

CORPORATE SOCIAL RESPONSIBILITY (CSR) AS A MODEL OF “EXTENDED” CORPORATE GOVERNANCE. AN EXPLANATION BASED ON THE ECONOMIC THEORIES OF SOCIAL CONTRACT, REPUTATION AND RECIPROCAL CONFORMISM

Lorenzo Sacconi

1. CSR: Much disorder under the sun

For a long time corporate ethical and social responsibility has been a subject of analysis and concern for highly active and committed but not exactly majority groups of business ethics scholars, NGO's activists, and firms with forward-looking managers. More recently, on the contrary, it seems that there is no trade or business association, local or national public authority, firm or university throughout Europe giving up the opportunity of organizing a conference or a meeting on the theme. Since the end of the last century, a variety of initiatives throughout Europe¹ and the US² have sought to formulate standards for management systems devoted to implementing corporate social responsibility (henceforth CSR) within corporations. These are initiatives which have sprung from joint projects amongst research centres and universities, business firms and the professions; and they involve cooperatives, non-governmental and

* I gratefully acknowledge the support obtained under the UE research project “developing a CSR framework to integrate Q-RES and other social and ethical standards”.

¹ By way of example, in the UK the *Accountability 1000* standard for ethical and social audit and reporting has been developed, followed by the *Sigma Project* supported by the Blair government. A *Values Management System* has been developed in Germany at the University of Konstanz (Wieland 2003), while CSR standards modelled on the example of quality management systems have been proposed by independent bodies of standardization in Spain (*Aenor*) and France (*Afnor*).

² In 2000 the US *Ethics Officer Association* submitted to the US ISO Committee (*BCMS*) a proposal for starting the setting of a new standard for corporate ethics and social responsibility (see Wieland 2003).

non-profit organizations, representatives of interest organizations, stakeholders associations and sometimes local and government authorities as well. In this regard Italy has for once not lagged behind, as testified by two voluntary initiatives.³

Many of these as well as other initiatives – which originally consisted in self-regulation models or best practices by firms – have been reflected in (but also encouraged by) the *Green Paper* (2001), and then the *Communication* (2002), issued by the European Commission to set out its strategy for the definition of concepts, benchmarking experiences, pooling of approaches, and development of a CSR framework at the European level. This has been accompanied by burgeoning public interest which has been indubitably fostered by the decision to put CSR on the agenda of governmental institutions, first those of the European Union (the UE Commission) and then those of the member-states. At the same time, however, unease has been expressed over the perceived superficiality of a number of these talks, together with a call for clarification so that misunderstandings may be prevented, and the consequent false expectations and disappointments forestalled.

Sometimes institutional initiatives do not help⁴. To begin with, in some governmental initiatives CSR risks to be confused with corporate giving and with the purpose of channelling of corporate donations into social programmes set up by the Government. The mistake in part resides in a muddling of the whole, which is CSR, with the part, namely corporate giving. But it consists above all in tampering with the structural mechanism of the donation-based funding of the nonprofit sector. Aside from differences in place and time, economic theory of nonprofit organisations sees the third sector as a response to ‘supra-median’ demand for quantity, quality and cultural differentiation in the supply of welfare goods and services that the governmental supply is unable to satisfy. The government in fact is structurally unable to cater to this demand with its own programmes (financed out of taxation), for these are by definition designed to satisfy demand by the ‘median voter’ – i.e. the voter whose consensus gives a political coalition or a party victory at elections (Weisbrod 1988). If *ad hoc* fiscal and financial incentives are used to channel donations towards social policies decided by the government (with the obvious intent of reducing even more fiscal pressure), there is a risk of distorting one of the structural mechanisms which fuel the development of the third sector and remedy inefficiencies in the overall supply of welfare services.

A second risk is misconceiving the incentives mechanisms that can promote voluntary CSR standards. This ensues if – as the initiative under examination suggests - the promise of tax relief on donations – provided these are pledged to government’s preferred social programs – is combined with the announcement of a system to ascertain how much companies are socially responsible - called “social

³ The *GBS* standard for social reporting issued in the spring of 2001, and the *Q-RES Project* aiming to define a quality standard of management systems for ethical and social responsibility of firms. The latter initiative led in October 2001 to issuing the *Q-RES Management Guidelines*, see (Sacconi et al. 2003), and more recently to the setting of the *Q-RES Norm for the Improvement of Corporate Ethical-Social Performances of Organisations* (March 2003) see (Sacconi, deColle, Baldin, Oakley, Wieland and Zadek 2003)

⁴ To exemplify the misunderstanding of CSR at governmental level, in the text I make some remarks on the Italian Government’s initiative (Ministry of Welfare) called CSR-SC project, proposed during the semester of Italian UE Presidency.

statement". By filling the *statement* companies would give rise to a public register of firms eligible for tax relief and entitled to publicize themselves as 'socially responsible' - admitted that they give their donation to a "public fund" managed by the Government. This induces perverse incentives, for all firms – and rightly so – seek to obtain tax abatement on donations to the third sector, so that in order to admit them to tax reduction a minimal 'quasi-compulsory' and 'fiscal' verification is needed. Instead, voluntary CSR certification should repay with reputation benefits (and only with such benefit) only those firms effectively able to demonstrate their compliance with high standards of CSR. The linkage between CSR certification aimed to reputation and ministerial appraisal aimed to concession of fiscal benefits may arouse such discontent in the business community (afraid that the standard would be too high to many companies only interested in obtaining some tax relief) that at end the Government is forced to reduce CSR certification to simple a self-declaration about a short list of compulsory items, mainly concerned with corporate community giving, and much less committing and informative than a proper social report. (Which is easy to foresee: if a 'label' is given by a Ministry in exchange for money contributed to a governmental fund, which the Ministry has interest to raise as high as possible, there are very poor incentives to be very selective in granting that 'label' – whose value result thus inflated.) At end, after much disorder and alarm, all that turn too indulgent towards those backward firms that see the 'voluntary approach to CSR' as mere discretion (the exact opposite of a moral *obligation* which firms are asked to account for).

Not only Governments and firms, however, but also distinguished scholars are seemingly reluctant to give the proper weight to a well-posed self-regulation approach to CSR. To be fair in accounting for the ongoing discussion on CSR, mention must be made of two opposed form of scepticism. The first can be identified with the position put forward by professor Michael Jensen (2001), which basically cast doubts on any multi-stakeholder approach to strategic management and governance of firms for these would destroy the uniqueness of goal and direction of corporate management, opening the route to managerial slack and self-dealing. It is sharply sympathetic with the dominant shareholder value maximisation view, and biased in favour of the idea that self-interest and profit-maximisation, even if in the long run, are able to accommodate any real problem which CSR is about. I will largely concentrate on these theses and I will attempt to counteract them in sec. 5, 6, 8. On the other hand, there is a second stream of scepticism which considers any voluntary approach to CSR, and the thesis itself that firms should endorse CSR as a criterion for their management and corporate governance, no more than *cheap talk* and hence invokes stronger regulations and imposition of hard law in order to oblige corporate actors to comply with even the minimal requirement of avoiding conflicts of interest. For example Guido Rossi, an outstanding Italian legal scholar and one of the consultants of the UE commission on the reform of corporate governance, is resolutely opposed to business ethics as the basis for introducing forms of ethical and social self-regulation by firms (for example, codes of ethics) (cf. Rossi 2003, chap. V). I will deal with this form of scepticism in the next section in order to be able of putting aside certain preliminary misunderstandings when I will step to the proper discussion of CSR in the following sections. In sections 9, 10 and 11 I will provides two arguments – one based on reputation, the second resting upon

conformism – which show that self-regulation, far from being without teeth, may resort to quite effective incentives and motivations, providing for its self-enforcement.

2. Business ethics and the lawyer's presumption

As a representative of the second kind of scepticism on CSR, Rossi needs some definition of what his scepticism is about. He finds a definition of *business ethics* (cf. *ibid.* p. 116) which locates firmly it in the sphere of rational ethics as justification (or refusal to justify) – by means of models of impartial collective choice – economic institutions at *macro* level (for example the mix between state and market in a given economic system), at *meso* level (the choice among alternative forms of enterprise), and *micro* level (the solution of professional ethical dilemma of managers).⁵ But then he entirely forgets this approach, in which business ethics is clearly rule-based ethics, and presumes that it can only be character-based ethics appealing to personal virtue and conscience and consequently devoid of any motivational efficacy and any social sanction (*ibid.* p. 123, 133-4). From this follows Rossi's presumption that, since ethics entirely lack clarity any effectiveness – which are instead typical of the law – any recourse to self-regulation based on business ethics would be misleading.

Yet this is to omit one of the most obvious definitions of ethics: namely that they constitute the set of social norms commonly accepted on the basis of agreements and conventions, and therefore sustained by the rational choices of multiple agents. Once those norms have passed the universalizability test,⁶ which is what meta-ethically distinguishes them from rules of mere prudence or from social norms not susceptible to moral meaning, they are effective by dint of their social function of solving cooperation and coordination problems. In sum, by forgetting the relation with social norms what is entirely lost is any reference to rational ethics, which besides Kant also owe a debt to Hume.⁷

It should be borne in mind that the relationship between ethics and legal rules is a complex one. From a prescriptive point of view (a legal theorist would say *de jure condendo*), *ethics precede the law*, in the sense that an enduring legal order always rests upon an “overlapping consensus” on constitutional ethical principles (Rawls 1993) which is impartial with respect to interests and particular conceptions of the good, so that the decision to promulgate a law must always pass a test of *impartial assessment* (that is, whether it complies with the content of the “overlapping consensus”) – unless it is believed that legislation is purely the domain of arbitrary decisions by the majority.⁸

Secondly, *ethics proceed together with the law*, especially as the spontaneous acceptance of the authority of the law and of the legitimate constitutional order. Without broad acceptance of the authority

⁵ The definition is taken from the introduction to Sacconi (1991).

⁶ Cf. Hare (1981).

⁷ This is rather paradoxical given the moralistic tone of Rossi's conclusions, where he seems to revert to the ethics of personal virtues (and vices) whose efficacy he has previously disputed. He concludes with Luther (who in turn cites St Paul) that “the love of money is the root of all evil”: cf. *ibid.* p. 143.

⁸ For Rawls (1993) “overlapping consensus” is the agreement stemming from public ethics defined as the exercise of “public reason”: that is, the minimum indispensable content of the social contract which corresponds to the (more exigent) notion of “justice as fairness”.

of the law, and especially of the constitution, these would be wholly impossible to enforce. Only a small number of breaches of the legal order can be sanctioned; while general disobedience would make it impossible for any legal order to survive.⁹ Moreover, via interpretations based on ethical theories of open norms like the 'good faith' rule, or with the help of the ethical rules arising from self-regulation, the judge extends and clarifies the range of application of the legal rules themselves.¹⁰

Finally, *ethics go beyond the law*. While numerous matters cannot be regulated by detailed rules, this can be done only by *general principles* or standards agreed by the parties concerned, or by private orders established by intermediate groups in society. These principles are *not* necessarily enforced by legal sanctions; used instead may be other incentive mechanisms like social approval, economic advantage, the threat of exclusion and being socially ostracised, and so on. The economy is typically the sphere in which complete contracts *ex ante* covering long-term transactions or all-compassing public regulations are not cognitively possible, or else would be too costly to apply. As a consequence, used in economic affairs are other forms of social regulation, self-constraining social norms, codes of behaviour, standards shared among groups or within social groups, and so on (Posner 2000).

It is evident that the term 'ethics' is used here according to at least two distinct meanings. Ethics can be understood as a *normative discipline* which applies the rules of moral reasoning, or as a *positive science* which describes norms and social rules shared in the form of social conventions (Lewis 1969) which *in addition* acquire moral meaning because they are universalizable. Thus identified are *two roles of business ethics*. The first is the furnishing of impartial justification for economic institutional arrangements and rules. The second concerns the social rules of economic self-regulation that possess the force imparted by voluntary compliance and effective social rules, even when these are not enforced by legal sanctions. The former function has only the force of impartial moral reasoning and lacks strong incentives deriving from material interests or psychological drives; the latter can instead rely on numerous individual incentives and efficacious motivational force (motivation may be driven by reputation and trust or by conformism and reciprocity, two distinct yet interdependent mechanisms).

The two aspects do not necessarily converge. For example, the construction of social contract as a choice behind the "veil of ignorance", described by John Rawls (the foremost political philosopher of the second half of the twentieth century), is no longer supported by personal incentives or individual motives once the 'veil' is skipped, unless one accepts the debatable psychological theory which holds that the "sense of justice" is innate and evolves naturally from the more intimate sphere of personal relationships to impersonal relations.¹¹ However if the theories of reputation effects and of conformist expectations

⁹ See in this regard the statement by Saverio Borrelli, former Chief Public Prosecutor of Milan, apropos of corruption: "[...] I do not believe – and I have never believed – that the criminal justice system is able to cure the ills of society [...] the magistrates do not cure the illness with their investigations. This should be done by the political system, the culture [...] Not only politics but also public opinion, society in short, has its responsibilities. I believe that the average Italian does not feel particularly upset by these offences because he or she is used to finagling and therefore accepts that others do it as well. And when this behaviour becomes a cultural fact, the general attitude, it is even more difficult to change..." interview with *La Repubblica*, 14/10/2003.

¹⁰ On this see Preite (1993), chap. 3, section 1.4.

¹¹ Cf. Rawls (1971).

(see sections 9, 10 and 11) are applied to codes of business ethics, it can be shown that the two classes of practical reasons (the ‘detached’ ones appealing to impartial reason, and the ‘personal’ ones consisting of incentives and motivational drives) may converge in a manner which ensures that moral rules, as well as codes of ethics and CSR voluntary management standards, are effective.

3. A definition of CSR

The EU Commission is quite demanding in its definition of CSR:

“By stating their social responsibility and voluntarily taking on commitments which go beyond common regulatory and conventional requirements, which they would have to respect in any case, companies endeavour to raise the standards of social development, environmental protection and respect of fundamental rights and embrace an *open governance, reconciling interests of various stakeholders in an overall approach of quality and sustainability*” (*Promoting a European Framework for Corporate Social Responsibility*, Green Paper, p.4, Brussels, 18.7.2001, emphasis added).

This quotation shows that the Commission regards CSR as a form of corporate strategic management that sets its standards of conduct at a level higher than legal constraints, and envisages CSR as a system for the governance of transactions and relations between the firm and its stakeholders. It is clear that here ‘governance’ is no longer the set of rules simply allocating property rights and defining the owners’ control over the management of a firm. Instead it resembles the neo-institutional view whereby the firm, like the contract and other institutional forms, is a ‘governance system’ which establishes diverse rights and obligations in order to reduce ‘transaction costs’ and the negative externalities of transactions.

I therefore propose the following definition of CSR: *a model of extended corporate governance whereby who runs a firm (entrepreneurs, directors, managers) have responsibilities that range from fulfilment of their fiduciary duties towards the owners to fulfilment of analogous fiduciary duties towards all the firm’s stakeholders.*

Definition is required of two terms in the foregoing proposition:

a) *Fiduciary duties.* It is assumed that a subject has a legitimate interest but is unable to make the relevant decisions, in the sense that s/he does not know what goals to pursue, what alternative to choose, or how to deploy his/her resources in order to satisfy his/her interest. S/he, the *trustor*, therefore delegates decisions to a *trustee* empowered to choose actions and goals. The trustee may thus use the trustor’s resources and select the appropriate course of action. For a fiduciary relationship – this being the basis of the trustee’s authority *vis-à-vis* the trustor – to arise, the latter must possess a claim (right) towards the former. In other words, the trustee directs actions and uses the resources made over to him/her so that results are obtained which satisfy (to the best extent possible) the trustor’s interests. These claims (i.e. the trustor’s *rights*) impose fiduciary duties on the agent who is entitled with authority (the trustee), which s/he is obliged to fulfil. The fiduciary relation applies in a wide variety of instances: tutor/minor and teacher/pupil relationships, and (in the corporate domain) the relation between the board of a trust and its

beneficiaries, or according to the predominant opinion, between the board of directors of a joint-stock company and its shareholders and then more generally between management and owners (if the latter do not run the enterprise themselves). By the term 'fiduciary duty', therefore, is meant the duty (or responsibility) to exercise authority for the good of those who have granted that authority and are therefore subject to it.¹²

b) Stakeholders. This term denotes individuals or groups with a major stake in the running of the firm and who are able to influence it significantly (Freeman and McVea 2002). However, a distinction should be drawn between the following two categories:

- i Stakeholders in the strict sense: those who have an interest at stake because they have made specific investments in the firm (in the form of human capital, financial capital, social capital or trust, physical or environmental capital, or for the development of dedicated technologies, etc.) – that is, investments which may significantly increase the total value generated by the firm (net of the costs sustained for that purpose) and which are made specifically in relation to that firm (and not in any other) so that their value is idiosyncratically related to the completion of the transactions carried out by or in relation to that firm. These stakeholders are reciprocally dependent on the firm because they influence its value but at the same time – given the specificity of their investment – depend largely upon it for satisfaction of their well-being prospects (lock-in effect).
- ii Stakeholders in the broad sense: those individuals or groups whose interest is involved because they undergo the 'external effects', positive or negative, of the transactions performed by the firm, even if they do not directly participate in the transaction, so that they do not contribute to, nor directly receive value from the firm.

It is evident that these two categories cannot be sharply separated. For example, a manufacturer in a developing country who supplies a component for an industrial good assembled in a Western European country is essentially dependent on his contract; and with his low labour costs (due to the customer's market power) he makes a crucial contribution to the European firm's profits. At the same time, however, if a mature technology is used, he is easily replaceable by the European firm, whose dependence on the supplier is therefore limited (in short, the reciprocal dependence relation is not symmetric). Likewise, a local community may not be party to the transactions performed by a company with a plant on its territory, but it is nevertheless subject to that plant's environmental and social externalities. However, if the community has representative institutions with the power to grant or withhold a 'licence to operate', it is able to influence the company's creation of value and negotiate a reduction in the negative externalities. These decisions – connected as they are with the furnishing of infrastructures – may be viewed as investments intended to select and attract production activities whose positive externalities outweigh their negative ones.

We are now able to appreciate the scope of CSR defined as an extended form of governance: it extends the concept of fiduciary duty from a mono-stakeholder setting (where the sole stakeholder

¹² On fiduciary duties see Flannigan (1989).

relevant to identification of fiduciary duties is the owner of the firm) to a multi-stakeholder one in which the firm owes fiduciary duties to *all* its stakeholders (the owners included). It is obvious that classification of stakeholders on the basis of the nature of their relationship with the firm must be regarded as important in gauging these further fiduciary duties.¹³

4. The economic bases of the idea that the firm has ‘further’ responsibilities towards its stakeholders

4.1. Theory of the firm.

Let us now inquire whether economic theory provides support for the thesis that the firm has ‘further’ responsibilities towards its stakeholders. According to neo-institutional theory (Williamson 1975, 1986; Grossman and Hart 1986; Hart and Moore 1990; Hart 1995; Hansmann 1996), the firm emerges as an institutional form of ‘unified transactions governance’ intended to remedy imperfections in the contracts that regulate exchange relations among subjects endowed with diverse assets (capital, labour, instrumental goods, consumption decisions, and so on). These assets, if used jointly, are able to generate a surplus over the cost of their use that is higher than in the case of their separate use by each asset-holder. However, contracts by which these asset-holders regulate their exchanges are incomplete: they do not include provisos covering unforeseen events, owing to the costs of drafting them, or because the cognitive limits of the human mind make it impossible to predict all possible states of the world. Yet for these assets to be used in the best manner possible, specific investments must be made: investments undertaken with a view to the value that they may produce within a idiosyncratic contractual relation. This entails that the surplus generated with respect to the costs sustained by each party to the exchange is determined by the undertaking of *specific* activities with *specific* counterparts (suppliers, customers, employees, financiers, etc.). Let us assume that parties behave opportunistically (that is, they are egoists who act with astuteness). Thus, once the investments have been made, contractual incompleteness means that the terms of the contract can be renegotiated, so that the party in a stronger *ex post* position is able to appropriate the entire surplus, thereby expropriating the other stakeholders. But if agents expect to be expropriated, they will have no incentive to undertake their investments at the optimal level. This expectation of unfair treatment gives rise to a loss of efficiency at the social level.

¹³ At first sight, it might be objected that many stakeholders, in both the ‘strict’ and ‘broad’ senses, do not have relations with a firm such that they formally delegate authority to those who run it (for example, they do not vote), with the consequence that the fiduciary duties as defined earlier do not apply to them. However, in the model of the social contract as a hypothetical explanation of the origin of the firm – see section 5.2 – all the stakeholders participate in the “firm’s second social contract”, with the consequence that their trust constitutes the authority of the firm’s owner and manager. This also explains how the authority of the latter may be accepted by these subjects. Moreover, the hypothetical social contract is typically used to explain how authority – that is, legitimate power – may come about at both the political and organizational levels: cf. Green (1990), Raz (1985), Watt (1982). For a discussion of managerial authority see MacMahon (1989) and Sacconi (1991).

The firm responds to this problem by bringing the various transactions under control of a hierarchical authority – the authority, that is, of the party which owns the firm and through ownership is entitled to make decisions over the contingencies that were not *ex ante* contractible. Unified governance supplements incomplete contracts with authority relations through the vertical and horizontal integration of the units that previously made separate contributions. The firm is therefore a special contractual form: when contracts lack provisos contingent upon unforeseen events, they can be ‘completed’ with the ‘residual right of control’ which entitles its holder to decide what should be done about decisions not *ex ante* contractible – that is, decisions ‘left over’ from the original contract and that become available only when unforeseen situations occur.

The residual right of control underpins authority: those parties entitled with residual right of control may threaten the other parties to the contract with exclusion from the physical assets of the firm, thereby ensuring that *ex ante* non-contracted decisions are taken *ex post* to their own advantage. They are thus safeguarded against opportunism by the other stakeholders, and they are able to protect the expected value of their investments in situations where contract incompleteness provides margins of discretion when residual decisions have to be taken. There is therefore an efficiency rationale for the idea of the firm as ‘unified governance’ of transactions: if one party (a class of stakeholders) has made a specific investment of greater importance than those made by the others at risk, or if its exercise of ‘unified governance’ discourages opportunism by the others to appropriate the surplus, then that party should be granted the property right and with it the right to take ‘residual’ decisions. This is also the basis for regulation of authority delegation from the owners to directors or managers by corporate governance rules, when the owners themselves are not able of directly exercising the entire residual right of control. Fiduciary duties owed to the owners must guarantee that delegated exercise of residual rights of control by the board of directors or managers will maintain or improve the efficiency of their original allocation to the selected class of stakeholders.

4.2. *The risk of abuse of authority.*

However, one should not underestimate the risks of the firm *qua* unified governance. There is not just one single stakeholder at risk because of contract incompleteness; it is usually the case that multiple stakeholders undertake specific investments (investments in human capital, investments of trust by consumers, investments of financial capital, investments by suppliers in raw materials, technologies and instrumental goods). Contracts with these stakeholders are also incomplete.

Yet if a firm brings its contracts with certain stakeholders (labour contracts, obligations towards and relations with minority shareholders) under the authority of a party to whom is allocated control over residual decisions (for example, the controlling shareholder group) – and more generally if a party is enabled by its *de facto* power to exercise discretion over *ex ante* non-contractible decisions concerning implicit or explicit contractual relations with the other stakeholders (consumers, customers, suppliers, creditors, etc.) – what, one may ask, is there to ensure protection of investments and interests other than those of the controlling stakeholder? It is evident that if fiduciary duties attach *only* to ownership, those

stakeholders *without* residual right of control will *not* be protected by the fiduciary duties of those who run the firm.

The inherent risk, therefore, is an abuse of authority (Sacconi 1997, 2000). Those wielding authority may use it to expropriate the specific investments of others by exploiting ‘gaps’ in contracts – which persist even under unified governance (in fact it simply allocates to only one stakeholder the right to ‘fill’ those gaps with its discretionary decisions). Those in a position of authority, in fact, are able to threaten the other stakeholders with exclusion from access to physical assets of the firm, or from the benefits of the contract, to the point that those other stakeholders become indifferent between accepting the expropriation and forgoing the value of their investments by withdrawing from the relation. Thus the entire surplus, included that part of it imputable to efforts and investments made by the non controlling stakeholders, will be appropriated by the controlling party. Again forward-looking stakeholders will be deterred from entering the hierarchical transaction with the controlling party. In general, this will produce a internal *crisis of legitimacy* between firm and stakeholders (a crisis in the relationships between the organizational authorities and participants in the organization) and an external *crisis of trust* (in relationships with stakeholders that have entered into contractual or external relations with the organization). Various stakeholders will *ex ante* have a reduced incentive to invest (if they foresee the risk of abuse), while *ex post* they will resort to conflicting or disloyal behaviour (typically possible when asymmetry of information is inherent in the execution of some subordinate activity) in the belief that they are being subjected to abuse of authority. In the economist’s jargon, this is a ‘second best’ state of affairs (less than optimum): all governance solutions based on the allocation of property rights to a single party may approximate social efficiency, but they can *never* fully achieve it. This much is acknowledged by the theoreticians of contractual incompleteness when they point out that the allocation of the residual right of control induces the party protected by that right to over-invest, while those not so protected are induced to under-invest, with a consequent shortfall with regard to the social optimum (Grossman and Hart 1986; Hart 1995).

On the other hand, if the stakeholder category entitled to exercise ownership (the double right of controlling residual decisions and claiming residual revenue; cf. Hansmann 1987, 1996) is selected on the basis of its ability to minimize total costs deriving from the summation of contractual costs borne by of various stakeholders and costs of exercising authority, it is by no means certain that a solution will be found which reduces *each* of those costs to the minimum (that is, reduces opportunism suffered by each stakeholder to the minimum). Sufficient for this solution to emerge is, for example, that the governance costs of one class (the capital-holders, for example) are low enough to counterbalance a relative increase in the contractual costs borne by another class (the workers, for example) compared to alternative cases (for instance the case in which there is no centralized governance, or the one in which it is a sub-set of workers that governs, or the consumers). In this case, too, some incentives are nullified, which distances the real-world solution from complete (Paretian) social efficiency. The fact is that the relative (in)efficiency depends on manifest or simply expected unfairness: separation between efficiency and fairness (a myth of neoclassical economics) is no longer feasible when we face the real-life problem of working out acceptable solution for the governance of transactions.

My suggestion is therefore that when CSR is viewed as ‘extended governance’, it completes the firm as an institution of transactions governance (cf. Sacconi 2000). The firm’s legitimacy deficit (whatever category of stakeholders is placed in control of it) is remedied if the residual control right is accompanied by further fiduciary duties towards the subjects at risk of abuse of authority and deprived of the residual control right. At the same time, this is a move towards greater social efficiency because it reduces the disincentives and social costs generated by the abuse of authority. From this perspective, ‘extended governance’ should comprise:

- *the residual control right* (ownership) allocated to the stakeholder with the largest investments at risk and with relatively low governance costs, as well as the right to delegate authority to professional directors and management;
- *the fiduciary duties* of those who effectively run the firm (administrators and managers) towards the owners, given that these have delegated control to them;
- *the fiduciary duties of those in a position of authority in the firm (the owner or the managers) towards the non-controlling stakeholders*: the obligation, that is, to run the firm in a manner such that these stakeholders are not deprived of their fair shares of the surplus produced from their specific investments, and that they are not subject to negative externalities.¹⁴

A number of recent economic and legal models of governance support this view of CSR. For example, the firm can be seen as a ‘nexus’ of specific investments regulated by incomplete contracts, rather than as a nexus of simple contracts, and therefore as a team of actors cooperating to produce a surplus from those specific investments (Rajan and Zingales 2000). Based on a similar view which combines different theories of the firm – the theory of incomplete contracts with that of team production – is the model of multi-stakeholder governance developed by Margaret Blair and Lynn Stout, and which sees the purpose of corporate governance structures as being prevention of opportunistic behaviour among the N members of the team that make specific investments. When applied to a public company, this model translates into a board of directors acting as a mediating hierarchy: an authority system charged with the task of finding

¹⁴ I have proposed in a previous work (cf. Sacconi 1991) a view of managerial ethics based on a similar analysis of the theory of firm, as well as on the cooperative game theory of the firm put forward by Mashairo Aoki (Aoki 1984) and the notion of extended fiduciary duties (cf. Sacconi 2000). Surprisingly, Guido Rossi (Rossi 2003, pp. 114-18), after polemically describing business ethics as “tampon ethics”, dismisses the idea that governance of a firm should comprise fiduciary duties towards multiple stakeholders as a variant of the “conflict of interest epidemic” (cf. Rossi 2003, 124). He thus confuses mediation by the trustee among the different trustors’ legitimate interests with a pathological conflict between the trustor’s interests, which the trustee is obliged to pursue, and the personal ones of the trustee himself. This confusion – so that a board of directors which seeks to mediate between shareholders value and other objectives would undergo to a conflict of interest (hardly a ‘progressive’ thesis) – may be due to Rossi’s ignorance of recent developments in the theory of the firm which at least suggest that the problem should be taken seriously. The contention – which springs precisely from internal criticism of neo-institutionalist theory – that firms have responsibilities towards their stakeholders (what I would call a ‘reformist’ view) may be ‘tampon ethics’, but Rossi is merely recycling Milton Friedman’s thesis that “the only social responsibility of business is to make profits” when he appears to accept the capitalism of shareholder value without any ‘tampon’ at all, menacing anyone who argues to the contrary with the same fate as the Weimar Republic (ibid. pp. 123-4).

the appropriate balance in the protection of diverse interests (cf. Blair, Stout 1999). The (controversial) legal basis for this form of “impartial governance” exercised by the board of directors and by management in the US joint-stock company is the ‘business judgment doctrine’: the manager’s use of a standard of professional conduct which insulates his/her choices against claims by shareholders (cf. Blair, Stout 1999, but also see Meese 2002).

However, a number of unanswered questions remain which the proponent of CSR as ‘extended governance’ must necessarily address. Does there exist a criterion with which to give more precise specification to these extended duties, and from which it is possible to derive a strategic management standard of sufficient clarity such that the ‘extended governance’ model cannot be accused to entail higher governance costs than the traditional ‘narrow’ corporate governance view? What norms are effective for the implementation of CSR? What is the role of company law with respect to other parts of the law that impose constraints on corporate behaviour? And what role can be played by self-regulation?

5. How can we identify CSR duties (without arbitrariness)?

5.1. *The Social Contract as a criterion for strategic management.*

If a firm is a team of participants with specific investments, then the metaphor of a ‘bargaining cooperative game’ among multiple stakeholders can be applied. These stakeholders must agree on a shared action plan (a joint strategy) which allocates tasks among the members of the team so that the contribution of each of them is efficient (because it produces the maximum surplus net of each stakeholder’s costs). The ‘bargaining cooperative game’ played by the stakeholders is typically one of mixed interests. Although it is in their common interest to cooperate, because this enables them to produce a surplus that would otherwise be impossible, conflict nevertheless persists among the stakeholders over the distribution of the value created. ‘Governance’ and strategic management consequently consist in the solution of two problems:

- a) Identifying the joint strategy that the stakeholders (as the players in the cooperative game) may utilize to coordinate themselves, in that they accept it *ex ante* as a voluntary agreement to cooperate – so that strategic management can reduce bargaining costs (time, conflict, etc.) and the costs of gathering information on the alternatives available and on the intentions of each players about cooperation.
- b) Ensuring *ex post* that each member of the team complies with the agreement on the joint strategy selected and does not act as a free rider with regard to the others.

Choosing the joint strategy (point a) is equivalent to select a bargaining equilibrium. It must therefore answer the question of what is due to each stakeholder and what each of them can expect from the firm in exchange for its contribution, so that each stakeholder may agree on that joint strategy. The question thus arises as to how the stakeholders’ interests can be balanced against each other, and what claims on the firm should be considered the appropriate basis for the management’s fiduciary duties. ‘Stakeholder’, in fact, is a descriptive term. It reminds us that a variety of classes of individuals have interests at stake in

the running of the firm, and that they may sometimes advance conflicting claims. The use of the term 'stakeholder', however, does not provide a *criterion* with which to balance claims when they are mutually conflicting.

To answer the question we consequently need a *criterion* able to identify the balance that *any whatever stakeholder* would accept as the basis for its voluntary cooperation with the firm: that is, an *impartial* criterion. It is here that ethics – understood as a set of *impartial criteria* for collective choice-making – come into play as part of the firm's governance and strategic management.

As an ethical criterion, therefore, I suggest the 'social contract' among the stakeholders of the firm (Sacconi 1997, 2000). By 'social contract' I mean not any whatever real-life bargain but a 'touchstone' from which point of view to assess the diverse outcomes of day by day practical running of the firm. In other words, the social contract is the agreement that would be reached by the representatives of all the firm's stakeholders in a hypothetical situation of impartial choice.¹⁵ Corresponding to the notion of 'social contract' is the following multi-stage deliberative procedure which generates impartially acceptable agreements.

- i Force, fraud and manipulation must be set aside.
- ii Each party comes to the bargaining table with only its capacity to contribute and its assessment of the utility of each agreement or non-agreement proposed (dispensing with any form of threat other than its possible refusal to agree).
- iii The bargaining *status quo* must be set at a level such that each stakeholder results immune against the cost of its specific investments – that is, each stakeholder must obtain from the social contract at least reimbursement of the cost of the specific investment with which it has contributed to the surplus (otherwise the bargaining process would permit opportunistic exploitation of the counterparty's lock-in situation). The distribution of the surplus is regulated by the social contract – and by the corresponding deliberative procedure – on the basis of 'initial endowments' thus defined.
- iv Each party in turn puts itself in the position of all the others, and in the position of each of them he can accept or reject the contractual alternatives proposed.
- v If solutions are found which are acceptable to some stakeholders but not to others, these solutions must be discarded and the procedure repeated (which reflects the assumption that cooperation by all stakeholders is recognized as necessary).

¹⁵ It is quite evident the debt of this contractarian view on the theory of firm to the works of both John Rawls (1971, 1993) and David Gauthier (1986). For the first formulation of the theory of the corporate social contract, based the revision of neo-institutionalist theory of firm and with reference to the problem of the abuse of authority *vis-à-vis* stakeholders, see however Sacconi (1991), and latterly Sacconi (1997, 2000). For a formulation external to economic theory see Donaldson (1982) and then Dunfee and Donaldson (1995). The social contract ethical theories - in both the 'ideal' (Kantian) and 'real' (Hobbesian) versions – has been put forward as keys to interpretation of the abuse of majority power in joint-stock companies by Disiano Preite (Preite 1992).

- vi The terms of the agreement reached are therefore those that each stakeholder is willing to accept from its particular point of view: that is, the non-empty intersection of the joint strategies and relative distributions acceptable to each of them.

Note that this intersection is necessarily *non-empty*, for otherwise the game would not allow a cooperative surplus. That is to say, it would not be the case that joint action by the parties may produce something more than their separate action and that at least one surplus distribution proves to be reciprocally advantageous (if it *must* be so, then there exists at least one agreement acceptable to all).

5.2. Social contract and the emergence of the firm.

Thus far, the social contract has been presented as a normative deliberative procedure by which to identify the terms of an agreement that would be acceptable from an impartial standpoint – that is, from the point of view of any whatever stakeholder – so that it can be adopted as a standard of behaviour by, for example, the mediating hierarchy proposed by Blair and Stout. However, the social contract can also furnish a reconstruction – understood as a ‘potential explanation’ – of how bargaining has given rise to a firm with *both* fiduciary duties towards the owners *and* social responsibility (i.e. further fiduciary duties) towards all the stakeholders.

Consider a ‘state of nature’ prior to the creation of the firm. Bilateral transactions among stakeholders regulated by incomplete contracts are subject to reciprocal opportunistic behaviour, with the consequence that prohibitive bargaining costs render them inefficient. At the same time, the parties to those transactions are entirely unconcerned about the negative external effects of their transactions on other agents, who although they do not participate, are nevertheless affected. This is a Hobbesian scenario in which the life of economic transactions among agents is “solitary, poor, nasty, brutish, and short”.¹⁶ The stakeholders thus address the problem of creating an association whereby all their transactions can be undertaken in accordance with agreed-to rules and are therefore not subject to contract-costs, while at the same time the negative effects on those who do not participate in the benefits from the transactions are reduced to the minimum. The ‘First Social Contract’ of the firm (*pactum unionis*) is nothing other than the agreement which the stakeholders reach *among themselves* to set up this association (the ‘*just* firm’). They negotiate on the association’s constitution, which consists in a common plan of action (joint strategy) to which each of them contributes either by carrying out a positive effort or by simply refraining from applying his/her veto. This *first social contract of the firm* stipulates as follows:

- a) rejection of shared plans of action which generate negative externalities for those not participating in the cooperative venture or, if these negative externalities are essential for the production of the cooperative surplus, a compensation of third parties so that they are rendered neutral;

¹⁶ Cf. Hobbes, *Leviathan*, (1651), part 1, chapter 13.

- b) production of the maximum surplus possible (difference between the value of the product for its consumers, who belong to the association, and the costs sustained by each stakeholder to produce it);
- c) a distribution of the surplus which is 'fair', or rationally acceptable to each stakeholder in a bargaining process free from force or fraud and based on an equitable *status quo*, that is, considering the surplus net of the specific investments.

However, if an attempt is made to reach this form of an ideal association (the 'just firm') which eliminates all the participants' contract-costs, they arrive in practice to an organisational form which is found to be inefficient from the point of view of its governance costs. The stakeholders discover, for example, that the general assembly of all members is unable to take coherent decisions in a reasonable amount of time. In the absence of a monitoring system, once the members of the association have established fair shares of the surplus to be distributed among them, they have an incentive to act opportunistically and not to play their part. Coordination problems arise on how the joint strategy can be implemented under changing circumstances, which may alter beliefs and reciprocal expectations asymmetrically. The stakeholders consequently draw up a *second social contract* of the firm (*pactum subjectionis*)¹⁷ by which they constitute, in the proper sense of the term, a governance structure for the association. It is only now that the association becomes a hierarchical structure.

The second social contract provides that authority should be delegated to the stakeholder most efficient in performing governance functions (the taking of residual decisions, devising coordination solutions as circumstances change, monitoring, the enactment of sanctions, excluding potential free riders, etc.). For this reason, it can also be seen as a contract *between the stakeholders and* those who is given control over the firm (social contract *with the firm*). After comparative examination of the governance costs of each stakeholder, the one with the lowest costs is selected and assigned ownership, and is therefore the one to which the right of governing the association is delegated (Hansmann 1996). This class, which is remunerated with the *residual* is authorised to delegate some discretionary decisions in regard to running the firm to professional director and managers, and to appoint those who are in the authority position of running the firm. *Prima facie*, their authority will be effectively constituted – that is, the delegation will remain valid – as long as they comply with what I call

Narrow fiduciary proviso: the owners are remunerated with the maximum residual revenue possible (in forms compatible with the diverse nature of the controlling stakeholder: profits, returns, discounts, improved conditions of service, improved conditions of employment, and so on) in the light of conditions obtaining in the firm's specific market.

However, it is evident that this proviso entails that the positions of the other stakeholders change (from the "just firm" to *just a firm*). Formerly co-equal members of the association, they are now subject in various ways to the discretionary decisions taken by the stakeholder entitled with authority, and by the administrators that it has appointed. Unlike in the standard economic theory of the firm, in the social

¹⁷ Interestingly, also Blair and Stout (1999) adopt the analogy between the firm and the two social contracts typical of the social contract tradition.

contract theory the risk of the abuse of authority can squarely be faced. The *second social contract* is therefore conceived in a manner such that this cost of hierarchy is forestalled as well. Hence, under the second social contract, the stakeholders agree to submit to authority, thereby rendering it effective, if the contract contains the which stipulates that the firm's new governance structure must comply with *fiduciary duties* towards all the stakeholders (owners and non-owners).

Extended fiduciary proviso:

- Towards the non-owners
 - The firm must abstain from activities which impose negative external effects on stakeholders not party to transactions, or compensate them so that they remain neutral;
 - The firm must remunerate the stakeholders participating in the firm's transactions with pay-offs (monetary or of other kinds, for example in terms of the quantity, quality and prices of goods, services, working conditions, etc.) which, taken for granted a fair status quo, must contain a part tied to the firm's economic performance such to approximate fair/efficient shares of the surplus (assuming that this is positive) as envisaged by the first social contract.¹⁸
- *Towards the owners:* The firm must remunerate the owners with the maximum residual compatible with fair remuneration – as defined by the first social contract – of the efficient contributions made by all the other stakeholders.

5.3. The “*Societary interest*” of the company and extended fiduciary duties.

What does this hypothetical explanation yield? It yields a definition of the ‘societary interest’ of the company – that is, the interest that the manager acting in the name of the company must serve – which is consistent with the contractarian model. According to this reconstruction, in fact, the manager (appointed through the second social contract) has a special fiduciary duty towards the owners (or the ‘residual claimant’) that has delegated authority to him/her (*via* narrow fiduciary proviso). This duty applies, however, only under the constraint that the *general* fiduciary duties are fulfilled towards *all* the stakeholders – which is defined *via* the extended fiduciary *proviso*. We may thus construct the corporate interest by means of a hierarchical decision-making procedure which moves from the most general conditions to the most specific ones:

- *First step:* minimize the negative externalities affecting stakeholders in the broad sense (perhaps by paying suitable compensation);
- *Second step:* identify the agreements compatible with the maximization of the joint surplus and its simultaneous fair distribution, as established by the impartial cooperative agreement among the stakeholders in the strict sense;

¹⁸ Note that meant here is remuneration in utility and not necessarily in money. Put in economic parlance, this remuneration consists of the consumer rent, the producer rent, the worker rent and so on, accruing to each of them from the firm's transactions. This means that some stakeholders may not want to receive monetary benefits from the firm, but rather improvements in working conditions or in purchasing power, in the quality of goods and services, of contractual conditions, etc., to which the shares of the surplus are in any case devoted.

- *Third step:* if more than one option is available in the above defined feasible set, choose the one that maximizes the *residual* allocated to the owner (for example, the shareholder).

Hence, the narrow corporate interest (the one usually advocated by supporters of the "shareholder value" view) results from a series of steps which select the admissible ways in which this interest can be satisfied – that is, those that are consistent with the various constraints imposed by the first social contract on the owner's behaviour. It should be emphasized that this concept cannot be reduced to that of value maximization for the 'residual claimant' (the owners) once constraints imposed by positive contractual obligations have been fulfilled. This is because we recognize all contracts are incomplete, and they are always susceptible to opportunism (even by those who run the firm), so that it is the entire hierarchical decision procedure which provides the basis for satisfying the corporate interest – i.e. the social contract identifies the goals or the internal (not merely external) moral constraints that channels managerial discretion. It results from satisfaction in sequence of the three requirements set out above and which can be summarized as follows: maximize the value for the residual claimant under the constraint of complying with the social contract between firm and stakeholders which defines the 'stakeholder value'.

6. An objective-function for the firm?

6.1. A computable solution to the firm's bargaining game.

The main objection brought against CSR is that the multi-stakeholder approach to firm's governance leaves the management without a clearly-stated and uniquely defined 'bottom line', to be used as a benchmark against which to evaluate its success or failure (Jensen 2001). The consequence, the argument runs, is that the management exploits this situation to pursue its personal interests. It comes up with every possible device to conceal its essentially self-dealing behaviour behind the interests of some or other stakeholder. Whereas, the critics of CSR maintain, it is easy to check the managerial strategy (among the alternatives available at any particular time) against the criterion of increasing as much as possible the firm's profits, and with which the management can be straightforwardly made to comply, this is not the case of 'stakeholder value', since this consists of numerous dimensions to maximize simultaneously (the interests of the various stakeholders). Consequently, stakeholder value contains an intrinsic contradiction – the pursuit of conflicting, or at any rate divergent, goals at the same time – so that the choice of which strategy to adopt is ultimately left to the mere managerial discretion.¹⁹

It is evident, however, that this objection does not apply to the model of the social contract of the firm proposed here – which by no means ignores the existence of a distributive conflict, but instead resolves it by identifying a bargaining equilibrium that permits mutual cooperation among the members of the team.

To be precise, I would add that counterpart to the philosophical contractarian model is a mathematical model of bargaining whose solution is as exactly computable as the profit function in microeconomic theory. Hence we can simply substitute profit maximization with maximization of the function which

¹⁹ This danger is also stressed by Tirole (2001), who however recognizes the relevance of the stakeholder approach to corporate governance.

assigns the solution to the bargaining game and assume this as the firm's (perfectly computable) objective-function.²⁰ This solution is simultaneously an answer to both the problem of cooperation and that of distributive conflict among the stakeholders. Moreover (for those able to appreciate the marvels of mathematics), if the bargaining set is well defined (and if one accepts the Nash, Harsanyi and Zeuthen postulates of bargaining theory), the solution is defined uniquely, so that the set of admissible solutions is reduced to one single alternative. Hence the best pursuit of the interest of the stakeholder controlling the firm is equivalent to the solution of the bargaining problem among all the stakeholders. (In any event, various theories of bargaining yield solutions which quite closely resemble each other – cf. Gauthier (1986) and Kalai and Smorodinski (1975) as they are slight changes of the basic Nash's solution; and for the purposes of this study, identifying a set of 'close' solutions compatible with the idea of rational bargaining seems good enough).

What does this solution say? On the assumption that the external effects on third parties with no influence on the transactions are minimized by selection of the forms of cooperation which damage them least, one may proceed as follows. Net of the initial (pre-bargaining) position or *status quo*, in which coverage of the costs of specific investments by each stakeholder must be included, calculation is made of the value for each stakeholder of each cooperative outcome from joint action plans. The set of these values (or better, the set of these vectors of values) is the outcomes space associating to each joint action plan (joint strategy) an allocation of the cooperative surplus (positive or nil) to the players. Rational bargaining takes place within this space among players endeavouring to obtain a share as high as possible of the cooperative surplus for themselves, to the detriment of the others' shares - once it is taken for granted the need for reciprocal agreement, for in its absence they would be unable to obtain anything more than the initial position (*status quo*). Calculable within this space is the Nash bargaining function, the product of the utilities of the various stakeholders with specific investments, that is, an aggregative function of their utilities. Where the product (the aggregation) is maximum, there is a bargaining equilibrium (i.e. a rational agreement among the participants in the social contract) corresponding to the Nash bargaining solution (Nash 1950; Harsanyi 1977). (see Fig. 1 for an example with two players).²¹

²⁰ For the theory of bargaining games see Harsanyi (1977). Mashairo Aoki uses the Nash bargaining solution in his theory of the firm, which envisages impartial governance by managers (Aoki 1984).

²¹ Specifically, let us consider a case with two players, 1 and 2, and let us assume that the solution is a point in space \mathbb{R}^2 enclosed between the positive Cartesian axes U_1 and U_2 , each of which measures the utility for a player of the outcomes of the cooperative game (see Figure 1). The space therefore represents the outcomes subject to bargaining in terms of their value in utility for the players (i.e. their payoffs). The standard analytical assumption is that the payoff space is convex and compact. The payoff space \mathbf{P} therefore has an efficient frontier (in the upper-right positive ortant) which represents the set of outcomes for which the players' utilities cannot be increased by an alternative agreement without reducing the utility of at least one other player. Below this frontier are agreements with respect to which gains are still possible for all; above it are outcomes unfeasible by any agreement or joint plan of action. All points in the space represents different possible values of the coalition among the two players. In fact, only when all of them agree on the solution of the game can they leave the *status quo* \mathbf{d} , which is represented by a point interior to the space, so that they may benefit form cooperation. The characteristic function of the coalition among all the players is therefore super-additive (it is better to agree than not to agree). Obviously, of interest are only those agreements for which there is an efficient allocation. But in what point among those on the frontier should the

This equilibrium does not need operational interpersonal comparisons of utility (which operationally are very problematic) in order to be calculated (interpersonal comparisons can be confined to the interpretive level²²). It obeys, in fact, simple axioms of individual rationality in bargaining – like the decision to grant a concession according to the expected personal utility given the probability that the counterparty will accept or refuse it, or that a player will not make a concession that he or she would not expect the counterparty will make in a similar situation – and conditions of mutually expected rationality, like expecting that the willingness of acceptance by the counterparty depends on a symmetric probabilistic assessment of the first party behaviour, and not to expect the counterparty to accept something that oneself would not accept, and so on (cf. Harsanyi 1977)].

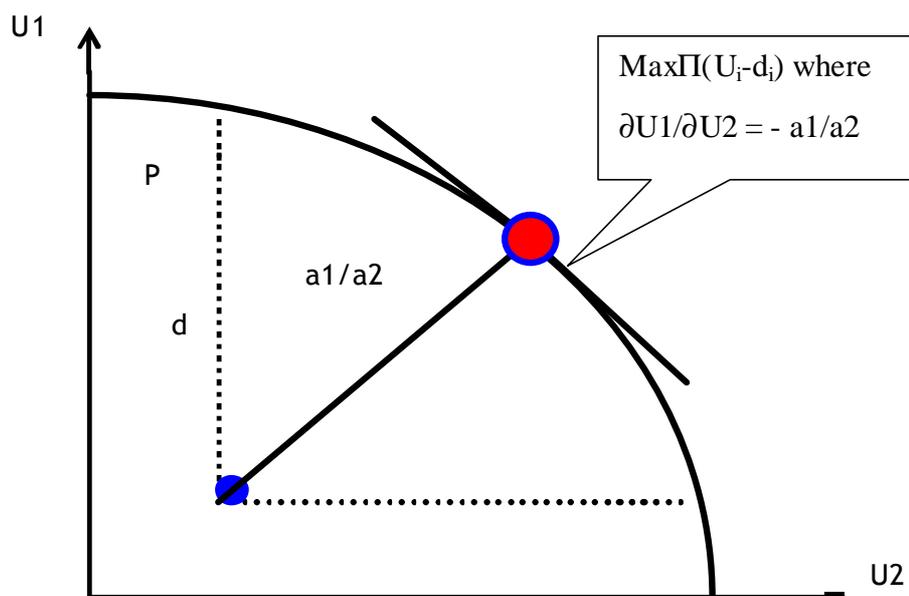
Of course, if these postulates are taken literally, they can be criticised as unrealistic; and it is likely that in the real world agents are unable to maximize or to estimate probabilities coherently, or to make accurate forecasts about the rational behaviour of others. But what matters for my purposes here is that these postulates are a good approximation of rational behaviour in a hypothetical (ideal) bargaining situation among stakeholders (for it is a *normative* model being developed here – one no less normative than that of profit maximization).

agreement fall? The Nash solution states that the players will agree on the joint strategy corresponding to the point of the frontier where the following holds:

$$\text{Max } \Pi_i (U_i - d_i) \quad (i=1,2 \text{ denotes the various participants in the bargaining})$$

where U_i is the utility of the generic stakeholder i for the cooperative transaction that it undertakes with the firm, and d_i is the cost of the specific investments made by i in order to participate in the joint action plan (that is, i always at least recoups the cost of its specific investment). The solution assumes that bargaining should provide each player with at least a small net advantage, which is the difference between the share of the surplus received and the *status quo* value. As a consequence of additional rationality postulates, these net individual advantages can be identified as such that the *product* of all of them is the maximum among those in the set of the possible outcomes of the cooperation. We may say that this is the collective choice function adopted by the members of the coalition, in light of their bargaining, to resolve the problem of their joint action. It is coherent with the proportionality of the remunerations with the relative utilities, because the ratio in which the shares of the surplus a_1/a_2 are distributed is proportional to the ratio between the marginal variations in the players' utilities $\partial U_1/\partial U_2 = - a_1/a_2$ (Brock 1979, Sacconi 1991, 1997). On the basis of Nash's postulates (1950) and those of the Zeuthen-Harsanyi (cf. Harsanyi 1977), this solution expresses a bargaining equilibrium based on individual rationality of the players.

²² See Brock (1979), and Sacconi (1991).



(Fig. 1. A cooperative bargaining game with two players and the Nash solution in a symmetrical case)

Moreover the outcome of the bargaining game can be interpreted as the solution which is coherent with a notion of distributive justice. On the assumption that it is possible to unify the units measuring utility, we find that the bargaining solution computed always distributes the advantages proportionally to ‘relative needs’ or to the relative marginal variations in the intensity of personal utilities (Brock 1979; Sacconi 1991, 1997). Because it is located on the upper-right frontier of the space of the bargaining outcomes, it fulfils the requirement of social efficiency (no advantage from cooperation is lost) and at the same time corresponds to an intuitive notion of fairness.

6.2. From general to particular, but not vice versa.

We may thus answer the question as to the relation between stakeholder value and shareholder value (or in general the value allocated to the owners). In order to justify social responsibility over shareholder value, old institutional theories concerned themselves with the firm’s interest *as such*, or with *survival* of the firm *as such*. But *as such* the firm is nothing. On the contrary it is a human artefact that derive its goals from human preferences and interests: it does not have an interest *per se* and its ‘survival’ is a misleading biological analogy with a single organism, rather than with a structure produced by the interdependent decisions of many individuals (moreover, the biological analogy is normatively debatable: is a firm never allowed to fail?).

New institutional theory, in the “social contract” version proposed here, does not need this biological analogy, for it can talk directly of the firm as an institutional/artificial solution to the problem of

coordination and cooperation among different individuals holding multiple interests. It is an *artificial construct* whose mission is an *intermediate* and *not* final end; in other words, it is a *means* by which to solve problems of coordination and cooperation among stakeholders *for* (i.e. to the advantage of) their interests. Thus, the firm persists and is successful if it realizes an equilibrium of cooperation among the stakeholders; it can create surplus wealth only if it achieves that equilibrium. However, contrary to the theory of the firm developed within the property rights school, and which to date has predominated in the neo-institutional approach, here methodological individualism does not entail identification of a single individual or class interest to serve, because the firm is explained as the *means* which enables a fair bargaining equilibrium to be achieved amongst *diverse* players in the players mutual advantage. The bargaining equilibrium is the point at which the satisfaction of various interests intersect; the one, that is, where the interest of each stakeholder is fulfilled, with each of them maximizing their objectives under the constraint that also the counterparty, viewed as free and informed, is able to do likewise.

We have thus established an implication relation that holds in only one direction: the impartial agreement *always* entails satisfaction of the shareholder's interest, given that this is conditioned to the achievement of an agreement of cooperation with others. However, the reverse is not necessarily the case: the maximum interest of the shareholder *does not always coincide* with the joint interest of the stakeholders, at least in the short term. Besides agreement, there are *other* ways in which a single stakeholder (for example the shareholder) could pursue his interest, but this presupposes either *conflict* (force) or *deceit* (not abiding by implicit pacts and promises). The prototype of these situations is depicted by the 'prisoner's dilemma' game, where each player has an interest in taking advantage of the other's cooperation without doing his part; and if he expects the other player not to play his part in the cooperation, he has even more reason for not playing his own. This way of pursuing individual interest is clearly self-destructive, because the outcome of mutual defection is always inferior to that of reciprocal cooperation. Nevertheless, as far as the players think that exactly this is the game they are bound to play, there is no escape from this self-defeating way of following their self-interest. Moreover in the real world, there are cases of asymmetry of knowledge or ability such that from time to time a stakeholder (typically that in the better position) may convince himself that he can get the better of the counterparty even in a situation which is fundamentally a 'prisoner's dilemma' (no wonder of course that once the DP like basic structure emerges, that is the counterparty becomes aware of its possibility of retaliation, all are back again to the mutually suboptimal solution.)

Therefore, should we wish to start from *one* stakeholder's pursuit of its interest, in particular the shareholder, we must establish *under what conditions* (i.e. non unconditionally) the 'particular' coincides with the 'general'. And this depends on the type of game that the stakeholders are required to play (in the prisoner's dilemma this would *not* come about, as we have seen), which concerns, amongst other things, the rules (not necessarily imposed from outside) that give some sort of structure to the games in question. One of the thesis of this paper is that *CSR management standards* are part of the (endogenous) rules that make possible that the game is mutually profitable and make possible to escape that the player are driven to conclude that their interaction is doomed to a self-destructive DP-like logic.

7. What norms can effectively bring about CSR? The role of company law.

We may therefore enquire as to the norms - legal or not - that can steer the interests of agents towards compliance with extended fiduciary duties. Welfare state regulations, labour-market and environmental regulations, establish a general legal framework, but they cannot regulate every detail of firms' decisions. They may lay down minimal compulsory conditions, but in many settings their application requires interpretation of a 'grey' zone; or else the conditions for verifying compliance with them may not be observable. Moreover, even when management decisions closely affect the stakeholders, the law cannot regulate them in every respect: the decision whether or not to restructure or down size a firm is always a business decision, notwithstanding the requirements of the law with regard to the protection of third parties or employees. Regulations aimed to dictate how such decisions should be made in every circumstance would inevitably be inefficient (as it has proved to be in any command economy).

Nor do contracts provide a solution. Indeed, the fact that contracts are generally incomplete has prompted the present discussion of CSR. As a consequence, social responsibility intervenes as a remedy to contractual gaps by introducing internal restraints within the very exercise of legitimate authority over non-controlling stakeholders by owners and management. On my hypothesis, the social contract between the firm and the stakeholders is the 'hypothetical contract' which works as a set of default or 'gap-filling' rules to complete incomplete contracts (Coleman 1992). The obligation to fulfil explicit but incomplete contracts does not guarantee compliance with the obligations deriving from the hypothetical contract, exactly because the explicit contract is mute on the matters that the hypothetical contract helps to clarify. Consequently, it is not a viable solution simply to ask a judge to protect the stakeholders by applying the default rules of the explicit contract. The judge will inevitably be less well-informed about unexpected contingencies than are the parties. Either s/he may cleave mechanically to the 'letter' of the incomplete contract, thereby failing to comprehend what the underlying hypothetical contract suggest upon the appearance of unexpected events, or s/he may be excessively discretionary, applying his or her notion of what is 'fair' irrespective of the hypothetical contract between the parties. Therefore, recourse to a judge as interpreter of the implicit agreement (construed in terms of 'good faith') requires some interpretative standard beyond the letter of the contract; a standard established by the best-informed parties – namely, the firm and its stakeholders.²⁵

One might ask whether this could not be done by a sharp reform of company law. Personally, I do not believe this to be necessary. Nor do I believe that any radical reform need be made to the notion of the 'societary interest' of the various types of firm (i.e. I do not deem necessary that the definition given in section 4.2 be directly enacted in law). I prefer an indirect approach. The law can create the conditions under which the shareholder's interest *coincides* with that of the stakeholders: doing so by imposing certain requisites on the governance and accountability criteria that do not entail the explicit replacement of any reference to 'shareholder value' with a symmetrical reference to 'stakeholder value', but instead

²⁵ This is how I interpret the corporate code of ethics in Sacconi (2000).

induce a game to be played by the firm so that this coincidence of the two be reached endogenously. To be more precise, I suggest that the following reforms could reasonably be made to company law:

- a) Compatibly with the recent reform of the Italian civil code (legislative decree 6/03), impartial representatives of the stakeholders' interests should be included among the independent members of the board of directors (in the Anglo-Saxon-like 'monist model') or the independent members of the supervisory committee (in the Germany-like 'dualist model'). More appropriate, however, to this end is the dualist model, whose supervisory committee (legislative decree 6/03, art. 2409, xii, xiii) could be a mouthpiece for the stakeholders – which typically do not seek to exercise control over management decisions but instead what to be ensured that managerial authority is not abused – which gives them representation and influence;
- b) the obligation to render transparent and objective account on all relevant matters that are material to the stakeholders (completely and objectively identified) by adopting social accounting and reporting rules;
- c) the obligation to safeguard the *trust* and *reputation* of the firm *vis-à-vis* all its stakeholders.

Indubitably – given that the presence of independent members would not be a particularly radical innovation – the most novel features of these proposals are points (b) and (c). If firms were obliged to give precise account of their operations to all their stakeholders, and also of the effects of those operations upon them, they would be compelled to reveal information that might effectively influence their reputation, especially if these stakeholders had representatives on the board of directors (or the supervisory committee) who could verify and make use of that information. This would indirectly induce firms to show greater concern for behaviour affecting fiduciary relations. However, it would not prevent situations from arising in which this concern is not deeply felt by the shareholder – for example, when a short-period or speculative business vision prevails, or when the owners are simply affected by myopia. In these cases, if the company's interest is defined in traditional terms, the owner can legitimately ignore the judgements of the stakeholders, even when social accounting and reporting has been properly done. Point (c), therefore, makes it compulsory for the firm to attend to its reputation construed *as a good in itself* and functioning as a proxy for the extended interests of the stakeholders. Behaviour negligent of the firm's reputation among its stakeholders would be in breach of a principle of good management. All together, these norms would compel firms always to play a game in which reputation is a payoff to be pursued, so that they adopt a long-run perspective typical of the search for reputation and discard short-run self-interested behaviour. Of course, this argument rests entirely on the hypothesis that the search for reputation is a private incentive providing enough force for CSR to come about. This, as will shortly become clear, introduces the theme of self-regulation.

8. Is it possible to affirm CSR by means of self-regulation?

8.1. *Self-regulation (1): the discretionary approach.*

We may distinguish between two approaches to the CSR self-regulation. The first I call the *discretionary approach* (which is the one preferred by the rear-guard of firms). Its basic tenet is that there is no reason to add any further specification or constraint apart from the *enlightened* self-interest of owners and those who runs the firm for their advantage. Respect for the stakeholders' claims will come about through free choice, or through the firm's free exercise of discretionality – the same discretionality which, as we have seen, constitutes the problem in the theory of residual rights of control.

Enlightened self-interest would thus be an endogenous force able to induce self-discipline because it induce to account for personal interest into the *long-run*. By virtue of the long-run, the firm, as it pursues the simple goal of profit maximisation, would be induced to respect the fiduciary relation with the stakeholders and take due consideration of their well-being. On this view, self-regulation is nothing other than *self-discipline* whereby the firm does not behave in a manner such to abuse the trust that stakeholders have placed in it. The firm does not self-impose any formal system of rules or adopt any explicit management system required to abide by standards or norms, even if voluntary: this is self-regulation *without* explicit rules (and therefore we may term it *weak self-regulation*). It would be too easy to discredit this thesis by considering in its favour the only argument based on the obviously unrealistic hypothesis of perfect competition and the 'invisible hand' of the market. If competition were perfect, no firms would exist in the sense of being alternative institutions of governance to minimize transaction costs. It is therefore obvious that the argument does not pertain to the ideal world in which the 'invisible hand' operates. I shall instead take this thesis at its best, although even in this case it fails.

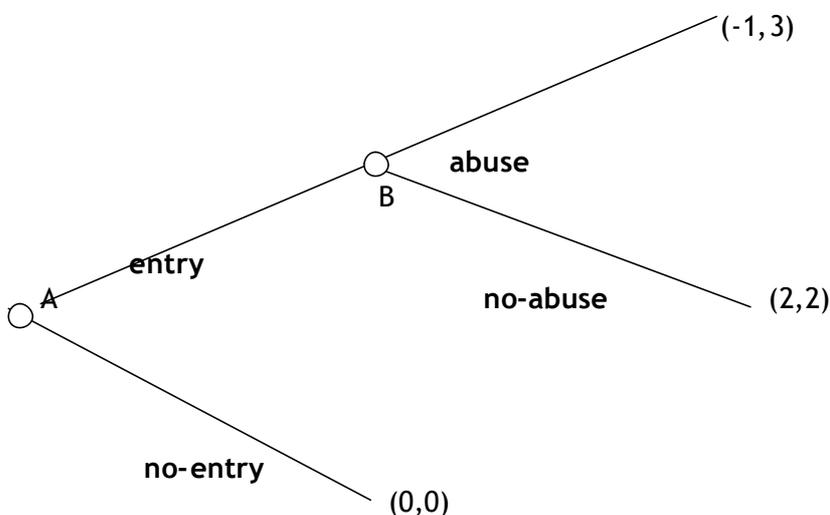
At best, the thesis maintains that enlightened self-interest leads to respecting trust of stakeholders, and therefore not to abuse them, in that the firm recognizes the importance of safeguarding and enhancing its **reputation**, which depends on non-abuse of the stakeholders. Reputation is one of the most valuable, albeit intangible, of the firm's assets. It is reputation that induces the stakeholders to trust the firm and consequently to cooperate with it, so that transactions come about at low costs of control or bargaining.

Moreover, reputation activates some sort of self-fuelling virtuous circle which leads to spontaneous compliance with social norms – that is, makes them self-enforcing. In fact, compliance with the norms *creates* reputation; reputation *induces* a cooperative response from the stakeholders; those who abide by the norms are thus *offered* a benefit; and this benefit acts as an *incentive* for complying with the norms. Unfortunately, however, reputation does not support the idea of *weak* self-regulation based on discretionality of the firm, which, without imposing any constraint or explicit rule upon itself, would decide unilaterally what actions to undertake in the best interest of both the firm itself and its stakeholders. To understand this, we must delve at least a little way into the theory of games of reputation.

8.2. The reputation game.

At the basis of the reputation mechanism lies a simple interactive situation (called *the trust game*, see Fig. 2) representing a transaction based on the fiduciary relation between a stakeholder A and the firm B. The stakeholder must decide whether or not to place its trust in the firm, entering or not into an exchange relation with it (assume that if it does so, necessarily he is making a specific investment). The firm then decides between abusing and not abusing. If, after the stakeholder has entered, the firm does not abuse its trust, there will be a reasonably good outcome for both. However, if the stakeholder places trust in the firm, the latter has an interest in abusing that trust, because in the current game this is the more remunerative option. Consequently, the stakeholder will not grant its trust and the transaction will not take place.

The idea underlying of a game of reputation is that there is an alternative solution which permits the transaction between the two parties to take place if the basic game is infinitely iterated, and if an incentive is thus created for the firm to protect its reputation (Fudenberg and Tirole 1991, ch. 9). We thus have an infinitely repeated game (which expresses the idea of long-run) whose stage-game is the trust game already defined. The players in the game are on the one hand an infinite series of stakeholders, called A_i (where i denotes the order of entry into the game), each of them lasting only for the stage-game in which they decide whether to enter or not enter (and are therefore short-run players), and on the other a firm (B), the long-run player, which last throughout all repetitions of the game.



(Figure 2. The trust game)

Information is crucial for the players: each A_i is uncertain about the *type* of B – in that B may be a *type* that never abuses trust or a *type* that always abuses it, or even a *type* that abuses with a certain probability and does not abuse with the residual probability. Let us assume that the various stakeholders believe that there exist only *pure types* which always abuse or do not abuse trust (and at most only two types of minimum deviations from these pure types, which represent the possibility of an *mistake* by the

pure types). Hence the *types* can be understood as commitments to the stereotyped use of a given basic game strategy (unless a mistake occurs). For every player A_i , all these types of B have some *a priori* positive probability (and, in particular, the type that never abuses - which for simplicity we may call 'honest' - is assigned positive, though very low, probability). At each stage game the current A_i player changes his beliefs (the probability assigned to types) according to what he has learned from the previous stage game. In each of these stages, the conditioned probabilities of the types are updated on the basis of the evidence furnished by the manner in which B has played the previous game.

Player B's reputation is the probability assigned by each player A_i in the current stage to the various types of player B. Player B's reputation of being a certain type increases as evidence is gathered which confirms that *type* (the probability of a *type* increases with the observation of those actions whose likelihood is positive, given the type), but it diminishes dramatically if a single observation is made that falsifies the type (if, in any stage, an action by B whose likelihood given its type is 0 is observed, the conditioned probability of that type is nullified as well). Player B, on the other hand, is perfectly rational and informed, so that its strategic reasoning scheme also includes awareness of the limitedly informed reasoning performed by players A_i , and in particular it enables player B to predict about the mechanism by which they learn and update beliefs about types. The players' interests are such that each stakeholder A_i maximizes its benefits in the current game (i.e. is short-sighted) while B is interested in long-run benefit. B may therefore be more or less far-sighted on the basis of a discount rate on future utilities which, in every period, increasingly reduces (even though at a marginally decreasing rate) payoffs associated to further outcomes of the repeated game.

These factors influence calculation of the players' rational choices. On the basis of its calculation of expected utility, each A_i chooses between **entry** and **no-entry** in light of the current conditioned probability of the types of B. Obviously, in the first stages-games, the probability of types is such that the first players A_i will usually not place their trust in B, because the expected utility for **entry** is less than that of the alternative. Sooner or later, however, some A_i will decide to trust B if they have observed a series of **non-abuse**, as a result of which the conditioned probability of the honest type has increased sufficiently to give the **entry** choice an expected utility greater than **no-entry**. There is always a calculable number of learning periods (in which learning must never contradict the hypothesis that B is an honest type) necessary for the probability of the type in question to reach the critical threshold p^* above which a short-run player A_i for the first time will rationally decide to trust B.

Analysing player B's choices requires consideration of the equilibrium strategies of the iterated game. At a first sight, B may opt for the equilibrium strategy of each stage-game, namely **abuse**, which is certainly the best response to the choices made by the players A_i in the first periods. This strategy by player B gives rise to an equilibrium profile in the iterated game which outcome consist of an infinite series of (**no-entry, abuse**). However, player B has a different strategy available, which consists in exploitation of its knowledge of the mechanism by which the beliefs of the various A_i are updated. It may choose to simulate the behaviour of the 'honest' type until the stage occurs in which the conditioned probability of this type reaches the critical level p^* at which the first A_i will enter. At this point, B calculates whether to play the **no-abuse** action, and consequently induce the players A_i to enter again, or

to profit from the first opportunity to defect by choosing **abuse**, thereby gaining a unilateral advantage from the (-1, 3) outcome on the first time but thereafter condemning itself to an infinite series of (0, 0) outcomes. If B is not impatient, and therefore if the discount rate of future utilities does not excessively reduce the value of the future prospects of cooperation, infinite outcomes of future cooperation (which begin once the first A_i has entered) are able to off-set the cost of the initial series of null outcomes (in which no A_i enters but B does not abuse), and to thwart the incentive to take advantage of an individual stakeholder as it enters. One possible B's rational strategy therefore is to sustain its reputation and to induce the sequence of stakeholders to trust it. The best response to this strategy by stakeholders A_i , from the time that the first of them has placed its trust in B, is to continue to be trustful until they observe a period in which B abuses. This gives rise to a game equilibrium profile in which a series of cooperative outcomes (**entry, non-abuse**) are observed from a certain point onwards.²⁶ Hence, the long-run search for reputation induces the firm to behave as if it wants to fulfil its fiduciary duties towards the stakeholders.

However, it is essential to understand the conditions under which this result holds:

- a) *signalling the types*: the firm must be able to signal the possibility that it is an honest type which does not abuse trust;
- b) *quasi-simultaneity*: the firm and the stakeholder must act and observe the result of each game simultaneously, for if the stakeholder acts first, the firm would have no reason to reveal its choice had the stakeholder not entered, so that there is no basis for *learning*;
- c) *observability of the results*: at the end of each stage-game, the stakeholder must be able to observe the outcome of the firm's choice without ambiguity, and it must be able to determine without ambiguity whether the firm has behaved according to a *type*. Because types can also be viewed as commitments (to a certain game action), the essential condition is that at the end of each stage-game each stakeholder (the current one) should be able to observe that 'what had to be done has been done';
- d) *shared knowledge among stakeholders*: each stakeholder must be able to transmit what it has learnt in a given period to the stakeholder that comes next: that is, all the stakeholders in succession must have the same judgement on the firm's fulfilment of its commitments;
- e) *absence of optimal mixed strategies types*: were the firm able to calculate the probability of *non-abuse* at which the stakeholder would be indifferent between entering and not entering, then it would abide by its *type* that refrain from abusing only at that minimum indispensable level of probability. This type would obviously not induce an equilibrium in which the firm does not abuse, but instead one in which it abuses with the maximum possible probability compatible with maintaining stakeholder indifference between entry and non-entry. Hence mixed equilibrium strategies must be excluded for the virtuous effect of reputation has to emerge.

In general, these conditions *are not* spontaneously fulfilled in situations relevant to the purposes of CSR, the consequence being that weak self-regulation (based on simple enlightened self-interest) normally fails.

²⁶ For an illustration of this model and result see Sacconi (1997, 2000).

9. The fragility of reputation and the cognitive role of norms

9.1. *Self-regulation (2): voluntary compliance with explicit norms.*

The main reason for weak self-regulation fails is the *cognitive fragility* of reputation. This is evinced by conditions a, b, c and d above, all of which refer to the knowledge that the players must possess if the model is to hold true. Accumulating reputation may be prohibitively difficult if in order to show that a commitment has been maintained it is needed enabling *each* stakeholder to observe that *concrete* actions have been undertaken, or that the *concrete* results have been obtained, so that they match their description established *ex ante* in a commitment announced by the firm (a possible *type*). Consider the following situations:

- *incomplete contracts*: the contract does not contain clauses covering unforeseen contingencies, so that there is no concrete benchmark against which to assess claims of renegotiation when unforeseen events occur;
- *unobservable quality*: the customer may not be able to verify the quality of a good or service on the basis of the information available to him or her by inspection or experience, so that s/he must 'trust' the opinion of an expert (usually the supplier);
- *organizational authority*: the 'boss' takes genuinely discretionary decisions with regard to tasks to be ordered the employees by virtue of the managerial authority granted to him/her by work contracts;
- *collusion*: the firm has resources by which is able to reach collusive agreements with agents of customers, suppliers or with official acting in the name of the 'public', which can induce them to act in conflict of interest to their 'principals'. Information about these agreements is reserved, so those not present when they were reached are unable to determine whether or not an illicit collusive exchange has taken place.

These are all settings in which information or knowledge about the firm's action is incomplete or highly asymmetric. Either commitments have not been defined in relation to unforeseen events, and therefore cannot be verified, or their fulfilment is not observable. The problem is that incomplete information makes it impossible to determine whether 'what had to be done has been done': either it was not established *ex ante*, so that there is nothing to verify, or it is impossible to observe results by which it can at least be inferred whether the commitment has been respected (since the result coincides with at least one of the possible results contemplated *ex ante*). Activation of the reputation mechanism is obstructed by a cognitive gap.

Self-regulation in the strict sense must therefore be viewed as the remedy for this cognitive gap. Rather than responding to concern over the enforcement of CSR norms by an external authority, it concentrates on the need to create the cognitive and informational bases that enable the social mechanism of reputation – with its endogenous rewards and punishments – to function properly. This comes about through the voluntarily-taken decision to accept *explicit norms with an appropriate structure* decided by the firm in the light of a multi-stakeholder social dialogue such to configure their impartial

acceptability.²⁷ For this reason, self-regulation is a voluntary but *not* discretionary approach. Voluntariness resides in the decision to endorse an *explicitly announced standard* for the firm's management system which is *ex ante* shared knowledge among the firm and its stakeholders.²⁸ This standard sets out general principles, whose contents are such to elicit stakeholder consensus, as well as explicit commitments to compliance with principles and rules which are to be known *ex ante* by stakeholders. It is clear that stakeholders' consensus can be more easily obtained if the standard relative to the strategic management system, intended to ensure CSR, is established by the firm through explicit dialogue with the stakeholders. This should come about both through the firm's direct relations with its most important stakeholders and through forms of multi-stakeholder social dialogue at the local, national or supranational levels.

However, dialogue does not detract from the voluntary nature of the agreed standards; nor does it preclude that compliance may then be obtained via the self-enforcement of the constraints and fiduciary duties established by the standard. How this occurs can once again be explained by referring to the reputation mechanism. The standard, and the procedures ensuring compliance with it, are announced *ex ante*; and it is *on these* – not in relation to particular (unforeseen) events or to particular (unobservable) actions or outcomes – that firm and stakeholders pass homogeneous judgement on *ex post* compliance with them. It is thus possible to activate the reputation reward and punishment mechanism, which generates endogenous incentives to comply with the standard. Everything rotates around the gap-filling function performed by the standard of CSR management system, whereby the firm's fiduciary duties towards its stakeholders are made explicit and announced. These duties assert – in the appropriate form – what is to be expected of the firm in *unexpected* situations, too, or in ones where the results of actions *are not observable*.²⁹

9.2. The logic of a CSR strategic management standard.

The *logic* of the management system for CSR (and of the standard that regulates it) is the logic that the firm's strategic behaviour must conform to in a context of incomplete information in order that reputation effects are reactivated. It has three components, as follows.

A. Generality and abstractness of principles

²⁷ A different explanation of the idea of self-imposition via individual rational choice of an explicit moral constraint, with its own content and structure identifiable with a social contract, consists in the idea of "constrained maximization" put forward by Gauthier (1986). Although for Gauthier compliance is not based on reputation effects, the idea of a self-imposed constraint whose structure and content is a social contract amounts to broadly the same I'm defending in this essay.

²⁸ This approach is the one preferred by the firms which, at national and European level, are endeavouring to define models and standards for the voluntary introduction of CSR management systems. Most notable in Italy are those firms that have participated in the Q-RES Project or have decided to apply the GBS and Accountability 1000 standards in their social accounting and reporting practices, or to promote and adopt SA8000.

²⁹ This argument was first put forward by Kreps (1990), who in effect talks of "corporate culture" (with a reference to "context-dependent" principles which I do not find convincing) and then developed, mainly with reference to codes of ethics, in Sacconi (1997) and Sacconi (2000).

The principles define the *vision* of the social contract that each firm proposes to its stakeholders (which must therefore be completely identified). These principles must consequently offer fair treatment acceptable to each stakeholder. They are abstract and general in form, so that they apply to a wide variety of events, including those which cannot be predicted or described beforehand. Consequently, their application does not require a *detailed description* of the situation; all that is necessary is recognition of the presence of certain abstract features which reflect a pattern established at the outset. Unlike detailed rules and contracts, principles never remain mute when unexpected situations arise: they cover both foreseen events and unforeseen ones (which instantiate the abstract feature identified from the outset). It is true that their range of application is vague with respect to unforeseen states of the world. Yet vagueness can be governed by the logic of moral reasoning and the application of modern fuzzy logic. Without going into technicalities, one may say that an abstract and general principle defines a set of states of the world as its application domain, membership of which is a ‘matter of degree’ (fuzziness). Once it is established the membership threshold above which it is conventionally accepted by default that a state satisfy the characteristics to be said a member of the domain, a pre-established conduct can be *ex ante* defined conditional simply on satisfaction of that threshold, without any request on the *ex ante* knowledge of the relevant states of the world.³⁰

B. Precautionary protocols of behaviour

Definition of the principles allows identification of areas of potential opportunism where interactions between stakeholders and firm put those principles at risk. For each of these risk areas, precautionary rules of behaviour can be established which assure the relevant stakeholder that a particular form of opportunism has been avoided. The distinctive feature of these rules is that their implementation is not conditional on the actual occurrence of concrete foreseen situations. More simply, they are applied when the risk that a principle will be breached exceeds a pre-announced threshold. They are consequently applied on the basis of a default logic (*in the absence of evidence to the contrary*): membership of the event in the domain of the principles need only exceed a given threshold of risk/vagueness for the rules to be applied (to say it in fuzzy logic terms, it is sufficient that it appears by a common fuzzy judgment that membership into the domain of a principle exceeds the threshold over which compliance with a given precautionary rule is due; see Figure 3 for a schematic illustration). Hence, their conditions of implementation can be established *ex ante* by the firm, and on these the stakeholder may legitimately form expectations about the firm’s behaviour. Their application constitutes the evidence that no principle has been intentionally breached, and consequently that the firm’s reputation is well-deserved and that trust in it is well-placed.

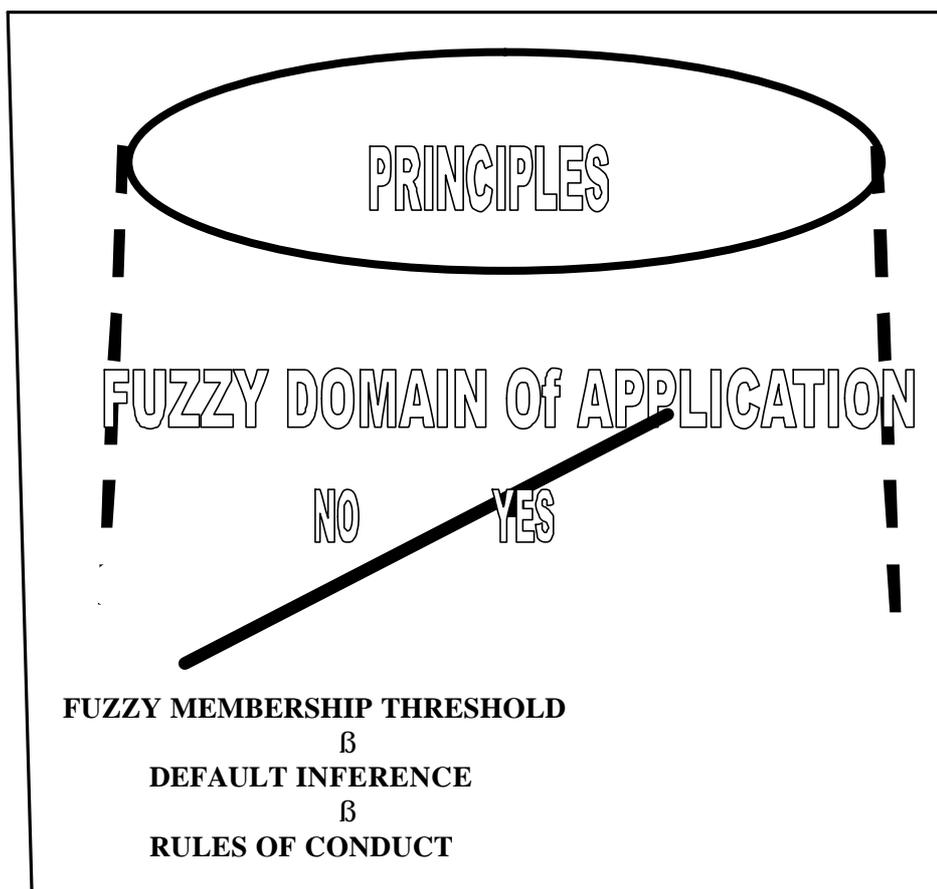
C. Communication and dialogue with the stakeholders

Principles and precautionary rules of behaviour must be communicated, given that reputation depends on them. The stakeholders base their judgements on the match among principles and rules announced *ex*

³⁰ See Sacconi (2000) and Sacconi (2003) for a technical description of the use of fuzzy logics and default logic in the case of general principles and procedures implementing a corporate code of ethics. For a schematic illustration see Figure 3.

ante, level of membership into the principles domain exhibited by any events have occurred, and the behaviour adopted. Essential, therefore, is *social accounting* and *reporting* of the firm’s performance in relation to the principles and rules announced. As a consequence, such accounting and reporting must do more than set out ‘social results’ as unintended by-products of economic activity undertaken for other purposes (profit, for example). If the firm does not give account of the intentional achievement of objectives defined in terms of duties fulfilment, which have been established *ex ante* in relation to the principles and the rules of behaviour, why should the stakeholder attach a fiduciary value to actions and results whose occurrence may have been entirely fortuitous?

Once the firm has communicated its principles, rules and social performance, incorporation of appraisal by the stakeholder into its corporate decision process is entirely in the firm’s interest. This equates to developing the firm’s capacity to judge as an impartial spectator from the point of view of the average stakeholder – neither malevolent nor benevolent – so that reactions to its behaviour can be anticipated and crises of reputation forestalled.



(Figure 3. Graphic representation of the relation between principles and rules of behaviour based on a fuzzy domain of application and default inference)

This capacity can be fostered by developing dialogue with the stakeholders in all the phases of the CSR management system,³¹ as follows.

- *'Enunciation' of the social contract stating the firm's fiduciary duties towards its stakeholders.* If this statement is formulated on consultation with the stakeholders, its acceptability is ensured and the parameters with which behaviour is assessed are known to both parties.
- *Internal management and implementation.* The presence of external members on internal board or committee – both corporate board of directors or those committees set up *ad hoc* to manage CSR, like an ethics committee – enables representation of the stakeholders' points of view and prevents divergence between the *ex post* assessments by the two parties.
- *Social accountability.* Dialogue with stakeholders identifies the areas of their effective interest for which account should be made, and therefore ensures the relevance of social communications.
- *Verification by an independent third-party.* Multi-stakeholder bodies – based on multi-stakeholder dialog – may control third-party verification and certification of CSR management standard compliance, preventing auditors conflicts of interest and granting credibility to the entire system.

³¹ It is easy to see that the *Q-RES management guidelines* developed by the *Q-RES Project* (cf. CELE, *Q-RES Project*, 2000), defining a standard for CSR management system, closely adhere to the logic set out in the main text. The Q-RES Standard lays down a process comprising the following six steps:

- 1) *Corporate ethical vision*: this is not a simple statement of the firm's mission but a *vision* of the social contract that the firm offers to its stakeholders, by which the company presents his own interpretation of the fair balance amongst of stakeholders' interests that will inspire day by day management of the firm;
- 2) *Code of ethics*: (i) general principles defining the firm's rights and duties *vis-à-vis* each category of stakeholder; (ii) preventive rules of conduct on every area of interaction between the firm and its stakeholders that is at risk of opportunism; these rule forestall typical forms of opportunism and state the standards of precautionary behaviour recommended;
- 3) *Ethical training*: enables the organisation members to give proper interpretation of organizational events in the light of their bearing on ethical principles and allows the transmission of a sense of commitment;
- 4) *Organizational systems of implementation*: an ethics committee which impartially represents the points of view of the various stakeholders; top-down control (auditing); development of bottom-up dialogue to integrate CSR into work tasks and objectives; systems for the assessment and material and non-material incentives of personnel;
- 5) *Social accountability*: external communication of principles and rules; a social report accounting for the relation between performance and commitments by means of proper illustration of relevant and material information expressed both by indicators or qualitatively; synoptic statement of the results achieved in relation to each stakeholder as regards both the economic value distributed among stakeholders and the other benefits or costs allocated among them; inclusion of the stakeholders' point of view;
- 6) *External verification and certification* of CSR by third-party (independent) auditors which examine evidence on each CSR tool and the results obtained in the various areas of management (resources, quality of products and services, etc.).

10. Stakeholders' moral preferences as a reinforcement factor of motivation

10.1. Conformist preferences.

Compliance with a CSR management standard can work effectively as a parameter with which to assess a firm's reputation. Yet a firm which has acquired a reputation as a relatively mild abuser of its stakeholders' trust can nonetheless gain their cooperation if the incomplete fulfilment of its duties generates to the stakeholder an expected payoff not less than that promised by his alternative choice of withdrawal from all transactions. In this case, the firm would fulfil its duties only to the minimum extent necessary to dissuade the stakeholder from exiting the relation (although it is quite unfair).³²

This introduces a second set of virtues attaching to explicit but self-imposed ethical norms (or standards for a CSR management system). Given that these entail ethical principles (set out in the social contract) and duties of behaviour – and are therefore couched in the language of *deontology* – they enable resort to stakeholder motivations which extend beyond the mere material advantage deriving from transactions with the firm. Many stakeholders, in fact, have preferences that are not purely self-interested or geared to material advantages (consequences). These stakeholders also place importance on the firm's fulfilment of duties in their regard deriving from the social contract, especially if the firm enunciates these duties in codes of ethics and communicates them externally. Hence, any deviation from the CSR standard, or from the firm's ethical commitments, is punished more than would be the case if simple material interest were concerned. The stock market panic that followed the American corporate scandals of 2002 can be construed as a punishment of the managements involved for breaching the principles of loyalty and accountability that they had declaredly endorsed. Another example is the boycotting by 'ethical investors' and 'responsible consumers' of certain multinationals whose operations in developing countries are in flagrant conflict with the ethical standards obtaining in their countries of origin. Although a traditional economist would tend to regard these phenomena as extraneous to economic behaviour; in the light of recent developments in economic theory the opposite view can be taken.

³² This requires examination of the possibility that condition (e) at the end of section 8.2 ('absence of optimal mixed strategies') may still obstruct the use of reputation to support self-regulation. Note that in the game illustrated by Fig. 2, the payoff to player A from the non-entry strategy is zero, so that B's optimal strategy to persuade A to enter is the offer of an infinitesimally positive payoff by the appropriate probability mixture of abuse and no-abuse. Consider, however, that the dependence of reputation on the 'general principles/precautionary procedures' pairing, where the latter are activated by default on the basis of satisfaction of a sufficient relevance threshold for the principle, would suffice in itself to render the firm unable to calculate the optimal mixed strategy with the minimum probability of conformity necessary to prevent the stakeholder's exit and enable the firm to act opportunistically in all other cases. In fact, the firm is unable to foresee *ex ante all* the possible states of the world in which it should implement a given action. Consequently, it is unable to calculate beforehand the probability with which it should abuse or not abuse trust optimally (given that the context of incomplete information would cause it to commit gross errors). The firm has a more rigid decision-making procedure available, which is to apply the rules in precautionary manner whenever a principle is involved to a sufficient extent (cf. Sacconi 2003).

The most recent analyses of preferences of economic agents (whether workers, consumers, financiers, or others) show that they comprise various motives for action.³³ It is possible to distinguish between *preferences for the consequences* of action, which may be as much egoistic and self-centred as they are altruistic, and *conformist preferences for actions in themselves*, in which case utility arises from fulfilment of the expectation that there will be reciprocal compliance with a mutually accepted ideal norm. If within an organization there is an agreed constitutive pact (an ideal, or better, a written moral code), then **Ego** derives utility from the expectation that **Alter** will conform if **Ego** does so as well. Put otherwise, if **Ego** expects that **Alter's** action, given **Ego's**, will generate closer proximity to the state coincident with the principle, then **Ego feels better** (conversely **Ego** feels *worse* if it generates greater distance). These preferences for reciprocal conformity are balanced with other material motives or preferences, so that when their weight is sufficiently great, there arise outcomes of agents' reciprocal best responses (equilibria) which would not exist if consequentialist preferences alone prevailed.³⁴

For example, consider an organization in which the internal members – who have influence over the decision-making process – are able to establish whether a part of the surplus produced can be used to improve the well-being of external stakeholders (e.g. consumers or the beneficiaries of a service, hired labour without decision-making power, suppliers or members of the surrounding community, etc.). Assume that by colluding among themselves they can appropriate all the value from which, if their behaviour were less egoistic, the other stakeholders would benefit. But also hypothesise that the internal members have explicitly admitted that there exists in the history of the organization a constitutive contract such that cooperation among the members must, according to a principle of fairness, also be to the advantage of the external stakeholders. Finally hypothesise that this constitutive pact establishes an ideal principle which determines the conformist preferences of the organization's members. Each of the latter derives utility from complying with the constitutive pact if they expect the other members to comply with it – given that the others expect that they themselves will do so. In other words, the expectation of reciprocal compliance induces in every agent a conformist preference for action that complies with that principle.

If this *conformist preference*, balanced with traditional egoistic *consequentialist preferences*, is sufficiently strong, as the members of the organization pursue their *overall* utility function, they may converge to an outcome of their interaction which is advantageous to the external stakeholders – that is, they refrain from appropriating the entire surplus. Note that if this component of preferences is lacking, or if it exerts insufficient weight in the agents' motivational system, the said outcome may not be achieved through the rational choices of the organization's members. On the contrary, because it results from

³³ There are numerous studies that go beyond the description of preferences in terms of mere self-interest alone: see Bernheim (1994), Rabin (1995), Chareness and Rabin (2002), Sugden (1998), Frey (1997), Falck and Fishbaker (1999), Fehr and Schmidt (2001), the studies collected in Sacco and Zamagni (2002) and also Sacconi (2002) and Grimalda and Sacconi (2003).

³⁴ See Sacconi (2002); for a proof based on the theory of psychological games see Grimalda and Sacconi (2003).

preference-based choices, a game equilibrium arises within the organization as a self-sustaining solution.³⁵

10.2. Motivational reinforcement.

We may now evaluate the implications of the foregoing analysis by considering the repeated trust game between the firm and its external stakeholders. In this game, the firm may seek to accumulate reputation in the eyes of the stakeholders by virtue of its compliance with an explicit CSR management standard – even though, as we have seen, it may seek also to reduce this compliance to a minimum necessary not to induce the stakeholder to stay out. The firm is a single player with a low level of conformist preferences, although it has an ethical code that may constitute the ideal around which such preferences can be generated. (Therefore, contrary to the argument that could be put forward concerning non-profit enterprises or cooperatives –on this see Sacconi (2002) – I do not argue here that in general conformist preferences are decisive in ensuring the onset of an equilibrium in which the firm spontaneously complies with CSR management standards.)

Unlike the firm, however, the external stakeholders have strong conformist preferences, in that if the firm has adopted a standard and a code of ethics which affirms an ideal of fairness and of mutually advantageous cooperation with them – and if they expect the firm to comply with the principles and they are expected to reciprocate – then the stakeholders associate high intrinsic utility with the fact that the firm is behaving consistently with the code of ethics and principles of mutual advantage and fairness. By contrast, if the stakeholders behave cooperatively and expect the firm to know that they do so, any deviation by the firm from its principles will affect them not only materially but also because it contradicts their conformist principles. (This seems to be the case of the ethical investors or responsible consumers which expect that the respect for human rights and concern over environmental impact currently required in their countries, and on which companies are usually accepted as corporate good citizens at the national level, should be shown by those same companies wherever they operate.)

This prompts the question whether this hypothesis is such to preclude the situation described at the beginning of the previous section: viz, the situation in which the firm seeks to calculate the strategy in which the probability of compliance is reduced to the minimum necessary to render the stakeholders indifferent between entry and non-entry, so that the firm is able to appropriate a large part of the surplus. By hypothesis, the firm will adopt this strategy only if it is able to induce the stakeholder to enter (see again the trust game illustrated in Figure 2). However, it can be shown that if the stakeholder is

³⁵ Note, however, this is only *one* of the possible equilibria arising from the interaction among the organization's members. But it is so in a very special manner: everything depends on the fact that the organization's members have appropriate beliefs about their reciprocal behaviour. Only if they effectively believe that they are reciprocally complying will the conformist component be activated and it will gain in force until it is capable to drive their behaviour. In the contrary case, egoistic preferences prevail, and in equilibrium the predominant behaviour will be that which induces them entirely to expropriate the external stakeholders. For a proof see Sacconi (2002) and Grimalda and Sacconi (2003).

conformist, it will refuse to enter, with the consequence that the firm's strategy does not induce an equilibrium.

Indeed, if one examines the significance (or the intention) of the stakeholder's decision whether or not to enter, given the hypothesis that the firm wants to adopt the mixed strategy, the stakeholder is entirely unable to alter the situation's proximity to the social contract ideal, nor the degree of conformity with the moral code, because the value resulting from its choice in terms of approximation is always close to zero (whether it decides to submit to the firm's strategy or whether it eschews the transaction). However, when it considers the firm's behaviour, given that the latter expects that in equilibrium the former will want to enter, the significance of the firm's choice of the mixed strategy is a marked distancing from the ideal of the social contract, because the intention is to expropriate the stakeholder precisely when it assumes cooperative behaviour. This, therefore, negatively affects the conformist preferences of the stakeholder, which will be doubly dissatisfied with the firm's behaviour, for its material remuneration is close to zero and at the same time the firm's action in itself increases the distance from the principle. The addition of a negative value to the utility function resulting from the dissatisfaction of conformist preferences – that is, because of the negative effects on compliance due to the firm's expected action – entails that the stakeholder will not form a relation with the firm, and indeed will punish the firm more harshly than the penalty that its calculated non-compliance would warrant in the absence of conformist preferences. Put otherwise, the firm must increase material payments to the stakeholder in order to off-set the negative impact of its incomplete compliance with the principle on the stakeholder's preferences. But this obviously also entails a choice of behaviour increasingly less distant from fairness, which therefore also reduces the stakeholder's negative evaluation of the firm's compliance.

These considerations provide an entirely rational explanation for why firms whose codes of ethics, however skimpy, affirm commitment to corporate integrity (see *Arthur Andersen* or *Enron*) have been harshly punished on the stock exchange by a crisis of investor confidence, and perhaps to an extent exceeding the economic damage directly caused by the embezzlement. They provide an equally rational explanation for why multinationals which invest large sums to construct an image reflecting the moral preferences of certain groups of consumers in the developed countries (see *Nike* or, to a lesser extent, *Benetton*) then pay the high price of boycotts of their products because human rights have been violated in their supply chain – sometimes disproportionately so, if one considers that similar violations are perpetrated by local producers less exposed to the impact of public opinion. In short, conformist preferences amplify positive or negative reputation effects and make it less easy for firms to adopt the strategy of guaranteeing respect for a system of shared values only to the minimum indispensable to make indifferent consumers or investors between keeping their relations to the firm and breaking them.

11. How conformism defeats opportunistic mixed strategy equilibria: a formal proof

In this section I give a formal proof of the foregoing section main proposition, i.e. if we assume that at least stakeholders have (*inter alia*) conformist preferences, then the firm's opportunistic resort to a mixed

strategy of minimal compliance with a codes of ethics will not be anymore a rational strategy to the firm itself. This result is based on a previous work (Grimalda and Sacconi 2003), which I summarize here briefly, before applying it to the present context. Let me assume that stakeholders have not only self-interested motives to prefer but also ideological motives to prefer and that their accepted ideology coincides with the social contract of the firm - i.e. the guiding principle of an extended corporate governance system. In Grimalda and Sacconi (2003) we suggest that these two classes of motives can be accounted for by two types of preferences of the self and by their relative mathematical representation in the corresponding utility function.

To begin with, it must be recognized that strategic interaction generates states of affairs which can be differently described according to their characteristics. A first description of states understands them as *consequences*. Consequences may be described as attributed just to the acting Self - what happens to the decision maker in any state - and this gives a basis to Self-interest: the Self has preference over consequence which are Self-referred. On the contrary, consequences may be attributed to every person (extended consequences) as they can be understood as what happens to whichever individual. This makes a case for some sort of impartial consequentialist ethics like utilitarianism, or altruism etc. In general, if a player defines his preferences only over states *described as consequences*, then he has *consequentialist personal preferences*.

Let me come to the second type of preferences, what we call *conformist personal preferences*. States description is no less important here, but them are now described as sets of interdependent actions characterised in terms of whether they conform or not to a given abstract principle or ideal. This ideal can be captured by means of a function of individual first type utilities attached to states, which measures fairness of any configuration of states of the world. Fix a pattern of behaviours (a vector of strategies) and define it as perfectly *deontological* if it fully conforms to the abstract principle of fairness - that is it maximises the function just defined. We call such a state the *ideal*. Then we look after a degree of conformity to the ideal appended to each strategy choice by seeing whether the *ideal* comes about through the choice of each player given what he believes about other parties' choice. For each strategy combination each player's conformist preferences will follow by a measure of expected reciprocal conformity, which is based on:

- (a) an index of how much the player himself contributes to carry out the ideal by conforming or deviating from it, given what he believes about the other player's choice;
- (b) an index of how much the other player contributes to carry out the ideal by conforming or deviating from it (as seen through the eyes of first player), given what second player believes (and first player believes that second player believes) that first player will do.

To be sure, there is a relation between the two kind of preferences - consequentialist and conformist. In fact, in order to define fairness, we look at distributions of payoffs, derived from *first type of preference* - i.e. material utilities. But this does not reduce second type preferences to the first one. First type utilities are no more than rough materials for the criterion defining second type preferences. What matters to the second description are not consequences or material payoffs as such, but a distributive property defined over payoffs, which is expressed by the function representing a principle of fairness.

Thus characterisation of second type preferences accords more with deontology than consequentialism. The more an expected state of affairs conforms to the ideal, the more it is preferred by a player (due to a measure of expected reciprocal conformity, which is at basis of each player's conformist preference over states). Moreover, there is no reason to link deontology to a belief that there is some objective source of value that would have an ontological reality "out there" (independent on the decision maker's affections). In fact, while conformist preferences depends on degrees of deontology, nonetheless deontology itself may be understood, as we do, simply as individual compliance with a fair distribution principle *that players could have rationally agreed in an ex ante hypothetical bargaining situation*.

Once these concepts are translated in a formal model, we can define a player's overall utility function that combines two parts (which we assume to be separable), i.e. the representations of consequentialist and conformist preferences. First, a form of the fairness-function T must be specified, which expresses formally the ideal. It has to be a mapping from the set of states (and first order utilities attached to them) to a fairness ordering ranging over states. Coherently with my discussion of contractarian principles basing the multiple-stakeholders model of corporate governance, a characterisation in contractarian terms of the ideal principle T is given by the Nash bargaining solution, i.e. *Nash social welfare function* N

$$T(\mathbf{s}) = N(U_1, \dots, U_N) = \prod_{i=1}^N (U_i - d_i)$$

where d_i represents the reservation utility that agents can get when the process of bargaining breaks down. In this case d_i coincides with the cost of each player's specific investments, which means that fair bargaining over the surplus may start only if parties are assured that they will end up at least with reimbursement of the cost they must bear in order to participate in cooperation.

Then let define the two personal indexes of conformity, which will be compounded in a measure of mutual expected conformity and will enter the utility function of players.³⁶ In this construction I take the point of view of player i (any other player j 's perspective is symmetrical).

A) Player i personal index of conformity:

Articolo II. It is player i degree of deviation from the ideal principle T (which varies from 0 to -1), due to player i choice given her expectation on the player j behaviour. It is normalised by the magnitude of difference between player i full conformity and no conformity at all conditional on player j choice

$$f_i(\mathbf{s}_i, b_i^1) = \frac{T(\mathbf{s}_i, b_i^1) - T^{MAX}(b_i^1)}{T^{MAX}(b_i^1) - T^{MIN}(b_i^1)}$$

where

b_i^1 is the belief of player i over player j action,

$T^{MAX}(b_i^1)$ is the maximum attainable by the function T given i 's belief,

$T^{MIN}(b_i^1)$ is the minimum attainable by the function T given i 's belief,

³⁶ In Grimalda and Sacconi (2003) we elaborate on Rabin (1993) in order to define our own model of reciprocity.

$T(\mathbf{s}_i, b_i^1)$ is the effective level attained by T when the player adopts strategy σ_i , given his belief on the other player's behaviour.

B) Estimation function of the second player index of conformity to the ideal:

Articolo III. It is player j degree of deviation form the ideal principle T (which also varies from 0 to -1), as seen through player i beliefs - as well normalised by the magnitude of difference between player j full conformity and no conformity at all given what he believes (and player i believes that he believes) about player i choice

$$\tilde{f}_j(b_i^1, b_i^2) = \frac{T(b_i^1, b_i^2) - T^{MAX}(b_i^2)}{T^{MAX}(b_i^2) - T^{MIN}(b_i^2)}$$

where

b_i^1 is player i *first order* belief over player j action (i.e. formally identical to a strategy of player j),

b_i^2 is player i *second order* belief over the belief of player j about the action adopted by player i (i.e. formally identical to a player i strategy predicted by player j).

These indexes are employed to build the following ideal component of the utility function

$$I_i [1 + \tilde{f}_j(b_i^2, b_i^1)] [1 + f_i(\mathbf{s}_i, b_i^1)]$$

where the weight λ_i - which may be any positive real number - is an exogenous psychological parameter that expresses how important in absolute deontology is within the motivational system of player i before expected reciprocity is considered. As a whole the formula works as follow: if player i perfectly conforms to the ideal, given her expectation, while the player j is also expected to perfectly conform, then the two individual indexes take values zero, so that the resulting utility value due to conformism is *1 time 1 time* I_i . Thus the maximum conformist utility value equals λ_i . On the contrary if a player does not entirely conform, while not expecting the other player also entirely to conform, then the two indexes take negative values (possibly -1). Thus the utility calculation for conformist reasons reduces to *(1-x) time (1-y)* (possibly both equal to zero) multiplied the weight λ_i and gives less than λ_i (possibly zero) as conformist utility value.

The overall utility function V_i is the linear combination of the two components

$$V_i(\mathbf{s}_i, b_i^1, b_i^2) = U_i(\mathbf{s}_i, b_i^1) + I_i [1 + \tilde{f}_j(b_i^2, b_i^1)] [1 + f_i(\mathbf{s}_i, b_i^1)]$$

It suggests that if a player predicts reciprocal conformism (as it enters the utility function), as far as weight λ_i is high, it is then possible that the overall utility of a strategy choice reverses the effect of player i simple consequentialist preferences represented by $U_i(\mathbf{s}_i, b_i)$. For example it may push the players to select strategies that they would never choose whether they rested on their material utility only.

That is enough for the theory of conformist preferences. Let now come back again to the infinitely repeated game of reputation between a long-run firm B and one short-run stakeholder at time A_t , whose stage-game is the game of trust (see fig. 2). The normal form of this stage-game is given in fig.4. To begin with, consider player's B type that adopts a mixed strategy $(2/3 \mathbf{a}, 1/3 \mathbf{no-a})$, i.e. B may try to

develop the reputation of being this type by playing the two pure strategies with the attached probability along all the repetitions of the game (let assume that these mixed strategy types can be understood and learned by the A_i players). This type, once the stakeholder believes it with probability one, makes each player A_i indifferent among all his pure or mixed strategies, since in every case he gets no more than 0. Typically player A_i 's best response is remaining indifferent between her pure strategies and using the mixed strategy ($\frac{1}{2} e$; $\frac{1}{2} no-e$), which gives as well A_i the expected payoff 0, and allows player B an expected payoff 1.33.

		Player B	
		<i>a</i>	<i>no-a</i>
Player A	<i>no-e</i>	0,0	0,0
	<i>e</i>	-1,3	2,2

fig. 4

Of course, this is not an equilibrium of the stage-game, for against any probability of entrance by players A_i player B's best response is to play abuse. However we are considering the repeated game, assuming that mixed-strategy-types of player B can accumulate reputation by playing repeatedly their characteristic strategy. This suggests the following repeated game equilibrium: B could even improve his stage-game payoff over the Stackelberg pure strategies payoff 2, by simply resorting to a commitment on the mixed strategy ($\frac{2}{3}-\epsilon$ *a*, $\frac{1}{3}+\epsilon$ *no-a*) (with ϵ as small as possible.) In fact, if player B is able to accumulate reputation that he is this type, player A_i necessarily enters in order to get the expected positive payoff 3ϵ (which, due to the infinitesimal ϵ , is practically nil), and gives player B a stage game expected payoff $2.66-\epsilon$. Notice that player B best response is then not to deviate to stable defection (**abuse**), for this would change his reputation till a player A_i will consider only possible a nearly complete abusive type, so that no more A_i players would enter thereafter. This suggests that there is a mixed strategy equilibrium in which, also by not complying entirely, player B (the firm) can get a reputation such that stakeholders enter and acquiesce to a firm appropriating large part of the surplus. But this implies that the firm would not carry out its fiduciary duties toward the stakeholders

This pessimistic result obtains when no role is played by conformist preferences. Let now the more complex representation of stakeholders' system of preferences enter the picture. I will be as parsimonious as possible in introducing motivations that make easier compliance with fiduciary duties on the part of the firm. Thus, as before, I assume that only stakeholders (for example employees or consumers, or investors) stick to the ideal of socially responsible firm as defined by a code of ethics or a voluntary CSR management system expressing the ideal model of a social contract amongst the firm and its stakeholders.

This presupposes that the firm has at least deliberated and signalled a commitment to such a code, but not necessarily that it has developed that kind of attitude that we call conformist preference toward the reciprocal compliance with it. Hence, put at 0 the λ_B parameter in the manager utility function by which I capture the utility function of player B, the "firm". On the contrary stakeholder A (I skip the individual index for simplicity of notation) has a positive weight λ_A and holds an *overall utility function* combining both consequentialist (self-interested) motives and conformist motives to act. Hence, the stakeholder's ideal of socially responsible governance of the firm may be calculated by means of the Nash social welfare function. Due to the very simplistic representation of the basic trust game, this welfare function is defined only over the firm's payoffs and the payoffs of the sole stakeholder taking part in the current stage-game, without considering that at any time there would be more stakeholders involved or that what the firm decides at one time may be relevant to other stakeholders at later times. I am confident that these more complex interactions would only reinforce the result of this section, as they would imply that more stakeholders' utility functions should be accounted for that include conformist preferences.

The relevant first order and second order player A's beliefs in this exercises are

$$\begin{aligned} b_A^1 &= (2/3a, 1/3 \text{ no-a}), & \text{in short } (2/3, 1/3) \\ b_A^2 &= \mathbf{e}, & b_A^2 = \mathbf{no-e} \end{aligned}$$

In other words I first define the overall utility function of player A for a situation in which she believes that player B will abuse with probability 2/3 and not abuse with probability 1/3, while she has the second order belief that players B predicts that she (player A) will enter. True, this is not the effective mixed strategy equilibrium of the repeated game without conformism, for this would be the case only for a infinitesimal ϵ added to the probability of **no-abuse** strategy. However we may disregard the small ϵ as it practically does not change the two players' payoffs, so that we may proceed as if the effective mixed strategy equilibrium were ($\mathbf{e}, (2/3, 1/3)$). Thereafter, I will contrast this utility value with the alternative utility value for the case that player A believes that player B will play the "equilibrium" mixed strategy (2/3, 1/3) but she will not play the entry strategy, so that her second order belief is that herself does not enter (formally $b_A^2 = \mathbf{no-e}$). In order to calculate player A's utility for the two alternatives (\mathbf{e} and **no-e**), we must consider :

- For strategy \mathbf{e}

A*: player A's conformity index when she chooses \mathbf{e} given that she believes player B is choosing (2/3, 1/3),

together with

B*: player B's conformity index based on player A's 2nd order beliefs, i.e. the index annexed to B choosing (2/3, 1/3) (and player A believing it) given that he believes that A chooses \mathbf{e} (and A believes that B believes it).

- For the strategy **no-e**

A**: player A's conformity index when she chooses **no-e** given that she believes that B plays (2/3, 1/3),

together with

B***: player B's conformity index based on player A's 2nd order beliefs, i.e. the conformity index resulting from B choosing (2/3 1/3) (and A believing it) given that he believes that A chooses *no-e* (and player A believes that B believes it).

Remember the form of player A's conformity index

$$f_A(s_A, b_A^1) = \frac{T(s_A, b_A^1) - T^{MAX}(b_A^1)}{T^{MAX}(b_A^1) - T^{MIN}(b_A^1)}$$

In a trust game with the given payoffs this index takes the following values:

Case A* (for player A strategy *e* given beliefs (2/3,1/3))

$$\frac{T(e, (2/3,1/3)) - T^{MAX}(2/3,1/3)}{T^{MAX}(2/3,1/3) - T^{MIN}(2/3,1/3)} = 0$$

Case A** (for player A strategy *non-e* given beliefs (2/3,1/3))

$$\frac{T(\text{no-e}, (2/3,1/3)) - T^{MAX}(2/3,1/3)}{T^{MAX}(2/3,1/3) - T^{MIN}(2/3,1/3)} = 0$$

In fact Nash social welfare function in case A* gives $T(e, (1/3,2/3)) = 0 \times 2.66$, and it is $T(\text{non-e}, (2/3,1/3)) = 0 \times 0$ in case A**. At the same time, $T_{max}(2/3, 1/3)$ and $T_{min}(2/3,1/3)$ are both 0 because player A's expected payoff is always 0 under player B strategy (2/3,1/3). The meaning of these 0-levels of the conformity index is better understood according to their interpretation as degrees of deviation from complete compliance with the ideal, conditioned on the other player expected choice. In both A* and A** cases the expected player B mixed strategy (2/3,1/3) nullifies any effort that payer A could make to enhance the level of ideal attainment. Whatever choice player A tries, in fact, the level of T is always 0. Thus A has not responsibility in any deviation from the maximum feasible level of T, given B's choice. Formally in fact the T level is at its maximum (which is also its minimum in this case) given player B opportunistic strategy that always reduces to 0 (or practically nil) player A's payoff.

Then consider the formula of player B conformity index based on player' A 2nd order beliefs

$$\tilde{f}_B(b_A^1, b_A^2) = \frac{T(b_A^1, b_A^2) - T^{MAX}(b_A^2)}{T^{MAX}(b_A^2) - T^{MIN}(b_A^2)}$$

This index takes the following values in the two cases (*e* and *no-e*) analysed concerning player B expected conformity in a trust game:

Case B* (for strategy (2/3,1/3) used by player B given his belief that A chooses *e*)

$$\frac{T((2/3,1/3), e) - T^{MAX}(\text{no-a}, e)}{T^{MAX}(\text{no-a}, e) - T^{MIN}(a, e)} = \frac{4}{7} = -0.57$$

$T^{MAX} = 2 \times 2$ is in fact the maximum value of *Nash product* that ensues when B does not abuse given that A enters, while $T^{MIN} = -1 \times 3$ is its minimum value ensuing when B abuses given that A enters. The ideal's value if B plays the mixed strategy when A enters is nevertheless zero due to $T = 0 \times 2.66$. Notice

that the index asks contrasting the T value brought about by player B opportunistic mixed strategy, in the case player A acquiesces with the T^{MAX} value that player B could attain given the same player A choice. But this implies a relevant deviation from maximal conformity conditional on A behaviour, which can be imputed entirely to player B decision to play his mixed strategy in place of his *no-a* strategy. In this case player B relevantly does not conform to the ideal and this results in the negative value taken by his conformity index.

Case B** (for strategy (2/3,1/3) used by player B when he believes that A chooses *no-e*, and player A believes that B believes it)

$$\frac{T((2/3,1/3), \mathbf{no-e}) - T^{MAX}(\mathbf{no-a}, \mathbf{no-e})}{T^{MAX}(\mathbf{no-a}, \mathbf{no-e}) - T^{MIN}(\mathbf{a}, \mathbf{no-e})} = 0$$

In fact, for whichever player B choice given the expected *no-e* by players A, T^{MAX} as well as T^{MIN} are both zero, and this is also true to the single payoff got by player A under (2/3,1/3). Given his belief *no-e*, players B cannot significantly deviate from the ideal, for also by playing his mixed strategy he is not accountable for a deviation from the maximal ideal's value given no entry by player A. Comparing B* and B**, it can be said that the intention of exploiting player A attitude to acquiesce implies a significant responsibility of B for a deviation (non conformity) to the ideal only conditional on the expectation that in effect player A will acquiesce, for it is exactly in this case that he does not reciprocate player A's conformity.

Finally, according to the formula of players A overall utility

$$V_A(\mathbf{s}_A, b_A^1, b_A^2) = U_A(\mathbf{s}_A, b_A^1) + I_A [1 + \tilde{f}_2(b_A^2, b_A^1)] [1 + f_A(\mathbf{s}_A, b_A^1)]$$

we can calculate overall utility values for the two alternative strategies *e* and *no-e* respectively, given that player A expects strategy (2/3,1/3) from player B, under the assumption that player B predicts player A choice, i.e. he believes (and player A believes that he believes) that player A uses either strategy *e* or strategy *non-e* respectively. The material payoff to player A if she enters under the equilibrium mixed strategy is 0 (accounting also for an additional infinitesimal probability ε attached to *no-a*, the payoff nevertheless is practically nil), while the conformist utility is based on the indexes A* and B*. Thus player A overall utility for strategy *e* is

$$V_A(\mathbf{e}, b_A^1, b_A^2) = 0 + \lambda_A(1+(-0.57))(1+0) = 0.43\lambda_A$$

On the other hand player A conformist utility of for strategy *no-e* is given by the indexes A** and B**, while her material payoff is again 0. Thus A's overall utility for strategy *no-e* is

$$V_A(\mathbf{no-e}, b_A^1, b_A^2) = 0 + \lambda_A(1+0)(1+0) = \lambda_A$$

A conclusion follows straightforwardly: player A with conformist preferences does not accept to play the mixed strategy equilibrium of the repeated trust game. As far as player B conformity index annexed to the equilibrium mixed strategy is negative and, *ceteris paribus*, weight λ_A the stakeholder attaches to conformity is positive, the logic of strategic choice under conformist preferences reverses the result of standard strategic calculation in a repeated trust game with mixed strategies. Consider that, granted λ_A is

positive, this result typically follows from the opportunistic nature of player B mixed strategy type. In fact it endeavours to minimize (till to nil) stakeholders' payoffs while convincing nevertheless them to enter in order enabling the firm to appropriate a part of the surplus larger than the "cooperative" payoff (2) in each stage-game. Of course were B ready to concede more to A, for example by rising the additional probability ϵ of **no abuse** to a substantial level, then A's firmness in rejecting **entry** would be lessened. However straightforward calculation shows that, *ceteris paribus*, A stops rejecting **entry** only when $T((2/3-\epsilon, 1/3+\epsilon), \epsilon)$ takes value 4, which means that probability of **no abuse** equates 1 - i.e. player B mixed strategy type degenerates to his "honest" type.

Summing up, it has been demonstrated the following remarkable result: if stakeholders share the ideal of a multi-stakeholder firm (i.e. the contractarian principle of a corporate code of ethics), and they get ideal utility from reciprocal expected conformity to that code -as far as it is assumed that the code asserts that Nash welfare function is to be maximised - even though it is just announced by the firm, *then* a mixed strategy equilibrium of the repeated trust game inducing stakeholders to forbear exploitation cannot be any more an equilibrium of the psychological game with *overall* utility functions on the part of stakeholders.

12. Intermediate social bodies, multi-stakeholder dialogue and CSR assurance and independent verifiability

I have shown the existence of a movement towards the standardization of corporate social responsibility management systems, and I have examined its theoretical rationale. The idea is also spreading that forms of third-party verification and independent certification can heighten the reputational benefit accruing to a firm from implementation of an ethical and social responsibility standard (of which SA8000 certification – even if mono-stakeholder and single-issue in character - is the best-known example). What is still needed for this process to be successful? In short, what is needed are institutions of the civil society able to give the necessary salience to corporate social responsibility; one with the competence to ascertain compliance with standards and to assure the independence, credibility and authoritativeness of declarations concerning compliance . The next stage in CSR can therefore be envisaged as the creation of institutes and organizational forms of civil society able to promote social dialogue which (i) creates broad consensus on standards, and (ii) promotes independent verification of compliance with those standards by means of appropriate monitoring and certification methods, these too endorsed at the social multi-stakeholder level.

If each firm has its own private model of CSR management, or if this model remains implicit, comparative evaluations become impossible. Necessary, therefore, is the public statement of standards, accepted by both firms and stakeholders, so that companies' conducts coherent to that publicly recognised benchmark may lead to the reputation rewarding or sanctioning of the deserting firms. However, a problem still persists even in the presence of a shared and accepted standard; a problem that revolves once again round the fragility of the reputation mechanism. Stakeholders may not possess the contingent information or the relevant reference criteria and knowledge they need to judge the firm's concrete

behaviour and communications. In the absence of a specialist agency able to collect, verify, , evaluate, benchmark and transmit information in summary form, thereby bridging the information gap between firms and stakeholders, the endogenous reputation mechanism may be too slow and imperfect. If this is so, we cannot prevent that reputation may in the short run accrue to or spread also over those firms who do not effectively comply with CSR criteria, without a precise enough discrimination between compliant and not compliant. But, if the reputation mechanism is highly imperfect (so that judgements are often erroneous), the entire CSR system will lose its credibility.

The point at issue, therefore, is what social organization or institution could undertake this task. There is an obvious risk of collusion between an agency in charge of performing this informative function in the behalf of the stakeholders and the firms that this same agency should subject to its scrutiny, by independent monitoring and verification. Which entails that a social institution must be created which optimal design should provides incentives against the collusions and conflicts of interest (in the economics jargon we say 'collusion-proof') so particularly insidious in contexts such as these, given the 'soft' nature of the information at hand. Optimal design of such an institution should therefore provide the following two properties:

1. *Multi-stakeholdership.* A multi-stakeholder body is directly functional to establish a broad consensus on the reference standards that provide the framework for corporate self-regulation. The problem, however, is to ensure that parochial interests do not take over the institution, as far as it is not captured by vested interests (those of who should be subjected to independent verification). Multi-stakeholdership is mostly important for, by appropriate design of the internal decision processes and symmetries in the distribution of decision rights and weights hold by every category of participants, it can help preventing collusion. The idea is simply resorting to the well known constitutional mechanism of 'checks and balances' among the various interests, so that in both standard setting and the monitoring of independent verification parochial interests are reciprocally eroded, freeing the institution from their control.

2. *Independence.* Multi-stakeholdership in itself will not suffice if these institutions turn into arenas of constant negotiation where those with the greatest bargaining power prevail or symmetric veto powers are able to drive them in dead-locks Also required is the adequate autonomy and separateness of ethical and techno-scientific professional infrastructure, able to ensure the independent working out of assessments which satisfy requirements of impartiality and competence.

An institution with these features could perform the following functions:

- setting CSR standards and the methodology for their implementation, holding their ownership, revising and updating them in a view of continuous improvement, promoting discussion for the progressive acceptance of CSR standards among the commonly accepted self-regulatory norms recognised by all the international bodies of standardisation and other international organisations;
- encouraging firms to comply with standards, and aiding them in performing pilot projects functional to fine tuning of the standards;
- disseminating information and promoting in-company training;

- agreeing with the accreditation bodies on the criteria and protocols that will be applied by auditors in carrying out their inspections concerning CSR compliance, and thereby establishing the procedures for the accreditation of third-party inspection and certification institutes collaborating in these activities in accordance with a pre-established protocol;
- especially in the initial phase, until CSR standards become commonsensical, maintaining close surveillance over the operations of the institutes of inspection that issue certificates of compliance with CSR standards;
- giving appropriate salience, by issuing surveys, reports and white papers, to the trend towards the adoption of ethical-social management systems;
- surveying and monitoring, by recollecting any source of information, the CSR profile of firms and furnishing the public with the information that it needs to form their judgements with objectivity, in particular to ethical finance operators and responsible consumers.

In summary, multi-stakeholder dialogue on CSR can be facilitated by the creation of civil society institutions endowed with competence, moral authority and independence, and at the same time considered representative of the interests at stake, so that they can reliably be addressed to the end of promoting corporate social responsibility and ascertain conformity to it against the reference point of shared criteria and standards. As multi-stakeholder bodies for the promotion, monitoring and independent verification of CSR, these institutes should take the form of non-profit organizations, with a broad base including business associations representing each type of enterprise (for-profit, cooperatives, non-profit) and the representative of their principal stakeholders: the trade unions, consumers, environmentalist associations, professional associations, non-profit associations advocating human right and social welfare, and local authorities – all of them supported by a network of rigorously independent research centres (the Universities have here important role to play). This would also lead to the strengthening of those intermediate social bodies that underpin a modern economic democracy and a democratic society in general.

REFERENCES

- AA.VV. (1993) "The Corporate Stakeholder Conference", *University of Toronto Law Journal*, XLII, n.3, pp. 297-793
- AOKI M. (1984), *The Cooperative Game Theory of the Firm*, Cambridge (Cambridge UP).
- BERNHEIM B. (1994): "A Theory of Conformity", *Journal of Political Economy*, Vol. 102, N. 5, pp.841-877.
- BLAIR. M AND L. STOUT (1999) "A Team Production Theory of Corporate Law", *Virginia Law Review*, Vol. 85, No. 2
- CELE (2001), *Progetto Q-RES, la qualità della responsabilità etico-sociale d'impresa, Linee guida per il management*, LIUC paper n.95.
- CHARNESS, G. E RABIN, M. (2000): "*Social Preferences: Some Simple Tests and a New Model*", mimeo, University of California at Berkeley
- COLEMAN J. (1992), *Risks and Wrongs*, Cambridge (Cambridge University Press).
- DONALDSON T. (1982) *Corporations and Morality*, Englewood Cliffs N J (Prentice Hall)
- DUNFEE T. AND T.DONALDSON (1995) "Contractarian Business Ethics", *Business Ethics Quarterly*, 5, pp.167-172
- FALK, A. E FISCHBACHER, U. (1999): "*A Theory of Reciprocity*", Institute for Empirical Research in Economics, University of Zurich, WP N. 6
- FEHR, E. E SCHMIDT, K. (2001): "*Theories of Fairness and Reciprocity – Evidence and Economic Applications*", Institute for Empirical Research in Economics, University of Zurich, WP N. 75
- FLANNIGAN R. (1989), "The Fiduciary Obligation", *Oxford Journal of Legal Studies*, 9, pp.285-294.
- FREEMAN R.E. AND J. MCVEA (2002) *A stakeholder approach to Strategic management*, working paper n.01-02 , Darden Graduate School of Business Administration
- FREY B. (1997), *Not Just for the Money*, Edward Elgar.
- FUDENBERG D. and TIROLE J. (1991): *Game Theory*, Cambridge Mass. (Mit Press).
- GAUTHIER D. (1986): *Morals by Agreement*, Oxford (Clarendon Press)
- GREEN L. (1990) *The Authority of the State*, Oxford, (Clarendon Press).
- GRIMALDA G. E L.SACCONI (2003) *The constitution of the non profit enterprise, ideals, conformism and reciprocity*, LIUC paper n.155
- GROSSMAN S. and O.HART (1986): "The Costs and Benefit of Ownership: A Theory of Vertical and Lateral Integration", *Journal of Political Economy*, 94, pp. 691-719.
- HARE R. M. (1981): *Moral Thinking*, Oxford (Clarendon Press).
- HARSANYI J.C. (1977): *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*, Cambridge (Cambridge University Press).
- HART O. (1995) *Firms, Contracts and Financial Structure*, Oxford (Clarendon press).
- HART O. and J.MOORE (1990): "Property Rights and the Nature of the Firm", *Journal of Political Economy*, 98, pp.1119-1158.
- HOBBS T. (1651), *Leviatano*, (trad. it.), Firenze (La nuova Italia), , 1975.
- LEWIS D. (1969), *Convention, A Philosophical Study*, Cambridge, Mass. Harvard University Press.
- JENSEN M.C. (2001) "Value Maximization, Stakeholder Theory, and the Corporate Objective Function" *Journal of Applied Corporate Finance*, Vol. 14, No 3, Fall.

- KREPS D. (1990): "Corporate Culture and Economic Theory" J. Alt and K. Shepsle (eds.), *Perspectives on Positive Political Economy*, Cambridge (Cambridge University Press).
- MCMAHON C. (1989) "Managerial Authority", *Ethics*, October, 1989.
- MEESE A.L. (2002), *The Team Production Theory of Corporate Law: A critical Assessment*, mimeo
- NASH J. (1950): "The Bargaining Problem", *Econometrica*, 18, pp.155-162.
- POSNER E. A. (2000) *Law and Social Norms*, Cambridge Mass (Harvard UP)
- PREITE D. (1992), *l'Abuso della regola di maggioranza nelle deliberazioni assembleari delle società per azioni*, Milano (Giuffrè).
- RABIN, M. (1993): "Incorporating Fairness into Game Theory", *American Economic Review*, Vol. 83, N. 5, pp. 1281-1302.
- RAJAN R. AND L.ZINGALES (2000) "The Governance of the New Enterprise", in Xavier Vives (ed.) *Corporate Governance, Theoretical and Empirical Perspective*, Cambridge (Cambridge UP)
- RAWLS J. (1971), *A Theory of Justice*, Oxford U.P.
- RAWLS J. (1993), *Political Liberalism*, New York (Columbia UP).
- RAZ J. (1985), "Authority and Justification", *Philosophy and Public Affairs*, 1985, pp.3-29
- ROSE-ACKERMAN, S. (1996): "Altruism, Non profits, and Economic Theory", *Journal of Economic Literature*, Vol. 34, N. 2, pp. 701-728
- ROSSI G. (2003) *Il conflitto epidemico*, Milano (Adelphi).
- SACCO P. E S.ZAMAGNI (2002)(a cura di), *Complessità relazionale e comportamento economico*, Bologna (Il Mulino).
- SACCONI L. (1991), *Etica degli affari, individui, imprese e mercati nella prospettiva dell'etica razionale*, Milano (Il Saggiatore)
- SACCONI L. (1997): *Economia etica e organizzazione*, Roma-Bari (Laterza).
- SACCONI L. (2000): *The Social Contract of the Firm. Economics, Ethics and Organisation*, Series: Studies in Economic Ethics and Philosophy, Springer Verlag, Berlin.
- SACCONI L. (2002); *The Efficiency of the Non Profit Enterprise: Constitutional Ideology, Conformist Preferences and Reputation* LIUC paper n.110, Varese, to appear in B.Hodgson (ed.) *The Invisible Hand and the Common Good*, Proceedings of the SEEP Conference, Series: Studies in Economic Ethics and Philosophy, Springer Verlag, Berlin, in print
- SACCONI L. (2003), "The Q-RES Project: The Quality of Social and Ethical Responsibility of Corporations" (with S. DeColle e E.Baldin), in Wieland, Josef (ed.), *Standards and Audits for Ethics Management Systems, The European Perspective*, Series: Studies in Economic Ethics and Philosophy, Springer Verlag, Berlin, pp.60-117
- SACCONI L. (2003), *Incomplete Contracts and Corporate Ethics: A Game Theoretical Model under Fuzzy Information*, in F.Cafaggi, A. Nicita and U. Pagano (eds.), *Legal Orderings and economic institutions*, London (Routledge), in print.
- SACCONI L., S. deColle, E.Baldin, J.Wieland, R.Oakley, S.Zadek (2003), *Developing a CSR framework to integrate Q-RES and other social and ethical standards*, final report, UE Research project N. VS 202/0406
- SEN A. (1993): *Moral Codes and Economic Success*, London (LSE, ST-ICERD discussion papers n.49).
- SUGDEN, R. (1998 a): "Normative Expectations: the Simultaneous Evolution of Institutions and Norms", in Avner Ben-Ner and Louis Putterman (eds): *Economics, Values, and Organization*, Cambridge (Cambridge University Press), pp. 73-100
- SUGDEN, R., (1998 b), *The motivating power of expectations*, mimeo
- TIROLE J. (2001), "Corporate Governance", *Econometrica*, vol 69, n.1, pp.1-35

WATT E.D. (1982) *Authority*, London (Croom Helm)

WEISBROD B. A., (1988), *The Non-profit Economy*, Cambridge Mass (Harvard University Press)

WIELAND J. (2003) (ed.) *Standards and Audits for Ethics Management Systems, The European Perspective*, Berlin (Springer),

WILLAMSON O.(1986), *The Economic Institutions of Capitalism*, New York (The Free Press), trad. it Franco Angeli, 1988.