

MPRA

Munich Personal RePEc Archive

Incomplete Contracts and Endogenous Enforcement. A Survey

Scoppa, Vincenzo

2003

Online at <https://mpra.ub.uni-muenchen.de/17284/>
MPRA Paper No. 17284, posted 09 Oct 2009 13:36 UTC

CONTRATTI INCOMPLETI ED ENFORCEMENT ENDOGENO. UNA RASSEGNA DELLA LETTERATURA

*Working Paper Version – pubblicato in *Economia Politica*, Anno XX, n. 3, dicembre 2003*

Vincenzo Scoppa *

Dipartimento di Economia e Statistica, Università della Calabria

“It is the effectiveness of the enforcement of agreements that is the single most crucial determinant of economic performance and the key difference that separates First from Third World economies. The ability to enforce contracts across time and space is the central underpinning of ‘efficient’ market” (North, 1993, p. 247).

1. Introduzione

Qualsiasi tipo di relazione economica che intercorre tra due o più parti è regolata da un accordo o contratto, che può essere formale o informale, stabilito tra le parti stesse. Affinché i soggetti economici accettino di entrare in relazione e realizzare transazioni mutuamente vantaggiose è necessaria l'esistenza di un meccanismo di enforcement, che garantisca il rispetto e l'applicazione dei contratti sottoscritti.

Nella teoria dell'equilibrio economico generale, il compito dell'enforcement è demandato al sistema legale e giudiziario (la “Corte”); per rendere esecutivi i contratti, la Corte deve disporre di un contratto ben definito in ogni aspetto, deve essere in grado di osservare ogni circostanza rilevante e di imporre sanzioni adeguate che disincentivino comportamenti difforni dagli accordi. In questa struttura, l'enforcement viene definito esogeno poiché è realizzato ad opera di una autorità esterna e non richiede una particolare attività dei contraenti. Tutti gli scambi efficienti possono essere effettivamente realizzati.

Tuttavia, come è stato messo in luce dal recente sviluppo delle teorie sui contratti incompleti, ogniqualvolta l'accordo non regola alcune circostanze – perché non previste o perché le parti hanno preferito lasciarle fuori dal contratto – oppure risulta formulato in modo ambiguo, l'autorità non può intervenire per imporre l'applicazione del contratto. D'altra parte, un contratto può essere applicato solo se è condizionato a informazioni che entrambe le parti possono osservare (osservabilità) e che la Corte può verificare (verificabilità): il ricorso all'autorità sarà quindi impossibile se le variabili rilevanti non sono osservabili e verificabili oppure se il costo per farlo è eccessivamente elevato.

Mentre alcune transazioni economiche possono essere realizzate come scambi spot – senza particolari difficoltà di esecuzione dei contratti – per molte altre la distanza nel tempo o nello spazio tra il quid e il quo o l'esistenza di asimmetria informativa sulle caratteristiche del bene o servizio scambiato generano problemi di opportunismo che mettono a rischio l'adempimento contrattuale e la realizzazione stessa della transazione. I mercati del lavoro, dei beni e del credito sono afflitti da numerosi e seri problemi di enforcement: “in many exchanges, including those between (a) employer

* *L'autore desidera ringraziare per gli utili commenti e suggerimenti ricevuti Maria De Paola, Maurizio Franzini, Antonio Nicita e due anonimi referee della Rivista.*

and employee; (b) lender and borrower; and (c) firm and customer, the agent on one side of the exchange gives money (employer, lender, customer), while the agent on the other side of the exchange gives a promise (employee, borrower, firm). The employee promises to work hard, the borrower promises to repay the loan, and the firm promises to provide high-quality products. Rarely, however, is this promise subject to a contract that can be enforced at reasonably low cost” (Gintis, 1999, p. 6).

L’osservazione della realtà economica mostra che anche quando gli individui sono soggetti a tentazioni opportunistiche, in contesti in cui le informazioni sono imperfette e non esiste una autorità in grado di imporre il rispetto del contratto, gli operatori economici sono in grado di creare schemi di incentivo per indurre le parti in relazione a rispettare fedelmente gli accordi sulle soluzioni efficienti ex-ante. In altri termini, i mercati sono in grado di realizzare sistemi di enforcement endogeni o privati e istituzioni, capaci di risolvere il problema dell’allineamento degli incentivi, grazie alla implementazione di un meccanismo, ad opera dei singoli soggetti in relazione (enforcement bilaterale) o di una pluralità di agenti sul mercato (enforcement multilaterale), che incentiva comportamenti onesti e sanziona l’opportunistico. L’enforcement si identifica sostanzialmente con il concetto di reputazione (con o senza informazione incompleta sulle caratteristiche degli agenti). I contratti stabiliti in queste situazioni sono definiti self-enforcing, poiché la loro applicazione è assicurata dall’autonomo comportamento degli agenti.

Il requisito fondamentale per la realizzazione dell’enforcement è che l’interazione tra i soggetti non sia unica ma si ripeta nel tempo. In secondo luogo, è necessario che la relazione con una specifica parte o la permanenza su un particolare mercato conduca alla creazione di un surplus e consenta pertanto l’acquisizione di quasi-rendite, cioè di remunerazioni che eccedono – a causa dell’esistenza di rilevanti costi di *switching* di varia natura – ciò che è possibile ottenere in relazioni alternative. Infine, si richiede che le parti incaricate del sanzionamento – che possono essere interne o esterne alla relazione – recepiscano le informazioni necessarie e siano dotate di adeguati incentivi per mettere in atto la sanzione.

Lo scopo principale di questo lavoro è una rassegna critica dei principali modelli che analizzano il funzionamento dei meccanismi di enforcement endogeni dei contratti stabiliti nell’ambito dei mercati del lavoro e dei beni, nelle situazioni in cui esiste – accanto alle asimmetrie informative tra le parti (enfaticizzate nei modelli di azzardo morale e agenzia) – anche un problema di *verificabilità* di variabili rilevanti.¹ L’apparato formale di cui si fa uso è quello della teoria dei giochi ripetuti, grazie alla quale i contratti *self-enforcing* possono essere visti come equilibri di Nash perfetti nei sottogiochi.

Mentre l’enforcement esogeno è collegato all’esistenza di “contratti completi”, le strutture di enforcement endogeno sono classificate in due categorie: 1) “enforcement bilaterale” (o *personal enforcement*); 2) “enforcement multilaterale” (o *community enforcement*).

Appartengono alla categoria dell’enforcement “bilaterale” quei meccanismi che assicurano l’applicazione del contratto sulla base delle strategie di sanzionamento messe in atto direttamente dalle parti coinvolte nella relazione consistenti nell’interruzione della relazione e del flusso di quasi-rendite derivante da essa. Le condizioni che garantiscono l’esecuzione del contratto sono: la possibilità di osservare (anche in modo imperfetto) il comportamento della controparte, l’esistenza di un surplus generato dalla relazione e la ripetizione della stessa per un numero molto elevato o infinito di volte. Non si richiede praticamente alcuna trasmissione di informazioni fuori dalla relazione e nessun intervento di parti esterne. Il sanzionamento è fondato sull’esistenza di costi di *switching* o frizioni di

¹ Si trascurano altre questioni a cui pure la teoria dei contratti ha dedicato molta attenzione, come quelle della ripartizione del rischio (*risk-sharing*), dell’opportunistico con investimenti specifici (*hold-up*), e della selezione avversa.

matching, sui possibili costi per la perdita di investimenti specifici o sull'esistenza di un disequilibrio di mercato.

L'enforcement "multilaterale" è necessario quando le relazioni tra le parti hanno durata limitata, o quando alla relazione con un particolare partner non è legata una quasi-rendita e quindi un agente può cambiare partner senza sopportare alcun costo di *switching*, anche grazie al fatto che il mercato è in una situazione di equilibrio walrasiano e un nuovo *matching* è prontamente realizzabile. Tale tipo di enforcement si realizza tramite il comportamento di parti esterne alla relazione, interessate solo indirettamente, che boicottano gli opportunisti escludendoli permanentemente da futuri scambi. Affinché un meccanismo reputazionale possa funzionare, è necessario che soggetti esterni – altri agenti sul mercato o individui socialmente collegati – osservino il comportamento delle parti in relazione, siano in grado di interpretarlo e abbiano interesse ad eseguire il sanzionamento. L'efficacia dipende ancora una volta dall'esistenza di un surplus o beneficio, connesso in questo caso all'operare in un particolare mercato o con un certo insieme di soggetti, oppure derivante dalla possibilità di poter usufruire nelle fasi finali della vita di un premio per i comportamenti cooperativi adottati. Le sanzioni che possono essere inflitte sono talvolta di natura sociale (ostracismo, stigma, ecc.).

Nell'enforcement multilaterale può sorgere un problema di *free-rider*, poiché il sanzionamento si configura come una sorta di bene pubblico. Per prevenire questo problema emergono in alcuni casi istituzioni di supporto ai soggetti privati, che svolgono un ruolo di raccolta, elaborazione e distribuzione delle informazioni o di coordinamento delle fasi di sanzionamento.

Nel corso dell'analisi si evidenzia come nei mercati in cui esistono problemi di enforcement e sorgono "meccanismi di disciplina" dei comportamenti, la funzione allocativa risulta alterata e possono venire meno sia la legge della domanda e dell'offerta, sia l'efficienza e la perfetta mobilità dei fattori.

Il lavoro è organizzato nel seguente modo. Nel paragrafo 2 si espone sinteticamente la teoria dei contratti incompleti e si presentano in generale le caratteristiche dei meccanismi di enforcement endogeni; il paragrafo 3 illustra il funzionamento dell'enforcement bilaterale e il paragrafo 4 è dedicato all'analisi dell'enforcement multilaterale. Seguono alcune considerazioni conclusive.

2. L'enforcement esogeno, i contratti incompleti e i meccanismi endogeni di enforcement

Nei modelli di impostazione neoclassica si suppone, quasi sempre implicitamente, che le transazioni tra gli agenti siano rette da un *contratto completo*, un contratto cioè che stabilisce, per ogni possibile contingenza (stato del mondo), i reciproci obblighi delle parti (prestazioni e pagamenti) e ne assicura il rispetto grazie alle capacità di verifica di una Autorità esterna (la "Corte") e alla sua capacità di imporre sanzioni eventualmente illimitate.²

In questo modo, il sistema legale e giudiziario, oltre ad assicurare il rispetto del sistema dei diritti di proprietà in essere, garantisce la pronta e perfetta esecuzione degli accordi contrattuali, senza nessun costo diretto per le parti. Poiché non richiede una particolare attività dei contraenti, l'enforcement può essere definito esogeno o esterno.

² "In general equilibrium theory, it is supposed that trade takes place through anonymous competitive markets and that every agent abides by the terms of any transaction he or she enters in" (Hart, 1995, p. 4). (Williamson, 1985, p. 93) sostiene: "La maggior parte degli studi dello scambio assume che le dispute contrattuali sono regolate da efficaci norme legali che la magistratura applica in modo competente, raffinato e poco costoso".

La necessità di un meccanismo di enforcement efficace – che fornisca la garanzia di reciproco adempimento contrattuale ai soggetti che entrano in relazione – è vitale se i soggetti economici si comportano in modo *opportunistic*, cioè se «perseguono con astuzia finalità egoistiche», ricorrendo anche a comportamenti quali «mentire, imbrogliare, rubare, fuorviare, travisare, offuscare, fingere, distorcere e confondere» pur di raggiungere i propri fini (Williamson, 1985, p. 134).³

Il controllo dell'opportunismo attraverso il ricorso al sistema legale richiede la definizione di contratti completi ed espliciti. Tuttavia, la distanza nel tempo e nello spazio tra il quid e il quo della relazione, la razionalità solo limitata degli agenti, l'esistenza di costi di transazione e l'informazione imperfetta e distribuita in maniera asimmetrica impediscono che le condizioni necessarie per stipulare tali contratti siano soddisfatte. In maggior dettaglio, le cause di incompletezza contrattuale che impediscono l'enforcement esogeno in una molteplicità di relazioni sono le seguenti (cfr. Hart - Holmstrom, 1987; Milgrom - Roberts, 1990; Tirole, 1999):

- a) l'impossibilità da parte dei contraenti di prevedere ogni possibile contingenza futura che dovesse verificarsi nel corso della relazione;
- b) gli elevati costi di contrattazione (*bargaining costs*) per accordarsi su ogni singola circostanza e il costo di descriverle (in modo non ambiguo) nel contratto;
- c) il costo di ricorrere al sistema legale per ottenere l'adempimento del contratto;⁴
- d) la difficoltà di ottenere l'enforcement del contratto, a causa delle informazioni asimmetriche esistenti tra le parti riguardo ad azioni, caratteristiche o stati del mondo (*non osservabilità*), oppure alla difficoltà di trasmettere queste informazioni – anche se condivise tra le parti – all'autorità esterna (*non verificabilità*);⁵ in alcune circostanze, il rischio di dare luogo a comportamenti distorti (*dysfunctional*) induce le parti a lasciare fuori dal contratto anche elementi verificabili ("incompletezza endogena").

Pertanto, l'opportunismo può sorgere direttamente, quando le parti cercano di approfittare della mancanza di osservabilità o di verificabilità, oppure indirettamente, quando i "gap" contrattuali, derivanti dalle difficoltà previsionali, dai costi di bargaining o di scrittura, permettono ad una delle parti di sfruttare a proprio vantaggio le situazioni future (Kreps, 1990).

³ La teoria standard non esclude, a rigore, la possibilità che gli accordi intercorsi tra le parti siano rispettati volontariamente dagli individui, senza la necessità dell'intervento dell'Autorità. In effetti, "il modello classico tace sulla rettitudine di comportamento di questi agenti. Non ci dice se si tratta di persone d'onore, condizionate dall'educazione a tenere sempre fede agli obblighi che hanno scelto di assumersi, oppure se esiste una autorità superiore dotata di credibilità sufficiente a fare rispettare i contratti minacciando di comminare sanzioni nel caso in cui non si mantengano gli impegni" (Dasgupta, 1988, p. 63). Quando i soggetti non sono mossi esclusivamente dal *self-interest* e dall'opportunismo, ma sono guidati nei loro comportamenti da valori quali l'onestà, la giustizia, i sentimenti di altruismo e di reciprocità, i principali problemi generati dall'incompletezza contrattuale sono attenuati, poiché non esiste una effettiva divergenza di interessi tra le parti e si parla di "enforcement morale" o di "first-party enforcement" (si vedano Kreps, 1997; Greif, 1994).

⁴ Altri motivi di impossibilità di enforcement esterno si hanno quando manca del tutto una Autorità, dotata di poteri coercitivi (ad esempio nei rapporti tra Stati sovrani), oppure quando l'ordinamento legislativo prevede un esplicito impedimento all'effettuazione di alcuni tipi di transazioni.

⁵ Il contratto di agenzia standard può essere incluso nella classe dei contratti completi. Nonostante la relazione tra Principale e Agente sia afflitta dall'asimmetria informativa (l'agente controlla una certa variabile che il Principale non può osservare), si suppone che esista un'altra variabile, correlata alla prima, che è verificabile dalla Corte, su cui l'intero contratto viene basato. Si assume che virtualmente ogni tipo di disposizione possa essere inserita in un tale contratto, il quale prevede ogni possibilità e dispone per essa i profili di azioni appropriati: un contratto di agenzia non dovrà mai essere rivisto o rinegoziato. "In principal-agent theory, it is supposed that it is costless to write a contract. An implication is that an optimal contract will be 'comprehensive'" (Hart, 1995, p. 4). Pertanto, il contratto di agenzia, fatto rispettare facendo ricorso alla Corte, è assimilabile ai contratti completi.

2.1. I meccanismi endogeni di enforcement e i contratti impliciti self-enforcing

Nelle circostanze in cui non è possibile stipulare un contratto completo sorge la necessità di utilizzare meccanismi di enforcement endogeni o privati, in alternativa al sistema legale: “where some aspect of the good or service supplied is both valuable to the buyer and costly to provide, the absence of third-party enforcement of claims gives rise to endogenous enforcement strategies” (Bowles - Gintis, 1993, p. 85). In pratica, la relazione contrattuale deve rispettare un vincolo di compatibilità degli incentivi: le parti devono avere interesse a tenere fede alle promesse anche in assenza di una autorità esterna che lo impone con metodi coercitivi.

A partire dai lavori classici di Macaulay (1963), Macneil (1978) e Williamson (1985), esiste ormai una vasta letteratura che dimostra empiricamente come, nelle economie avanzate e in misura maggiore nei paesi meno sviluppati, la maggior parte degli accordi nel mondo degli affari non è basata su contratti espliciti e non è fatta rispettare facendo ricorso al sistema legale, ma utilizzando altri meccanismi informali, privati, basati sulla fiducia personale, sulla reputazione, sull'appartenenza a particolari comunità o network, su istituzioni private, sulla “disciplina dei mercati”. Si vedano, per esempio, i lavori di: Bernstein (1992) sul mercato internazionale dei diamanti; De Soto (1989) e Portes (1992) sul settore informale in America Latina; Fafchamps, (1993) per il Ghana e altri paesi africani; Greif - Kandel (1995) e Johnson - MacMillan - Woodruff (2002) per la Russia e altre economie in transizione; Banerjee - Duflo (2001) per il mercato del software in India.⁶

L'enforcement endogeno viene generalmente realizzato attraverso la costruzione di strutture e istituzioni che prevedono un incentivo per coloro che si conformano agli accordi sottoscritti e impongono un sanzionamento di qualche natura per i soggetti che non si attengono agli impegni assunti. La minaccia di sanzioni deriva non più da una autorità esterna ma dal comportamento della controparte direttamente in relazione al soggetto deviante (enforcement bilaterale) oppure da altri soggetti esterni alla relazione – ma in qualche forma connessi alle parti contrattuali – che ricevono le informazioni rilevanti e sono interessati in via indiretta (enforcement multilaterale).⁷

La caratteristica fondamentale, necessaria per l'attivazione di schemi di enforcement endogeno, è che la relazione economica tra le parti sia ripetuta nel tempo oppure che il loro orizzonte temporale si estenda al di là della specifica relazione. La continuità della relazione assicura ad ogni soggetto la possibilità di intraprendere azioni o strategie condizionate ai comportamenti tenuti in precedenza dalle controparti (reputazione).

Un altro requisito essenziale è che i soggetti incaricati di attuare le sanzioni dispongano delle informazioni necessarie. Pertanto, il comportamento di ognuno deve essere direttamente osservabile, o deve poter essere inferito sulla base di qualche variabile osservabile. Inoltre, le parti devono essere in grado di giudicare ex-post se il comportamento è stato corretto e per tutti deve risultare chiaro quali azioni bisogna seguire nelle varie circostanze.⁸

I meccanismi di enforcement endogeni sono basati su un *contratto implicito* stabilito tra le parti, che consiste in un accordo informale, generalmente non scritto, che delinea i comportamenti a cui

⁶ Altrettanto numerosi i lavori che mostrano l'impiego di vari meccanismi e istituzioni private per il rispetto dei contratti nel corso della storia economica (si veda il paragrafo 4).

⁷ Le definizioni di enforcement *endogeno* e *esogeno* sono di Bowles - Gintis (1993). La distinzione tra enforcement bilaterale e multilaterale è alla base del lavoro di Greif (1993), *personal* e *community* enforcement sono state coniate da Kandori (1992a). Distinzioni analoghe sono presenti in Hart - Holmstrom (1987), Klein, 1992) e Stiglitz (1992) che per il mercato del lavoro distingue tra: 1) “reaction of current employees”; 2) “reaction of potential employees”.

⁸ Come osserva Kreps (1990, p.105): “compliance must be observable” e “we suppose that we understand what compliance means”.

ciascuna di esse deve conformarsi; l'accordo, per i motivi anzidetti, non ha nessun valore legale: si tratta di un semplice scambio di promesse (e di minacce!) e di un insieme di aspettative condivise.⁹ Tuttavia, sotto certe condizioni, il contratto implicito è sufficiente per garantire la realizzazione di scambi efficienti poiché permette di stabilire gli incentivi che inducono ogni parte ad adempiere alle proprie obbligazioni. "A self enforcing agreement between two parties remains in force as long as each party believes himself to be better off by continuing the agreement than he would be by ending it" (Telser, 1980).¹⁰

L'analisi nei termini della teoria dei giochi

Per il modo in cui sono strutturate, le relazioni contrattuali sono analizzate con grande efficacia ricorrendo all'apparato formale della teoria dei giochi. Il gioco costituente (*stage-game*) alla base della relazione tra le parti può essere rappresentato dal gioco standard del dilemma del prigioniero, con mosse simultanee, oppure dal gioco sequenziale del "dilemma del prigioniero unilaterale". Nel primo caso si ha la seguente matrice dei *pay-off*:

	2	Rispetta contratto	Defeziona
1			
Rispetta contratto		Π^C, Π^C	Π^F, Π^M
Defeziona		Π^M, Π^F	Π^D, Π^D

Tabella 1

dove $\Pi^M > \Pi^C > \Pi^D > \Pi^F$. Entrambe le parti hanno la tentazione di agire opportunisticamente, poiché le loro scelte non sono immediatamente note alla controparte.

Nel dilemma del prigioniero unilaterale (definito il "gioco della fiducia"), **rappresentato nella fig. 1**, un giocatore muove per primo, scegliendo tra due possibilità: a) entrare in relazione e offrire fiducia; b) non entrare in relazione. In questo secondo caso, il gioco finisce e entrambi i giocatori ottengono la loro utilità di riserva Π^D (che può essere normalizzata a zero). Se, invece, il giocatore 1 sceglie di offrire fiducia, il secondo giocatore può onorarla (con *pay-off* per entrambi di Π^C) oppure approfittarne, ottenendo Π^M e costringendo il primo ad ottenere Π^F . In questo caso, a causa della struttura sequenziale del gioco solo la parte che agisce per prima rischia di subire l'opportunismo dell'altra.

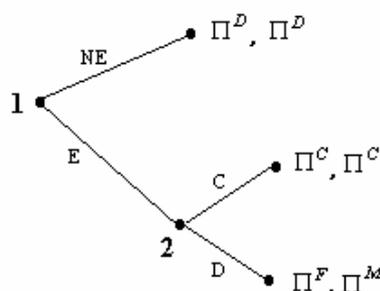


Figura 1

⁹ Saranno definiti contratti impliciti tutti quegli accordi che contengono elementi importanti che non sono *enforceable* per mezzo della Corte. Tuttavia, anche all'interno di un contratto implicito possono esistere degli aspetti che sono invece pienamente verificabili e sono fatti rispettare tramite la Corte.

¹⁰ I contratti self-enforcing sono definiti anche "contratti relazionali" (Baker - Gibbons - Murphy, 2002).

In entrambi i giochi, le parti avrebbero interesse a entrare in relazione se potessero impegnarsi e avere garanzia che le azioni scelte da tutti saranno quelle cooperative. Nei cosiddetti “giochi cooperativi” esiste la possibilità per le parti di effettuare un *binding commitment*, che vincola in modo assoluto le parti a seguire una data strategia concordata. Al contrario, nei “giochi non cooperativi” non si può fare affidamento su meccanismi esterni per vincolare le scelte dei giocatori e una eventuale soluzione cooperativa deve emergere dalle autonome decisioni dei giocatori. Come è evidente, la distinzione tra giochi di natura cooperativa e giochi non cooperativi corrisponde alla distinzione tra enforcement esogeno e endogeno.

In assenza di enforcement esogeno, le scelte dei soggetti non sono vincolate da accordi o comunicazioni intervenute prima del gioco vero e proprio e ognuno tende a effettuare la scelta migliore per il proprio benessere, data la scelta effettuata dagli altri (equilibrio di Nash). Se la relazione è unica, cioè se il gioco è *one-shot*, nei due giochi presentati l'equilibrio di Nash risultante è quello non cooperativo e Pareto-inefficiente, in cui le parti ottengono solo Π^D o non entrano in relazione. Se, invece, il gioco è ripetuto per un certo numero di volte (supergioco) la conclusione può mutare drasticamente.

In questi giochi, al termine di ogni periodo le parti osservano i comportamenti adottati da tutti e possono, quindi, stabilire se l'accordo è stato rispettato o meno. E' possibile allora dimostrare – tramite il noto *Folk Theorem* – che la ripetizione del gioco, attraverso l'impiego di strategie condizionate alla “storia del gioco”, che costituiscono parte integrante del contratto implicito, rende possibile l'emergere di un equilibrio cooperativo (enforcement contrattuale).

I contratti impliciti *self-enforcing*, nelle versioni più generali, sono quindi rappresentabili come equilibri di Nash, perfetti nei sottogiochi, di un gioco ripetuto tra le parti.¹¹ Un problema serio di questo approccio è che possono esistere molteplici equilibri di questo tipo. Una ipotesi ragionevole usualmente adottata per limitare il numero degli equilibri è che i contratti impliciti *self-enforcing* costituiscano equilibri Paretoefficienti.

3. L'enforcement bilaterale

La categoria dell'enforcement bilaterale – o *personal* (cfr. Kandori, 1992a) o *internal* (cfr. Hart - Holmstrom, 1987) – è basata sull'applicazione di strategie di «punizione diretta», tramite le quali un soggetto può infliggere direttamente alla propria controparte una serie di costi quando quest'ultima si dimostra in qualche modo inadempiente. La caratteristica peculiare di questo genere di sanzione è che essa viene comminata senza la necessità che altri soggetti terzi, al di fuori della relazione, acquisiscano informazioni o adottino particolari comportamenti sanzionatori: solo la parte che subisce l'opportunismo è impegnata a sanzionare.

In situazioni in cui il ricorso alla Corte non è possibile, spesso l'unica alternativa a disposizione delle parti è quella di interrompere una relazione se si è insoddisfatti o di evitare di entrare in rapporto con soggetti ritenuti inaffidabili.¹² Infatti, in assenza di una misura verificabile, viene meno la possibilità che i termini dello scambio siano usati in modo condizionale all'effettivo svolgimento della

¹¹ In seguito, si chiarirà in quali condizioni è necessario che il gioco sia ripetuto infinitamente e in quali è sufficiente la ripetizione solo per un numero finito di volte.

¹² “Most of the literature assumes that, without verifiable performance, contracts must be of the termination type, in which an employee receives an explicitly agreed wage as long as performance in the past has been satisfactory but is dismissed if it has not. Such contracts look more like actual labor contract than those in standard principal agent model” (MacLeod - Malcomson, 1989, p. 448).

performance, poiché meccanismi di questo tipo darebbero luogo semplicemente all'opportunismo della controparte. Un contratto con remunerazione legata alla performance potrebbe risolvere l'opportunismo del lavoratore, ma in assenza di verificabilità della variabile da cui dipendono i compensi dell'agente (e quindi con "valutazione soggettiva della performance"), sposterebbe il rischio di opportunismo dal lavoratore all'impresa (si veda però il paragrafo 4.3). Per la stessa ragione, non è possibile il ricorso a strumenti quali l'imposizione di penalità, come proposto in Becker - Stigler (1974), o la confisca di *performance bond* (Carmichael, 1985).¹³

Una tipica strategia, adottata diffusamente nella realtà e impiegata in teoria anche per la sua semplicità analitica, che, sotto assunzioni mediamente restrittive, riesce ad assicurare l'applicazione del contratto nei contesti in esame, è quella che viene definita "strategia di rinnovo contingente" (Bowles - Gintis, 1993), che consiste nella minaccia di un soggetto di interrompere la relazione se l'accordo sottoscritto non viene rispettato dalla controparte ("*termination-type contract*"). "If agents do not keep their promises, principals dump them: employers fire workers who shirk, lenders refuse future loans to borrowers who have defaulted, and customers switch to new suppliers when dissatisfied" (Gintis, 1999, p. 6).

Il contratto con rinnovo contingente prescrive termini di scambio indipendenti dalle variabili che non sono pubblicamente osservabili e prevede l'interruzione della relazione in caso di performance insoddisfacente. In tal modo, esso consente di superare i problemi di opportunismo dell'Agente senza crearne di nuovi a carico del Principale. In secondo luogo, essendo questo contratto bilaterale, non è attaccabile da forme di corruzione: «a [...] reason for using two-party contracts is that contracts involving more parties give scope for collusion and corruption in ways that two-party contracts do not» (MacLeod - Malcomson, 1988, p. 850).¹⁴

Nei termini della teoria dei giochi, la "strategia di rinnovo contingente" corrisponde all'adozione di una "*trigger strategy*", che prevede il rispetto dell'accordo fino ad un eventuale comportamento opportunistico della controparte, a partire dal quale si innesca l'interruzione perenne della cooperazione. L'efficacia dell'enforcement bilaterale può essere dimostrata formalmente mediante l'ausilio delle varie versioni del «Folk Theorem» (ad esempio, Friedman, 1971; Fudenberg - Maskin, 1986), che, nell'ambito di giochi come il dilemma del prigioniero ripetuto un numero infinito di volte, dimostrano che la soluzione Pareto-efficiente di continua mutua cooperazione, può emergere come equilibrio di Nash (perfetto nei sottogiochi), se i giocatori attuano la minaccia di passare all'equilibrio di Nash statico in caso di defezione¹⁵ e se il loro tasso di sconto è relativamente basso (cioè, se essi valutano sufficientemente i benefici futuri).

In termini economici, le condizioni necessarie affinché la minaccia delle sanzioni bilaterali sia efficace e il contratto sia rispettato dalle parti sono le seguenti: (1) l'osservabilità tra le parti stesse delle azioni intraprese; (2) la relazione deve avere un orizzonte potenzialmente molto lungo e con termine incerto; (3) dalla relazione deve scaturire un *surplus* rispetto alle opportunità alternative.

L'applicazione di un contratto tramite la minaccia di sanzioni bilaterali richiede semplicemente che tra le parti vi sia osservabilità (perfetta o imperfetta) delle variabili rilevanti e la conoscenza e la comprensione degli accordi contrattuali. Come dimostrato da Green - Porter (1984), è spesso sufficiente anche un grado di osservabilità imperfetta, anche se in questo caso le strategie ottimali

¹³ A volte sono disposizioni legali che rompono di fatto il legame tra il salario e la performance (vietando, ad esempio, che il salario possa essere negativo o al di sotto di un certo livello).

¹⁴ Sull'efficacia delle strategie di *termination* per i mercati del lavoro e del credito, si veda anche Stiglitz - Weiss (1983).

¹⁵ La perfezione nei sottogiochi è automaticamente garantita dal fatto che la minaccia consiste nel giocare strategie che fanno parte dell'equilibrio di Nash statico.

possono essere diverse da quelle “trigger”.¹⁶ Non è necessaria la verificabilità o alcuna trasmissione di informazioni al di fuori della relazione.

Il secondo requisito fondamentale per l’efficacia dell’enforcement bilaterale è che la relazione non abbia un termine definito e che la probabilità che continui per un ulteriore periodo sia sufficientemente elevata.

E’ possibile mostrare che la struttura di questo gioco ripetuto con termine incerto è analoga ad un gioco ad orizzonte infinito (particolarmente semplice da analizzare tecnicamente), se il tasso di sconto dei pay-off futuri è opportunamente ridefinito per tenere conto della probabilità di interruzione (q). Se r' rappresenta il puro tasso di preferenza intertemporale, allora l’equivalente tasso di sconto r in un orizzonte infinito è pari a: $r = (r' + q)/(1 - q)$. In pratica, il tasso di sconto risulta semplicemente più elevato a causa delle possibilità di interruzione della relazione.

Se, al contrario, il gioco è ripetuto un numero T finito di volte, e il periodo terminale è noto con certezza, Selten (1978) ha dimostrato con il noto *chain store paradox*, che non esistono possibilità di fare emergere comportamenti cooperativi, anche se T è un numero molto grande.¹⁷

L’enforcement del contratto richiede, pertanto, se deve basarsi unicamente su sanzioni bilaterali, che l’orizzonte temporale sia potenzialmente infinito. La teoria ha anche accertato che questa drastica conclusione non è valida in due circostanze: 1) se lo *stage-game* prevede equilibri di Nash multipli, ordinabili parzialmente. Benoit - Krishna (1985) dimostrano che in questo caso la soluzione Pareto-superiore può essere ottenuta anche se il gioco è ripetuto un numero finito di volte, se le parti usano la minaccia di passare a un equilibrio Pareto-inferiore; 2) se esiste informazione incompleta sulle caratteristiche dei soggetti (Kreps - Milgrom - Roberts - Wilson, 1982) (si veda paragrafo 3.1).

Una terza condizione indispensabile affinché il contratto sia *self-enforcing* è che esista un interesse delle parti a continuare la relazione intrapresa. Ciò richiede che la relazione sia in grado di generare un surplus, costituito da un flusso di quasi-rendite,¹⁸ di cui le parti possano appropriarsi (MacLeod - Malcomson, 1989). Bowles - Gintis (1993) sostengono che per il rispetto dei contratti sia necessario il pagamento di una “enforcement rent”, definita come “a payment in excess of at least one agent’s best alternative” (si veda anche Stiglitz, 1992).

¹⁶ Il modello di Green - Porter (1984) analizza gli accordi di collusione tra imprese oligopolistiche le quali raggiungono un accordo sulla quantità da produrre. Poiché tale variabile non è direttamente osservabile, le imprese non sono in grado di stabilire con certezza se l’accordo viene rispettato o meno. Ogni impresa osserva invece il prezzo, determinato congiuntamente dalle quantità prodotte e da uno shock casuale, e decide le proprie azioni sulla base di esso. Se il prezzo scende sotto una certa soglia si fa scattare la fase punitiva, poiché ciò segnala, con elevata probabilità, la defezione della controparte. Poiché può anche darsi che l’evento si verifichi per un fatto accidentale (per esempio, domanda fiacca), è opportuno non punire per sempre, ma solo per un certo numero di periodi T , al termine dei quali si torna a cooperare. L’entità della sanzione (con quale frequenza viene innescata, quanto tempo dura, ecc.) deve essere stabilita in base a due esigenze contrapposte: se è troppo blanda, i contraenti tendono ad approfittarsene, comportandosi con frequenza opportunisticamente; d’altra parte, se è troppo dura, si rinuncia spesso ai vantaggi derivanti dalla cooperazione.

¹⁷ La dimostrazione è basata sul metodo dell’induzione a ritroso. Nell’ultimo periodo, siccome non esistono periodi successivi nei quali adottare promesse o minacce, entrambe le parti defezionano dagli impegni contrattuali. Nel penultimo periodo, ($T-1$), dato che è nota la comune defezione in T , la strategia dominante è ancora quella di defezionare. Questo ragionamento si applica a ritroso su tutti i periodi precedenti.

¹⁸ “Una quasi-rendita è costituita dalla frazione dei proventi che eccede l’ammontare minimo necessario per evitare che un lavoratore abbandoni un certo lavoro o un’impresa abbandoni una certa industria” (Milgrom - Roberts, 1992, p. 400). Mentre le rendite sono definite in termini di una decisione di entrata, le quasi-rendite fanno riferimento a una decisione di uscita. La differenza tra le rendite e le quasi-rendite è costituita dalla presenza di costi da sostenere per entrare in una certa relazione, che non sono recuperabili (*sunk-cost*) se si decide di uscire.

Così, di fronte alla minaccia di interruzione della relazione, ogni soggetto è portato ad aderire agli impegni per non subire la perdita dei guadagni futuri generati dalla relazione, quando questi sono superiori ai benefici immediati derivanti da un comportamento opportunistico.

In termini analitici, facendo riferimento alla Tabella 1, affinché sia garantito il rispetto del contratto è necessario che il guadagno di breve periodo realizzato attraverso una defezione, pari a $(\Pi^M - \Pi^C)$, sia inferiore al valore attuale del flusso di quasi rendite future, pari a $\frac{1}{r}(\Pi^C - \Pi^D)$, dove r rappresenta il tasso di sconto (che include la probabilità di interruzione della relazione), che implica:

$$r \leq \frac{\Pi^C - \Pi^D}{\Pi^M - \Pi^C}$$

Pertanto, l'enforcement contrattuale basato su sanzioni bilaterali è tanto più probabile quanto:

- 1) più elevati sono i vantaggi della cooperazione (maggiore è Π^C rispetto a Π^D)
- 2) minori sono i benefici di una eventuale defezione (minore è Π^M rispetto a Π^C);
- 3) minori sono la probabilità di interruzione e il puro tasso di sconto intertemporale (poiché riducono il valore dei benefici futuri).

In sintesi, come si mostrerà in seguito anche attraverso l'esposizione di alcuni modelli, l'enforcement bilaterale tende a prevalere quando esiste osservabilità tra le parti, ma la trasmissione di queste informazioni all'esterno è proibitivamente costosa, se le relazioni originano surplus o quasi-rendite sufficientemente elevate e se i rapporti hanno potenzialmente una durata molto lunga, cioè se il turnover esogeno è limitato.

Le fonti delle quasi-rendite nelle relazioni economiche

Per aggiungere un po' di struttura economica alla scarsa rappresentazione fornita dalla teoria dei giochi è necessario individuare nelle effettive relazioni economiche quali sono le cause che determinano l'emergere di un surplus superiore a quello che le parti possono conseguire all'esterno. Si noti che surplus di questa natura sono generalmente in contraddizione con il funzionamento di mercati perfettamente concorrenziali, poiché l'operare della forze di mercato tende ad annullarli. Tuttavia, l'assenza di quasi-rendite costituirebbe un impedimento all'esistenza di una serie di relazioni economiche e di interi mercati.

Le fonti delle quasi-rendite collegate a una relazione con uno specifico partner comunemente considerate possono essere raggruppate in tre classi. Innanzitutto, a causa di frizioni di varia natura e delle informazioni imperfette esistenti nei mercati, le parti, uscendo dalla relazione, possono incorrere in costi per cambiare il partner (costi di *switching*), che consistono in costi di informazione, di ricerca, di turnover, ecc..

In secondo luogo, è possibile che per instaurare la relazione un soggetto abbia dovuto effettuare un investimento specifico, abbia cioè acquisito risorse che hanno valore solo fin quando sono impiegate all'interno della relazione stessa, mentre perdono significativamente di valore al di fuori di essa (cioè per realizzare la transazione si sono sostenuti dei *sunk cost*).

Infine, il mercato nel quale la transazione si svolge può trovarsi in disequilibrio. Quando il lato del mercato di cui un agente fa parte è razionato, l'interruzione della relazione impone su di esso un costo, derivante dalla probabilità positiva di non riuscire a formare prontamente un nuovo match e di dover rinunciare alla differenza tra l'utilità derivante dalla relazione e la sua utilità di riserva.¹⁹

¹⁹ Nell'ambito del mercato del lavoro, ad esempio, l'impresa infligge un costo licenziando un lavoratore opportunisto se: a) esistono frizioni o costi di switching per cambiare lavoro; b) il lavoratore ha effettuato investimenti specifici in capitale umano; c) sul mercato esiste un eccesso di offerta di lavoro (disoccupazione). D'altra parte, se il lavoratore abbandona il proprio posto, nei casi di opportunismo dell'impresa, quest'ultima

In generale, un soggetto è penalizzato dall'interruzione della relazione se le sue opportunità alternative sono ridotte. Il sanzionamento derivante dalla perdita delle quasi-rendite originate nelle tre modalità esaminate, è considerato di natura bilaterale poiché scatta con la semplice interruzione della relazione, decisa unicamente per iniziativa della controparte danneggiata.

3.1. Il funzionamento dell'enforcement bilaterale in alcuni modelli tipici

L'enforcement attraverso la disoccupazione nei modelli con salario di efficienza

In uno dei più noti modelli della letteratura sui salari di efficienza, Shapiro - Stiglitz (1984) affrontano il problema dell'enforcement del contratto stipulato tra impresa e lavoratori, in presenza di osservabilità imperfetta e non verificabilità dello sforzo prestato dal lavoratore.

La soluzione che essi prospettano, in un contesto con orizzonte temporale infinito, fa affidamento sulla combinazione tra il pagamento di un salario in eccesso rispetto al salario di riserva (una "rendita da enforcement" determinata endogenamente) e la minaccia di una sanzione inflitta direttamente dall'impresa ai lavoratori opportunisti, consistente nel licenziamento e in un conseguente periodo di disoccupazione. La disoccupazione rappresenta il "meccanismo di disciplina dei lavoratori" e permette l'enforcement del contratto attraverso il rischio di perdita del flusso di quasi-rendite associato all'occupazione.

Il modello di Shapiro e Stiglitz descrive quei mercati del lavoro in cui la diffusione delle informazioni è problematica, non esistono investimenti specifici e altri costi di *switching*²⁰ e la necessità di creare quasi-rendite impone di pagare ai lavoratori un salario di efficienza sufficientemente elevato,²¹ che altera l'equilibrio di mercato e genera disoccupazione involontaria, annullando i "gains from trade" di molte relazioni potenzialmente efficienti.

Disequilibri sul mercato e struttura dei contratti self-enforcing

Un risultato fondamentale che stabilisce un legame tra l'enforcement dei contratti di lavoro e le condizioni del mercato sul quale le relazioni si svolgono è raggiunto nei lavori di MacLeod - Malcomson (1989; 1998). Essi dimostrano che il tipo di contratto che può essere reso *enforceable* dalle strategie dei soggetti varia sostanzialmente a seconda della situazione prevalente sul mercato.

Nel modello si assume che esiste piena osservabilità della prestazione lavorativa, ma è impossibile la *verifica* presso una autorità esterna, e che gli agenti vivono indefinitamente. Il contratto analizzato da MacLeod e Malcomson è composto da una parte esplicita e da una implicita. La parte esplicita prevede il pagamento di un salario fisso w . La possibilità per l'impresa di osservare le azioni del lavoratore le consente di pagare un salario connesso alla prestazione: il contratto implicito stabilisce che l'impresa paghi un bonus b in aggiunta a w (e continui a occupare il lavoratore) solo se l'impegno prestato è stato elevato. Il lavoratore, dal canto suo, presta lo sforzo alto se ha sempre ricevuto il bonus e abbandona l'impresa se, avendo prestato l'impegno concordato, non riceve il bonus promesso.

sopporta un costo se ha sostenuto almeno in parte dei costi per assumere il lavoratore, se ha investito risorse nella sua formazione, o se esistono dei posti di lavoro vacanti che le imprese non riescono prontamente a coprire.

²⁰ Un recente modello di Ramey - Watson (2001) analizza l'enforcement contrattuale basato su frizioni di matching. Malcomson (1999) discute informalmente della possibilità di enforcement attraverso investimenti specifici e costi di turnover. Per un modello in cui i costi di turnover e gli investimenti specifici svolgono il ruolo di meccanismo di enforcement sia consentito il rinvio a Scoppa (2001; 2003).

²¹ Il salario di efficienza risulta tanto più elevato quanto più difficile è l'osservabilità e quanto più frequente è il turnover.

MacLeod - Malcomson (1989; 1998) dimostrano che tali strategie costituiscono un equilibrio di Nash perfetto. Tale equilibrio è garantito dall'esistenza di un surplus, generato all'interno della relazione, rispetto a quello che le parti riescono ad ottenere all'esterno.

La forma effettiva del contratto – che determina anche come il surplus viene suddiviso tra impresa e lavoratore – dipende comunque dalla situazione complessiva sul mercato del lavoro, che MacLeod e Malcomson considerano esogenamente determinata.

Si supponga che esista disoccupazione. In assenza di verificabilità, l'unica forma contrattuale che garantisce il lavoratore è che tutta la sua remunerazione sia fissata a priori nel contratto esplicito, mentre il bonus pagato discrezionalmente non può che essere nullo. In caso contrario, infatti, siccome l'impresa può facilmente sostituire il lavoratore con un disoccupato, essa avrebbe sempre convenienza a licenziarlo al termine del periodo, evitando di pagare il bonus anche se la prestazione è stata soddisfacente. Mentre il lavoratore deve ottenere un surplus rispetto alla condizione di disoccupazione, come incentivo all'impegno, tale surplus deve essere assente per l'impresa, altrimenti avrebbe l'interesse a sostituire il lavoratore e ricominciare il gioco con un altro. Questa situazione è analoga a quella descritta da Shapiro - Stiglitz (1984), anche se nel modello di MacLeod e Malcomson è l'esistenza di disoccupazione che induce le imprese a pagare un salario di efficienza e non viceversa.

I modelli con salari di efficienza costituiscono utili rappresentazioni di quei mercati del lavoro "anonimi", in cui è difficile la diffusione di informazioni all'esterno, e in cui i rapporti tra impresa e lavoratore tendono ad essere relazioni di lunga durata. Infatti, in mercati con elevato turnover il salario di efficienza per scoraggiare l'opportunismo risulterebbe proibitivamente elevato.

La situazione opposta all'esistenza di disoccupazione è quella in cui esiste un eccesso di domanda di lavoro, cioè quando esistono dei posti di lavoro vacanti poiché le imprese non riescono a trovare tutti i lavoratori necessari. In questo caso, il contratto deve essere differente dal salario di efficienza. Infatti, se si pagasse un salario fisso, come nella situazione precedente, nessun lavoratore avrebbe interesse ad impegnarsi. Occorre, invece, che l'impresa paghi una remunerazione collegata alla prestazione, cosicché il lavoratore riceve il bonus solo se si è impegnato ("valutazione soggettiva della performance"). Il rispetto del contratto da parte dell'impresa (che deve pagare il bonus, decidendo discrezionalmente *ex-post*) è assicurato concedendo ad essa il surplus derivante dalla relazione, che può venir meno in caso di separazione dal lavoratore. In altre parole, all'impresa non conviene defezionare poiché il lavoratore può sanzionarla decidendo di dimettersi: data la prevalenza di un eccesso di domanda di lavoro, per l'impresa sarebbe notevolmente costoso sostituire il lavoratore.

Pertanto, il contratto con la remunerazione legata alla performance può emergere solo in mercati del lavoro *tight*, in cui le imprese si contendono i lavoratori (mentre questi ultimi non hanno difficoltà a cambiare occupazione). Ciò deriva crucialmente dall'ipotesi che non è possibile (o eccessivamente costosa) una accurata diffusione delle informazioni nel mercato sui comportamenti adottati dalle parti. Bull (1987) e Baker - Gibbons - Murphy (1994) abbandonano questa ipotesi e studiano l'enforcement di un contratto con bonus attraverso la reputazione dell'impresa (si veda paragrafo 4.3).

Come si desume dal modello, il contratto è congegnato in modo tale che la parte che ha la facoltà di esimersi dal rispetto senza sopportare costi (l'impresa, quando esiste disoccupazione; il lavoratore, se ci sono dei posti vacanti) debba eseguire i propri impegni per prima (o sia obbligato a farlo da una esplicita disposizione);²² la parte che adempie «residualmente» è incentivata a farlo dal fatto che riceve tutto il surplus dalla relazione e può essere penalizzata dalla reazione della controparte.

²² Siccome il pagamento del salario è considerato perfettamente *enforceable* il momento dell'effettivo pagamento non è realmente importante.

Poiché in questo modello non esistono altre fonti di surplus, quali costi di mobilità o investimenti specifici o rendite reputazionali, l'enforcement si basa sulle quasi-rendite generate a favore dei soggetti presenti sul «lato corto» del mercato in presenza di disequilibrio tra domanda e offerta.

L'inefficacia dell'enforcement bilaterale in situazioni di market-clearing

In assenza di quasi-rendite di altra natura, la necessità di una divergenza tra beneficio e costo marginale, al fine di assicurare l'efficacia dell'enforcement, è dimostrata formalmente da Greif - Milgrom - Weingast (1994). Essi mostrano che in un mercato in equilibrio non è possibile garantire l'applicazione del contratto se l'unico provvedimento possibile contro chi si comporta opportunisticamente è costituito da sanzioni di natura bilaterale, cioè se si può sanzionare solo con l'interruzione della relazione in atto e il rifiuto a intrattenere ulteriori affari.

L'obiettivo del loro modello è quello di illustrare l'emergere di alcune istituzioni commerciali, come le "gilde mercantili", nell'Europa continentale del Dodicesimo e Tredicesimo secolo, ma la loro idea fondamentale è utile anche per la spiegazione del funzionamento degli attuali mercati.

La relazione di interesse è quella tra le autorità di una città o centro commerciale e i mercanti che vi si recano per effettuare scambi. Le città ospitanti ricevono dei benefici fiscali e commerciali dalla presenza dei mercanti, ma, agendo opportunisticamente – espropriando i mercanti o non fornendo loro adeguata protezione (poiché è costoso farlo) – possono ottenere un beneficio immediato molto maggiore. L'ipotesi chiave è che i benefici marginali del commercio per la città decrescono al crescere del volume degli scambi.

Greif - Milgrom - Weingast (1994) dimostrano che se il mercato raggiunge il livello di scambio efficiente, cioè se il commercio è spinto fino al punto in cui il beneficio marginale è esattamente pari al costo marginale, allora il mercante «marginale» fornisce alla città un flusso di profitti praticamente nullo. Pertanto, se quest'ultimo viene espropriato, e come reazione si astiene dall'effettuare ulteriori scambi, la città, a fronte del «bottino» estorto, non sopporta alcuna perdita. In altri termini, se la ritorsione è unicamente di natura bilaterale – cioè, se solo il mercante espropriato si astiene dal commercio – il costo che si può infliggere alla città opportunistica è del tutto trascurabile.

Nonostante quindi esista tra ogni mercante e la città una relazione di durata potenzialmente infinita, l'effettiva assenza di una sanzione, causata dal fatto che il surplus conseguibile con i mercanti marginali è prossimo a zero (poiché costi e benefici marginali sono uguali), conduce all'annullamento di un certo numero di relazioni efficienti.

Si noti comunque che, essendo il beneficio marginale crescente al diminuire del numero dei soggetti in relazione, si viene a determinare un livello di scambio non nullo, ma inferiore a quello efficiente, in corrispondenza del quale da ogni relazione si consegue un surplus significativamente maggiore di zero. La minaccia della perdita di questo surplus, mediante l'adozione di strategie di punizione bilaterale, è un deterrente sufficiente per scoraggiare ulteriori abusi della città. Quindi, sebbene una strategia di punizione bilaterale non sia in grado di garantire l'esecuzione contrattuale di tutti gli scambi efficienti, essa può servire a sostenere un certo livello limitato di scambi.²³

Questa analisi dimostra che i mercati in cui esistono problemi di enforcement non sono in grado di raggiungere le condizioni di efficienza allocativa (cfr. anche nel paragrafo 4.3 le analisi di Klein - Leffler, 1981, e Shapiro, 1983).

²³ Nel paragrafo 4.2 questo modello sarà ripreso per illustrare alcuni dei problemi inerenti all'enforcement multilaterale.

Enforcement bilaterale con orizzonte finito e informazione incompleta

In alcune situazioni è possibile realizzare l'enforcement sulla base di sanzioni bilaterali anche se l'orizzonte temporale degli agenti è finito e noto con certezza. A tal fine è necessario che le parti dispongano di un certo grado di incertezza (informazione incompleta) sulle effettive caratteristiche delle controparti. Questo importante risultato è stato dimostrato da Kreps - Milgrom - Roberts - Wilson (1982).

Si consideri un dilemma del prigioniero unilaterale (Figura 1), ripetuto per un certo numero T (finito) di periodi. Se esiste anche una piccola probabilità che l'agente 2 (che muove successivamente) non sia opportunistica (perché, ad esempio, i suoi pay-off sono diversi da quelli descritti), l'agente 1 potrebbe avere interesse ad accordargli fiducia. Il giocatore 2 può decidere di rispettare l'accordo anche se è opportunistica (sopportando quindi un costo come investimento in reputazione), per indurre il primo giocatore a credere di essere onesto, allo scopo di usufruire dei guadagni della cooperazione anche nei periodi successivi.

Con l'approssimarsi del periodo terminale del gioco, il giocatore opportunistica non ha più interesse a nascondere le proprie vere caratteristiche e decide di approfittare della fiducia concessa. Ciò avviene quando i guadagni dell'opportunismo risultano maggiori dei benefici attesi in futuro dalla cooperazione, i quali si riducono progressivamente con il passare del tempo. Si dimostra che il numero dei periodi finali (N) in cui il contratto non viene rispettato dipende inversamente dalla probabilità di imbattersi in un individuo onesto e dal fattore di sconto, ma non dipende dal numero totale dei periodi di gioco (T). Così, se T è grande, l'enforcement del contratto può essere realizzato anche per un elevato numero di periodi ($T-N$), superando il paradosso di Selten. Il limite di questa impostazione è costituito dal fatto che la cooperazione è assicurata solo fino a un certo periodo. Se T non è molto grande si defeziona in tutti o nella maggior parte dei periodi.²⁴

4. L'enforcement multilaterale

Come mostrato nel paragrafo precedente, esistono situazioni in cui l'enforcement bilaterale non è praticabile o non ha efficacia, poiché non c'è la possibilità di sanzionare la controparte o le sanzioni dirette non sono sufficienti a prevenire l'opportunismo. Ciò accade in tutti i casi in cui l'eventuale opportunistica può uscire senza costi da una relazione contrattuale, perché non ha sostenuto *sunk cost* o effettuato investimenti specifici e un *matching* con un nuovo partner è prontamente realizzabile, grazie al fatto che domanda e offerta sul mercato in questione sono in equilibrio e non sono necessari costi di ricerca. In secondo luogo, se la relazione tra le parti ha un orizzonte finito e noto, l'impossibilità di sanzionare eventuali inadempienze nello stadio finale conduce all'annullamento della cooperazione in tutti i periodi.

Nelle economie contemporanee, con elevata divisione del lavoro e specializzazione, la maggior parte delle transazioni ha natura breve e predeterminata e le controparti commerciali generalmente cambiano con una elevata frequenza. Inoltre, gli individui hanno un orizzonte di vita e di lavoro finito, anche se esistono istituzioni, come le imprese, il cui orizzonte è potenzialmente infinito.

In queste situazioni in cui o non esistono quasi-rendite per le parti o il rapporto tra le parti ha una durata finita possono emergere equilibri nei quali il rispetto degli accordi e la cooperazione tra le parti è sostenuta dalla minaccia di sanzionamento *inflitta all'esterno della relazione dagli altri agenti sul mercato o nella società*, a condizione però che le informazioni sugli accordi intrapresi e sul comportamento di ognuno siano accessibili all'esterno.

²⁴ Reny (1993) ottiene un risultato cooperativo in un dilemma del prigioniero con orizzonte finito in una impostazione che ammette la violazione dell'ipotesi di "common knowledge".

Questa categoria di meccanismi di enforcement endogeni è definita “multilaterale” (Greif, 1993) o “*external or market-based*” (Hart - Holmstrom, 1987) o ancora “*community enforcement*” (Kandori, 1992a). Altri autori fanno semplicemente riferimento all’idea di reputazione.²⁵

Il *community enforcement* può essere basato su sanzioni di duplice natura: (1) sanzioni di mercato, operate dagli altri potenziali partner che si rifiutano di stipulare accordi con coloro che in passato sono stati riconosciuti colpevoli di inadempienza; oppure (2) sanzioni sociali, nel senso che l’individuo subisce l’ostracismo degli altri nella società o comunità di appartenenza.

A livello di mercato, la sanzione più comune consiste, analogamente al caso dell’enforcement bilaterale in cui si interrompe la relazione, nell’astenersi dall’entrare in rapporti futuri con i soggetti che violano gli accordi contrattuali comportandosi opportunisticamente. La punizione per chi defeziona si configura ancora come un costo-opportunità, cioè la perdita di un flusso di quasi-rendite conseguibile continuando ad operare sul mercato.

La sanzione sociale è, invece, una vera e propria disutilità diretta, inflitta a colui che defeziona, ad opera dei membri del gruppo sociale al quale appartiene, disutilità derivante da qualche forma di comportamento ostile o stigma sociale o semplicemente dalla frustrazione dell’universale desiderio umano di essere socialmente accettati, rispettati e benvenuti all’interno della comunità in cui si vive (Sugden, 1989, p. 95).

Il concetto di reputazione

Per *reputazione*, nel suo significato più comune, si intende una valutazione o una aspettativa sul comportamento presente e futuro di un certo soggetto, effettuato da altri soggetti in relazione con esso, sulla base del suo comportamento osservato in passato. In termini più tecnici, la reputazione di un giocatore può essere definita come una funzione dalla storia passata a una distribuzione di probabilità sulle sue strategie (Greif, 1993, p. 530). Entro questa definizione si possono comprendere due categorie di reputazione.

La prima si ha in presenza di *informazione incompleta* su alcune caratteristiche di un soggetto: la reputazione è legata al fatto che le scelte di questi nel corso della relazione trasmettono informazioni sul suo essere di un determinato tipo, fornendo indicazioni sul suo comportamento futuro.

La seconda categoria fa riferimento ad un contesto con informazione completa, in cui la reputazione viene a coincidere con l’adozione continuata delle strategie di equilibrio cooperativo del gioco ripetuto, cioè si considera dotato di reputazione un soggetto che si comporta in modo da evitare il meccanismo che conduce al sanzionamento dell’opportunismo, messo in atto da parte dei soggetti presenti sul mercato attraverso l’adozione di particolari strategie punitive (tipicamente, il rifiuto di intraprendere affari in futuro).²⁶

Gli agenti impegnati in una relazione contrattuale possono dunque evitare di comportarsi opportunisticamente e mantenere una reputazione di affidabilità allo scopo di: (1) non trasmettere informazioni sfavorevoli sul proprio conto, che farebbero scattare reazioni negative degli altri; (2) non

²⁵ Sebbene in letteratura non manchi il ricorso a concetti come «reputazione interna» o «reputazione bilaterale» per descrivere le strategie decise nei rapporti bilaterali sulla base del comportamento passato della controparte, la definizione di reputazione è prevalentemente impiegata nei casi in cui è coinvolta una pluralità di soggetti esterni alla relazione che, sulla base dell’informazione diffusa nel mercato o nella comunità sociale di riferimento, stabilisce l’atteggiamento da tenere nei confronti delle parti. Tale significato appare più in linea con il linguaggio comune.

²⁶ Fudenberg - Tirole (1991) contestano l’uso del termine reputazione in quest’ultimo caso, poiché non in grado di dare conto del fatto che la reputazione è qualcosa che si apprende nel corso della relazione. Si è scelto di applicare il termine anche in questa accezione poiché gran parte della letteratura definisce reputazione anche le strategie di equilibrio in giochi con informazione completa (cfr. ad esempio, Carmichael, 1984 e 1989; Greif, 1993; Stiglitz, 1992; Klein - Leffler, 1981; Shapiro, 1983, Kreps, 1990).

incorrere nelle ritorsioni future degli altri giocatori, che attuano strategie condizionate alle azioni scelte. Il soggetto che nelle relazioni intrattenute non ha mai tradito la fiducia accordatagli è considerato «affidabile», e altri soggetti saranno disposti in futuro a instaurare con esso nuove relazioni economiche. Chi, invece, si comporta in modo opportunistico perde la reputazione e non può continuare a scambiare poiché gli viene negata la fiducia dai nuovi potenziali partner.

La reputazione è un meccanismo che induce a rinunciare ai vantaggi di breve periodo connessi ad un comportamento fraudolento allo scopo di continuare a stabilire anche in futuro relazioni vantaggiose. Pertanto, essa è mantenuta intatta solo fintantoché i benefici attesi in futuro da essa sono maggiori dei profitti di breve periodo conseguibili rovinandola.

I vantaggi dell'enforcement multilaterale rispetto al sistema legale

Una delle ragioni principali che spiegano la prevalenza dell'enforcement bilaterale sull'enforcement attuato dalla Corte è connessa, come si è visto, alla mancanza di verificabilità delle variabili rilevanti da parte della Corte stessa, in quanto parte esterna. Tuttavia, se l'osservabilità all'esterno della relazione è preclusa, neppure un meccanismo di enforcement multilaterale può essere attivato: l'inefficacia dell'enforcement multilaterale si accompagna a quello ad opera del sistema legale, poiché entrambi i meccanismi si basano sulle azioni di soggetti esterni che sono ostacolate dalla carenza di informazioni.²⁷ L'enforcement bilaterale, richiedendo che le informazioni rilevanti siano a disposizione della sola diretta controparte pone, da questo punto di vista, requisiti meno stringenti.

Tuttavia, ammesso che l'informazione filtri all'esterno, esistono diverse ragioni per cui l'esecuzione dei contratti mediante l'operare di soggetti privati può risultare più efficiente del sistema legale.

Una prima spiegazione si basa sulle competenze necessarie per formulare giudizi. Frequentemente, per giudicare adeguatamente i fatti occorre una certa preparazione, bisogna essere esperti o possedere conoscenze professionali specifiche ad un particolare campo. Mentre questo è possibile per agenti che si trovano nello stesso mercato o settore, per la Corte le difficoltà possono essere notevoli. In secondo luogo “poiché [il sistema legale] è di tipo generale, esso si basa su regole che possono risultare ben poco adatte al particolare settore nel quale è sorta la disputa” (Milgrom - Roberts, 1992, p. 396). Inoltre, mentre il cambiamento formale delle leggi e delle procedure legali – spesso necessario per effetto di innovazioni tecnologiche – richiede tempi molto lunghi, il sistema privato può mutare più rapidamente e mostrare maggiore flessibilità. In terzo luogo, l'enforcement multilaterale presenta dei vantaggi rispetto alla Corte quando permette di risparmiare gli eccessivi costi dei ricorsi legali o di evitare la lentezza dei procedimenti, l'aleatorietà del giudizio, il rischio di corruzione o condizionamento della Corte (Tirole, 1999).²⁸ Infine, nelle relazioni commerciali internazionali, le Corti potrebbero non avere nessuna autorità per imporre le proprie decisioni.

In certe situazioni, per far fronte ai problemi di intervento della Corte le parti ricorrono alla istituzione di un giudice o arbitro privato. Questa soluzione permette di alleviare alcuni problemi –

²⁷ Si noti tuttavia che spesso le relazioni non sono uno-a-uno, cioè ogni parte non ha solo una controparte, ma un numero N , maggiore di uno (uno-a-molti). Per esempio, sul mercato del lavoro l'impresa di frequente assume un certo numero di lavoratori, sul mercato dei beni, ogni impresa vende i suoi prodotti a una pluralità di consumatori. Questa circostanza rende realistico immaginare che le informazioni sulle singole relazioni bilaterali, ancorché di difficile verificabilità, possano essere osservate dagli altri agenti contemporaneamente in relazione.

²⁸ Come sostiene Tirole (1999, p. 761): “Judges are subject to both moral hazard (they may not put enough effort into reading and understanding the details of the case) and adverse selection (they may not have the proper background to understand what the parties tell them; they may also have their own preferences, in the form of legal precedents and principles, which they may embody in their decisions, sometimes regardless of what the contract says or what the parties want). Judges and arbitrators may also collude with the parties”.

sopperisce ai problemi di mancanza di esperienze professionali e di aggiornamento e garantisce una maggiore celerità – ma può essere molto costosa e non elimina il rischio di corruzione o collusione con qualcuna delle parti. Un giudice privato, inoltre, può mancare del potere coercitivo per imporre le decisioni assunte.

Il funzionamento dell'enforcement multilaterale

Il funzionamento dell'enforcement multilaterale richiede che i terzi esterni – i potenziali futuri partner o individui in relazione sociale – conoscano lo svolgimento della relazione, siano consapevoli degli accordi intercorsi tra le parti, siano in grado di valutare eventuali inadempienze, abbiano interesse a sanzionare la parte inadempiente e si aspettino che anche gli altri sanzioneranno. La realizzazione di questa forma di enforcement richiede quindi di: (1) raccogliere e disseminare le informazioni; (2) giudicare e interpretare i fatti; (3) coordinare le aspettative (*beliefs*) dei vari soggetti; (4) imporre le sanzioni.²⁹

Per quanto riguarda la disponibilità di informazioni si sono esaminate in precedenza le ragioni del vantaggio dei meccanismi di enforcement endogeni rispetto a quello esogeno, sia per la materiale disponibilità delle informazioni, che per la diversa capacità di comprensione e trattamento di esse e per la maggiore flessibilità.

Occorre comunque valutare se le soluzioni di sanzionamento decentralizzate sono in grado di fornire gli incentivi ai singoli agenti affinché essi abbiano effettivamente interesse a mettere in atto la punizione. Il problema degli incentivi può avere valenza differente per l'enforcement bilaterale e quello multilaterale.

Nell'enforcement bilaterale corrisponde sempre all'interesse della parte che viene ingannata troncarsi il rapporto con l'opportunist. In altri termini, una deviazione unilaterale dalla fase di sanzionamento non è conveniente poiché la controparte continuerebbe ad agire opportunisticamente: ciò implica che la minaccia di punizione è credibile e l'equilibrio è perfetto.

La situazione è più complicata nel meccanismo di enforcement multilaterale: poiché l'onere del sanzionamento degli opportunisti ricade sulle parti esterne, gli incentivi a sanzionare possono mancare. I costi per imporre il sanzionamento possono essere di due tipi: 1) costi di acquisizione delle informazioni sui potenziali nuovi partner; 2) costi derivanti dal rifiuto di intrattenere rapporti con un partner e costi di ricerca e matching di uno nuovo (questi costi dipendono anche dalla situazione sul mercato: se trovare un nuovo partner è semplice, essi possono essere irrilevanti, e viceversa).

Dato che il sanzionamento è costoso, si può avere interesse a non eseguire la punizione, nell'aspettativa che siano gli altri a punire. Ma se tutti si comportano in questo modo, le minacce non sono credibili (l'equilibrio non è perfetto nei sottogiochi) e l'esecuzione degli accordi è impossibile.³⁰

²⁹ Come afferma Williamson (1985, p. 574-5), «[g]li effetti di reputazione possono dissuadere dal disattendere la lettera e lo spirito di un accordo, nella misura in cui 1) gli inadempimenti possono essere resi di pubblico dominio, 2) le conseguenze dell'inadempimento possono essere completamente accertate (il che consente, fra l'altro, di distinguere le genuine pretese di indennizzo per inadempimento da quelle artefatte) e 3) le parti che subiscono o rilevano l'inadempimento penalizzano completamente il trasgressore e/o i suoi subentranti».

³⁰ Si consideri una economia in cui esistono N agenti che giocano un dilemma del prigioniero bilaterale, dopodiché si separano e ognuno trova un nuovo partner (il quale è informato del comportamento precedente). La semplice strategia che prevede di cooperare con chi in passato ha cooperato e di non scambiare con chi prima ha invece ingannato, potrebbe non garantire l'equilibrio, anche con un basso tasso di sconto. Nel sentiero fuori dall'equilibrio, l'agente che si imbatte in un precedente opportunist, dovrebbe rinunciare ai guadagni della cooperazione, ma se tutti gli altri seguono la strategia prescritta, continuando a punire chi defeziona, si può perdonare l'imbroglione precedente, riabilitando il precedente opportunist e questi avrà interesse a non imbrogliare. Ma tale deviazione unilaterale, poiché annulla la punizione, distrugge l'equilibrio cooperativo.

In pratica, la messa in atto del sanzionamento nell'enforcement multilaterale si configura come un «bene pubblico»: nel complesso ognuno beneficia del sanzionamento praticato dagli altri, ma per il singolo esiste un costo nell'eseguire la punizione e quindi un incentivo a comportarsi da free rider.

Tale problema non si presenta quando c'è informazione incompleta, in quanto l'opportunismo passato rivela caratteristiche negative che inducono automaticamente i nuovi partner a non accordare fiducia.

Allo scopo di far fronte alla possibilità di deviazione dal sanzionamento, nei mercati emergono particolari istituzioni o organizzazioni ausiliarie all'attività di sanzionamento degli agenti privati, che svolgono funzioni di raccolta e diffusione delle informazioni e di coordinamento dei "Principali". In alcuni contesti, è necessaria l'applicazione di una meta-norma che impone di sanzionare anche coloro che non sanzionano.

Nei paragrafi successivi si esaminano i principali modelli con enforcement multilaterale proposti nella letteratura economica allo scopo di evidenziare i contesti economici in cui questo tipo di enforcement è applicato e le particolari modalità di funzionamento nelle varie situazioni.

4.1. I principali modelli con *enforcement* multilaterale

I numerosi lavori che analizzano soluzioni di enforcement multilaterale sono tutti caratterizzati dall'assunzione che ogni singola relazione che si svolge tra le parti abbia durata finita. Essi tuttavia differiscono sulla base di un aspetto non secondario quale la lunghezza dell'orizzonte degli agenti, che usualmente non coincide con la durata delle relazioni economiche. I lavori esistenti possono essere, pertanto, classificati in tre gruppi:

- 1) modelli con orizzonte infinito per tutti gli agenti e relazioni di durata finita a causa del *turnover* tra i vari partner;
- 2) modelli con relazioni finite poiché agenti a vita finita interagiscono con soggetti aventi orizzonte infinito;
- 3) modelli con agenti tutti a vita finita.

4.2. Relazioni finite fra agenti a vita infinita

Salari di efficienza e reputazione dei lavoratori: il modello di Greif

Il modello presentato da Greif (1993), simile nell'impostazione a quello di Shapiro - Stiglitz (1984), è utile per evidenziare la differenza di funzionamento esistente tra l'enforcement multilaterale e quello bilaterale.

Nel modello di Shapiro e Stiglitz è assente qualsiasi meccanismo reputazionale poiché le imprese non conoscono le precedenti esperienze dei lavoratori. Greif (1993) dimostra che l'enforcement multilaterale è più efficiente dell'enforcement bilaterale se ogni impresa: a) osserva o viene a conoscenza delle performance degli agenti impiegati anche dalle altre imprese; b) segue la strategia di non assumere i lavoratori che in precedenza si sono comportati opportunisticamente (*multilateral punishment strategy*); c) si aspetta che anche le altre imprese seguano questa strategia.³¹

³¹ La (c) è una condizione necessaria in quanto se la politica delle assunzioni delle altre imprese non è condizionata alla storia precedente dei lavoratori, l'informazione a disposizione dell'impresa è virtualmente inutile, poiché le opportunità alternative dei lavoratori – che sono decisive per l'enforcement – rimangono inalterate. In un lavoro successivo, Greif (1994) dimostra che il passaggio da un sistema basato su strategie di punizione bilaterali a uno con strategie di punizione multilaterali richiede, oltre alla disponibilità informativa, il

Il modello intende descrivere una precisa esperienza storica, cioè le relazioni commerciali tra mercanti magrebini dell'Undicesimo secolo e i loro agenti all'estero sui vari mercati del Mediterraneo. Il problema centrale è quello di prevenire l'opportunismo dell'agente, che può appropriarsi dei beni del mercante o del suo capitale.

Greif mostra come la «coalizione» formata da tutti i mercanti si configura come una istituzione utile per l'enforcement contrattuale. Gli stretti legami comunitari che uniscono i mercanti permettono una rapida e accurata diffusione dell'informazione e i valori culturali comuni rendono agevole l'interpretazione degli avvenimenti e il coordinamento delle aspettative.

La reputazione degli agenti viene mantenuta finché essi operano correttamente nell'ambito della loro occupazione corrente. Colui che non rispetta l'accordo viene sanzionato immediatamente con il licenziamento e, in seguito, con la perdita della reputazione e l'emarginazione (punizione collettiva).

Greif dimostra che tale strategia di punizione multilaterale costituisce un equilibrio perfetto poiché corrisponde all'interesse di ogni principale portare a termine la minaccia rifiutandosi di impiegare un agente che si è comportato male in passato, se tutti gli altri adottano la strategia di emarginazione: data la natura di relazioni di lavoro con durata definita, chi impiega un agente che ha la reputazione rovinata presso la collettività dovrebbe offrirgli un salario elevatissimo per assicurarsi la sua onestà, dato che l'agente non ha più molto da perdere. Da ciò deriva la non convenienza a intrattenere rapporti di affari con chi subisce l'ostracismo degli altri.³²

Un risultato fondamentale del modello è che una strategia di punizione multilaterale (MPS) (senza considerare i costi per la trasmissione delle informazioni), garantisce sempre risultati migliori, dal punto di vista dell'efficienza, di una strategia di punizione bilaterale (BPS). Infatti, è possibile dimostrare che il salario di efficienza da pagare se l'unica punizione disponibile è il licenziamento (BPS) è sempre maggiore del salario che bisogna garantire se la sanzione è costituita, oltre che dal licenziamento, anche dall'impossibilità di occupazioni future.³³ Per il lavoratore il costo atteso dell'opportunismo è molto più alto con una MPS e quindi il salario che induce all'onestà è più basso. La differenza tra i due livelli salariali è tanto più accentuata quanto maggiore è il turnover nei rapporti di lavoro. Poiché un salario più elevato scoraggia le imprese dall'impiego di lavoratori se il surplus è inferiore al premio salariale, con una BPS si hanno maggiori perdite di efficienza.

Il modello di Greif (1993) rappresenta una descrizione dei mercati del lavoro «trasparenti» in cui, per la natura del lavoro, le informazioni sulle performance dei lavoratori e sui rapporti di lavoro si diffondono abbastanza facilmente nel mercato e le imprese decidono le assunzioni sulla base della reputazione dei lavoratori. In questa situazione, l'enforcement contrattuale può essere realizzato anche in presenza di un elevato turnover dei lavoratori.

Il Folk Theorem nel matching game e l'equilibrio «contagioso»

Al contrario del modello di Greif, l'analisi effettuata da Kandori (1992a) è astratta ed esclusivamente tecnica. Egli presenta una estensione e una generalizzazione del *Folk Theorem*, valido per i giochi ripetuti standard, ai giochi ripetuti con *matching*, cioè quelli in cui il partner di ogni giocatore (a vita

coordinamento di tutti i «principali». Senza coordinamento, nessuno ha individualmente interesse a mutare strategia.

³² Fama (1980) e Holmstrom (1999) presentano dei modelli in cui la prestazione dei lavoratori è osservata sul mercato ed essi sono indotti ad impegnarsi poiché la remunerazione futura e le prospettive di carriera (all'interno o all'esterno dell'impresa) dipendono dalla loro performance corrente.

³³ La prova è immediata se si suppone che le relazioni durino per un solo periodo. Con una BPS, ogni lavoratore eluderà l'impegno, per qualsiasi livello del salario corrisposto, perché non subisce nessuna reale sanzione. Al contrario, se è in atto una MPS, esiste un salario definito che assicura che l'agente rispetterà il contratto, per evitare la perdita della reputazione e l'annullamento delle opportunità di lavoro future.

infinita) può cambiare in ogni periodo (in ciascun periodo ogni agente è accoppiato casualmente con un altro). L'analisi definisce le condizioni necessarie per l'efficacia dell'enforcement tramite le reazioni degli agenti esterni alla relazione (*community enforcement*).

Il *community enforcement* è retto da una *norma sociale* che stabilisce quali sono i comportamenti che i soggetti sono tenuti ad osservare, quali sono le sanzioni in caso di deviazione e chi è incaricato di sanzionare.³⁴ Kandori (1992a) esamina il funzionamento del meccanismo di enforcement sotto tre diverse ipotesi sulla disponibilità delle informazioni per le parti esterne.

La prima situazione è quella con trasmissione perfetta dell'informazione. Nonostante ogni relazione bilaterale abbia luogo solo per un periodo, Kandori dimostra che se tutti gli agenti nella comunità ricevono informazioni perfette su ogni rapporto e quindi la storia di ogni agente è pubblicamente nota, l'enforcement degli accordi cooperativi, attuato mediante il ritiro della cooperazione futura da parte di ogni nuovo partner ai danni del deviante, è pienamente efficace e supplisce alle difficoltà operative del *personal enforcement*.

Come in tutti i meccanismi con imposizione «diffusa» delle sanzioni è necessario prevedere un meccanismo di sanzionamento per chi non sanziona, poiché rifiutare la cooperazione a un precedente opportunisto è costoso per i nuovi partner: il meccanismo che induce questi ultimi a non deviare è basato su una applicazione ricorsiva della minaccia originaria, cioè chi non sanziona viene a sua volta sanzionato.³⁵

L'ipotesi che ognuno possa osservare con cura quanto avvenuto nelle relazioni che coinvolgono altri soggetti è estrema. La situazione opposta esaminata da Kandori (1992a) è quella in cui gli agenti possiedono un set informativo estremamente povero: ognuno conosce solo la storia dei propri incontri, non è in grado di osservare gli altri rapporti e nessuna comunicazione tra gli agenti è possibile. In questo secondo scenario, a condizione che il tasso di sconto sia sufficientemente basso e la popolazione non troppo numerosa, un equilibrio di cooperazione è ancora raggiungibile, se si attua la minaccia di un processo di defezione definito «contagioso». Alla base dell'equilibrio contagioso è presente il seguente schema di punizione: ogni agente che dovesse subire una defezione comincerebbe dal turno successivo a comportarsi in modo opportunistico indiscriminatamente con tutti i partner che incontra. Siccome anche questi ultimi reagiscono allo stesso modo, la defezione si diffonde rapidamente, come una malattia contagiosa, attraverso tutta la popolazione.

In questo contesto la fiducia è legata all'intera società, piuttosto che al singolo individuo. Ognuno sa che approfittare nel presente della cooperazione altrui porta come risultato ultimo alla distruzione della cooperazione in tutta la comunità, e quindi anche se la ritorsione non può essere immediata (poiché nessuno dei nuovi partner osserva alcunché), in un futuro più o meno lontano anche i propri partner si comporteranno egoisticamente.

L'efficacia di una minaccia di questo tipo dipende, oltre che dall'usuale tasso di sconto, anche dalla numerosità della popolazione. Infatti, se il numero dei componenti è molto elevato, dopo una

³⁴ Alcuni autori usano il termine di norma sociale per far riferimento all'enforcement multilaterale e di contratto implicito per l'enforcement bilaterale.

³⁵ Un risultato negativo relativamente all'efficacia dell'enforcement multilaterale è ottenuto nell'articolo di Bendor - Mookherjee (1990). Nel loro modello esistono agenti a vita infinita che intrecciano contemporaneamente e ripetutamente relazioni bilaterali. L'informazione sulle azioni intraprese in qualsiasi rapporto è immediatamente nota a tutti. Essi mostrano che, nonostante il *community enforcement* sia in linea teorica una alternativa disponibile, in giochi standard come il dilemma del prigioniero la minaccia delle sanzioni esterne non svolge alcun ruolo. Ma la causa di questi sorprendenti risultati è la particolare struttura del gioco proposto dagli autori: in esso, tutti gli agenti sono contemporaneamente in relazione bilaterale gli uni con gli altri. La minaccia di sanzioni future da parti esterne non è credibile, poiché si presta a una ritorsione presente: il giocatore minacciato da reazioni future da altri è in grado di adeguarsi immediatamente, adottando nei loro confronti comportamenti ostili. In pratica, in questa struttura non esistono delle parti effettivamente esterne.

defezione la probabilità di incontrare un soggetto «contagiato» dall'opportunismo è ridotta, e occorre un lungo periodo di tempo prima che tutti arrivino a giocare la strategia di defezione: nel frattempo, è possibile approfittare della cooperazione di tutti quelli che si incontrano. Ma la consapevolezza di ciò conduce alla completa distruzione della cooperazione fin dall'inizio.

Un equilibrio quale quello contagioso ha altre proprietà indesiderabili: da un lato, è inefficiente ed iniquo, poiché comporta una punizione anche per i soggetti "innocenti"; dall'altro, è estremamente instabile, poiché un piccolo disturbo o errore, facendo scattare la strategia punitiva, portano alla completa scomparsa della cooperazione.

Tra i due casi estremi di diffusione delle informazioni, Kandori individua un caso intermedio, in cui è all'opera un meccanismo o istituzione che permette agli agenti di accedere a una specifica informazione sulla storia passata del proprio partner («*local information processing*»), che costituisce la sua «label». L'istituzione, dopo ogni rapporto, aggiorna la *label* degli agenti. La *label* permette, anche in assenza di informazioni accurate sul comportamento passato di ogni membro, di sapere se il partner è affidabile oppure no. Un esempio di *label* è il possesso di una particolare licenza o di una carta di credito, o l'appartenenza a qualche organizzazione. L'adozione delle usuali appropriate strategie garantisce un equilibrio cooperativo anche in questo caso. L'effettivo funzionamento dell'enforcement con «*label*» è illustrato sotto, con applicazione a uno specifico caso economico.

Il ruolo delle istituzioni per la diffusione delle informazioni e il coordinamento degli agenti

Una situazione simile a quella studiata da Kandori nell'equilibrio «contagioso» e in quello con «*label*» è quella verificatasi tra i mercanti delle fiere medievali del Dodicesimo e Tredicesimo secolo (le Fiere della *Champagne*), analizzata da Milgrom - North - Weingast (1990).

Nel loro modello – molto utile per esaminare il ruolo della disponibilità informativa nell'applicazione dell'enforcement multilaterale – esiste un numero elevato N di agenti a vita infinita che in ogni periodo entrano in relazione a due a due e giocano un dilemma del prigioniero. Ogni agente incontra in successione tutti gli altri, dopodiché il giro ricomincia. Esiste, quindi, un continuo turnover, anche se ogni agente è destinato ad incontrare lo stesso partner ogni N periodi (i risultati sono equivalenti se il matching avviene casualmente). L'informazione a disposizione di ogni membro è solo quella relativa ai propri incontri passati.

Poiché l'orizzonte è infinito, è in linea teorica possibile adottare strategie che prevedono in caso di defezione le usuali sanzioni bilaterali di interruzione della cooperazione: chi viene ingannato attua la ritorsione contro l'opportunist in tutti i loro prossimi incontri, cioè ogni N periodi.

Tuttavia, si può facilmente dimostrare (usando i *pay-off* della Tabella 1) che il tasso di sconto necessario per l'enforcement bilaterale tende esponenzialmente a zero, al crescere della popolazione:³⁶

$$r \leq \left[\frac{\Pi^M - \Pi^D}{\Pi^M - \Pi^C} \right]^{\frac{1}{N}} - 1$$

Per popolazioni numerose l'enforcement bilaterale diventa completamente privo di efficacia se il tasso di sconto è positivo, poiché a fronte dei guadagni immediati della defezione, le penalizzazioni future avvengono in tempi troppo lontani per servire allo scopo.

Si può dimostrare inoltre che anche l'adozione delle strategie che caratterizzano l'equilibrio «contagioso» – cioè di non cooperare mai più con nessuno dopo avere subito una defezione – conduce alle stesse condizioni dell'equazione precedente. Pertanto, al crescere della numerosità degli agenti viene meno sia la possibilità di un enforcement bilaterale che l'equilibrio contagioso.

³⁶ Infatti, se nei confronti di un particolare giocatore si coopera sempre si ottiene: $\Pi^C + \Pi^C (1/r)^N$. Defezionare invece comporta: $\Pi^M + \Pi^D (1/r)^N$.

Sulla base di questa situazione, Milgrom - North - Weingast (1990) spiegano l'emergere delle Corti mercantili nelle economie del Medioevo, ma soprattutto le ragioni dell'esistenza di alcune istituzioni contemporanee che svolgono funzioni analoghe. La Corte mercantile è una istituzione che svolge il ruolo di giudice privato, anche se è priva di poteri coercitivi, e ha il compito di raccogliere le informazioni dei vari mercanti, di formulare un giudizio su quello che è successo nelle dispute e di comunicare queste informazioni ai mercanti che le richiedono. Con l'ausilio di questa istituzione, è possibile stabilire un sistema di enforcement con sanzionamento multilaterale basato sulle strategie dei singoli agenti, rispettando condizioni molto meno stringenti dell'equazione precedente.

Come nel meccanismo della *label* presente in Kandori (1992a), l'istituzione è incaricata unicamente di diffondere le informazioni e coordinare le risposte dei mercanti, ma non di mettere direttamente in pratica sanzioni. Il vantaggio di questo sistema rispetto all'enforcement legale è che gli agenti privati sono in grado di stabilire le norme più adatte al particolare settore in cui operano, di dotarlo delle necessarie professionalità e adeguarlo rapidamente ad ogni nuova esigenza. Il limite che caratterizza i sistemi di arbitri privati, cioè la mancanza del potere di sanzionamento, viene superato facendo ricorso alle sanzioni inflitte direttamente dagli agenti, che evitano di entrare in relazione con coloro che sono stati giudicati inadempienti.

Le istituzioni private di questo tipo non sono una peculiarità della storia economica, ma sono presenti in varie forme anche negli attuali mercati, nei settori in cui l'informazione è fondamentale, ma defluisce con difficoltà dai singoli che instaurano una relazione al mercato nel suo complesso (si veda Klein, 1997). Esempi di questo tipo di istituzioni sono le agenzie di *rating* e i *credit bureau*, le associazioni di consumatori che pubblicano propri bollettini informativi, gli albi e gli ordini professionali, le associazioni commerciali, le agenzie che rilasciano licenze,³⁷ ecc. Queste istituzioni raccolgono e diffondono le informazioni rilevanti sul comportamento delle parti in relazione, mentre lasciano il sanzionamento all'iniziativa dei singoli.

Landa (1994) argomenta convincentemente che anche gli intermediari commerciali (rivenditori, dettaglianti) svolgono un ruolo cruciale per l'enforcement. Ponendosi tra i produttori e gli acquirenti, gli intermediari stabiliscono relazioni ripetute con entrambi e «creano un ponte di fiducia» tra questi soggetti che avrebbero altrimenti rapporti poco frequenti: "By providing trust, the middleman reduces costs associated with search, getting-to-know-you, negotiation, enforcement, anxiety, and so on. The middleman mediates between traders who do not trust each other but mutually trust the middleman. The middleman also acts as a knower - by recognizing brand names and seals of approvals, by conducting his own tests and inspections, by keeping track of customer complaints and purchase returns, by hiring testing services, by following trade literature or consumer literature, by studying whether other retailers carry the product and so on" (Landa, 1994, p. 201).

Il problema dell'enforcement multilaterale analizzato nei modelli di Kandori (1992a) e di Milgrom - North - Weingast (1990) è relativo alla disponibilità delle informazioni per le parti esterne: una istituzione che raccoglie e processa le informazioni e coordina l'intervento dei singoli può essere necessaria se questi ultimi non sono perfettamente informati.

Nell'articolo di Greif - Milgrom - Weingast (1994) il problema che si affronta non è solo quello informativo, ma anche quello di rendere credibile la minaccia di sanzionamento, cioè di assicurare che

³⁷ Nei mercati in cui si richiede una licenza (o una certificazione) per operare, il possesso della licenza - oltre ad assicurare che la qualità è adeguata e che in passato non si sono rilevati comportamenti fraudolenti - genera una rendita, poiché l'accesso al mercato viene limitato. La perdita di questa rendita connessa alla sottrazione della licenza agisce da deterrente contro comportamenti opportunistici.

tutte le parti esterne abbiano interesse a realizzare il sanzionamento se si presenta la necessità, nei casi in cui esiste un elevato costo opportunità per farlo.

Nel consueto stile che contraddistingue questo gruppo di autori, il modello proposto nasce dal tentativo di spiegare e interpretare il funzionamento di particolari istituzioni emerse nel corso della storia. Le caratteristiche del modello di Greif, Milgrom e Weingast sono state esposte nel paragrafo 3.1, in cui si è evidenziato che l'enforcement degli accordi tra una città e i mercanti che vi si recano, basato su strategie di sanzionamento bilaterali, anche se riesce a sostenere un certo numero di scambi, non può garantire il raggiungimento del livello efficiente.

Un altro risultato fondamentale del modello è che anche l'enforcement multilaterale è inefficace, se privo di istituzioni di supporto o di coordinamento. Tale meccanismo poggia sul rifiuto di entrare in relazione da parte di tutti i mercanti, se la città ha agito opportunisticamente ai danni di qualcuno di essi. In altri termini, occorre attuare un *embargo* contro la città. La ragione dell'inefficacia è duplice: innanzitutto, le incertezze e le ambiguità che possono sorgere sull'avvenuto rispetto dei contratti o su eventuali violazioni o su quali erano gli intendimenti contrattuali rendono difficoltosa l'opera dei soggetti esterni, che possono avere differenti interpretazioni su ciò che è avvenuto; in secondo luogo, quando viene dichiarato un embargo, esistono forti incentivi per ogni singolo mercante a non attuare la minaccia di interruzione dei traffici commerciali, poiché la città può offrire condizioni molto favorevoli a chi rompe l'embargo. Infatti, a un livello ridotto di scambi, il beneficio marginale per la città di ogni relazione diventa molto elevato, e parte di questi benefici possono essere offerti ai mercanti. In aggiunta, l'efficacia delle sanzioni bilaterali a livelli ridotti di commercio garantisce i mercanti che scambiano nonostante l'embargo da un'ulteriore defezione del governo.

Per assolvere alle due funzioni necessarie al sanzionamento multilaterale, storicamente sono emerse le *gilde mercantili* (tra le quali la famosa Hansa o Lega anseatica), che sono associazioni di mercanti che forniscono una interpretazione degli avvenimenti e rendono chiaro quando l'inesco della sanzione è necessario e, soprattutto, mettono in atto delle ritorsioni nei confronti degli agenti che vengono meno alla fase punitiva continuando ad intrattenere relazioni commerciali con le città opportuniste. La possibilità di sanzionare i singoli mercanti deriva dal fatto che l'affiliazione alla gilda rende ad ogni mercante una quasi-rendita (poiché gli permette di sfruttare posizioni monopolistiche, ecc.), che sarebbe annullata in caso di comportamento scorretto.

Recentemente Hogan (2001) – prendendo spunto dai lavori di MacLeod - Malcomson (1989) e Greif - Milgrom - Weingast (1994) – ha mostrato che tale struttura è applicabile al mercato del lavoro in cui il sindacato agisce come istituzione che comunica ai lavoratori le informazioni sull'affidabilità delle imprese e coordina l'eventuale sanzionamento, riuscendo ad assicurare un livello di occupazione più elevato.

4.3. Relazioni tra agenti a vita infinita con agenti a vita finita

Diversamente dai modelli considerati finora con agenti tutti a vita infinita, in questo paragrafo si analizzano alcuni modelli che risolvono il problema dell'enforcement nelle relazioni tra un soggetto a vita perenne e una parte contrattuale a vita finita. La caratteristica fondamentale di questi modelli – che permette di superare il paradosso di Selten – è la rappresentazione della relazione come dilemma del prigioniero *unilaterale*: solo una delle parti ha la possibilità di mettere in atto comportamenti opportunistici, e questa parte è fatta coincidere con il soggetto a vita infinita.³⁸ Quest'ultimo, poiché continua a rimanere sul mercato, è interessato a mantenere una buona reputazione – non commettendo

³⁸ Tipicamente, si analizzano relazioni tra una impresa (soggetto considerato a vita infinita) e i suoi lavoratori o consumatori.

azioni scorrette – a condizione che la permanenza sul mercato consenta di guadagnare una quasi-rendita.

Le relazioni tra impresa e consumatori quando la qualità non è osservabile

Klein - Leffler (1981) e Shapiro (1983) forniscono delle applicazioni di meccanismi reputazionali di questo genere riferite al mercato dei beni. In questi modelli esiste una impresa con orizzonte infinito che entra in relazione con una miriade di piccoli consumatori che possono essere considerati a vita finita.³⁹

L'impresa ha la possibilità di fornire un bene di qualità elevata oppure uno di qualità scadente, a causa dell'impossibilità da parte dei consumatori – o, *a fortiori*, di un Tribunale – di osservarne contestualmente la qualità (*bene esperienza*). La qualità del prodotto è però osservabile dal consumatore con un periodo di ritardo e le informazioni sulla qualità sono trasmesse correttamente e senza ritardo tra tutti i consumatori.

Il costo di produzione del bene di buona qualità (c_h) è maggiore di quello del bene scadente (c_s). Il prezzo p è superiore al costo in entrambi i casi: $p > c_h > c_s$. L'utilità per il consumatore è positiva se il bene è di qualità elevata, $(u_h - p) > 0$, e negativa se il bene è di qualità scadente, $(u_s - p) < 0$. L'utilità e i profitti sono nulli se impresa e consumatori non effettuano alcuno scambio.

L'enforcement bilaterale del contratto che prevede la vendita di un bene di qualità elevata è impossibile a causa dell'orizzonte finito di ogni relazione. I singoli consumatori non possono credibilmente attuare alcuna minaccia nei confronti dell'impresa che dovesse ingannarli.

Al contrario, se attraverso comunicazioni tra i consumatori, l'informazione su ogni rapporto diventa nota, l'applicazione del contratto può essere basata sulla reputazione dell'impresa. La strategia dei consumatori consiste semplicemente nell'aspettativa che la qualità fornita dall'impresa in futuro sia uguale a quella fornita in passato e quindi nella decisione di continuare ad acquistare solo se si prevede che la qualità sia elevata.

Data questa strategia, l'impresa ha due alternative:

- (1) offrire sempre qualità elevata, mantenendo una buona reputazione che consente di avere anche in futuro consumatori disposti a comprare, ottenendo:

$$(p - c_h)Q \left[1 + \frac{1}{1+r} + \left(\frac{1}{1+r} \right)^2 + \dots \right] \text{ (dove } Q \text{ rappresenta la quantità venduta), oppure}$$

- (2) ingannare tutti i consumatori, guadagnando un profitto di breve periodo più elevato, ma rinunciando ai profitti futuri: $(p - c_s)Q$;

E' agevole dimostrare che un equilibrio in cui si realizza l'enforcement del contratto (in cui l'impresa vende beni di qualità elevata e le aspettative dei consumatori sono realizzate) può esistere solo se il prezzo è adeguatamente superiore al costo di produzione:

$$p \geq c_h + r(c_h - c_s)$$

Questa condizione conferma il risultato generale che stabilisce che il contratto è rispettato solo se chi deve adempiere ottiene una rendita. La strategia dei consumatori si configura come una minaccia di interruzione della rendita se l'impresa si comporta scorrettamente.

Questa analisi ha delle conseguenze interessanti anche per il funzionamento dei mercati concorrenziali. Come è possibile conciliare la necessità di mantenere un prezzo più elevato del costo, per garantire la qualità, con la libertà di entrata in un mercato concorrenziale che spinge il prezzo uguale ai costi e tende ad annullare i profitti? In altri termini, l'equilibrio impone che una impresa

³⁹ L'impresa può essere rappresentata come un giocatore *long-run* che entra in relazione con una sequenza infinita di giocatori *short-run*, ognuno dei quali vive per un solo periodo (cfr. Fudenberg - Tirole, 1989, p.169).

potenziale entrante nel mercato abbia la prospettiva di profitti nulli e che, nello stesso tempo, le imprese già presenti sul mercato conseguano un profitto positivo. Questa duplice condizione è soddisfatta solo se per entrare nel mercato le imprese devono sostenere un investimento specifico (*sunk cost*), di ammontare pari ai profitti conseguibili successivamente grazie al divario tra prezzo e costo. In tal modo, gli incentivi dell'impresa a comportarsi onestamente sorgono dal timore di perdere l'investimento effettuato se si rovina la reputazione e si è costretti ad uscire dal mercato. Due esempi di tale investimento sono le spese in pubblicità non informativa (Nelson, 1974), oppure la necessità che nelle fasi iniziali il bene sia venduto ad un prezzo inferiore al costo (Shapiro, 1983).

Sulla base di questo modello sono interpretabili istituzioni come il marchio (*brand name*) o le catene di negozi in franchising. Il nome dell'impresa che unifica oggetti e strutture differenti amplia gli strumenti di rivalsa a disposizione di ogni consumatore. Si vedano Akerlof (1970), Klein (1997) e Cabral (2000).

Le relazioni tra impresa e lavoratori e la valutazione della performance soggettiva

L'idea della reputazione legata ad un soggetto a vita infinita nelle sue relazioni con partner con orizzonti più brevi è stata applicata anche da Bull (1987), Baker - Gibbons - Murphy (1994) e Kreps (1990), come interpretazione della relazione tra impresa e lavoratori.⁴⁰

Quando la prestazione lavorativa è particolarmente complessa e non è verificabile dall'esterno, ma può essere valutata soggettivamente dall'impresa stessa ("subjective performance evaluation"), un modo comune per incentivare l'impegno dei lavoratori è di collegare la remunerazione alla performance ottenuta ("performance related pay"), ad esempio attraverso il pagamento di un bonus o con una promozione discrezionale.

Come nel modello di MacLeod - Malcomson (1998), il problema di questa struttura contrattuale è l'opportunismo dell'impresa, che potrebbe rinnegare la promessa, dichiarando ex-post che la performance ottenuta è stata scadente.⁴¹ Bull (1987) e Baker - Gibbons - Murphy (1994) illustrano una soluzione alternativa a quella basata sull'esistenza di posti di lavoro vacanti, che fa affidamento sulla reputazione dell'impresa presso le generazioni coesistenti dei lavoratori: in caso di inganno ai danni di qualche lavoratore, infatti, l'impresa rovinerebbe la propria reputazione nei confronti degli attuali e potenziali lavoratori, i quali non sarebbero disposti a lavorare presso di essa o ad accordarle fiducia per future intese contrattuali. Il contratto implicito è rispettato dall'impresa se i guadagni futuri attesi dalla relazione lavorativa sono superiori ai vantaggi immediati derivanti dal risparmio dell'incentivo salariale. Baker - Gibbons - Murphy (1994) determinano endogenamente la dimensione del bonus che può essere garantito attraverso la reputazione, mostrando, ad esempio, che se il tasso di interesse è alto o le quasi-rendite future sono ridotte, il vincolo di compatibilità degli incentivi è più stringente e solo un livello basso del bonus può essere corrisposto, incoraggiando così solo uno sforzo limitato del lavoratore.⁴²

⁴⁰ Simon (1951) considera la relazione tra reputazione dell'impresa e uso dell'autorità.

⁴¹ "Where performance is subjective and performance pay thus not enforceable in court, firms have an incentive to renege on a bonus at the end of the period" (MacLeod - Malcomson, 1998, p. 389).

⁴² Come soluzione al problema dell'azzardo morale del lavoratore, in alternativa ai salari di efficienza o alla remunerazione collegata alla performance, Lazear (1981) ha studiato il meccanismo dei salari crescenti con l'anzianità (*life-cycle incentive*), in cui al lavoratore nelle fasi iniziali della carriera si pagano salari minori della produttività, per diventare maggiori a partire da un certo periodo, accoppiato alla minaccia di licenziamento in caso di shirking. Analogamente alla struttura con valutazione soggettiva della performance, tale meccanismo si presta all'opportunismo dell'impresa – che potrebbe licenziare il lavoratore quando il salario diventa maggiore della produttività – e richiede un sistema di reputazione dell'impresa.

L'interazione tra contratti espliciti e impliciti

I rapporti di lavoro rappresentano relazioni complesse, caratterizzate da una molteplicità di variabili, alcune delle quali pienamente verificabili e altre solo osservabili. Dal punto di vista dell'enforcement contrattuale, la presenza contemporanea di elementi osservabili e verificabili dà luogo a interessanti interazioni tra contratti espliciti e impliciti (questi ultimi applicati tramite meccanismi reputazionali), analizzate recentemente nei lavori di Bernheim - Whinston (1998), Holmstrom - Milgrom (1991), Baker (1992) e Baker - Gibbons - Murphy (1994).

Il contributo di Bernheim - Whinston (1998) sull'"ambiguità strategica", strettamente collegato ai lavori di MacLeod - Malcomson (1989) e di Holmstrom - Milgrom (1991), offre una spiegazione originale dell'esistenza di incompletezza contrattuale, non originata da razionalità limitata o da costi di transazione. Bernheim - Whinston (1998) dimostrano che nelle situazioni in cui alcune misure della performance non sono verificabili, le parti possono decidere deliberatamente di non includere nel contratto disposizioni su altri aspetti che sono, invece, verificabili. La parte implicita di un accordo può funzionare meglio se il set delle scelte discrezionali risulta più ampio, attraverso la mancata specificazione di aspetti verificabili della relazione. I "gap" contrattuali così originati conferiscono flessibilità al contratto e permettono di adattare ex-post le prestazioni sulla base di variabili osservabili (ma non verificabili), permettendo di premiare discrezionalmente una buona performance (o di sanzionarne una cattiva).⁴³ Il meccanismo è per molti versi analogo a quello analizzato da MacLeod - Malcomson (1989) per i contratti in cui si rinuncia a fissare esplicitamente l'intero salario e l'impresa concede un bonus discrezionale sulla base dell'osservazione dello sforzo del lavoratore.

Mentre nel modello di Bernheim - Whinston (1998) è conveniente per gli agenti lasciare dei gradi di libertà nel contratto per poter meglio condizionare alcuni aspetti al comportamento osservato della controparte, in altre analisi si evidenzia come variabili verificabili non vengono regolate nei contratti per evitare che ciò dia luogo a comportamenti distorti o controproducenti da parte del lavoratore.

Il modello del multitasking di Holmstrom - Milgrom (1991) rappresenta una varietà di situazioni in cui un lavoratore deve svolgere un insieme complesso di compiti (task), alcuni dei quali sono verificabili mentre altri non lo sono. In questo contesto, essi dimostrano che nonostante per lo svolgimento di alcuni task possano essere utilizzati incentivi espliciti (ad esempio, la remunerazione legata a una variabile oggettivamente misurabile), può essere preferibile non adottare schemi di questa natura per evitare "comportamenti distortivi" da parte del lavoratore, che potrebbe dare eccessivo peso alle variabili verificabili a detrimento delle funzioni per le quali non esiste un meccanismo di incentivazione esplicita. Così, alcune azioni utili potrebbero essere completamente ignorate dall'agente, mentre altre potrebbero risultare sovraincentivate.

Un tale problema si riscontra in tutti i casi nei quali è molto importante la qualità del prodotto oppure è necessario un uso diligente del capitale fisico. Incentivare l'impegno del lavoratore con una remunerazione dipendente dalla quantità prodotta, poiché relativamente facile da misurare, può condurre ad effetti deleteri sulla qualità o sull'uso del capitale poiché tali aspetti verrebbero trascurati dal lavoratore in quanto non misurabili.

In tali circostanze, può risultare ottimale un contratto con un salario fisso piuttosto che collegato a qualche misura della performance. Più in generale, in occupazioni con multitasking è preferibile l'uso di incentivi "a basso potenziale".

Nello stesso spirito di Holmstrom - Milgrom (1991), Baker (1992) evidenzia che nei rapporti di lavoro in cui le misure disponibili della performance possono essere distorte può essere opportuno

⁴³ La convenienza a lasciare aspetti fuori dal contratto si riscontra solo in un contesto dinamico, mentre non vale in un gioco statico.

evitare contratti con incentivi espliciti o limitare il peso di questi ultimi. Nel modello è rappresentata la distinzione tra una misura della performance verificabile p e il vero contributo dato all'impresa dal lavoratore y , non osservabile o verificabile. Il sistema di incentivazione risulta problematico se l'impatto dello sforzo del lavoratore e sulle due variabili è diverso. Sia $y=\theta e$ e $p=\mu e$ dove θ e μ rappresentano parametri osservabili in ogni periodo dal lavoratore, ma non dall'impresa. Se la remunerazione è condizionata a p , si ottengono comportamenti distorti nella misura in cui θ e μ risultano differenti. Per l'impresa l'impegno è efficiente quando θ assume valori elevati, mentre il lavoratore è indotto ad impegnarsi maggiormente quando μ è elevato. Da tale situazione scaturisce la scarsa convenienza dell'impresa a fornire incentivi espliciti quando θ e μ tendono ad essere diversi.

Baker - Gibbons - Murphy (1994) notando che molte imprese impiegano sia contratti espliciti che contratti con valutazione soggettiva, studiano le conseguenze per l'enforcement derivanti dall'interazione tra queste due tipologie di contratti. Da un lato, i contratti espliciti possono dar luogo a distorsioni, come mostrato da Baker (1992) e da Holmstrom - Milgrom (1991), ma consentono di limitare il bonus discrezionale da pagare al lavoratore. Dall'altra parte, i contratti impliciti permettono di remunerare il lavoratore sulla base di una migliore valutazione (soggettiva) della performance, ma sono soggetti al rischio di appropriazione dell'impresa.

Baker - Gibbons - Murphy (1994) mostrano che possono presentarsi due situazioni opposte a seconda del valore assunto dai parametri del modello: 1) in un primo caso, l'esistenza di buoni contratti espliciti, grazie a misure oggettive della performance poco distorte, può compromettere l'applicazione dei contratti impliciti, poiché rende troppo allettante per l'impresa la situazione successiva alla defezione dalla promessa del bonus; 2) in un caso opposto, i due tipi di contratto risultano complementari, nel senso che una combinazione dei due conduce alla produzione di un surplus positivo, mentre sarebbe impossibile l'implementazione della sola parte implicita o della sola parte esplicita.⁴⁴

4.4. Orizzonte di vita finito per tutti gli agenti

I lavori che presentano soluzioni con enforcement multilaterale esposti in precedenza compiono un importante passo in avanti sulla strada di una maggiore aderenza delle costruzioni teoriche alla realtà, costruendo dei modelli in cui le relazioni tra gli agenti non durano all'infinito, ma sono ben delimitate nel tempo e si verificano frequenti cambiamenti di partner. In essi, tuttavia, si continua ad assumere che la vita di uno o di tutti gli agenti sia virtualmente infinita, e questa ipotesi può essere molto limitativa.

Una serie di modelli innovativi riesce a superare questo limite coniugando la finitezza della relazione tra due partner con quella dell'orizzonte temporale degli agenti. Le soluzioni adottate sono analoghe a quelle proposte da Kreps - Milgrom - Roberts - Wilson (1982) in un contesto con informazione incompleta sulle caratteristiche degli agenti (paragrafo 3.1). Altri lavori più recenti mostrano che attraverso strategie di sanzionamento multilaterale la cooperazione è un risultato possibile anche con informazione completa. Essi sfruttano l'idea che anche se l'orizzonte è finito, il periodo di vita di ogni agente si *sovrappone* a quello degli altri, nel senso che in ogni periodo esistono generazioni giovani che coesistono con generazioni più anziane (*overlapping generations*), e quindi non esiste un periodo terminale comune, bensì scaglionato.

⁴⁴ Una analisi analoga sull'uso contemporaneo di strumenti espliciti e impliciti è quella di Pearce - Stacchetti (1998) che enfatizzano la distinzione tra variabili osservabili e verificabili all'interno di un contratto di agenzia.

Sulla base di questa caratteristica, la gran parte dei modelli presenta soluzioni in cui nella fase finale della vita ad ogni agente è assegnato, da parte dei giocatori più giovani, un incentivo o premio, se la sua condotta precedente è stata sempre cooperativa. Nell'aspettativa del premio futuro, i giovani sono così incentivati a cooperare (e a tener fede alla promessa di premiare gli anziani). In pratica, i benefici che si conseguono con l'approssimarsi della fine della vita svolgono il ruolo che nei modelli con agenti a orizzonte infinito è svolto dal flusso continuo di quasi-rendite. Ma vediamo in maggior dettaglio l'impostazione seguita nei modelli più rappresentativi.

Kreps (1990) fornisce una spiegazione dell'esistenza dell'impresa delineando innanzitutto una struttura, analoga a quella proposta da Bull (1987), caratterizzata da una impresa a vita infinita in cui una generazione di lavoratori (ma un ragionamento analogo si applica ai rapporti con i clienti o con le imprese fornitrici) le accorda fiducia solo se essa nei rapporti con le generazioni precedenti ha sempre agito correttamente. Successivamente, l'analisi viene sviluppata considerando che l'impresa è pur sempre controllata e gestita da individui a vita finita e dimostrando la possibilità di un equilibrio di cooperazione anche quando sia i lavoratori che i proprietari-manager hanno vita finita.

In questa impostazione, la cooperazione si realizza grazie al fatto che il valore derivante dal possedere una buona reputazione è trasferibile da una generazione di proprietari ad un'altra. L'impresa è lo strumento che permette di compiere questo trasferimento e come tale essa rappresenta una attività intangibile, un semplice capitale di reputazione.

Ogni lavoratore considera la reputazione dell'impresa e decide di accordare ad essa fiducia solo se in passato si è comportata onestamente con tutti. Fintantoché l'impresa gode di questa fiducia, essa ha un preciso valore di mercato, derivante dalla sua capacità di generare reddito grazie alle relazioni instaurabili con i lavoratori, presenti e futuri. Se i lavoratori perdono la fiducia nell'impresa, il valore di quest'ultima si annulla, poiché mai più nessuno è disposto ad intrattenere rapporti con essa.

Ipotizzando che esistano dei soggetti che vivono almeno due periodi, nel primo dei quali acquistano l'impresa sul mercato e mettono in atto la relazione con il lavoratore, e nel secondo vendono l'impresa e consumano i guadagni accumulati, è facile dimostrare che i contratti tra agenti a vita finita sono sempre rispettati (se il tasso di sconto è sufficientemente basso). Nessuno dei manager ha interesse ad approfittarsi del lavoratore, poiché in questo caso distruggerebbe il valore dell'impresa.⁴⁵ L'enforcement è così realizzato grazie sia alle strategie adottate dai lavoratori, sia a quelle dei potenziali acquirenti dell'impresa, che non sono disposti ad acquistare una impresa che ha ingannato.⁴⁶

Sempre in questo lavoro Kreps si sofferma a lungo su un altro aspetto centrale per l'enforcement: la possibilità di stabilire con chiarezza se le disposizioni contrattuali sono state rispettate o meno. Il problema sorge quando è impossibile prevedere gli eventi futuri (*unforeseen contingencies*) e, quindi, non si possono stabilire le azioni che le parti sono tenute a svolgere in determinate circostanze. La «cultura dell'impresa» (*corporate culture*) può aiutare a fare chiarezza, costituendo un paradigma di riferimento (per i lavoratori, i consumatori, e tutti gli altri «*stakeholders*»), cioè un insieme di principi e regole di comportamento seguite usualmente dall'impresa, che serve a definire come verranno affrontate le situazioni per le quali non ci sono

⁴⁵ L'idea di Kreps dell'impresa come portatrice di reputazione è presente anche nel lavoro di Williamson (1985, p. 575): "Un'impresa caratterizzata da una certa struttura di proprietà/management non può eludere le penalizzazioni imposte alla struttura precedente chiedendo clemenza: le colpe dei padri vengono invece fatte ricadere sui figli, con la conseguenza che le valutazioni delle risorse di un'impresa rifletteranno costantemente il comportamento precedente".

⁴⁶ Recentemente, Tadelis (1999) e Mailath - Samuelson (2001) hanno sviluppato l'analisi di Kreps, analizzando il mercato dei "nomi" delle imprese, in presenza di selezione avversa e di monitoring imperfetto, spiegando come la reputazione può formarsi (e perdersi) gradualmente.

previsioni nei contratti espliciti e cosa si debbano aspettare le parti. La *corporate culture* permette, quindi, di giudicare *ex-post* se il comportamento seguito è stato in linea con gli accordi e di adottare le eventuali contromisure.

Nel modello di Cremer (1986) esiste una impresa o organizzazione che dura infinitamente all'interno della quale operano un numero T di agenti che giocano un dilemma del prigioniero collettivo, scegliendo simultaneamente se cooperare o defezionare. In essa in ogni periodo entra un nuovo agente, che vive per T periodi, dopo i quali esce definitivamente di scena. Pertanto, in ogni periodo l'impresa è costituita da individui tutti di età differente, da quello più anziano che si appresta a partecipare al suo ultimo gioco a quello appena entrato che resterà per altri T periodi.

La cooperazione comporta per il singolo un costo pari ad e ; se cooperano un numero n di giocatori, l'output aggregato prodotto è $Y = n\Pi^C$. L'output è diviso tra tutti i T giocatori. Si assume che $\Pi^C > e$. Se n giocatori stanno cooperando, l' i -esimo giocatore ottiene: $\left[\frac{(n+1)\Pi^C}{T} - e \right]$ se coopera; $n\Pi^C/T$ se defeziona.

Se il gioco è *one-shot*, conviene sempre defezionare, a meno che il singolo contributo non sia estremamente produttivo (cioè si coopera solo se $\Pi^C > eT$). Al contrario, se il gioco è ripetuto infinitamente, si è in grado di ottenere un comportamento cooperativo in tutti i periodi di vita dei giocatori tranne l'ultimo se tutti seguono una *trigger strategy* così congegnata: defezionare nel periodo T , cooperare in tutti i periodi precedenti a meno che non si osserva una defezione di qualcuno più giovane di T , nel quale caso defezionare sempre.

La generazione più vecchia T ha sicuramente interesse a seguire questa strategia. Ma anche alla generazione di età $(T-1)$ conviene cooperare, se tutti seguono la strategia prescritta. Se coopera, nei due periodi restanti ottiene (assumendo un tasso di sconto nullo): $\left[\frac{(T-1)\Pi^C}{T} - e \right] + \left[\frac{(T-1)\Pi^C}{T} \right]$, mentre se defeziona: $\left[\frac{(T-2)\Pi^C}{T} \right] + 0$. La cooperazione conviene poiché, per ipotesi, $\Pi^C > e$.

Si può facilmente dimostrare che, con condizioni ancora meno restrittive, conviene cooperare anche in tutti i periodi precedenti. Il premio che conduce alla cooperazione è la facoltà che gli agenti concedono al più anziano dell'organizzazione di defezionare ed avvantaggiarsi degli sforzi altrui.

Una generalizzazione del risultato di Cremer è stata fornita attraverso il lavoro, esclusivamente tecnico, di Kandori (1992b) in cui si presenta un *Folk Theorem* per giochi ripetuti all'infinito tra individui con orizzonte finito, ma con periodi di vita che si sovrappongono. Kandori dimostra che è possibile ottenere equilibri cooperativi con *pay-off* efficienti, a condizione che il periodo di vita degli agenti e quello di coesistenza con gli altri siano sufficientemente lunghi.

Durante il periodo di coesistenza, le generazioni più anziane, maggiormente tentate dall'opportunismo, sono minacciate di ritorsioni dai più giovani se il loro comportamento non è stato quello desiderabile e sono premiate se invece hanno agito correttamente (essi sono anche tenuti a sanzionare coloro che avevano deviato dalla cooperazione).

Salant (1991) ottiene un *Folk Theorem* analogo a quello di Kandori facendo esplicito riferimento ad un gioco con un periodo finale non produttivo (ad esempio, il periodo di pensionamento dopo l'attività lavorativa) durante il quale i giocatori possono essere premiati oppure sanzionati a seconda del comportamento tenuto nei precedenti periodi di attività.

Un modello con generazioni coesistenti strutturalmente simile a quello di Kandori, ma con una strumentazione tecnica notevolmente più semplice, è stato proposto da Guttman (1992).

Gli agenti entrano nella comunità, per ipotesi molto vasta, in tempi diversi e vi restano per T periodi. Nessuno, a parte lo stesso individuo, conosce però il momento di ingresso e, di conseguenza, per tutti è ignoto il momento del termine del rapporto (*unknown endpoint*). In ogni periodo ogni agente gioca un dilemma del prigioniero con un partner scelto casualmente. Le azioni scelte possono essere osservate da tutti gli altri giocatori.

Guttman dimostra che un particolare equilibrio cooperativo in cui gli individui cooperano in tutti i periodi, con l'unica eccezione del periodo finale in cui defezionano, può essere sostenuto dalla seguente strategia «reputazionale»: nei periodi precedenti all'ultimo, si coopera con chi in passato si è comportato correttamente, mentre ci si rifiuta di giocare con chi ha agito opportunisticamente.⁴⁷ Pertanto, chi imbrogliando perde immediatamente la reputazione e non ha più possibilità di stabilire relazioni vantaggiose. Chi coopera corre invece un certo rischio, che consiste nella possibilità di incontrare (con probabilità $1/T$) un partner giunto al suo periodo finale e di subire il suo opportunismo. Se i benefici della cooperazione sono sufficientemente alti, gli agenti scelgono di cooperare e di affrontare questo rischio.

Hirschleifer - Rasmusen (1989) studiano le strategie in grado di determinare la cooperazione tra un gruppo di N individui a vita finita (T periodi), impegnati congiuntamente in un gioco del dilemma del prigioniero. Questo modello è l'unico in cui non esistono generazioni coesistenti.

La misura che gli agenti adottano come forma di sanzionamento contro comportamenti opportunistici è «l'ostracismo», cioè la pratica di espellere gli individui colpevoli dalle interazioni con il gruppo dei restanti membri. Il punto cruciale di questa impostazione è che subire il sanzionamento mediante ostracismo risulta più oneroso della mancata cooperazione degli altri agenti. L'ostracismo di un individuo è però costoso anche per tutto il gruppo che continua a cooperare, poiché si suppone che esistono «economie di aggregazione», sia nella sfera produttiva che in quella sociale, che fanno crescere i benefici di ognuno al crescere del numero dei partecipanti al gruppo.

Le strategie che i giocatori seguono prescrivono di cooperare in tutti i periodi precedenti l'ultimo e di ostracizzare chi non coopera o chi non ostracizza gli opportunisti. Chi ostracizza non sopporta un costo diretto, ma solo indiretto, derivante dalla riduzione della cooperazione congiunta, diffuso su tutti i membri che cooperano. La mancata ostracizzazione però è molto costosa, perché chi si astiene dal farlo sarà a sua volta ostracizzato.

Nell'ultimo periodo la cooperazione tra i membri viene a mancare completamente, poiché nessuno può essere più punito. Ma questo fatto – lungi dall'impedire l'emergere della cooperazione nei periodi precedenti come nel paradosso di Selten – consente di rendere credibile la minaccia di ostracizzare chi non coopera. Infatti, nell'ultimo periodo non è costoso sanzionare, poiché comunque nessuno coopera, e attraverso l'induzione a ritroso la credibilità di questa minaccia si può far risalire fino al primo periodo. Hirschleifer e Rasmusen dimostrano che con queste strategie la cooperazione tra i membri può essere ottenuta in tutti i periodi eccetto quello finale.

Il problema relativo a questo modello è che nella realtà può mancare la possibilità – senza l'intervento di una autorità – di imporre una sanzione che vada al di là, e sia più costosa, del rifiuto della cooperazione. Una punizione ulteriore rispetto alla non cooperazione è plausibile se legata agli aspetti sociali, come argomentato nel paragrafo successivo.

4.5. Enforcement sociale

Le relazioni economiche non si svolgono in un «vuoto», come spesso ipotizzato nei modelli economici, ma all'interno di un denso tessuto sociale. I contraenti sono legati ad altre persone – da relazioni familiari, di amicizia, di vicinato, comunitarie (all'interno di gruppi etnici, religiosi, ecc.) – che osservano, giudicano e reagiscono ai loro comportamenti. Secondo Granovetter (1985) le transazioni economiche sono "incorporate" (*embedded*) nella struttura sociale.

⁴⁷ Chi deve sanzionare non sopporta alcun costo poiché si assume che colui che rifiuta di giocare con un precedente opportunistica possa essere accoppiato immediatamente con un altro giocatore.

La perdita della reputazione di lavoratore o venditore onesto presso l'ambiente sociale in cui si vive, della benevolenza e stima dei colleghi, dell'amicizia, di una serie di relazioni sociali utili o piacevoli possono essere forme ugualmente importanti e dolorose di punizione, spesso sufficienti a scoraggiare la messa in atto di comportamenti fraudolenti.

In molti casi, l'enforcement sociale richiede che l'utilità dei soggetti sia influenzata da variabili diverse da quelle legate semplicemente al benessere materiale (come il reddito o la disutilità dello sforzo), quali la considerazione sociale o la benevolenza degli altri individui nei rapporti sociali.⁴⁸ Come spiega Polanyi, 1944: "He [man] does not act so as to safeguard his individual interest in the possession of material goods; he acts so as to safeguard his social standing, his social claims, his social assets. He values material goods only in so far as they serve this end".

Alcuni dei modelli esaminati in precedenza presentano casi in cui la punizione per gli opportunisti non consiste necessariamente nella mancata cooperazione nella sfera produttiva. Guttman (1992) accenna alla distinzione tra giochi produttivi e giochi sociali e sostiene che comunità animate da spirito cooperativo possono fornire un punto focale per la scelta degli equilibri di cooperazione; in Hirschleifer - Rasmusen (1989), l'ostracismo è una penalizzazione che va oltre la mancata cooperazione.

Ma a parte fugaci accenni in passato gli economisti non hanno dedicato molta attenzione a fenomeni di questo genere, oggetto più specifico di studio di sociologi, psicologi e politologi (cfr., ad esempio, Coleman, 1990; Putnam, 1993; Fukuyama, 1995). Recentemente, tuttavia, si è assistito a una inversione di tendenza e numerosi lavori hanno analizzato l'enforcement sociale in vari contesti. Ad esempio, ha riscosso grande interesse la politica della Grameen Bank, la famosa banca del Bangladesh, che richiede ai prenditori a prestito di raggrupparsi in gruppi di cinque membri provenienti dallo stesso villaggio, ritenendo il gruppo congiuntamente responsabile per la restituzione dei prestiti: le relazioni sociali all'interno del gruppo favoriscono l'applicazione del contratto (cfr., per esempio, Besley - Coate 1995).

Un approccio che promette di essere particolarmente fruttuoso è quello basato sul concetto di *linked games* (cfr. Spagnolo, 1999; Bernheim - Whinston, 1990), che rappresentano un insieme di individui impegnati, contemporaneamente o sequenzialmente, in un gioco produttivo e in un gioco sociale. La cooperazione nella fase produttiva può essere assicurata dalla minaccia di sanzionamento nel gioco sociale, nonostante gli agenti siano a vita finita.⁴⁹

Il collegamento tra aspetti sociali e politici e *performance* economica è analizzato nel lavoro di Putnam (1993). Egli fa esplicito riferimento all'influenza della *civic community* o *civicness* sulle possibilità di enforcement degli accordi, e quindi sullo sviluppo economico e politico. Le comunità che sono ricche di *civicness* – definita come "patterns of social cooperation based on tolerance, trust and widespread norms of active citizen participation" – dispongono di un «capitale sociale»: "le reti di impegno civico, come le associazioni di quartiere, i cori, le cooperative, i circoli sportivi, i partiti di massa... sono l'espressione di interazioni orizzontali e rappresentano una componente essenziale del capitale sociale.[..] Queste reti di impegno civico accrescono la potenziale sanzione che il trasgressore

⁴⁸ L'enforcement sociale ha caratteristiche simili all'enforcement morale (paragrafo 2) anche se esiste una profonda differenza tra i due meccanismi. Nell'enforcement morale si rispettano spontaneamente gli impegni sottoscritti poiché si ritiene giusto comportarsi in questo modo; nell'enforcement sociale, gli impegni non sono elusi solo perché si teme la reazione degli altri in campo sociale. La distinzione è evidente se si pone in relazione con il contesto informativo: se non ci sono possibilità che un atto opportunistico venga scoperto, l'enforcement morale continua ad essere pienamente efficace, e gli individui si asterranno comunque dall'approfittare, anche se protetti dallo schermo informativo; al contrario, in questi casi l'enforcement sociale non ha validità e si agisce opportunisticamente, sapendo che la reazione degli altri non si manifesterà in assenza di informazioni diffuse (si veda Kreps, 1997).

⁴⁹ Si veda anche Ben-Porath (1980), Kranton (1996), Clay (1997), La Ferrara (1998).

deve pagare in caso di defezione. Cedere all'opportunismo significa quindi mettere a repentaglio i benefici che ci si aspetta di ottenere da ogni altra transazione in cui si è impegnati, oltre che quelli potenzialmente derivanti da future contrattazioni.... Le reti di impegno civico incrementano l'iterazione e l'interconnessione dei giochi stessi" (Putnam, 1993, p. 204). "I cittadini che interagiscono in molteplici contesti sociali tendono a sviluppare comportamenti accettabili da tutti. [...] Le reti facilitano le comunicazioni e migliorano il flusso delle informazioni riguardanti l'affidabilità di una persona, in quanto la reputazione di ciascuno può essere conosciuta più facilmente e con maggiore precisione" (ibidem). Sul ruolo del capitale sociale si veda anche Knack - Keefer (1997).⁵⁰

Un risultato comune che emerge da queste analisi è che l'enforcement sociale si configura spesso come una esternalità positiva, poiché i network informali basati su relazioni personali facilitano gli scambi economici senza essere esplicitamente diretti a tale scopo.

5. Conclusioni

Le asimmetrie informative, i costi di transazione e la razionalità limitata degli agenti costituiscono le cause fondamentali che impediscono la stipula di un contratto completo tra le parti e ostacolano l'intervento della Corte diretto a far rispettare gli accordi contrattuali. In queste circostanze, le relazioni economiche possono essere minate dal rischio di opportunismo. Tuttavia, come conferma una ampia evidenza empirica, allo scopo di prevenire l'opportunismo e realizzare transazioni efficienti, le parti sono spesso in grado di delineare schemi contrattuali con enforcement endogeno, che basano il loro funzionamento sulla prospettiva di interazioni dinamiche tra gli stessi agenti in relazione (enforcement bilaterale) o tra gli agenti sul mercato (enforcement multilaterale). La continuità delle relazioni permette di sanzionare i comportamenti opportunistici o di premiare quelli cooperativi e garantisce l'applicazione dei contratti, a condizione che il valore attuale dei benefici di lungo periodo superi i vantaggi immediati derivanti dall'opportunismo.

La letteratura ha mostrato che il meccanismo di enforcement effettivamente adottato dagli agenti dipende: 1) dalla disponibilità di informazioni da parte degli agenti interni ed esterni alla relazione; 2) dalla misura e dal genere di costi che gli agenti sono in grado di imporre sugli opportunisti; 3) dall'interesse a imporre le sanzioni quando queste sono necessarie.

L'enforcement bilaterale, realizzato con strategie messe in atto dalle dirette controparti, tipicamente di interruzione della relazione, richiede che le informazioni siano disponibili solo tra i soggetti in relazione. In questo contesto, la parte lesa ha in genere interesse a mettere in atto la sanzione per evitare che la controparte continui a comportarsi opportunisticamente. L'efficacia delle sanzioni dipende innanzitutto dalla continuità della relazione – l'orizzonte temporale deve essere potenzialmente molto lungo – e dall'assenza di un termine certo del rapporto. In secondo luogo, per la realizzazione dell'enforcement è necessario un flusso di quasi-rendite derivante dalla continuazione della relazione, generate da investimenti specifici, costi di *switching* o situazioni di squilibrio tra domanda e offerta. Da questa condizione scaturisce anche una relazione tra l'applicabilità di una certa forma contrattuale (ad esempio, un salario fisso piuttosto che legato alla performance) e la situazione di eccesso di domanda o di offerta del mercato.

A differenza del meccanismo di natura bilaterale, il funzionamento dell'enforcement multilaterale impone la diffusione delle informazioni anche agli agenti che non sono coinvolti direttamente nella relazione. Il comportamento delle parti deve essere osservato all'esterno, interpretato e comunicato. Inoltre, è necessario fornire incentivi agli agenti esterni per mettere in atto

⁵⁰ Nelle società antiche, la limitata ampiezza della comunità, la fede negli stessi valori, l'affinità e gli stretti legami rendevano prevalente l'enforcement sociale.

le sanzioni e realizzare il coordinamento delle aspettative sulle strategie da impiegare. In compenso, in questo caso l'enforcement è efficace nonostante la durata delle relazioni sia breve e predeterminata e l'orizzonte temporale degli agenti sia finito. Inoltre, le quasi-rendite non sono associate a una particolare relazione, ma derivano dalla permanenza sul mercato.

In molti casi istituzioni private ausiliare (come le associazioni di consumatori o di imprese, le agenzie di rating, gli albi professionali, i sistemi di arbitraggio privati, ecc.) sono necessarie per raccogliere e comunicare le informazioni, fornire una interpretazione degli avvenimenti e risolvere il problema del coordinamento e dell'applicazione delle sanzioni per gli eventuali comportamenti devianti. In questa ottica può essere interpretata anche l'istituzione impresa, che svolge il ruolo di agente con orizzonte infinito per il quale è cruciale la reputazione nei rapporti con le generazioni coesistenti di agenti con orizzonti più brevi (i lavoratori, i consumatori, gli investitori, ecc.). Inoltre, la realizzazione dell'enforcement dipende anche da aspetti culturali, storici, sociali, etnici. Le norme sociali agiscono da supporto delle relazioni di mercato, contribuendo a stabilire il comportamento più consono da adottare e a sostenere il sanzionamento degli opportunisti, specialmente quando le sanzioni economiche sono poco efficaci.

L'enforcement multilaterale tende pertanto a prevalere rispetto all'enforcement bilaterale quando è contenuto il costo della diffusione delle informazioni relativamente agli accordi e alle azioni intraprese, se il coordinamento degli agenti esterni è agevole e se questi ultimi hanno interesse a sanzionare gli opportunisti. Inoltre, il meccanismo multilaterale è più efficace quando sui mercati la mobilità (turnover) è elevata e quindi la durata delle specifiche relazioni è relativamente breve e quando sono modesti i costi di *switching* e di *matching*.

Nonostante i meccanismi di enforcement endogeno permettano la realizzazione di transazioni mutuamente vantaggiose, essi non sono esenti da costi o inefficienze, poiché il funzionamento dei mercati risulta in questi casi profondamente diverso dai mercati concorrenziali. In assenza dell'enforcement legale, gli agenti possono realizzare scambi solo con soggetti ben identificati, non anonimi, ad essi legati da relazioni durature o membri della stessa comunità (commerciale o sociale). Avviare nuovi contratti relazionali è problematico e richiede in primo luogo investimenti in acquisizione di informazioni nonché in attività specifiche o il sostenimento di costi di turnover. Per queste ragioni, sul mercato si instaurano relazioni potenzialmente di durata prolungata e la perfetta mobilità dei fattori viene meno; i mercati, separati da elevate barriere all'entrata, assumono caratteristiche non concorrenziali. Le economie tendono ad essere segregate e la divisione del lavoro risulta limitata.

Inoltre, i mercati non svolgono solo funzioni allocative ma anche una funzione di disciplina dei comportamenti (Bowles - Gintis, 1993; Stiglitz, 1992). Essi sono caratterizzati generalmente da assenza di *market-clearing*, cioè domanda e offerta non sono in equilibrio e il prezzo non si aggiusta per annullare i disequilibri; di conseguenza, gli agenti sono spesso razionati, cioè non possono effettuare tutte le transazioni desiderate. Il prezzo risulta diverso dal costo marginale, il salario corrisposto dalle imprese è maggiore del salario di riserva. Le rendite in questi mercati sono perciò pervasive e comportano perdite di efficienza, cioè "gains from trade" non sfruttati.

Riferimenti bibliografici

- Akerlof G. (1970), The Market for «Lemons»: Quality Uncertainty and the Market Mechanism, *Quarterly Journal of Economics*, vol. 84, n. 3, pp. 488-500.
- Alt, J. - Martin, L. (1994), Contracting and the Possibility of Multilateral Enforcement, *Journal of Institutional and Theoretical Economics*, vol. 150, n. 1, pp. 265-71.

- Baker G. (1992), Incentive Contracts and Performance Measurement, *Journal of Political Economy*, vol. 100, n. 3, pp. 598-614.
- Baker G. - Gibbons R. - Murphy K. (1994), Subjective Performance Measures in Optimal Incentive Contracts, *Quarterly Journal of Economics*, vol. 109, n. 4, 1125-56.
- Baker G. - Gibbons R. - Murphy K. (2002), Relational Contracts and the Theory of the Firm, *Quarterly Journal of Economics*, vol. 117, n. 1, pp. 39-84.
- Banerjee A. - Duflo E. (2000), Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry, *Quarterly Journal of Economics*, vol. 115, n. 3, pp. 989-1017.
- Becker G. - Stigler G. (1974), Law Enforcement, Malfeasance and Compensation of Enforcers, *Journal of Legal Studies*, vol. 3, n. 1, pp. 1-18.
- Bendor J. - Mookherjee D. (1990), Norms, Third-Party Sanctions and Cooperation, *Journal of Law, Economics and Organization*, vol. 6, n. 1, pp. 33-63.
- Benoit J. - Krishna V. (1985), Finitely Repeated Games, *Econometrica*, vol 53, n. 4, pp. 905-22.
- Ben-Porath Y. (1980), The F-Connection: Families Friends and Firms and the Organization of Exchange, *Population and Development Review*, vol. 6, n. 1, pp. 1-30.
- Bernheim D. - Whinston M. (1990), Multimarket Contact and Collusive Behavior, *Rand Journal of Economics*, vol. 21, n. 1, pp.1-26.
- Bernheim D. - Whinston M. (1998), Incomplete Contracts and Strategic Ambiguity, *American Economic Review*, vol. 88, n. 4, pp. 902-32.
- Bernstein L. (1992), Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry *Journal of Legal Studies*, vol. 21, n. 1, pp. 115-57.
- Besley T. - Coate S. (1995), Group Lending, Repayment Incentives and Social Collateral, *Journal of Development Economics*, vol. 46, n. 1, pp. 1-18
- Bowles S. - Gintis H. (1993), The Revenge of Homo Economicus: Contested Exchange of Political Economy, *Journal of Economic Perspectives*, vol. 7, n. 1, pp. 83-114.
- Bull C. (1987), The Existence of Self-Enforcing Implicit Contracts, *Quarterly Journal of Economics*, vol. 102, n. 1, pp. 147-159.
- Cabral L. (2000), Stretching Firm and Brand Reputation, *Rand Journal of Economics*, vol. 31, n. 4, pp. 658-673.
- Carmichael L. (1984), Reputations in Labor Market, *American Economic Review*, vol. 74, n. 4, pp. 713-25.
- Carmichael L. (1985), Can Unemployment Be Involuntary? Comment, *American Economic Review*, vol. 75, n. 5, pp. 1213-4.
- Carmichael L. (1989), Self-Enforcing Contracts, Shirking, and Life Cycle Incentives, *Journal of Economic Perspectives*, vol. 3, n. 4, pp. 65-83.
- Clay K. (1997), Trade Without Law: Private-Order Institutions in Mexican California, *Journal of Law, Economics, and Organizations*, vol. 13, n. 1, pp. 202-231.
- Coase R. (1937), The Nature of the Firm, *Economica*, vol. 16, n. 4, pp. 386-405.
- Coleman J. (1990), *Foundations of Social Theory*, Harvard University Press, Cambridge (Mass.).
- Cremer, J., 1986), Cooperation in Ongoing Organization, *Quarterly Journal of Economics*, vol. 101, n. 1, pp. 33-50.
- Dasgupta P. (1988), La fiducia come bene economico, in Gambetta, D. (a cura di), *Le strategie della fiducia. Indagini sulla razionalità della cooperazione*, Einaudi, Torino, pp. 63-93.
- Dawes R. - Thaler R. (1988), Anomalies: Cooperation, *Journal of Economic Perspectives*, vol. 2, n. 3, pp. 187-97.
- De Soto H. (1989), *The Other Path*, New York, Harper & Row.
- Ellickson R. (1991), *Order without Law*, Cambridge: Harvard University Press.
- Ellickson R. (1989), A Hypothesis of Wealth-Maximizing Norms: Evidence from the Whaling Industry, *Journal of Law, Economics and Organizations*, vol. 5, n. 1, pp. 83-97.
- Fafchamps M. (1996), The Enforcement of Commercial Contracts in Ghana, *World Development*, vol. 24, n. 3, pp. 427-48.
- Fama E. (1980), Agency Problems and the Theory of the Firm, *Journal of Political Economy*, vol. 88, n. 2, pp. 288-307.
- Friedman J. (1971), A Non Cooperative Equilibrium for Supergames, *Review of Economic Studies*, vol. 38, n.113, pp. 1-12.

- Fudenberg D. - Maskin E. (1986), The Folk Theorem in Repeated Games with Discounting and with Incomplete Information, *Econometrica*, vol. 54, n. 3, pp. 533-54.
- Fudenberg D. - Tirole J. (1991), *Game Theory*, MIT Press, Cambridge (Ma).
- Fudenberg D. (1992), Explaining Cooperation and Commitment in Repeated Games, in Laffont J. J. (a cura di), *Advances in Economic Theory*, Cambridge University Press, pp. 89-129.
- Fukuyama F. (1995), *Trust*, The Free Press, New York.
- Gintis H. (1999), Some implications of Endogenous Contract Enforcement for General Equilibrium Theory, *International School of Economic Research*, Siena, July 1999.
- Granovetter M. (1985), Economic Action and Social Structure: The Problem of Embeddedness, *American Journal of Sociology*, vol. 91, n. 3, pp. 481-510.
- Green E. - Porter R. (1984), Noncooperative Collusion under Imperfect Price Information, *Econometrica*, vol. 52, n. 1, pp. 87-100.
- Greif A. - Kandel E. (1995), Contract Enforcement Institutions: Historical Perspective and Current Status in Russia in Lazear, E. (a cura di), *Economic Transition in Eastern Europe and Russia: Realities of Reform*. Stanford, CA: Hoover Institution Press.
- Greif A. (1993), Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition, *American Economic Review*, vol. 83, n. 3, pp. 525-48 .
- Greif A. (1994), Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies, *Journal of Political Economy*, vol. 102, n. 5, pp. 912-50.
- Greif A. (1997), Contracting, Enforcement, and Efficiency: Economics beyond the Law in Bruno, M., Pleskovic, B. (a cura di), *Annual World Bank Conference on Development Economics*, Washington, D.C.: World Bank, pp. 239-65.
- Greif A. - Milgrom P. - Weingast B. (1994), Coordination, Commitment and Enforcement: The Case of the Merchant Guild, *Journal of Political Economy*, vol. 102, n. 4, pp. 745-76.
- Grillo M. (1994), Teoria economica dell'organizzazione, *Economia Politica*, vol. 11, n. 3, pp. 405-37.
- Guttman J. (1992), The Credibility Game: Reputation and Rational Cooperation in a Changing Population, *Journal of Comparative Economics*, vol. 16, n. 4, pp. 619-632.
- Hart O. - Holmstrom B. (1987), The Theory of Contracts, in Bewley, T. (a cura di), *Advances in Economic Theory*, Cambridge University Press, Cambridge, pp. 71-155.
- Hart O. (1995), *Firms, Contracts and Financial Structure*, Oxford University Press, Oxford.
- Hirshleifer D. - Rasmusen, E. (1989), Cooperation in a Repeated Prisoners' Dilemma with Ostracism, *Journal of Economic Behavior and Organization*, vol. 12, n. 1, pp. 87-106.
- Hogan C. (2001), Enforcement of Implicit Employment Contracts through Unionization, *Journal of Labor Economics*, vol. 19, n. 1, pp. 171-95.
- Holmstrom B. (1999), Managerial Incentive Problems: A Dynamic Perspective, *Review of Economic Studies*, vol. 66, n. 1, pp. 169-82.
- Holmstrom B. - Milgrom P. (1991), Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design, *Journal of Law, Economics, and Organization*, vol. 7, Special issue, pp. 24-52.
- Johnson S. - McMillan J. - Woodruff C. (2002), Courts and Relational Contracts, *Journal of Law, Economics and Organizations*, vol. 18, n. 1, pp. 221-77.
- Kandori M. (1992a), Social Norms and Community Enforcement, *Review of Economic Studies*, vol. 59, n. 1, pp. 61-80.
- Kandori M. (1992b), Repeated Games Played by Overlapping Generations of Players, *Review of Economic Studies*, vol. 59, n. 1, pp. 81-92.
- Klein B. - Leffler K. B. (1981), The Role of Market Forces in Assuring Contractual Performance, *Journal of Political Economy*, vol. 89, n. 4, pp. 615-41.
- Klein B. (1985), Self-Enforcing Contracts, *Journal of Institutional and Theoretical Economics*, vol. 191, n. 4, pp. 594-600.
- Klein B. (1992), Contracts and Incentives: The Role of Contract Terms in Assuring Performance, in Werin, L. e Wijkander, H. (a cura di), *Contract Economics*, Blackwell, Oxford, pp. 149-172.
- Klein D. (1997) (a cura di), *Reputation: Studies in the Voluntary Elicitation of Good Conduct*, University of Michigan Press, Ann Arbor.

- Knack S. - Keefer P. (1997), Does Social Capital Have an Economic Payoff? A Cross-Country Investigation, *Quarterly Journal of Economics*, vol. 112, n. 4, pp. 1251-88.
- Kranton R. (1996), Reciprocal Exchange: A Self-Sustaining System, *American Economic Review*, vol. 86, n. 4, pp. 830-51.
- Kreps D. (1990), Corporate Culture and Economic Theory, in Alt J. - Shepsle K. (a cura di), *Perspectives on Positive Political Economy*, Cambridge University Press, Cambridge, pp. 90-143.
- Kreps D. (1997), Intrinsic Motivation and Extrinsic Incentives, *American Economic Review - Papers and Proceedings*, vol. 87, n. 2, pp. 359-64.
- Kreps D. - Milgrom P. - Roberts J. - Wilson R. (1982), Rational Cooperation in the Finitely Repeated Prisoner's Dilemma, *Journal of Economic Theory*, vol. 27, n. 2, pp. 245-52.
- La Ferrara E. (1998), Ethnicity and Reciprocity: A Model of Credit Transactions in Ghana, *mimeo*.
- Landa J. (1994), *Trust, Ethnicity, and Identity: Beyond The New Institutional Economics of Ethnic Trading Networks, Contract Law, and Gift-Exchange*, University of Michigan Press, Ann Arbor.
- Lazear E. (1981), Agency, Earnings Profiles, Productivity, and Hours Restrictions, *American Economic Review*, vol. 71, n. 4, pp. 606-20.
- Macaulay S. (1963), Non Contractual Relations in Business: A Preliminary Study, *American Sociological Review*, vol. 28, n. 1, pp. 55-70.
- Macleod B. - Malcomson J. (1988), Reputation and Hyerarchy in Dynamic Models of Employment, *Journal of Political Economy*, vol. 96, n. 4, pp. 832-54.
- Macleod B. - Malcomson J. (1989), Implicit Contracts, Incentive Compatibility and Involuntary Unemployment, *Econometrica*, vol. 57, n. 2, pp. 447-480.
- MacLeod B. - Malcomson J. (1998), Motivation and Markets, *American Economic Review*, vol. 88, n. 3, pp. 388-411.
- Macneil I. (1978), Contracts: Adjustments of Long-Term Economic Relations under Classical, Neoclassical, and Relational Contract Law, *Northwestern University Law Review*, vol. 72, pp. 854-906.
- Mailath G. - Samuelson L. (2001), Who Wants a Good Reputation?, *Review of Economic Studies*, vol. 68, n. 2, pp. 415-441.
- Malcomson J. (1999), Incentive Contracts in Labor Markets, in Ashenfelter, O. e D. Card (a cura di), *Handbook of Labor Economics*, Vol. III, New York: North Holland.
- Milgrom P. - Roberts J. (1992), *Economia, organizzazione e management*, Il Mulino, Bologna.
- Milgrom P. - Roberts J. (1990), Bargaining Costs, Influence Costs and The Organization of Economic Activity, in Alt J. - Shepsle K. (a cura di), *Perspectives on Positive Political Economy*, Cambridge University Press, Cambridge, pp. 57-89.
- Milgrom P. - North D. - Weingast B. (1990), The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs, *Economics and Politics*, vol. 2, n. 1, pp. 1-23.
- Nelson P. (1974), Advertising as Information, *Journal of Political Economy*, vol. 82, n. 4, pp. 729-54.
- North D. (1990), *Istituzioni, cambiamento istituzionale, evoluzione dell'economia*, Il Mulino, Bologna.
- Pearce D. - Stacchetti E. (1998), The Interaction of Implicit and Explicit Contracts in Repeated Agency, *Games and Economic Behavior*, vol. 23, n. 1, pp. 75-96.
- Pearce D. (1992), Repeated Games: Cooperation and Rationality, in Laffont, J. J. (a cura di), *Advances in economic theory*, Cambridge University Press, Cambridge, pp. 132-174.
- Polanyi K. (1944), *The Great Transformation*, Rinehart, New York.
- Portes A. (1994), The Informal Sector and Its Paradoxes, in Portes, A., (a cura di), *The Economic Sociology of Immigration*. New York: Russell Sage Foundation.
- Prendergast C. (1999), The Provision of Incentives in Firms, *Journal of Economic Literature*, vol. 37, n. 1, pp. 7-63.
- Putnam R. (1993), *Making Democracy Work. Civic Traditions in Modern Italy*, Princeton, Princeton University Press.
- Ramey G. - Watson J. (2001), Bilateral Trade and Opportunism in a Matching Market, *Contributions to Theoretical Economics*, vol. 1, n. 1, Article 3.

- Reny P. (1993), Common Belief and the Theory of Games with Perfect Information, *Journal of Economic Theory*, vol. 59, n. 2, pp. 257-74.
- Salant D. (1991), A Repeated Game with Finitely Overlapping Generations of Players, *Games and Economic Behavior*, vol. 3, n. 2, pp. 244-59.
- Scoppa V. (2001), I vincoli al licenziamento come ‘commitment device’ nei contratti impliciti incentivanti, *Politica Economica*, vol. 17, n. 1, pp. 47-72.
- Scoppa V. (2003), The Role of Turnover Costs in the Enforcement of Performance Related Pay Contracts, *Metroeconomica*, vol. 54, n. 1, pp. 60-78.
- Selten R. (1978), The Chain Store Paradox, *Theory and Decision*, vol. 9, pp. 127-159.
- Shapiro C. - Stiglitz J. (1984), Equilibrium Unemployment as a Worker Discipline Device, *American Economic Review*, vol. 74, n. 3, pp. 433-444.
- Shapiro C. (1983), Premiums for High Quality Products as Returns to Reputations, *Quarterly Journal of Economics*, vol. 98, n. 4, pp. 658-79.
- Simon H. (1951), A Formal Theory of the Employment Relationship, *Econometrica*, vol. 19, n. 3, pp. 293-305.
- Spagnolo G. (1999), Social Relations and Cooperation in Organizations, *Journal of Economic Behavior and Organization*, vol. 38, n. 1, pp. 1-25.
- Stiglitz J. - Weiss A. (1983), Incentive Effects of Termination, *American Economic Review*, vol. 73, n. 5, pp. 912-27.
- Stiglitz J. (1987), Imperfect Information and the Product Market, in Schmalensee, R. - Willig, R. (a cura di), *Handbook of Industrial Organization*, North Holland.
- Stiglitz J. (1992), Contract Theory and Macroeconomic Fluctuations, in Werin, L. e Wijkander, H. (a cura di), *Contract Economics*, Blackwell, Oxford, pp. 292-322.
- Sugden R. (1989), Spontaneous Order, *Journal of Economic Perspectives*, vol. 3, n. 4, pp. 85-97.
- Tadelis S. (1999), What’s in a Name? Reputation as a Tradeable Asset, *American Economic Review*, vol. 89, n. 3, pp. 548-563.
- Telser L. (1980), A Theory of Self Enforcing Agreements, *Journal of Business*, vol. 53, n. 1, pp. 27-44.
- Tirole J. (1996), A Theory of Collective Reputations (With Applications to the Persistence of Corruption and to Firm Quality), *Review of Economic Studies*, vol. 63, n. 1, pp. 1-21.
- Tirole J. (1999), Incomplete Contracts: Where do we Stand?, *Econometrica*, vol. 67, n. 4, pp. 741-81.
- Williamson O. (1985), *Le istituzioni economiche del capitalismo*, Franco Angeli, Milano.
- Wilson R. (1985), Reputations in Games and Markets, in Roth, A. (a cura di), *Game Theoretic Models of Bargaining*, Cambridge University Press, Cambridge, pp. 27-62.

Summary. Incomplete Contracts and Endogenous Enforcement. A Survey. (J.E.L. L14, J41, D82)

When Court enforcement is excessively difficult or costly, agents are often able to create «endogenous enforcement mechanisms», that is, to design agreements so that each party finds it optimal to carry them out (self-enforcing implicit contracts), thanks to the threat of sanctions affecting the defaulting party, imposed by direct counterparts in the relationship (bilateral enforcement) or by external agents, future potential partners (multilateral enforcement). In this paper it is shown, in a game theoretical framework, the main features and constraints of these endogenous enforcement devices through a survey of the models proposed in the literature. Self-enforcing contracts are viewed as Nash equilibria of repeated games sustained by termination-type and boycotting strategies. The prevailing enforcement mechanism depends on: the information available to internal and external agents; the kind of costs that agents are able to impose on the cheating part; the incentives to apply sanctions. The bilateral mechanism requires only that information is available to the direct counterpart; the effectiveness of sanctions depends on the existence of a flow of quasi-rents from continuing the relationship, caused by switching costs, specific investments or market disequilibrium. The working of a multilateral enforcement mechanism depends on an efficient external transmission of information. However, contractual enforcement is possible even if each single relationship does not generate quasi-rents, has a known endpoint and the agents’ time horizon is finite. In some cases, auxiliary private institutions can emerge to gather and spread information and solve problems of punishment coordination. The working of markets and their efficiency is altered by enforcement problems.

