Jocs repetits

PID_00268967

Ignacio Sánchez Cuenca

Temps mínim de dedicació recomanat: 2 hores

Universitat
Oberta
de Catalunya

Ignacio Sánchez Cuenca

L'encàrrec i la creació d'aquest recurs d'aprenentatge UOC han estat coordinats pel professor: Albert Batlle (2019)

Segona edició: setembre 2019

Av. Tibidabo, 39-43, 08035 Barcelona
Realització editorial: FUOC

Índex

Introducció

És molt habitual que una situació estratègica, tant si es representa com un joc en forma normal com en forma extensiva, es repeteixi al llarg del temps. Quan és el cas, els models estàtics no serveixen, són una simplificació abusiva de la realitat. La repetició del joc pot introduir canvis en els equilibris respecte al joc jugat una sola vegada.

En aquest mòdul s'analitzen els canvis que es produeixen quan un joc es repeteix al llarg del temps. Es veurà de seguida que el que és veritablement important és que el joc es repeteixi indefinidament, és a dir, que els jugadors no sàpiguen quan acabaran d'interactuar. Si el joc es repeteix un nombre de vegades determinat, no canvia res respecte al joc jugat una sola vegada.

Quan el joc es repeteix indefinidament, hi ha la possibilitat que els jugadors estableixin tàcitament o explícitament relacions de col·laboració, en el sentit que cada un condicioni l'estratègia que juga al que ha fet fins llavors el rival. L'aparició d'estratègies condicionals augmenta considerablement els equilibris del joc. De fet, hi ha un teorema, conegut com el «teorema popular», que estableix que sota certes circumstàncies els jocs repetits indefinidament tenen equilibris infinits.

És especialment interessant aquesta anàlisi aplicada al dilema del presoner. Com vam veure en el mòdul 2, aquest joc resulta molt desfavorable per a la cooperació i, tanmateix, hi ha moltes situacions socials que es corresponen amb aquest joc i en què els agents, en contra del que sembla que la teoria estableix, cooperen. Una explicació raonable de per què cooperen és simplement que el joc es repeteix de manera indefinida.

L'aprenentatge del càlcul d'equilibris en joc repetits serveix també per a abordar els models de negociació, en què dos o més agents s'han de posar d'acord, amb ofertes i contraofertes successives, en com s'han de repartir un bé d'un cert valor. En la mesura que l'estructura de la negociació es repeteix al llarg del temps fins que s'assoleix un acord, es pot enfocar el problema com un joc repetit. Els models de negociació són importants perquè es poden aplicar en moltes situacions polítiques.

Objectius

En aquest mòdul s'ofereixen ensenyaments substantius, però no s'avança gaire en el maneig d'instruments d'anàlisi o de càlcul. En aquest sentit, l'única novetat veritablement important és el càlcul de la utilitat esperada al llarg del temps quan s'aplica un factor de descompte temporal. Tanmateix, es continua parlant de nocions d'equilibri ja familiars, com l'equilibri de Nash i l'equilibri de perfecció en el subjoc.

Els principals objectius de l'aprenentatge substantiu són els següents:

1. Distingir entre jocs repetits n vegades i indefinidament.

2. Entendre els efectes de la repetició sobre el dilema del presoner.

3. Conèixer els continguts bàsics del teorema popular.

4. Conèixer els models de negociació.

1. La naturalesa dels jocs repetits

Fins ara s'han estudiat jocs estàtics i jocs dinàmics. Els primers corresponen als jocs en forma normal, en què les decisions es prenen simultàniament o bé es prenen sense coneixement del que han escollit els altres, mentre que els segons corresponen als jocs en forma extensiva. Els jocs en forma extensiva són dinàmics perquè hi ha una seqüència o ordre de moviments i, per tant, podem descriure la «història» dels moviments dels jugadors.

Els jocs repetits consisteixen en el fet que una estructura d'interacció estratègica es repeteixi al llarg del temps. Dit d'una altra manera: un joc repetit és el que es juga més d'una vegada. El joc que es repeteix pot ser en forma normal o extensiva.

En llenguatge tècnic, el joc que es repeteix se sol anomenar el ***joc de referència*** (*stade game*). Per exemple, podem considerar què es dóna si el dilema del presoner (DP) es repeteix al llarg del temps, és a dir, si els actors en cada període de temps han de jugar un DP (Quadre 16). El joc de referència en cada fase o en cada etapa és el mateix, però analitzat globalment, des de la perspectiva del temps, l'equilibri o els equilibris del joc repetit no han de coincidir necessàriament amb l'equilibri o els equilibris del joc de referència considerats en un únic moment.

Quadre 16. El dilema del presoner

		*J***2**	
		C	*D*
*J***1**	*C*	2, 2	0, 3
	D	3, 0	1, 1

Quan un joc es repeteix, poden sorgir **estratègies condicionals**. En una estratègia condicional, un jugador tria una estratègia o una altra segons el que el seu rival hagi fet fins llavors. El jugador condiciona la seva estratègia al que faci l'altre jugador. EL fet que hi hagi estratègies condicionals és el que produeix l'aparició de nous equilibris respecte al joc de referència. En el joc repetit hi *pot* haver equilibris diferents dels equilibris del joc de referència perquè es poden triar estratègies condicionals, cosa que és impossible en el joc de referència jugat una sola vegada.

Situació d'interacció continuada

Hi ha una multitud de situacions en què els actors interactuen de manera continuada al llarg del temps: dos països que any rere any decideixen els seus graus d'obertura comercial (i que apugen o abaixen els aranzels), la interacció continuada entre un estat i una organització terrorista, la relació que s'estableix al llarg del temps entre un partit polític i els seguidors o votants, etc. Vegem aquest últim cas.

S'ha dit una vegada i una altra, des de Downs endavant, que la celebració periòdica d'eleccions exerceix una poderosa influència sobre l'acció dels partits que arriben al govern: una vegada escollits, els partits i els seus membres es podrien desentendre de les promeses fetes en campanya i dedicar-se a enriquir-se personalment o a gaudir dels múltiples privilegis que comporta l'exercici del poder polític.

En les democràcies no hi ha cap mecanisme legal que obligui els polítics a complir els programes. Però si actuessin al marge dels programes, saben que comprometrien les futures possibilitats de ser reelegits (els votants els castigarien). Si es volen mantenir en el poder o continuar tenint expectatives de tornar a guanyar les eleccions, els partits no es poden desviar gaire del que van prometre als electors, ja que condicionaran el futur vot a la gestió anterior dels partits.

Bibliografia

Downs, A. (1957). *An Economic Theory of Democracy*. Nova York: Harper and Row.

En analitzar els jocs repetits, el que es pretén és esbrinar si hi ha equilibris nous basats en estratègies condicionals que no hi havia en el joc de referència. Pel que fa a això, resulta fonamental la diferència entre jocs repetits un nombre de vegades determinat i jocs repetits indefinidament. Aviat veurem que, quan hi ha un final conegut per tots dos jugadors, la repetició del joc gairebé no canvia, mentre que, si no hi ha un final establert, és a dir, si el joc es continua jugant sempre (o no se sap quan acabarà), els equilibris del joc repetit són molt diferents respecte als del joc de referència.

2. El temps i el factor de descompte

Quan es fan càlculs sobre els pagaments que reben els jugadors en cada període, ronda o repetició del joc, s'ha de tenir en compte que el temps no passa debades. No sempre podem suposar que la utilitat final que obté cada jugador és simplement la suma dels pagaments aconseguits en cada període. Això és així per dues raons que es complementen entre si:

1) Una de tècnica

La raó tècnica és molt senzilla

Si el joc es juga indefinidament, la suma dels pagaments en cada període dóna una quantitat infinita. És a dir, si en cada període el jugador rep un pagament de dues unitats d'utilitat, el pagament total en un joc d'aquestes característiques és infinit. Però això, clarament, no té sentit. La utilitat serà infinita positiva amb cada estratègia que proporcioni pagaments per sobre de 0, cas en què no es pot distingir entre els resultats que produeixen les diverses estratègies.

2) Una altra de substantiva

Raó substantiva

La raó substantiva es pot expressar de diferents maneres, encara que la idea subjacent és sempre la mateixa: valorem menys un futur pagament idèntic en quantitat que no un pagament present (ens proporciona menys utilitat). És a dir, entre rebre un euro avui i rebre'l d'aquí a un any, ens solem estimar més rebre'l avui. Això pot ser o bé perquè l'agent sigui impacient (potser perquè realment necessita els diners immediatament), o bé perquè temi el futur, que es presenta incert (potser tem que hi hagi certa probabilitat que el joc es pugui acabar en el període següent).

En cas que els pagaments siguin monetaris, la justificació encara és més clara: si ens donen l'euro, el podem invertir perquè ens proporcioni algun tipus d'interès. En canvi, si el rebem al cap d'un any, durant l'any d'espera hem deixat de guanyar el benefici que podríem haver aconseguit si n'haguéssim disposat des del començament.

Per a entendre com es produeix aquesta pèrdua en valor present d'un pagament futur a mesura que el pagament s'allunya en el temps, suposarem que és monetari. D'aquesta manera, podem ser més precisos i introduir la taxa d'interès. Tenim una taxa d'interès, r, de l'1%. Això significa que si avui posem al banc 100 euros, al cap d'un any en tenim 101 ($101 = 100\,(1 + r)$, $r = 0{,}01$). Per tant, no podem ser indiferents entre rebre 100 euros avui i rebre'n 100 d'aquí a un any. Serem indiferents més aviat entre rebre avui 100 euros i rebre d'aquí a un any 101 euros. Podem entendre la taxa d'interès r més generalment com a *taxa de descompte*, i es pot interpretar com la quantitat extra d'una unitat de pagament que necessitem per a compensar el retard amb què rebem el pagament. En l'exemple, la taxa de descompte seria de 0,01 euros per euro.

2.1. El factor de descompte

Aquesta mateixa idea es pot reflectir de manera alternativa: en lloc d'establir quant hauria d'augmentar el futur pagament perquè l'agent fos indiferent entre rebre'l avui i rebre'l en el pròxim període, podem calcular quant valora l'agent en el present un futur pagament. Ho aconseguim amb el *factor de descompte* δ, que definim com el següent:

$$\delta = \frac{1}{1+r}$$

Així, un pagament de 100 euros d'aquí a un any valdria avui aquesta quantitat:

$$100\left(\frac{1}{1+0{,}01}\right) = 100 \times 0{,}99 = 99$$

El futur pagament expressat en valor present té un descompte o una depreciació. Com més elevada és la taxa de descompte, més baix és el factor de descompte, la qual cosa implica, atès que δ varia entre 0 i 1, que descomptem més valor amb el pas del temps. Si δ està pròxim a 1 (i *r*, per tant, pròxim a 0), com en l'exemple que utilitzem, això significa que el descompte és molt petit. Si δ s'allunya d'1 i s'acosta a 0, l'agent descompta molt el futur: és molt impacient o la seva incertesa sobre el final del joc és molt alta.

Aquí utilitzarem el factor de descompte δ, deixant de banda la taxa de descompte. El que queda per resoldre ara és com es descompta el valor per a períodes més allunyats en el temps que en el període següent. Recordeu que el valor de δ és el factor de descompte aplicat al període immediat. Quan es consideren diferents (o infinits) futurs períodes, s'ha de decidir la manera en què s'aplica δ. El suposit que gairebe sempre s'empra en la teoria de jocs és que *el descompte és exponencial*, de manera que el valor present d'un pagament π és $\delta^{0\pi}$ a t =0, $\delta^{1\pi}$ a t = 1, $\delta^{2\pi}$ a t = 2, $\delta^{3\pi}$ a t = 3 i així successivament. Matemàticament, el descompte exponencial per a un període infinit es pot representar així respecte a un pagament qualsevol π:

$$\delta^0\pi + \delta^1\pi + \delta^2\pi + \ldots = \sum_{t=0}^{\infty} \delta^t \pi$$

Aquesta suma no és infinita, sinó que convergeix en una quantitat determinada sempre que 0 < δ < 1. Amb això es resol la dificultat tècnica abans apuntada que en un joc indefinit els pagaments al llarg del temps es tornin infinits. Vegem per què això és així. Suposem primer, per al cas més senzill, que π = 1. En aquest cas, la sèrie convergeix de la manera següent:

$$1 + \delta + \delta^2 + \ldots = \frac{1}{1-\delta}$$

La demostració és molt senzilla: si δ està delimitat entre 0 i 1, en cada període successiu, en ser elevat a un exponent cada vegada més elevat, es torna més petit, amb la qual cosa necessàriament cada nou terme de la sèrie serà més petit i es garantirà, així, la convergència. Per a determinar el valor concret en què convergeix, diguem de moment que aquest valor és *s*. Traient factor comú en la sèrie, arribem a això:

$$s = 1 + \delta + \delta^2 + \delta^3 + \ldots = 1 + \delta(1 + \delta + \delta^2 + \ldots) = 1 + \delta s$$

Per tant, si passem δ *s* a l'esquerra restant i treiem factor comú,

$$s(1 - \delta) = 1$$

Ara només falta aïllar *s*:

$$s = \frac{1}{1 - \delta}$$

Si π és qualsevol altra quantitat que no sigui 1, llavors, aplicant la mateixa lògica, la sèrie convergeix de la manera següent:

$$\pi + \delta\pi + \delta^2\pi + \delta^3\pi + \ldots = \frac{\pi}{1 - \delta}$$

La demostració segueix els mateixos passos que l'anterior.

$$\delta\pi + \delta^2\pi + \delta^3\pi + \ldots = \frac{\delta\pi}{1 - \delta}$$

Una vegada exposada la idea del descompte exponencial del futur, podem analitzar els jocs repetits.

3. Jocs repetits *n* vegades

La teoria de jocs repetits *n* vegades es basa en un resultat o proposició bàsica. Si el joc de referència que es repeteix té un únic equilibri de Nash, llavors aquest equilibri és l'original del joc de referència. Per tant, la repetició del joc un nombre de vegades fix no altera l'equilibri original, no introdueix cap canvi. Així, un DP repetit cinc vegades es juga igual que un DP jugat una sola vegada. No hi ha possibilitat que sorgeixin estratègies condicionals.

La lògica que hi ha darrere d'aquest resultat no és més que la lògica de retroinducció.

Situem-nos en l'última ronda o període del DP (en la cinquena). És evident que, com que el joc no continuarà, l'única elecció racional consisteix a escollir l'estratègia dominant, 'defraudar'. Els dos jugadors defrauden en la ronda final. El DP s'hi juga igual que si es jugués una sola vegada. En la quarta ronda, els jugadors són capaços d'anticipar què passarà en l'última i, sabent, per tant, que tots dos defraudaran, entenen que el millor que poden fer ara també és defraudar. No té sentit que es plantegin cooperar en la quarta ronda per a condicionar la cooperació del rival en la roda següent, ja que cada un sap que l'altre defraudarà amb seguretat en la ronda última i no tenen cap incentiu per a cooperar, per molt que l'altre hagi cooperat en el passat. Però aquest mateix raonament es pot traslladar a la tercera ronda. Ara *J*1 i *J*2 saben que tots dos defraudaran en les rondes quarta i cinquena, amb la qual cosa tornen a decidir defraudar. Arribem així fins a la primera ronda, en què els dos jugadors defrauden.

És l'expectativa del que succeirà en l'última ronda del joc repetit el que arruïna la possibilitat que aparegui alguna forma de cooperació condicional entre els jugadors. En treure les conseqüències lògiques del coneixement cert que en l'última ronda tots dos dos defraudaran, inicien el joc en la primera ronda defraudant. Es tracta del mateix argument que s'aplicava en el cas del joc del centpeus.

Conclusió

La conclusió es dóna tant si el joc es repeteix dues vegades com si ho fa dues mil. Fins i tot quan el joc es juga durant 2.000 períodes, és la certesa sobre el que succeirà en la ronda 2000 el que impedeix que en les 1.999 rondes anteriors hi pugui haver alguna forma de cooperació condicional.

Ara bé, de la mateixa manera que en el cas del centpeus concloïem que el criteri de retroinducció va massa lluny, ja que a la pràctica les persones recolzen cadenes de cooperació durant bona part del joc, en el cas del DP o de qualsevol altre joc repetit *n* vegades la predicció de la teoria també s'allunya molt del que s'aprecia experimentalment.

S'ha comprovat de manera sistemàtica que la gent està disposada a cooperar en un DP repetit *n* vegades, sense fer cas de l'argument retroinductiu que, si s'anticipa que en l'última ronda es defraudarà, llavors no compensa cooperar abans.

En qualsevol cas, aquest resultat tan poc convincent només serveix si, com s'ha vist abans, el joc de referència que es repeteix *n* vegades té un únic equilibri de Nash. Si el joc de referència té múltiples equilibris, poden sorgir estratègies condicionals, encara que moltes vegades donen lloc a equilibris poc raonables.

4. El dilema del presoner repetit indefinidament

El DP és un dels jocs que, com vam veure en el mòdul 3, reflecteixen el problema de la cooperació (Quadre 16). Concretament, es va explicar que el DP representa la configuració de pagaments més desfavorable possible per a la cooperació, perquè l'estratègia de defraudar domina fortament l'estratègia de cooperar. Hi ha una multitud de situacions socials i polítiques que queden representades pel DP. Si escollim el DP no és solament perquè es doni sovint, sinó perquè l'anàlisi del que succeeix quan es repeteix contribueix a dissipar en part la sospita que hi ha algun element paradoxal o absurd en la conclusió que la racionalitat individual sempre obliga en un DP a sacrificar els guanys col·lectius pels personals, amb el resultat que finalment els dos jugadors estan pitjor que si no haguessin actuat segons els dictats d'aquesta racionalitat individual. La paradoxa només es dóna quan el DP no es repeteix.

Quadre 16. El dilema del presoner

		*J*2	
		C	*D*
*J*1	*C*	2, 2	0, 3
	D	3, 0	1, 1

Refrescar el problema

Per a refrescar el problema, en el quadre 1 es torna a reproduir el DP, definit com un ordre de preferències particular en relació amb el joc genèric de la cooperació. L'estratègia *D* domina *C* perquè $T > R$ i $P > S$. En conseqüència, l'únic equilibri de Nash és (*D*, *D*).

Quadre 1. El dilema del presoner

		*J*2	
		C	*D*
*J*1	*C*	*R, R*	*S, T*
	D	*T, S*	*P, P*

R= recompensa; *T*= temptació; *P*= penalització; *S*= *sucker* (babau); Ordre de preferències: $T > R > P > S$

Encara que no es demostrarà aquí, resulta que, si el joc es repeteix indefinidament, el DP passa a tenir infinits equilibris. En lloc d'entrar en la demostració d'aquest sorprenent resultat, examinarem simplement alguns dels equilibris possibles i ens centrarem en un d'especialment interessant, l'equilibri basat en l' **estratègia Tit for Tat** ('tal faràs, tal trobaràs').

En primer lloc, l'equilibri original del joc de referència es manté com a equilibri del joc repetit indefinidament. És a dir, una possibilitat de jugar el joc consisteix que tots dos defraudin permanentment. Ens pot semblar que això presenta un panorama una mica gris, però això no treu que sigui un equilibri de ple dret. Si *J*1 espera que *J*2 respongui a la cooperació defraudant, i *J*2 espera el mateix de *J*1, cap dels dos no té raons per a cooperar: les seves expectatives es reforcen mútuament i fan que defraudar sigui la resposta òptima davant de l'elecció de defraudar de l'altre. Els pagaments esperats de tots dos jugadors en aquest equilibri són els següents:

$$P+\delta P+\delta^2 P+\delta^3 P+\ldots=\frac{P}{1-\delta}$$

Resulta xocant que això pugui ser un equilibri, ja que és evident que, si cooperessin entre si, podrien aconseguir pagaments molt superiors. Tanmateix, aquest és el resultat més lògic si domina la desconfiança absoluta entre els jugadors. En qualsevol cas, és important adonar-se que la combinació oposada d'estratègies, cooperar sempre, no és un equilibri de Nash ni en el joc de referència ni en el joc repetit indefinidament. Si en la versió repetida *J*1 espera que *J*2 cooperarà sempre, la millor resposta possible de *J*1 a aquesta expectativa no és cooperar també, sinó defraudar. Si els dos jugadors juguessin l'estratègia de cooperar sempre, el pagament esperat de cada un seria el següent:

$$\frac{R}{1-\delta}$$

Però cada un podria estar millor defraudant sempre si l'altre jugués aquesta estratègia de sempre cooperar. Concretament, si *J*1 defrauda sempre mentre que *J*2 coopera sempre, *J*1 obtindria

$$\frac{T}{1-\delta}$$

essent

$$\frac{T}{1-\delta}>\frac{R}{1-\delta}$$

ja que hem partit d'un ordre de preferències en què $T > R$. Per tant, *J*1 tindria un incentiu per a desviar-se del parell d'estratègies de cooperar sempre, amb la qual cosa cooperar sempre no pot ser un equilibri de Nash.

Tant les estratègies de defraudar sempre com de cooperar sempre són incondicionals. L'acció del jugador no depèn del que hagi fet fins llavors el rival. La primera estratègia, defraudar sempre, és un equilibri, mentre que la segona, cooperar sempre, no ho és.

A part d'aquestes dues estratègies incondicionals, hi ha una infinitat d'estratègies condicionals possibles. Són aquestes estratègies condicionals les que expandeixen l'espai d'equilibris possibles.

4.1. Exemple d'una estratègia condicional

Vegem alguns exemples d'estratègies condicionals. Tenim **l'estratègia «desoladora»** (***grim-trigger strategy***), que consisteix a començar cooperant, continuar cooperant sempre que el rival cooperi i, en cas que en algun moment el rival defraudi, defraudar sempre a partir d'aquí. Per a determinar si aquesta estratègia, jugada per tots dos jugadors, representa un equilibri, s'ha d'esbrinar si, atesos els pagaments que produeix, algun dels jugadors tindria una raó per a desviar-se'n. Si els dos jugadors juguen l'estratègia «desoladora», tots dos comencen cooperant i continuen cooperant indefinidament, sense necessitat que hagin de posar en pràctica el càstig que es conté en la seva estratègia (defraudar fins a l'eternitat en cas que el rival defraudi en una ocasió). Els pagaments que obtindrien serien, per tant, els següents:

$$\frac{R}{1-\delta}$$

Per a veure si algun dels jugadors té incentius per a desviar-se de la seva estratègia, s'ha de comparar el pagament que obtindria desviant-se'n amb el que acabem de veure que obtindria si continués jugant l'estratègia «desoladora». Suposem que $J1$ en un període qualsevol t defrauda. A t, $J1$ rep el pagament màxim, T, el pagament de la temptació, però a partir d'aquell moment (a partir de t +1), sap que $J2$ defraudarà sempre i, per tant, el millor que pot fer $J1$ a partir d'ara és defraudar sempre. Esquemàticament, podem representar la història del joc així:

Quadre 2. Una desviació de *J1* en t de l'estratègia desoladora

	$t-1$	T	$t+1$	$t+2$	...	Pagament esperat a partir de t
$J1$	C	D	D	D	...	$T + \delta P + \delta^2 P + ... = T + \frac{\delta P}{1-\delta}$
$J2$	C	C	D	D	...	$S + \delta P + \delta^2 P + ... = S + \frac{\delta P}{1-\delta}$

Ara ja podem saber sota quines condicions a $J1$ no li compensa desviar-se de l'estratègia «desoladora». Concretament, no el compensarà quan el pagament esperat de continuar-la jugant sigui superior al pagament esperat de desviar-se una vegada. És a dir, quan

$$\frac{R}{1-\delta} > T + \frac{\delta P}{1-\delta}$$

Es tracta d'aïllar δ en aquesta equació. Ho podem fer de la manera següent: multipliquem a tots dos costats per (1 – δ) per a deslliurar-nos de les fraccions i després posem en un costat tots els termes que multipliquen per δ i en l'altre tots els termes que són lliures de δ:

$$R > T(1-\delta)+\delta P$$
$$R > T-\delta T+\delta P$$
$$\delta T-\delta P > T-R$$
$$\delta(T-P) > T-R$$
$$\delta > \frac{T-R}{T-P}$$

Com que el joc és simètric i considerem, per simplicitat, que tots dos jugadors tenen el mateix factor de descompte, podem concloure que, sempre que els jugadors siguin prou pacients, és a dir, sempre que

$$\delta > \frac{T-R}{T-P}$$

l'estratègia «desoladora», que és una estratègia condicional, configura un equilibri. Això significa que, si cada un dels jugadors juga l'estratègia «desoladora» i el seu factor de descompte és superior a la quantitat indicada, cap dels dos no té incentius per a desviar-se.

4.2. Variacions d'estratègia desoladora: Tit for Tat

L'estratègia «desoladora» és solament una de les infinites estratègies condicionals possibles. Hi ha innombrables variacions de l'estratègia «desoladora». Per exemple, començar cooperant, continuar cooperant mentre l'altre cooperi i, en cas que l'altre defraudi una vegada, castigar-lo a partir de llavors defraudant en les 1.423 rondes següents. Com abans, si el factor de descompte satisfà certa condició, aquesta estratègia pot ser un equilibri.

Entre totes les variants de l'estratègia «desoladora» n'hi ha una que destaca clarament, l'estratègia Tit for Tat, identificada per primera vegada per Anatol Rapoport. Aquesta estratègia es pot formular així:

1) El jugador comença cooperant en la primera ronda.

2) En totes les altres rondes, el jugador fa el que va fer el seu rival en la ronda anterior.

D'aquesta manera, el jugador coopera en la ronda t si el seu rival va cooperar en la ronda $t-1$, i defrauda en t si el seu rival va defraudar en $t-1$. Suposem que el rival, $J2$, va defraudar en $t-1$. Llavors, $J1$ defrauda en t. Si en t $J2$ coopera, $J1$ coopera en $t+1$; si $J2$ defrauda en t, $J1$ torna a defraudar en $t+1$. I així successivament.

Quina relació té Tit for Tat amb l'estratègia «desoladora»? Tit for Tat és una variant de l'estratègia «desoladora» perquè, de la mateixa manera que aquesta estratègia, produeix cooperació mentre el rival coopera, i es distingeixen en el fet que el càstig que aplica quan el rival defrauda és el màxim de suau possible, defraudar en ronda la següent, mentre que amb l'estratègia «desoladora» el càstig és el màxim possible, defraudar fins a l'eternitat.

Sota quines condicions representa un equilibri que els dos jugadors juguin l'estratègia Tit for Tat? Quan continuar jugant-la sigui millor que desviar-se'n.

1) Desviar-se només en una ronda de Tit for Tat defraudant una vegada després que el rival hagi cooperat en la ronda anterior i tornant després a l'estratègia original.

2) Desviar-se per sempre, defraudant de manera permanent a partir de la primera desviació.

És possible demostrar que sota certes condicions Tit for Tat produeix millors resultats que qualsevol de les dues desviacions, vol dir que aquesta estratègia és un equilibri. En rigor es demostra que, de la mateixa manera que no compensa ni una ni infinites desviacions, tampoc compensa un número intermedi n de vegades, $1 < n < \infty$.

4.3. Característiques de l'estratègia Tit for Tat

En qualsevol cas, què té d'especial Tit for Tat davant d'altres estratègies condicionals? Robert Axelrod, en el seu seu famós llibre *La evolución de la cooperación*, va fer un torneig de lligueta entre estratègies per a un DP repetit indefinidament.

Els participants en la lligueta enviaven el que creien que seria la millor estratègia possible, l'estratègia que maximitzaria els pagaments. Axelrod va traduir totes aquestes estratègies a un llenguatge de programació i va posar a competir cada estratègia contra totes les altres. Concretament, cada estratègia jugava dues-centes rondes del DP amb cada una de les altres estratègies. L'estratègia guanyadora va ser precisament la Tit for Tat. Era la més senzilla de totes les estratègies proposades. Segons Axelrod, les raons d'aquest inesperat èxit de Tit for Tat s'expliquen per algunes de les seves característiques:

Bibliografia

Axelrod, R. (1984). *The evolution of cooperation*. Nova York: Basic Books. (Està traduït a l'espanyol a Alianza, 1986).

1) És una estratègia **decent**, entenent per decència que comença cooperant, que no és la primera a defraudar.

2) És una estratègia **indulgent**, en el sentit que reprèn la cooperació amb relativa rapidesa després d'haver defraudat el rival.

3) És una estratègia **venjativa**, ja que, si el seu rival defrauda, no passa per alt una ofensa semblant i també defrauda.

4) És una estratègia **clara** el funcionament de la qual és transparent i fàcil d'entendre. Aquesta combinació de característiques fan de Tit for Tat una estratègia guanyadora en el tipus de torneig que va organitzar Axelrod.

Això no obstant, perquè Tit for Tat tingui èxit cal que hi hagi altres estratègies decents. Tit for Tat obté molt bons resultats si es fonamenta en altres estratègies decents amb què pugui establir relacions duradores de cooperació condicional. En canvi, funciona pitjor davant d'estratègies no decents: per exemple, Tit for Tat davant d'una estratègia que comenci defraudant i a continuació cooperi llevat que l'altre defraudi, cas en què defrauda per sempre, produeix una seqüència indefinida de defeccions mútues, ja que en la primera ronda Tit for Tat coopera i l'altra defrauda; en la segona ronda, Tit for Tat defrauda i l'altra coopera, i a partir d'aquí l'altra ja defrauda sempre.

En qualsevol cas, que Tit for Tat tingui unes propietats tan atractives no significa necessàriament que, quan els jugadors juguin un DP, seleccionin l'equilibri corresponent a aquesta estratègia. Al cap i a la fi, Tit for Tat és tan sols un equilibri possible dins d'un conjunt d'equilibris infinit. En realitat, no és evident que tot jugador racional hagi de ser conscient de les característiques especials d'aquesta estratègia.

D'altra banda, quan veiem casos reals de cooperació condicional reeixida, és difícil saber quines estratègies sostenen la cooperació. Fixeu-vos que qualsevol variant de l'estratègia «desoladora», incloent-hi Ti tfor Tat i la mataeixa estratègia «desoladora», produeixen cooperació condicional. Amb altres paraules, la cooperació condicional és compatible o sostenible amb múltiples estratègies possibles.

Davant dels resultats de l'anàlisi del DP repetit indefinidament, podem sentir la temptació de concloure que les estratègies condicionals són el que expliquen la presència d'un grau de cooperació en les societats més elevat del que es podria esperar des de la teoria de jocs. Tanmateix, convé subratllar i insistir que estratègies com TIt for Tat només són un equilibri quan dues persones juguen el DP. En canvi, les formes de cooperació col·lectiva en la societat, quan es donen, solen involucrar un gran nombre de jugadors.

> A mesura que augmenta el nombre de jugadors que intervenen en un DP, les possibilitats que sorgeixi cooperació condicional disminueixen ràpidament, per molt que es repeteixi el joc i per molt pacient que sigui cada un dels jugadors.

Senzillament, si el grup és molt gran, l'únic equilibri realista d'un únic DP jugat entre tots els membres del grup consisteix a fer que tots defraudin permanentment.

5. El teorema popular (*the folk theorem*)

Fins ara hem vist que, quan el DP es juga indefinidament, sorgeixen múltiples equilibris de Nash (l'estratègia «desoladora» i les seves múltiples variants). Es pot demostrar que la repetició indefinida afecta igual tots els jocs, no solament el DP.

> L'anomenat *teorema popular* demostra que els jocs repetits indefinidament tenen infinits equilibris.

Encara que aquest teorema resulta molt interessant, ja que revela l'enorme diferència que hi ha entre els jocs jugats una vegada i els jugats indefinidament, en el fons compromet moltíssim la capacitat predictiva de la teoria. La teoria no pot aclarir quin dels infinits equilibris seleccionaran els jugadors.

Ens limitarem a presentar la idea general que guia la demostració del teorema.

Bibliografia

Kreps, D. (1990a). *A course in microeconomic theory* (pàg. 507-508). Princeton: Princeton University Press.

Dutta, P.K. (1999). *Strategies and games* (pàg. 232-234). Cambridge, Massachusetts: The MIT Press.

Per a això, necessitem dos conceptes nous:

1) El de cicle

Concepte de cicle

Es defineix tècnicament un cicle com qualsevol combinació possible d'accions al llarg del temps.

Per exemple, en el DP un cicle podria ser: (C, C) durant T_1 rondes, (C, D) durant T_2 rondes, (D, C) durant T_3 rondes i (D, D) durant T_4 rondes. El cicle $T_1 + T_2 + T_3 + T_4$ quedaria llavors caracteritzat pels valors de T_i, que en ocasions poden ser zero. Per exemple, la combinació d'estratègies de defraudar sempre seria un cicle amb aquests valors: $T_1 = T_2 = T_3 = 0$, $T_4 = 1$.

2) El de pagament minimax

Concepte de pagament

D'altra banda, el pagament minimax és el millor que pot obtenir un jugador quan el seu rival té com a única finalitat enfonsar-li els pagaments.

Es tracta del pagament màxim entre els mínims que li deixa el rival, d'aquí ve el nom de *minimax*. Ara és irrellevant considerar si el rival té bones raons per a buscar abans de tot fer la vida impossible al nostre jugador. Només importa establir el concepte.

Suposem que *J*1 i *J*2 es troben en un joc qualsevol. Podria ser que cap d'ells no tingui incentius per a desviar-se de l'estratègia d'equilibri, sigui quina sigui, si la conseqüència d'aquesta desviació consisteix a rebre el pagament minimax. És a dir, els equilibris es poden sostenir sobre la base que si *J*1, per exemple, es planteja desviar-se de l'equilibri, anticipa que la seva desviació farà realitat l'amenaça de *J*2 d'aplicar el càstig del pagament minimax. *J*2 amenaça *J*1 de condemnar-lo a rebre el seu pagament minimax si *J*1 es desvia de l'equilibri. Si *J*1 fa la mateixa amenaça, castigar *J*2 amb el pagament minimax en cas de desviació, *J*2 tampoc no tindrà raons per a desviar-se de l'estratègia que s'hagi configurat en equilibri.

Doncs bé, el que estableix el teorema popular és simplement que qualsevol cicle d'estratègies que proporcioni pagaments superiors al pagament minimax pot ser un equilibri de Nash gràcies al mecanisme de les amenaces mútues de forçar aquest pagament en cas que es produeixi una desviació.

En basar la demostració en la idea de cicle d'estratègies, el que s'aconsegueix és posar de manifest que qualsevol combinació d'estratègies, per molt absurda o arbitrària que ens sembli, pot arribar a ser un equilibri de Nash sempre que els pagaments que proporcioni el cicle siguin millors que el pitjor pagament possible, el pagament que rebria el jugador si se li ocorregués desviar-se en una ocasió del cicle d'estratègies proposat. El pitjor pagament possible és el pagament minimax, és a dir, el que resulta quan el rival no té més propòsit que fastiguejar l'altre jugador.

El teorema popular té gran importància. D'una banda, assenyala la increïble varietat d'equilibris possibles en els jocs repetits indefinidament. Gairebé qualsevol combinació d'estratègies pot arribar a ser un equilibri. Però, de l'altra, aquest teorema revela les fortes limitacions de la teoria de jocs com a teoria predictiva. La teoria estableix un marge molt gran de resultats possibles i no pot especificar quin equilibri se seleccionarà finalment en el conjunt de tots els equilibris.

6. El model de negociació de Rubinstein

En aquesta secció final, veurem el model de negociació de Rubinstein com un joc repetit tres vegades. L'anàlisi d'aquest model ens permetrà d'utilitzar tot l'instrumental analític que hem après fins ara, tant la idea d'equilibri de perfecció en el subjoc com el descompte temporal dels jocs repetits.

En una negociació, dues o més parts s'han de posar d'acord en com s'han de repartir un bé. Aquí ens limitarem al cas més simple, en què només hi ha dos negociadors. Correspon a Nash el mèrit d'haver ofert per primera vegada el 1950 una solució als problemes de negociació basada en la teoria de jocs (en aquest sentit, convé no confondre la idea d'equilibri de Nash amb la idea de solució negociadora de Nash). El que va fer Nash va ser establir una sèrie de condicions que hauria de complir una proposta de solució, i va demostrar axiomàticament que la seva proposta les satisfeia. Davant d'aquesta solució axiomàtica, **Ariel Rubinstein** va plantejar el 1982 un model que no es basava en condicions ideals, sinó que partia directament de la racionalitat dels agents i del seu poder de negociació. En realitat, el mateix Nash havia plantejat també de manera embrionària una solució basada només en la racionalitat, però el desenvolupament complet va haver d'esperar fins al treball de Rubinstein. El model de Rubinstein consisteix en un joc repetit indefinidament, tot i que aquí se n'exposarà una versió simplificada en només tres períodes en què s'arriba, en essència, al mateix resultat.

Bibliografia

Osborne, M.; Rubinstein, A. (1990). *Bargaining and markets* (pàg. 29-49). San Diego: Academic Press.

6.1. El poder negociador

La clau per a entendre un problema de negociació consisteix en com s'han treure conseqüències del poder negociador de les parts. Suposem que el que hi ha en joc és simplement un euro. Dues persones s'han de posar d'acord en com se l'han de repartir: si no ho aconsegueixen, tots dos es queden sense res.

Per a Nash, el poder negociador queda reflectit en l'actitud envers el risc de l'individu.

Bons resultats

Com més necessitat tingui una de les parts d'aconseguir bons resultats, més avers al risc serà l'agent, és a dir, menys s'arriscarà que hi hagi un desacord final i menys capacitat d'amenaçar tindrà en l'intercanvi d'ofertes i contraofertes.

Rubinstein parteix del mateix supòsit, només que identifica el poder negociador amb el factor de descompte dels individus.

> Com més baix sigui el factor de descompte, és a dir, com més impacient sigui l'agent, menys poder negociador tindrà, ja que més pressa tindrà per a aconseguir un acord, fins i tot a costa d'obtenir potser el pitjor resultat possible.

En el model hi ha dos jugadors, *J*1 i *J*2. A diferència del que hem fet fins ara, considerarem que cada jugador té el seu factor de descompte, que pot ser diferent del de l'altre jugador. Així, parlarem de δ_1 i δ_2 respecte a *J*1 i *J*2 respectivament. Els períodes de negociació són tres. Cada un es compon de dues parts, una oferta (x, $1 - x$) i una resposta a l'oferta, en què x és la part que s'emporta *J*1 i $1 - x$, la que s'emporta *J*2. Si un jugador accepta l'oferta del seu rival, el joc acaba. Si no l'accepta, es passa al període següent. *J*1 proposa ofertes en els períodes senars i *J*2 en els parells.

Els pagaments del joc corresponen a la mateixa oferta que finalment s'accepti (llevat que els actors no es posin mai d'acord, cas en què cada un obté 0). El valor en el present dels pagaments que s'obtenen en el període t s'ha de descomptar exponencialment, de manera que una oferta (x, $1 - x$) acceptada en el període t val, en termes presents,

$$\left(\delta_1^t x,\ \delta_2^t(1-x)\right)$$

Aquest joc té múltiples equilibris de Nash, la majoria dels quals no tenen gaire sentit.

Per exemple, la següent combinació d'estratègies és un equilibri de Nash

(J1 sempre demana 1 i rebutja qualsevol oferta inferior; *J*2 sempre ofereix 1 i accepta qualsevol oferta.)

Aquest parell d'estratègies es traduiria així en aquest context: *J*1 fa l'oferta (1, 0) en la primera ronda, per la qual *J*1 es queda amb tot l'euro i *J*2 amb res, i *J*2 ho accepta, amb la qual cosa s'acaba el joc. Fixeu-vos que si *J*2 està convençut que *J*1 rebutja qualsevol altra cosa i *J*1 creï que *J*2 es conforma amb qualsevol oferta, cada estratègia és una resposta òptima a l'altra.

6.2. El equilibri de perfecció en el subjoc

Rubinstein va demostrar que, malgrat que hi ha una multiplicitat d'equilibris de Nash, hi ha un únic equilibri de perfecció en el subjoc.

Aquesta demostració només depèn d'un supòsit: que quan un jugador sigui indiferent entre acceptar i rebutjar una oferta, tria acceptar-la.

En el cas més extrem, això implica que, si un jugador, en l'últim període, ha d'escollir entre una oferta que li deixa 0 euros i rebutjar-la i provocar la desaparició de l'euro, accepta l'oferta. El jugador s'estima més acceptar l'oferta (1, 0) a rebutjar-la.

Per a calcular l'equilibri de perfecció en el subjoc, s'ha de procedir per retroinducció, d'enrere cap endavant. Es desenvoluparà l'argument en tres fases, a, b i c.

Esquemàticament, el procés es podria representar com apareix en el quadre 5.

Quadre 5. El model de negociació de Rubinstein en tres períodes

	Oferta de *J*1	Oferta de *J*2
t= 1	$(1 - \delta_2 (1 - \delta_1 1), \delta_2 (1 - \delta_1 1))$	
t = 2		$(\delta_1 1, 1 - \delta_1 1)$
t = 3	(1, 0)	

a) Comencem pel període final, $t = 3$. Havent-hi només tres períodes d'ofertes i fent ofertes *J*1 en els períodes senars, *J*1 té capacitat d'ultimàtum en l'última ronda. *J*1 pot forçar al màxim *J*2, ja que, si *J*2 rebutja l'oferta final de *J*1, *J*2 es queda sense res. D'aquesta manera, a $t = 3$ *J*1 fa l'oferta (1, 0) i *J*2, per la raó anteriorment exposada, accepta.

b) A $t = 2$, li toca a *J*2 fer una oferta. *J*1 només l'acceptarà si és almenys tan bona com el que pot aconseguir en el període següent, $t = 3$. El que pot aconseguir a $t = 3$ és 1, que en el període $t = 2$ ho valora com a $\delta_1 1$. Per tant, l'oferta de *J*2 serà $(\delta_1 1, 1 - \delta_1 1)$. Està oferta fa *J*1 indiferent entre acceptar i rebutjar. Suposem que *J*1 l'accepta.

c) A $t = 1$, *J1* fa una oferta que *J*2 acceptarà només si el que li toca és almenys tan bo com el que podria aconseguir rebutjant l'oferta i passant al període següent. El que *J*2 pot aconseguir a $t = 2$ és $1 - \delta_1 1$ i, per tant, *J*2 serà indiferent a $t = 1$ entre aconseguir $\delta_2(1 - \delta_1 1)$ a $t = 1$ o aconseguir-ne $1 - \delta_1 1$ a $t = 2$. Per tant, l'oferta de *J*1 serà $(1 - \delta_2(1 - \delta_1 1), \delta_2(1 - \delta_1 1))$. Quan *J*1 fa aquesta oferta, *J*2 l'accepta immediatament.

L'únic equilibri de perfecció en el subjoc consisteix, per tant, que *J*1 i *J*2 segueixin les estratègies que s'acaben d'especificar. En la primera ronda *J*1 ofereix $(1 - \delta_2(1 - \delta_1 1), \delta_2(1 - \delta_1 1))$ i *J*2 accepta. El joc acaba allà. O si es prefereix, aquesta és la ruta d'equilibri. Això no obstant, aquesta ruta d'equilibri se sosté sobre les respostes òptimes calculades fora de la ruta d'equilibri, és a dir, la ruta d'equilibri és així perquè els jugadors anticipen que en el segon període l'oferta de *J*2 seria $(\delta_1 1, 1 - \delta_1 1)$ i que en el tercer període l'oferta de *J*1 seria (1, 0).

Suposem que $\delta_1 = \delta_2 = 0{,}8$. Tenint el mateix poder negociador, la diferència o asimetria que es produeixi en l'oferta inicial de *J*1 serà conseqüència únicament i exclusivament del poder d'ultimàtum que té *J*1 en l'últim període. Amb aquest factor de descompte comú, l'oferta de *J*1 seria (0,84, 0,16). Si a més variem el factor de descompte, el poder negociador, de manera que *J*1 sigui més pacient que *J*2, la diferència serà encara més important. Tenim $\delta_1 = 0{,}9$ i $\delta_2 = 0{,}7$. Ara l'oferta de *J*1 seria (0,93, 0,07), encara més asimètrica.

Quan el joc es juga indefinidament, havent-hi una successió ininterrompuda d'ofertes i contraofertes fins que s'assoleix un acord, el poder d'ultimàtum de *J*1 desapareix i qualsevol diferència que es produeixi en l'oferta d'equilibri en el primer període es deurà gairebé en la seva totalitat als diferents factors de descompte dels jugadors.

6.3. La informació i la negociació

La propietat més sorprenent del model de Rubinstein, o bé en la versió simplificada en tres períodes, o bé en el joc repetit indefinidament, és que la negociació mai no s'arriba a desenvolupar, sempre s'acaba en la primera ronda amb una oferta d'equilibri que llança *J*1 i que *J*2 accepta. Tanmateix, en el món real veiem que de vegades les negociacions es prolonguen durant llargs períodes de temps. Això no té sentit en termes del model de Rubinstein, ja que els dos jugadors saben que, com més temps passi, més valor perd el bé sobre el qual negocien. Anticipant aquesta pèrdua de valor segons el factor de descompte de cada individu, les parts aconsegueixen arribar a un acord en la primera ronda del joc. Si malgrat això les negociacions no s'acaben sempre en la primera ronda, és perquè la situació real és més complexa que el model de Rubinstein, concretament, perquè hi haurà informació incompleta, és a dir, els jugadors no tindran tota la informació rellevant sobre els rivals i podran avançar en la negociació amb la intenció de demanar part d'aquesta informació que els falta. Aquest tipus de jocs d'informació incompleta s'estudien en el mòdul següent.

Exemples de ciències socials

El model de Rubinstein no té solament aplicacions econòmiques. Vet aquí un parell d'exemples en ciències socials:

- El model s'ha utilitzat per a entendre per què i com es donen les vagues en els enfrontaments entre treballadors i empresaris.
- També s'ha utilitzat per a modelitzar el funcionament intern del Congrés americà.

Bibliografia

Lange, P.; Tsebelis, G. (1993). «Wages, strikes and power: an equilibrium analysis». A: Booth, James i Meadwell (ed.). *Politics and Rationality*. Cambridge: Cambridge University Press.

Baron, D.P.; Ferejohn, J. (1987). «Bargaining and agenda formation in legislatures». *American Economic Review* (72, 2, pàg. 303-309).

Resum

Un joc es pot repetir un nombre determinat de vegades o indefinidament. Si es repeteix n vegades, l'equilibri del joc és el mateix que quan es juga una única vegada. La repetició finita del joc no introdueix cap canvi a causa de la lògica retroinductiva: com que els jugadors saben que l'última ronda és estratègicament equivalent al joc jugat una sola vegada, jugaran en l'última ronda com si juguessin una sola vegada. Anticipant això, en la penúltima ronda succeeix el mateix, i també en l'avantpenúltima, i fins a arribar a la primera ronda.

Si el joc es juga indefinidament, l'equilibri del joc de referència continua essent un equilibri, encara que ara hi ha molts altres equilibris possibles. Per exemple, en un dilema del presoner, la parella d'estratègies (defraudar sempre; defraudar sempre) és un equilibri de Nash. Però també hi ha molts altres equilibris basats en estratègies condicionals. Un d'aquests equilibris destaca sobre els altres: el que té lloc quan els dos jugadors juguen l'estratègia Tit for Tat ('tal faràs, tal trobaràs'). Aquesta estratègia és fàcil d'aplicar, té propietats úniques i serveix per a explicar múltiples casos de cooperació condicional.

Anant més enllà del dilema del presoner, el teorema popular demostra que, si el factor de descompte és prou alt, hi ha infinits equilibris de Nash en un joc repetit indefinidament. Qualsevol combinació d'estratègies que proporcioni pagaments superiors al pagament minimax pot ser un equilibri de Nash. Aquesta multiplicitat d'equilibris fa que la teoria sigui incapaç de determinar què passarà quan el joc es repeteixi.

Tanmateix, hi ha vegades que la repetició del joc no impedeix que hi hagi un únic equilibri. Així, en el model de negociació de Rubinstein, malgrat que hi ha infinits equilibris de Nash, hi ha un únic equilibri de perfecció en el subjoc. Aquest equilibri reflecteix el poder negociador de les parts. El poder negociador es reflecteix en el factor de descompte. Com més impacient és una persona, menys poder tindrà en una negociació, ja que més urgència té per a aconseguir un acord com més aviat millor, fins i tot si és a costa d'aconseguir un acord en pitjors termes.

Bibliografia

Alesina, A. (1988). «Credibility and policy convergence in a two-party system with rational voters». *American Economic Review* (núm. 78, pàg. 796-805).

Axelrod, R. (1984). *The Evolution of Cooperation*. Nova York: Basic Books. (Hi ha traducció espanyola a Alianza, 1986.)

Baron, D.P.; Ferejohn, J. (1987). «Bargaining and agenda formation in legislatures». *American Economic Review* (vol. 72, núm. 2, pàg. 303-309).

Calvert, R. (1995). «Rational actors, equilibrium and social institutions». A: J. Knight (ed.). *Explaining Social Institutions*. Ann Arbor: University of Michigan Press.

Downs, A. (1957). *An Economic Theory of Democracy*. Nova York: Harper and Row.

Dutta, P.K. (1999). *Strategies and Games*. Cambridge, Massachusetts: The MIT Press.

Gibbons, R. (1992). *Game Theory for Applied Economists*. Princeton: Princeton University Press.

Kreps, D. (1990). *A Course in Microeconomic Theory*. Princeton: Princeton University Press.

Lange, P.; Tsebelis, G. (1993). «Wages, strikes and power: an equilibrium analysis». A: Booth; James; Meadwell (ed.). *Politics and Rationality*. Cambridge: Cambridge University Press.

Osborne, M.; Rubinstein, A. (1990). *Bargaining and markets*. San Diego: Academic Press.