

INFORMAZIONE A PRIORI ED INSORGENZA DELL'EQUILIBRIO:

un modello di simulazione

1. Introduzione

Lo studio dei processi decisionali è un elemento fondamentale della scienza economica; la piena comprensione dei meccanismi che essi coinvolgono è tuttavia un obiettivo ancora lontano ed attualmente si dispone solo di modelli che rappresentano tali processi in modo schematico al fine di prevederne gli esiti e valutare gli effetti che su di questi esercitano cambiamenti nel contesto o nelle ipotesi di partenza.

Il modello di scelta più consolidato ed affermato è quello fondato sulla teoria dell'utilità attesa, che si basa su alcune ipotesi molto forti sulla razionalità e sull'informazione degli agenti. Già dagli anni 50 queste ipotesi sono ripetutamente state messe in discussione, e da molte voci è stato affermato che il modello tradizionale di uomo economico non costituisce una base solida su cui costruire una teoria delle scelte in grado di descrivere fedelmente come gli agenti economici si comportino, ma solo, eventualmente, come “dovrebbero” comportarsi. Un primo fondamentale intervento su questo problema è costituito da un articolo di Herbert Simon pubblicato nel 1955, in cui l'autore afferma che “il concetto di uomo economico necessita di una revisione piuttosto drastica”, nel senso che le ipotesi che su di esso vengono tradizionalmente fatte sono troppo forti e non sono sufficientemente supportate dall'evidenza empirica, e questo lascia pensare che la differenza tra il processo decisionale effettivo e quello descritto dalla teoria sia piuttosto marcata. In particolare emerge una sensibile divergenza tra modello teorico e comportamento reale, sia dal punto di vista delle informazioni disponibili che da quello della loro elaborazione da parte dei decisori.

Per quanto riguarda l'informazione di cui gli agenti dispongono a priori, si nota che:

- in molte situazioni gli agenti non conoscono tutte le alternative disponibili; ci sono casi, ad esempio, in cui le opzioni possibili non si presentano simultaneamente ma in sequenza e l'agente deve di volta in volta decidere se accettare o rifiutare;
- è realistico supporre poi che il decisore non sia sempre in grado di mettere in corrispondenza precisa gli elementi dell'insieme delle azioni con elementi o almeno con sottoinsiemi dell'insieme delle conseguenze; per questo può essere necessario introdurre nel modello di decisione anche la rappresentazione del processo di acquisizione di informazioni o di apprendimento.

Inoltre le capacità di calcolo e di elaborazione delle informazioni attribuite dalla teoria al decisore sono probabilmente sovrastimate: si assume infatti che il soggetto economico sia dotato di un'abilità computazionale tale da permettergli di valutare in anticipo, per ogni alternativa a sua disposizione, l'intero corso delle conseguenze e di capire quale sia quella che gli permette di raggiungere l'esito migliore possibile secondo il suo ordinamento di preferenze. In contesti dinamici, ad esempio, il *payoff* che si ottiene da una scelta particolare dipende non solo dalla scelta

stessa, ma anche da tutte le alternative precedentemente adottate in altri problemi di decisione. L'ottimizzazione richiederebbe allora di scegliere, all'inizio della sequenza, una strategia per l'intera serie di scelte da effettuare, ma "una procedura del genere complicherebbe rapidamente il problema oltre la capacità computazionale dell'organismo" (Simon, 1955).

I limiti informativi e computazionali che in realtà sembrano affliggere il decisore ed impedirgli di avere una visione completa del contesto economico in cui è inserito minano le fondamenta di alcune delle conclusioni fondamentali della teoria economica tradizionale. Per esempio, Aumann e Brandenburger (1995) dimostrano che in un contesto statico, per giustificare l'equilibrio di Nash è necessario che gli agenti come minimo conoscano la struttura del gioco e l'insieme delle alternative di cui ciascun soggetto dispone, inoltre bisogna che condividano la medesima *prior* sugli stati del mondo futuri e siano in grado di formulare congetture sulle possibili azioni altrui, e che queste congetture siano conoscenza comune.

Ci si chiede ora cosa accada quando questo complesso di informazioni non è disponibile. Secondo Salmon (1995) "il senso comune [...] suggerisce che i limiti informativi imposti da una conoscenza parziale dell'ambiente economico modellano gli obiettivi degli individui, la loro attività di apprendimento, e in ultimo, ci aspetteremmo, anche il tipo di regola di decisione che essi adottano". Dunque l'informazione disponibile a priori influenza tutto il meccanismo di scelta e ne condiziona in modo significativo l'esito finale. Per questo motivo è importante riuscire a capire a quante e quali informazioni abbiano accesso gli agenti *ex ante*, quando affrontano un particolare problema decisionale per la prima volta e come, nel tempo, essi acquisiscano nuove informazioni e sulla base di esse costruiscano le proprie congetture e credenze sugli stati del mondo. Sotto l'ipotesi di razionalità limitata l'apprendimento diventa così una parte integrante di tutto il processo decisionale e risulta importante trovare modelli in grado di rappresentarlo in modo appropriato.

Lo scopo del presente articolo è quello di mostrare che, adottando un modello di apprendimento che attenua notevolmente i presupposti relativi all'informazione a priori e alle capacità computazionali degli agenti, è comunque possibile formulare un modello di interazione strategica tra agenti economici in cui nel lungo periodo prevale un equilibrio di Nash con caratteristiche di ottimalità sociale.

La struttura dell' articolo è la seguente: nel paragrafo 2 si illustrano brevemente i motivi che hanno portato a rifiutare i modelli di apprendimento utilizzati più comunemente, nel paragrafo 3 viene esposto il modello di apprendimento adattivo evolutivo che si è deciso di adottare al posto dei modelli tradizionali, il paragrafo 4 descrive il modello di simulazione utilizzato, il paragrafo 5 ne presenta i risultati ed infine il paragrafo 6 contiene le conclusioni.

2. Principali modelli di acquisizione ed elaborazione delle informazioni

Uno dei più comuni strumenti di rappresentazione analitica del processo di acquisizione ed

elaborazione di informazioni è il metodo dei minimi quadrati ordinari. Nell'ambito dei modelli di regressione lineare, il metodo dei minimi quadrati fornisce risultati molto soddisfacenti poiché produce gli stimatori lineari non tendenziosi più efficienti.

Tuttavia per diversi motivi sembra difficile che questo metodo possa rappresentare efficacemente il processo di apprendimento degli agenti economici (Salmon, 1995). Innanzitutto bisogna notare che la convergenza degli stimatori verso il loro "vero" valore è piuttosto lenta e per ottenere una valutazione sufficientemente precisa può essere necessario un numero di prove troppo elevato rispetto al tempo che gli agenti hanno a disposizione, anche perché si suppone che l'agente si limiti ad accogliere passivamente i dati disponibili, senza effettuare esperimenti ad hoc per accelerare il processo di apprendimento e per indagare aspetti del fenomeno studiato che altrimenti rimarrebbero in ombra.

Ancora, si fa notare che l'ipotesi di conoscenza a priori della forma funzionale appropriata da dare al modello è piuttosto forte, e che tra l'altro, se essa viene meno, la convergenza dello stimatore al vero valore del parametro non ha luogo e anzi si verificano fenomeni di divergenza ciclica che impediscono all'agente di pervenire a conclusioni definitive sulla forma del fenomeno studiato; siccome poi questo metodo non prevede meccanismi automatici che permettano di modificare in itinere la specificazione del modello, esso non è in grado di rappresentare i processi di aggiustamento che nella realtà vengono svolti dagli agenti, i quali sanno dall'inizio che il modello cui prevengono attraverso l'apprendimento sarà sempre solo un'approssimazione, più o meno precisa, della realtà.

L'ultimo e forse il più grave problema connesso all'adozione del metodo dei minimi quadrati è quello del *turn off* cioè della riduzione progressiva dell'impatto della nuova informazione sul valore dei parametri stimati col tendere a zero della loro varianza. La ragione essenziale di questo è che l'algoritmo di apprendimento in questione assume la coincidenza tra il modello stimato e il processo reale che genera i dati; se però tale processo non fosse costante "potrebbe essere meglio impiegare un tasso di apprendimento costante, piuttosto che uno che decresce nel tempo. In questo modo l'agente potrebbe conservare la flessibilità per imparare continuamente e rispondere alle innovazioni nell'informazione che emergessero da un processo [...] del genere"¹

Un secondo schema di apprendimento molto utilizzato è il modello bayesiano. La regola di Bayes descrive un meccanismo di aggiornamento delle probabilità a priori in seguito all'esperienza che si integra facilmente nell'analisi dei problemi di decisione sia in ambito statistico che economico. Tuttavia presenta anch'essa alcune problematiche di cui bisogna tener conto. Per maggiore chiarezza, conviene partire da un esempio concreto di applicazione dell'approccio bayesiano in un gioco non cooperativo ripetuto (Marimon e McGrattan, 1995). Si supponga che

1 Salmon, 1995

ogni giocatore conosca solo i propri *payoff*, ma non quelli degli avversari, di cui può solo osservare le mosse; osservando la frequenza delle mosse degli avversari e partendo dal proprio *prior* $\mu_i = x_{i,-1}$ sulle loro strategie miste, in ciascun periodo t il giocatore i -esimo rivede la propria distribuzione di probabilità $x_{i,t}$ su queste strategie secondo la formula:

$$x_{i,t}^{\bar{c}} = x_{i,t-1}^{\bar{c}} + \alpha_{i,t} [\chi_{\bar{c}}(c_{-i,t-1}) - x_{i,t-1}^{\bar{c}}]$$

in cui $x_{i,t}^{\bar{c}}$ è l'elemento del vettore $x_{i,t}$ corrispondente alla strategia congiunta \bar{c} che rappresenta la combinazione delle strategie pure effettivamente giocate da ciascun avversario,

$$\chi_{\bar{c}}(c_{-i,t-1}) \text{ è un indicatore che assume valore 1 se } c_{-i,t-1} = \bar{c}, \text{ mentre } \alpha_{i,t} = \frac{1}{(q+t+1)}, q > 0.$$

Questo significa che se la strategia congiunta \bar{c} viene in effetti giocata al tempo $t-1$, la probabilità che al tempo t il giocatore i le attribuirà sarà maggiore, mentre se non viene giocata le verrà associata una probabilità minore. Il ritmo di questo aggiustamento tenderà a diminuire al crescere di t , dando luogo al fenomeno di *turn off* che si era già notato per i minimi quadrati: anche in questo caso si suppone che gli agenti reagiscano come se il comportamento degli avversari fosse stazionario.

Un secondo problema dell'approccio bayesiano è che in generale il risultato del processo di apprendimento dipende fortemente dalle aspettative a priori del soggetto che però sono estremamente difficili da determinare. È per questo motivo che Fine (1973) afferma che “le statistiche bayesiane possono procurare un falso senso di sicurezza. Non è né razionale né saggio forzare nella forma ingannevolmente dettagliata di una distribuzione le poche briciole di informazione che possiamo avere”.

Oltre a questi problemi “tecnici” ce n'è un altro che viene sottolineato da molti autori², ed è la mancanza di riscontro empirico: alcuni esperimenti di Kahneman e Tversky ad esempio, mostrano che, dovendo valutare la probabilità che un evento appartenga o no a una data categoria, le persone tendono a far riferimento a quanto l'evento è rappresentativo della categoria trascurando le probabilità a priori anche quando esse sono esplicitate chiaramente, in violazione palese della regola di Bayes. Altri esperimenti mostrano che le valutazioni intuitive non tengono conto della dimensione del campione, che invece dovrebbe avere un ruolo cruciale nella determinazione delle probabilità a posteriori.

La portata descrittiva della regola di Bayes sembra dunque essere piuttosto limitata, non è detto però che non possa aumentare in futuro. Infatti, siccome l'approccio bayesiano rimane prevalente nell'ambito della teoria delle decisioni e dei processi cognitivi, non è da escludere che si verifichino fenomeni di “evoluzione culturale”, cioè che questo approccio venga insegnato ai futuri dirigenti e questi lo interiorizzino e di seguito lo applichino realmente nel prendere decisioni.

² Salmon, 1995; Selten, 1991; Tversky e Kahneman, 1974.

Bisogna inoltre riconoscere alla teoria bayesiana una notevole importanza a livello normativo-ideale, cioè come strumento per chiarire come sarebbe razionale agire se non esistessero i limiti che caratterizzano gli agenti reali. Tuttavia, dal momento che qui si intende esaminare proprio l'impatto di questi limiti sull'esito delle decisioni, si è preferito non adottare quest'approccio nel modello che verrà sviluppato più oltre.

3. Il modello di apprendimento adattivo evolutivo

Alla luce di queste considerazioni sui modelli di apprendimento più comunemente adottati, abbiamo deciso di studiare un modello alternativo, che imponesse ipotesi minime sull'informazione iniziale degli agenti e fosse al riparo dalle principali critiche generalmente mosse al modello classico di razionalità. L'algoritmo prescelto a questo scopo appartiene alla categoria degli algoritmi di apprendimento adattivo evolutivo, sviluppata da Marimon e McGrattan (1995) a partire da un lavoro di Milgrom e Roberts (1991).

In questo approccio si assume che i giocatori all'inizio del gioco conoscano solo le strategie a propria disposizione e che poi possano osservare e ricordare i *payoff* ottenuti; non conoscono invece le mosse fatte dagli avversari né i loro *payoff*, e non sono nemmeno in grado di attribuire ad essi delle probabilità. Per questo non possono stabilire le proprie mosse attraverso un processo di massimizzazione dell'utilità attesa; devono dunque adottare metodi diversi. In particolare si ipotizza che la scelta di ogni giocatore dipenda da tre fattori, che sono:

- l'adattamento, cioè la tendenza a preferire le strategie che in passato hanno garantito i *payoff* più elevati;
- la sperimentazione, ovvero la propensione dei giocatori a provare tutte le strategie a propria disposizione;
- l'inerzia, che in qualche modo controbilancia la sperimentazione inducendo gli agenti a evitare di cambiare la strategia che hanno adottato nel periodo precedente.

Per la definizione formale di questi tre concetti si rimanda all'articolo citato di Marimon e McGrattan.

Il modello di apprendimento che adotteremo prevede che i giocatori elaborino la strategia mista che giocheranno nel periodo successivo sulla base di informazioni relative unicamente alle strategie pure di cui dispongono, alle strategie che loro stessi hanno giocato nei periodi precedenti ed ai *payoff* ottenuti. Il modello ci è parso particolarmente interessante perché non richiede alcun tipo di conoscenza sulla struttura del gioco, né sul comportamento passato degli altri giocatori o sui *payoff* da loro conseguiti e non richiede che i giocatori abbiano aspettative coerenti sull'altrui comportamento. In particolare, esso ipotizza che ciascun giocatore valuti ciascuna strategia pura $a^i \in A^i$ di cui dispone sulla base dei *payoff* $\pi_i(a^i, a^{-i})$ ottenuti giocando quella particolare

strategia nei periodi precedenti; per fare questo egli attribuisce ad ogni strategia pura una “forza”

$$S_{i,t}(a^i) = \begin{cases} S_{i,t-1}(a^i) - \frac{1}{\eta_{i,t-1}(a^i)} [S_{i,t-1}(a^i) - \pi_i(a^i, a^{-i})] & \text{se } i \text{ gioca } a^i \text{ al tempo } t \\ S_{i,t-1}(a^i) & \text{altrimenti} \end{cases}$$

In pratica, se al tempo t il giocatore i gioca a_i la forza attribuita a questa strategia aumenterà o diminuirà a seconda che il *payoff* ottenuto sia maggiore o minore della forza di a^i al periodo precedente, e tale variazione sarà tanto più sensibile quanto più piccolo è il valore di $\eta_{i,t}(a^i)$, che non è altro che un contatore che rappresenta il numero di volte in cui la strategia a_i è stata giocata dall’ultima volta in cui il giocatore i ha modificato la propria strategia mista.

Più precisamente si ha che: $\eta_{i,0}(a^i) = 1 \forall i \in I, a^i \in A^i$ e, se la strategia mista

$\sigma_t^i = (\sigma_t^i(a_1^i), \dots, \sigma_t^i(a_{m_i}^i))$ rimane invariata ($\sigma_t^i = \sigma_{t-1}^i$),

$$\eta_{i,t}(a^i) = \begin{cases} \eta_{i,t-1}(a^i) + 1 & \text{se } i \text{ gioca } a^i \text{ al tempo } t \\ \eta_{i,t-1}(a^i) & \text{altrimenti} \end{cases}$$

mentre se $\sigma_t^i \neq \sigma_{t-1}^i \Rightarrow \eta_{i,t}(a^i) = 1 \forall a^i \in A^i$. L’azzeramento dei contatori consente di superare il problema del *turn off* che, come si è visto nel primo paragrafo, caratterizza i modelli di apprendimento basati su un approccio bayesiano o sul metodo dei minimi quadrati.

Al termine di ciascun *round* di gioco ogni giocatore aggiorna ciascuna delle m_i componenti della propria strategia mista con probabilità $\rho_{i,t}$; in questo senso, si dice che $(1 - \rho_{i,t})$ rappresenta il livello di inerzia del giocatore i al tempo t . La strategia mista evolve nel modo seguente:

$$\sigma_t^i(a^i) = \begin{cases} \epsilon_{i,t} & \text{se } \tilde{\sigma}_t^i(a_i) \leq \epsilon_{i,t} \\ \frac{\tilde{\sigma}_t^i(a^i)}{\sum_{\tilde{\sigma}_t^i(\hat{a}^i) > \epsilon_{i,t}} \tilde{\sigma}_t^i(\hat{a}^i)} (1 - \tilde{\epsilon}_{i,t}) & \text{altrimenti} \end{cases}$$

dove $\tilde{\epsilon}_{i,t} = \epsilon_{i,t} \times |\{\sigma_t^i(\hat{a}^i) < \epsilon_{i,t}\}|^3$ in cui $\epsilon_{i,t} \in (0, 1) \forall i, t$ rappresenta la propensione alla sperimentazione del giocatore i nel periodo t , ovvero la probabilità minima che ogni strategia pura venga giocata, mentre⁴

3 Secondo una convenzione comunemente in uso, con $|E|$ si indica la cardinalità dell’insieme E .

4 Il modello originale di Marimon e McGrattan prevede una formula diversa:

$$\tilde{\sigma}_t^i(a_i) = \begin{cases} \sigma_{t-1}^i(a^i) \frac{S_{i,t-1}(a^i)}{\sum_{\hat{a}^i} \sigma_{t-1}^i(\hat{a}^i) S_{i,t-1}(\hat{a}^i)} & \text{con probabilità } \rho_{i,t} \\ \sigma_{t-1}^i(a^i) & \text{con probabilità } 1 - \rho_{i,t} \end{cases}$$

che tuttavia non risulta applicabile a giochi che ammettano anche *payoff* negativi. Infatti in questo caso è possibile che il denominatore diventi negativo; ma a questo seguirebbe che la probabilità $\sigma^i(a^i)$ associata ad una strategia pura a^i che abbia forza negativa tenderebbe a crescere, anziché a diminuire.

$$\tilde{\sigma}_t^i(a_i) = \begin{cases} \sigma_{t-1}^i(a^i) \frac{\exp[S_{i,t-1}(a^i)]}{\sum_{\hat{a}^i} \sigma_{t-1}^i(\hat{a}^i) \exp[S_{i,t-1}(\hat{a}^i)]} & \text{con probabilità } \rho_{i,t} \\ \sigma_{t-1}^i(a^i) & \text{con probabilità } 1 - \rho_{i,t} \end{cases}$$

A parole, in questo modello di apprendimento si assume che con l'esperienza gli agenti tendano a giocare con probabilità maggiore le strategie da cui nel passato hanno ottenuto i payoff più elevati, ma si ipotizza anche che tutte le strategie pure mantengano comunque una probabilità positiva – per quanto piccola – di essere giocate; si può pensare che questo elemento rappresenti la propensione alla sperimentazione degli agenti, o anche la possibilità che per errore essi possano scegliere una strategia diversa da quella preferita.

Nel seguito, adotteremo questo modello assumendo che $\epsilon_{i,t} = \epsilon$, $\epsilon \in (0,1)$ \wedge $\rho_{i,t} = \rho$, $\rho \in (0,1) \quad \forall i, t$; in questo modo possiamo dimostrare che esso costituisce una regola di apprendimento adattiva. Infatti il secondo lemma enunciato da Marimon e McGrattan (1995) garantisce che un algoritmo di apprendimento è coerente con l'apprendimento adattivo se

$$(4) \quad \sum_{t=0}^{\infty} \epsilon_{i,t} = +\infty \quad \wedge \quad \sum_{t=0}^{\infty} (1 - \rho_{i,t}) = +\infty$$

condizione evidentemente rispettata nel caso in cui ϵ e ρ siano costanti positive. La dimostrazione del lemma è immediata, se si considera che in questo caso la condizione da soddisfare perché la regola di apprendimento possa essere definita “adattiva” è che per ogni t esista un m tale che, se

$$\bar{a}^i \notin B_i[f_{i,m}(a_t^-, \dots, a_{t+m-1}^-)] \quad \wedge \quad \hat{a}_t^i \in B_i[f_{i,m}(a_t^-, \dots, a_{t+m-1}^-)]$$

dove $B_i[f_{i,m}(a_t^-, \dots, a_{t+m-1}^-)]$ è l'insieme delle risposte ottime rispetto distribuzione di frequenza osservata sulle mosse compiute in passato dall'avversario, allora

$$\rho \frac{\exp[S_{i,t+m}(\bar{a}^i)] \sigma_{i,t+m}(\bar{a}^i)}{\exp[S_{i,t+m}(\hat{a}^i)] \sigma_{i,t+m}(\hat{a}^i)} + (1 - \rho) \frac{\sigma_{i,t+m}^i(\bar{a}^i)}{\sigma_{i,t+m}^i(\hat{a}^i)} < \frac{\sigma_{i,t+m}^i(\bar{a}^i)}{\sigma_{i,t+m}^i(\hat{a}^i)}$$

che, se vale la (4), equivale alla diseguaglianza:

$$S_{i,t+m}(\bar{a}^i) < S_{i,t+m}(\hat{a}^i)$$

la quale è sempre verificata se in un periodo compreso tra t ed m il giocatore i modifica la propria strategia mista.

Nel modello utilizzato per la simulazione, infine, si assumerà che la condizione iniziale sia:

$\sigma_1^i(a^i) = 1/m_i \quad \wedge \quad S_{i,0}(a^i) = 0 \quad \forall i \in I \quad \wedge \quad a^i \in A^i$. Tale ipotesi può significare semplicemente che il giocatore non ha alcun tipo di conoscenza o di aspettative a priori sul gioco.

4. Il modello di simulazione

Il modello di apprendimento prescelto è stato applicato a un gioco simmetrico a due giocatori presentato in forma normale nella tabella 1.

In questo gioco i due soggetti contrapposti debbono accordarsi sulle proporzioni secondo cui dovranno spartire l'utile che potrebbero trarre da un lavoro svolto assieme, che è pari a $(m-1)p$ dove m è il numero di strategie pure

	0	1	2	3	4
0	0, 0	0, 25	0, 50	0, 75	0, 100
1	25, 0	25, 25	25, 50	25, 75	-25, -25
2	50, 0	50, 25	50, 50	-25, -25	-25, -25
3	75, 0	75, 25	-25, -25	-25, -25	-25, -25
4	100, 0	-25, -25	-25, -25	-25, -25	-25, -25

Tabella 1

che appartengono all'insieme di scelta di ciascun giocatore, mentre p è un premio fissato in modo esogeno. Nel caso presentato in tabella si è fissato $m=5$ e $p=25$; in questo modo la somma da spartire è pari a 100. Come si può osservare, il *payoff* del giocatore i quando questi gioca contro il giocatore j è definito dalla formula:

$$\pi_i(a_i, a_j) = \begin{cases} p a_i & \text{se } a_i + a_j \leq m - 1 \\ -p & \text{altrimenti} \end{cases} .$$

Questo significa che se i due raggiungono un accordo, cioè se le quantità richieste da ciascuno sommano a un valore inferiore o uguale a $(m-1)p$, ognuno ottiene quanto domandato; altrimenti l'accordo salta e ciascuno ci rimette un'uguale quantità p , che rappresenta l'impegno e il tempo dedicato a formulare il contratto che poi non viene concluso.

Si assume che in ogni periodo t tutti gli n agenti della popolazione I vengano estratti a sorte ed accoppiati tra loro, e che ciascuno scelga la propria strategia pura $a^i \in A^i$, $A^i = \{0, 1, 2, 3, \dots, m-1\} \forall i$ secondo il modello di apprendimento adattivo evolutivo esposto nel paragrafo 3.

Il modello è stato costruito in questo modo allo scopo di verificare simultaneamente se, dopo un numero elevato di *round* di gioco

- le strategie miste σ_i scelte da ciascun giocatore i tendono ad assumere valori costanti;
- la distribuzione osservata sulle $n/2$ strategie congiunte giocate ad ogni *round* si stabilizza; in particolare ci interessa scoprire se emerge una strategia congiunta (a_i, a_j) che viene giocata con una frequenza spiccatamente superiore alle altre e nel caso verificare se essa coincide con un equilibrio di Nash in strategie pure.

Questo gioco presenta cinque equilibri di Nash in strategie pure: è interessante capire se il modello di apprendimento fin qui sviluppato ed analizzato seleziona uno di questi equilibri oppure produce un ciclo all'interno del quale l'esito predominante oscilla tra due o più di essi; alternativamente potrebbe infatti accadere anche che nessun equilibrio riesca ad emergere e ad imporsi a livello sociale.

L'evoluzione del comportamento individuale e sociale generata dall'applicazione a questo gioco del

modello di apprendimento considerato è troppo complessa per poter essere studiata in modo analitico, ad esempio facendo ricorso alla teoria delle catene markoviane; per questo motivo si è deciso di analizzare il comportamento asintotico dei giocatori ricorrendo ad una simulazione al computer. Prima di descriverla in dettaglio però vorremmo presentare un ultimo ragionamento di tipo analitico sul modello in questione, utile per descrivere alcune caratteristiche che i risultati della simulazione dovrebbero avere.

Se il numero n di giocatori è sufficientemente elevato, la probabilità $\alpha_t^i(a_k)$ che l'avversario del giocatore i giochi a_k al tempo t dovrebbe essere circa uguale per ogni giocatore:

$$\alpha_t^i(a_k) \simeq \alpha_t^j(a_k) \simeq \alpha_t(a_k) = \frac{\sum_i \sigma_t^i(a_k)}{n} \quad \forall i, j \in I \quad .$$

Supponiamo che, come auspicato, nel lungo periodo emerga a livello sociale una situazione di equilibrio; questo dovrebbe implicare che:

$$\text{per } t > \tau \quad \alpha_t = (\alpha_t^i(a_1), \alpha_t^i(a_2), \dots, \alpha_t^i(a_m)) = \bar{\alpha}$$

Secondo le definizioni date da Marimon e McGrattan (1995) una regola di apprendimento è “adattiva”, se data la frequenza relativa alle mosse degli avversari, la strategia di gioco di ciascun giocatore si muove in direzione della risposta ottima. Dal momento in cui viene raggiunto l'equilibrio in poi, tale frequenza dovrebbe tendere asintoticamente ad $\bar{\alpha}$. Ma se $\bar{\alpha}$ è uguale per tutti, allora anche la risposta ottima di ciascun giocatore dovrebbe essere la medesima, dunque

$$\sigma_t^i \rightarrow \bar{\sigma} \quad . \text{ Se poi si ricorda la definizione di } \alpha_t(a_k) \text{ data più sopra si noter\`a che dev'essere anche } \bar{\alpha} = \bar{\sigma} \quad .$$

In conclusione. Questo ragionamento conduce ad aspettarsi che, se si raggiunge una situazione di equilibrio statico, essa sarà caratterizzata da una strategia mista comune a tutti i giocatori e tale da costituire una risposta ottima a se stessa, ovvero:

$$\bar{\sigma} \text{ t.c. } E_{\bar{\sigma}, \bar{\sigma}}(\pi) \geq E_{\hat{\sigma}, \bar{\sigma}}(\pi) \quad \forall \hat{\sigma} \neq \bar{\sigma} \quad .$$

A questo punto procediamo alla descrizione del modello di simulazione vero e proprio. Questo in pratica riproduce il gioco di *bargaining* presentato più sopra, e il modello di apprendimento prescelto. Permette tuttavia di impostare, ad ogni nuova simulazione, il numero di giocatori n , il numero di *round* di cui la simulazione dev'essere composta, e i parametri ϵ e ρ . Inoltre è possibile anche modificare il gioco vero e proprio, scegliendo il numero di strategie m a disposizione dei giocatori e il premio p che determina i diversi *payoff*. Facendo variare uno alla volta questi parametri è possibile indagare come ciascuno di essi influenzi l'esito di lungo periodo del gioco.

Nel modello di cui ci si sta occupando la formazione delle coppie, il processo con cui in ogni periodo ciascun giocatore sceglie la strategia pura che adotterà e il meccanismo di aggiornamento

delle probabilità sono determinati dall'intervento del caso. All'interno della simulazione questo viene riprodotto utilizzando un generatore di numeri casuali inizializzato con un seme definito in modo arbitrario prima che inizi la simulazione vera e propria. Dato il seme, la sequenza dei numeri casuali prodotti dal generatore non cambia; questo elemento è importante, perché permette di studiare l'impatto di una variazione in uno dei parametri menzionati prima $(m, n, p, \epsilon, \rho)$ sull'esito del processo, isolandola dall'intervento del caso.

5. i risultati

In questa simulazione si parte da una situazione molto semplificata, assumendo che i giocatori siano tutti perfettamente identici, che abbiano cioè il medesimo modo di apprendere dalle proprie esperienze e che all'inizio non dispongano di alcuna informazione né sul gioco né sul proprio avversario.

Nel seguito di questo paragrafo verranno presentati in modo più approfondito i risultati della simulazione, relativamente al comportamento del singolo giocatore, al comportamento "medio" di tutti i giocatori e ad alcuni valori aggregati che ci sono sembrati significativi, come la somma delle vincite di tutti i giocatori in ciascun *round*, il numero di coppie da cui ogni strategia congiunta viene giocata, e la frequenza con cui viene adottata ciascuna strategia pura. Prima di tutto si prenderà in considerazione un caso "base", in cui tutti i parametri sono fissati in modo arbitrario; questo sarà poi assunto come punto di riferimento per osservare uno ad uno i cambiamenti che conseguono alla variazione di ogni singolo parametro.

5.1 Il caso di riferimento

Nell'analisi svolta sulle dinamiche che scaturiscono dalla ripetizione del gioco, abbiamo scelto di assumere una situazione di riferimento caratterizzata da una popolazione composta da 100 giocatori, che dispongono di cinque strategie pure ed ottengono *payoff* determinati da un premio p pari a 25 (come nel gioco in forma normale presentato al paragrafo 4). La probabilità di errore ϵ è fissata all'1% e il livello di inerzia, $1 - \rho$, è posto uguale a 0,8.

Si ricorda poi che all'inizio del gioco, tutti i giocatori hanno la medesima strategia mista:

$$\sigma_0^i = \left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5} \right)$$

Questi valori sono evidentemente molto arbitrari; proprio per questo si ritiene importante verificare che nel lungo periodo cambiamenti anche molto sensibili di questi parametri non influenzino eccessivamente l'esito del gioco.

Per prima cosa, si osserverà come nel tempo cambi il modo di giocare del giocatore 1, preso come riferimento (l'osservazione degli altri giocatori conduce alle medesime conclusioni e ci è

parsa superflua). Il grafico a punti riportato in Figura 1 presenta le strategie pure effettivamente scelte dal giocatore in ciascun *round* di gioco.

Si nota che durante i primi round il giocatore ha modo di sperimentare tutte e cinque le

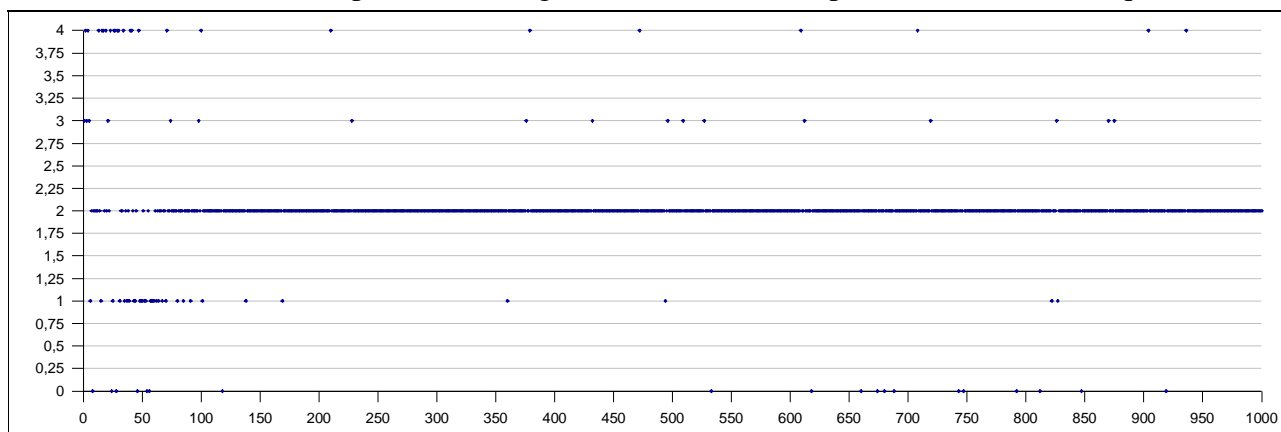


Figura 1: strategie pure effettivamente giocate

strategie a sua disposizione; attraverso questa esperienza, valuta la “forza” di ciascuna strategia (figura 2) e sulla base di questo seleziona progressivamente la strategia che si rivela più proficua.

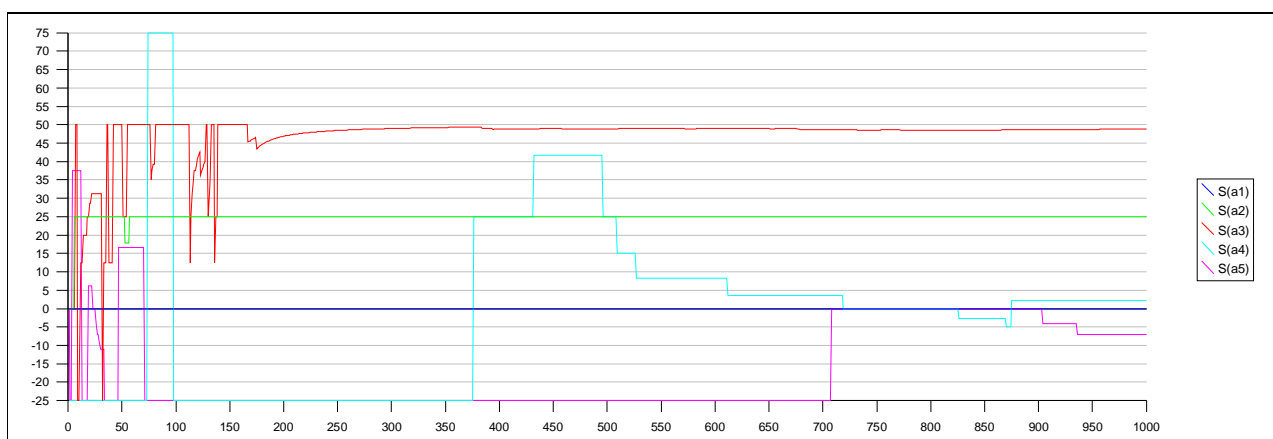


figura 2: forze attribuite alle diverse strategie

Come si può vedere, anche dopo un numero elevato di partite, le forze possono subire delle modifiche notevoli perché tutte le strategie continuano ad essere occasionalmente giocate; è questo elemento che consente al giocatore di rivedere le proprie scelte strategiche, nel caso in cui ad esempio cambi il contesto di gioco.

Consideriamo ora l’evoluzione della strategia mista del giocatore 1, σ_t^1 . Il grafico presentato in figura 3 presenta l’andamento delle sue diverse componenti $\sigma_t^1(a_k)$ rispetto al numero di partite giocate.

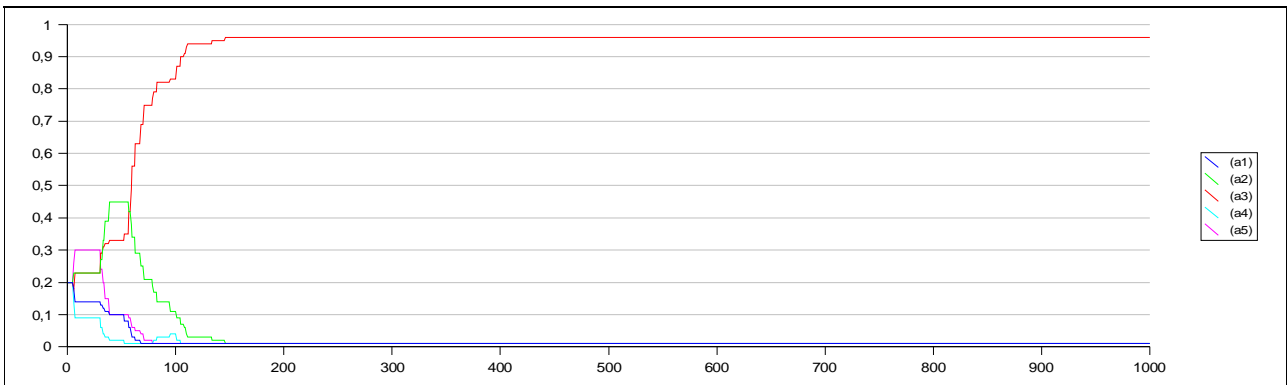


Figura 3: evoluzione della strategia mista del giocatore 1

Oltre ai valori relativi al singolo individuo ci è sembrato importante considerare anche i valori medi valutati sull'intera "popolazione" dei cento giocatori introdotti nella simulazione. Si è osservata quindi la probabilità media che ciascuna strategia pura ha di essere giocata in ciascun periodo

$\alpha_t(a_k)$, presentata nella Figura 4:

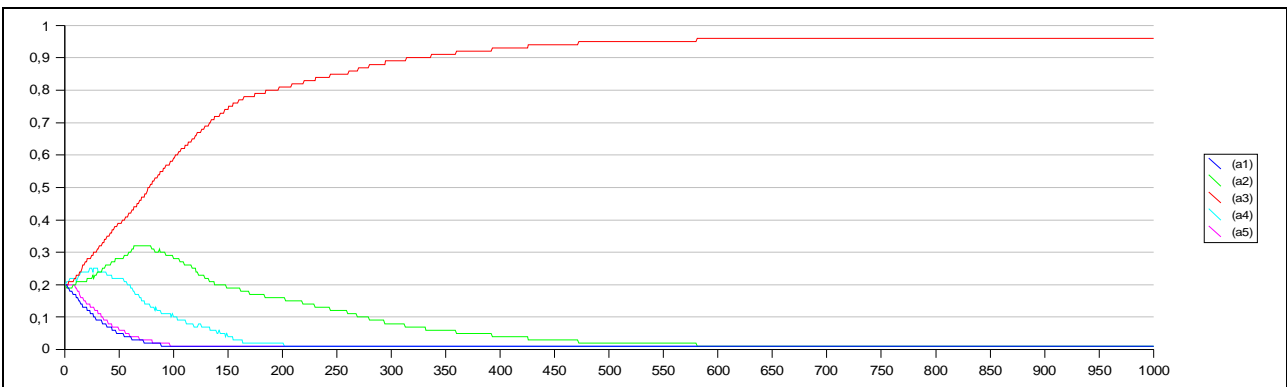


Figura 4: probabilità medie

Questo grafico mostra che, come previsto, nel lungo periodo le probabilità medie assumono il medesimo valore delle probabilità assegnate alle diverse strategie pure dai singoli giocatori, perché queste diventano tutte eguali fra loro. La strategia 3, che prevede la richiesta della metà esatta della quantità da spartire, prevale nettamente a livello sociale e conseguentemente diventa sempre più conveniente. Infatti, il *payoff* ottenuto in media giocando la strategia a_3 :

$$\bar{\pi}_t = \frac{\sum_i \pi_t^i(a_3)}{n_t(a_3)} \quad \text{dove } n_t(a_3) \text{ rappresenta il numero di giocatori che in } t \text{ hanno giocato la strategia } a_3,$$

crece fino a quasi a raggiungere il proprio valore massimo, ovvero 50. (si veda la Figura 5).

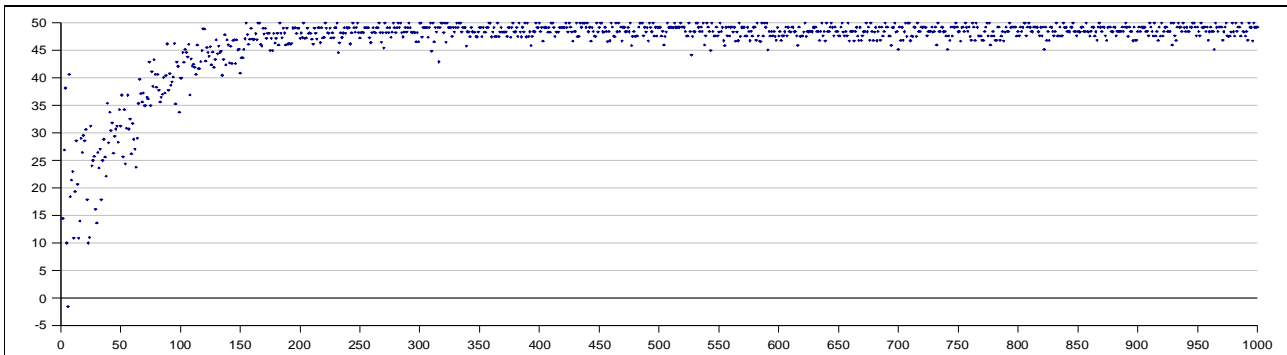


Figura 5: evoluzione del *payoff* medio della strategia a_3

Questo è possibile perché l'esito costituito dalla strategia congiunta (a_3^i, a_3^i) si impone come "convenzione sociale", intendendo con questo termine un concetto molto vicino a quello usato da Young (1993 b).

Da ultimo, ci sembra interessante mostrare come tenda ad aumentare notevolmente anche il "guadagno" complessivo, a livello aggregato $\Pi_t = \sum_i \pi_t^i$, rappresentato nella Figura 6.

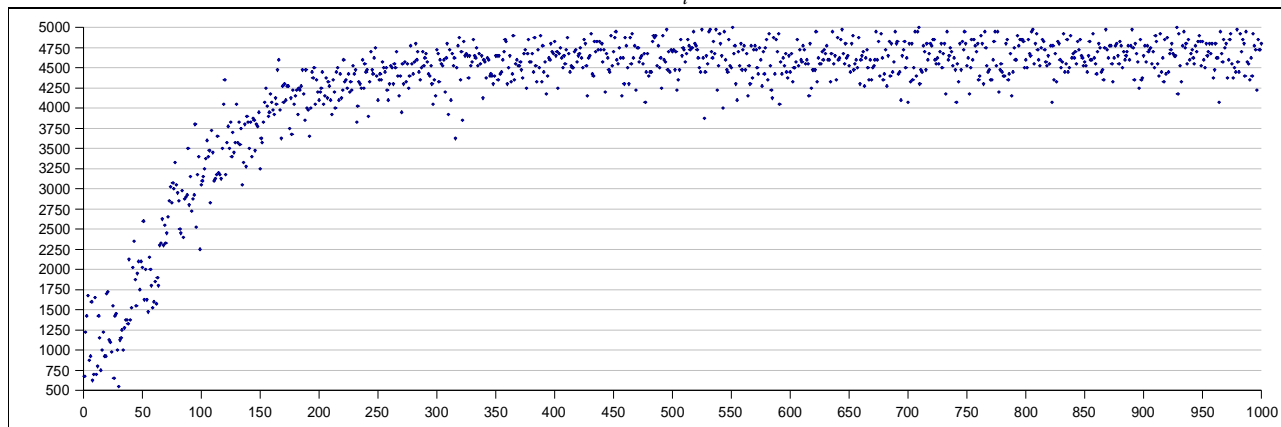


Figura 6: somma dei punteggi realizzati da tutti i 100 giocatori, nel corso dei diversi round di gioco

Alla luce di questo elemento, sembra possibile affermare che l'esito del gioco nel lungo periodo si avvicina anche a una situazione di ottimalità sociale, poiché, nel caso descritto, la vincita complessiva massima è pari a 5000, ed il livello di Π_t raggiunto nel lungo periodo si avvicina molto a questo valore.

I risultati esposti fino a questo punto sembrano confortanti: l'algoritmo di apprendimento adottato conduce tutti i giocatori a selezionare una strategia da cui traggono un profitto quasi costante e mediamente elevato, e a livello sociale emerge come convenzione uno dei cinque equilibri di Nash – che sono tutti socialmente ottimali nel senso esposto prima – ed in particolare quello che prevede la spartizione più equa delle vincite tra i vari giocatori.

Per verificare che questo risultato non sia conseguenza solo della scelta arbitraria dei parametri, e per capire meglio quali effetti questi parametri abbiano sull'esito del gioco, verranno ora presentati alcuni risultati interessanti, ottenuti variando uno ad uno i valori attribuiti a ciascuno di essi.

5.2 Variazioni sul valore di ϵ

Come si è già chiarito in precedenza, ϵ rappresenta la propensione alla sperimentazione, o all'errore, che si suppone sia presente all'interno del processo di apprendimento. Osservando le figure 5 e 6 si nota che, anche dopo il raggiungimento dell'equilibrio, permane una certa variabilità nei *payoff*. Questo è dovuto al fatto che, con un ϵ pari all'1%, ogni giocatore devierà dalla strategia a_3 con probabilità del 4%. Riducendo il valore di ϵ , ad esempio, fino a 0.001, si ottengono in effetti valori molto più concentrati e il *payoff* medio risulta molto vicino al massimo

raggiungibile.

Se ϵ assume un valore così basso tuttavia alcuni giocatori possono impiegare molto tempo per individuare la strategia migliore, dunque la probabilità che in media vengano giocate strategie diverse (nella fattispecie, la strategia a_2) rimane a lungo diversa da ϵ (figura 7).

Occorrono quasi 5000 partite perché si raggiunga l'equilibrio.

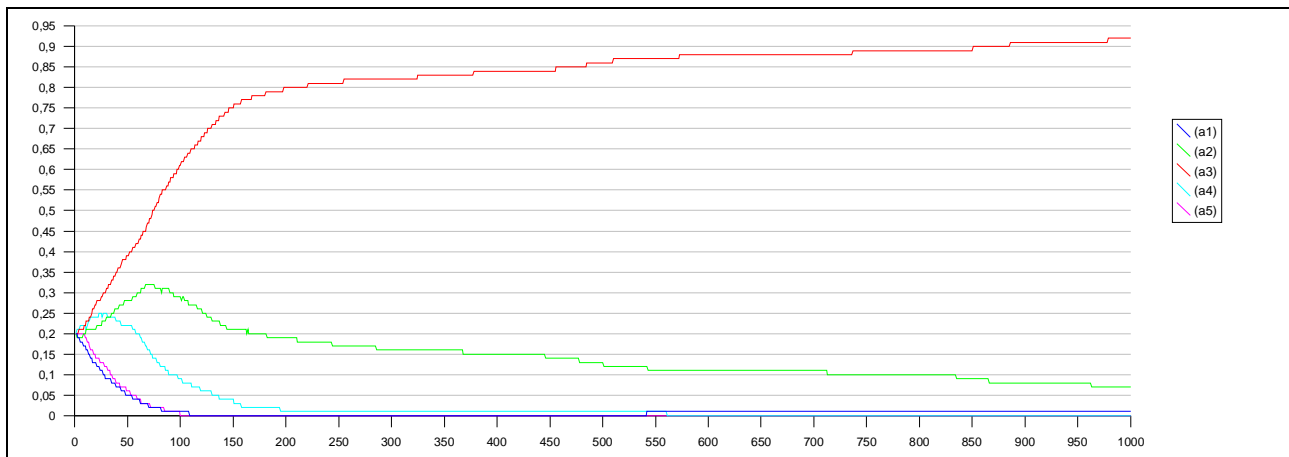


Figura 7: andamento delle probabilità medie con ϵ pari a 0,001

Si è affermato in precedenza che se ϵ è pari a 0 il modello di apprendimento non può più essere

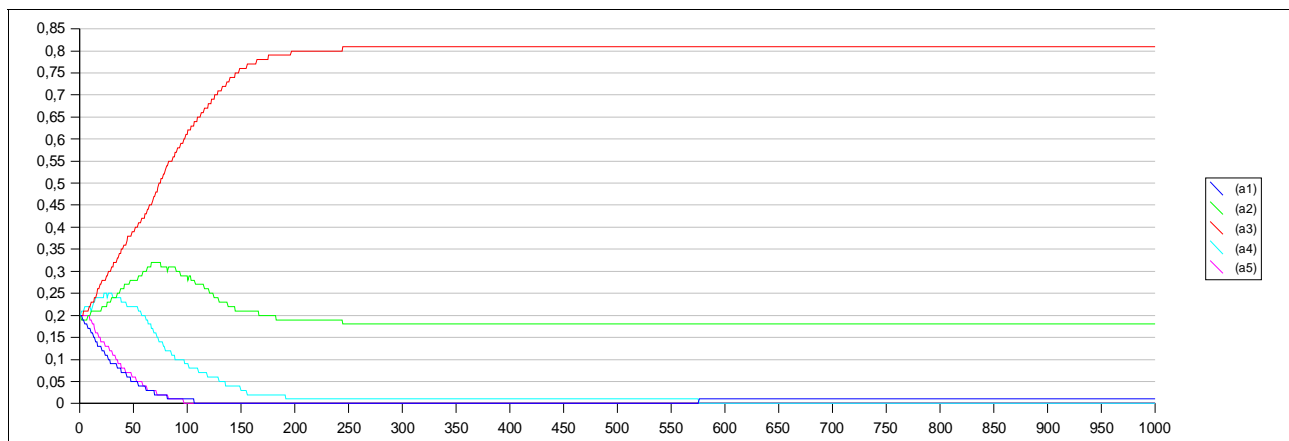


Figura 8: probabilità medie, con ϵ pari a 0

definito adattivo: a riprova di questo presentiamo ora il grafico (Figura 8) che rappresenta l'andamento delle probabilità medie sotto questa ipotesi. Sono necessarie meno di 300 partite per raggiungere l'equilibrio, che però è costituito da una situazione in cui la strategia a_2 viene giocata con probabilità pari al 18%.

Osservando l'evoluzione delle strategie miste dei giocatori, si comprende la causa di questa situazione: il processo di apprendimento può condurre alcuni giocatori ad attribuire alla strategia a_2 una probabilità pari a 1 (figura 9); questa scelta risulta strettamente sub-ottimale in una situazione in cui le strategie a_4 e a_5 sicuramente non verranno mai giocate, poiché la strategia a_3 garantisce un *payoff* certo pari a 50, doppio rispetto a quello che si ottiene giocando a_2 . Tuttavia, non sperimentando mai strategie alternative, i giocatori non hanno modo "accorgersene" e di adeguare conseguentemente la propria strategia mista al contesto di gioco.

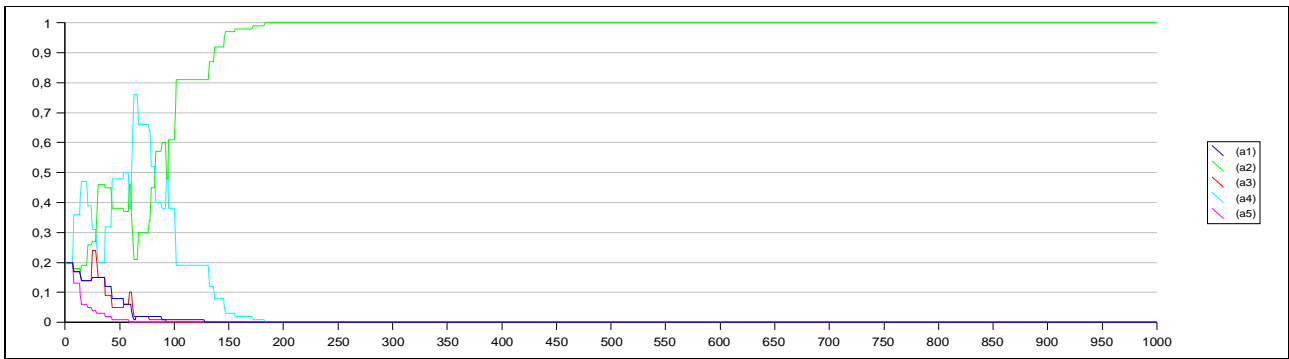


Figura 9: evoluzione della strategia mista di uno dei giocatori, nel caso in cui ϵ sia nullo.

D'altra parte, se ϵ assume valori troppo elevati non può emergere alcuna convenzione sociale. Assumiamo ad esempio che per essere considerata "convenzione sociale" una strategia congiunta (a_k, a_k) debba verificarsi con probabilità superiore al 70%; perché emerga una convenzione, bisognerà allora che la probabilità massima che la strategia pura a_k venga giocata sia $[1 - \epsilon(m-1)] > \sqrt{0,7}$, da cui risulta che:

$$\epsilon < \left(\frac{1 - \sqrt{0,7}}{m-1} \right) \text{ ovvero, nel caso considerato, } \epsilon < 0.0408 .$$

Dunque, affinché il modello di apprendimento conduca ai risultati sperati bisogna imporre delle condizioni su ϵ limitandone le possibili variazioni al di sotto di un valore tanto più piccolo quante più sono le strategie pure disponibili.

5.3 Variazioni sul valore di ρ

Il parametro ρ rappresenta la propensione al cambiamento dei giocatori: più il suo valore è elevato, più è probabile che ciascun agente riveda la propria strategia mista e la aggiorni sulla base delle proprie esperienze. Secondo Marimon e McGrattan "l'inerzia è una forma particolare di ciò che chiamiamo 'apprendimento correlato in modo imperfetto'. Questa caratteristica degli algoritmi di apprendimento produce una forma debole di stazionarietà nell'ambiente. Se i giocatori modificano costantemente le loro strategie, nessun giocatore riesce a effettuare dei "test" per trovare una strategia che porti buoni risultati. L'inerzia genera stazionarietà garantendo che, con probabilità positiva, la strategia (mista) dell'avversario di i rimarrà fissa nel periodo immediatamente successivo a una revisione della strategia di i con probabilità positiva."

Dal momento che questa affermazione di Marimon e McGrattan si riferisce ad un gioco in cui si affrontano ripetutamente sempre gli stessi due giocatori, si è ritenuto opportuno, innanzitutto, verificare l'impatto di modifiche a ρ in questa situazione (si pone $n=2$, mentre tutti gli altri parametri vengono tenuti costanti al valore di riferimento). I risultati ottenuti da questa prova contrastano con quanto affermato dai due autori: infatti, ponendo $\rho=1$, si raggiunge l'equilibrio di Nash (a_3, a_3) come nel caso in cui $\rho=0,2$ ma impiegando un numero molto minore di partite.

Modificando il seme del generatore di numeri casuali, però, nel caso a due giocatori si

ottengono risultati anche piuttosto diversi tra loro (nel caso in cui i giocatori siano in numero maggiore, vedremo, questo non avviene). In particolare, abbiamo osservato che in ogni caso si raggiunge un equilibrio di Nash, ed il numero di partite necessario per conseguire questo risultato comunque diminuisce al crescere di ρ ; tuttavia non sempre l'equilibrio di Nash che viene raggiunto è costituito dall'esito (a_3, a_3) .

Queste osservazioni sono relative al caso in cui due soli giocatori si affrontano ripetutamente tra loro, che è piuttosto diverso da quello fin qui considerato, nel quale 100 giocatori formano di volta in volta delle coppie assortite in modo casuale. In effetti nel nostro caso la simulazione mostra che i risultati di lungo periodo, specie per quel che riguarda le variabili relative all'intera popolazione – come la probabilità media che ogni strategia venga giocata, il numero di coppie che giocano la strategia congiunta (a_3, a_3) o la somma delle vincite Π_t – sono influenzati in modo non significativo dal valore di ρ . Consideriamo ad esempio l'evoluzione delle probabilità medie nel caso in cui sia $\rho=0.8$ (figura 10): confrontandolo col grafico in figura 4, si nota che la differenza è minima.

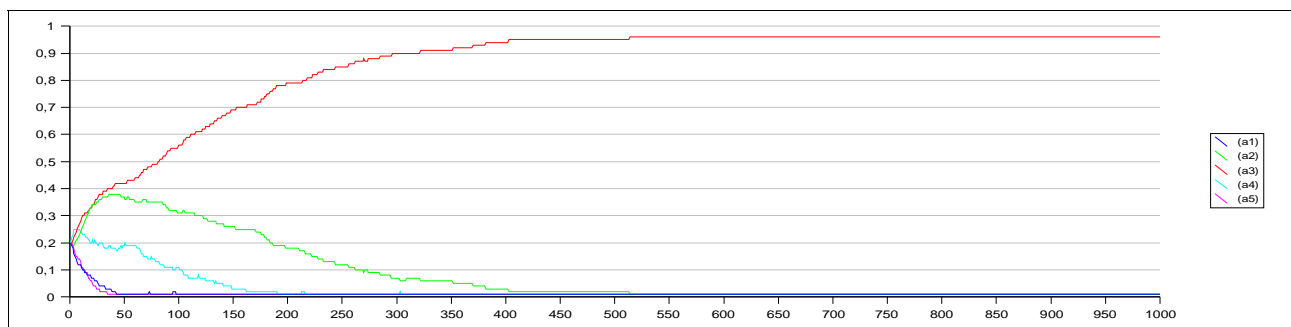


Figura 10: probabilità medie nel caso in cui $\rho=0,8$

Le variazioni di ρ influenzano in modo più marcato le strategie adottate dai singoli giocatori durante le partite giocate prima del raggiungimento dell'equilibrio, pur senza modificare quest'ultimo: infatti, più elevato è ρ , più è frequente che ciascun giocatore modifichi la propria strategia mista, dunque le oscillazioni nelle probabilità che egli attribuisce a ciascuna strategia pura sono di maggiore entità (figura 11).

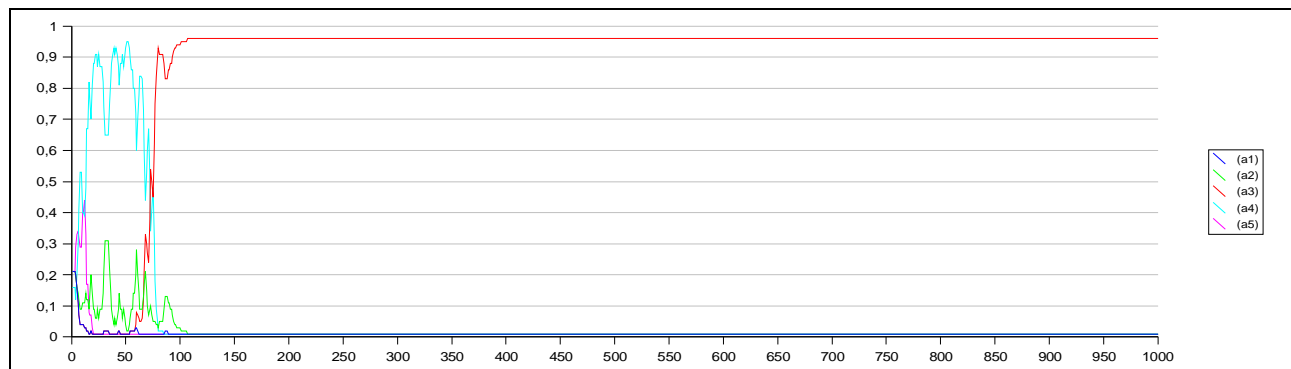


Figura 11: evoluzione della strategia mista del giocatore 1 nel caso in cui $\rho=0,8$

5.4 Variazioni sul numero di giocatori n

Si è già notato in precedenza come l'ipotesi che esistano due soli giocatori produca risultati diversi da quelli ottenuti fissando $n=100$; tuttavia effettuando numerose prove si è potuto concludere che si tratta di un caso isolato: se n è superiore a 2 ed è un numero pari e se tutti gli altri parametri esogeni mantengono i valori attribuiti loro nel caso base nel lungo periodo emerge sempre come convenzione sociale l'equilibrio (a_3, a_3) , e le variazioni nel numero di giocatori non sembrano avere conseguenze importanti né sulle variabili relative ai singoli giocatori (forza delle strategie pure, strategia mista) né su quelle relative all'intera popolazione. Di quest'ultima categoria di variabili, consideriamo ad esempio le probabilità medie $\alpha(a_k)$: le simulazioni effettuate modificando ripetutamente il valore di n hanno mostrato che la variazione di questo parametro non influenza in modo significativo l'evoluzione delle probabilità medie relative alle diverse strategie pure di gioco: mentre il numero di partite necessarie a raggiungere l'equilibrio è tanto minore quanti di meno sono i giocatori, l'andamento relativo a ciascuna strategia non cambia.

5.5 Variazioni sul numero di strategie, m

L'ultima prova effettuata per verificare la solidità dei risultati ottenuti è volta a controllare che, anche variando il numero di strategie pure di cui ciascun giocatore dispone, il risultato non cambia, ovvero prevale comunque l'esito nel quale i giocatori si spartiscono la vincita equamente.

La prova è stata fatta con $m=6$ ed $m=11$; il premio p è stato posto rispettivamente pari a 20 e a 10, in modo che la vincita massima rimanesse comunque pari a 100.

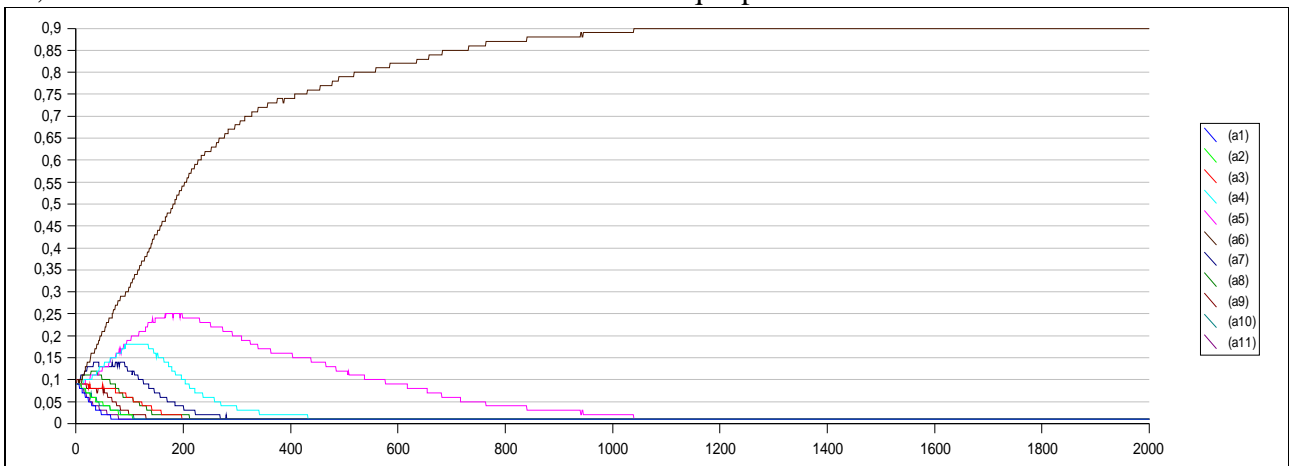


Figura 12: probabilità medie nel caso in cui $m=11$

I risultati ottenuti ponendo $m=11$ sono molto simili a quelli ottenuti nel caso base: la figura 12 mostra come nel lungo periodo prevalga la strategia pura a_6 che corrisponde alla richiesta di metà esatta della somma da spartire, anche se è necessario un numero molto maggiore di partite per raggiungere l'equilibrio.

Il caso in cui, invece, $m=6$ produce risultati piuttosto diversi, e rappresenta un buon esempio di cosa accada in generale quando il numero di strategie è pari, cioè quando non esiste una strategia

pura cui corrisponda la richiesta di metà esatta del totale da spartire. Il grafico in figura 13 mostra che in questa situazione non si riesce a raggiungere un vero e proprio equilibrio statico: la strategia che viene giocata con maggior probabilità è la strategia a_3 – che in questo caso corrisponde alla richiesta del 40% della somma totale – ma anche la a_4 (richiesta del 60% del totale) viene giocata con una probabilità che oscilla poco al di sotto del 30%. Diversamente da quanto accade negli altri casi sin qui osservati, le probabilità medie associate a queste due strategie non raggiungono mai un valore costante ma continuano ad oscillare, anche se entro un range limitato.

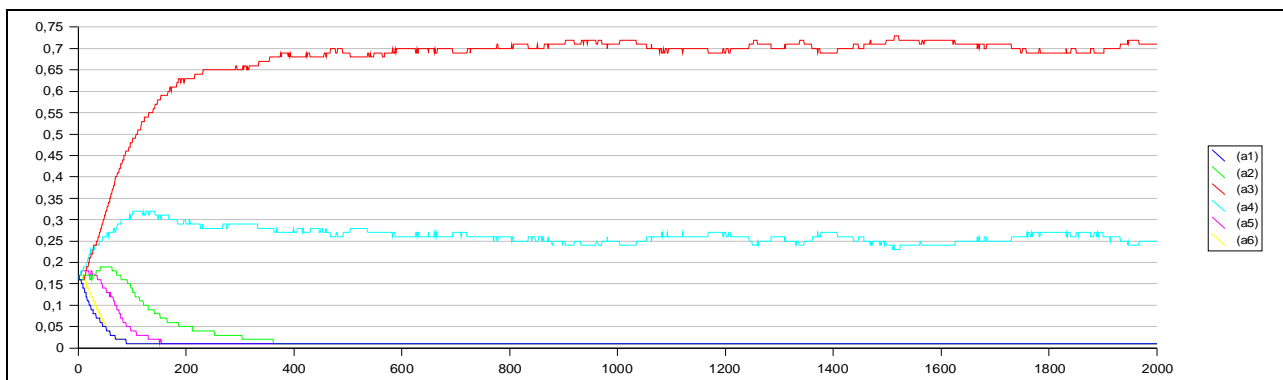


Figura 13: probabilità medie nel caso in cui $m=6$

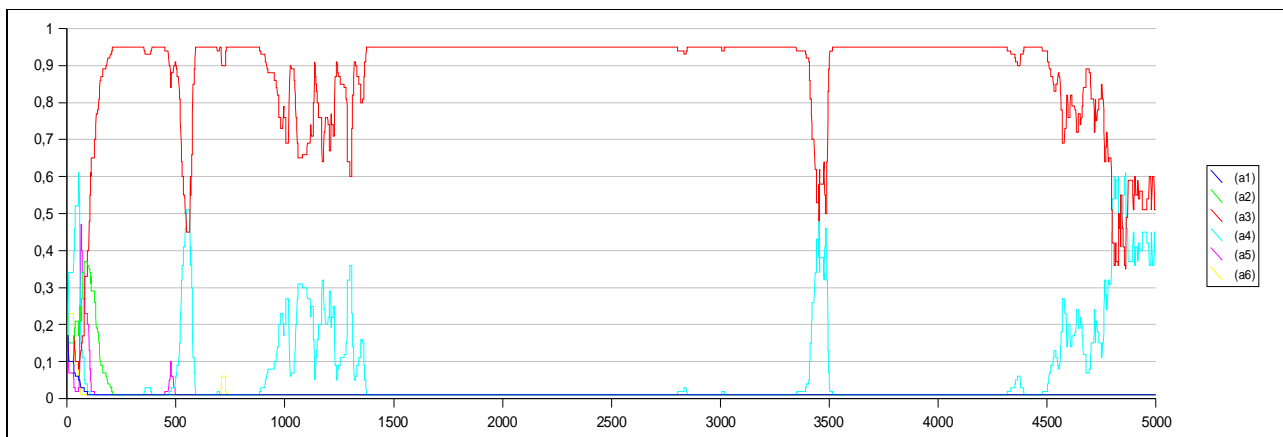


Figura 14: evoluzione della strategia mista di un giocatore, nel caso in cui $m=6$

La ragione di questo andamento delle probabilità medie si comprende meglio osservando come evolve la strategia mista di uno dei giocatori (figura 14): si nota infatti che anche dopo un numero elevato di partite il giocatore non riesce a definire stabilmente la propria strategia mista ma la modifica ripetutamente quando accade che, grazie alla tendenza alla sperimentazione, egli devia dalla strategia preferita a_3 e sceglie invece a_4 da cui riesce, occasionalmente, ad ottenere un punteggio maggiore; questo provoca un incremento della forza $S^i(a_4)$ che a sua volta genera un aumento della probabilità $\sigma^i(a_4)$.

Anche questo caso comunque testimonia come il meccanismo di apprendimento conduca i giocatori a prediligere una spartizione il più possibile equa della vincita totale, quindi non costituisce una reale smentita ai risultati cui si giunge attraverso l'esame del caso di riferimento.

5. Conclusioni

Tutte le prove di simulazione svolte consentono di affermare che il modello di comportamento studiato ed applicato in questa tesi conduce ciascun giocatore ad individuare una strategia pura che egli giudica nettamente più conveniente rispetto alle altre e che giocherà quindi con maggiore frequenza; è poi sufficiente che il numero di strategie pure m sia dispari affinché ogni giocatore possa stabilire, dopo un numero più o meno elevato di partite, una strategia mista da mantenere costante nel tempo. Se poi la tendenza alla sperimentazione – rappresentata dal parametro ϵ – è ragionevolmente bassa, si è visto che emerge come convenzione sociale l'equilibrio di Nash in cui si ha la spartizione della vincita in parti uguali, che costituisce un esito molto vicino all'ottimo sociale sia dal punto di vista dell'efficienza che da quello dell'equità.

Con questo lavoro si è voluto mostrare che la rimozione di alcune delle ipotesi più forti relative all'informazione a priori e alle capacità computazionali degli agenti non impedisce di formulare modelli di scelta in condizioni di incertezza e di interazione strategica tra agenti economici. Si è visto inoltre che questi modelli, nonostante le radicali differenze rispetto a quelli tradizionali, portano comunque ad una conferma – piuttosto che ad un rifiuto – di alcuni risultati fondamentali della teoria tradizionale, come la convergenza all'equilibrio di Nash e la massimizzazione del benessere sociale, inteso come somma della ricchezza di tutti gli individui.

Sarebbe interessante proseguire il lavoro di ricerca svolgendo esperimenti empirici con soggetti reali, per ottenere stime più ragionate dei parametri inseriti nel modello studiato e per confrontare le previsioni teoriche con i risultati che si ottengono in realtà.

BIBLIOGRAFIA

- R. Aumann; A. Brandenburger (1995): "Epistemic conditions for Nash equilibrium", *Econometrica*, 63, pp. 1161-1180.
- K. Binmore (1987), "Modeling rational players", *Economics and philosophy*, 3, pp. 179-214
- D. Canning (1995): "Learning and social equilibrium in large populations", in *Learning and rationality in economics*, a cura di A. P. Kirman e M. Salmon, Blackwell, Oxford.
- R.P. Cubitt; R. Sudgen (1994): "Rational justifiable play and the theory of non-cooperative games", *The Economic Journal*, 104, pp.798-803.
- T.L. Fine (1973); "Theories of probability: an examination of foundations", Academic Press, New York.
- R. Marimon; E. McGrattan (1995): "On adaptive learning in strategic games", in *Learning and rationality in economics*, a cura di A. P. Kirman e M. Salmon, Blackwell, Oxford.
- P. Milgrom; J. Roberts (1991): "Adaptive and sophisticated learning in normal form games", *Games and Economic Behavior*, 3, pp.82-100.
- P. Milgrom; N. Stokey (1982): "Information, trade, and common knowledge", *Journal of Economic Theory*, 26, pp. 17-27.
- M. Salmon (1995): "Bounded rationality and learning; procedural learning", in *Learning and rationality in economics*, a cura di A. P. Kirman e M. Salmon, Blackwell, Oxford.
- R. Selten (1991): "Evolution, learning, and economic behavior", *Games and Economic Behavior*, 3, pp.3-24.
- H. Simon (1955): "A behavioral model of rational choice", *The Quarterly Journal of Economics*, 69, pp. 99-118.
- A. Tversky; D. Kahneman (1974): "Judgement under uncertainty: heuristics and biases", *Science*, 185, pp.1124-1131
- A. Tversky; D. Kahneman (1979): "Prospect theory: an analysis of decision under risk", *Econometrica*, 47, pp.263-292.
- A. Tversky; D. Kahneman (1981): "The framing of decisions and the psychology of choice", *Science*, 211, pp.453-458.
- H. P. Young (1993): "The evolution of conventions", *Econometrica*, 61, pp. 57-84.