer som skall spela ett spel dragits från en och samma population. Vi skall nu se vad som sker om vi istället har två populationer och drar radspelarna från den ena populationen (radpopulationen) och kolumnspelarna från den andra populationen (kolumnpopulationen). Låt $x(t)$ beteckna andelen hökspelare i kolumnpopulationen vid tidpunkten t, så att $x(t)$ är sannolikheten för en radspelare att möta en individ som spelar hök. Låt $y(t)$ beteckna andelen hökspelare i radpopulationen vid tidpunkten t.

Vi modifierar tumregeln för inlärning på så sätt att en individ som funderar på att byta strategi jämför sig med en slumpvist vald annan individ från *samma population som individen själv*. Om den individen följer en strategi som ger en

15 Oprea m.fl. (2011) studerar hök-duva i ett experiment där försöksdeltagarna upprepade gånger får välja mellan hök och duva och belönas med pengar i enlighet med siffrorna i spelmatrisen, beroende på hur de andra försökspersonerna beter sig. Forskarna finner att populationen av försöksdelatagare så småningom kommer mycket nära den jämvikt som förutsägs av teorin.

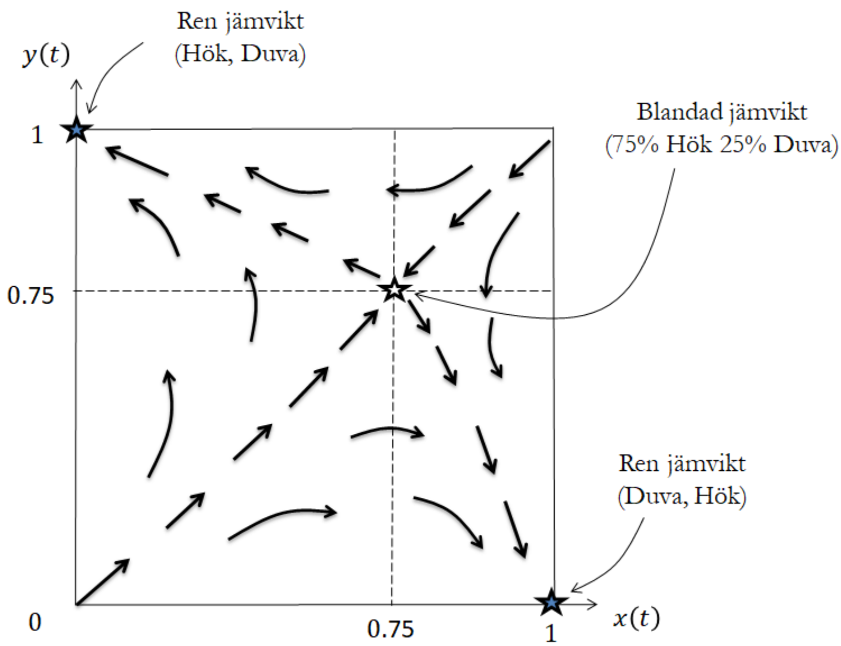
högre förväntad nytta så är den jämförande individen benägen att byta till den mer framgångsrika strategin.

Den förväntade nyttan av att spela hök för en radspelare är $U_{Hök} = 6 - 7x(t)$ och den förväntade nyttan av att spela duva är $U_{Duva} = 3 - 3x(t)$. Hökspelarna i radpopulationen tjänar mer än duvspelarna i radpopulationen, och ökar därmed som andel av radpopulation, om och endast om $x(t) < 0{,}75$. På motsvarande sätt tjänar hökspelarna i kolumnpopulationen mer än duvspelarna i kolumnpopulationen, och ökar därmed som andel av kolumnpopulation, om och endast om $y(t) < 0{,}75$.

Genom att gå igenom de fyra olika möjligheterna som ges av huruvida $x(t)$ eller $y(t)$ är större eller mindre än $0{,}75$, och jämföra vilken strategi som ger mest i vardera populationen, kan vi skapa oss en bild av hur andelen hök- och duvspelare förändras inom endera populationen som en funktion av sammansättningen av den andra populationen. Ett sådant, så kallat fasdiagram, återfinns i figur 5.

Den horisontella axeln mäter andelen hökspelare i kolumnpopulationen och den vertikala axeln mäter andelen hökspelare i radpopulationen. De tjocka svarta pilarna indikerar hur populationsandelarna rör sig över tid. Om vi startar på diagonalen så förs vi mot den blandade jämvikten men om vi startar någon annanstans så leder inlärningsprocessen istället mot någon av de rena jämvikterna. Vilken av de rena jämvikterna vi hamnar i beror på vilken sida av diagonalen vi börjar. Till skillnad från i fallet med en enda population är den blandade jämvikten här instabil medan de båda rena jämvikterna är stabila.

Figur 5. Inlärning i hök och duva-spelet när rad- och kolumnspelare kommer från var sin population.



KÖNSROLLER I "DE FRIVILLIGAS DILEMMA"

Låt oss nu något spekulativt tolka analysen med två populationer i det abstrakta hök-duva-spelet i termer av något så konkret som könsroller. Vi antar då att kvinnor imiterar, och jämför sig med, andra kvinnor, medan män imiterar, och jämför sig med, andra män.

Betrakta följande exempel. Två medarbetare delar arbetsplats. Kopiatorn har gått sönder och för att få den lagad måste någon göra en felanmälan via leverantörens krångliga webformulär. Det räcker med att en person gör felanmälan för att maskinen skall lagas. Vi kan generellt tänka på en grupp av två personer där det räcker att *en* person utför uppgiften för att problemet skall vara löst. Om båda anmäler sig frivilligt så singlar man slant om vem som skall utföra uppgiften. Var och en av medarbetarna värderar det tre olika utfallen enligt följande. Att själv utföra den tråkiga uppgiften ger nytta 2. Att den andre medarbetaren utför den tråkiga uppgiften ger nytta 4. Att ingen utför den tråkiga uppgiften ger nytta 1. Vi kan kalla spelet, som sammanfattas i figur 6, för *de frivilligas dilemma*. Notera att nyttan när båda erbjuder sig att hjälpa till är $(4 + 2)/2 = 3$ för var och en, eftersom man i detta fall singlar slant om vem som skall utföra arbetsuppgiften.

Figur 6. De frivilligas dilemma.

	Inte hjälpa till	**Hjälpa till**
Inte hjälpa till	1 , 1	4 , 2
Hjälpa till	2 , 4	3 , 3

Notera den strukturella likheten med hök-duva-spelet. Det finns två Nashjämvikter i rena strategier: en där radspelaren hjälper till och kolumnspelaren inte hjälper till, och en annan där kolumnspelaren hjälper till och radspelaren inte hjälper till. Den blandade jämvikten består i att båda spelare lägger exakt lika stor vikt vid att hjälpa till som att inte hjälpa till.[16] Inlärning, av den sort vi diskuterat, skulle i de frivilligas dilemma ge upphov till samma fasdiagram som för hök-duva-spelet i figur 5, med den enda skillnaden att den instabila, blandade jämvikten ligger vid $(x , y) = (1/2 , 1/2)$.

Babcock m.fl. (2017) genomförde ett experiment som är snarlikt det jag beskrev som de frivilligas dilemma ovan.[17] En skillnad var att grupperna bestod av tre medarbetare. En annan skillnad var att man kunde avvakta och se om de

16 Notera att $U_{inte\ hjälpa} = q + 4(1 - q) = 2q + 3(1 - q) = U_{Hjälpa}$ implicerar $q = 1/2$.

17 Holm (2000) är en tidig studie med ett liknande tema.

andra skulle hjälpa till. Om ingen anmält sig frivilligt inom två minuter räknades det som att ingen hjälpte till.

Experimentdeltagarna var helt anonyma för varandra när de fattade sina beslut sittande framför datorskärmar i en stor sal. Innan man gick in i experimentsalen kunde man dock se vilka andra som deltog i experimentet. I vissa sessioner var det en jämn könsfördelning medan det i andra ingick enbart män eller enbart kvinnor. Experimentet gick ut på att jämföra hur beteendet skilde sig beroende på denna könssammansättning. Man fann att män i grupper som bestod av enbart män var *lika* benägna att ställa upp och hjälpa till som kvinnor i grupper som bestod av enbart kvinnor. I könsblandade grupper var män däremot *mindre* benägna att ställa upp och hjälpa än kvinnor.

Det ligger nära till hands att tolka resultaten som att män och kvinnor tycker lika illa om att utföra uppgiften, men att både män och kvinnor förväntar sig att kvinnor i allmänhet hjälper till och att män i allmänhet låter bli. Dessa förväntningar om könsroller är självuppfyllande i den meningen att det är rationellt för kvinnor att ställa upp och rationellt för män att låta bli (beteendena i fråga utgör en Nashjämvikt).[18]

Analysen av inlärning i de frivilligas dilemma (vilken som sagt är kvalitativt identisk med den för analysen av konfliktspelet hök-duva i figur 5) indikerar att könsrollerna kan vara mycket stabila och svåra att ändra. Låt oss anta att kvinnorna utgör kolumnpopulationen och männen radpopulationen. Jämvikten där kvinnor hjälper till och män låter bli motsvarar då den punkt i diagrammet där $(x, y) = (0, 1)$. För att inte halka tillbaka till denna jämvikt skulle man i ett enda svep behöva ändra beteendet hos en så stor dela av männen och kvinnorna att man hamnade på en punkt under diagonalen (45-graderslinjen). Om man lyckades med detta skulle dock inlärningsdynamiken (i denna enkla modell) föra oss till en punkt där männen alltid hjälpte till och kvinnorna aldrig hjälpte till. Det enda som skulle kunna leda fram till ett stabilt jämställt läge där både män och kvinnor turades om vore ifall män och kvinnor slutade bete sig som två separata populationer antas göra i modellen. Det vill säga, det som krävs är att kvinnor och män blir könsblinda så att båda könen imiterar, och jämför sig med, både kvinnor och män, när de uppdaterar sitt val av strategi.

18 Mer allmänt kan det vara fruktbart att tolka sociala normer som jämvikter i spel med flera jämvikter. När normen är en jämvikt så har spelarna inga incitament att bryta normen så länge de förväntar sig att andra följder den. Om detta säger Binmore & Samuelson (1994) följande: "In consequence, one should make no sharp distinction between *homo economicus* and *homo sociologicus*. In using a social norm in a situation to which it is well adapted, *homo sociologicus* behaves as though he were optimizing. Similarly, when optimizing, *homo economicus* behaves as though he were employing a social norm that is well adapted to his problem".

Inlärning och upprepade spel

ETT SOCIALT DILEMMA: FÅNGARNAS DILEMMA

Det kanske mest kända exemplet från spelteorin är fångarnas dilemma. Spelet beskriver en situation där deltagarna skulle tjäna på att hjälpas åt, men där var och en av deltagarna föredrar att själva låta bli att bidra. Låt oss strunta i hur fångarnas dilemma fått sitt namn och istället tänka oss följande situation. Ask och Embla arbetar med ett projekt tillsammans. Båda har två handlingsalternativ, eller strategier, nämligen att anstränga sig eller att lata sig. Båda föredrar att någon anstränger sig så att projektet blir färdigt i tid. Men båda vill helst slippa anstränga sig och vill istället åka snålskjuts på den andres ansträngning. Problemet är att om ingen anstränger sig blir projektet inte klart i tid. Ask och Embla har därför motstridiga intressen. Vi kan sammanfatta dessa preferenser med matrisen nedan i figur 7.

Figur 7. Socialt dilemma: fångarnas dilemma.

	Anstränga sig	**Lata sig**
Anstränga sig	3 , 3	0 , 5
Lata sig	5 , 0	1 , 1

Oavsett vad Ask gör så når Embla en högre nytta genom att lata sig (eftersom 5 är större än 3 och 1 är större än 0). På samma vis är det bästa Ask kan göra att låta bli att anstränga sig, oavsett vad Embla gör. Spelet har därmed en enda Nashjämvikt, vilken består i att båda latar sig. Det hade varit bättre för båda spelarna om de båda valt att anstränga sig, eftersom de då fått 3 vardera istället för 1 vardera, vilket blir fallet om båda latar sig. Det sociala dilemmat består i att det kollektivt fördelaktiga utfallet (att båda anstränger sig) inte är förenligt med att var och en av spelarna beter sig (individuellt) rationellt.[19]

Vad säger vår inlärningsmodell om fångarnas dilemma? Låt $x(t)$ vara andelen samarbetsspelare vid tidpunkten t. Den förväntade nyttan av att spela samarbete (anstränga sig) är $U_{Samarbete} = 3x(t)$ och den förväntade nyttan av att spela icke-samarbete (lata sig) är $U_{Icke\text{-}samarbete} = 4x(t)+ 1$. De som inte samarbetar har därför högre förväntad nytta än de som samarbetar, oavsett vilket värde $x(t)$ tar. Det betyder att $x(t)$ alltid rör sig mot noll. Populationen konvergerar mot den enda jämvikten – ingen samarbetar.

19 Vi skulle kunna ha antagit att Ask och Embla är mera generösa och föredrar att städa om den andra personen städar, exempelvis för att de får dåligt samvete annars. Då hade vi kunnat byta ut siffran 5 mot en lägre siffra och få ett annat spel, med andra jämvikter. Men givet våra antaganden, som de sammanfattas i figur 7, är det ovedersägligen så att om spelarna är rationella så kommer de att välja att lata sig och det leder till ett utfall som är sämre än om båda hade ansträngt sig.

UPPREPADE SPEL

I beskrivningen av hur inlärning går till antogs att individer matchas ihop med en ny motspelare i varje period. Eftersom befolkningen förutsätts vara stor så är sannolikheten liten för att två individer ska mötas snart igen. När individerna möts väldigt sällan kan de låta blir att bry sig om de framtida konsekvenserna av sina handlingar. Detta är avgörande för vår slutsats att samarbete inte kan överleva i fångarnas dilemma.

Låt oss nu undersöka vad som händer om individerna spelar ett spel flera gånger med samma motpart. Vi modellerar detta genom att anta att det, varje gång två spelar möts, finns möjlighet att de spelar fångarnas dilemma flera gånger med varandra. Mer exakt tänker vi oss att de två matchade individerna med säkerhet spelar fångarnas dilemma en gång. Därefter spelar de en andra gång med sannolikheten δ. Om de spelar fångarnas dilemma en andra gång så spelar de sedan en tredje gång med sannolikheten δ. Varje gång de spelat en omgång så är sannolikheten δ att de ska spela ytterligare en gång. Vi säger att de spelar ett obestämt upprepat spel, med upprepningssannolikheten δ.

Det finns oändligt många sätt att betinga sitt val mellan samarbete och icke-samarbete på vad som hänt hittills i ett obestämt upprepat spel. För enkelhets skull begränsar vi oss till att undersöka två enkla tumregler, eller strategier, för det obestämt upprepade fångarnas dilemma-spelet.

- *Aldrig samarbeta.* Som namnet anger väljer en individ som följer denna strategi aldrig att samarbeta.
- *Samarbete tills vidare, sedan aldrig mer ("Grim trigger")* . En individ som följer denna strategi börjar med att samarbeta och fortsätter att samarbeta så länge som alla andra hittills alltid samarbetat. Om någon inte samarbetar så byter man till att aldrig mera samarbeta.[20]

När två som aldrig samarbetar möts för att spela fångarnas dilemma med de nyttor som anges i figur 7 så tjänar de 1 i varje spelomgång. Den förväntade nyttan är därför $1 + \delta + \delta^2 + \delta^3 + \cdots$. Man kan visa att detta är lika med $1/(1-\delta)$. När två som spelar *samarbete tills vidare*-strategin möts samarbetar de i varje omgång de möts (eftersom de båda börjar med att samarbeta och fortsätter därmed så länge ingen gjort något annat). Den förväntade nyttan är därför $3 + 3\delta + 3\delta^2 + 3\delta^3 + \cdots = 3/(1-\delta)$. När en som följer strategin *samarbete tills vidare* möter en som följer strategin *aldrig samarbeta* så samarbetar *samarbete tills vidare*-spelaren endast i första perioden. Den förväntade nyttan för *samarbete tills vidare*-spelaren är därför $0 + \delta + \delta^2 + \delta^3 + \cdots = \delta/(1-\delta)$, och den förväntade nyttan för den andre är $5 + \delta + \delta^2 + \delta^3 + \cdots = 5 + \delta/(1-\delta)$. Vi

20 Resultaten skulle vara desamma om vi istället använde strategin tit-for tat, berömd från Axelrods turnering.

sammanfattar detta i figur 8(a). Om vi sätter $\delta = 2/3$ så erhåller vi spelmatrisen i figur 8(b)

Figur 8. Upprepat fångarnas dilemma.

(a) Upprepningssannolikhet δ

	Samarbete tills vidare	Aldrig samarbeta
Samarbete tills vidare	$\frac{3}{1-\delta}$, $\frac{3}{1-\delta}$	$\frac{\delta}{1-\delta}$, $5+\frac{\delta}{1-\delta}$
Aldrig samarbeta	$5+\frac{\delta}{1-\delta}$, $\frac{\delta}{1-\delta}$	$\frac{3}{1-\delta}$, $\frac{3}{1-\delta}$

(b) Upprepningssannolikhet δ = 2/3

	Samarbete tills vidare	Aldrig samarbeta
Samarbete tills vidare	9 , 9	2 , 7
Aldrig samarbeta	7 , 2	3 , 3

Om vi studerar matrisen i figur 8(b) närmare, är det lätt att se att den har en struktur som påminner om koordinationsspelet i figur 2. Det bästa man kan göra är att välja samma strategi som sin motspelare. Därför finns det två rena jämvikter i spelet, där man kan välja mellan att följa *samarbete tills vidare* eller *aldrig samarbeta* i ett obestämt upprepat spel med upprepningssannolikheten $\delta = 2/3$. I den ena jämvikten spelas bara *samarbete tills vidare*, vilket betyder att alla samarbetar. I den andra samarbetar ingen. Om vi jämför med det icke-upprepade spelet, det vill säga det vanliga fångarnas dilemma, så ser vi att jämvikten utan samarbete visserligen finns kvar som en möjlighet, men att det på grund av upprepningen också tillkommit en möjlighet att samarbeta. Detta sker genom att samarbete idag belönas med samarbete imorgon och att icke-samarbete idag straffas med icke-samarbete imorgon.[21] Samarbetet förutsätter här ingen omtanke om motparten; genom att betinga framtida beteende på dagens beteende skapar spelarna ömsesidiga incitament att samarbeta.

21 Denna insikt, att obestämd upprepning skapar möjligheter till samarbete, går under benämningen folkteoremet i spelteorin, eftersom den var en del av spelteorins ”folkvisdom” innan resultatet formaliserades. Den introducerades i den evolutionsbiologiska litteraturen av Trivers (1971). Studiet av evolution i det upprepade fångarnas dilemma-spelet stimulerades kanske särskilt av Axelrods (1984) datorturneringar. Man kan även studera scenarier där individer inte spelar upprepade fångarnas dilemma i par utan varje omgång möter en ny motpart i fångarnas dilemma. Även under dessa omständigheter kan evolution leda till samarbete, ifall spelarna får veta något om hur motparten brukar spela, så att spelarna har möjlighet att skapa ett visst rykte om sig. För analyser av detta scenario, se Kandori (1992), Nowak & Sigmund (1998) och Heller & Mohlin (2016).

Liksom i koordinationsspelet så är de båda rena jämvikterna i det upprepade fångarnas dilemma-spelet stabila. Vi kan tänka på det som två olika normer som kan råda i samhället. I den ena är alla optimistiska och förtroendefulla – de räknar med att motparten samarbetar och eftersom det är precis vad som sker så fortsätter alla att samarbeta och vara optimistiska. I den andra jämvikten är alla pessimistiska och misstänksamma – de utgår helt korrekt från att ingen samarbetar och har därför inga skäl att samarbeta.[22] Vilken norm som blir rådande beror, enligt den enkla inlärningsmodell vi studerat, på utgångsläget. Om samhället till att börja med mestadels består av dem som spelar *samarbete tills vidare* så kommer de att ta över, men om samhället innehåller en tillräckligt stor andel som spelar *aldrig samarbeta* så kommer istället dessa att ta över. [23]

Inlärning som inte leder till jämvikt

I alla spel vi har studerat hittills leder inlärningsprocessen till att en jämvikt spelas. I koordinationsspelet spelas visserligen inte den blandade jämvikten och i hök–duva-spelet inte de rena jämvikterna (när det bara finns en population), men någon av jämvikterna spelas i alla dessa spel. Man kan fråga sig om det alltid är så att den inlärningsprocess vi beskrivit leder till en jämvikt. Det vore ju bra om det vore så eftersom så mycket av ekonomisk teori förutsätter jämvikt. Tyvärr är det inte så.[24]

Du har antagligen spelat sten, sax och påse som barn. Det är ett spel med två spelare där påse vinner mot sten som vinner över sax som i sin tur vinner över påse. Låt oss anta att en vinst ger nytta 1, en förlust ger nytta −1 och om det är oavgjort så får man nyttan 0. Figur 9 avbildar spelmatrisen för detta spel.

22 Det bör noteras att koordinationsproblemen grovt underskattas av att vi begränsar oss till två strategier i det upprepade spelet. Det finns som sagt oändligt många strategier i det upprepade fångarnas dilemma-spelet. Om vi betraktar sociala dilemman med fler än två handlingar, som svarar mot olika grader av samarbete, kanske med mindre symmetriska utfall, växer koordinationsproblemen ytterligare.

23 Litteraturen om evolution av samarbete, särskilt i fångarnas dilemma, är mycket omfångsrik. För en introduktion, se exempelvis Novak (2006). En översikt över den experimentella forskningen rörande inlärning i det upprepade fångarnas dilemma-spelet ges av Dal Bó & Fréchette (2018).

24 Man kan fråga sig om det åtminstone är så att inlärning leder till att dominerade strategier, eller iterativt dominerade strategier inte används. Inte heller detta är generellt sant. Se Hofbauer & Sandholm (2011) och Bernergård & Mohlin (2019) samt referenser däruti.

Figur 9. Sten, sax och påse.

	Sten	Påse	Sax
Sten	0 , 0	-1 , 1	1 , -1
Påse	1 , -1	0 , 0	-1 , 1
Sax	-1 , 1	1 , -1	0 , 0

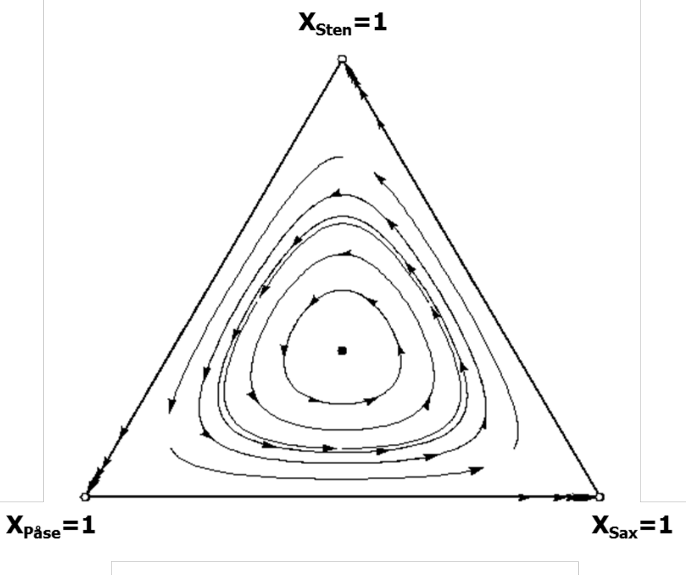
Detta spel har en enda Nashjämvikt. I denna jämvikt spelar båda spelare alla sina tre strategier med lika stor sannolikhet. En spelare väljer alltså vart och ett av sina handlingsalternativ med sannolikhet 1/3.

Vad händer om detta spel spelas i en population av individer, som anpassar sig med hjälp av den tumregel för inlärning som vi arbetat med hittills? Man kan visa, att populationen aldrig kommer att röra sig mot jämvikt. Detta illustreras i figur 10. Varje punkt i triangeln står för en blandad strategi: en kombination av strategierna sten, påse och sax. I triangelns översta hörn spelar alla sten, i det nedre vänstra hörnet spelar alla påse och i det nedre högra hörnet spelar alla sax. Pilarna visar att populationen beter sig cykliskt. Vi ser att var vi än börjar så rör vi oss så småningom tillbaka mot ursprungspunkten.

Om de flesta spelar sten så tjänar de som istället spelar påse mest och växer som andel av befolkningen. När de flesta börjat spelar påse så tjänar de som spelar sax mest och växer som andel av befolkningen. När de flesta spelar sax så är det de som spelar sten som tjänar mest och på så vis återgår populationen till utgångsläget där de flesta spelade sten.[25]

25 Inlärning i sten, sax, och påse har studerats experimentellt av Cason m.fl. (2014). De finner stöd för att beteende cirkulerar på det sätt som förutsägs av teorin. Det här spelet verkar dessutom ha biologisk relevants: hanar av ödlearten *Uta stansburiana* ("side-blotched lizars") använder sig av tre reproduktionsstrategier som står i samma relation till varandra som strategierna i sten, sax och påse (Sinervo m.fl 1996). De olika hanarna har olika färg på halsen. Orangehalsade hanar kontrollerar stora territorier med flera honor. Gulhalsade hanar håller till i utkanten av de orange-halsades territorier och smyger in och parar sig med honor på de orangea hanarnas territorier när tillfälle ges. Blåhalsade hanar vaktar en enskild hona. De kan försvara sig mot intrång från gulhalsade hanar men inte mot de orangehalsade. Om populationen främst består av orangehalsade hanar är det fördelaktigt att vara gulhalsad. Om populationen istället domineras av gulhalsade hanar är det fördelaktigt att vara blåhalsad. Om de blåhalsade dominerar är det fördelaktigt att vara orangehalsad (bilden kompliceras sedan ytterligare av att honorna har två olika strategier).

Figur 10. Inlärning i sten, sax och påse.[26]



En fallstudie

Fram till nu har vår diskussion huvudsakligen varit teoretisk, med vissa hänvisningar till laboratorie-experiment. Vi skall nu titta närmare på en studie av inlärning utanför laboratoriet (Mohlin m.fl. 2020, se även Östling m.fl. 2011).[27]

Under 49 dagar våren 2007 drev företaget Svenska Spel ett lotteri som de kallade Limbo. Något förenklat var spelet uppbyggt kring följande regler. Varje dag fick alla som ville delta. Deltagarna betalade en avgift och fick sedan välja ett nummer mellan 1 och 99 999. När alla hade gjort det identifierade man det enda – och lägsta – nummer som hade valts av endast en enda person. Vederbörande vann en stor summa pengar (i storleksordningen 100 000 kronor). Spelarna kände till dessa regler. I genomsnitt deltog mer än 50 000 spelare varje dag. Det går att bevisa att spelet har en unik jämvikt. I denna följer alla spelare samma blandade strategi, vilken uttrycks som den streckade linjen i figur 11(a) och 11(b). Denna blandade strategi lägger positiv sannolikhet på alla nummer mellan 0 och 99 999. Sannolikheten för att välja höga nummer är dock lägre

26 This phase diagram was created using the sofware Dynamo (Sandholm & Dokumaci 2007).

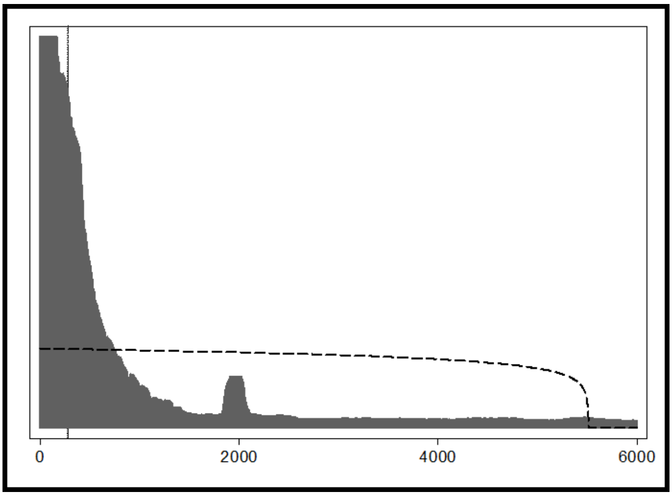
27 Att studera inlärning i spel utanför laboratoriet är svårt, bland annat eftersom man måste kunna avgränsa vad deltagarna uppfattar som det relevanta spelet. Därför är studier baserade på helt naturliga inlärningssituationer ovanliga. Doraszelski m.fl. (2018) studerar dock den engelska energimarknaden och finner stöd för inlärningsdynamiker som är besläktade med dem vi diskuterat här.

än sannolikheten för att välja låga nummer. Sannolikheten för att välja de första 5 500 numren är avsevärt högre än sannolikheten för att välja högre nummer.

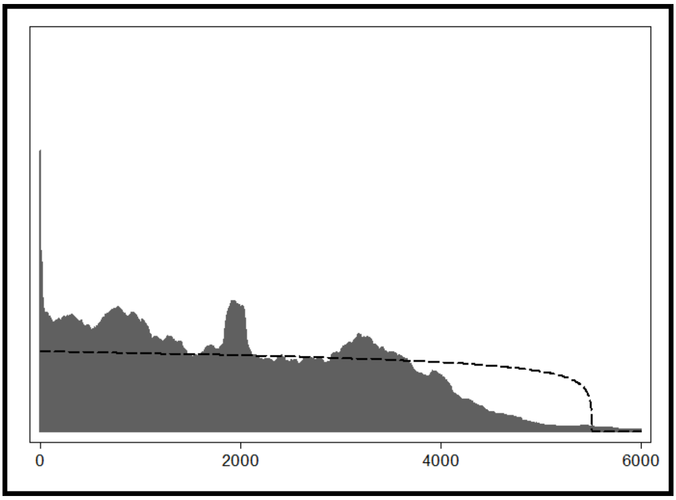
Det verkar orimligt att anta att spelarna skulle kunna resonera sig fram till vad jämvikten är. Och även om de hade kunnat det, så ligger det inte i deras intresse att bete sig i enlighet med jämvikten om de inte också tror, att alla andra kommer att bete sig på samma sätt. I figur 11(a) ser vi mycket riktigt att fördelningen av nummer som valdes den andra dagen (den grå massan i figur 11) av spelet avviker kraftigt från jämvikten. Men i figur 11(b) ser vi att förhållandet förändras avsevärt fram till den sista gången spelet genomfördes. Fördelningen av valda nummer under dag 49 ligger betydligt närmare jämvikten, även om det är en bit kvar till full jämvikt.

Figur 11. Inlärning och jämvikt i Limbo-spelet.

(a) Beteende dag 2



(b) Beteende dag 49 (sista dagen)

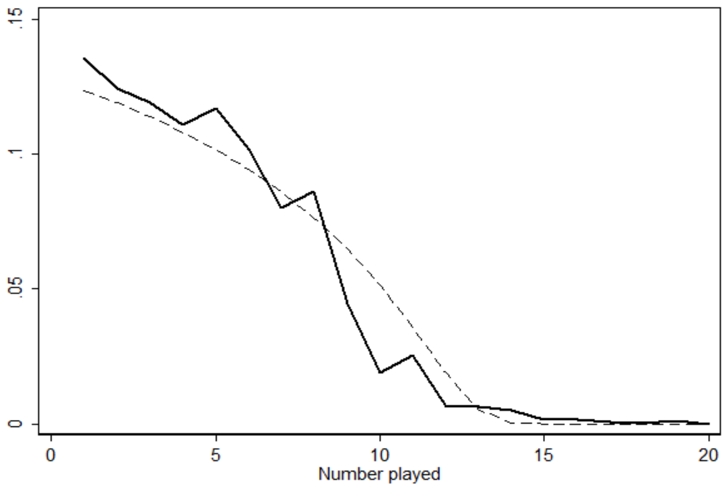


Hur går detta till? Spelarna har tillgång till information om vilka nummer som vunnit tidigare. Det verkar som att spelarna använder denna information som en indikation på vilka nummer som kan tänkas vinna i framtiden. De väljer nämligen nummer som ligger nära tidigare vinnare, och de är mer benägna att välja ett nummer, ju närmare de tidigare vinnarna detta nummer ligger.

För att undersöka samma spel under mera kontrollerade former, genomfördes ett experiment där studenter rekryterades och fick spela en version av Limbo-spelet på datorer 49 gånger i rad. Antalet nummer att välja mellan minskades från 99 999 till 99 och det genomsnittliga antalet deltagare minskade från över 50 000 per dag till i genomsnitt 26 stycken per omgång. Prissumman för varje omgång var 7 dollar (ungefär 50 kronor). Den enda information deltagarna hade var vilka nummer som hade vunnit i förgående spelomgångar.

Figur 12 jämför den teoretiska jämvikten (streckad linje) med det faktiska beteendet (svart linje) efter en period av inlärning som motsvarar de första 20 omgångarna. Vi ser att inlärning fört beteendet mycket nära jämvikten. Att experimentet använder färre nummer och ett mindre antal spelare, som uteslutande fokuserar på just detta spel under en stund i en datorsal, verkar leda till snabbare inlärning – det krävs färre omgångar av spelet för att beteendet skall hamna mycket nära jämviktstillståndet.

Figur 12. Experimentversionen av Limbo-spelet.



Komplikationer

EXTRAPOLERING MELLAN SPEL

I vår modell av inlärning har vi utgått från att individer från en population möts för att spela ett och samma spel gång på gång. De lär sig spela detta spel utan hänsyn till erfarenheter från andra spel. I verkligheten erfar vi sällan exakt samma typ av interaktion mer än en gång. Istället är vi delaktiga i ett otal mer eller mindre olika spel samtidigt. De innebär at våra möjligheter att genom inlärning hamna i jämvikt i ett enskilt spel beror på hur vi extrapolerar från erfarenheter i tillräckligt likartade spel. Även om spelteoretiker förstår detta så har det nästan inte alls studerats explicit.[28]

Ett intressant undantag är Grimm & Mengel (2014). De låter experimentdeltagare spela antingen två eller sex olika spel i slumpvis ordning vid sammanlagt 100 tillfällen. Varje enskilt spel har en enda Nashjämvikt som det skulle vara lätt att konvergera mot, om experimentdeltagarna endast mötte en och samma strategi under de 100 tillfällena. Att hålla reda på två olika speltyper torde heller inte vara så svårt, men att däremot försöka överblicka sex olika spel skulle kunna vara kognitivt påfrestande. De sex olika spelen kunde emellertid klassificeras i tre par, där spelen i varje par hade samma strategiska struktur. Spel inom ett par kunde därför lämpligen användas för att lära sig det andra spelet i paret.

I experimentet visade det sig att en hög andel deltagare förmådde upptäcka att de sex olika spelen väsentligen kunde behandlas som tre olika kategorier av spel, och därmed lära sig att spela jämvikt. Detta förutsatte dock att deltagarna erhöll tillräckligt rik återkoppling om resultatet i tidigare perioder. När deltagarna fick mer begränsad återkoppling lärde de sig inte att spela jämvikt.[29]

INTERAKTIONER UTAN MÖJLIGHET TILL INLÄRNING

Inte alla interaktioner ger deltagarna möjlighet att ackumulera erfarenhet som gör att de gradvis kan anpassa sitt beteende till dess att de eventuellt hamnar i en jämvikt. Många viktiga interaktioner beskrivs kanske till och med bäst som engångsföreteelser för de inblandade. Om jämviktsteorier endast är användbara för att beskriva beteende som grundats på erfarenhet och inlärning genom upprepade försök, vad skall vi använda för teori för strategiskt beteende innan inlärning ägt rum, eller i situationer där det saknas möjligheter att ackumulera tillräckliga erfarenheter att basera sina beslut på?

28 Förmodligen är huvudorsaken till att det inte studerats är att det är svårt. Dels är det svårt att bedöma vad om är en rimlig modell av hur erfarenheter extrapolerats (de teoretiska valmöjligheterna är mycket större än när man fokuserar på ett enskilt spel), dels är det svårt bestämma vad som är en realistisk klass av spel som gör inlärning ”lagom” svårt.

29 Relaterat finner Mengel & Sciubba (2014) att spelare lär sig fortare om de tidigare spelat ett spel med likartad strategisk struktur och lär sig långsammare om de tvärtom tidigare spelat ett spel med annorlunda (motsatt) strategisk struktur.

En teori som försöker göra reda för beteende innan spelare hunnit skaffa sig erfarenhet av ett spel är så kallad *level-k* teori. Teorin stipulerar att olika individer använder sig av olika tumregler som tillsammans bildar en hierarki. En individ på nivå 0 antas ofta välja likformigt slumpmässigt mellan sina strategier. En individ på nivå 1 beter sig som om den trodde att alla andra var på nivå 0, dvs en individ på nivå 1 väljer den strategi som är nyttomaximerande när motparten väljer likformigt slumpmässigt mellan sina strategier. En individ på nivå 2 beter sig som om den trodde att alla andra var på nivå 1, dvs som om alla andra trodde att alla andra var på nivå 0. Mer allmänt väljer en individ på nivå *k* en strategi som är nyttomaximerande, givet att alla andra beter sig som nivå *k-1*.[30]

Avslutning

Vi har sett att inlärning baserad på enkla tumregler i många viktiga spel leder till att en grupp individer så småningom beter sig i enlighet med teoretiska antaganden om jämvikt. Genom att studera de dynamiska processerna bakom kan vi avgöra vilka jämvikter som torde vara stabila och vilka som är instabila. Om ett spel har flera jämvikter så kan inlärning och evolution ge en vink om vilken eller vilka jämvikter som det är troligast att man kommer att röra sig mot i det långa loppet. I vissa fall leder inlärning och evolution dock inte till jämvikt. Det kan hända även om spelet har en enda jämvikt. Evolutionär spelteori och experimentella undersökningar av inlärning i spel har därmed avsevärt förbättrat vår förmåga att bedöma hållbarheten i tillämpningar av klassisk spelteori.

Litteratur

Axelrod, R., 1984. *The Evolution of Cooperation*. Basic Books.

Babcock, L., Recalde, M. P., Vesterlund, L., & Weingart, L., 2017. ”Gender differences in accepting and receiving requests for tasks with low promotability”, *American Economic Review*, 107(3), s. 714–47.

Bernergård, A., & Mohlin, E., 2019. “Evolutionary Selection against Iteratively Weakly Dominated Strategies”, *Games and Economic Behavior* 117 (Sept.), s. 82–97.

Binmore, K. & Samuelson, L., 1994. ”An economist’s perspective on the evolution of norms”, *Journal of Institutional and Theoretical Economics* 150 (1), s. 45–63.

Björnerstedt, J. & Weibull, J., 1996. “Nash equilibrium and evolution by imitation”, i K. J. Arrow, E. Colombatto, M. Perlman & C. Schmidt (red.), *The Rational Foundations of Economic Behaviour*. London: MacMillan.

30 Level-*k*-teorin introducerades av Nagel (1995) och Stahl & Wilson (1995). Camerer m.fl. (2004) utvecklade en besläktad teori kallad cognitive hierarchy. Crawford m.fl. (2013) är en grundlig genomgång av dessa modeller, såväl teori som evidens och tillämpningar. Man skulle kunna fråga sig vad som sker om fördelningen av nivåer i en population är föremål för inlärning eller evolution, se Stahl (2000) och Mohlin (2012).

Börgers, T. & Sarin, R.,1997. "Learning through reinforcement and replicator dynamics", *Journal of Economic Theory* 77(1), s. 1–14.

Camerer, C., 2003. *Behavioral Game Theory*. Princeton: Princeton UP.

Camerer, C., Ho, T-H., & Chong, J.K., 2004. "A Cognitive Hierarchy Model of Games", *Quarterly Journal of Economics* 119(3), s. 861–898.

Cason, T. N., Friedman, D., & Hopkins, E., 2014. "Cycles and Instability in a Rock–Paper–Scissors Population Game: A Continuous Time Experiment", *Review of Economic Studies*, 81(1), s. 112–136.

Crawford, V., Costa-Gomes, M., & Iriberri, N., 2013. "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications", *Journal of Economic Literature* 51 (March), s. 5–62.

Dal Bó, P. & Fréchette, G. R., 2018. "On the determinants of cooperation in infinitely repeated games: A survey", *Journal of Economic Literature*, 56(1), s. 60–114.

Doraszelski, U., Lewis, G., & Pakes, A., 2018. "Just starting out: Learning and equilibrium in a new market", *American Economic Review*, 108(3), s. 565–615.

Fudenberg, D. & Levine, D. K., 1998. *The theory of learning in games*. Cambridge MA: MIT Press.

Grimm, V. & Mengel, F., 2012. "An experiment on learning in a multiple games environment", *Journal of Economic Theory*, 147(6), s. 2220–2259.

Heller, Y., & Mohlin, E., 2017. "Observations on cooperation", *The Review of Economic Studies*, 85(4), s. 2253–2282.

Hofbauer, J. & Sandholm, W. H., 2011. "Survival of dominated strategies under evolutionary dynamics", *Theoretical Economics* 6(3), s. 341–377.

Hofbauer, J. & Sigmund, K., 1988. *The Theory of Evolution and Dynamical Systems*. Cambridge: Cambridge University Press.

Holm, H. J., 2000. "Gender-based focal points", *Games and Economic Behavior*, 32(2), s. 292–314.

Kandori, M., 1992. "Social Norms and Community Enforcement", *Review of Economic Studies* 59, s. 61–80.

Kandori, M., Mailath, G. & Rob, R., 1993. "Learning, Mutation and Long-Run Equilibria in Games", *Econometrica* 61, s. 29–56.

Mengel, F., & Sciubba, E., 2014. "Extrapolation and structural similarity in games", *Economics Letters*, 125(3), s. 381–385.

Mohlin, E., 2012. "Evolution of theories of mind", *Games and Economic Behavior*, 75(1), s. 299–318.

Mohlin, E., Östling, R. & Wang, J. T.-Y., 2020. "Learning by Similarity-weighted Imitation in Games", *Games and Economic Behavior* 120 (March), s. 225–245.

Nash, J., 1950. *Non-cooperative Games*, PhD Thesis, Princeton University.

Nagel, N., 1995. "Unraveling in Guessing Games: An Experimental Study", *American Economic Review* 85 (5), s. 1313–1326.

Nowak, M., & Sigmund, K., 1998. "Evolution of indirect reciprocity by image scoring", *Nature* 393, s. 573–577.

Nowak, M., 2006. *Evolutionary Dynamics*. Cambridge MA: Harvard Belknap Press.

Oprea, R., Henwood, K. & Friedman, D., 2011. "Separating the Hawks from the Doves: Evidence from continuous time laboratory games", *Journal of Economic Theory,* 146(6), s. 2206–2225.

Sandholm, W. H., 2010. *Population games and evolutionary dynamics*. Cambridge MA: MIT Press.

Sandholm, W. H. & Dokumaci, E., 2007. *Dynamo: Phase diagrams for evolutionary dynamics* (software suite), http://www.ssc.wisc.edu/ whs/dynamo.

Schlag, K. H., 1998. "Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits", *Journal of Economic Theory* 78(1), s. 130–156.

Stahl, D. & Wilson. P., 1995. "On Players' Models of Other Players: Theory and Experimental Evidence", *Games and Economic Behavior* 10, s. 218–254.

Stahl, D. O., 2000. "Rule learning in symmetric normal-form games: theory and evidence", *Games and Economic Behavior*, 32(1), s. 105–138.

Sinervo, B. & Lively, C. M., 1996. "The rock–paper–scissors game and the evolution of alternative male strategies", *Nature* 380, s. 240–243.

Taylor, P. D. & Jonker, L. B., 1978. "Evolutionary stable strategies and game dynamics", *Mathematical Biosciences* 40(1–2), s. 145–156.

Trivers, R.,1971. "The evolution of reciprocal altruism", *Quarterly Review of Biology*, 46, s. 35–56.

Weibull, J. W., 1995. *Evolutionary Game Theory*. Cambridge MA: MIT Press.

Young, P., 1993. "The Evolution of Conventions", *Econometrica* 61, s. 57–84.

Young, H. P., 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton NJ: Princeton University Press.

Östling, R., Wang, J. T.-Y., Chou, E. Y. & Camerer, C. F., 2011. "Testing game theory in the field: Swedish LUPI lottery games", *American Economic Journal: Microeconomics* 3(3), s. 1–33.

Appendix: Biologisk evolution och social inlärning

Vi antar en stor men ändlig population. Vid varje tidpunkt spelar varje individ en av de rena strategierna i ett givet spel. Antalet individer som spelar strateg i vid tid t är $p_i(t)$ och den totala populationens storlek är $p(t) = \Sigma_{i=1}^{k} p_i(t)$. Andelen som spelar strategi i är $x_i(t) = p_i(t)/ p(t)$. Populationens tillstånd är en vektor $x(t) = (x_1(t) , x_2 , , ... , x_k(t)$. Notera att ett populationstillstånd svarar mot en blandad strategi $x(t) \in \Delta$. Låt e^i beteckna enhetsvektorn som motsvarar den "degenererade" blandade strategin som lägger sannolikheten ett på den rena strategin i. Förväntad nytta för en som spelar strategi i mot en slumpvist vald motspelare är $u(e^i , x)$. Den genomsnittliga förväntade nyttan i populationen är $u(x , x) = \Sigma_{i=1}^{k} x_i u(e^i , x)$.

BIOLOGISK EVOLUTION

Under en biologisk tolkning är varje individ "programmerad" att spela en av de rena strategierna. Vi antar ny att det vi tidigare kallat nytta nu representerar effekter på en individs "fitness" mätt som förväntad avkomma per tidsenhet. Varje avkomma ärver sin enda förälders strategi. Låt $\beta \geq 0$ stå för bakgrunds-"fitness" (oberoende av utfallet i spelet vi studerar). Om reproduktion äger rum kontinuerligt över tid så är fertiliteten (födelsetakten) vid tidpunkt

t, för individer som är programmerade till den rena strategin i, $\beta + u(e^i, x(t))$. Låt mortaliteten $\delta \geq 0$ vara samma för alla individer. När $p(t)$ är stor nog, så kommer det genomsnittliga antalet avkomma för en individ som spelar strategi i att vara nära det förväntade antalet avkomma $\beta + u(e^i, x(t))$, och den faktiska andelen som dör kommer vara nära δ. Om vi använder punkter för att indikera derivata med avseende på tid så kan vi skriva ned den resulterande populationsdynamiken:

$$\dot{p}_i(t) = [\beta + u\,(e^i, e(t)) - \delta]\, p_i(t) = 3 - 3q$$

Notera att $p(t)x_i(t) = p_i(t)$ och derivera med avseende på tid på båda sidor.

$$\dot{p}(t)x_i(t) + p(t)\dot{x}_i(t) = \dot{p}_i(t)$$

Om vi flyttar om detta så får erhåller vi (efter lite arbete)

$$\dot{x}_i(t) = [\beta + u(e^i, x(t)) - \delta]\, x_i(t) - [\beta + u(x(t), x(t)) - \delta]\, x_i(t)$$

Om vi förenklar detta uttryck (och tar bort hänvisningen till tid) så får vi fram den så kallade *replikatordynamiken*:

$$\dot{x}_i = [u(e^i, x) - u(x, x)]x^i$$

SOCIAL INLÄRNING

Under en social tolkning följer varje individ för närvarande en av de rena strategierna. Det vi tidigare kallade nytta bör nu tolkas som ett mått på individers observerbara "framgång" per tidsenhet. Framgång kan bestå i vadhelst spelarna strävar efter – exempelvis välstånd, status eller hälsa. När en aktör erhåller en möjlighet att revidera sin strategi (för enkelhetsskulle antar vi att de följer en Poissonfördelning med parameter 1 som är oberoende av individ och tidpunkt) så väljer hon slumpmässigt en annan individ att ha som sin (potentiella) förebild. Hon observerar förebildens strategi och nytta/framgång. Om hennes förebilds nytta är högre än hennes egen nytta så imiterar hon förebilden med en sannolikhet som står i proportion mot skillnaden i nytta mellan aktören och hennes förebild. Om aktören har en högre nytta än förebilden så byter hon inte strategi. Om en aktör som för närvarande använder strategi observerar en förebild som för närvarande använder strategi så kan vi uttrycka sannolikheten för att aktören byter till förebildens strategi som

$$[u(e^j, x(t)) - u(e^i, x(t))]_+ .$$

Här indikerar det nedsänkta plustecknet att vi endast bryr oss om uttrycket inom hak-parenteserna ifall det är icke-negativt. Det följer att andelen av populationen som byter från strategi i vid tid t är

$$x_i(t)\, \Sigma_{j=1}^{k}\, x_j(t)\, [u(e^j, x\,(t)) - u(e^i, x(t))]_+ ,$$

och andelen av befolkningen som byter till vid tidpunkten är

$$\Sigma_{j=1}^{k} x_j(t)\, x_i(t)\, [u(e^i, x(t)) - u(e^j, x(t))]_+ .$$

Om vi sätter samman dess uttryck, och tar bort hänvisning till tid, så erhåller vi

$$\dot{x}_i = x_i\, \Sigma_{j=1}^{k} x_j[u(e^i, x) - u(e^j, x)]_+$$

$$-x_i\, \Sigma_{j=1}^{k} x_j[u(e^j, x) - u(e^i, x)]_+$$

$$= x_i\, \Sigma_{j=1}^{k} x_j[u(e^i, x) - u(e^j, x)]$$

$$= x_i\, u(e^i, x) - u(x, x)]$$

Återigen har vi härlett *replikatordynamiken*.