EconomEtica
Centro interuniversitario per l'etica economica
e la responsabilità sociale di impresa
promosso dalla Fondazione Italiana Accenture

N.27 March 2011

Lorenzo Sacconi

From individual responsibility to 'shared' social responsibilities: concepts for a new paradigm

Working papers

From individual responsibility to 'shared' social responsibilities: concepts for a new paradigm[1]

by Lorenzo Sacconi

1. Why a new allocation of social responsibility is needed

Any discussion of the idea of 'shared social responsibility' on a European level should be introduced by an explanation as to why a new allocation of social responsibilities between public and private actors throughout Europe is needed: how it challenges received wisdom and how difficult it is – also considering the possible pitfalls into which the search for a new model could fall.

Putting things in very schematic terms, the old model – from which shared responsibility is a clear departure – was the liberal paradigm (widely accepted within economic theory for example) based on a simple allocation of social responsibility between the two main social institutions, the market and the state, where:

- the market, and agents operating within the market (firms, consumers etc.), are assigned responsibility for the efficient allocation of resources;
- the state is allocated responsibility for providing public goods and for reshaping the initial endowments with which individual agents enter the market process in order that a final efficient allocation of resources may be arrived at which also reflects the initial political choice on basic endowments.

The typical feature of the liberal paradigm is that individuals or organized agents operating within the confines of the market are required to fulfil a very limited set of moral and social obligations. The optimal allocation of resources does not require them to account for the outcomes of their collective actions or their compliance with any overarching ethical principle, other than the need to respect commercial law and specific contracts. Under the typical conditions of idealized perfect competition, the rational and farsighted pursuit of self-interest is the sole behavioural requirement in order for the market to function properly. There

[1] To be published in *Towards a Europe of shared social responsibilities. Challenges and strategies*, Trends in Social Cohesion Series, N. 26, 2011, Council of Europe Publishing, Strasbourg

are two versions of this minimalist view of social responsibility assigned to private market actors.

(i) The first maintains that broad moral responsibility is superfluous or even impossible for agents operating within the marketplace because they cannot intentionally affect the market outcome. Nevertheless this moral neutrality of market operations does not prejudice its outcome, which is a social optimum that fully respects negative freedoms.

(ii) The second envisages a standard of greed and self-interested behaviour exactly as the unique role-specific moral responsibility that individuals, professionals or organized private (corporate) actors are required to respect in order to guarantee the proper functioning of the market (according to a typical assumption of the First Theorem of Welfare Economics). Not being personally greedy (for individual entrepreneurs or consumers) or not seconding the personal greed of owners and shareholders (for managers), would amount to a violation of the 'ethics of the market', thereby jeopardizing its socially beneficial functioning.

The old paradigm has passed through different phases. After the Second World War, up until the extension of civil and social rights during the 1960s and the 1970s, the paradigm also ideologically survived a significant expansion of the Welfare State, since the requirement of social responsibility was still mainly limited to the government and public servants (even though the remit and number of the latter increased). It must be acknowledged that in the late 1960s the idea of corporate social responsibility arose again and inspired the "business and society" movement - as it had also previously been an important participant in the academic debate of the 1930s on corporate governance and the fiduciary duties of the modern corporation (see Bearle and Means 1932, Freeman and McVea, 2002, Kaufman 2002, Kaufman, A., L. Zacharias, M. Karson, 1995). Nevertheless as a matter of fact the main focus (at least in Europe) was on expanding the social responsibility of the public sector rather than on sharing social responsibility between different areas of society, including the private agents operating through the market and the third (non profit) sector.

Over the last thirty years of the neo-liberal or libertarian age, the old paradigm reached its apogee, since it was perfectly consistent with the rolling back of the welfare state and a re-allocation of many decision making rights and of broad discretion to private actors interacting through the market, in the belief that the social costs of public decision making processes were higher than the transaction costs associated with the exercise of decision making rights by private actors constrained by the logic of market competition (often assumed to be "efficient markets").

It should be noted that, according to the prevailing view, this was again intended to comply with the *the principle of separability of responsibility*: responsibility for efficiency (except for public goods in the strict sense) conferred on market actors as against responsibility for social justice and fairness assigned to the "inefficient" political process. Hence, the neo-liberal age stressed even more so the belief that the minimal social responsibilities should be borne by private agents (see Friedman 1970 for the classical formulation of this thesis).

This is rather paradoxical because, in general, the removal of decision making rights from one agent and their conferral on another also entails the transfer of the relative responsibility from the former to the latter. Not so for libertarians however. Due to the doctrine of the 'invisible hand' (efficient markets as impersonal mechanisms spontaneously reaching efficient equilibria) the idea spread that not responsibility but intelligent and farsighted greed was the key to the proper functioning of the market and the achievement of socially desirable outcomes. Thus the only requirement in terms of social responsibility was the undemanding obligation that self-interest should be pursued, albeit in an intelligent and far-sighted manner.

However, perhaps as an unintended consequence of the neo-liberal age, the increasing allocation of decision making rights, power and discretion to private organizations operating according to the rules of the market raised calls that also the allocation of social responsibility should be reshaped in accordance with the new configuration of decision making rights and powers. The same process that has transferred so many decision making powers from the public sector to the private sector on mere efficiency grounds reasonably raises the question concerning the social justice and social welfare consequences of private sector decisions. The requirement follows that – where relevant – fairness and social welfare standards should be met not just by public institutions but also by the decisions of powerful private agents, even though they operate through the market.

In fact there are social cohesion pitfalls brought about by the effect of private actors' decisions that the state does not cause and which it cannot face up to on its own. In such decisions, private actors exercise power and discretion mediated by market incentives. Some examples may concern:

- *Immigration of workers* from poor countries on “low cost” employment conditions. While providing benefits to domestic companies, private sector industries (both manufacturing and services) and families, immigrant workers’ demand for public welfare services is seen as a crowding of these serivice and a way to ‘free ride’ on the costs of their provision – especially when they operate on the irregular labour market or their wages are so low that their capacity to pay tax is limited. This gives rise adverse social effects such as the repugnant racial hostility that jeopardises social cohesion in European societies. Social integration and support for the growing costs of welfare systems associated with the immigration of low-income earners would be much better dealt with *ex ante* by preventive responsible behaviour on the part of private actors – which may contribute to the immigrants’ social integration – rather than *ex post* by the state.
- *Global effects of financial turmoil* related to perverse managerial incentives resulting from the principle of shareholder value maximization. These generate the financial externalities on economic systems that have been widely apparent in the most recent financial and economic crisis (reduced access to credit for companies, reduced demand by consumers etc.). States have responded to these effects with strong stimulus policies, but they may not be able to repeat this success in the near future because of the level of public debt incurred. These effects must then be prevented by responsible behaviour from the private-sector.
- *Distributive inequalities* generated through allocation mechanisms. Companies are ‘team production’ organizations within which many stakeholders cooperate by making investments and contributing to wealth production. However, authority is vested in shareholders and boards of directors that are accountable to shareholders and are incentivized by returns conditional on the share value. Existing corporate governance systems do not give significant protection to corporate stakeholders. Nevertheless stakeholders also make firm-specific investments while being imperfectly protected by incomplete contracts. At the same time, they receive no guarantees from the exercise of residual rights of control, as owners, shareholders or directors do (see Sacconi 2000, 2006a,b).

These considerations should preliminarily show what shared responsibilities do not entail. It should be made clear that in proposing that responsibilities be shared, we are not embracing the view of shifting social risks, and hence responsibility for such risks, to individuals, who are the weakest players and hence the least able to bear such risks. To cite only a few examples:

- Company employees are required to bear the entrepreneurial risk in terms of flexibility of their employment conditions and liability to dismissal; at the same time, there is no corresponding proposal that employees may receive protection similar to that deriving from the fiduciary duties owed by managers to owners.

- Citizens are required to bear the risk of adverse social events and to insure themselves against such events by means of private insurance mechanisms, even though they are not best able to bear such risks; however, private insurance schemes are less efficient than public insurance (universalist welfare-state) schemes, and the recent financial crisis suggests that financial markets are unable to re-insure the systematic social risks associated with poor people's requirements for housing, health services etc.

At the same time, the recognition that individual agents have only bounded rationality and are subject to systematic cognitive biases has made the case for some degree of 'liberal paternalism' (see Sunstein 2002, Thaler and Sunstein 2008). This is intended to give people cognitive cues in order that they may take care of themselves, and also to prevent imprudent risk taking. However, it also suggests that they are insufficiently rational and well-informed to bear significant social risks. In general, this change of viewpoint on the cognitive capability of individual agents stresses the role of 'social responsibility duties' that should be borne by those institutional actors in the private or public sectors that act as trustees for the people and must comply with the fiduciary duties placed on them.

Moreover the 'shifting' of social responsibilities (as opposed to their 'sharing') from the government to private individuals, families or the market for social service provision bring with it the risk of not only reducing the extent of coverage of positive social rights but also of changing the nature of those rights, causing detriment both to the universality of social citizenship as well as the impartiality and equality of treatment of people by service providers.

2. A minimal definition of responsibility and the challenge of its extension to shared responsibility

Before embarking on any deeper discussion of 'shared responsibility' it is necessary to provide a clear definition by elaborating on some minimal but unproblematic definition of responsibility. Hence, what at minimum does 'responsibility' mean? According to a standard view in philosophy of law (see Hart 1968), being responsible means *having the capacity to be* *subjected to blame or praise* due to an action or the outcome of an action in terms of some norm (legal or moral) from which a duty is derived.

Duties can be associated with any right according to the logical construct of 'claim-right' (in fact, contrary to the notion of a right as liberty/permission, a right may consist in a claim to some benefit, action, inaction or state which entails a correlative duty incumbent upon another agent to provide that benefit, action, inaction or state – see Hohfeld 1923). This is therefore a quite basic and overarching definition. Responsibilities can be established for compliance with duties deriving from negative but also *positive* claims – i.e. social rights where the claim concerns not just refraining from action (as in *negative* rights such as ownership) but the provision of a service (this point will be returned to in the next section).

2.1. The allocation and disclaimer of responsibility

The above definition of responsibility is useful because it leads us directly to the question as to what it means to be in the condition of 'having the capability for being subjected to blame or praise'. Philosophers (and common sense) answer this question with the postulate "ought implies can" (see Hare 1963, Danley 1988). Hence one cannot be attributed responsibility for an act if one cannot make a choice regarding that action. Admittedly, this is rather obvious, but nonetheless it immediately raises a basic challenge to the definition of "shared social responsibilities": we cannot *share* any responsibility with another (natural or legal) person if that person cannot make any choice concerning the matter (I have already used this idea when criticizing the alternative program of *shifting* social responsibility to private citizens or families who are the least able to bear these risks).

To be somewhat more precise about the conditions for the attribution of responsibility, two more qualifications must be made:

i) in order for an agent to be held responsible for a state of affairs S, it must exert *causal force* over S by means of an action that (at least in part) causally produces S;
ii) moreover that action must be at least to some extent *intentional*, so that the agent concurs in the production of state of affairs S by means of an *intentional* action.

Intentionality can be characterized in various ways. According to economic methodology, an action is intentional for a given decision making problem if it can be construed as utility maximizing in that context. Thus the agent acts intentionally if he has a complete and coherent system of preferences over a set of courses of action (and their consequences), whereby the behaviour observed may be derived as a best choice given the agent's ordering of preferences (i.e. it is ranked as the most preferred action vis-à-vis the decision making problem concerned).

Alternatively, an agent can be said to act intentionally in a given context if his or her observed behaviour can be interpreted in such a way as to satisfy a mental representation (which the agent happens to hold) of a goal and of an act that is an effective means of achieving that goal. If the behaviour described cannot be said to satisfy, through a consistent mind/world correspondence, a mental model that the agent holds regarding its goal and the action which is conductive to that goal, then it cannot be said to be intentional (see Searle 2001).

However, a proper understanding of the conditions for the attribution of responsibility requires that they should not be too demanding, in order not to extend or restrict unreasonably the range of subjects to whom responsibility may be attributed. For example, linear causality, or uniquely determining casual force, should not be required. Otherwise, practically no social event could be attributed to the responsibility of any agent, because it is obvious that the causal determination of social events always involves multiple variables. For example, it is usually said in organizations that "multiple hands" are the cause of a given state of affairs (see Thompson 1985), or in games that the outcome is brought about by the interdependent decisions of the participants.

Moreover *direct* intention, such as having the mental representation of an outcome S as the proper goal of the individual's action, also seems to be excessive. In order to be responsible for S, the agent need not have represented S as his true goal; he/she can simply have represented that state as a by-product of the action that he/she intentionally performs in pursuit of another goal, so that he/she is aware of S simply as a possible undesirable side effect of his/her action. The same also holds for the 'preference explanation' of intentionality. In order to attribute responsibility to an intentional actor understood in a preferential sense, the outcome need not be represented as if it were the end state that he/she most prefers. It can be considered merely as a cost that he/she is ready to pay in order to achieve the desired outcome.

As a matter of fact, the agent will not have a complete representation in his/her mind of all the characteristics of any particular state of affairs, so that in preferring an action that includes the state S as a possible consequence he/she would also reveal a preference ordering of all these states. He/she will explicitly apply the preference ordering only to some salient features of outcomes – those characteristics that are grasped as explicit decision making variables of interest to the decision maker within the *frame* of the decision making problem that comes to the agent's mind ('framing' will be discussed again below in this section). Nevertheless, for the purposes of attributing responsibility it is sufficient that the agent be aware of it as a related state of affairs even though it is not the specific object of his or her desires. What counts for awareness is the *effective* cognitive mental representation of S as a state of affairs associated with the course of action.

This will reduce the number of states of affairs that may fall within an agent's area of responsibility. Since an agent is boundedly rational, he/she is effectively unable to foresee every possible state of the world. On the other hand, requiring a *complete preference ordering* including *all the possible states of the world* would exempt practically all real-world agents from responsibility, given that none of them would in practice be able to satisfy such a high standard of decision making consistency (we impose a preference ordering on states only in terms of the relevant features, which makes it possible to factor out many non-salient details).

To sum up, an unintended consequence – i.e. one not envisaged *ex ante* as the agent's goal or preferred outcome – deriving as a composition effect (that nobody specifically wanted) from the interaction between many intentional players can be attributed to the responsibility of those agents, provided that each of them acted according to some intentional goal and was aware of that outcome as a possible unintentional by-product of the composition process.

Conditions for the attribution of responsibility are typically invoked by private and or corporate agents as grounds for disclaiming responsibility, as well as by individual citizens. For example, companies disclaim all responsibility for inhumane working conditions or child labour in the 'low cost' plants where they have delocalized the production of components necessary for their own manufacturing process, or for the employment of immigrant workers on discriminatory conditions (or without any contractual protection at all).[2] In doing so, they resort to arguments like the following:

[2] The attribution of moral responsibility to corporate actors has been extensively discussed; see for example Danley (1988), Danley (1990), DeGeorge (1982), French (1984), Goodpaster (1984) Ladd (1984), Velasquez (1983) and moreover Sacconi (2005).

- No information was available about what was going on in the plant (which was not supervised), and hence no intentionality can be imputed to the action that contributed to bringing about a bad state of affairs.
- No intentionality may be recognized in the action even under conditions of perfect information about the employees' labour conditions, because the company was not free to choose to improve the working conditions in the supplier's plant, since it must remain competitive on the market, which in turn means that the best strategy is to reduce the costs of components acquired through the supply chain.
- No causality can be attributed to the company's action (even if it is aware of the consequences) because the abominable working conditions in the supplier's plant would have been exactly the same had the company decided to cancel the supply contract, or not to conclude it at all. Given the ongoing competitive conditions of the market, another company would have been induced to buy similar components at exactly at the same price conditions from the same supplier.

Similarly, individual citizens disclaim all responsibility for the depletion of natural resources or their insufficient contribution to the provision of local public goods (such as a reduction in pollution, energy saving and greenhouse emission reduction programs, insufficient participation in responsible consumption initiatives etc.) by denying that some of the conditions for attributing responsibility are satisfied in the case in point. For instance, pollution is an unintended collective composition effect of the decisions of many individuals, whereas each of them takes his/her individual decision whilst bearing in mind a personal goal or intended outcome that is completely different from the composition effect. Thus there is no intentionality. Moreover, no causality on the composition outcome can be imputed because each individual action makes only a minute contribution to the global outcome. The same would be the case even without the contribution of the single individual.

Of course, many of these claims can be rebutted simply by a more accurate consideration of the conditions for attributing responsibility. For example, it could be stressed that the company voluntarily decided not to supervise the supplier's plant, and that in general it cannot pretend that it was unaware that labour conditions cannot be anything but poor in similar factories at that price level. Moreover, the decision to remain competitive on the market by carrying out these kinds of cost-saving policies is an intentional and deliberate choice concerning one's ultimate professional goals, just like any other. Thus the entrepreneur cannot disavow responsibility for the consequences of that deliberate decision as if the action had been carried out under duress (“it *must necessarily* remain efficient on the market in order that the cost-saving policy automatically follows”).

Finally, the market structure in the relationship between the supplier and the buyer is probably not “perfectly” competitive. The buyer decided to relocate to that plant the production of components that it previously made itself in its home country. There was probably something like an exclusive contract that ‘locked’ the supplier into a privileged business relationship with the buyer because the latter was the former’s only client for that specialized component, which amounted to the largest part of the plant’s production. There was something like a bilateral duopoly situation between the two parties, under which the buyer had the strongest bargaining position. Since the company had bargaining power over the supplier, if the buyer wanted, it could have demanded that working conditions be improved in the plant, perhaps at the same time permitting the supplier to take advantage of improved contractual conditions. But by contrast, the company exploited its bargaining power (for example, the threat not to honour an incomplete contract) with the goal of reaping as much of the surplus resulting from the exchange as possible. Given the part of the surplus left to the supplier, the latter could not do a great deal to improve labour conditions at the plant.

For private citizens, responsibility disclaimers can be rejected on the grounds that the individual cannot claim not to be aware that the cumulative outcome of many individual actions, even if *prima facie* aimed at different goals, ultimately results in the depletion of a natural resource or an environmental common good. Moreover, at least at the level of local public goods or commons, the contribution of one single agent may not be miniscule and may not make a negligible contribution to the causation of a suboptimal outcome. In fact, the situation resembles more a multi-player prisoner’s dilemma or a free-rider problem with a finite number of players than the pure public good paradox involving an infinite number of agents.

All this only proves that individual responsibility can be expanded to account for social outcomes that are partly within the discretionary range of certain private agents. It should in fact be noted that every rebuttal of a responsibility disclaimer refers to the existence of conditions in such a way such that the disclaiming individual nevertheless retains some discretion over a social outcome, and hence he/she can be attributed responsibility for that state of affairs. However nothing has been said thus far about the possibility that individuals may *share* social responsibilities for the implementation of certain *desirable social goals* (welfare objectives or norms) decided at the level of a collective body or society at large.

2.2. From individual to shared responsibility

Defining shared social responsibilities requires the identification of a collective body (a group, the society) that *shares social welfare goals or principles of social justice among its members*, exercises intentionality in pursuing them, and also has causal force over the determination of states of affairs that are consistent with such intentions. It is also necessary to have an argument which proves that, starting from a goal or a set of principles shared by a collective body, responsibility for that goal or those principles can be allocated to its members, such that they share responsibility for fulfilling the goal or principles of the collective entity.

Before discussing how this step can be achieved, let me try to define a view of society that can legitimate the idea that there is something like a *shared social responsibility* among its members, or that suggests that individuals share responsibility for some *social* situation in terms of a goal or a set of principles established on societal level.

Shared responsibility seems to be associated with a duty to behave according to certain principles or goals that are of value *along with other agents* (citizens, organizations, institutions) and/or to pursue the achievement of some common goals or states of affairs. It suggests treating (European) society as a *cooperative venture for the mutual advantage* of its members (Rawls 1971). In order to support the continuous participation and contribution of all its individual members, such an entity must be grounded on agreed principles of social justice concerning the production and distribution of primary goods and basic capabilities (Sen 2009) necessary for the well-being of all its member citizens.

When these principles are *agreed to*, and hence *shared* by the citizens, this entails that they are responsible to behave in accordance with these principles. This may be interpreted as a *shared* responsibility borne by them (even though it is differentially allotted between them), their private organizations and associations, companies and firms, and public institutions on different levels. From this perspective, shared responsibility means that public authorities, citizens, non-profit and for-profit organizations and their stakeholders are included within areas of responsibility under the terms of some principle or standard of social justice and welfare, and that collective action or the generation of a social outcome is required. Different outcomes associated with the notion of social cohesion may be the object of agreements and may flow from the agreement process properly defined.

To be precise therefore, shared social responsibilities require (as a prerequisite for their allocation) that corresponding duties be identified and attributed to members of society (or some relevant subset of it) because they are derived from values or goals that are shared by the same members understood as a collective body and who have made a choice (an agreement) about them. At the same time, the members of society must also be able to discharge these responsibilities effectively according to the conditions establishing their responsibility: viz. their capacity to act, to have causal force, and to exercise some degree of intentionality. In other words, the merely formal assignment of duties from an independent 'collective' level may not be effective in engendering true responsibility if the members of the collective do not perceive themselves as having the capacity as agents to agree on those duties and hence to discharge the duties from which their responsibilities stem.

Since a number of paradoxes of collective action (such as the Free Rider problem or the Prisoners' Dilemma) that sever the link between collective moral goals or principles and the individual's actual capacity to accomplish the corresponding duties are well known, the conditions for the effective ascription of shared social responsibility remain problematic. To summarize the paradoxes for the benefit of the reader, let us consider first the *Free Rider problem*. Many players participate in collective action for the provision of a public good. When some amount of the public good is provided its consumption by some individual does not make physically impossible that the same amount also benefits other individuals (what entails non rivalry in consumption), whilst participation in production or paying a price for the good cannot be a basis for claiming a special exclusive stance for the appropriation of a share of the public good, since nobody can be excluded from benefiting from it (i.e. non-excludability from consumption, for example by imposing a price). However participation, even where minimal, amounts to a cost for the individual participant. Thus, each individual member of the relevant group, by acting on the basis of his/her mere rational self-interest, will decide to abstain from effective participation in the collective action, since by this individual strategy (no participation) s/he nevertheless still has access to the fruits of the collective action, and at the same time s/he dispenses her/himself from bearing the cost. Since none may predict the case in which s/he is determinant in reaching a threshold of participants in the collective action which is necessary for the good provision, this individual decision is not conditional on the prediction of how many people will participate in the collective action. Consequently, participation should come to an end and the public good provision should disappear.

The Prisoners' Dilemma is an even more skeletal instance of strategic interaction. At least two players may gain a mutual benefit from reciprocal cooperation, but in the event that one fails to cooperate, while the other plays his/her part, the former will fare much better. Moreover, where the other player does not cooperate, it would not make sense for either player to cooperate, since it is more beneficial to adopt the defensive strategy of non-cooperation as well, whereas cooperating would subject the player to unilateral exploitation by the counterpart. Whatever the predicted behaviour of the other party, non-cooperation is the best strategy for each player, assuming that each reasons from the perspective of individually rational self-interest. The paradox lies in the fact that , while non-cooperation is the dominant strategy for both players, the resulting outcome for both of them is worse than the outcome that would obtain had they both decided to cooperate against their rational self-interest.

These 'social dilemmas' (Ostrom) typically lead individual agents to disavow their individual responsibility for the collective harm which is the outcome of their interaction. Even though their actions are intentional and they exercise (together with other agents' actions) some casual power over the social outcome, agents fail to see the rationale for accepting responsibility for the failure to achieve the socially optimal outcome. In the Free Rider case each agent argues that her/his individual contribution cannot by itself make the difference in the provision of the public good. The player could only accomplish the task through success in coordinating a large number of contributors. But according to rational self-interest, no other player will contribute. Therefore it is not the case that he/she may be attributed responsibility for the failure to bring about the social optional outcome.

Even though the causal role of the individual player in the Prisoner's Dilemma game is much more relevant in determining the outcome, he/she follows similar line of reasoning in declaiming his/her responsibility for the failure to achieve the Pareto optimal outcome of reciprocal cooperation. He cannot avoid recognizing that the suboptimal outcome is the joint responsibility of both the payers. Nevertheless, since non-cooperation is the dominant strategy for both of them, there will be no chance for him to predict that other players will coordinate in order to achieve the cooperative outcome. By acting unilaterally each can only generate a situation of unilateral exploitation to his/her own disadvantage – but of course putting him/herself in the situation of falling prey to unilateral exploitation may go far beyond his/her reasonable obligations.

These arguments do not provide a valid refutation of the players' individual responsibility in determining the social harm. But what matters here is whether there is also a solid basis for making an argument in favour of accepting individual responsibility for bringing about the social good. 'Ought implies can', but as long as the behaviour of other players is determined according to the requirements of the standard theory of rational self-interest, the player cannot modify the suboptimal outcome through his/her individual decision, and thus he/she cannot feel responsible for not having satisfied an 'ought' which goes beyond the reach of his/her individual decision.

Quite obviously, the demand for responsibility may be shifted to a different level, viz. the responsibility for favouring a change in the form of the game played or in the rules whereby the same game is played.[3] This would allow for some coordinating mechanism, credible cooperative agreement or the availability of additional sanctions against non cooperative behaviour such that players could agree on a joint plan of action permitting a different social outcome. Thereafter, individual responsibility could be predicated on the individual's decision to carry out an individual action that concurs with other agents in order to determine the best social outcome. In sum, individual responsibility can be rejected as far as each player perceives him/herself (i) as an individual agent that makes his/her best individual decision given a rational prediction of the other agents' non cooperative behaviour, and not (ii) as a component of a collective agency unit that settles (and is committed to) a common goal and undertakes the action consistent with pursuit of the goal itself. In particular, the PD's individual players fail to recognize their responsibility for cooperation until there is no identification with an agency unit , a collective goal or a shared principle requiring that both players play the game according to the joint strategy (*cooperate, cooperate*) .

[3] Different approaches have been proposed over decades of research into game theory, which may be formulated in terms of suggestions to the players : 'play the game *repeatedly* so to allow *reputation* effects', 'change the *given* rules of the game so to pass from a *non-cooperative* to a *cooperative* game', 'insert the game in a *meta-game* where *rules* for playing a particular game can be chosen within a larger game', 'allow deciding on *rules* for playing the game', 'introduce a pre-play (apparently) *cheap talk* stage in which player may make an agreement under the *veil of ignorance* on the game's rules, and then see what happens to their *preferences* in the ensuing game', and finally (as discussed in the main text) 'trigger the *frame "We"* to bring it to the players' minds, and then let the players play the game *as a group'*.

A highly effective way of explaining the shift from collective values to individual duties and responsibility within such social dilemmas is suggested by the theory of "We Thinking" and "Team Reasoning" (Bacharach 1999, 2006, Sugden 2000, 2003, Tuomela 1995). Team reasoning, according to Bacharach in particular, requires the following premises: for each individual member of a group, it is true that s/he knows (and s/he knows that every other individual member also knows) that (i) s/he is a member of the group, (ii) the group has a goal, (iii) every member of the group identifies with the group and hence takes the group's goal as her/his own goal, (iv) there is a joint action that satisfies the group's goal better than any other, and members of the group calculate it; (v) from that calculation each member knows the individual action which is the projection of the common action onto the individual strategy set of each player - i.e. the part of the collective action that is under the control of each individual player. It *then* follows that each individual member knows that s/he *should* carry out the individual action as the component under her/his control of the collective action satisfying the group's goal. Moreover, according to the same authors, this inference is a mere piece of *valid reasoning* according to the logic of 'successful instrumental rationality' – that is, it does not presuppose any further condition other than the above *plus logic* (see Bacharach 2006) .

The key assumption behind team reasoning is *group identification,* which entails that for each individual there is a cognitive change in the understanding of the *unit of agency*. It is no longer the individual but the group as such that is considered as *the agent*. Once the agency unit has changed, the individual's goal (or payoff function) changes. S/he no longer identifies her/his objective-function with her/his own *personal* utility function. Instead, her/his objective-function is now identical with the objective-function representing the common interest of the group: that is, henceforth the individual takes the group's goal as the premise for her/his practical reasoning.

The conclusion as to what each individual should do follows as a simple valid inference according to the logic of practical reasoning. Since the group's goal is her/his own goal, and since collective action is the best way of achieving that goal, and moreover since s/he is able to infer from that action (a vector of individual actions) the projection onto the action set of each individual (the vector component belonging to each individual), then the individual must know that her/his best way of satisfying *her/his* goal (i.e. the *group* goal) is to implement her/his component of the collective action. Therefore s/he *should* carry out this action.

From this it is only a small step towards responsibility: the individual who team reasons can recognize that s/he has been attributed responsibility for the performance of her/his collective action component. If s/he team reasons, s/he also understands that s/he can be praised or blamed from the group's point of view for having played or not played her/his part in implementing the collective action conducive to the group's goal.

It should however be noted that team reasoning is not valid in general but only contingently. It is based on a matter pertaining to cognitive social psychology – *identification* – which is not necessary but contingent in itself. An ever possible alternative in fact is that the individual does not identify with the group and continues to see himself as the unit of agency.

What triggers group identification according to Bacharach is the cognitive psychological mechanism of 'framing'. If the individual enters the "We" frame, then he will see the group as the unit of agency, and the rest of the team reasoning will follow. But if he remains within, or enters into, the "I" frame, he will see himself as the unit of agency and the logics that he will follow will be the canons of typical strategic reasoning used in game theory (albeit subject to some rationality bound). In this case, the logical consequences are the collective-action paradoxes, or (so to speak) the typical calculation of an individually dominant strategy in the Prisoners' Dilemma that leads to the collective suboptimal outcome (i.e. agents individually do not discharge their responsibility toward the collective goal).

Neither framing is unique or voluntary: *it happens* that a frame comes to the individual's mind. But it may also occur that another frame emerges – in the same way as the adoption of some (internally consistent but mutually exclusive) alternative way of perceiving things is always possible given certain figures according to *Gestalt Psychologie*. If it *is the case that* the "We" frame comes to the individual's mind, then it circumscribes the reasoning that the individual can implement given what he can see and account for within the bounds of that frame. For example he does not keep a mental account of the rational opportunity of unilateral defection in the Prisoners' Dilemma because this is not rational within that frame, where he identifies with the group's goal and hence strives to achieve the 'cooperative' outcome (which is the best outcome from the team's point of view). When however the "I" frame comes to the individual's mind, then he reasons in terms of the individual's best response and hence is perfectly capable of accounting for the decision to defect as a rational individual action with respect to his present goal (to maximize his personal utility).

Triggering the “We” frame is therefore essential for the effectiveness of team reasoning. Even if framing is contingent and not voluntary, a given frame is triggered by some clue or salient element occurring in the situation and that attracts the individual’s attention through some experience that helps the relevant frame to come to the individual’s mind. And this occurs exactly at the same time that another frame for the same situation is still available. This should be borne in mind when seeking to understand the implications for shared responsibility.

In fact, on this view, responsibility follows the same route as team reasoning. If the “We” frame comes to my mind, I will identify with the group goal, which results in my duty to play my part in the collective action that is instrumental to that goal, assuming that it is calculable. Moreover, if this is the case for all the members of the group, they will all share the responsibility to play their part in achieving the group’s goal. Thus each individual shares the goal and hence *shares* responsibility for carrying out *her component* of the collective course of action conducive to the group’s goal. It follows that shared responsibility can be allocated to each group member who identifies with the group goal due to the fact that the “We” frame comes to her mind.

However, it must be remembered that responsibility for the group’s goal can be shared only if the cognitive phenomenon of framing does indeed occur. Hence – assuming that Bacharach’s explanation of team reasoning is true (what is not completely certain) – it is possible to understand how challenging it is to satisfy the conditions for attributing shared social responsibility: what is needed is to discover on a social level what conditions/clues (so to speak) favour the adoption of the “We” frame.

This must be the case not for the members of a single group alone (which would entail group responsibility but also parochialism); it is also necessary that the *identical* “We” frame be triggered for individuals belonging to *different* groups: for example, individuals belonging to *different stakeholder categories* and who are normally able to perceive themselves as distinct from other stakeholders because they have distinct sectional interests at stake (and different locations) within a particular decision making domain. In other words, what should be triggered are frames such as “We the society”, or some relevant part of it, “We the territory X”, “We the town Y”, “We Europe”, etc.

This relates both to the *ex ante* identification with the collective and its objectives (which is not at all obvious given the existence of distributive conflicts between different stakeholders), and the *ex post* generation of the motivational forces that support execution of the individual action by each single agent. It may be useful to think of some sort of 'law of conservation' for the *sense of commitment* needed in the move from identification with collective action (assuming it exists) to its projection onto the individual's action (coinciding with the collective action component under the exclusive 'responsibility' of one particular agent).

According to team reasoning, once the individual component of the best collective action has been calculated, and assuming group identification, the second step should be a mere matter of valid reasoning (i.e. a syllogism). But we may doubt that this is the case (Sugden 2003). In addition to identification on the societal level with some common goal and principle, it is also necessary for the causal conditions (cognitive and motivational) that give us the specific impulse to carry out our individual component of the collective action (to play our part in it) to be satisfied.

The idea suggested here is that activation of the "We" frame on the social level – resulting in shared responsibility for the common goal (with its projection onto the individual of responsibility to play one's own part) – is dependent on whether it is possible to conduct the *thought experiment* of the "social contract under the veil of ignorance in an original position" (see Rawls 1971). Even though some moral philosophers would see this as a pure exercise of rational autonomy and hence as perfectly voluntary, from a social science point of view it must be admitted that it conflicts with numerous cognitive biases and specific particular interests (it is simple to say in theory, but difficult to implement in practice, when all the agent-related beliefs and motivations must be taken into account). It is therefore necessary to identify and accurately sketch out the institutional, organizational, cultural and governance-related conditions that constitute the potential social clues favouring the establishment in people's minds of the relevant frame of "We the parties in the social contract".

The most challenging aspect of this task can be put as follows: identifying the conditions – those that affect beliefs and motivations – that allow for the preservation of the "We the society" frame (or a relevant part of it) when the agent enters a decision making or interactional domain in which it is obvious that (a) carrying out an action in accordance with the common goal or principle is an *individual choice,* and (b) an *individual* decision to *free ride* on the shared responsibility is an *open* possibility - as is typical in compliance contexts. These challenges however go beyond the scope of this chapter (although see the related essay on multi-stakeholder governance, Sacconi 2010, *infra*). Yet before identifying these conditions, we have still to consider a preliminary question concerning the possibility of such a social contract from which the norms of shared social responsibility are derived, given the diversity of the interested parties' claims and the stakeholders involved.

3. The multi-stakeholder agreement whence shared social responsibilities ensue

In order to bring about SSR, it is necessary to identify a collective unit of agency capable of establishing agreement on shared (among the members) principles and goals which can be transferred into the shared responsibility of each member not only during the deliberative stage but also, and principally, during the stage when each member is called upon to contribute to implementing those shared principles and goals. Here the "we thinking" mode of reasoning (and motivation) may play an important part (where there is one).

The task is not easy. At whatever possible level of analysis, the relevant collective agency unit will consist of some multi-level (local, regional, national, continental) and multi-stakeholder group. In fact, without this multi-level and multi-stakeholder composition of the relevant group of agents called upon to share social responsibility, the word "shared" itself will be meaningless, or almost so. Complete homogeneity of the group of agents involved in sharing responsibility for a collective goal and action can be excluded even in the simplest representations of the problem provided by simple models of 'social dilemmas' – such as the Free Rider or the Prisoners' Dilemma games. Even in such abstract contexts, where the existence of a best joint action is obvious, there is in fact a conflict of interest (resulting in divergent actions). Each player prefers his/her individual defection over cooperation (or non-contribution) once he/she expects that the other player will cooperate.

However, when such an abstract and simple representation of the problem is accepted – which minimizes the scope for differences lying behind distributive conflicts – it may be conceded that there is a unique Pareto optimal course of collective action which collectively dominates over any other course of action and outcome. Consequently, if reasoning in terms of what would be better for the players' group as a whole is allowed, all agents belonging to the group will unanimously prefer that course of action and the corresponding outcome whereby they all "fare better" as a group. In the PD case this is of course the joint action consisting in the strategy pair (*cooperation, cooperation*). In the event of a 'public good' game, it is the joint level of contribution that maximizes the joint benefit of the public good provision net of the costs of its provision. Hence, if the possibility to act in a cooperative and coordinated manner is assumed (i.e. the players consider their group to be an agency unit and they do not consider during the agreement stage that they could individually breach any agreement), it is quite obvious that all of the group's members will agree to adopt such a plan of action and its outcome as their common goal. The only problem persisting in this case would relate to "implementation", which consists in the need to overcome the incentive to depart from the optimal collective course of action when it is realized at the implementation stage that, by cooperating, each member incurs a cost that he/she as an individual player could save by acting opportunistically. But this is exactly where the "we thinking" can play its role. It consists in assuring consistency for all the group members between the agreed collective plan and the individual course of action in the implementation stage.[4]

Nevertheless, when a less abstract perspective is taken on the SSR content, conflicts between different stakeholders' interests are likely to emerge. This is moreover natural, given that claims of a different nature may clash even if they are held by the selfsame individual at different points in time throughout his/her life, so as to classify him/her as a different stakeholder. (For example at different stages the same individual may be a young person who has not yet entered the productive life-stage and needs skills and resources to do so successfully, then a participant in a productive venture occupying a specific professional role and claiming recognition of her/his efforts, and finally a vulnerable person in need of health care).

The social contract is suggested as an answer to the demand for a cognitive mechanism capable of providing the idea of a group unit of agency and the 'we' frame that can solve the 'ought implies can' problem underlying the allocation of shared responsibility. The idea of the social contract must then satisfactorily deal with conflicting interests and different claims. Here I do not consider possible analytical developments in the idea of the social contract, which is deferred until a later chapter in this book (Sacconi 2011b, *infra*). Here I am only concerned with the possibility that stakeholders with different stakes and claims may agree to settle their shared obligations according to the moral priority of those different stakes and claims. The focus is on the extent of moral legitimacy and priority that an impartial agreement

[4] Note that a compatible and perhaps more convincing explanation can be given by the idea of a "sense of justice" (Rawls 1971), i.e. the idea that conformity preferences supporting voluntary compliance with an agreed principle of fairness, will develop when there is public knowledge that all the participants have endorsed the fair agreement and moreover it is expected by all parties that all participants will reciprocate compliance with the agreement itself (Grimalda and Sacconi 2005, Sacconi and Faillo 2010, Sacconi 2011a).

may grant to different stakes and claims, so that their ordered satisfaction may be translated into the 'shared responsibility' of all parties to the agreement. "Moral" is understood here in the contractarian sense as indicating an impartial and fair term of agreement. In other words, my concern here is with the core idea of the social contract as a rational and impartial agreement on principles of justice whereby stakeholders' claims may be ordered according to certain priority principles.

The discussion cannot yet be confined to the 'ideal theory' of the social contract, so that the subject matter of the agreement can be assumed to be already well defined due to the very nature of the 'original position' where the agreement ideally takes place. (For example, according to Rawls, given the nature itself of the 'original positions', only 'primary goods' are considered). On the contrary, we must admit that from a 'non-ideal' perspective, stakeholders called upon to share social responsibilities may bring both fundamental claims to the negotiating table (that is, basic rights over primary goods and basic functioning) and claims that seem less fundamental, though still legitimate, such as the claim for the fair remuneration of efforts devoted to whatever productive contribution to general wealth and welfare. In fact, it would be unrealistic to assume that these different claims are not staked in any real life situation where shared responsibilities for the accomplishment of some common goals are established.

Moreover, from a 'non-ideal' perspective, the power of different stakeholders may be unequal and not necessarily aligned with the relative legitimacy and priority of their claims. What can thus be expected is that, according to their threatening position, those stakeholders who control the most valuable resources and are better organized and concentrated will control the group's decision, and thereby affect the allocation of responsibility in a manner favourable to themselves. That is to say, they will shift some responsibility onto weaker stakeholders – by denying some of their legitimate claims – and at the same time disclaim responsibility for the most burdensome tasks or goals by emphasizing some other conflicting claim. It is quite likely that large groups of people, characterized by a state of need concerning some basic good (i.e. education or civil rights protection) may, precisely because they are needy, not have resources to spend on political organization, while at the same time may be so dispersed across a large territory as to render political coordination and organization difficult. At the same time, since smaller professional or business groups are well concentrated, easy to coordinate, and do not illegitimately hold highly valuable assets, they may coalesce effectively in order to exercise their influence over public allocatory decisions. A reasonable view of the social contract that is capable of accommodating different claims on different levels would thus also have important consequences in terms of the prerequisites for the design of an institution able of preventing strong but illegitimate (or less legitimate) stakeholders from exploiting the decision making process that resolves on the allocation of shared responsibility.

The importance of this point cannot be overstated. If the agreement does not reflect the proper priority ordering of morally legitimate claims, it may fail to induce those who have the most urgent claims to identify with the group, thereby dooming the idea itself of 'sharing' social responsibility to failure. In fact, even though the capacity to take advantage of decisions on the allocation of responsibility may differ significantly, each stakeholder, whether strong or weak, has some strategy for avoiding proper compliance with its responsibility, perhaps by taking advantage of limited monitoring and control and concealing its behaviour.

Put as simply as possible, the aim here is to propose an impartial acceptability test of the agreement whereby principles and goals are settled in order to assign responsibilities to each stakeholder. Stakeholders' claims may be differentiated according to their level (the type of stakes) or – within the same type or level – according to the degree of incompatibility between claims of the same type put forward regarding the distribution of a given set of scarce goods or a bundle of rights over scarce resources, control over which may be complementary only to a very limited degree. The idea of an agreement must be able to accommodate conflicts deriving from both of these two sources of differentiation, and what must be accomplished first is the priority ordering of stake-types or claim-levels whereby stakeholders may put forward their different and perhaps conflicting claims.

4. Different stakeholders' claims

In the following section three types of stakes or claim-levels put forward by stakeholders will be analyzed in sequence: *need-based claims*, *merit-based claims* and *externality based claims*.

4.1. Need-based claims

Here the interests at stake – and which prompt agents to stake claims – are conceived as *needs* for *primary goods* that are necessary in order for individuals to accomplish whatever life-plan they may have. Moreover, they are *needs* for *capabilities* to transform such fundamental goods into successful (or at least acceptable) functionings within some activity or sphere of human flourishing which is also consistent with any life-plan. Such goods or capabilities enable individuals to achieve states of well-being understood as voluntary achievements of a level of decency or excellence in some condition or activity according to how they are understood through the public use of reason and discussion in any given society. As is clear, this definition accommodates both Rawls's and Sen's views of well-being and justice (Sen 2009), and also a mild Aristotelian interpretation of 'functionings'. 'Needs', understood to be more basic than mere 'desires', can be identified as the means required by anyone in order to pursue any life-plan, or the capabilities that anyone needs in order to function successfully within a wide array of activities and conditions. When all individuals equally require the same primary goods or capabilities with respect to any life-plan or broad array of basic functionings, it may be said that *their needs are the same*.

Cardinal utility may be used as a measure of needs by significantly departing from standard utility theory, where it is defined as mere representation of subjective preferences. Utility in our case may be taken as a measure of how much (technically) a good or capability is "instrumental" for the life-plan of a given agent – i.e. as effectiveness measure of the relevant good or 'capacity to function' *indispensability* as a *means* to the *end* of accomplishing a life-plan (or some part of it). Intuitively, 'utility' here means the extent to which a good is useful in order to achieve a goal, and in the event that a good or a capability were in many senses nearly indispensable as a means to reach the end of a life-plan, it could be said that they are 'needed' for that plan, and hence their 'utility' (in an instrumental sense) are very high. This may be regarded as the total probability associated with the event that a given good or 'capability to function' may happen to be the effective means for accomplishing (in the sense of a *means–end causation)* the individually chosen life-plan across the various possible states of the world and situations in which it can be used (see Roemer 1996). Thus *utility* is a measure of the extent to which an agent *needs* a good or capability given his/her life-plan.

Even though utility in this case is not a representation of subjective preferences, but stands for 'expected effectiveness', it can still account for some of the diversity between individuals. Two individuals with different life-plans may have different instrumental utilities for certain primary goods or basic capabilities, according to the differences between their life-plans and the related differences in the probabilities that some good or capability will be an effective means for fulfilling those life-plans. Thus, if we 'fine tune' the analysis of the use of goods or capabilities as a function of the expected success in fulfilling different aspects of life-plans, any two individuals' needs for the same goods or functionings may be different. However, since no acceptable realization of a life-plan can dispense with certain primary goods and basic capabilities to transform goods into basic functionings consistent with it, it follows that all agents *equally need* such *primary* goods and *basic* capabilities.

Consequently, these individuals also have similar understandings of the relationship of instrumentality (or causality) between these means and the ultimate ends of life-plans, and hence of the extent to which a person will need these goods and capabilities. There is therefore nothing to prevent an interpersonal measures of utility of these means from being agreed to in order that a common unit of measurement can be deployed to express the extent to which a given good or functioning can help achieve some aspect of the life-plan pertaining to any individual in general. Put differently, searching for a metric of the *causal relationship* (*utility* as *usefulness* in reaching some goal) between primary goods and functionings on the one hand, and the realization of the ultimate ends of life-plans on the other, does not encounter the standard difficulties relating to interpersonal comparability that on the contrary arise when utility is defined as a measure of subjective preferences. Hence, given the life-plans of two different individuals *j* and *i*, we can in principle say that some primary good or capability has more utility for the achievement of agent *j*'s life-plan ends than for the achievement of agent *i*'s life-plan ends. We are able to understand in interpersonal terms the extent to which a good or capability is instrumental to agent *j*'s life-plan as against the extent to which it is instrumental to agent *i*'s life-plan. In other words, we can determine whether individual *j* needs a good or capability more than individual *i* (even if, in order to reach a decent level of success in any life plan, primary goods and basic capabilities in general are needed in equal measure in order to accomplish both individual *i's* and individual *j's* life-plans).

Formally, this relationship between the needs of different agents can be expressed in terms of *relative needs*: that is, as the *ratio* between the positive marginal variation of agent *i's* utility for a positive marginal increase in the use or possession of a given good or capability (in terms of variation in the degree of attainment of his/her life plan) and the negative marginal variation of agent *j*'s utility for the corresponding marginal decrease in the same good or capability used by agent *j* (in terms of variation in the attainment degree of agent *j*'s life-plan). The greater the need of agent *j* (compared to agent *i*) for the good or 'capability to function' x, expressed in terms of its marginal utility for him/her, the greater will also be the marginal increase in its utility for him/her with respect to the marginal decrease in the utility of agent *j* for the same good or capability x. We can thus estimate different distributions of primary goods or capabilities across different stakeholders in terms of those stakeholders' relative needs for them (different distributions will be associated with different *ratios* of the marginal variations in utility for the agents).

In order to be clear about the measurement of *relative* needs, let us assume that an amount of a given primary good or 'capability to function' is fully distributed between two stakeholders, so that there are no further allocable shares that can further enhance the utility (needs satisfaction) of both the agents involved. In other words, by allocating the good or functioning capability we reach the *Pareto-frontier* (in economic jargon). Along this frontier only Pareto-indifferent distributions can be found (i.e. allocations that may satisfy one stakeholder's needs slightly more, but only at the cost of slightly reducing the satisfaction of another stakeholder's needs).

It should be noted that, whilst a given amount of some primary good or 'capability to function' may be allocated in various positive shares of different magnitudes to two stakeholders, when the Pareto-frontier is reached *relative needs* will react to these different allocations by different marginal increases (or decreases respectively) in agent *i*'s utility compared to different marginal decreases (or increases respectively) in agent *j*'s utility. In fact, agents *i* and *j* can both gain as long as their shares can both be augmented – which may occur insofar as their shares can both be moved from some inferior *status quo* towards the Pareto frontier (where all the different allocations of the entire amount available are represented). However, once this frontier has been reached, only *trade-offs* (involving conflicts of interests) between different stakeholders' needs will be feasible.

Nevertheless, agreement must be reached on some distribution located along the Pareto frontier. A distribution according to the criterion of *relative needs* is the one that allocates goods and capabilities to stakeholders according to the *proportion* (inverse) resulting from the ratio between marginal variations in the agents' utilities (needs) *which is associated* (along the Pareto frontier) *with the allocation itself.*[5] This occurs in the locus of the Pareto frontier

[5] Although it may not be obvious, this condition is not trivial: given a unique measure of relative needs (i.e. the inclination of the tangent to the Pareto frontier) , there is *one single* proportion in the distribution of goods and capabilities that can equate the relevant ratio between the marginal variations of needs. Moreover, when the measure of relative needs changes continuously along a *convex* Pareto frontier, there is *only one* possible value

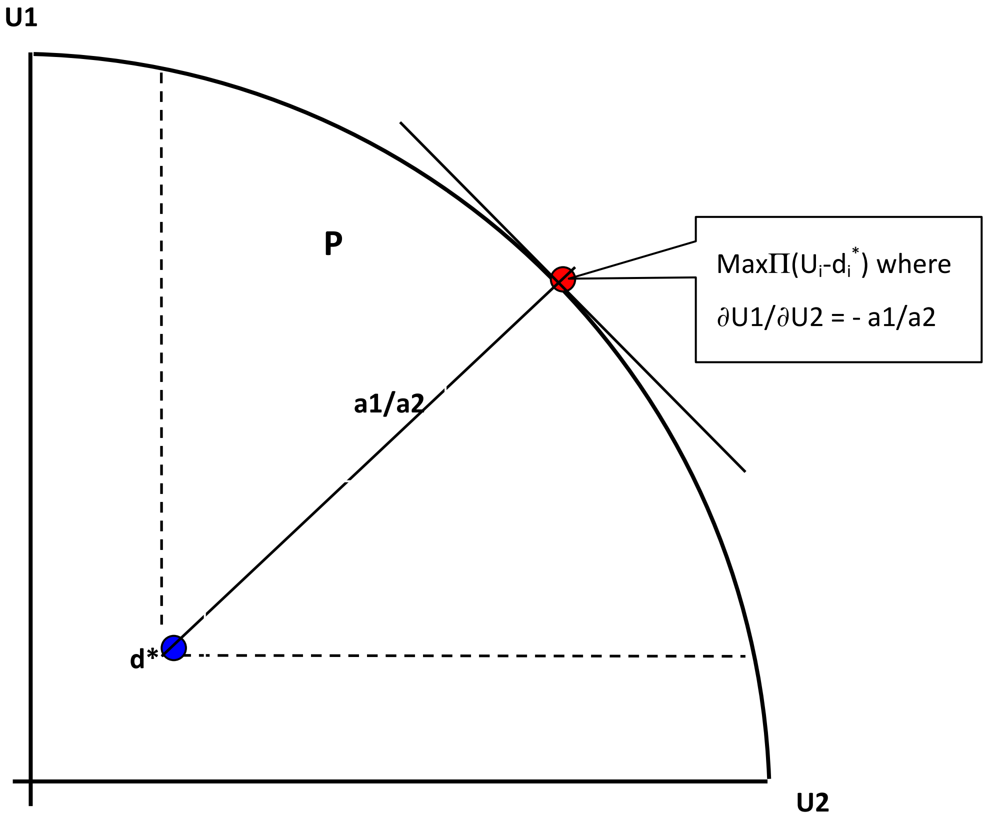
where the Nash bargaining product is maximized, i.e. the point at which the multiplication of the stakeholders' utilities – deriving from agreements on the allocation of shares –, after deduction of what they would have gained anyway in the absence of any agreement, is maximized. This, in fact, is the mathematical condition for a 'rational solution' of the cooperative bargaining problem over the distributions of primary goods or capabilities between different (i.e. at least two) stakeholders represented through utility-payoffs. Moreover it coincides with *proportionality to relative needs* (see fig.1).[6] It identifies a unique solution for the bargaining problem between stakeholders proposing need-based claims.

Primary goods and 'capabilities to function' may be understood as the contents of basic citizenship rights. Thus, needs-based claims are the appropriate basis for the allocation of citizenship rights over scarce primary goods and entitlements to capabilities that enable an agent to use goods in order to transform them into achievements of some relevant functionings. In order to give every stakeholder equal consideration and respect, rights commanding control over such goods and capabilities must *equate* their needs; that is, in order to make them *equal* the distribution must be proportional to *relative needs*. But it has already been stated that, even though certain details of individual life-plans may differ, and hence may require some fine tuning in the distribution of goods and capabilities in relation to different individuals' life-plans, primary goods and basic capabilities are nevertheless *equally needed* by every individual living under similar external (social, economic, technological and environmental) conditions. In fact, under similar conditions they are equally necessary for the appropriate realization of all life-plans. Hence, under similar conditions, a distribution of rights (entitlements) between primary goods and basic capabilities that can be recognized as proportional to relative needs is basically *egalitarian*.

---

of this ratio of marginal needs that can be produced through a distribution of goods and capabilities according to a proportion reproducing the same (inverse) ratio (see fig. 1 in the main text).

[6] As a matter of fact, the unique proportion of shares of goods and capabilities that may be allocated in order to equate the ratio of marginal variations of needs occurs at exactly the place on the Pareto frontier where the Nash bargaining product is maximized (see Brock 1978/9, Sacconi 1991, 2000, 2006a).



(Fig. 1. Nash Bargaining Solution of a two-player cooperative bargaining game in the event of a symmetrical payoff space P. The point selected on the Pareto frontier maximizes the multiplication of the players' utility differences ($U_i - d_i^*$); the bisector starting from the status quo d allocates payoffs (representing needs for goods and capabilities) according to a ratio - a_1/a_2 - identical to the ratio* $\partial U_1/\partial U_2$ *of marginal utilities (relative needs) represented by the inclination of the tangent to the frontier exactly at the point where the Nash bargaining product is maximized. Hence, by selecting the point where the Nash bargaining product is maximized, the distribution of payoffs is also proportional to relative needs. See Brock 1978/9)*

4.2. Merit-based claims.

In this case, interests at stake are understood according to the 'interest in gaining access to some benefit in proportion to one's personal contribution or effort'. Merit is therefore understood simply as *contribution* to some social surplus. The claim to a share of surplus is fair insofar as it reflects a *personal* contribution.

Moral merit is not invoked here, because it is an empty notion that must be defined according to some further moral concept. For example, someone may morally deserve a share in a given pie in proportion to his/her *needs* for the pie or to his/her contributions to the pie's production.

Merit-based claims may also implicitly refer to *talent* or *natural endowment* with some *skill* and biological *capability*, since it is obvious that talent affects the capability to contribute or to produce effort at a lower psychological cost. Unless remunerations are to be proportional to the 'pain' of producing effort, merit-claims may therefore consist in claims to higher remuneration for more talented persons (otherwise, were it the 'pain' of putting effort into a job that is considered important, then -- assuming an equal level of effort -- talented persons would be paid less than untalented ones because producing the same effort level would "cost" the talented person less "pain" than the untalented one).

However, an implicit reference to talents does not seem to provide a sound foundation for meritocracy. Rawls's (1971) criticism of talent-based principles of justice is compelling: talents are the fruit of a morally arbitrary natural (and socio-biological) lottery, and nobody can claim either to deserve her/his natural talents or to possess any merit for 'having been given' such talents. Hence, if the casual distribution of talents were reproduced by the distribution of goods or rights, also the final distribution of outcomes and the corresponding social structure would become morally arbitrary.

Insofar as talent affects the natural endowment of capabilities, it seems to interfere with the distribution of capabilities according to needs. Assume that basic goods and capabilities are distributed according to relative needs. Then the people who are most talented have the least need for capabilities. They should therefore receive less, so that a wide distribution of capabilities according to relative needs could attempt to level out these differences in capabilities by giving more primary goods and learnable capabilities to the less talented. However this egalitarian redistribution of capabilities seems idealistic and unreal. There is probably no way, at least in the short-run, to equalize natural capability endowments through their redistribution, even when needs constitute the basic criterion for allocating entitlements over resources that can be used to 'learn' capabilities.

Moreover, the exploitation of talents must be incentivized through some reward for their use and cultivation. At the same time, inequalities resulting from the exploitation of talents would be unjustified unless such exploitation were justified by some principle. This is the reason why a regulation of inequalities generated through the differential exercise of talents is required according to a principle of justice such as the Rawlsian 'difference principle'. Since inequalities that result directly from talents are not justifiable, even if some reward may be causally necessary to incentivize their use and cultivation, the differential remuneration due to talents may be based only on the benefits for all resulting from their use (including the worst-off). Since inequalities incentivizing the use of talents favour the talented, the only relevant criterion by which to decide on the magnitude of these inequalities is the extent to which they affect the position of the worst-off. Thus, the maximum deviation from equality that is acceptable for incentivizing the use of talents and their cultivation is that associated with the maximization of the absolute level of the satisfaction of the needs of the worst-off, and hence the overall capacity to contribute to mutual benefit (not just the talented's benefit). Any wider inequality is unacceptable.[7] Thus there cannot be any justification on talent grounds (for example the talent in using financial tools) for the huge increase in social inequalities throughout the world over the last three decades. This is certainly the case for developed countries such as the US, the UK, Ireland and Italy, as well as most of the developing ones – with particular reference to managerial bonuses and shareholder returns compared against average wages. To summarize, only contributions, and not talents, can be regarded as a further legitimate source for claims to shares in some social surplus.

Contributions are given to coalitions of agents that jointly produce marketable goods (enterprises) or that through collective action provide some universal public goods, local public goods, commons or club good. These are cooperative activities that often involve specific investments by their participants. 'Specific investments' are decisions to specialize in some asset or resource in order subsequently to increase its value within a given specific bilateral exchange or transaction. Such investments are sunk costs which are not recoverable outside a specific relationship. They create a reciprocal (but not symmetrical) relationship of dependence for mutual advantage. They may have different natures: human capital may become more specialized by learning specific techniques, languages or codes of behaviour. Creative and inventive work entails the specific investment of human capital in a specific project before any result can be achieved. If the worker is removed from the project before the result is obtained, the sunk cost of all of the preparatory activity – which is essential for the invention – cannot be recovered by embarking on another project. Specialist training may be a specific investment, but even workers who are not highly-skilled may carry out specific investments, sunk costs and idiosyncratic work relations – for example immigrant workers are idiosyncratically locked into specific employment relationships if a breach of their contract of

[7] This is a moral inference of the distributive 'difference principle' from a more basic idea of moral equality; for a different 'game theoretical' deduction see Binmore 2005 and Sacconi 2010a

employment may entail their expulsion from the country and the loss of their sunk immigration costs. Specific investments may also be made by the providers of services and technologies specifically dedicated to the productive process of their client's firm or organization. Financial risk capital may be idiosyncratically invested in the acquisition and development of technologies, plant and equipment so that their costs cannot be recovered before some long period of production activity has elapsed. Social capital, i.e. fiduciary relations and trust, is a specific investment that cannot be useful outside a given relation-network, and so if an entrepreneur, worker or consumer endowed with high social capital is expelled from his/her relation-network, then his/her social capital will also be significantly devalued, and consequently so too all of his/her trust-based transactions. Consumers also invest specifically in research, information gathering and relation-building in order to establish idiosyncratic relationships with professionals who are selected in order to establish a long-term service relationship based on trust (consider the cases of doctors and lawyers, but also financial, banking and insurance professionals).

Specific investments are often multilateral, complementary and conducive to team production. The typical feature of team production is that by increasing the number of contributors who join a coalition and make specific investments, the coalition's productivity will increase more than proportionally. In other words, the production function is super-additive. Moreover, the team's additional surplus compared to the alternative scenario of separate production cannot be ascribed to any particular individual and split into separate shares attributable to any of them, because it is only the cooperative interaction within the team that makes the surplus possible. Thus the good side of team production is that cooperation between the team members produces more than the sum of their separate activities. The surplus (or parts of it) cannot be attributed to any of the team members, but only allocated to the group as a whole. However, there is also a problematic side: because there is no separable measure of personal productivity, it is impossible to remunerate individuals in such a way that reflects their personal contributions. When an individual's contribution cannot be separated and rewarded as such, opportunistic behaviour may arise within the team, without being detected and directly sanctioned. Given a suboptimal collective output with respect to the potential best joint output, it is impossible to say who has contributed efficiently and who has not; and it is also impossible to pay team members in proportion to their actual individual marginal productivities.

However, to assert that productivity and contributions are joint and investments are specific and complementary is not to imply that there is no measure of how much individuals contribute to cooperation. In fact, the contribution of any additional member is the marginal variation of the coalition's value as a function of each new additional member who joins the coalition. Put simply: this is not a measure of the separable contribution of any particular member and cannot be attributed only to his/her merit. Exactly the same marginal increase of surplus can in fact be observed also by reversing the order in which individuals join the team – so that by interchange the 'marginal individual' joining the team according to any order, the same increase in team productivity will still result. In other words, whoever is marginally added to a given team dimension will induce the same marginal variation of the coalition value. This means that all the group members can be allocated responsibility for any variation of the team value. Thus their contributions are equal. Summing up, a measure of the individual's overall expected contribution is the marginal variation of the value of each coalition that he/she may enter as a member (in whatever order) *multiplied by* the probability that any coalition structure will occur, assuming that each member's contribution is equal when specific investments are multilateral and complementary because it is equally essential for the surplus as a whole.

Contribution-based claims – understood as both individual contributions and team production contributions – are legitimate claims to shares of the wealth surplus created by cooperative ventures in companies, organizations and collective action in the production of marketable private goods or public or quasi-public non-marketable goods (club goods, local commons etc.). The moral basis for these claims is simply that whoever participates in the production of a given surplus should receive a share in proportion to his/her contribution to its production. It is then intuitive that many apparently 'meritocratic' reward systems that bring about large inequalities in developed and developing countries cannot be justified on the basis of the 'proportionality to contribution' principle. In fact there is no reason to think that the tenfold or greater multiplication of the remuneration differentials between CEOs and the average employees in the typical capitalist company over the last two or three decades has been related to a proportional increase in the top managers' efforts or their personal contributions to the overall value created by the company compared to the contribution provided by typical employees (which would have decreased in proportion tenfold or more). Nor there is any evidence to support the belief that the largest remuneration differentials and highest bonuses paid to top managers occur in companies where the shareholder value has increased most in the long run.

There is also a theoretical reason for not believing that the striking growth in inequalities and pay differentials in many companies has any relationship to any notion of merit. In fact, incentive mechanisms – such as bonuses and stock options – introduced in order to link managers' remunerations to share-value performance, are based on the realistic hypothesis that information held by principals and agents is asymmetric. Thus assessing managers' contributions directly in terms of observation of their behaviour and efforts is impossible for their principals. Managers hold private information on their efforts and hence cannot be rewarded on the basis of their own reports of their efforts. For this reason, managerial pay is based on financial outcomes which are only probabilistically and indirectly correlated to unobservable effort. Share value is taken as an indirect proxy indicator of the consequences on shareholder equity of managers' behaviour and efforts. But it should be noted that in order to encourage managers to work harder, rather than adopting a line of managerial slack in situations with higher productivity, an outcome-based remuneration structure must pay them more for different outcomes that would occur with exactly the same effort level, but under different exogenous productivity conditions. When high effort under unfavourable exogenous productivity conditions produces the same output as low effort under good productivity conditions, only this kind of incentive may encourage a manager to choose high effort under good conditions in order to obtain the best output. But this involves paying different compensation for the same effort (merit). Premiums paid for better outcomes constitute an 'information rent' that managers are able to extract by threatening not to make an effort when it is impossible to distinguish between cases of high effort under unfavourable productivity conditions from those of low effort under favourable productivity conditions. Thus, according to an honest reading of economic theory, the level of bonuses and financial incentives paid is proportional to the managers' threat to resort to opportunist and lazy behaviour when their efforts cannot be directly measured. However, the principle "to each according to his/her threat power and informative advantage" sounds quite different from "to each according to his/her contribution (merit)".

4.3. Externality based claims

Interests in this case are claims to *redress* for social costs and externalities that spill over onto stakeholders – individuals, groups or entire communities – as the external cumulative effects of multiple individual decisions.

Externalities originate from the interference of market transactions with the allocation of public or common goods, or failures of collective actions aimed at their production or preservation. Basically, they occur according to the logic of the Free Rider: one or both of the parties to a market transaction may consume some public or common resource, thereby generating social costs (i.e. external to the transaction) that spill over onto third parties – or also onto the parties themselves insofar as they are interested in concluding their transaction within certain constraints on the consumption of common resources (excessive consumption generates a social cost while reducing the private costs borne by participants in a particular market transaction). Some participants in collective action aimed at the maintenance of a public good or a common may refrain from playing their part, profiting unfairly from the other parties' contributions. But also in a purely private domain, some members of a team may exploit the unverifiability of individual contributions to team production in order to refrain from more efficient efforts, so as to gain advantages from the other members' work.

The main source of difficulty where there is a large number of actors involved in decisions giving rise to an externality is that it is a cumulative effect that spills over onto other agents, or onto the same agents, although these decisions are not intended to induce such effects, but are instead aimed at gaining profit from a private transaction involving only a subset of the externality's stakeholders (regarded as participants in a private exchange). Since externalities are the cumulative effects of many decisions, the individual agent may not recognize his/her causal power over them. In actual fact, the decision may be deliberately intended to yield some advantage from an apparently 'freely available' resource or an opportunity to reduce costs. But because the individual decision is not directly intended to achieve that cumulative effect, the agent may not recognize the intentional nature of the decision contributing to the cause of the external effect. Hence the individual agent will disavow any responsibility for it. This difficulty is even greater when an externality may be brought about by decisions whereby some merit claim is met in a parallel domain (private production of some marketable good). Moreover, needs-based claims may already have been met for the same common or public good in the past, so that the externality subsequently interferes only with its marginal use or enjoyment by third parties, without calling the entire basic entitlement on such goods into question. Thus the infringement of needs principles of justice is not apparent.

The sharing of social responsibility for these claims is not immediate, since those who are typically responsible for externalities do not perceive themselves as the relevant decision makers. Indeed, these claims are not usually raised within a process of collective decision making concerning cooperative joint action, such as when a primary good must be distributed or a coalition must be formed to develop and exploit it. But it is under these circumstances that agents may accept a "we frame" as cooperators. Externality, on the contrary, emerges only as an unintentional interference with the implementation of these decisions. Externality-claims must be brought against individuals participating in a market decision who do not regard themselves as involved in decisions concerning needs and the reproduction or development of public goods or commons, and who *prima facie* disclaim their responsibility for them.

Overall, however, these situations obviously make distribution inconsistent with contribution, if contribution is considered also to concern the joint maintenance, development and exploitation of some commons or public goods. Moreover, they may contradict a previous distribution proportional to relative needs. In fact, even if two individuals may cooperate within a bilateral transaction and share a surplus proportional to their relative contributions, they may also collude in order to profit from the appropriation of some common resource to the detriment of other parties who do not participate in their private transaction, and who thus see their access to the commons reduced (or crowded out), or their involuntary consumption of some public disadvantage (i.e. pollution) increased.

Accordingly, externality-based claims may be regarded as not independent of those based on the two above principles, namely proportionality to needs and proportionality to contributions. They should be better understood as claims for redress brought because previous need-based claims and merit-based claims have not been met. This is part of what is understand here as externality-based claims: they are claims for redress or compensation for the failure to honour more basic moral claims based on needs or merits. Nevertheless, due to their individual characteristics they should be considered as distinct from other claims because they may be made after some initial decision concerning primary goods or capabilities has been taken according to needs; and they may be made collaterally to allocations and distributive decisions made in some domain where proportionality to contributions standard appears to be met. Thus, externality-based claims are made as further stakes *prima facie* independently of, and sometimes in additional to or against, other claims based on the foregoing two principles (typically *against* some special case of distribution according to contributions in the market domain).

For example, access to a common environmental resource may have been initially allocated in proportion to needs between A, B and C. Then private activities begin wherein a subset of stakeholders (say A and B) participate in the production of some private good and are remunerated according to their contributions, but their use of the environmental resource creates externalities for the agent C who does not participate in it. As a result, over the following period of time this third party may draw on a smaller endowment of natural resources, though at the same time has not benefited from the private business between A and B. In this description, need-claims are respected at time 1, contribution-based claims are met together for A and B at time 2, although time 3, because the externality spills over onto C, the principle of relative needs in the allocation of the common good is no longer effective. This justifies agent C's externality-based claim for redress.

Second example: at time 1, resources to provide collective security or basic education are allocated to A, B and C according their equal need for security and basic education. They are then under the obligation to participate in a collective action of ongoing maintenance and the development of security or education. Their relative contributions may result in wealth differences, though without interfering with their equal claims for security and basic education. However, A and B may develop a technology for their private exchanges that enables them both also to improve their private security and the education of their children through their private exchange. They thus reduce their participation in the collective action providing collective security and basic education, yet nevertheless gain from C's continuing participation in the collective effort (it should be remembered that private security is worse than public security if also the third individual stops providing public security, and also overall education is in general worsened if agent C stops his/her contribution to basic education). Overall, C obtains less security and lower-quality basic education than what he/she deserves, and this provision is also inadequate for his/her needs. A and B instead fairly profit from their contributions to their private transaction, but they also enjoy additional security and education due to C's contribution to public security and education, so that their level of security and their children's education is somewhat higher than the original level (low-quality basic education is supplemented by private specialist education). Thus, need-claims are met at time 1, and merit-claims are also apparently respected within a stakeholders' sub-domain at time 2. But at time 3, from a broader perspective, merit claims to the maintenance of security and basic education are not respected, and the distribution of security and education no longer respects agent's C needs. Hence C makes an externality-based claim for redress.

Thus, even if externality-based claims are not logically independent, they may arise empirically and temporally as claims separate from needs- and contributions-based claims. *Prima facie*, externality-based claims may clash with contribution-based claims that only on further reflection reveal themselves to be illegitimate changes compared to previous distributions according to need-based claims. Moreover, they may also clash with contribution-based claims related to the functioning of some sectional productive coalition which is shown on closer analysis to infringe also other contribution-based claims regarded, from a broader perspective, as being based on fair contributions to the maintenance, development and production of some public good or commons. Externality-based claims thus make it possible to include consideration of a dynamic element within our representation of possible types of claims and stakes: how previously stated needs-claims or merit-claims can in logical terms be turned against themselves, and how they can be restored in different form through redress and compensation after other merit claims, mainly exercised through market transactions, have been made and met.

5. Spheres of distributive justice and the priority ranking of claims

The problem of the priority of the types of moral claims considered above may be resolved as follows. First, it is necessary to determine the proper context within which each type of claim predominates in the argumentation aimed at reaching impartial agreement. Second, the relative priority of each context is settled in terms of a model of constitutional and post-constitutional social contract. In other words, we must first identify the spheres of application within which each principle of justice is intuitively appropriate, and then use the idea of a social contract in order to settle the logical priority and succession between different spheres so as to derive the overall agreement that sequentially satisfies each principle (type of claim) in due order. The relevant spheres of justice are

a) *Manna from heaven context*. This is a context where unanimous agreement must be reached on rights that give each stakeholder a privileged relationship with shares (if divisible) or use-rights (if not divisible) of primary goods or basic capabilities seen as "manna from heaven". 'Manna' is taken as given, not produced. Before any productive activity begins and any contribution is made to the manna's maintenance or development, and before any production is undertaken through its use, a collective agreement on the criterion for the manna's distribution must be endorsed by all of the parties involved. A constitution of rights must be established before any further cooperative productive activity is undertaken whereby rights granting control over goods and capabilities may be employed as assets. Cooperation at the 'manna level' occurs prior to further cooperative employment of the manna, because stakeholders can use it for any goal only if they first agree on a principle of fair and impartial manna distribution. Otherwise primary goods and capabilities will vanish, or better will be destroyed by the eruption of conflicts among stakeholders. The idea is that manna is 'cooperative in nature', since it is only useful for furthering the society members' life-plan, while it is unsuited to living in a 'state of nature' characterized by radical conflict, which would instead entail its rapid depletion.

b) *Non-manna from heaven context, type I*. These are contexts within which 'manna goods and capabilities' are actively employed as inputs for the production of further goods and services and to bring about various cooperative surpluses. Surplus, of course, is not manna. In this context individuals join coalitions of various sizes in which they contribute their manna endowment as inputs, and cooperation within coalitions makes complementary-specific investments and team production possible. Effort (based on capabilities) is provided on an individual basis and also as inseparable productivity. Individuals expect some return from the various coalitions in which they participate in the production of many social surpluses. Coalitions are mainly seen here as firms and productive organizations operating in the market and providing marketable goods and services.

c) *Non-manna from heaven context, type II*. Many primary goods are public goods (universal, local or club goods) or commons, or include components of them (for example 'education' has a public good component). Whilst they may be initially distributed as manna, the maintenance, reproduction and development of these goods is nevertheless the result of collective action. Many individuals contribute their basic endowments of such primary goods as resources to an all-inclusive coalition or to sub-coalitions including subsets of the population aimed respectively at the maintenance, reproduction or development of universal public goods or local public goods and commons. Type II non-manna contexts include mainly non-market productive organizations providing types of social surplus through cooperation. Non-manna activities of type I, however, interfere with the maintenance of public goods or commons through external effects. Thus, an individual who does not participate in some non-manna productive type I coalitions may see her/his manna endowment allocated to collective action in the non-manna type II context reduced or devalued as a result of the externalities spilling over onto her/him from these collateral market activities.

It seems natural to order these contexts according to a logical sequence whereby, as the first step, the manna context arises and rights are allocated over basic goods and capabilities that have not yet been produced or elaborated through cooperation, whereas they are appropriated as they are given exogenously. These are resources that agents can voluntarily employ for the purposes of their life-plans only after their entitlement has been settled. Secondly, the non-manna context of type I subsequently occurs and the manna endowments are exploited as resources for producing something extra (surpluses) through investments in cooperative activities and exchanges determined by private transactions between various individuals – and which can take place only according to voluntary agreements concerning the rightful use of manna endowments. Subsequent to the manna context, but in parallel with the first non-manna context, the non-manna context of type II also occurs in step 3: now public goods and commons, which are part of what was initially allocated as manna through rights assignments, are preserved (in fact manna, if not cared for, may deteriorate), improved and developed, and then used for the purposes of individual life-plans. These are also cooperative and voluntary activities carried out through coalitions of varying degrees that arise only because individuals confer their initial manna endowments. Hence also this step is strictly logically subsequent to the first step in the manna context. In this non-manna setting, market activities carried out in type I non-manna contexts may interfere through externalities that diminish the available level of manna, reduce its usefulness for right-holders and cause cooperative activities devoted to the manna's preservation and improvement to fail. However, since such cooperation is subject to the typical Prisoners Dilemma and Public Good games paradoxes, it may also fail by itself.

It can thus be ideally argued that, in the initial step, participants in a collective decision face a problem of distributive justice concerning the manna's allocation: that is, the allocation of resources that are available before any contribution has yet been made, and that can be used to enter mutually advantageous cooperation only after allocation. At this stage therefore, the only acceptable principle is *proportionality to relative needs*, since no claim based on merit can have been put forward yet – since no contribution has yet been provided and natural talents do not matter as a basis for distributive principles. Rights settled by agreement – which are logically claim-rights (Hohfeld) – may only have the content of *claims based on relative needs*, so that the principle for the allocation of rights is "*to anyone according to his/her relative needs*".

Thereafter, individuals use primary goods and capabilities over which they have control to make investments and to establish further cooperative relations by participating in various joint production and exchange activities which generate surpluses in terms of both private goods and public goods or commons. Parties enter into multiple cooperative coalitions to which they make contributions within the limits of their original endowments aimed at increasing the value of what they may enjoy compared to how they would fare by relying on their initial endowments. This can sometimes be done through complementary specific investments and team production. Participants join each coalition by agreeing on the rights that they will legitimately claim later on over the distribution of the surplus that the coalition produces. Since the needs have already been met and participants make contributions, the distribution of the benefits must be proportional to the relative contributions. At this stage, rights to distributive shares can only take as their content claims based on merit understood as relative contributions. Thus the rights allocation principle is "*to anyone according to her/his relative contributions*".

However, since each individual deploys at this stage exactly the resources that he/she has received at the first step, and since these resources were allocated according to the relative needs principle, the end result (by definition consistent with the proportionality to contributions principle) must also reflect the proportionality to needs principle. In fact, rational participation in any cooperative venture is regulated on an individual level by utility maximization, which in our context is essentially a measure of needs satisfaction (i.e. utility maximization means that a resource is deployed efficiently by the individual as a means needed to accomplish her/his life-plan). Assuming that opportunistic behaviour does not prevail, each individual pushes his/her cooperative effort in any coalition insofar as it is conducive to the satisfaction of his/her needs. Of course, this cannot prevent inequality of needs satisfaction if individuals have arbitrary or strongly unequal endowments. But this is not the case here, because the non-manna context is reached only after the distribution of basic goods and capabilities at the first step has been carried out in proportion to relative needs. Therefore, considering that agents maximize their utility (in our needs-related sense), the final distribution according to contribution (and merit) is also proportional to relative needs. Meritocracy in this sense is literally secondary compared to the priority of the needs principle, and it is necessary in order to achieve a final distribution that reflects relative needs.

This still only occurs if productive coalitions, using the basic endowments of their participants, do not negatively affect the basic endowments of other stakeholders: that is, if they only deploy the shares of basic endowments which agents voluntarily confer in order to participate in collective action or voluntarily exchanges, without causing external negative effects on other agents that do not participate. Special attention must therefore be paid to the third step, where the non-manna context is entered into with regard to the preservation, development and deployment of public goods and commons. Since this too is a non-manna context, the proper principle of justice is proportionality to relative contribution, even though in contexts where individual contributions are not separable, remuneration based on contribution must be egalitarian. However, the essential feature is that in these contexts externalities deriving from cooperative activities undertaken in the non-manna context type I are possible. Therefore, in order to maintain consistency with the basic principle of relative needs and the condition relating to relative contribution, at this stage the distributive principle must be "*redress anyone according to the external effects to which he/she has been subjected*". Thus the contents of rights settled at this stage are claims based on externalities.

A natural way of encapsulating the operation of the ordering of spheres of justice and the priority of the relative needs principle with respect to other principles of justice is to represent the subject by means of a *social contract model* organized as a two-step collective bargain: the *constitutional contract* and *post-constitutional* contracts (see Brock 1978/9, Sacconi 1991, 2000, 2006a, 2010). The *first* stage is a collective choice on the constitution of rights modelled by a bargaining game amongst all cooperating stakeholders. The second stage is a coalition game that the stakeholders play within the rules of the game (the constitution) selected during the first stage. This second step generates a final allocation of payoffs.

Because they are linked sequentially, the two games can be resolved by reasoning backwards. Constitutions are regarded as restrictions on the strategies available to the players in the second-stage game (restrictions on one agent's freedom of action correspond to other agents' rights, protections and endowments). Each second-stage game has a solution in terms of payoff allocation, so that from the perspective of the first-stage, a Constitution may be selected according to the final allocation of payoffs associated with the second stage game as foreseen from the *ex ante* constitutional perspective. Because this is a bargaining game, the first-stage game is resolved by the most accredited solution concept for such games, i.e. the *Nash bargaining solution*, which prescribes the maximization of the product of the players' payoffs for agreements after deduction of status quo payoffs. Formally, this coincides with a distribution of the surplus proportional to relative marginal variations of the players' utility. Under the additional assumption of interpersonal utility comparability, this may be interpreted as distribution proportional to a measure of their *relative needs*. The second-stage game is a coalitional cooperative game played within a given institutional framework (a given constitution of rights) that assigns each player certain rights and obligations. It must be solved by way of a solution concept that allocates a quantum of utility to each player related to his/her importance for each possible coalition. This brings us to a distributive principle based on proportionality to relative contribution. But it should be noted that the institutional arrangement – a structure of rights and duties that influences the level of each player's contribution to each coalition – is chosen on the constitutional level, so that what players are able to gain on the basis of their contributions is also a distribution acceptable from the constitutional point of view according to the relative needs principle.

In addition to the two-step bargaining game outlined above, a third step must be introduced in order to account for the *ex post* decision on the amount of redress that it is agreed to pay after a coalitional game has been selected via a constitution, if the game's cooperative solution relative to some domain causes external negative effects on the participants' endowments that gives rise to a deviation from the bargaining solution as foreseen from the initial constitutional perspective (as a matter of fact, the solution to the selected coalition game due to externalities will not coincide with the Nash bargaining solution in the constitutional step because, within the given constitution, a coalition game does not prevent external effects affecting the basic endowments of some players). Thus in the post-constitutional stage, a third solution agreed to by participants is added to each constitution concerning fair redress. These rules associated with each constitution can also be envisaged from the constitutional agreement perspective, so that in the first stage only a constitution with the *proper* fair redress procedure added to the second stage coalition game will be selected. This corresponds to a specific constitution that, thanks to the proper redress rule, allows only for the formation of coalitions that either do not induce external effects or incorporate fair redress for social costs, so that the end result again approximates a distribution according to relative needs.

6. Concluding remarks

The above pages have set out an abstract model of constitutional and post-constitutional contracts. By adhering to the logic of agreement appropriate at each step of the model (bargaining, coalition cooperative games and agreement over redress rules) we can infer the required priority order of the different stakeholders' claims discussed in the previous sections. However, the most interesting aspect here is the implications this will have for a non-ideal view of the agreements that can be settled on SSR policies. The deliberative process and implementation mechanism for principles and goals whereby shared social responsibility is allocated to the different stakeholders must reflect the priority ordering of principles of justice that entail fair treatment of the stakeholders' different claims. SSR can follow from the legitimacy of the claims satisfied by the policies. However, nothing will follow in terms of sharing responsibility if claims are not satisfied in the proper order according to the social contract model.

We may therefore conclude that a '*de facto*' agreement on policies achieved by stakeholders that puts merit claims before the satisfaction of the needs claims of the relevant stakeholders (for example, consider health programs designed to yield the maximum profit for private insurance companies or to improve doctors' pay, or rebuilding plans after earthquakes that favour construction companies, etc.) is illegitimate, and hence unable to bring about any effective sharing of social responsibility. In this case, the policy would not be consistent with subordination of the 'remuneration according to contributions' principle to the goal of achieving outcomes that satisfy need-based claims.

In real-world situations, *de facto* bargains resting upon power relations between stakeholders participating in a deliberative process may give rise to decisions achieved by stakeholders only by pursuing merit claims. In fact, it is simpler for them to join a collective action, since they are already concentrated into a homogeneous professional group, are less dispersed across a broad territory, and have valuable resources to be dedicated to political pressure. By contrast, it is simple to predict that externality-based claims may be underrepresented in the deliberative process leading to the settlement of principles and goals to be implemented by means of a call for SSR. In fact, externalities (for example environmental ones, but also the depletion of the public good component of education) typically spill over onto a large number of dispersed and not necessarily communicating and coordinated people who are unable to affect the decision process. In all these cases however, since actual deliberation procedures and implementation mechanisms do not comply with the social contact model, no proclamation of SSR can be effective in eliciting a real sharing of social responsibility. This means it is even more important that the governance mechanisms and deliberative processes of SSR that can incorporate the normative requisites deriving from the social contract model (see Sacconi 2011 *infra*) be properly designed.

REFERENCES

Bacharach M. (1999): "Interactive Team Reasoning: A Contribution to the Theory of Cooperation," *Research in Economics*, 53, 30.

Bacharach M. (2006), *Beyond Individual Choice. Teams and Frames in Game theory*, edited by R.Sugden and N.Gold, Princeton U.P., Princeton and Oxford.

Berle A., G. Means (1932), *The Modern Corporation and Private Property*, Mac Millan, New York.

Binmore, K. (2005), *Natural Justice*, Oxford, Oxford University Press.

Buchanan J. (1975), *The Limits of Liberty,* The Univ. of Chicago Press

Brock H.W. (1978), "A New Theory of Social Justice Based on the Mathematical Theory of Games", in Ordeshook (ed.) *Game Theory and Political Science*, N.Y. University Press.

Brock H.W. (1979), "A Game Theoretical Account of Social Justice", *Theory and Decision,* 11, pp. 239-265.

Danley J. (1988), "Ought" Implies "Can" or The Moral Relevance of the Theory of the Firm, *Journal of Business Ethics*, 7, pp.23-28

Danley J. (1990), "Corporate Moral Agency: The Case for Anthropological Bigotry", in Hoffman M. and Mills Moore J. (eds.), *Business Ethics*, McGaw-Hill Publ., N.Y.

De George R. (1982), "Can Corporation Have Moral Responsibility?", in M.Paven. (ed.), *Collective Responsibility in the Professions*, University of Dayton Review, 1981-1982, pp.3-15

Dezau, A. and D. North (1994), 'Shared mental models: Ideologies and institutions', *KIKLOS*, **47** (1), pp. 1–31.

French P. (1979), "The Corporation as a Moral Person", *Amer. Phil. Quarterly*, 3, pp.207-215,

French P. (1984), *Collective and corporate responsibility*, Columbia University Press, N.Y.

Frankena W. (1963), *Ethics*, Prentice-Hall, Englewood Cliffs, (trad. it.Etica,Comunità, Milano, 1981).

Freeman R.E. and J. McVea (2002) *A stakeholder approach to Strategic management,* Working paper n.01-02 , Darden Graduate School of Business Administration

Friedman M.,(1970), "The Social responsibility of Business is to Increase its Profit", *New York Magazine*, September 13,

Gauthier D. (1986), *Morals by Agreement*, Clarendon Press, Oxford..

Goodpaster K.,(1984), "The Concept of Corporate Responsibility", in T.Regan (ed.) *Just Business*, Random House, N.Y.

Grimalda, G. and L. Sacconi (2005), 'The constitution of the not-for-profit organisation: reciprocal conformity to morality', *Constitutional Political Economy*, **16** (3), pp. 249–76.

Hare R. (1981), *Moral Thinking*, Oxford (Clarendon Press).

Hare R.(1963), *Reason and Freedom,* Oxford (Oxford University Press).

Harsanyi J.C. (1977): *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*, Cambridge (Cambridge University Press).

Hart H. (1982), "Legal Rights", in *Essays on Bentham*, Oxford Clarendon Press.

Hart H.(1968), *Punishment and Responsibility*, Oxford, Oxford U.P.

Hohfeld W. (1923), *Fundamental Legal Conceptions,* New Haven, Yale University Press.

Kaufman, A., L. Zacharias (1992), 'From trust to contract: The legal language of managerial ideology, 1920-1980', *Business History Review*, vol. 66, pp.523-572.

Kaufman, A., L. Zacharias, M. Karson (1995), *Managers vs. owners: The struggle for corporate control in American democracy*, Oxford U.P, Oxford.

Ladd J. (1984), "Corporate Mythology and Individual responsibility", *International Journal of applied philosophy*, 2, pp.1-21

Nash J. (1950), "The Bargaining Problem", *Economertica,* 18, pp.155-162

Rawls J. (1971), *A Theory of Justice,* Oxford U.P., Oxford.

Rawls J. (1993), *Political Liberalism*, Columbia U.P. New York.

Roemer J. E. (1996), *Theories of Distributive Justice*, Harvard U.P, Cambridge Mass.

Sacconi L. (1991), *Etica degli affari, individui, imprese e mercati nella prospettiva dell'etica razionale,* Milano (Il Saggiatore)

Sacconi L. (2000), *The Social Contract of the Firm. Economics, Ethics and Organisation,* Springer Verlag, Berlin.

Sacconi L. (2005), “Responsabilità morale”, in Sacconi L. (a cura di), *Guida critica alla responsabilità sociale e al governo di impresa*, Bancaria editrice, Roma

Sacconi, L. (2006a), ‘A Social Contract Account For CSR as Extended Model of Corporate Governance (Part I): Rational Bargaining and Justification’, *Journal of Business Ethics*, Volume 68, Number 3 / October, 2006, pp.259-281

Sacconi L. (2006b) “CSR as a model of extended corporate governance, an explanation based on the economic theories of social contract, reputation and reciprocal conformism”, in F. Cafaggi (ed.), *Reframing self-regulation in European private Law*, Kluwer Law International, London.

Sacconi L. (2010a), “A Rawlsian view of CSR and the Game Theory of its Implementation (Part II): Fairness and Equilibrium”, in L. Sacconi, M. Blair, R.E. Freeman and A.Vercelli (eds.) *Corporate Social Responsibility and Corporate Governance: The Contribution of Economic Theory and Related Disciplines*, Palgrave Macmillan, Basingstoke.

Sacconi L. (2011a),“A Rawlsian View of CSR and the Game Theory of its Implementation (Part III): Conformism and Equilibrium Selection” in *Social Capital, Corporate Social Responsibility, Economic Behavior and Performance*, edited by L. Sacconi and G. Degli Antoni , Palgrave Macmillan Basingstoke.

Sacconi L. (2011b), “Multistakeholder governance for effectively sharing social responsibility: social contracts, deliberative democracy and endogenous conformity” , this volume.

Sacconi L., Faillo M. (2010); “Conformity, Reciprocity and the Sense of Justice. How Social Contract-based Preferences and Beliefs Explain Norm Compliance: the Experimental Evidence” *Constitutional Political Economy*, Volume 21, N. 2 / June, pp.171-201

Searle J.R (2001), *Rationality in action*, The MIT press, Cambridge Mass.

Sen A.(1985), *Commodities and Capabilities*, North Holland.

Sen A. (1997), *On Economic Inequality. Expanded Edition*, Oxford (Clarendon Press),

Sen A. (2009) , *The idea if Justice*, Harvard U.P., Cambridge Mass.

Sugden R. (2000), "Team Preferences," *Economics and Philosophy*, 16, pp.175-204.

Sugden R. (2003), "The Logic of Team Reasoning," *Philosophical explorations*, 16, pp. 165-181.

Sunstein C. (2002), *Risks and reason,* Cambridge U.P., Cambridge.

Thaler R. and C. Sunstein (2008) *Nudge, Improving Decisions About Health, Wealth, and Happiness,* Yale University Press.

Thompson D. (1985), “The possibility of administrative ethics”, *Public Administration Review*, 45, 5, pp.555-561.

Tuomela R. (1995), *The Importance of Us: A Philosophical Study of Basic Social Notions,* Stanford University Press

Velasquez M.G. (1983), “Why Corporation are not Morally Responsible for What They Do”, *Business & Professional Ethics Journal,* 2 , Spring, pp.1-17