Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Jocs Repetits Infinitament

Autor: Guillem Tena Rodriguez

Director: Dr. Josep Vives i Santa-Eulàlia

Realitzat a: Universitat de Barcelona

Departament de Matemàtiques i Informàtica

Barcelona, 20 de juny de 2021

Abstract

In 1951, John Forbes Nash defined a solution concept for games with two or more players which today is known as Nash equilibrium. We will study these equilibria for games that are repeated infinitely, and we will see the proof of the existence of these solutions, made by James W. Friedman in 1971. We will focus on two of the most known games, the prisoner's dilemma and the Cournot's oligopoly. Finally, we will perform simulations that will help us to confirm the equilibrium results.

Resum

L'any 1951, John Forbes Nash va definir un concepte de solució per a jocs amb dos o més jugadors que avui dia coneixem com equilibri de Nash. En aquest treball estudiarem aquests equilibris per jocs que es repeteixen infinitament, i veurem la prova de l'existència d'aquests, feta per James W. Friedman el 1971. Ens centrarem ens dos dels jocs més coneguts, el dilema del presoner i l'oligopoli de Cournot. Finalment, realitzarem unes simulacions que ens ajudaran a confirmar els resultats d'equilibri.

2020 Mathematics Subject Classification. 91A20,91A26, 91A80

Agraïments

En primer lloc, m'agradaria agrair al meu tutor, en Josep Vives, per haver tutoritzat i enfocat el projecte com és degut. M'ha sabut guiar en tot moment i ha estat sempre disponible. Agrair també a tots els professors que he tingut, tant a la universitat com fora d'ella, per haver fomentat el desenvolupament de la meva curiositat.

Per últim, però no per això menys important, gràcies a la meva família i amics que han suposat en tot moment un suport moral molt important.

Índex

1 Formes de representació d'un joc

En aquest primer capítol, introduirem els conceptes bàsics per entendre que és un joc i com funciona, això ho trobarem explicat de manera més detallada a [1], [2], [3], [4] i [5], cinc llibres per a novells en la teoria de jocs.

1.1 Introducció

En llenguatge ordinari, el concepte de joc fa referència a diversió i també a una activitat on els participants, sota certes regles que han de complir, intenten guanyar, tot i que també poden perdre. Els més coneguts són els jocs de taula, com el pòquer o els escacs, o els jocs esportius com el futbol o el bàsquet. Solen tenir diversos jugadors, tot i que a vegades amb un de sol n'hi ha prou.

En aquests jocs, cada jugador intenta obtenir el millor resultat possible, maximitzant la seva utilitat, però sempre tenint en compte que el resultat final del joc no depèn només de les seves pròpies accions, sinó també de les accions de la resta de jugadors. És aquesta característica dels jocs -prendre les decisions que més convinguin per guanyar, havent de complir unes normes, i sabent que la resta de jugadors també influeixen en els resultats amb les seves decisions- la que més valor té per al seu estudi sistemàtic, ja que moltes situacions d'interès per l'economia i per altres ciències (polítiques, sociologia o biologia), i que no tenen res a veure amb els jocs mencionats anteriorment, comparteixen amb ells aquesta característica. La teoria de jocs s'ocupa de l'anàlisi rigorosa i sistemàtica de dites situacions. Així doncs, podríem referir-nos a la teoria de jocs, com la teoria de la decisió interactiva, que és diferent de la teoria de la decisió individual.

Tot i que la teoria de jocs no se centra especialment en els jocs corrents, si que els fa servir com a exemples per facilitar la comprensió de certs conceptes. El camp d'estudi de la teoria de jocs és molt general, no és necessari que hi hagi entreteniment, però si interacció. Tot i que les aplicacions millor estudiades de la teoria de jocs suposen que els jugadors són agents (persones, empreses o governs) racionals (la seva capacitat de raonament i càlcul per identificar les accions i estratègies que els conduiran a resultats més desitjables, és infinita), en altres casos, els jugadors no necessiten ser persones ni racionals. En economia sovint, s'estudien situacions de decisió individual en les que l'agent intenta maximitzar la seva utilitat, sense importar el que faci la resta. Per exemple:

- L'elecció de quantitats d'un bé a produir per part d'una empresa preu acceptant.[2] Se suposen coneguts els preus del bé i dels factors de producció i coneguda la funció de producció.
- L'elecció del preu d'un bé per un monopolista. Se suposen donats els preus dels factors de producció i la corba de demanda de dit bé i coneguda la funció de producció.

[2] Preu acceptant: Venedor o comprador que no pot afectar el preu i la decisió del qual està restringida a la quantitat que ha de vendre o comprar al preu que fixa el mercat.

No obstant, hi ha moltes altres situacions en les quals la utilitat del resultat final no depèn només de l'acció d'un agent, sinó també de les accions d'altres agents. Per exemple:

- L'elecció per una certa empresa A de la quantitat a produir d'un bé o del preu de dit bé, si també el produeix una empresa B i cap més (duopoli).
- L'elecció per una empresa d'automòbils d'un nivell de despesa en publicitat. Les conseqüències finals de tal despesa dependran de la despesa realitzada en publicitat per les empreses competidores.

1.2 Breu història de la teoria de jocs

Es considera que el naixement de la teoria de jocs com a disciplina va ocórrer l'any 1944 amb la publicació de *Game Theory and Economic Behaviour* de John Von Neumann i Oskar Morgenstern, tot i que hi ha treballs anteriors com els dels matemàtics Zermelo (1913), Borel (1921) i del propi Von Neumann (1928), en els que ja s'anticipava part de les bases de la teoria de jocs. També cal destacar els treballs pioners d'economistes com Cournot (1838) i Edgeworth (1881). Von Neumann i Morgenstern van establir les bases del que actualment es coneix com a Teoria de Jocs clàssica, proporcionant una solució per a jocs de suma zero (aquells en els que els jugadors es troben en conflicte absolut), establint els fonaments per a l'anàlisi de jocs amb més de dos jugadors. En aquest sentit, van crear una teoria unificada i sistemàtica que va incloure com a casos particulars, les aportacions anteriors, i que va fer factible el seu desenvolupament posterior. Ja en els anys cinquanta, John Nash va aportar alguns dels conceptes més importants (equilibri de Nash i solució de negociació de Nash) per a una gamma més àmplia de jocs (no només aquells que modelitzaven el conflicte pur), i en els anys setanta, investigador com Selten (en jocs dinàmics) i Harsanyi (en jocs amb informació incompleta), van desenvolupar els conceptes que permetrien l'aplicació fructífera de la teoria de jocs a l'economia i altres disciplines.

1.3 Elements d'un joc

- **Jugadors:** Són els participants en el joc que prenen decisions amb la finalitat de maximitzar la seva utilitat.

$$J = \{1, 2, \ldots, n\} \text{ (conjunt finit)}$$

- **Accions de cada jugador:** Són les decisions que pot prendre cada jugador en cada moment que li toca jugar. El conjunt d'accions d'un jugador pot ser finit o infinit.

- **Resultats del joc:** Són els diferents modes en què pot concloure un joc. Cada resultat porta aparellades unes conseqüències per a cada jugador.

- **Estratègies i perfils d'estratègies:** Una estratègia d'un jugador és un pla complet d'accions amb les quals aquest podria participar en un joc. Un perfil d'estratègies és un conjunt d'estratègies, una per cada jugador. Designarem amb S_i al llistat d'estratègies d'un jugador arbitrari i i $s_i \in S_i$ a una estratègia del jugador i.

$$S = S_1 \text{ x } S_2 \text{ x } \ldots \text{ x } S_i \text{ x } \ldots \text{ x } S_n$$

  és el conjunt de tots els perfils d'estratègies. Un possible escenari pel jugador i, és un vector d'estratègies de la resta de jugadors. Denotarem per S_{-i} al conjunt de tots els possibles escenaris del jugador i, on:

$$S_{-i} = S_1 \text{ x } S_2 \text{ x } \ldots \text{ x } S_{i-1} \text{ x } S_{i+1} \text{ x } \ldots \text{ x } S_n.$$

- **Pagaments:** Cada jugador rep un pagament en acabar el joc, que depèn de quin hagi estat el resultat del joc. El significat d'aquest pagament és la utilitat que cada jugador atribueix al dit resultat, és a dir, la valoració que pel jugador tenen les conseqüències d'arribar a un determinat resultat en el joc.

$$\begin{array}{rccc} u_i\colon & S & \longrightarrow & \mathbb{R}^n \\ & s & \longmapsto & u_i(s) \end{array}$$

1.4 Funcions d'utilitat

Sigui S_i un conjunt d'estratègies possibles, mútuament excloents, entre les que ha de triar un jugador. A S_i suposem definida una relació binària $\gtrsim$, anomenada relació de preferència, de manera que per s_i, $s_i' \in S_i$, $s_i \gtrsim s_i'$, vol dir que l'alternativa s_i és preferida o indiferent a l'alternativa s_i'. D'aquí, podem definir la relació de preferència estricta $>$:

$$s_i > s_i' \Leftrightarrow s_i \gtrsim s_i', \text{ però no } s_i' \gtrsim s_i.$$

Donat un joc i un jugador i, direm que l'estratègia s_i domina estrictament una altre estratègia s_i', si aquesta té preferència estricta, és a dir, si és una estratègia millor independentment de l'escenari en què es trobi.

Exemple 1.1. Imaginem un joc amb dos jugadors: Jugador 1 i Jugador 2, on les possibles estratègies de cada un d'ells són:

- El jugador 1 pot triar entre dues estratègies: pot jugar "Dalt"o "Baix".
- El jugador 2 pot triar entre tres estratègies: pot jugar "Esquerra", "Mig"o "Dreta".
- Així doncs, hi ha 6 casos possibles: "Dalt Esquerra", "Dalt Mig", "Dalt Dreta", "Baix Esquerra", "Baix Mig"o "Baix Dreta".

Jugador1 / Jugador 2	Esquerra	Mig	Dreta
Dalt	$1,0$	$1,2$	$0,0$
Baix	$0,3$	$0,1$	$2,0$

Per cada una d'aquestes 6 combinacions, la matriu indica el pagament que rep cada jugador, on a cada casella, el primer pagament es correspon al Jugador 1, i el segon element al Jugador 2. Un joc escrit en forma matricial es coneix com un Joc de Forma Normal. Fem servir aquesta representació quan els dos jugadors trien simultàniament.

Ara bé, com resolem un joc de forma normal com el que tenim a dalt? Ho farem a partir de la racionalitat dels jugadors, és a dir, cada jugador sempre triarà l'estratègia que li representi pagaments més alts, sense pensar en la resta de jugadors. Anem a mirar l'exemple anterior.

Comencem pel Jugador 1, on les seves estratègies són "Dalt"o "Baix". Mirarem si alguna d'aquestes dues estratègies domina estrictament l'altra, ho farem comparant les utilitats del jugador 1, suposant fixada l'estratègia del jugador 2. En els dos primers casos, fixats "Esquerra"o "Mig", l'estratègia "Dalt"del Jugador 1 domina a "Baix", ja que $1 > 0$ i $1 > 0$, però en el tercer cas, "Dreta", $0 < 2$, així que de moment cap estratègia del jugador 1 domina estrictament.

Si mirem ara al Jugador 2, veiem que l'estratègia "Mig"domina estrictament l'estratègia "Dreta", ja que $2 > 0$ i $1 > 0$, per tant podem eliminar l'estratègia "Dreta"del Jugador 2, obtenint així una matriu 2x2.

Jugador 1 / Jugador 2	Esquerra	Mig
Dalt	$1,0$	$1,2$
Baix	$0,3$	$0,1$

En aquest joc simplificat, l'estratègia "Dalt"del Jugador 1, dominaria estrictament "Baix", i la podríem eliminar. Finalment el Jugador 2 hauria de triar entre "Esquerra"o "Mig", on triaria "Mig"ja que $0 < 2$. Aleshores, en aquest joc, l'estratègia d'equilibri seria *(Dalt,Mig)* i el pagament d'equilibri $(1,2)$.[3]

Ara bé, no sempre tenim estratègies estrictament dominants, i és per això que tenim altres mètodes per a trobar aquests equilibris.

[3] En aquest exemple hem pogut arribar a un equilibri, ja que sempre hem trobat una estratègia que dominava estrictament a una altra. Ara bé, la racionalitat no ens permet eliminar estratègies que dominin dèbilment, és a dir, on almenys una de les utilitats sigui igual en les dues estratègies a comparar, ja que fer això podria suposar l'eliminació de l'estratègia d'equilibri.

1.5 Jocs en Forma Extensiva

1.5.1 Elements d'un joc

A continuació es defineixen els elements que caracteritzen a un joc en forma extensiva.

1. Denotarem per $J = \{0, 1, 2, \ldots, n\}$ al conjunt de jugadors. Si no hi ha moviments d'atzar o de la naturalesa, aleshores $J = \{1, 2, \ldots, n\}$.

2. Denotarem per X al conjunt de nodes. Un node representa una possible situació del joc. Entre els nodes hi ha un d'ells que és l'arrel del joc, punt de començament del joc. Tal node es representa per O (referent a l'origen). A continuació es defineix la següent funció:

$$\begin{array}{rccc} \sigma\colon & X & \longrightarrow & X \\ & x & \longrightarrow & \sigma(x) \end{array}$$

   on $\sigma(O) = O$, i per cada $x \neq O$, $\sigma(x)$ és el node immediatament predecessor de x. Sigui $\sigma^2(x) = \sigma(\sigma(x))$, $\sigma^3(x) = \sigma(\sigma(\sigma(x)))$, i així successivament, pel que iterant $\sigma(x)$ s'obtenen tots els nodes predecessors de x. Sigui $s(x) = \sigma^{-1}(x)$, el conjunt de nodes que segueixen immediatament a x. Un node és terminal si no el segueix cap altre node sent un node de decisió si el segueix algun altre node. Es defineixen els següents conjunts:

$$T(X) = \{x \in X : s(x) = \varnothing\}$$

   el conjunt de nodes terminals del joc i

$$D(X) = \{x \in X : s(x) \neq \varnothing\}$$

   el conjunt de nodes de decisió del joc.

3. Denotarem per A al conjunt de totes les possibles accions. Es defineix la funció:

$$\begin{array}{rccc} \alpha\colon & X \setminus \{O\} & \longrightarrow & A \\ & x & \longrightarrow & \alpha(x) \end{array}$$

   que fa correspondre a cada node diferent de l'origen aquella acció $\alpha(x)$ que porta des del node immediat predecessor $\sigma(x)$ al node x. Es verifica que si $x', x'' \in s(x)$, sent $x' \neq x''$, aleshores $\alpha(x') \neq \alpha(x'')$. És a dir, accions que parteixen del mateix node i condueixen a nodes diferents, han de ser diferents. Per qualsevol node de decisió $x \in D(X)$, representem el conjunt d'accions disponibles a partir de x per:

$$A(X) = \{a \in A : \exists x' \in s(x), a = \alpha(x')\}.$$

4. Per cada jugador i, sigui X_i el conjunt de nodes de decisió en els que el jugador i ha de triar una acció. En un node particular de decisió només mou un dels jugadors. Es té que:

$$\bigcup_{i \in J} X_i = D(X)$$

$$\forall i, j, \text{ a,b } i \neq j, \text{ es verifica que } X_i \cap X_j = \varnothing$$

Observem, per tant, que la familia $\{X_i\}_{i\in J}$ constitueix una partició, per jugadors, del conjunt de nodes de decisió $D(X)$.

5. Una família de conjunts d'informació H, i una funció:

$$\begin{array}{rccl} h: & X & \longrightarrow & H \\ & x & \longrightarrow & h(x) \end{array}$$

que assigna a cada node x un conjunt d'informació $h(x)$ al que pertany. Els conjunts d'informació formen una partició de $D(X)$. Com hem vist anteriorment, tots els nodes de decisió que pertanyen a un mateix conjunt d'informació tenen les mateixes accions disponibles, és a dir:

$$A(x) = A(x'), \text{ si } h(x) = h(x')$$

Sigui $h = h(x)$, un conjunt d'informació pertanyent a H. Per tant, podem representar per $A(h)$ el conjunt d'accions disponibles en el conjunt d'informació h.

$$A(h) = \{a \in A : a \in A(x), x \in h\}$$

Sigui H_i el conjunt de tots els conjunts d'informació del jugador i. Sigui H el conjunt que conté tots els conjunts d'informació continguts en els H_i, per $i \in J$. És a dir,

$$H = \bigcup_{i\in J} H_i$$

6. Una funció:

$$\begin{array}{rccl} \rho: & H_0 \times A & \longrightarrow & [0,1] \\ & (h,a) & \longrightarrow & \rho(h,a) \end{array}$$

que assigna probabilitats a accions en conjunts d'informació on el moviment correspon a la naturalesa o a l'atzar.[4] S'ha de verificar que:

$$\rho(h,a) = 0, \text{ si } a \notin A(h) \text{ i } \textstyle\sum_{a\in A(h)} \rho(h,a) = 1, \; \forall h \in H_0$$

7. Una funció de pagaments:

$$\begin{array}{rccl} r: & T(x) & \longrightarrow & \mathbb{R}^n \\ & x & \longrightarrow & u(x) = (u_1(x), \ldots, u_n(x)) \end{array}$$

on $u_i(x)$ indica el pagament o utilitat que rep el jugador i si s'ha arribat al node terminal x.

1.6 Equilibri de Nash

Definició 1.2. *En el joc $G = \{S_1, \ldots, S_n; u_1, \ldots, u_n\}$, diem que el perfil d'estratègies pures $(s_1^*, s_2^*, \ldots, s_i^*, \ldots, s_n^*)$ és un equilibri de Nash (EN), si per cada jugador i,*

$$u_i(s_1^*, \ldots, s_{i-1}^*, s_i^*, s_{i+1}^*, \ldots, s_n^*) \geq u_i(s_1^*, \ldots, s_{i-1}^*, s_i, s_{i+1}^*, \ldots, s_n^*) \text{ per a tot } s_i \text{ de } S_i.$$

[4] H_0 és el conjunt de tots els conjunts d'informació del jugador 0, on recordem que són les accions de l'atzar o la naturalesa.

És a dir, per cada jugador i, s_i^* és una solució del problema:

max $u_i(s_1^*, \ldots, s_{i-1}^*, s_i, s_{i+1}^*, \ldots, s_n^*)$ on s_i és l'estratègia de decisió, i pertany a S_i.

Dit d'una altra manera, per cada jugador i, s_i^* és una resposta òptima a s_{-i}^*.

D'aquesta definició es dedueix que un equilibri de Nash (EN), és un perfil d'estratègies del qual cap jugador desitjaria desviar-se unilateralment, és a dir, cap es penedeix de la decisió presa, donades les estratègies triades per la resta de jugadors. Un EN està format per estratègies que són òptimes per cada jugador donades les estratègies de la resta de jugadors.

Això no significa que en un EN cada jugador estigui obtenint el millor resultat possible, sinó el millor resultat possible condicionat pel fet que la resta de jugadors, han jugat les estratègies indicades per ells en el perfil esmentat.

Pot haver-hi múltiples equilibris de Nash en un joc, i anomenarem S^{EN} al conjunt de perfils que són equilibris de Nash.

1.6.1 Correspondència de resposta òptima

El procés de càlcul dels EN en estratègies pures d'un joc, depèn de les característiques de tal joc. En els jocs finits i de mida reduïda, com l'exemple vist anteriorment, és fàcil comprovar de manera detallada totes les possibilitats, mentre que en els jocs infinits sol ser necessari un plantejament més analític.

Tanmateix, en tots els casos és convenient organitzar la recerca dels EN de manera sistemàtica, calculant les estratègies òptimes que cada jugador podria triar en resposta a qualsevol combinació d'estratègies que poguessin triar els altres jugadors.

Es busca, per tant, donat un jugador i i per cada combinació s_{-i} d'estratègies de la resta de jugadors, un conjunt d'estratègies d'aquest, que anomenarem $R_i(s_{-i})$. La regla que a cada s_{-i} (el que podrien fer la resta) li assigna $R_i(s_{-i})$ (el que li convé fer a ell), rep el nom de correspondència de resposta òptima del jugador i.

Definició 1.3. *En el joc $G = \{S_1, \ldots, S_n; u_1, \ldots, u_n\}$, i per cada jugador i, anomenem correspondència de resposta òptima del dit jugador a la regla que assigna, a qualsevol combinació d'estratègies $s_{-i} = (s_1, s_2, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$, el conjunt $R_i(s_{-i})$ d'estratègies de i que són resposta òptima a s_{-i}, és a dir, que compleixen:*

$$\begin{array}{rccl} R_i\colon & S_{-i} & \longrightarrow & S_i \\ & R_i(s_{-i}) & \longmapsto & R_i(s_{-i}) = \{s_i^* \in S_i \mid u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i}), \forall s_i \in S_i\} \end{array}$$

Exemple 1.4. El següent exemple està basat en un programa de la televisió anglesa anomenat *Golden Balls*. Inicialment són diversos concursants, però ens centrarem en el joc final on només en queden dos. Després de diverses proves, els jugadors han acumulat unes desenes de milers de lliures. A l'escenari final cada jugador rep dues boles daurades, en una posa "Split"(Dividir) i en l'altra "Steal"(Robar), i de manera secreta han de seleccionar una d'aquestes dues boles. Abans de prendre una decisió final, tenen l'oportunitat de negociar amb l'oponent. En la majoria de casos, ambdós participants s'intenten convèncer a triar l'opció "Dividir", que és la més mútuament beneficiosa. Malgrat això, l'estratègia de "Dividir-Dividir", tot i maximitzar el benestar dels dos, no és una estratègia estable. La majoria de jocs acaben amb almenys una de les dues parts desertant. O bé per una banda un s'emporta tot (deixant a l'altre sense res) o bé els dos tornen a casa de buit. Hi ha tres escenaris possibles:

- Si els dos trien "Dividir", es reparteixen el pot en dues parts iguals.
- Si un tria "Robar", i l'altre "Dividir", el jugador que ha robat s'emporta tot el pot, mentre que l'altre se'n va amb les mans buides.
- Si els dos trien "Robar", cap dels dos s'emporta premi.

Concursant 1 / Concursant 2	Dividir	Robar
Dividir	$50, 50$	$0, 100$
Robar	$100, 0$	$0, 0$

Per resoldre el joc, podríem intentar buscar estratègies estrictament dominants, però en aquest cas no hi ha cap. Anem a fer-ho a partir de la correspondència de resposta òptima, assignarem el color blau a les millors respostes del concursant 1 i el color vermell a les del concursant 2. Per cada caixa que tingui els dos colors, tindrem un equilibri de Nash del joc.

Concursant 1 / Concursant 2	Dividir	Robar
Dividir	$50, 50$	$\color{blue}{0}, \color{red}{100}$
Robar	$\color{blue}{100}, \color{red}{0}$	$\color{blue}{0}, \color{red}{0}$

Les correspondències de resposta òptima són:

- $R_1(Robar) = \{Dividir, Robar\}$
- $R_2(Robar) = \{Dividir, Robar\}$

Pels dos concursants podem veure que Robar $\gtrsim$ Dividir, és a dir, que Dividir es una estratègia dominada dèbilment per Robar. De la taula treiem, que el joc té tres equilibris de Nash:

$S^{EN} = \{(Robar, Dividir), (Dividir, Robar), (Robar, Robar)\}.$

1.6.2 Eficiència de Pareto

Definició 1.5. *Donat el joc* $G = \{S_1, \ldots, S_n; u_1, \ldots, u_n\}$,

1. *Diem que el perfil d'estratègies* $s = (s_1, s_2, \ldots, s_i, \ldots, s_n)$ *està dominat en el sentit de Pareto pel perfil* $s' = (s'_1, s'_2, \ldots, s'_i, \ldots, s'_n)$ *si i només si la desigualtat* $u_i(s'_1, \ldots, s'_n) \geq u_i(s_1, \ldots, s_n)$ *es compleix per tot jugador i, i per algun d'ells es compleix de manera estricta.*

2. *Diem que el perfil d'estratègies* $s = (s_1, s_2, \ldots, s_i, \ldots, s_n)$ *és un òptim de Pareto si i només si no està dominat en el sentit de Pareto per cap altre perfil. Direm que és ineficient en el sentit de Pareto si algun altre el domina.*

És a dir, si un perfil d'estratègies és eficient, no es pot canviar a cap altre perfil, de manera que cap jugador surti perdent i algun surti guanyant estrictament.

Aquest concepte de dominació, fa referència a perfils complets que al·ludeixen a tots els jugadors alhora, mentre que el concepte d'estratègia dominada es referia a cada jugador individualment. Mentre la dominació de Pareto és un concepte d'anàlisi de l'eficiència col·lectiva, rellevant pel grup de jugadors com tal grup, la dominació d'estratègies és un concepte d'anàlisi de l'eficiència individual, rellevant per cadascun dels jugadors com agent individual. No és d'estranyar, per tant, que ambdós conceptes resultin totalment independents, com veurem en l'exemple següent.

Exemple 1.6. El dilema del presoner és, probablement, el joc més simple i famós, i es basa en el següent relat il·lustratiu:

Dos delinqüents habituals són atrapats quan acaben de cometre un delicte greu. No hi ha cap prova clara contra ells, però si indicis forts del dit delicte i a més hi ha proves d'un delicte menor. Són interrogats simultàniament en habitacions separades. Els dos saben que si callen, seran absolts del delicte principal per falta de proves, però els condemnaran pel delicte menor (1 any de presó), que si els dos confessen, seran condemnats pel principal, però se'ls hi rebaixarà una mica la pena per confessar (4 anys), i finalment, que si un sol confessa, aquest es lliurarà i a l'altre li cauran 5 anys.

La representació en forma estratègica és la següent:

Presoner 1 / Presoner 2	Callar	Confessar
Callar	$-1, -1$	$-5, 0$
Confessar	$0, -5$	$-4, -4$

Tenint en compte el significat dels pagaments, podem aplicar a l'escala de pagaments de cada jugador una transformació afí positiva, per exemple sumarem 5 unitats a tots els pagaments del joc. Això ens serà útil per més endavant, quan repetim el joc infinitament. Així doncs, ens centrarem en el fet que el jugador que tingui la utilitat més gran, és el que entrarà menys anys a presó.

La nova representació en forma estratègica, l'anomenarem dilema del presoner en escala estàndard i és la següent:

Presoner 1 / Presoner 2	Callar	Confessar
Callar	$4, 4$	$0, 5$
Confessar	$5, 0$	$1, 1$

Si resolem el joc mirant les millors respostes, ens queda:

Presoner 1 / Presoner 2	Callar	Confessar
Callar	$4, 4$	$0, 5$
Confessar	$5, 0$	$1, 1$

De manera aparentment sorprenent, l'únic perfil que no estaria constituït per estratègies no estrictament dominades (sent per tant l'únic EN), és *(Confessar,Confessar)*, el qual és ineficient en el sentit de Pareto. Efectivament, el perfil *(Callar,Callar)* el domina en el sentit de Pareto.

1.6.3 Comentaris sobre cooperació i conflicte

L'anàlisi del dilema del presoner ens mostra una possible debilitat del concepte de solució EN (i també del resultant de l'eliminació iterativa d'estratègies estrictament dominades). El perfil proposat com a solució és ineficient en el sentit de Pareto. Dit d'una altra manera, com jugaran *(Confessar,Confessar)* tal com prescriu la solució EN, si jugant *(Callar,Callar)* els dos aconsegueixen un millor pagament?

La clau per entendre aquesta aparent paradoxa, radica en comprendre que el resultat *(Callar,Callar)* és l'apropiat des del punt de vista social, és a dir, del conjunt dels jugadors, mentre que el perfil *(Confessar,Confessar)*, és l'apropiat des del punt de vista individual.

Hi ha qui se sorprèn d'aquesta contradicció entre l'individual i el social, però aquest senzill joc posa de manifest que tal contradicció existeix. I bé, quin és el resultat predictible del joc si dues persones el juguen efectivament en la pràctica? La resposta depèn del context particular. Suposem que el context és el més semblant al del model matemàtic al qual anomenem *Dilema del presoner*, és a dir, es juga un sol cop i els dos jugadors saben, que no poden comunicar-se abans de decidir la seva jugada i que els dos són racionals (saben raonar i es preocupen només pel seu pagament). En la nostra opinió, el resultat predictible seria el perfil EN. Malgrat això, si modifiquem alguna de les suposicions anteriors, el resultat cooperatiu guanyaria força. Per exemple, suposem que el joc el juguen un pare i un fill, el resultat predictible seria *(Callar,Callar)*. Més endavant veurem que passa quan es juga més d'un cop.

2 Jocs Dinàmics amb informació completa

En aquest apartat analitzarem els jocs dinàmics amb informació completa, que són aquells en què els jugadors poden prendre decisions en diferents instants del temps i on cada jugador coneix amb tota certesa les funcions de pagament de tots els jugadors. La seva forma de representació natural és l'extensiva. Per veure de manera més detallada els jocs dinàmics, consultar [6], [7], [9] i [12].

En incloure en l'anàlisi les característiques temporals del joc, es fa possible que un jugador tingui, en el moment de jugar, un coneixement de les decisions preses per altres jugadors en moments anteriors, i prevegi la possibilitat que les seves decisions actuals siguin conegudes pels altres jugadors, si aquests juguen amb posterioritat.

2.1 Elements d'un joc representat en forma extensiva

Recordem que la representació en forma extensiva d'un joc requereix respondre les preguntes: qui?, quan?, com?, que saben? I quant? Concretant el significat d'aquestes preguntes, és necessari precisar:

1. Qui són els jugadors.
2. (a) Quan ha de jugar cada jugador.
   (b) Quines coses pot fer quan li toca jugar (accions factibles).
   (c) Que sap tal jugador, quan li toca jugar, del desenvolupament previ del joc.
3. A quant pugen els guanys de cada jugador per cada possible desenvolupament del joc.

Anem a veure el dilema del presoner repetit n cops. Dos jugadors juguen n vegades consecutives el dilema del presoner. Se suposa que en començar cada una d'elles és de domini públic com han jugat ambdós jugadors en les etapes anteriors i que els pagaments finals del joc es calculen sumant els pagaments corresponents en les diferents etapes.

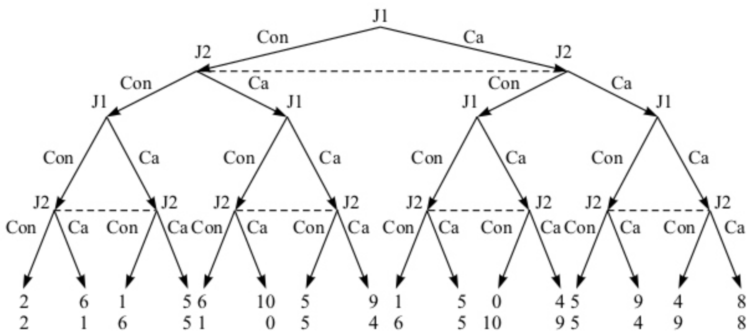


Figura 1: Dilema del Presoner repetit 2 cops

Definició 2.1. *Un conjunt d'informació d'un jugador és unitari, si al jugador al qual li toca decidir, coneix tota la història prèvia del joc fins aquell moment.*

Definició 2.2. *Sigui G un joc d'informació completa (l'estructura i els pagaments són de domini públic). Diem que G és d'informació perfecta si cada conjunt d'informació de qualsevol dels jugadors és unitari. Si, per contra, existeix un jugador amb algun conjunt d'informació no unitari, diem que el joc G és d'informació imperfecta.*

2.2 Equilibri de Nash Perfecte en Subjocs (ENPS)

L'estructura temporal d'un joc dinàmic amb informació completa obliga cada jugador a tenir en compte que les seves decisions en cada moment influeixen en les possibilitats i pagaments posteriors per a ell i per a la resta, i que al mateix temps les decisions futures previsibles seves i de la resta condicionen les seves decisions presents. Apareix així un element de vital importància, la credibilitat que es pot donar a les decisions futures a l'hora de determinar les decisions presents, és a dir, la credibilitat de les possibles amenaces o promeses que es poden plantejar sobre el comportament futur per condicionar el comportament present dels jugadors.

2.2.1 Subjocs

Quan en un punt del desenvolupament d'un joc en forma extensiva es dóna la circumstància de què l'ocorregut fins aquell moment en el joc (quines jugades han realitzat els diferents jugadors i quins resultats s'han obtingut) és de domini públic, això fa que els jugadors puguin veure la part que falta per jugar com un joc en si mateix. Sembla natural pensar que els criteris d'actuació que els guiaran a partir d'aquell moment coincidiran amb els que ja tenien abans d'iniciar el joc. Aquesta idea ens permetrà definir, un nou concepte d'equilibri. Tal concepte afegeix, a les característiques del ja conegut equilibri de Nash, l'exigència de coherència entre el mode de comportament abans d'iniciar-se el joc, i el mode de comportament quan el joc ja s'ha iniciat. És precís, per tant, identificar amb precisió quines parts poden raonablement considerar-se com un joc en si mateix. A aquestes parts les anomenarem subjocs.

Definició 2.3. *Donat un joc G amb informació completa en forma extensiva, i un node de decisió x de G, diem que G' és un subjoc de G amb inici en x si G' és una part de G que compleix:*

1. *Conté al node x i a tots els nodes que segueixen a x, i només a ells.*
2. *El node x és un conjunt d'informació unitari.*
3. *Si el node y pertany a G', també pertanyen a G' tots els nodes del conjunt d'informació al que pertany y, és a dir, G' no trenca cap conjunt d'informació.*

Observació 2.3.

1. És evident que el propi G és sempre un subjoc de G. Anomenem subjocs propis de G a aquells subjocs que no coincideixen amb G (és a dir, que no comencen en el node inicial de G).
2. Si G és d'informació perfecta, qualsevol part que comenci en un node de decisió i contingui tots els nodes que el segueixen, és un subjoc (pel fet que tots els conjunts d'informació són unitaris).
3. També es consideren subjocs aquells en els que només intervindrà un dels jugadors.

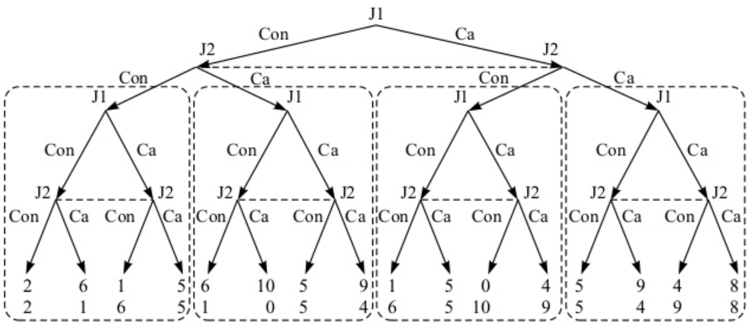


Figura 2: Subjocs propis del dilema del presoner, repetit 2 cops.

2.2.2 Equilibri de Nash Perfecte en Subjocs

Estem ja en condicions de definir el nou concepte d'equilibri, que va ser proposat per Richard Selten l'any 1965. Aquest autor va rebre el Premi Nobel d'economia el 1994 per les seves contribucions a la teoria de jocs, especialment les referides a aquest i altres conceptes d'equilibri.

Definició 2.4. *Sigui G un joc en forma extensiva.*

Sigui s un perfil d'estratègies de G que és un EN. Diem que s és un Equilibri de Nash Perfecte en Subjocs (ENPS) de G si la restricció de s a qualsevol subjoc de G és un EN de tal subjoc.

A continuació veiem un exemple d'un ENPS i d'un EN que no és ENPS.

Exemple 2.5. Ens trobem davant d'un joc no simultani, on inicialment el jugador 1 tria l'opció A o B i tot seguit el jugador 2 tria entre les opcions C i D.

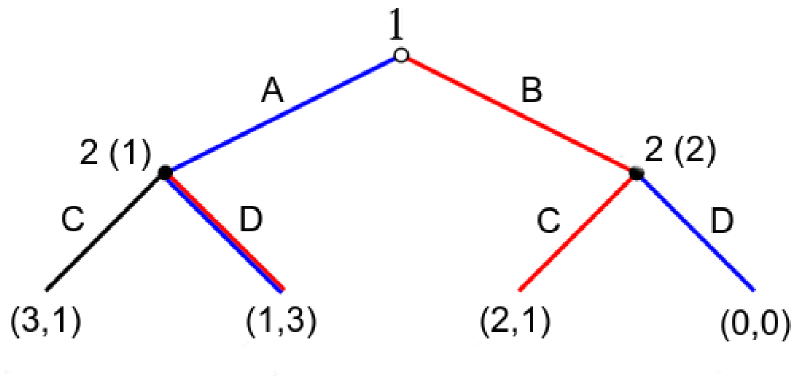


Figura 3: Diferència entre ENPS i EN

Jugador1 / Jugador 2	(C,C)	(C,D)	(D,D)	(D,C)
A	$\underline{3}, 1$	$\underline{3}, 1$	$\underline{1}, \underline{3}$	$1, \underline{3}$
B	$2, \underline{1}$	$0, 0$	$0, 0$	$\underline{2}, \underline{1}$

Fixem-nos que tenim dos EN seleccionant les millors respostes per cadascun dels jugadors. Ara bé, podem observar que l'equilibri en color blau és un EN, però no un ENPS, ja que el jugador 2 fa una amenaça no creïble quan al segon node prefereix triar D (independentment que el joc es desenvolupi per la banda esquerra, la resposta del jugador 2 manca de raó quan tria guanyar 0 davant d'1). Per altra banda l'equilibri en color vermell sí que és un ENPS, ja que totes les restriccions a subjocs són EN.

3 L'oligopoli de Cournot

Una de les aplicacions més fructíferes de la teoria de jocs és la relativa a l'estudi de l'organització industrial en entorns amb un nombre d'agents no gaire gran, en particular, l'estudi de models de mercat amb un nombre reduït d'empreses. Els models de duopoli constitueixen una aplicació pionera d'aquest tipus.

Cournot va ser un dels precursors de la teoria de jocs. En un treball realitzat l'any 1838, va proposar el que ara coneixem com el model clàssic de Cournot, en el que un petit nombre d'empreses competeixen en el mercat d'un producte homogeni, decidint simultàniament quines quantitats de producció aportaran al mercat, i el preu de mercat queda determinat per la quantitat total aportada, d'acord amb la funció de demanda inversa. Tot i que el mecanisme mitjançant el qual es buida el mercat venent tota la producció aportada no s'especifica, podem imaginar una subhasta entre compradors de la producció total. A l'equilibri resultant de tal model se l'anomena equilibri de Cournot o equilibri de Cournot-Nash, indicant que es tracta de l'equilibri de Nash del joc definit pel model de Cournot. Per veure oligopolis de Cournot amb diferents nombres de jugadors, veure [5] i [21], i per una representació més gràfica del problema, consultar [4].

3.1 Oligopoli de Cournot model simplificat

En un mercat hi ha n empreses $E_1, E_2, \ldots, E_n$ que fabriquen un determinat producte homogeni i competeixen en quantitats. Sigui q_i la quantitat que produeix l'empresa E_i per $i = 1, 2, \ldots, n$. Suposem que la funció de demanda inversa és decreixent i lineal en l'interval $[0, a/b]$,[5] que els costos marginals de cada empresa són constants, menors que a i iguals per totes, sense que hi hagi costos fixos, és a dir, si una empresa produeix zero unitats, no tindrà pèrdues i que en aquest mercat es ven tota la producció. En concret, sigui la funció de demanda inversa:

$$\mathrm{P(Q)} = \begin{cases} a - bQ & \text{si bQ <a} \\ 0 & \text{si bQ} \geq \text{a} \end{cases} \text{(on } b > 0 \text{ i } Q = q_1 + q_2 + \ldots + q_n\text{)}$$

i les funcions de cost són:

$$C_i(q_i) = cq_i, \text{ on } c < a, \forall i \in \{1, 2, \ldots, n\}$$

Per tant, les funcions de benefici seran:

$$u_i(q_1, q_2, \ldots, q_n) = q_i \cdot P(q_i + Q_{-i}) - C_i(q_i) = q_i(a - b(q_i + Q_{-i}) - c)$$

on $Q_{-i} = q_1 + \ldots + q_{i-1} + q_{i+1} + \ldots + q_n$, $\forall i \in \{1, 2, \ldots, n\}$.

[5]Per valors de $Q = 0$, la funció de demanda inversa pren el valor a, és a dir, el preu màxim que pot assolir un producte, es donarà quan la producció d'aquest sigui 0, mentre que quan $Q = a/b$, prendrà el valor mínim, que és 0, ja que no tindria sentit definir preus negatius. El valor a indica el preu de reserva (preu al qual es deixa d'adquirir el producte) i el valor b la sensibilitat dels consumidors a les variacions en la producció total.

3.2 Càlcul de l'equilibri de Nash

La resposta òptima d'una empresa qualsevol E_i a una combinació d'accions prefixades, $q_{-i} = (q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n)$ de la resta de les empreses, s'obtindrà resolent:

$$\max_{q_i} u_i(q_1, \ldots, q_i, \ldots, q_n) = q_i(a - b(q_i + Q_{-i}) - c)$$

$$\text{subjecte a: } 0 \leq q_i \leq a/b$$

Suposem, que la solució és interior, és a dir, la suma $q_1 + \ldots + q_i + \ldots + q_n$ pertany a l'interval obert $(0, a/b)$. Derivant la funció de beneficis i igualant a zero, obtenim que la condició de primer ordre és:

$$\partial u_i(q_i, q_{-i})/\partial q_i = P(q_i + Q_{-i}) - \partial C_i(q_i)/\partial q_i + q_i \partial P(q_i + Q_{-i})/\partial q_i = 0$$

que es concreta en:

$$a - b(q_i + Q_{-i}) - c + q_i(-b) = 0;$$

$$a - 2bq_i - bQ_{-i} - c = 0$$

I la condició de segon ordre és $\partial^2 u_i(q_i, q_{-i})/\partial q_i^2 = -2b < 0$ (condició suficient de màxim ja que $b > 0$).

Val la pena veure que l'expressió de la condició de primer ordre té una interpretació interessant. La resta dels dos primers termes del primer membre, $P(q_i + Q_{-i}) - \partial C_i(q_i)/\partial q_i$ és la rendibilitat directa per l'empresa E_i de produir una unitat addicional, mentre que el tercer terme, $q_i \partial P(q_i + Q_{-i})/\partial q_i$, és l'efecte (negatiu) que aquesta unitat addicional causa en la rendibilitat de les unitats ja produïdes per E_i. Així doncs, al produir una unitat més, l'empresa compensa exactament amb la dita rendibilitat positiva directa, la rendibilitat negativa indirectament ocasionada. Per altra banda, el fet que E_i compensi mitjançant la seva rentabilitat positiva directa únicament l'efecte negatiu que a ella li causa tal producció addicional, $q_i \partial P(q_i + Q_{-i})/\partial q_i$), en lloc de compensar l'efecte negatiu causat a tota la indústria, $(q_i + Q_{-i})(\partial P(q_i + Q_{-i})/\partial q_i)$, posa de manifest l'externalitat[6] negativa creada entres les empreses al decidir la seva producció.

També pot fer-se la següent interpretació de l'equació de la condició de primer ordre. La resta $P(q_i + Q_{-i}) - \partial C_i(q_i)/\partial q_i$ és l'excés del preu de mercat P respecte al cost marginal de E_i. Aquest excés és proporcional a la quantitat q_i produïda per E_i, i el factor de proporcionalitat és $-\partial P(q_i + Q_{-i})/\partial q_i$, és a dir, el pendent canviat de signe de la funció inversa de demanda.

Tornant al càlcul de l'equilibri, $(q_1^*, \ldots, q_i^*, \ldots, q_n^*)$ és un equilibri de Nash si compleix les n condicions de primer ordre:

$$a - 2bq_i^* - bQ_{-i}^* - c = 0, \; \forall i \in \{1, 2, \ldots, n\}$$

Tenint en compte que, per raons de simetria, $q_i^* = q_j^*$, i per tant $(n-1)q_i^* = Q_{-i}^*$, obtenim:

[6] Perjudici o benefici experimentat per un individu o una empresa a causa d'accions executades per altres persones o entitats.

$$a - 2bq_i^* - b(n-1)q_i^* - c = 0;\ b(n+1)q_i^* = a - c;$$
$$q_i^* = \frac{a-c}{(n+1)b},\ \forall i \in \{1, 2, \ldots, n\}$$

En conclusió, hi ha un únic equilibri, que és:

$$(q_1^* = \tfrac{a-c}{(n+1)b}, \ldots, q_i^* = \tfrac{a-c}{(n+1)b}, \ldots, q_n^* = \tfrac{a-c}{(n+1)b})$$

La quantitat total produïda en equilibri és $Q^* = n\frac{a-c}{(n+1)b}$, el preu en equilibri és $P^* = a - bQ^* = \frac{a+nc}{n+1}$, els beneficis en equilibri són:

$$u_i^* = q_i^*(a - b(q_i^* + Q_{*i}^*) - c) = q_i^*(a - n\tfrac{a-c}{n+1} - c) = \tfrac{(a-c)^2}{(n+1)^2 b},\ \forall i \in \{1, 2, \ldots, n\}$$

i el benefici total en equilibri és $U^* = u_1^* + \ldots + u_n^* = \frac{n(a-c)^2}{(n+1)^2 b}$.

3.3 Comparació dels resultats anteriors amb els corresponents als casos de monopoli i de competència perfecta en quantitats

Les dues situacions hipotètiques extremes quant al grau de competència en el mercat del producte homogeni considerat, i en les quals no són aplicables els conceptes d'equilibri de la teoria de jocs, són la de monopoli i la de competència perfecta. Veiem, no obstant, que els resultats que la microeconomia obté per les dues situacions, podrien obtenir-se com casos particulars dels a dalt obtinguts, mitjançant l'equilibri de Nash, pel cas de l'oligopoli.

3.3.1 Monopoli

Si hi ha una sola empresa produint aquest bé homogeni per aquest mercat, no tenim un veritable joc, sinó una situació de decisió individual. Dita empresa prendrà decisions òptimes de producció, tenint en compte que les condicions de demanda en tal mercat determinaran un preu al qual es vendrà tota la producció. Suposant les mateixes condicions de demanda i de costos, tindrem que:

La funció de demanda inversa és:

$$\mathrm{P(Q)} = \begin{cases} a - bQ & \text{si bQ <a} \\ 0 & \text{si bQ} \geq \text{a} \end{cases} \text{ (on } Q \text{ és la quantitat produïda)}$$

La funció de costos és:

$$C(Q) = cQ \text{ on se suposa } c < a$$

Els beneficis seran, per tant:

$$U(Q) = \pi(Q) = Q(a - bQ) - cQ = Q(a - c - bQ)$$

Calculem la decisió òptima del monopoli, que serà produir la quantitat Q^* que maximitza $\pi(Q)$.

Condició de primer ordre (suposant que l'òptim és interior):

$$\partial\pi(Q)/\partial Q = Q(-b) + (a - c - bQ) = 0; \quad a - c - 2bQ = 0; \quad Q = \tfrac{a-c}{2b}$$

Condició de segon ordre:

$$\partial^2\pi(Q)/\partial Q^2 = -2b < 0 \text{ (Màxim)}$$

Així doncs la quantitat produïda òptima és $Q^* = \frac{a-c}{2b}$, que determina un preu $P^* = a - bQ^* = \frac{a+c}{2}$, obtenint un benefici màxim $U^* = \frac{(a-c)^2}{4b}$.

3.3.2 Competència perfecta

Si hi ha moltes empreses, de tal manera que la quantitat produïda per cadascuna d'elles té un pes insignificant en la quantitat produïda agregada, i per tant, no afecta al preu determinat per la igualtat entre la demanda i aquesta quantitat agregada (podríem dir que hi ha infinites empreses petites), cada empresa produirà una quantitat q_i^* que maximitzi els seus beneficis suposant fixat el preu. Sigui P^* tal preu, i sigui la funció de costos de cada empresa $C_i(q_i) = c_i q_i$. El problema d'optimització de E_i és:

$$\max_{q_i} \pi_i(q_i) = (P^* - c_i)q_i$$

L'única possibilitat que, sent compatible amb aquesta situació, tingui una certa estabilitat, es produeix quan $P^* = c_i$. En efecte, la possibilitat $P^* < c_i$ implica beneficis negatius per l'empresa E_i que obligarien a aquesta a abandonar el mercat ($q_i^* = 0$), mentre que la possibilitat $P^* > c_i$ implica que l'empresa E_i només aconsegueix apropar-se a la seva quantitat òptima ($q_i^* = +\infty$ produïnt quantitats tan grans que contradiuen la seva suposada insignificança). Per tant, $P^* = c_i$ per qualsevol empresa.

Suposant les mateixes condicions de demanda i de costos que en els models anteriors, la funció de demanda inversa és:

$$\mathrm{P(Q)} = \begin{cases} a - bQ & \text{si bQ <a} \\ 0 & \text{si bQ} \geq \text{a} \end{cases} \text{ (on } Q \text{ és la quantitat produïda)}$$

i la funció de costos de cada empresa és $C_i(q_i) = cq_i$ (on $c < a$). Així doncs, tindríem el preu d'equilibri $P^* = c$, la quantitat total d'equilibri $Q^* = \frac{a-c}{b}$ (donat que $P(Q^*) = a - b\frac{a-c}{b} = c$), i uns guanys globals en equilibri nuls. Resumint, cada empresa produirà una quantitat tal que el seu cost marginal iguali a P^*, obtenint així un benefici net nul, i la suma Q^* de les quantitats així produïdes, serà precisament aquella que determina, a través de la funció de demanda inversa, el preu P^* que buida el mercat.

3.4 Característíques dels EN segons el nombre d'empreses

Val la pena que ens fem la següent pregunta: ¿Els resultats en quantitats, preus i guanys que acabem d'obtenir pels casos de monopoli i competència perfecta, podrien trobar-se com els casos particulars en què n = 1 i n = ∞, respectivament, dels equilibris de Nash deduïts per l'oligopoli de Cournot amb n empreses? A la següent taula veurem que la resposta és afirmativa, en donar les característiques de l'equilibri segons els valors de n. Podem observar que a major nombre d'empreses (augmentant així la competència al mercat), menor és el benefici total d'aquestes, i major és, en conseqüència, l'excedent del

Taula 1: Resultats d'equilibri, segons el nombre d'empreses

	M^* (n=1)	O_n^* (n qualsevol)	CP^* (n=∞)
Producció Individual	$q_i^* = \frac{a-c}{2b}$	$q_i^* = \frac{a-c}{(n+1)b}$	
Producció total	$Q^* = \frac{a-c}{2b}$	$Q^* = n\frac{a-c}{(n+1)b}$	$Q^* = \frac{a-c}{b}$
Preu	$P^* = \frac{a+c}{2}$	$P^* = \frac{a+nc}{n+1}$	$P^* = c$
Benefici individual	$u_i^* = \frac{(a-c)^2}{2^2b}$	$u_i^* = \frac{(a-c)^2}{(n+1)^2b}$	$u_i^* = 0$
Benefici total	$U^* = \frac{(a-c^2)}{2^2b}$	$U^* = \frac{n(a-c)^2}{(n+1)^2b}$	$U^* = 0$
Excedent del consumidor	$E^* = \frac{(a-c)^2}{2\cdot 2^2b}$	$E^* = \frac{n^2(a-c)^2}{2\cdot(n+1)^2b}$	$E^* = \frac{(a-c)^2}{2b}$

consumidor. Observar que en un oligopoli de Cournot amb n empreses, l'equilibri de Nash, on cada empresa produeix $q_i^* = \frac{a-c}{(n+1)b}$ i obté uns beneficis $u_i^* = \frac{(a-c)^2}{(n+1)^2b}$, no és un òptim de Pareto. Aquest òptim (des del punt de vista de la indústria, sense considerar als consumidors) s'aconseguiria per col·lusió simulant un monopoli, de manera que cada empresa produís $q_i^* = \frac{a-c}{2nb}$ i obtingués uns beneficis de $u_i^* = \frac{(a-c)^2}{2^2nb}$. No obstant, aquesta combinació estratègica no es un equilibri de Nash, perquè cada empresa de manera individual té incentius per augmentar la seva producció si preveu que la resta compliran l'acord de col·lusió.

En la següent figura s'ha representat la corba de demanda inversa $P(Q) = a - bQ$, situant sobre la gràfica els punts (Q, P) corresponents a monopoli (M^*), duopoli (D^*), oligopoli amb tres i n empreses (O_3^* **i** O_n^*) i el cas de competència perfecta (CP^*). S'observa també aquí, que a mesura que augmenta la competència (a través del nombre d'empreses), augmenta la quantitat total produïda i disminueix el preu de mercat que aquesta quantitat determina. Dit d'una altra manera, a mesura que augmenta la competència, l'equilibri va situant-se sobre zones cada cop més elàstiques de la funció de demanda.

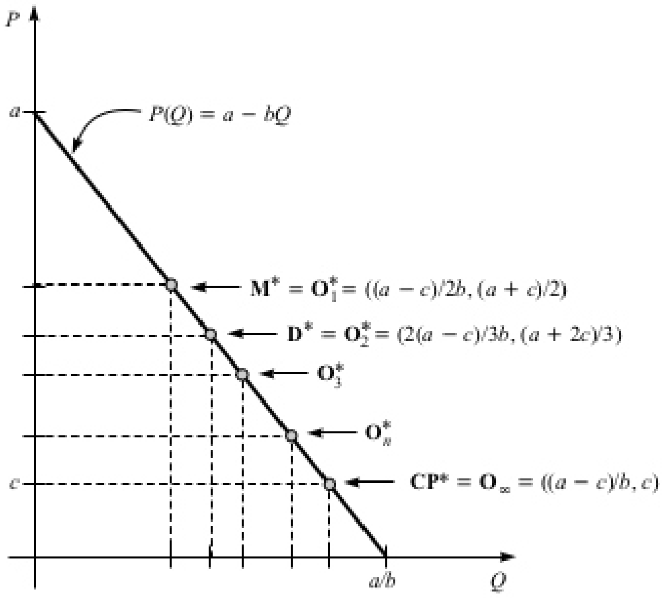


Figura 4: Quantitat i preu d'equilibri, segons el nombre d'empreses

Exemple 3.1. Suposem que ens trobem en un duopoli ($n = 2$), on $a = 130$ i $b = 1$ que determinen la funció de demanda inversa:

$$\mathrm{P(Q)} = \begin{cases} 130 - Q & \text{si } \mathrm{Q} < 130 \\ 0 & \text{si } \mathrm{Q} \geq 130 \end{cases} \quad (\text{on } Q = q_1 + q_2)$$

i suposem que a les dues empreses produir una unitat els hi representa el mateix cost, $c = 10$, sent les funcions de cost:

$$C_i(q_i) = 10q_i, \, \forall i \in \{1, 2\}$$

Si mirem a la taula dels resultats d'equilibri, observem que, per simetria, les quantitats a produir individualment seran $q_i^* = 40$, a un preu de $P^* = 50$, generant un benefici individual de $u_i^* = 1600$.

Ara bé, hi ha alguna manera en la que les dues empreses puguin obtenir beneficis per un valor superior a 1600?

Anem a veure ara el cas on les dues empreses formen un monopoli (càrtel) i prenen decisions com si fossin un únic agent.

El nostre objectiu és maximitzar la funció de beneficis $u_{1+2}(q_1, q_2) = u_1(q_1) + u_2(q_2)$.

$$\max_{q_i \in [0,130]} (130 - (q_1 + q_2)) \cdot (q_1 + q_2) - 10q_1 - 10q_2$$

Per obtenir les condicions de primer ordre, fem $\nabla = 0$;

$$\begin{cases} \frac{\partial u_{1+2}}{\partial q_1} = 120 - 2q_1 - 2q_2 = 0 \\ \frac{\partial u_{1+2}}{\partial q_2} = 120 - 2q_1 - 2q_2 = 0 \end{cases}$$

La solució d'aquest sistema és $q_1 = q_2 = 30$, fent així que el preu d'una unitat fos de $P = 130 - 60 = 70$, representant uns beneficis individuals de $u_1 = u_2 = 1800$. Com hem vist en altres exemples, el perfil d'estratègies $(30, 30)$ domina al perfil $(40, 40)$ en el sentit de Pareto, ja que cada empresa obté uns beneficis superiors, però el perfil no és d'equilibri. Això es deu al fet que les dues empreses saben que l'altre produirà 30 unitats, aleshores una d'elles pot incomplir l'acord i produir-ne 45, maximitzant així els seus beneficis.

Més endavant veurem com la repetició infinita de la competència en el model de Cournot pot donar lloc a una situació de cooperació, en la que les empreses condicionen el seu comportament present a la possibilitat de ser recompensats o castigats en el futur.

4 Jocs repetits en infinites etapes

Un dels problemes als quals ens enfrontem quan tenim en compte l'horitzó temporal en què pot desenvolupar-se un joc és la valoració dels resultats en diferents moments del temps. Té el mateix valor o utilitat per un individu una unitat monetària avui que dins d'un any?, És constant la utilitat que obtenim d'un mateix resultat o d'un mateix bé quan les rebem en diferents moments del temps? En general, la resposta és no. Per veure més a fons els jocs repetits tant finitament, com infinitament des d'un vessant més econòmic, consultar [8] i [10], i per un vessant més matemàtic [13] i [20].

El valor o utilitat que els jugadors atribueixen a rebre una quantitat de diners (o qualsevol altre factor al qual atribueixin utilitat) té un caràcter temporal, depèn del moment en què hagi de produir-se la recepció.

Definició 4.1. *Donat un jugador i un període temporal τ, anomenem factor de descompte d'aquest agent, en aquest període, al nombre real positiu entre 0 i 1, δ tal que dit agent és indiferent entre cobrar la quantitat δC en un instant t o la quantitat C en l'instant $t+\tau$.*

Dit d'una altra manera, el factor de descompte δ és el valor, en unitats de pagament actuals, d'una unitat de pagament que es farà efectiva un període més tard. També podem mirar a δ com la probabilitat de què un joc es repeteixi acabada una etapa. Per exemple, si tenim $\delta = \frac{1}{2}$, la probabilitat de jugar una segona etapa d'un joc és $\delta = \frac{1}{2}$, la d'una tercera $\delta^2 = \frac{1}{4}$ i així amb totes les possibles futures etapes. Els valors extrems, $\delta = 0$ i $\delta = 1$, corresponen, respectivament, al cas en què l'agent només valora el present (impaciència extrema) i al cas en què valora en igual mesura el futur que el present (paciència extrema).

Definició 4.2. *Sigui $\{q_t\}_{t=1,2,\ldots}$ una seqüència o flux de pagaments, finita o infinita.*

Donat una factor de descompte δ, el valor present descomptat de la seqüència de pagaments $\{q_t\}_{t=1,2,\ldots}$ es denota $VP[\{q_t\}_{t=1,2,\ldots}, \delta]$, i es defineix com:

$$VP[\{q_t\}_{t=1,2,\ldots}, \delta] = q_1 + \delta q_2 + \delta^2 q_3 + \ldots + \delta^{t-1} q_t + \ldots = \sum_t q_t \delta^{t-1}$$

Hem vist que en alguns jocs estàtics com el dilema del presoner o el duopoli de Cournot, la cooperació de tots els jugadors és un resultat desitjable des del punt de vista del conjunt dels jugadors (és un òptim de Pareto), però no és previsible, ja que l'únic equilibri de Nash és el comportament egoista de tots. Seguirà passant el mateix si el joc es repetís? O pel contrari és possible que la contradicció entre el benestar individual i el social desaparegui quan el joc es repeteix?

Definició 4.3. *Sigui un joc d'etapa qualsevol $G = \{A_1, \ldots, A_n; u_1, \ldots, u_n\}$, on A_i són els conjunts d'estratègies de cada jugador i g_i les funcions de pagament de cada jugador, on els pagaments estan acotats (existeix un valor $a \in \mathbb{R}$, tal que $|u_i(a_1, \ldots, a_n)| < a$, $\forall\, i$, $\forall (a_1, \ldots, a_n)$). Sigui $\delta = (\delta_1, \delta_2, \ldots, \delta_n)$ un vector de factors de descompte, on $\delta_i < 1$, $\forall\, i$. Anomenarem joc repetit infinitament a $G^\infty(\delta)$ al joc que compleix:*

1. *Abans de començar el joc, és de domini públic:*
    - *El factor de descompte de cada jugador i és δ_i.*
    - *Després de qualsevol etapa k, el joc pot prosseguir en l'etapa següent.*
    - *Els pagaments de $G^\infty(\delta)$ són, per a cada jugador i, el valor present (pel factor de descompte δ_i) de la successió infinita dels seus pagaments d'etapa.*

2. *Abans de començar qualsevol etapa, són de domini públic les jugades realitzades en totes les etapes anteriors.*

Definició 4.4. *Donat el factor de descompte $\delta < 1$, el valor present d'una seqüència infinita de pagaments $\{u_i^t\}_{t=1}^{\infty}$ d'un jugador i és:*

$$u_i = u_i^1 + \delta u_i^2 + \delta^2 u_i^3 + \ldots + \delta^{t-1} u_i^t + \ldots = \sum_{t=1}^{\infty} \delta^{t-1} u_i^t$$

A diferència dels jocs finitament repetits, necessitem que el factor de descompte sigui menor a 1 perquè aquesta expressió estigui ben definida. Això es dona, perquè per $\delta = 1$ aquesta suma podria no estar ben definida si hi ha infinits pagaments d'etapa positius. En aquest cas la suma valdria infinit. Si els pagaments u_i^t estan acotats, aleshores amb un factor de descompte $\delta < 1$ aquesta suma estarà ben definida.[7]

Aquest descompte és una suposició econòmica natural per a una sèrie de pagaments. La qüestió és, però, si és raonable assumir que els jugadors s'endinsaran en una seqüència infinita de joc. Podríem dir que una resposta sensata, és que no ho farien. Afortunadament, podem donar una motivació raonable per aquesta idea en lloc de fer servir la interpretació de l'interès del descompte, considerant una interpretació alternativa d'un futur incert.

Imaginem que els jugadors juguen una etapa avui, i amb una certa probabilitat $\delta < 1$ tornaran a jugar demà altre cop, i amb probabilitat $1 - \delta$ s'acaba el joc. Ara, suposem que en aquest esdeveniment, els jugadors continuen jugant, els pagaments del jugador i venen donats per la sèrie $\{u_i^t\}_{t=1}^{\infty}$. Si no hi ha cap descompte addicional, podem pensar en els pagaments esperats d'aquesta seqüència com,

$$Eu = u_i^1 + \delta u_i^2 + (1-\delta)\cdot 0 + \delta^2 u_i^3 + \delta(1-\delta)\cdot 0 + \delta^3 u_i^4 + \delta^2(1-\delta)\cdot 0 = \sum_{t=1}^{\infty} \delta^{t-1} u_i^t$$

És a dir, podem pensar en el valor present dels pagaments com el valor esperat obtingut en jugar a aquest a joc amb aquesta sèrie de pagaments. De fet, hi ha alguna cosa més sobre la interpretació de $\delta < 1$ com la probabilitat que els jugadors no trenquin el seu joc continu. Concretament, quina és la probabilitat que els jugadors juguin infinitament aquest joc? La resposta és $\delta^{\infty} = 0$. És a dir, el joc acabarà en un temps infinit amb probabilitat 1, però el futur possible sempre és infinit.

Abans de continuar amb l'anàlisi dels jocs repetits infinitament, ens serà útil introduir la següent definició:

Definició 4.5. *Donat $\delta < 1$, el pagament mitjà de una seqüència infinita de pagaments $\{u_i^t\}_{t=1}^{\infty}$ és:*

$$\overline{u_i} = (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i^t$$

És a dir, el pagament mitjà d'una seqüència és una normalització del valor net present: reduïm el valor net actual per un factor de $1 - \delta$.

[7] Per veure això, suposem que existeix una cota $b > 0$ tal que $u_i^t \leq b$ per tot t. Primer observem que $u_i = \sum_{t=1}^{\infty} \delta^{t-1} u_i^t \leq \sum_{t=1}^{\infty} \delta^{t-1} b$. Ens queda veure que $B = \sum_{t=1}^{\infty} \delta^{t-1} b$ és finit. Podem posar $B(1-\delta) = b$ perquè $\lim_{t\to\infty} \delta^{t-1} b = 0$, el que implica que $B = \frac{b}{1-\delta} < \infty$.

Això és convenient per la següent propietat. Considerem un jugador que té una seqüència finita de pagaments v en cada període $t \in \{1, 2, \ldots, T\}$. El seu valor present és la suma:

$$s(v, \delta, T) = v + \delta v + \delta^2 v + \ldots + \delta^{T-2} v + \delta^{T-1} v$$

Considerem aquesta suma multiplicada per δ:

$$\delta s(v, \delta, T) = \delta v + \delta^2 v + \delta^3 v \ldots + \delta^{T-1} v + \delta^T v$$

Restant les dues expressions anteriors obtenim $s(v, \delta) - \delta s(v, \delta) = v - \delta^T v$ que implica:

$$s(v, \delta, T) = \frac{v - \delta^T v}{1 - \delta}$$

Si ara prenem el límit quan T tendeix a infinit (quan el joc es repeteix infinitament):

$$s(v, \delta) = \lim_{t \to \infty} s(v, \delta, T) = \frac{v}{1 - \delta}$$

Com és relaciona això amb el pagament mitjà? Si mirem al valor net present d'un seguit de pagaments fixats v, aleshores la definició de pagament mitjà és:

$$\overline{v} = (1 - \delta) s(v, \delta) = v$$

És a dir, el pagament mitjà d'un conjunt infinit d'algun valor v, és ell mateix. Com veurem una mica més endavant, aquest fet ens ajudarà a identificar quins resultats, definits pel pagament mitjà obtingut, poden ser equilibris perfectes d'un subjoc, d'un joc repetit.

4.1 Històries, estratègies i subjocs

Per poder definir amb rigor els conceptes d'estratègia i de subjoc en un joc repetit, és útil precisar en primer lloc, el concepte d'història del joc en un període o etapa concret.

Definició 4.6. *Donat el joc d'etapa $G = \{A_1, \ldots, A_n; u_1, \ldots, u_n\}$ en forma estratègica i el joc repetit $G^\infty(\delta)$, les t-històries o històries del joc fins el moment t, denotades h_t, recullen tota l'expèriencia passada del joc fins arribar a aquell moment, és a dir, totes les decisions preses pels n jugadors en les t-1 etapes anteriors. El conjunt de totes les t-històries h_t és:*

$$H_t = \left\{ \{a_1^k, a_2^k, \ldots, a_n^k\}_{k=1,2,\ldots,t-1} | a_i^k \in A_i \right\}$$

on a_i^k, pertanyent a A_i, ha d'interpretar-se com l'acció realitzada pel jugador i en l'etapa k.

Les estratègies s_i del jugador i en $G^\infty(\delta)$ són plans d'accions que determinen quina acció realitzarà tal jugador en cada etapa per cada possible història fins aquell moment, i poden descriure's simbòlicament així:

$$s_i = (a_i^1(h_1), a_i^2(h_2), \ldots, a_i^k(h_k), \ldots).$$

Un perfil d'estratègies $\sigma = (\sigma_1, \ldots, \sigma_i, \ldots, \sigma_n)$ del joc repetit $G^{\infty}(\delta)$ determina (a través d'un procés iteratiu) la successió de combinacions d'accions

$$\left\{a_1^k(h_k(\sigma)), a_2^k(h_k(\sigma)), \ldots, a_n^k(h_k(\sigma))\right\}_{k\geq 1}$$

una en cada etapa k, que es posaran en pràctica de manera efectiva i que constitueix la trajectòria del joc determinada per σ. I els corresponents pagaments d'etapa determinen, d'acord amb el factor de descompte, els pagaments globals per a cada jugador del joc repetit infinitament.

Els subjocs de $G^{\infty}(\delta)$ comencen en iniciar-se cada etapa (existeixen tants que comencin en l'etapa t com t-històries siguin possibles) i són idèntics en la seva estructura al joc global.

4.2 ENPS en el dilema del presoner repetit infinitament

El joc infinitament repetit que és més conegut i estudiat, és justament el que té al Dilema del Presoner (DP) com joc d'etapa. L'anomenarem $\mathrm{DP}^{\infty}(\delta)$, i veurem que la seva anàlisi dóna lloc a resultats molt diferents dels que coneixem pel cas de repetició finita.

Alguns perfils que no són EN del dilema del presoner repetit infinitament, són els següents:

- *Perfil d'estratègies incondicionades de cooperació ingènues.* El perfil de cooperació (*Callar sempre*, *Callar sempre*) determinaria una trajectòria de cooperació completa del joc, consistent en (*Callar*, *Callar*) en cada etapa. Però tal perfil no és un EN, independentment del valor de δ. En efecte, la millor resposta a *Callar* sempre del jugador 1, per part del jugador 2, no és *Callar* sempre, que li representaria un pagament mitjà de 4, sinó *Confessar* sempre, que li representaria un pagament mitjà de 5.

- *Perfil d'estratègies condicionades, però errònies.* El perfil $\sigma = (\sigma_1, \sigma_2)$ en el que $\sigma_1 = \sigma_2 =$ jugar en l'etapa $t > 1$ l'acció que el contrari ha jugat en l'etapa $t - 1$, i en la primera etapa *Confessar*, no és un EN, sempre que el valor de δ sigui suficientment alt. En efecte, la millor resposta a σ_1 del jugador 1, no és σ_2, ja que li representa un flux de pagaments $(1, 1, \ldots, 1, \ldots)$, amb valor present $\frac{1}{1-\delta}$, al jugador 2, mentre que $\sigma_2' =$ *Callar sempre* li representaria un flux de pagaments $(0, 4, 4, \ldots, 4, \ldots)$, amb valor present $\frac{4\delta}{1-\delta}$, que és estrictament major que $\frac{1}{1-\delta}$ sempre que $\delta > \frac{1}{4}$.

Passem a analitzar un perfil que, per valors suficientment alts de δ, és EN, però no ENPS.

Perfil d'estratègies ***Ull per ull***

Considerem l'estratègia Ull per Ull = jugar en l'etapa $t > 1$ l'acció que el contrari ha jugat en l'etapa $t - 1$, i en la primera etapa *Callar*. El perfil $\sigma = (\sigma_1, \sigma_2)$ en el que $\sigma_1 = \sigma_2 =$ Ull per Ull, determinaria una trajectòria de cooperació completa del joc, consistent en (*Callar*, *Callar*) en cada etapa, ja que ambdós jugadors comencen no acusant, el que implicaria que continuïn no acusant en cada etapa posterior. Per altra part, el perfil σ, si és un EN, sempre que el valor de δ sigui suficientment alt, però no és un ENPS. Provem-ho:

- σ és un EN. En efecte, $\sigma_1 =$ Ull per ull, és resposta òptima per part del jugador 1 a $\sigma_2 =$ Ull per ull del jugador 2 (i viceversa), ja que seguir σ_1 assegura un pagament de 4 en cada etapa, a causa d'un resultat (*Callar*, *Callar*) en cada etapa, mentre que qualsevol desviació de σ_1 el perjudicarà. Observem: si es desvia de l'acció de *Callar* en alguna etapa, aconsegueix en ella un pagament de 5, però a costa de ser castigat i no ser perdonat fins que triï de nou l'acció *Callar*. Suposem que la tanda de desviació dura dues etapes (que podríem denominar traïció i penediment). El jugador 1 obtindria els pagaments:

  - En l'etapa de traïció rebria un pagament de 5.
  - En l'etapa de penediment rebria un pagament de 0.

    El valor present d'aquest interval és:

  $$\text{Amb desviació } (\sigma_1') \qquad v' = 5 + \delta \cdot 0 = 5$$

  $$\text{Sense desviació } (\sigma_1) \qquad v = 4 + \delta \cdot 4$$

  Així doncs, serà millor σ_1 sempre i quan $4 + \delta \cdot 4 \geq 5$, el que succeix si $\delta \geq 1/4$.

  Suposem ara que la tanda de desviació dura $n+2$ etapes (una de traïció, n de càstig, i una de penediment i perdó). El jugador 1 obtindria els pagaments:

  - En l'etapa de traïció rebria un pagament de 5.
  - En les n etapes de càstig rebria un pagament d'1.
  - En l'etapa de penediment rebria un pagament de 0.

  El valor present d'aquest interval és:

  $$\text{Amb desviació } (\sigma_1') \qquad v' = 5 + \delta \cdot 1 + \ldots + \delta^n \cdot 1 + 0 = 5 + (\delta(\delta^n - 1)/(\delta - 1)$$

  $$\text{Sense desviació } (\sigma_1) \qquad v = 4 + \delta \cdot 4 + \ldots + \delta^n \cdot 4 + \delta^{n+1} \cdot 4 = 4(\delta^{n+2} - 1)/(\delta - 1)$$

  Així doncs, és millor σ_1 sempre i quan:

  $$4(\delta^{n+2} - 1)/(\delta - 1) \geq 5 + (\delta(\delta^n - 1)/(\delta - 1);$$

  $$4\delta^{n+2} - 4 \leq 4\delta - 5 + \delta^{n+1};$$

  $$4\delta(\delta^{n+1} - 1) \leq \delta^{n+1} - 1;$$

  $$\delta \geq 1/4$$

Suposem, un últim cop, que la tanda de desviació és per sempre (una de traïció, i en endavant càstig). El jugador 1 obtindria els pagaments:

- En l'etapa de traïció rebria un pagament de 5.
- En les infinites etapes de càstig rebria un pagament d'1.

El valor present d'aquesta successió infinita de pagaments és:

$$\text{Amb desviació } (\sigma_1') \qquad v' = 5 + \delta \cdot 1 + \ldots + \delta^n \cdot 1 + \ldots = 5 + (\delta/(1-\delta))$$

$$\text{Sense desviació } (\sigma_1) \qquad v = 4 + \delta \cdot 4 + \ldots + \delta^n \cdot 4 + \ldots = 4/(1-\delta)$$

Així doncs, és millor σ_1 sempre i quan:

$$4/(1-\delta) \geq 5 + (\delta/(1-\delta));$$

$$4 \geq 5(1-\delta) + \delta;$$

$$\delta \geq 1/4$$

En conclusió, el perfil d'estratègies Ull per ull és un EN si $\delta \geq 1/4$.

- σ no és un ENPS. En efecte, en el subjoc en el que el resultat d'etapa anterior és (*Callar*, *Confessar*) al jugador 1 li convé desviar-se de σ_1 si el jugador 2 continuarà amb σ_2. De fet, si no es desviés, es produiria una cadena infinita (*Confessar*, *Callar*), (*Callar*, *Confessar*), (*Confessar*, *Callar*), (*Callar*, *Confessar*), ... , amb pagaments:

$$5 + 0 \cdot \sigma + 5 \cdot \sigma^2 + 0 \cdot \sigma^3 + 5 \cdot \sigma^4 + \ldots = 5(1 + \delta + \delta^2 + \delta^4) + \ldots = 5/(1-\delta^2)$$

  Li interessaria més *Callar* d'ara endavant, el que el conduiria a la cadena infinita (*Callar*, *Callar*), (*Callar*, *Callar*), (*Callar*, *Callar*),..., amb pagaments $4+4\delta+4\delta^2+\ldots = 4/(1-\delta)$, que és major o igual sempre que $4 \geq 5/(1+\delta)$, que succeeix per tot $\delta \geq 1/4$. En definitiva:

  - Si $\delta \geq 1/4$ el perfil Ull per Ull (*Ull per ull*, *Ull per ull*) és un EN, però no és un ENPS.
  - Si $\delta < 1/4$ el perfil Ull per Ull (*Ull per ull*, *Ull per ull*) no és un EN.

Fixem-nos que s'ha aconseguit, per primer cop, mantenir la cooperació indefinida en el dilema del presoner, mitjançant un EN (tot i no ser un ENPS), que utilitza, diguem així, l'amenaça del càstig a partir de l'acusació, equilibrant la balança de la cooperació.

4.2.1 ENPS

Passem a veure ara dos ENPS del dilema del presoner repetit infinitament.

1. *Perfil d'estratègies incrèdules.* El perfil $\sigma = (\sigma_1, \sigma_2)$ en el que $\sigma_1 = \sigma_2 =$ Confessar sempre, que determinaria una trajectòria de nul·la cooperació del joc, consistent en (*Confessar*, *Confessar*) en cada etapa, si és un ENPS. Provem-ho:

- σ és un EN. En efecte, és evident que σ_1 és resposta òptima per part del jugador 1 a σ_2 del jugador 2 (i viceversa).
- A més, la restricció del perfil σ a qualsevol subjoc, torna a ser aquest mateix perfil, aleshores és un EN en tot subjoc. En definitiva, és ENPS.

2. *Perfil d'estratègies de tirador.* Considerem l'estratègia *Callar* en la primera etapa, i a més *Callar* en l'etapa t si la història del joc fins a aquell moment ha estat sempre *Callar* per part d'ambdós jugadors, i *Confessar* en cas contrari. Dit d'altra manera, començar cooperant i continuar així fins que la cooperació es trenqui. A partir d'aquest moment, no cooperar mai més. Aquesta estratègia, que incorpora una amenaça de penalització permanent davant de qualsevol desviació de la conducta cooperadora *Callar*, rep el nom d'estratègia de tirador.[8] El perfil $\sigma = (\sigma_1, \sigma_2)$ en el que $\sigma_1 = \sigma_2 =$ ET són les estratègies de tirador de cada jugador, determinaria una trajectòria de completa cooperació dels jocs (ambdós jugadors comencen cooperant, i després continuen cooperant, ja que la cooperació no es trenca si se segueixen aquestes estratègies). Doncs bé, el perfil $\sigma = (\sigma_1, \sigma_2)$ és un ENPS si el factor de descompte δ és suficientment alt. Provem-ho:

   - σ és un EN. Veiem si l'estratègia $\sigma_i =$ ET és resposta òptima del jugador i a l'estratègia $\sigma_j =$ ET del jugador j.

     Responent amb $\sigma_i =$ ET, el jugador i obté els següents pagaments corresponents als següents jocs d'etapa:

Etapa	t	$t+1$
Perfil jugat	*(Callar, Callar)*	*(Callar, Callar)*
Pagaments de i	4	4

     i actualitzant els pagaments al moment t, el jugador i obté:

     $$\mathrm{VP}_i(\sigma_i, \sigma_j) = 4 + 4\delta + 4\delta^2 + 4\delta^3 + \ldots = 4/(1-\delta)$$

     Si pel contrari, i respongués amb qualsevol altra estratègia σ_i', els pagaments que obtindria a partir del moment en que es desviés jugant *Confessar* serien:

Etapa	t	$t+1$
Perfil jugat	*(Confessar, Callar)*	*(Confessar, Confessar)*
Pagaments de i	5	1

     (Un cop i ha decidit no cooperar, j dispara la penalització i a partir d'aquest moment, a i sempre li convindrà no cooperar (*Confessar*), perquè j passarà a jugar sempre *Confessar*).

     Actualitzant els pagaments al moment t (moment en el que i decideix no cooperar), i obté:

     $$\mathrm{VP}_i(\sigma_i', \sigma_j) = 5 + \delta + \delta^2 + \delta^3 + \ldots = 5 + \delta/(1-\delta) = (5-4\delta)/(1-\delta)$$

     Per quins valors de δ és l'estratègia $\sigma_i =$ ET, del jugador i una resposta òptima a l'estratègia $\sigma_j =$ ET, del jugador j? Ho serà si $4/(1-\delta) \geq (5-4\delta)/(1-\delta)$, és a dir, si $\delta \geq 1/4$.

[8] L'expressió original en anglès és *Grim Trigger Strategy*, traduït com *Estrategia de apretar el gatillo* en castellà o *Estratègia de prémer el gatell* en català. Com no hi ha un conveni per la traducció, l'anomenarem estratègia del tirador.

- σ és un ENPS. Veiem per quins valors de δ el perfil σ determina un EN en cada subjoc. Per poder respondre a aquesta pregunta, és necessari considerar tots els possibles subjocs que presenta el joc repetit infinitament. Des del punt de vista de les estratègies de tirador, dits subjocs poden classificar-se en dos tipus:
  - Subjocs pels quals la història passada ha estat sempre de cooperació, és a dir, aquells en els que amb anterioritat sempre s'ha jugat *(Callar,Callar)* en cada joc d'etapa. Aquests subjocs són equivalents al joc original, tenen estructura de $DP^{\infty}(\delta)$ (Dilema del Presoner infinit) i comencen en les mateixes circumstàncies que en la primera etapa del joc original sobre el qual s'aplica el perfil d'estratègies de tirador. De manera que en aquests subjocs, el perfil σ d'estratègies de tirador, genera un EN si i només si $\delta \geq 1/4$, com acabem d'analitzar.
  - Subjocs en els que en algun moment de la història passada no s'ha cooperat (algun jugador ha decidit *Confessar* en alguna etapa prèvia). En aquests subjocs s'exhibeix un comportament no cooperatiu per part d'ambdós jugadors com millor resposta en cada etapa, ja que la resposta òptima a l'estratègia de tirador (quan la penalització s'està executant) és seguir la pròpia estratègia de tirador o l'estratègia incondicionada que hem anomenat incrèdula, consistent en *Confessar* sempre (ambdues estratègies són equivalents si en alguna etapa prèvia algun jugador no ha cooperat). Així doncs, també en aquest tipus de subjocs σ genera un EN per tot δ.

En conclusió:

En el dilema del presoner repetit infinitament, el perfil d'estratègies de tirador és un ENPS si i només si $\delta \geq 1/4$.

4.2.2 Teorema de Friedman

Hem vist que quan es repeteix infinitament un joc com el dilema del presoner, apareixen nous EN i nous ENPS, inclús alguns que sustenten de manera permanent la cooperació. Cal pensar que si això succeeix amb un joc d'etapa tan especial com el dilema del presoner, on l'únic EN és un equilibri en estratègies estrictament dominants, quelcom semblant hauria de succeir amb jocs en què la cooperació no és tan difícil com en aquest.

Per altra banda, sembla estar sorgint un aspecte negatiu de la repetició infinita, la possibilitat que existeixin molts ENPS del joc repetit, fins al punt de poder fer problemàtica la tria d'un d'ells.

Val la pena, en conseqüència, esbrinar quins resultats (en termes de vectors de pagaments) són assolibles a través d'equilibris de Nash perfectes en subjocs, ja que això ens indicarà quines possibilitats hi ha de millora en el sentit de Pareto, i si hi ha moltes o poques possibilitats.

El Teorema de Friedman, que el propi Friedman va establir l'any 1971 i s'enuncia a continuació, és un dels anomenats Teoremes de Tradició Popular (*Folk Theorems*) i ens dóna una resposta precisa a la qüestió anterior. Trobem diverses maneres de demostrar aquest teorema a [14], [16], [18] i [19], d'una manera menys formal i tres de més rigoroses a [11] i [17].

Teorema 4.7. *Sigui G un joc finit, estàtic i amb informació completa, sigui el perfil $\alpha^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_n^*)$ un EN del joc G i sigui $g(\alpha^*) = (g_1(\alpha^*), g_2(\alpha^*), \ldots, g_n(\alpha^*))$ el vector de pagaments corresponent a α^*. Si $v = (v_1, v_2, \ldots, v_n)$ és un vector de pagaments factible de G que compleix $v_i > g_i(\alpha^*)$ per qualsevol jugador i, aleshores existeix un factor de descompte δ_0 tal que per tot $\delta > \delta_0$ existeix un ENPS de $G^\infty(\delta)$, que obté v com vector de pagaments mitjà (és a dir, v_i és el pagament mitjà del jugador i pel resultat del joc quan s'executa dit perfil d'equilibri).*

Demostració. Només considerarem el cas en que existeix un perfil $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ d'estratègies (pures o mixtes) del joc d'etapa G que produeix el vector de pagaments $v = (v_1, v_2, \ldots, v_n)$.

Considerem ara el perfil estratègic $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ en $G^\infty(\delta)$, en el que σ_i és l'estratègia següent del jugador i:

«En la primera etapa jugar α_i. En qualsevol etapa posterior jugar α_i si fins el moment s'ha jugat el perfil d'etapa $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ en totes les etapes, i jugar α_i^* en cas contrari ».

El perfil estratègic σ és un perfil en estratègies de tirador, ja que cada jugador i comença jugant l'estratègia cooperadora α_i, però amenaça amb jugar per sempre l'estratègia penalitzadora $\alpha_i *$.

Veiem que σ és un EN de $G^\infty(\delta)$ si δ és suficientment gran.

- Suposem que tots els jugadors excepte i es conformen al perfil σ. Si en alguna etapa hi ha hagut desviació respecte al perfil d'etapa α, és clar que jugar σ_i d'ara endavant és resposta òptima de i a σ_{-i} (doncs prescriu respondre d'ara endavant amb α_i^* en cada etapa a α_{-i}^*).

- Suposem ara que estem en la primera etapa, o en una etapa tal que no hi ha hagut desviació prèvia respecte al perfil d'etapa α, i comprovarem que també σ_i és resposta òptima a σ_{-i}.

Sigui α'_i la desviació òptima que podria adoptar i en una etapa t sense desviacions prèvies respecte a α. Dita desviació li proporcionarà a i un benefici addicional en l'etapa de d_i, és a dir, rebria un pagament equivalent a v_i+d_i, però ocasionaria el desencadenament de les penalitzacions de la resta de jugadors, que a partir d'aquesta etapa jugarien les seves estratègies de càstig en α^*, obligant a i a defensar-se (òptimament) amb la seva pròpia estratègia α_i^* en qualsevol etapa posterior. Per tant, el flux de pagaments òptim que obtindria i en cas de desviar-se seria:

Etapa	t	$t+1$	$t+2$	...
Perfil jugat	(α'_i, α_{-i})	α^*	α^*	...
Pagament de i	$v_i + d_i$	$g_i(\alpha^*)$	$g_i(\alpha^*)$	...

el valor present del qual és:

$$VP' = v_i + d_i + \delta g_i(\alpha^*) + \delta^2 g_i(\alpha^*) + \ldots = v_i + d_i + \delta g_i(\alpha^*)/(1-\delta).$$

Mentre que si i es conformés al perfil σ obtindria el flux de pagaments:

Etapa	t	$t+1$	$t+2$	...
Perfil jugat	$\alpha = (\alpha_i, \alpha_{-i})$	α	α	...
Pagament de i	v_i	v_i	v_i	...

el valor present del qual és $VP = v_i + \delta v_i + \delta^2 v_i + \ldots = v_i/(1-\delta)$.

Doncs bé, σ_i és resposta òptima de i a σ_{-i} si VP $\geq$ VP', és a dir, si:

$$v_i/(1-\delta) \geq v_i + d_i + \delta g_i(\alpha^*)/(1-\delta);$$

$$v_i \geq (1-\delta)(v_i + d_i) + \delta g_i(\alpha^*);$$

$$0 \geq -\delta v_i + d_i - \delta d_i + \delta g_i(\alpha^*);$$

$$\delta \geq d_i/(v_i + d_i - g_i(\alpha^*))$$

En conclusió, si $\delta \geq max\,\{d_i/(v_i + d_i - g_i(\alpha^*))\}_{i=1,2,\ldots,n}$, σ és un EN de $G^\infty(\delta)$. Veiem ara que σ és un EN per qualsevol subjoc de $G^\infty(\delta)$, i ho és per aquests mateixos valors de δ. Tal com vam raonar en el cas del dilema del presoner, des del punt de vista de les estratègies de tirador de σ, els subjocs de $G^\infty(\delta)$, tots ells iguals a $G^\infty(\delta)$, poden classificar-se en dos tipus:

- Subjocs pels que la història passada ha estat sempre de cooperació, jugant el perfil α en cada etapa. Aquests subjocs comencen en les mateixes circumstàncies que en la primera etapa del joc original sobre el qual s'aplica el perfil σ. De manera que en aquests subjocs el perfil σ genera un EN si i només si

$$\delta \geq max\,\{d_i/(v_i + d_i - g_i(\alpha^*))\}_{i=1,2,\ldots,n}.$$

- Subjocs en els que en algun moment de la història passada no s'ha cooperat (algun jugador i ha decidit desviar-se de σ_i en alguna etapa prèvia). En aquests subjocs σ genera un EN consistent en jugar α^* en cada etapa. En conclusió, si $\delta \geq max\{d_i/(v_i + d_i - g_i(\alpha^*))\}_{i=1,2,\ldots,n}$, σ és un ENPS de $G^\infty(\delta)$.

□

Observació 4.8. Tot i que aquest teorema s'enuncia per jocs d'etapa finits, també és vàlid en jocs d'etapa infinits, com l'oligopoli de Cournot i altres jocs importants en economia.

Per apreciar visualment el significat del teorema de Friedman, representem les regions de pagaments factibles corresponents al dilema del presoner i els pagaments d'equilibri corresponents a l'únic EN.

El teorema de Friedman ens diu, pel que fa al dilema del presoner, que qualsevol vector de pagaments factible que es trobi per sobre i a la dreta del punt $(1, 1)$ és assolible com vector de pagaments mitjans per algun ENPS de $DP^\infty(\delta)$ sempre que δ sigui suficientment gran. La regió assolible de pagaments és la següent:

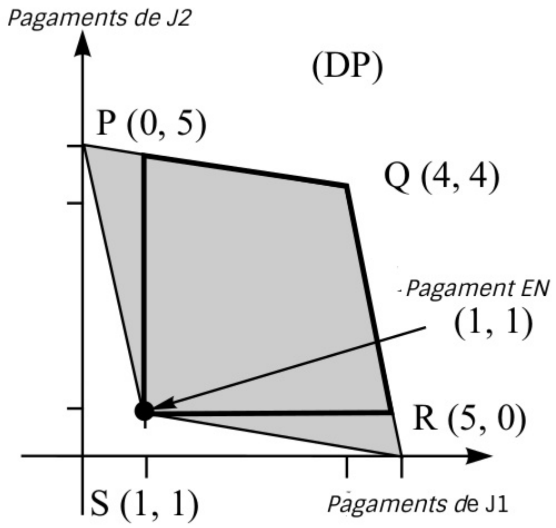


Figura 5: Vectors de pagaments factibles i EN en el dilema del presoner

Observem, per altre banda, que si $v_i < g_i(\alpha^*)$, v no pot assolir-se amb un ENPS ni amb un EN de $G^\infty(\delta)$. En efecte, li bastaria al jugador i amb jugar sempre la seva estratègia en el perfil d'etapa α^* per aconseguir $g_i(\alpha^*)$ de pagament mitjà.

4.2.3 Teorema de Friedman per jocs infinits

L'equivalent al teorema de Friedman per jocs infinits, també podria incloure's en els Teoremes de Tradició Popular (*Folk Theorems*). A partir del teorema de Friedman podem provar-lo fàcilment. Ens centrarem en el cas particular de dos jugadors, donat l'exemple que estem estudiant. La demostració que es troba a continuació està basada en [15].

Teorema 4.9. *Sigui G un joc finit, estàtic i amb informació completa, sigui el perfil $\alpha^* = (\alpha_1^*, \alpha_2^*)$ un EN del joc G i sigui $g(\alpha^*) = (g_1(\alpha^*), g_2(\alpha^*))$, que denotarem com (g_1^*, g_2^*), el vector de pagaments corresponent a α^*. Per cada parell racional (v_1, v_2) existeix un δ_0 tal que per tot $1 > \delta > \delta_0 >$ hi ha un ENPS amb pagaments (v_1, v_2).*

Demostració: Sigui (α_1^*, α_2^*) el perfil d'etapa Equilibri de Nash que reporta uns pagaments de $(g_1(\alpha^*), g_2(\alpha^*))$. Ara definim el perfil (σ_1, σ_2) que en cada etapa dóna (v_1, v_2).

Considerem la següent estratègia: «Començar jugant σ_i i continuar fent-ho mentre cap dels dos jugadors canviï. Si un dels dos es desvia, jugar α_i^* a totes les etapes futures.»

Comencem provant que l'estratègia anterior és un Equilibri de Nash.

Sense pèrdua de generalitat, si el jugador 1 es desvia a $\sigma_1^{'} \in S_1$, una altra estratègia tal que $u_1(\sigma_1^{'}, \sigma_2) > v_1$, en l'etapa k, aleshores:

$$U_1^{(k)} = \textstyle\sum_{t=1}^{k-1} \delta^{t-1} v_1 + \delta^{k-1} g_1(\sigma_1^{'}, \sigma_2) + g_1^* (\frac{1}{1-\delta} - \sum_{t=1}^{k} \delta^{t-1}).$$

Recordant que el jugador 1 rebria v_1 a cada etapa sense desviació, el benefici més gran que pot obtenir és a la primera etapa (tots els beneficis posteriors per desviar-se tenen una major càrrega del factor de descompte (són més petits)). Aleshores, si podem trobar δ_0 tal que $\delta > \delta_0$ implica que $U_1^{(1)} \leq \frac{v_1}{1-\delta}$, el jugador 1 no tindrà incentius a desviar-se.

$$U_1^{(1)} = g_1(\sigma_1^{'}, \sigma_2) + g_1^* \tfrac{\delta}{1-\delta} \leq \tfrac{v_1}{1-\delta}$$

$$(1-\delta) u_1(\sigma_1^{'}, \sigma_2) + g_1^* \delta \leq v_1$$

$$g_1(\sigma_1^{'}, \sigma_2) - v_1 \leq \delta(g_1(\sigma_1^{'}, \sigma_2) - g_1^*)$$

ja que $(\sigma_1^{'}, \sigma_2) > v_1 > g_1^*$, agafant $\delta_0 = \frac{g_1(\sigma_1^{'}, \sigma_2) - v_1}{u_1(\sigma_1^{'}, \sigma_2) - u_1^*}$ ens dóna el resultat requerit pel jugador 1 i repetint el mateix argument pel jugador 2, completa la prova de que l'estratègia mencionada anteriorment és un Equilibri de Nash.

Per construcció, aquesta estratègia és també un Equilibri de Nash Perfecte en Subjocs. Donada qualsevol història, ambdós jugadors actuaran de la mateixa manera i no tindran incentius a desviar-se:

- Si considerem un subjoc just després que un jugador s'hagi desviat de $\sigma_i^{'}$ aleshores els dos juguen α_i^*.
- Si considerem un subjoc on cap dels dos s'ha desviat de $\sigma_i^{'}$, aleshores els dos continuen jugant $\sigma_i^{'}$.

4.3 Oligopoli de Cournot simplificat repetit infinitament

Recordem l'oligopoli de Cournot. Ens trobem en un mercat on n empreses $E_1, \dots, E_n$ produeixen un mateix bé. Sigui $q_{i,t}$ la quantitat que produeix l'empresa E_i en l'etapa t, per $i = 1, 2, \dots, n$. Suposem ara que en cada període t, les empreses s'enfronten a la funció de demanda inversa:

$$P_t(q_t) = \begin{cases} a - bQ_t & \text{si } bQ_t < a \\ 0 & \text{si } bQ_t \geq a \end{cases} \quad (\text{on } b > 0 \text{ i } Q_t = q_{1,t} + q_{2,t} + \ldots + q_{n,t}, \forall t = 1, 2, \ldots)$$

i que les funcions de cost són:

$$C_{i,t}(q_{i,t}) = cq_{i,t}, \text{ on } c < a,\ \forall i \in \{1, 2, \dots, n\},\ \forall t = 1, 2, \ldots$$

Per tant, les funcions de beneficis seran:

$$u_{i,t}(q_{1,t}, q_{2,t}, \dots, q_{n,t} = q_{i,t} \cdot P(q_{i,t} + Q_{-i,t}) - C_{i,t}(q_{i,t}) = q_{i,t}(a - b(q_{i,t} + Q_{-i,t}) - c)$$

on $Q_{-i,t} = q_{1,t} + \ldots + q_{i-1,t} + q_{i+1,t} + \ldots + q_{n,t}, \forall i \in \{1, 2, \dots, n\}$, $\forall t = 1, 2, \ldots$.

Recordem que el joc d'etapa posseeix un únic equilibri de Nash consistent a fer que cada empresa E_i produeixi una quantitat $q_i^C = \frac{a-c}{(n+1)b}$, el que els generarà un benefici de $u_i^C = \frac{(a-c)^2}{(n+1)^2 b}$.

Anem a veure que el perfil cooperatiu en què cada empresa produeix la n-èsima part de la quantitat de monopoli $q_i^{m/n} = \frac{a-c}{2nb}$ amb la finalitat d'aconseguir un repartiment equitatiu dels beneficis del monopoli $u_i^{m/n} = \frac{(a-c)^2}{4nb}$, no és un EN, doncs qualsevol empresa E_i té incentius per incomplir l'acord. En efecte, la resposta òptima de E_i davant aquest suposat equilibri s'obté resolent el problema:

$$\max_{u_i}(q_i, q_{-i}^{m/n}) = [a - bq_i - (n-1)bq^{m/n}]q_i - cq_i = q_i[a - c - bq_i - (n-1)bq^{m/n}] = 0 \quad (4.1)$$

i obtenint, a través de les condicions de primer i segon ordre:

$$\frac{\partial u_i}{\partial q_i} = q_i(-b) + [a - c - bq_i - (n-1)bq^{m/n}] = 0$$

$$\frac{\partial^2 u_i}{\partial q_i^2} = -2b < 0 \text{ (màxim)}$$

el següent resultat:

$$q_i^r = \frac{(a-c)(n+1)}{4nb}$$

que proporciona a l'empresa que es desvia de l'acord uns beneficis de:

$$u_i^r = \frac{(a-c)^2(n+1)^2}{16n^2 b} > \frac{(a-c)^2}{4nb} = u_i^{m/n},\ \forall n > 1$$

4.4 Equilibris cooperatius, mitjançant estratègies de tirador, del joc repetit infinitament

Quan l'horitzó temporal és infinit o indeterminat, pot sostenir-se com ENPS un comportament en el qual cada empresa decideix produir $q_{i,t}^{m/n}$.

Aquest resultat s'aconsegueix quan totes les empreses segueixen la denominada estratègia de tirador, ja estudiada en el dilema del presoner i en el cas més general del teorema de Friedman. En el cas de l'oligopoli de Cournot podem definir l'estratègia del tirador en presència de n empreses de la següent manera:

«Començar en $t = 1$ triant $q_{t=1}^{m/n} = \frac{a-c}{2nb}$, enèsima part de la quantitat de monopoli. En cada etapa $t > 1$ produir $q_t^{m/n} = \frac{a-c}{2nb}$ si la història passada h_t ha donat lloc al perfil d'accions $(q_{1,k}^{m/n}, q_{2,k}^{m/n}, \ldots, q_{n,k}^{m/n}$, $\forall k > t$, produint la quantitat de l'equilibri de Cournot, $q_t^C = \frac{a-c}{(n+1)b}$, en cas contrari.»

Com en el cas del dilema del presoner analitzat en la secció anterior, per determinar si el perfil d'estratègies de tirador $(ET_1, ET_2, \ldots, ET_n)$ constitueix un ENPS, haurem de provar que constitueix un EN del joc repetit i que genera un EN en cada subjoc, és a dir, que davant de qualsevol història passada o desenvolupament del joc, $h_t \in H_t$.

Sense pèrdua de generalitat assumirem que totes les empreses presenten el mateix factor de descompte, és a dir, $\delta_1 = \delta_2 = \ldots = \delta_n = \delta < 1$. Haurem d'analitzar quan ET_i és una resposta òptima a ET_{-i}. Suposem que totes les empreses, excepte E_i, adopten l'estratègia del tirador. Donada una etapa k qualsevol, podem trobar les següents respostes òptimes depenent del desenvolupament o història prèvia a la etapa k, h_k:

1. Històries de no cooperació. En aquest cas totes les empreses, seguint la seva estratègia del tirador decidiran produir el vector $q_{-i,t}^C = (q_{1,t}^C, q_{2,t}^C, \ldots, q_{i-1,t}^C, q_{i+1,t}^C, \ldots, q_{n,t}^C)$, en tot $t \geq k$. En conseqüència l'empresa E_i produirà la quantitat de l'equilibri de Cournot $q_{i,t}^C$ en totes les etapes $t \geq k$, ja que fent això maximitzarà els beneficis de cada etapa, i per tant també la suma descomptada dels seus beneficis presents.

2. Històries de cooperació. Si pel contrari en la primera l'etapa i en tota etapa $t < k$ s'ha aconseguit el perfil d'accions $(q_{1,k}^{m/n}, q_{2,k}^{m/n}, \ldots, q_{n,k}^{m/n})$, les empreses determinaran, segons l'estratègia ET, una producció $q_{j,t=k} = q_{j,t=k}^{m/n}, \forall j \neq i$, en $t = k$ i una producció depenent de la decisió de E_i en $t > k$. Així, el comportament òptim de l'empresa E_i en $t \geq k$ dependrà de quin sigui el flux màxim de beneficis que pot aconseguir.

Donada l'estratègia ET_j, $\forall j \neq i$ si E_i segueix ET_i, aconseguirà un flux de beneficis:

Etapa	k	$k+1$	$k+2$	$k+3$	...
Pagaments de E_i	$u_i^{m/n}$	$u_i^{m/n}$	$u_i^{m/n}$	$u_i^{m/n}$	...

que li proporciona el següent valor present actualitzat en $t = k$:

$$VP_i(ET_i, \delta) = \textstyle\sum_{t=k}^{\infty} \delta^{t-k} u_{i,t}^{m/n} = (\sum_{t=k}^{\infty} \delta^{t-k}) u_i^{m/n} = \frac{1}{1-\delta} \frac{(a-c)^2}{4nb}$$

Pel contrari, si en l'etapa $t = k$ decideix una quantitat diferent $q_{i,t}^{m/n} \neq q_{i,t\neq k}^{m/n}$, provocarà que la resta d'empreses en les següents etapes decideixin una quantitat $q_{j,t} = q_{j,t}^C$ davant la qual, l'única resposta òptima es $q_{i,t} = q^C_{i,t}$, $\forall t > k$, de manera que per maximitzar el flux de beneficis presents i futurs haurà de decidir en $t = k$ la quantitat $q^r_{i,k} = \frac{(a-c)(n+1)}{4nb}$, que hem vist que maximitza el benefici de l'etapa donada la decisió de la resta d'empreses de produir:

$$q^{m/n}_{-i,k} = (q^{m/n}_{1,k}, q^{m/n}_{2,k}, \ldots, q^{m/n}_{i-1,k}, q^{m/n}_{i+1,k}, \ldots, q^{m/n}_{n,k}).$$

Així, donat el perfil ET_{-i}, si E_i es desvia de la manera indicada en $t = k$ (li direm NET_i a l'estratègia resultant), aconseguirà el següent flux de beneficis:

Etapa	k	$k+1$	$k+2$	$k+3$	...
Pagaments de E_i	u^r_i	u^C_i	u^C_i	u^C_i	...

I donat el factor de descompte $\delta < 1$, obtindrà el valor present actualitzat en $t = k$

$$VP_i(NET_i, \delta) = u^r_{i,k} + \textstyle\sum_{t=k+1}^{\infty} \delta^{t-k} u^C_{i,t} = u^r_{i,k} + u^C_i (\textstyle\sum_{t=k+1}^{\infty} \delta^{t-k})$$

$$= \tfrac{(a-c)^2(n+1)^2}{16n^2b} + \tfrac{\delta}{1-\delta}\tfrac{(a-c)^2}{(n+1)^2b}$$

Així doncs, l'estratègia de tirador serà la resposta òptima d'E_i en tota etapa que segueixi a una historia de cooperació si i només si:

$$VP_i(ET_i, \delta) \geq VP_i(NET_i, \delta);$$

$$\tfrac{1}{1-\delta}\tfrac{(a-c)^2}{4nb} \geq \tfrac{(a-c)^2(n+1)^2}{16n^2b} + \tfrac{\delta}{1-\delta}\tfrac{(a-c)^2}{(n+1)^2b};$$

$$\delta \geq \tfrac{(n+1)^2[(n+1)^2-4n]}{(n+1)^4-16n^2}$$

Això significa que el perfil en estratègies de tirador $(ET_1, ET_2, \ldots, ET_n)$ és un EN, i també un ENPS, de l'oligopoli de Cournot repetit infinitament si i només si:

$$\delta \geq \tfrac{(n+1)^2[(n+1)^2-4n]}{(n+1)^4-16n^2}$$

Es pot comprovar fàcilment que conforme augmenta el nombre d'empreses, el perfil d'estratègies del tirador requereix un major factor de descompte ($\lim_{n\to\infty} \delta = 1$).

5 Simulacions del Dilema del Presoner

Ambdues simulacions que trobarem a continuació, estan inspirades en diferents experiments realitzats a diverses universitats europees i americanes. La simulació real està principalment basada en la feta a la Universitat de Bonn [23], tot i que també pretén seguir els passos de les universitats de Califòrnia [22] i Viena i Zuric [25]. Per altra banda, la simulació a ordinador segueix el model de la universitat de Liverpool [24], però utilitza [26], [27] i [28] per a comprovar els resultats.

5.1 Simulació del Dilema del Presoner real

A través dels formularis de Google, s'han realitzat una sèrie de preguntes per intentar modelar la influència de l'experiència en el comportament final en el superjoc[9] del Dilema del Presoner. Compararem els resultats teòrics observats en el capítol anterior amb els resultats de l'experiment. Han participat 468 persones de diferents rangs d'edat[10]. Un cop recollides les dades, s'han anat creuant de manera aleatòria perquè a cada torn juguessin contra oponents anònims diferents (intentant semblar-se al màxim al joc original).

5.1.1 Introducció

En el superjoc del Dilema del Presoner finit, el mateix joc es repeteix un nombre fixat de cops, conegut pels dos jugadors. En aquest experiment, com simulava la repetició del joc infinits cops, els jugadors desconeixien quants cops es repetiria.

Els primers experiments que van intentar simular el superjoc del Dilema del Presoner van ser Lave (1965), Rapoport-Dale (1966) i Morehous (1966), tots ells en universitats dels Estats Units, on mostraven que els individus triaven més freqüentment l'opció de cooperar.

Més recentment, a la dècada dels 2000, es van realitzar experiments on els individus jugaven el mateix Duopoli de Cournot o Dilema del Presoner molts cops, contra oponents anònims que canviaven a cada ronda i contra els que no es tornarien a creuar. Els resultats mostraven un patró del comportament que es descriuria com cooperació tàcita, fins als últims torns on es canviava a opcions no cooperatives. Un cop un dels jugadors es desviava a la no cooperació, l'altre reaccionava d'igual manera i mai més se superava la cooperació. Aquest patró de cooperació, seguit d'un desviament s'observa en gairebé tots els jocs entre jugadors experimentats. Òbviament, el raonament tèoric simple del joc, no pot explicar el comportament d'individus amb experiència en Dilema del Presoner repetit. Un podria intentar explicar això assumint que la utilitat dels jugadors és diferent d'una recompensa monetària. Els jugadors poden, per exemple, valorar la cooperació com a tal, i per tant abstenir-se de comportaments no cooperatius malgrat dels incentius monetaris. Tals raonaments no són convincents en vista d'un possible desviament que indica que els incentius monetaris són més forts que el desig de cooperar per aquells que es desvien davant de la no cooperació.

Un intent d'explicar la cooperació seguida d'un desviament com a resultat del comportament completament racional, pot basar-se en la idea de tenir informació incompleta sobre

[9] En teoria de jocs, un superjoc és un joc de forma extensiva que es repeteix.

[10] Com no és possible estar repetint el joc infinits cops, els resultats d'aquest experiment estan basats en perfils creats a partir de les respostes de cada enquestat, duplicant les seves respostes.

els pagaments de l'altre jugador. Tanmateix, aquestes teories prediuen el patró d'individus sense experiència. Això no concorda amb les observacions experimentals. Els subjectes primer han d'aprendre a cooperar i només després descobreixen l'efecte de desviar-se. Les teories descriptives no poden ignorar el fet que la racionalitat humana és limitada.

En aquesta simulació, intentarem presentar una teoria de l'aprenentatge que expliqui el comportament dels jugadors en el Dilema del Presoner repetit. Assumirem que la motivació dels jugadors és entrar a la presó el mínim nombre d'anys possibles. Ens basarem en el model de Markov per fer-ho, on els individus canviaran la seva intenció de desviar-se de la cooperació en un període concret amb transició de probabilitats depenent de la seva experiència en l'últim superjoc.

5.1.2 Procediment experimental

Abans de comentar els resultats i com s'ha realitzat aquesta simulació, comentarem breument l'experiment original en el qual pretenia basar-me a l'hora de portar-lo a terme. Aquest es va realitzar a la universitat alemanya de Bonn fa vint-i-cinc anys. Cada individu va jugar 25 superjocs on a cada superjoc es jugaven 10 etapes. Es van jugar al llarg de dues sessions de 4 hores cadascuna. Cada individu es trobava aïllat i durant tot el transcurs de l'experiment no podia comunicar-se amb el seu rival. A cada sessió hi havia 12 subjectes, que van ser dividits en dos grups de 6. A la segona sessió, va fallar un dels individus i va ser substituït per una estratègia fixada pels professors que duien a terme l'experiment. L'estratègia consistia a mantenir una actitud cooperadora fins que l'oponent es desviés en algun dels nou períodes següents. Només dos dels dotze participants no van desviar-se del camí de la cooperació. Aquest experiment no només tenia fins matemàtics i econòmics, també en tenia de filosòfics i psicològics. És per això que abans de prendre cada decisió se'ls hi demanava als jugadors, que donessin motius pels quals triaven una opció o una altra a cada etapa.

L'experiment que veureu a continuació pretén semblar-se al mencionat anteriorment. Com no disposava dels recursos suficients per fer-ho, a través dels formularis de Google s'ha preguntat als diferents enquestats el següent:

Inicialment se'ls hi presentava la situació del dilema del presoner.[11]

- Si et trobessis en aquesta situació, ignorant el fet que coneixes l'altre detingut, què faries?
- Canviaria la teva decisió si coneguessis a l'altra persona?

La pregunta que ens interessa és la primera, ja que ens indica que faria cada individu si el joc es jugués només un cop. La segona pertanyeria al camp de Jocs Cooperatius, que no es tracten en aquest treball.

[11] El teu company i tu, sou dos delinqüents habituals que acabeu de ser atrapats cometent un delicte greu. No hi ha cap prova clara contra vosaltres, però si indicis forts d'aquest delicte i a més hi ha proves d'un delicte menor. Sou interrogats simultàniament en habitacions separades. Els dos sabeu: - Si calleu, sereu absolts del delicte principal per falta de proves, però us condemnaran pel delicte menor (1 any de presó). - Si els dos confesseu, sereu condemnats pel principal, però se us rebaixarà una mica la pena per confessar (4 anys de presó). - Si un sol confessa, aquest es lliurarà i a l'altre li cauran 5 anys.

Just després se'ls informava que es tornarien a trobar en aquesta situació anys després amb el mateix company, però ara coneixent quina decisió va triar el rival.

- T'informen que el primer cop, l'altre detingut va decidir confessar, quina decisió prendries?
- I si l'altra persona va optar per callar, què faries?

Finalment, per acabar de concretar quin tipus d'estratègia seguiria cadascú, se'ls hi oferien diverses possibilitats: callar sempre, confessar sempre, caòtica, ull per ull, estratègia del tirador o igualitari.[12]

5.1.3 Resultats experimentals

Durant les diferents etapes, els individus haurien après un patró de comportament on inicialment cooperarien, però acabarien sense fer-ho. Per poder precisar això, s'introdueixen les següents definicions:

Definició 5.1. *Anomenarem cooperativa al conjunt de decisions d'un joc si les següents condicions es satisfan:*

1. *Ens els primers m períodes, on m és al menys 20, els dos jugadors trien l'alternativa cooperadora callant els dos.*
2. *Entre els períodes $m+1$ i $m+5$ (on $m < 50$), al menys un dels dos jugadors tria no cooperar.*
3. *Ens els últims períodes, des del $m+6$ fins al 50 (si es que n'hi ha algun), els dos jugadors trien la opció de no cooperar.*

Notem que aquesta definició no exclou el cas on $m = 50$, on els dos jugadors cooperen de principi a fi. Certament, el requisit $m > 20$ és fins a cert punt arbitrari. Tanmateix, és necessari introduir algun criteri que ens permeti distingir jugades on inicialment hi havia cooperació, però s'acaba desviant de les que no s'ha cooperat en cap moment. És més, en l'experiment cap cas addicional hauria de ser classificat com cooperació si la primera condició on $m \geq 20$ és rebaixada a $m \geq 1$.

Definició 5.2. *Direm que hi ha hagut un desviament a la no cooperació si:*

1. *Ambdós jugadors trien la opció cooperativa al menys en vint períodes consecutius $k, \ldots, m$.*

Les altres dues condicions són les mateixes que a la definició de jugada cooperativa. Per definició, una jugada cooperativa també és una jugada amb desviament. La diferència és que la primera condició de la definició de jugada cooperativa, indica que és en els primers períodes del joc, mentre que en la de desviament pot ser en qualsevol moment.

[12] Ull per ull: jugar el mateix que el teu rival l'etapa anterior. Tirador: començar cooperant i en el moment en què el rival deixi de cooperar, canviar a confessar i no canviar mai més. Igualitari: Callar a la primera ronda. A la resta de rondes examinar la història de l'altre jugador i comptar el total de confessions i silencis. Si les confessions són més que els silencis, el jugador confessarà i si no farà silenci.

Direm que un superjoc pertany a la ronda n si va ser la n-èsima ronda jugada pels individus. Com hi ha 468 jugadors, a cada ronda tindrem 234 jocs repetits. La següent taula recull en quantes jugades s'ha observat desviament i quantes són cooperatives. Això vindrà indicat per cadascun dels 6 grups de 78 persones que interactuen per separat.

A la taula s'observa que els jugadors amb experiència tendeixen a ser cooperatius, mentre que hi ha altres que intenten guanyar avantatge triant l'alternativa no cooperativa en el primer període, esperant que el rival no contraataqui. Aquest comportament en la majoria dels casos acaba en jocs on s'observa desviament, però no arriben a ser cooperatius en el sentit de la definició donada anteriorment. Només un dels 468 jugadors sembla que tenia dificultats per entendre la situació, ja que el seu comportament és bastant irregular.

El model que estem intentant fer, contempla la intenció de desviar-se en un període concret com un estat intern i independent. En tots els casos on un individu s'ha desviat abans o simultàniament que l'oponent, la intenció de desviar-se en un període específic no és més que el període de desviació observat. No obstant, si el rival es desvia abans, aquest període no ve únicament determinat per les seves pròpies decisions.

	Grup													
	I		II		III		IV		V		VI		Total	
Ronda	DEV	CP	DEV	CP	DEV	CP	DEV	CP	DEV	CP	DEV	CP	DEV	CP
1	5	0	4	0	3	1	3	1	12	6	12	4	39	12
2	2	1	2	2	3	3	1	1	4	3	2	2	14	12
3	11	8	7	6	6	6	13	12	29	15	13	6	79	53
4	7	4	14	7	14	13	18	15	29	18	22	20	104	77
5	4	4	7	3	10	6	13	8	18	8	13	10	65	39
6	10	8	11	7	10	8	20	18	27	13	13	11	91	65
7	8	6	13	7	13	11	21	17	22	16	15	9	92	66
8	13	9	13	6	27	21	26	18	39	19	26	17	144	90
9	39	36	39	30	13	12	26	24	28	18	25	22	170	142
10	26	24	38	38	13	10	13	11	26	22	27	24	143	129
11	25	24	39	34	12	12	26	23	13	11	26	24	141	128
12	26	21	39	31	26	25	26	20	26	10	26	23	169	130
13	26	26	26	26	39	39	39	39	39	39	39	39	208	208
14	26	25	26	20	39	37	39	37	39	38	39	37	208	194
15	26	26	39	13	39	39	39	39	39	39	39	39	221	195
16	26	26	39	26	39	39	39	39	39	39	39	39	221	208
17	26	26	39	26	39	39	39	39	39	39	39	39	221	208
18	26	26	39	26	39	39	39	39	39	39	39	39	221	208
19	26	26	39	26	39	39	39	39	39	39	39	39	221	208
20	26	26	39	26	39	39	39	39	39	39	39	39	221	208
21	39	36	39	26	39	36	39	34	39	38	39	39	234	209
22	39	39	39	26	39	39	39	39	39	39	39	39	234	221
23	39	39	39	39	39	39	39	39	39	39	39	39	234	234
24	39	39	39	39	39	39	39	39	39	39	39	39	234	234
25	39	39	26	25	39	39	39	39	39	39	39	39	221	220

Taula 1: Nombres de jocs amb desviaments (DEV) i de jocs cooperatius (CP) per rondes i grups

5.1.4 Un model d'aprenentatge sobre el comportament dels desviaments

El nostre model d'aprenentatge conté el període k de desviament intencionat com l'estat intern de l'individu. Un individu canvia el seu estat intern de la ronda t a la ronda $t+1$ d'acord amb probabilitats de transició constants. Cada individu ve caracteritzat per tres paràmetres α, β, γ.

Si un individu observa que en l'etapa t el seu oponent s'ha desviat abans del que ell tenia pensat desviar-se, aleshores amb probabilitat α, ell canviarà el seu període de desviament intencionat de k a $k-1$. La probabilitat que l'estat intern del subjecte quedi fix a k és de $1-\alpha$.

Si un individu observa que a l'etapa t el seu oponent s'ha desviat al mateix període k que ell ho ha fet, aleshores amb probabilitat β el seu període de desviament intencionat canviarà a $k-1$ i amb probabilitat $1-\beta$ es quedarà igual.

Si un individu observa que en l'etapa t s'ha desviat abans que el seu oponent, aleshores amb probabilitat γ canviarà a $k+1$, recordant que $k < 50$. Amb probabilitat $1-\gamma$ no canviarà el seu estat intern. Tot això es veu resumit en la taula següent:

	Període intencionat de desviament a la ronda t+1		
Període intencionat de desviament a la ronda t	Un període abans	No canvia	Un període després
Després de l'oponent	α	1-α	
A la vegada	β	1-β	
Abans de l'oponent		1-ɣ si k<50 1 si k=50	ɣ si k<50 0 si k=50

Taula 2: Probabilitats de transició entre rondes del període de desviament intencionat d'un individu on k aquest període en una ronda t

Ara procedirem a discutir les motivacions darrere de les probabilitats de transició. Un individu que ha observat que el seu oponent s'ha desviat abans del que ell tenia pensat fer-ho, pensarà que hauria estat millor desviar-se abans. El mateix és cert, amb un grau més baix, si l'oponent es desvia al mateix que ell ho fa. Aleshores, és raonable assumir que $\alpha \geq \beta > 0$. En ambdós casos no hi ha motius de canviar la intenció de desviar-se a períodes posteriors.

Ara considerem un individu que en la ronda t s'ha desviat en el període $k < 50$[13] i observa que el seu oponent ha cooperat fins al període k. Ell no sap exactament en quin període el rival tenia intenció de desviar-se. Aleshores, podria haver estat millor fer-ho en un període més tard. Anem a mirar l'etapa $k = 48$. L'individu no sap si el rival pretenia desviar-se al període 49,50 o no fer-ho. Si ho fes en algun dels dos últims, desviar-se en el període $k = 9$ hauria estat millor. És possible que aquesta incertesa produeixi una

[13]Recordar que es van jugar 25 rondes, on a cada ronda, cada parella repetia el Dilema del Presoner 50 cops, sense ells saber que després de l'etapa 50 s'acabaria.

tendència a canviar el període intencionat de desviament. És clar, per $k = 50$ no hi ha tal incertesa i l'individu ha de concloure que el correcte era desviar-se en l'últim període si l'oponent coopera fins al final, però no sabien quan acabaria el joc.

L'especificació de les idees generals explicades a dalt content certes suposicions que simplifiquen el model. Excloem la possibilitat que el període intencionat de desviament canviï en més d'un període. Es podria fer un model més general, però entre la falta de dades i que els perfils han estat creats a partir de certes preguntes, ens obliga a restringir la nostra atenció a models amb el nombre mínim de paràmetres possibles.

En una situació on l'individu es desvia abans que el seu oponent, la seva incertesa sobre la naturalesa de la seva experiència és més gran com abans es desviés. Com més períodes hi hagi després del desviament, abans d'acabar el superjoc, major és la probabilitat que el desviament fos massa d'hora. Aleshores, un pot pensar de fer el paràmetre γ depenent de k de tal manera que γ creixi quan k decreixi. Aquesta seria una modificació teòricament atractiva del model, però també per això, la necessitat d'augmentar el nombre de paràmetres ens impedeix comparar amb tals models, amb les dades de les quals disposem.

5.1.5 Consideracions teòriques

A continuació mirarem les conseqüències del nostre model teòric en una situació ideal que no es correspon amb la de l'experiment.

Considerem una mostra molt gran d'individus on els paràmetres α, β i γ són iguals per tots. Amb aquesta població, imaginem un experiment fictici sobre una seqüència molt llarga de superjocs. A cada ronda els individus s'enfronten de manera aleatòria.

Podem preguntar-nos com evolucionen en aquest sistema les probabilitats dels períodes intencionats de desviament. Per tal de descriure el procés que controla l'evolució d'aquestes probabilitats, introduïm les següents notacions:

Denotarem per p_k^t a la probabilitat que a la ronda t un individu escollit aleatòriament tingui intenció de desviar-se en el període k, on $k = 51$ implica que no te intenció de desviar-se.

Denotarem per $S_k^t = \sum_{m=1}^{k} p_m^t$ a la probabilitat que en la ronda t un individu triat aleatòriament tingui la intenció de desviar-se en els períodes $1, \ldots, k$, i α, β i γ són els paràmetres de la taula anterior.

És útil mirar aquesta situació d'una manera semblant a una cadena de Markov. Ens podem fer la següent pregunta: quines són les probabilitats que un individu tingui la intenció de desviar-se en el període $k-1, k$ o $k+1$ en la ronda $t+1$ si tenia la intenció de desviar-se en el període k en la ronda t? Aquestes probabilitats de transició poden col·locar-se en una matriu, on les columnes corresponguin als períodes intencionats de desviament a la ronda t i les files als períodes intencionats de desviament a la ronda $t+1$. Una part d'aquesta matriu es pot veure en el sistema de sota.

Amb l'ajuda de la taula anterior es pot veure fàcilment que les probabilitats de transició, són, de fet, les mostrades a la taula de baix. Pel que hem comentat fins ara, és clar que les probabilitats p_k^{t+1} estan determinades pel següent sistema d'equacions:

$$
\begin{cases}
p_{51}^{t+1} = [(1-\alpha)S_{50}^t + (1-\beta)p_{51}^t]p_{51}^t \\
p_{50}^{t+1} = [\alpha S_{50}^t + \beta p_{51}^t]p_{51}^t + [(1-\alpha)S_{49}^t + (1-\beta)p_{50}^t + (1-\gamma)(1-S_{50}^t)]p_{50}^t + \gamma(1-S_{49}^t)p_{49}^t \\
. \\
. \\
. \\
p_2^{t+1} = [\alpha S_2^t + \beta p_3^t]p_3^t + [(1-\alpha)p_1^t + (1-\beta)p_2^t + (1-\gamma)(1-S_2^t)]p_2^t + \gamma(1-p_1^t)p_1^t \\
p_1^{t+1} = [\alpha p_1^t + \beta p_2^t]p_2^t + [p_1^t + (1-\gamma)(1-p_1^t)]p_1^t
\end{cases}
$$

Començant amb una distribució inicial $(p_1^1, \ldots, p_{51}^1)$ el vector de probabilitats $(p_1^t, \ldots, p_{51}^t)$ pot ser calculat a cada ronda t. Ens podríem preguntar si aquest vector de probabilitats convergeix a una distribució en equilibri estable, però no intentarem donar una resposta teòrica rigorosa a la pregunta de la convergència. Malgrat això, tenim un llarg nombre de càlculs numèrics, els resultats dels quals ens indiquen un patró fixat que definirem a continuació. Primer cal assenyalar que les equacions en diferències[14] compleixen la següent propietat: Si $p_k^t = 0$ es manté per $k = m, \ldots, 51$ per algun $t = t_0$, aleshores les mateixes condicions se satisfan per tot $t > t_0$. Només per aquest motiu, el resultat de la simulació no pot ser completament independent de les condicions inicials.

Ara bé, si p_{50}^1 i p_{51}^1, són suficientment grans, els resultats dels nostres càlculs no dependran en les condicions inicials exactes. Els resultats obtinguts de $p_{51}^1 = 1$ no canvien mentre que les condicions inicials es mantinguin en un entorn d'aquest cas extrem. La mida d'aquest entorn depèn dels paràmetres, però per la majoria de casos sembla ser prou gran.

La taula 3 mostra els resultats numèrics per diferents combinacions de paràmetres. Tots aquests càlculs s'han fet agafant com a condició inicial $p_{51}^1 = 1$. Els resultats de l'experiment suggereixen que els individus aprenen a cooperar abans d'aprendre a desviar-se. És per això que assumir $p_{51}^1 = 1$ és prou raonable.

Tots els càlculs fets amb $p_{51}^1 = 1$ convergeixen a una distribució estacionària que quasi sempre estava majoritàriament concentrada o bé al final o bé al començament del superjoc. A la taula 3, els tres primers períodes o els tres últims períodes obtenien almenys un 97 per cent de la massa total de probabilitat.

Les combinacions de paràmetres estan organitzades en grups amb β i γ constants i α creixent. Si α és petita en comparació a $\gamma - \beta$, la distribució estarà majoritàriament concentrada al final del superjoc. Amb α creixent, aquesta concentració esdevé menys pronunciada fins que un valor crític d'α s'assoleix i a partir del qual, la distribució estacionària es concentra majoritàriament al començament del superjoc. Com es pot veure a la taula 3, el valor crític d'α es troba una mica per sota de $\gamma - \beta$. Es pot comprovar analíticament que per $\alpha + \beta = \gamma$, la distribució $p_1 = p_2 = 0.5$ és estacionària. De fet, en casos on $\alpha + \beta = \gamma$ el procés convergeix a aquesta distribució.

[14] En matemàtiques, es defineix una equació de diferències lineal o relació de recurrència lineal com una successió a_n definida en funció dels elements anteriors d'aquesta mateixa successió per tot n no negatiu i els seus termes únicament poden tenir grau 0 i 1 per complir que sigui lineal.

Els resultats d'aquests càlculs suggereixen un canvi abrupte de la distribució estacionària al valor crític d'α. A la taula 3, els valors crítics d'α estan continguts en intervals de longitud 10^{-3}. Es pot observar que dins d'aquest interval petit, la distribució estacionària assolida pel procés canvia dràsticament. El canvi és en certa manera menys pronunciat si reduïm l'interval a una longitud de 10^{-7}, però encara així les probabilitats de les etapes al mig de cada ronda són pràcticament zero.

A l'experiment, a cada grup teníem 78 individus i els valors dels paràmetres variaven considerablement. El model aplicat a la situació experimental es pot considerar com una cadena de Markov amb un estat d'espai definit adequadament.

El període intencionat de desviament més alt entre tots els individus no pot augmentar d'un superjoc al següent, però si α i β són positius i γ és més petit que 1 per tots els individus, aleshores sempre hi ha una probabilitat positiva que el període de desviament intencionat més alt disminuirà en 1. Aleshores, amb probabilitat 1, el període de desviament més alt acabarà decreixent a 1 en una seqüència infinita de superjocs. Això vol dir que a la llarga el comportament convergeix completament a la no cooperació.

Inclús si la possibilitat de convergir a una distribució estacionària queda exclosa del model, si s'aplica a una població finita d'individus, els càlculs per a la situació amb una població infinita no està exempta d'interès. Suggereixen que en la situació finita, la convergència a la no cooperació pot ser molt lenta si els paràmetres α i β són relativament petits i γ relativament gran. A més a més, en els càlculs del cas infinit, un ha de considerar la possibilitat que la conclusió sobre la convergència a la no cooperació no és robusta a causa de petites especificacions errònies en el model d'aprenentatge. Suposem que les probabilitats per les transicions excloses no són realment zero, però són molt petites. És raonable esperar que sota aquesta condició, una distribució estacionària mostrant un punt de desviament estable, podria ser obtingut amb la combinació de paràmetres adients en el cas finit.

No obstant, es pot esperar que els resultats obtinguts pel cas del grup gran amb els mateixos paràmetres són indicatius del que podem preveure que passaria en la situació experimental si el model és el correcte.

α	β	γ	p_1	p_2	p_3	...	p_{48}	p_{49}	p_{50}
0.100	0.1	0.1	1.000	—	—		—	—	—
0.100	0.1	0.4	—	—	—		0.017	0.250	0.733
0.200	0.1	0.4	—	—	—		0.038	0.346	0.615
0.264	0.1	0.4	—	—	—		0.135	0.520	0.340
0.265	0.1	0.4	0.119	0.509	0.371		—	—	—
0.300	0.1	0.4	0.500	0.500	—		—	—	—
0.300	0.1	0.5	—	—	—		0.034	0.356	0.610
0.355	0.1	0.5	—	—	—		0.105	0.520	0.373
0.356	0.1	0.5	0.098	0.512	0.390		—	—	—
0.400	0.1	0.5	0.500	0.500	—		—	—	—
0.500	0.1	0.5	0.990	0.010	—		—	—	—
0.200	0.2	0.5	—	—	—		0.086	0.400	0.511
0.300	0.2	0.5	0.500	0.500	—		—	—	—
0.400	0.2	0.5	0.667	0.333	—		—	—	—
0.100	0.1	0.6	—	—	—		0.005	0.167	0.829
0.200	0.1	0.6	—	—	—		0.007	0.201	0.791
0.300	0.1	0.6	—	—	—		0.013	0.256	0.731
0.400	0.1	0.6	—	—	—		0.030	0.364	0.606
0.449	0.1	0.6	—	—	—		0.095	0.533	0.370
0.450	0.1	0.6	0.087	0.522	0.391		—	—	—
0.200	0.2	0.6	—	—	—		0.044	0.333	0.622
0.300	0.2	0.6	—	—	—		0.083	0.417	0.498
0.353	0.2	0.6	—	—	—		0.188	0.516	0.282
0.354	0.2	0.6	0.151	0.500	0.349		—	—	—
0.400	0.2	0.6	0.500	0.500	—		—	—	—
0.500	0.2	0.6	0.667	0.333	—		—	—	—
0.600	0.2	0.6	0.993	0.007	—		—	—	—
0.100	0.1	0.7	—	—	—		0.003	0.143	0.854
0.200	0.1	0.7	—	—	—		0.004	0.167	0.828
0.300	0.1	0.7	—	—	—		0.006	0.203	0.791
0.400	0.1	0.7	—	—	—		0.011	0.259	0.730
0.500	0.1	0.7	—	—	—		0.028	0.370	0.602
0.543	0.1	0.7	—	—	—		0.073	0.517	0.409
0.544	0.1	0.7	0.072	0.515	0.413		—	—	—
0.600	0.1	0.7	0.500	0.500	—		—	—	—
0.700	0.1	0.7	0.990	0.010	—		—	—	—
0.440	0.2	0.7	—	—	—		0.164	0.525	0.302
0.441	0.2	0.7	0.133	0.506	0.361		—	—	—
0.529	0.2	0.8	—	—	—		0.144	0.530	0.320
0.530	0.2	0.8	0.119	0.509	0.371		—	—	—
0.445	0.3	0.8	—	—	—		0.213	0.510	0.257
0.446	0.3	0.8	0.165	0.495	0.340		—	—	—
0.538	0.4	0.9	—	—	—		0.227	0.504	0.244
0.539	0.4	0.9	0.173	0.492	0.335		—	—	—
0.100	0.1	1.0	—	—	—		—	0.100	0.900

Taula 3: Distribucions de probabilitats estables sobre els períodes intencionats de desviament amb diferents combinacions de paràmetres

5.2 Simulació del Dilema del Presoner computacional

En aquest apartat contrastarem els resultats teòrics del dilema del presoner obtinguts en el capítol 4, amb els resultats numèrics d'un programa en llenguatge C[15], que simularà els diferents encreuaments d'estratègies entre els dos jugadors, fixat un factor de descompte δ. Per tal d'optimitzar el programa, els sumatoris fins a infinit dels quals coneixem el resultat en funció de δ apareixeran com a fórmula, mentre que els altres, assignarà la utilitat corresponent a cada jugador i ho multiplicarà per la potència del factor de descompte $t-1$ corresponent a l'etapa t. Aquest últim s'aturarà quan δ^{t-1} sigui inferior a una tolerància arbitrària de 10^{-8}.

Cada jugador disposa de 6 estratègies diferents:

- Confessar sempre
- Callar sempre
- Random: el programa tria aleatòriament a cada etapa el nombre 0 o 1, i assigna el 0 a confessar en aquella etapa o l'1 a callar.
- Igualitari: Callar a la primera ronda. A la resta de rondes examina la història de l'altre jugador i compta el total de confessions i silencis. Si les confessions són més que els silencis, el jugador confessarà i si no farà silenci.
- Ull per ull: Callar a la primera ronda. A la resta de rondes imita el comportament de l'altre jugador en la ronda prèvia.
- Tirador: Cooperar totes les rondes mentre el rival també cooperi. En el torn posterior en què el rival no cooperi, canvia a confessar i no canvia mai més.

Taula 2: Combinacions d'estratègies, fixat $\delta = 0.999$.

J1/J2	Confessar	Callar	Random
Confessar	$(1000, 1000)$	$(5000, 0)$	$(2977.61, 505.60)$
Callar	$(0, 5000)$	$(4000, 4000)$	$(1977.61, 4505.60)$
Random	$(505.60, 2977.61)$	$(4505.60, 1977.61)$	$(2463.27, 2500.75)$
Igualitari	$(999, 1004)$	$(4000, 4000)$	$(2627.65, 1905.43)$
Ull per Ull	$(999, 1004)$	$(4000, 4000)$	$(2482.70, 2485.23)$
Tirador	$(999, 1004)$	$(4000, 4000)$	$(2975.61, 513.6)$

J1/J2	Igualitari	Ull per Ull	Tirador
Confessar	$(1004, 999)$	$(1004, 999)$	$(1004, 999)$
Callar	$(4000, 4000)$	$(4000, 4000)$	$(4000, 4000)$
Random	$(1905.43, 2627.65)$	$(2485.23, 2482.70)$	$(1486.14, 2732.47)$
Igualitari	$(4000, 4000)$	$(4000, 4000)$	$(4000, 4000)$
Ull per Ull	$(4000, 4000)$	$(4000, 4000)$	$(4000, 4000)$
Tirador	$(4000, 4000)$	$(4000, 4000)$	$(4000, 4000)$

[15] Consultar l'annex al final del document.

Podem observar que tenim 10 EN diferents. Ara bé, només 2 són ENPS, que són els que havíem vist anteriorment. Els 8 EN que rebutgem com ENPS tenen una cosa en comú, almenys un dels dos jugadors tria una estratègia on hi ha una dependència infinita respecte a les decisions del rival, és a dir, on la tria de decisió de la pròxima etapa, és causada per les decisions del rival en l'etapa actual en el cas d'Ull per ull o en totes les etapes fins al moment, en el cas d'Igualitari.

Ara anem a veure què passa en el cas de tenir una $\delta < 1/4$.

Taula 3: Combinacions d'estratègies, fixat $\delta = 0.2$.

J1/J2	Confessar	Callar	Random
Confessar	$(1.25, 1.25)$	$(6.25, 0)$	$(5.45, 0.2)$
Callar	$(0, 6.25)$	$(5, 5)$	$(4.2, 5.2)$
Random	$(0.2, 5.45)$	$(5.2, 4.2)$	$(0.98, 5.98)$
Igualitari	$(0.25, 5.25)$	$(5, 5)$	$(4.2, 5.2)$
Ull per Ull	$(0.25, 5.25)$	$(5, 5)$	$(4.24, 5.04)$
Tirador	$(0.25, 5.25)$	$(5, 5)$	$(4.25, 5)$

J1/J2	Igualitari	Ull per Ull	Tirador
Confessar	$(5.25, 0.25)$	$(5.25, 0.25)$	$(5.25, 0.25)$
Callar	$(5, 5)$	$(5, 5)$	$(5, 5)$
Random	$(5.2, 4.2)$	$(5.04, 4.24)$	$(4.24, 4.44)$
Igualitari	$(5, 5)$	$(5, 5)$	$(5, 5)$
Ull per Ull	$(5, 5)$	$(5, 5)$	$(5, 5)$
Tirador	$(5, 5)$	$(5, 5)$	$(5, 5)$

Ara, només tenim un ENPS, que és el perfil d'estratègies incrèdules. Fixem-nos que per valors de δ molt petits, implicaria que la probabilitat que el joc es repetís fos molt baixa, és a dir, el nostre joc seria el dilema del presoner amb una única etapa, on ja sabem que l'únic EN és on ambdós jugadors trien confessar.

6 Conclusions

Diversos cops al llarg d'aquests quatre anys he sentit la paraula joc, i no ha estat fins a l'últim que he pogut descobrir la relació entre aquest concepte i les matemàtiques, en cursar l'assignatura de teoria de jocs, que no és més que una breu introducció a la infinitat de situacions amb què ens podem creuar, des de les més comunes, del dia a dia, fins a les més vitals com pot ser quants anys entrar a la presó.

Si bé és cert que no era el primer cop que veia el dilema del presoner, sí que ha estat el primer on l'he intentat enfocar des d'un punt de vista matemàtic i no només filosòfic. És curiós, com un problema tan fonamental, mostra com dues persones poden arribar a no cooperar inclús si va en contra dels seus interessos. Hem pogut observar al llarg del treball les diferents solucions d'aquest joc i en el cas de la seva repetició infinita, com variaven aquests equilibris en funció del valor que cada jugador li donava al present. També hem pogut veure els resultats d'un altre problema clàssic de la teoria de jocs amb un vessant més econòmic com era l'oligopoli de Cournot, on l'equilibri del monopoli és factible un cop desapareix la restricció temporal.

Finalment, s'ha intentat crear un model d'aprenentatge per intentar predir el comportament dels individus davant una situació hipotètica del dilema del presoner, on s'ha pogut notar que no tots seguien la línia esperada. Això en part es pot deure al fet que la mostra de l'experiment no era suficientment gran, però per altra banda a què la racionalitat humana i la capacitat de pensament a un horitzó infinit és limitada.

Referències

[1] Sánchez - Cuenca, I.: *Teoría de Juegos*, 2a edició ampliada i revisada, Centre d'Investigacions Sociològiques, 2009.

[2] Osborne, M. J.: *An Introduction to game theory*, International edition, Oxford University Press, 2009.

[3] Binmore, K.: *La Teoría de juegos: una breve introducción*, 2a edició, Alianza, 2011.

[4] Pérez, J.; Jimeno, J. L.; Cerdá, E.: *Teoría de juegos*, 2a edició, Garceta 2013.

[5] Gibbons, R.: *Un primer curso de teoría de juegos*, Bosch, 1993.

[6] Karlin, S.: *Mathematical Methods and Theory in Games, Programming and Economics*, Volume II, The Theory of Infinite Games, Stanford University, 1959.

[7] Axelrod, R.: *The Evolution of Cooperation*, Estats Units, 1984.

[8] Hurtado. C: *Infinitely Repeated Games*, Department of Economics, University of Illinois, 2015.

[9] Ozdaglar, A.: *Game Theory with Engineering Applications: Repeated Games*, MIT, 2010.

[10] Ozerturk, S.: *Infinitely Repeated Games*, Southern Methodist University, 2014.

[11] González, J.: *A generalized Nash Folk Theorem*, Department of Statistics and Operations Research, Faculty of Mathematics, Universidad de Santiago de Compostela, 2015.

[12] Abréu, D.; Dutta, P. K.; Smith, L.: *The Folk Theorem for Repeated Games: A NEU Condition*, Econometrica, Vol. 62, pàg 939 - 948, 1994.

[13] Benoît, J. P.; Krishna, V.: *Finitely Repeated Games*, Econometrica, Vol. 53, pàg 905 - 922, 1985.

[14] Fudenberg, D.; Maskin, E.: *The Folk Theorem in Repeated Games with Discounting*, Econometrica, Vol. 54, pàg 533 - 554, 1986.

[15] Gossner, O.: *The Folk Theorem for Finitely Repeated Games with Mixed Strategies*, International Journal of Game Theory, vVol. 24, pàg 95 - 107, 1995.

[16] Smith, L.: *Necessary and Sufficient Conditions for the Perfect Finite Horizon Folk Theorem*, Econometrica, Vol. 63, pàg 425 - 430, 1995.

[17] Wen, Q.: *The Folk Theorem for Repeated Games with Complete Information*, Econometrica, Vol. 62, pàg 949 - 954, 1994.

[18] Gossner, O.; Tomala, T.: *Repeated Games*, London School of Economics, 2007.

[19] Huang, H.: *Introduction to Repeated Games*, University of California, 2011.

[20] Dal Bó, P.; Fréchette, G. R.: *The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence*, Brown University & New York University, 2009.

[21] Fudenberg, D.; Tirole, J.: *Game Theory*, MIT, 1991.

[22] Aumann, R. J.; Shapley, L. S.: *Long-Term Competition - a Game-Theoretic Analysis*, Hebrew University of Jerusalem & University of California at Los Angeles, 2013.

[23] Selten, R.; Stoecker, R.: *End Behaviour in Supergames*, Rheinische Friedrich-Wilhelms-Universität Bonn, 1985.

[24] Wooldridge, M.: *Computation and the Prisoner's Dilemma*, University of Liverpool, 2012.

[25] Aumann, R. J.: *Essays in game theory and mathematical economics in honor of Oskar Morgenstern*, pàg 11 - 42, Wissenschaftsverlag, Mannheim, Wien & Zurich, 1981.

[26] Mertens, J. F.; Sorin, S.; Zamir, S.: *Repeated Games*, CORE discussion paper 9420 - 9422, 1994.

[27] Mertens, J. F.: *Repeated Games. In proceedings of the international congress of Mathematics*, pàg 1528 - 1577, Berkeley, California, 1986.

[28] Mailath, G. J.; Samuelson, L.: *Repeated Games and Reputations: Long - Run Relationships*, Oxford University Press, 2006.

Annex

En aquest annex es troba el codi *presoner.c* utilitzat al treball Jocs Repetits Infinitament a l'apartat de simulació del dilema del presoner computacional.

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>

int recompensesA(char,char);
int recompensesB(char,char);

int main(void){

int etapa=0,num,callaA=0,confA=0,callaB=0,confB=0;
double delta,acumulat=1,utilA=0,utilB=0,sumA,sumB,tol=1e-08;
char presonerA,presonerB,accioA,accioB;

printf("Quin és el factor de descompte d'ambdós jugadors?\n");
scanf("%lf",&delta);

/*Igualitari  -> Calla a la primera ronda. A la resta de rondes examina
la història de l'altre jugador i compta el total de confessions i silencis.
Si les confessions són més que els silencis, el jugador confessarà i sinó
farà silenci.*/

/*Ull per ull -> Calla a la primera ronda. A la resta de rondes imita el
comportament de l'altre jugador en la ronda prèvia.*/

/*Tirador -> Coopera totes les rondes mentre el rival també cooperi.
En el torn posterior a que el rival no cooperi, canvia a confessar i
no canvia mai més.*/

printf("Estratègies: Confessar sempre (C), Silenci sempre (S), Random (R),
Igualitari (I), Ull per Ull (U), Tirador (T)\n");

printf("Quina estratègia seguirà el presoner A?\n");
scanf(" %c",&presonerA);

printf("Quina estratègia seguirà el presoner B?\n");
scanf(" %c",&presonerB);

/*A = Acusar (Confessar), N = No acusar (Callar)*/


/*El presoner A jugarà Confessar sempre*/

if(presonerA == 'C'){
```

```
accioA = 'A';

if(presonerB == 'C'){
utilA = 1/(1-delta);
utilB = utilA;
}

if(presonerB == 'S'){
utilA = 5/(1-delta);
utilB = 0;
}

if(presonerB == 'R'){
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioB = 'A';
confB++;
}else{
accioB = 'N';
callaB++;
}

printf("Acció del jugador B a l'etapa %d: %c\n",etapa,accioB);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
utilB = utilB + acumulat*sumB;
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador B ha confessat %d cops i ha callat %d cops\n",confB,callaB);

}

if(presonerB == 'I'){
utilA = 5 + (delta)/(1-delta) ;
utilB = (1*delta)/(1-delta);
}

if(presonerB == 'U'){
utilA = 5 + delta/(1-delta);
utilB = delta/(1-delta);
}

if(presonerB == 'T'){
utilA = 5 + delta/(1-delta);
utilB = delta/(1-delta);
```

```
}

printf("La utilitat del jugador A és de %12.6lf\n",utilA);
printf("La utilitat del jugador B és de %12.6lf\n",utilB);

}

/*El presoner A jugarà Callar sempre*/

if(presonerA == 'S'){
accioA = 'N';

if(presonerB == 'C'){
utilA = 0;
utilB = 5/(1-delta);
}

if(presonerB == 'S'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'R'){
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioB = 'A';
confB++;
}else{
accioB = 'N';
callaB++;
}

printf("Acció del jugador B a l'etapa %d: %c\n",etapa,accioB);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
utilB = utilB + acumulat*sumB;
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador B ha confessat %d cops i ha callat %d cops\n",confB,callaB);
}

if(presonerB == 'I'){
utilA = 4/(1-delta);
utilB = utilA;
}
```

```
if(presonerB == 'U'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'T'){
utilA = 4/(1-delta);
utilB = utilA;
}

printf("La utilitat del jugador A és de %12.6lf\n",utilA);
printf("La utilitat del jugador B és de %12.6lf\n",utilB);

}

/*El presoner A jugarà aleatòriament*/

if(presonerA == 'R'){

if(presonerB == 'C'){
accioB = 'A';
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioA = 'A';
confA++;
}else{
accioA = 'N';
callaA++;
}

printf("Acció del jugador A a l'etapa %d: %c\n",etapa,accioA);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
utilB = utilB + acumulat*sumB;
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador A ha confessat %d cops i ha callat %d cops\n",confA,callaA);
}

if(presonerB == 'S'){
accioB = 'N';
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
```

```
accioA = 'A';
confA++;
}else{
accioA = 'N';
callaA++;
}

printf("Acció del jugador A a l'etapa %d: %c\n",etapa,accioA);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
utilB = utilB + acumulat*sumB;
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador A ha confessat %d cops i ha callat %d cops\n",confA,callaA);
}

if(presonerB == 'R'){
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioA = 'A';
confA++;
}else{
accioA = 'N';
callaA++;
}
num = rand() % 2;
if(num == 0){
accioB = 'A';
confB++;
}else{
accioB = 'N';
callaB++;
}

printf("Acció del jugador A a l'etapa %d: %c\n",etapa,accioA);
printf("Acció del jugador B a l'etapa %d: %c\n",etapa,accioB);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
utilB = utilB + acumulat*sumB;
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador A ha confessat %d cops i ha callat %d cops\n",confA,callaA);
```

```
printf("El jugador B ha confessat %d cops i ha callat %d cops\n",confB,callaB);
}

if(presonerB == 'I'){
accioB = 'N';
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioA = 'A';
confA++;
}else{
accioA = 'N';
callaA++;
}

printf("Acció del jugador A a l'etapa %d: %c\n",etapa,accioA);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
utilB = utilB + acumulat*sumB;
if(confA > callaA){
accioB = 'A';
}else{
accioB = 'N';
}

acumulat = acumulat*delta;
etapa++;
}
printf("El jugador A ha confessat %d cops i ha callat %d cops\n",confA,callaA);
}

if(presonerB == 'U'){
accioB = 'N';
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioA = 'A';
confA++;
}else{
accioA = 'N';
callaA++;
}

printf("Acció del jugador A a l'etapa %d: %c\n",etapa,accioA);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
```

```
sumB = recompensesB(accioA,accioB);
utilB = utilB + acumulat*sumB;
accioB = accioA;
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador A ha confessat %d cops i ha callat %d cops\n",confA,callaA);
}

if(presonerB == 'T'){
accioB = 'N';
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioA = 'A';
confA++;
}else{
accioA = 'N';
callaA++;
}

if(accioB != 'A' && accioA == 'A'){
accioB = 'A';
}

printf("Acció del jugador A a l'etapa %d: %c\n",etapa,accioA);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
utilB = utilB + acumulat*sumB;
accioB = accioA;
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador A ha confessat %d cops i ha callat %d cops\n",confA,callaA);
}

printf("La utilitat del jugador A és de %12.6lf\n",utilA);
printf("La utilitat del jugador B és de %12.6lf\n",utilB);

}

/*El presoner A jugarà de manera igualitària*/

if(presonerA == 'I'){
accioA = 'N';

if(presonerB == 'C'){
```

```
utilA = delta/(1-delta);
utilB = 5 + delta/(1-delta);
}

if(presonerB == 'S'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'R'){
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioB = 'A';
confB++;
}else{
accioB = 'N';
callaB++;
}

printf("Acció del jugador B a l'etapa %d: %c\n",etapa,accioB);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
  utilB = utilB + acumulat*sumB;
if(confB > callaB){
accioA = 'A';
}else{
accioA = 'N';
}
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador B ha confessat %d cops i ha callat %d cops\n",confB,callaB);
}

if(presonerB == 'I'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'U'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'T'){
utilA = 4/(1-delta);
```

```
utilB = utilA;
}

printf("La utilitat del jugador A és de %12.6lf\n",utilA);
printf("La utilitat del jugador B és de %12.6lf\n",utilB);

}

/*El presoner A jugarà Ull per Ull*/

if(presonerA == 'U'){
accioA = 'N';

if(presonerB == 'C'){
utilA = delta/(1-delta);
utilB = 5 + delta/(1-delta);
}

if(presonerB == 'S'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'R'){
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioB = 'A';
confB++;
}else{
accioB = 'N';
callaB++;
}

printf("Acció del jugador B a l'etapa %d: %c\n",etapa,accioB);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
utilB = utilB + acumulat*sumB;
accioA = accioB;
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador B ha confessat %d cops i ha callat %d cops\n",confB,callaB);
}

if(presonerB == 'I'){
utilA = 4/(1-delta);
```

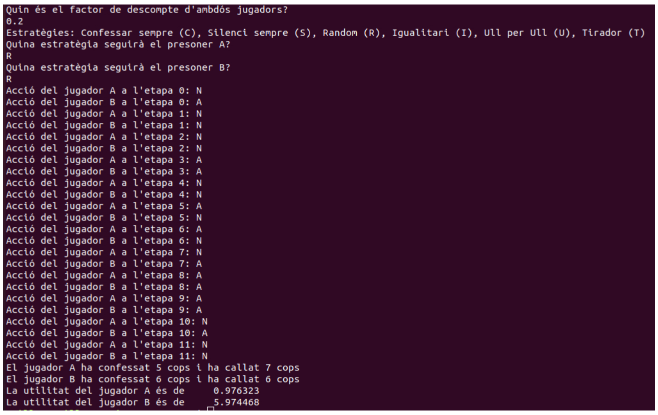
```
utilB = utilA;
}

if(presonerB == 'U'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'T'){
utilA = 4/(1-delta);
utilB = utilA;
}

printf("La utilitat del jugador A és de %12.6lf\n",utilA);
printf("La utilitat del jugador B és de %12.6lf\n",utilB);

}

/*El presoner A jugarà Tirador*/

if(presonerA == 'T'){
accioA = 'N';

if(presonerB == 'C'){
utilA = delta/(1-delta);
utilB = 5 + delta/(1-delta);
}

if(presonerB == 'S'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'R'){
while(acumulat > tol){
num = rand() % 2;
if(num == 0){
accioB = 'A';
confB++;
}else{
accioB = 'N';
callaB++;
}

printf("Acció del jugador B a l'etapa %d: %c\n",etapa,accioB);

sumA = recompensesA(accioA,accioB);
utilA = utilA + acumulat*sumA;
sumB = recompensesB(accioA,accioB);
```

```
utilB = utilB + acumulat*sumB;
if(accioB == 'A'){
accioA = 'A';
}
acumulat = acumulat*delta;
etapa++;
}
printf("El jugador B ha confessat %d cops i ha callat %d cops\n",confB,callaB);
}

if(presonerB == 'I'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'U'){
utilA = 4/(1-delta);
utilB = utilA;
}

if(presonerB == 'T'){
utilA = 4/(1-delta);
utilB = utilA;
}

printf("La utilitat del jugador A és de %12.6lf\n",utilA);
printf("La utilitat del jugador B és de %12.6lf\n",utilB);

}

return 0;

}

int recompensesA(char accioA,char accioB){

if(accioA == 'A' && accioB == 'A'){
return 1;
}else if(accioA == 'A' && accioB == 'N'){
return 5;

}else if(accioA == 'N' && accioB == 'A'){
return 0;

}else if(accioA == 'N' && accioB == 'N'){
return  4;
}
exit(1);
}
```

```
int recompensesB(char accioA,char accioB){

if(accioA == 'A' && accioB == 'A'){
return 1;

}else if(accioA == 'A' && accioB == 'N'){
return 0;

}else if(accioA == 'N' && accioB == 'A'){
return 5;

}else if(accioA == 'N' && accioB == 'N'){
return 4;
}

exit(2);
}
```

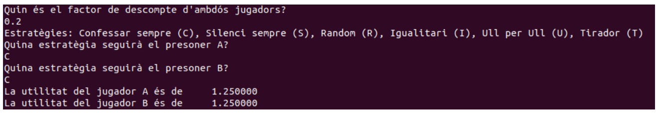
A continuació, es mostren algunes captures de pantalla del terminal per poder veure d'on s'han extret les dades utilitzades a les taules 6x6 amb les diferents combinacions d'estratègies. Recordar que el programa s'aturarà quan $\delta^x \leq 10^{-8}$, una tolerància fixada arbitràriament. Dos exemples amb $\delta = 0.2$:

Figura 6: Dilema del presoner amb $\delta = 0.2$ on els dos jugadors trien l'estratègia Random.



```
Quin és el factor de descompte d'ambdós jugadors?
0.2
Estratègies: Confessar sempre (C), Silenci sempre (S), Random (R), Igualitari (I), Ull per Ull (U), Tirador (T)
Quina estratègia seguirà el presoner A?
R
Quina estratègia seguirà el presoner B?
R
Acció del jugador A a l'etapa 0: N
Acció del jugador B a l'etapa 0: A
Acció del jugador A a l'etapa 1: N
Acció del jugador B a l'etapa 1: N
Acció del jugador A a l'etapa 2: N
Acció del jugador B a l'etapa 2: N
Acció del jugador A a l'etapa 3: A
Acció del jugador B a l'etapa 3: A
Acció del jugador A a l'etapa 4: N
Acció del jugador B a l'etapa 4: N
Acció del jugador A a l'etapa 5: A
Acció del jugador B a l'etapa 5: N
Acció del jugador A a l'etapa 6: A
Acció del jugador B a l'etapa 6: N
Acció del jugador A a l'etapa 7: N
Acció del jugador B a l'etapa 7: A
Acció del jugador A a l'etapa 8: A
Acció del jugador B a l'etapa 8: A
Acció del jugador A a l'etapa 9: A
Acció del jugador B a l'etapa 9: A
Acció del jugador A a l'etapa 10: N
Acció del jugador B a l'etapa 10: A
Acció del jugador A a l'etapa 11: N
Acció del jugador B a l'etapa 11: N
El jugador A ha confessat 5 cops i ha callat 7 cops
El jugador B ha confessat 6 cops i ha callat 6 cops
La utilitat del jugador A és de     0.976323
La utilitat del jugador B és de     5.974468
```

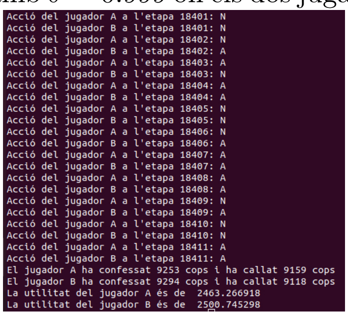
Figura 7: Dilema del presoner amb $\delta = 0.2$ on els dos jugadors trien l'estratègia Confessar Sempre.



```
Quin és el factor de descompte d'ambdós jugadors?
0.2
Estratègies: Confessar sempre (C), Silenci sempre (S), Random (R), Igualitari (I), Ull per Ull (U), Tirador (T)
Quina estratègia seguirà el presoner A?
C
Quina estratègia seguirà el presoner B?
C
La utilitat del jugador A és de     1.250000
La utilitat del jugador B és de     1.250000
```

Un exemple més amb $\delta = 0.999$:

Figura 8: Dilema del presoner amb $\delta = 0.999$ on els dos jugadors trien l'estratègia Random.



```
Acció del jugador A a l'etapa 18401: N
Acció del jugador B a l'etapa 18401: N
Acció del jugador A a l'etapa 18402: N
Acció del jugador B a l'etapa 18402: A
Acció del jugador A a l'etapa 18403: A
Acció del jugador B a l'etapa 18403: N
Acció del jugador A a l'etapa 18404: A
Acció del jugador B a l'etapa 18404: A
Acció del jugador A a l'etapa 18405: N
Acció del jugador B a l'etapa 18405: N
Acció del jugador A a l'etapa 18406: N
Acció del jugador B a l'etapa 18406: A
Acció del jugador A a l'etapa 18407: A
Acció del jugador B a l'etapa 18407: A
Acció del jugador A a l'etapa 18408: A
Acció del jugador B a l'etapa 18408: A
Acció del jugador A a l'etapa 18409: N
Acció del jugador B a l'etapa 18409: A
Acció del jugador A a l'etapa 18410: N
Acció del jugador B a l'etapa 18410: N
Acció del jugador A a l'etapa 18411: A
Acció del jugador B a l'etapa 18411: A
El jugador A ha confessat 9253 cops i ha callat 9159 cops
El jugador B ha confessat 9294 cops i ha callat 9118 cops
La utilitat del jugador A és de  2463.266918
La utilitat del jugador B és de  2500.745298
```