

Contractarianism and Personalisms in dialogue about the ethical foundation of CSR (comment on Helen Alford's "Stakeholder theory")

paper prepared for the international conference "The Good Company" ,
Rome, 4-7 October 2006

By Lorenzo Sacconi (Trento University and EconomEtica)

0. Introduction

In her insightful paper Helen Alford makes several interesting points over the current state of stakeholder theory and the contribution of a neo-Thomist and personalist perspective to the foundation of CSR. Let me split some of them into a *pars destruens* and a *pars construens*, not to give a complete reconstruction of her arguments, but just to focalize on what I'm going to discuss in this comment.

Pars destruens

I put aside Alford's criticism of the business case for CSR, on which we completely agree , and also the criticism of the enlightened self-interest and utilitarian accounts, on which we almost agree, in order to concentrate on her criticisms on what matter to me: individualistic ethics and the social contract approach to CSR.

a) according to Helen, every individualistic ethics, the contractarian included, being based on an inappropriate understanding of the very "nature" of the company, is unable to account for shared goods and common interests inside it. The contractual analogy , equated to Alchian-Demestz's idea of a *nexus of contracts*, is inappropriate to represent the company reality and unable to recognise its genuine nature as a "community".

b) Social contract theory is unable to provide an ethical foundation for CSR as far as it is incapable to suggest a sound principle for balancing stakeholders' conflicting claims, which is due to how they are seen within the contractarian perspective as autonomous individuals or groups with different interests, plans of life, different views of their personal good etc. This is so for a number of reasons:

b1) as far as stakeholders are autonomous groups of individuals it is impossible to find a basis for balancing their different interests without doing violence to their autonomy and rights;

b2) the social contract sees CSR (and also moral responsibility) only as *instrumental*, that is as a *mean* for achieving ends without any ethical content, given that our "ethical positions" (our ideas of the good) remain private and "out of the picture". The agreement is seen only as a useful *mean* for achieving further ends and not as

an intrinsic moral value *per se*. In other words the social contract sees *morality* as *instrumental*, not based on an *intrinsic* value – i.e. an idea of the good;

b3) since the social contract is just an hypothetical an abstract model of potential choice, it is incapable to provide a practical basis for balancing concrete claims and interests: why should we accept to balance according to an abstract hypothetical contract that we have not agreed in practice?

b4) the social contract does not provide a solid motivational basis for putting in practice CSR: if it is just instrumental to the private ends of stakeholders, what is its the qualitative difference or advantage with respect to more traditional claims like as those based on property rights, or other form of economic egoism? Why should the claim to CSR override them?

Pars construens

As an ethical foundation for CSR, Catholic Social Thought would have four advantages over its individualistic competitors:

- a) Personalist anthropology gives a fuller picture of human beings, based on the tension between the human as an individual and as a person, where the “person” is not only capable of but also naturally aimed to sociality and Relationality.
- b) As a specialisation of the Thomist idea of “common good”, the firm may be seen as a “community of work”, hence providing some shared value and goods, which underlies the constellation of stakeholders’ different and possibility conflicting interests.
- c) The two last ideas are connected by mean of how persons (as opposed to individuals) are related to the social “whole” - and the “small society” constituted by a company - within Maritain’s philosophy. This guarantees that the community of work is a moral community giving a moral foundation (according to a universalistic understanding of the “common good”) to CSR, while other communitarian accounts may fail in doing that (this is quite implicit, but I want to make it explicit here).
- d) This would be not only conceptually more satisfying, but also practically capable to better provide an explanation of the concrete motivational drives that push the company’s members to carry out CSR practices and organisational forms.

In what follows I will argue that the foregoing allegations against individualism in general, and the social contract approach to CSR in particular, are unwarranted. Secondly, even if I admit to have been fascinated by reading Maritain’s personalist account of what could be called the complexity of the Self’s motivations to act (which has so many prophetic analogies with many contemporary models in

behavioural economics – the mine included), I will question whether personalist anthropology by itself may keep the promise of an ethical grounding for CSR, without assuming the much less unquestionable Thomist metaphysical view of the “common good”, which in fact is inherent to Maritain’s view of the person freedom (with respect to any temporal societal organisation). At the same time, underwriting these assumptions - with their theological implications – seems unnecessary, given what we can reach by a proper contractarian understanding of the complexity of the Self’s motivation to act, which is also relevant for CSR implementation.

1. Is methodological individualism really incapable to account for shared interests and social interactions ?

In this and the second section of this comment I will defend individualism, not as an ontological thesis (“only individual are out there, not collectives”), but as methodological and ethical position.

To start with methodological individualism, let me ask: is it really incapable to account for shared interest, common ends and interactions? If we recognise that game theory is the basic theoretical approach within which the main individualist methodological explanations of social facts have been given over the last three decades, we must recognize that even if we restrain our discussion to the most traditional understanding of individualism, i.e. to the typical instrumental version of rationality, the answer cannot be positive. On the contrary game theory suggests us a number of models in which conflict of interests and common interests are interlocked and can be studied jointly. Non-cooperative mixed motives games (like the “battle of sexes”) are models of interactive situations where coordinating in general is a common interest that players share, but at the same time they may disagree about the particular coordination mode to be chosen, and hence they may hold conflicting preferences about what mode of coordination has to be selected among the many possible.

Cooperative games, on the other hand, offer the players the opportunity to join coalitions (even the total coalition of all of them) with super-additive characteristic function. That is, the value of the coalition (or the production function of a team), is the result of intrinsically common action and cooperation throughout the coalition, and cannot be reduced to the simple addition of individual separate productive efforts. At an ex ante stage then players are involved in a bargaining game in which they agree on the coalition structure in the light of how much each of them may gain from joining every possible coalition. The coalition, seen as having an intrinsically joint production function and a set of common plans of action at its disposal, becomes the object of an interdependent bargaining game by which players decide to enter or not coalitions according to their own personal objectives and their individual preferences.

No doubt this is an instrumental explanation of cooperation (and I will go beyond it by considering also *ethical* individualism), but nevertheless it is a perfectly understandable account of common action and common ends, reconciled with the *ex ante* individual decision to subscribe to a common plan of action generating genuine common goods (not replicable by separate production). The players definitively want to reach an agreement over a joint action plan, for they recognise that only by a joint action a surplus is attainable. Moreover, in the *ex ante* bargaining session they outguess one another decision, and they reciprocally adapt their decisions in order to reach an agreement on the coalition formation. They form reciprocal expectations on their relationships, so that they not only *interact* with one another, but they are also *aware* of their interaction, and assume some degree of *de facto* rational symmetry or equality among the interacting parties.

Institutions, coalitions, organisations, joint action plans are worthwhile for they provide unique sources of value, but ultimately they are means for maximizing individual utility. This might have a desirable ethical implication however of for there is no room here for idolatry of organisations, teams, companies or institutions. They may be “good” only in so far as they are conducive to individual well being or goal-attainment or preferences satisfaction, so that individuals remains free centres of decision. (I guess that personalists would be attracted by this picture of mundane organisations, seen as functional to persons retaining their freedom).

Not only a coalition (and its super-additive value) is a value *for* the individual members (and hence the coalition members are not mere means *for* the coalitional value), but also a consideration is given to self-enforceability of the coalition. This aspect carries us back to the non-cooperative branch of game theory, where the simple underwriting of an agreement is not sufficient to predict the players’ effective behaviour, i.e. their compliance with a joint action plan. It is needed moreover that the agreement be also *ex post* self-enforceable and this is not the same as to say that there are reason for agreeing on it *ex ante*. It means that a coalition needs incentives, motivation or reason to act that make the agreement self-supporting, which in turn implies that it cannot be simply enforced by an external authority. As a consequence some equilibrium concepts (mainly Nash equilibrium) is used, which means that players remain free to choose whether to comply or not with a given coalition structure or agreement. Therefore, it is possible to predict that the agreement will be put in practice only when their reciprocal, and reciprocally expected, actions supporting it will adapt one another as reciprocal best responses. Again, society cannot entirely overrule the individual, and he/she retains a fundamental freedom with respect to any institution or organisation.

Classical game theoretical models view decisions according to the typical instrumental model of economic rationality. Beliefs provide an epistemic support to an individual’s choice as far as they give a representation of the other players’ choice that enter the utility calculation only in so far as they serve as probability

weights attached to the utility of any outcome (according to the expected utility functional form). They are not direct subject matter for our preferences.

However in recent behavioural economic models, expectations also enter the scope of preference. Reciprocal expectations about one another's behaviour become a subject matter for the utility function, for they are an intrinsic source of preference and utility of the self (this approach is typically formalised in the branch of game theory called "psychological games" - see for example Rabin 1993). Since expectations refer to reciprocity in the players' behaviours in terms of mutual consistency to some norm or principle of payoffs distribution (like as minimizing direct inequality or maximizing kindness), they represent relationships amongst players (how much one reciprocates the other behaviour). And, since what we know about the individual's identity is represented by his/her utility function, individual identity (the individual's preferences) reflects social relationships amongst players.

Summing up, individual preference nowadays are more complex than in the traditional economic view, and a utility function reflects more than just *one* motive to act. It may include the desire for mutual conformity to some principle of justice, kindness or relative inequality etc. But at the end of the day the explanation of social artefacts, institutions and collective actions depends on individual preferences and rational (even if boundedly so) individual choices of the Self. This amounts to methodological individualism nevertheless, because we explain norms and institutions as solutions of games corresponding to equilibria reached through individual rational choices, even if based on players' enlarged preferences.

2. Does ethical individualism really attach only instrumental value to the social contract and CSR?

Let me consider now *ethical* individualism. It would be incorrect to state that the meaning of, and the consequent foundation of CSR on, the agreement among stakeholders (as categories of individuals) would be entirely "instrumental" – which means devoid of any ethical content. Ethical individualism attaches intrinsic ethical value to the *procedure* or the *rule of treatment* by which individual characteristics of moral value are treated. To say that according to any individualist ethics (as well as the contractarian approach) moral values are left outside the picture as personal private position, while the agreement is only an instrumental relationship functional to ends whose ethical content remain uncovered, simply implies to presuppose that all ethical values must be conflated into an idea of the "good". But of course we know that individualistic ethics may be neutral with respect any personal idea of the good, by being focused otherwise on

- a) *other* individual characteristics of *moral value* that are irreducible to an idea of the "good", like as for example autonomy, rights, duties - in addition to pleasure or pain, and wellbeing;

- b) the *moral* treatment of these characteristics (whoever may hold them) through some *impartial* procedure of collective decision is recognized as having *moral* value.

Generally speaking, in order to generate any ethical model, we must answer two questions: *first*, which *characteristics* related to a given practice, decision or institution *has moral value*? *Second*, what is the *moral treatment* of the characteristics we have recognized to have value, whoever might hold them?

To the first question the individualist will answer by quoting characteristics *held by individuals*. For example, “a decision must be taken by *autonomous* individuals”, “an institutions must respect *individuals’ rights*”, “a social practice must be based on the *individuals’ rational adhesion*”, “interpretation of a given professional role must discharge the *individual’s obligation* owed to other individuals”. Autonomy, rightfulness, rationality and obligations are characteristics of value that can be predicated as properties of individuals, and decisions, practices or institutions must satisfy them as predicates of individuals.

At the second step, however, what enters the picture is a qualification concerning the *mode of treatment* to which characteristics of value are subjected, *whoever* might hold them. That is a treatment that may be generalised and universalised with reference to all the individual who possess (at the same extent) the characteristic of value. What is needed is a specification of the general idea of *equality of treatment*, or impartiality of treatment, with reference to a particular value, whoever may possess or is affected by it. A contractarian, who answers the first question by saying “autonomy” or “rationality”, would answer the second by working out a procedure of impartial treatment, or equal treatment, fitting moral characteristics like as “rationality” or “autonomy”. Therefore, the appropriate (moral) treatment in this case is an *impartial agreement* amongst all the rational involved agents. In fact a “social contract”, by asking unanimous (i.e. equal) rational acceptance by all the autonomous agents, does *equally* respect the autonomy of *all* of them.

Hence, you may be correct in saying that a contractarian will see society and institutions as means for achieving the individual ends of those who agree the contract. But this would not be an appropriate account for what really carries normative force in the contractarian argument. Institutions or societies are not important *per se* because what *has moral value* is the individuals’ rationality and autonomy, and hence the universalistic treatment of this value consist of a rule of impartial treatment for rationality (or autonomy) – that is a general, unanimous agreement among rational choosers. Institutions, or the society itself, is “ethical” as it is rationally agreed upon by all the members by means of an impartial exercise of their rational decision and autonomy.

As you can see, the social contract (and its implication, as for example, CSR social norms), has intrinsic moral value as far as it is an impartial exercise of rationality by all the involved individuals. This is not a teleological account for ethics, of course,

but teleology is not the sole form of ethical justification. Were consequences (in terms of pleasure or pain) the characteristics of value involved by a practice, a decision or an institution, the answer on the second question could have been to treat them through the summation of pleasures and pains of all the involved patients *equally weighted* – that is utilitarianism.

As Thomas Nagel suggested some years ago, the request of equality has a sort of methodological primacy over any other moral consideration, since it is the general modality of moral treatment for the processing of whatever moral characteristics held by individuals. Were liberty the main substantive value (characteristic of value), then liberty should be treated equally with respect to whatever individual may be characterised in terms of liberty in relation to a given decision, practice, or institution. Equality of treatment (to be specified by some more elaborated rule of collective choice like an agreement, a convention or an additive social welfare function etc.) must be applied to whatever value. At a different level of reasoning, it replicates the role of “universalisability” of moral statements, as it is a general requirement for moral reasoning independently of the substantive characteristics of value that identifies a particular ethical theory.

Summing up, to come back to the instrumentality of the social contract, it not to pursue our private end that we recommend a social contract. It is on the contrary because we recognize moral value to the equal treatment of whichever individuals' autonomy/rationality.

Here is appropriate to remind the distinction between what I call the “context of justification” and the “context of implementation” of any ethical principle. As far as we are interested in justification, we play the game of moral reasoning and moral language, and hence we are committed to respect their rules. One of them is that whatever the substantive value we recognize to be involved in a given practice or decision, we must treat it equally whoever holds that characteristics. Lacking this condition, we could not be said to have given a justification of the practice or decision at hand. This amounts to making judgements that are able to retain their moral significance under the permutation of different concrete situations (where involved are different individuals) as far as the individual's characteristic of value (considered anonymously) remains unchanged.

A completely different matter is how we become able to carry out such a moral reasoning and to speak within moral language. Probably the answer is connected to the millenary evolution that fixed in our brain the cognitive capability to put ourselves into the shoes of others, and to grasp symmetries and invariance under the permutation of individual positions, point of views or roles. To be sure, it is more a matter connected to the emergence of a capacity for “empathy” than to “sympathy”, i.e. the cognitive ability to take the point of view of other persons and understanding their own evaluation according to their own terms, more than the capability to identify with their sentiments (which probably presupposes a certain commonality of feelings amongst the two parties – see Sugden 2005).

This capability is probably an evolutionary trait that has benefited the specie over a millenary evolutionary process, for it allowed members to win high payoffs if capable to do that under cooperation or coordination situations where they were selected at random to occupy different position or roles over times (Binmore 2005). However the explanation of *why* we possess these capabilities does not answer the question about *why* equal treatment justifies a given practice or institution. I would not say that an institution is “good” or “fair” because the rule to make this judgment shows high evolutionary fitness (unless I do not recognize moral value to the reproduction of the species).

3. A Social contract justification for CSR: an univocal, normative and operational basis for an alternative institutional governance structure

Now I will defend a particular version of ethical individualism in business ethics i.e. the social contract approach to CSR. My claim is that the social contract offers a sound basis for defining CSR as an institutional form of productive organization, and is able to give a solution the *vexata quaestio* of how balancing the stakeholders’ possibly (at least partly) conflicting interests. Consequently it amounts also to rejecting the allegation that, lacking a concept of common good, contractarianism is mute in front of the “stakeholders balancing problem”.

At first, let me remind what is the domain where this discussion is placed: we are in the domain of *justification* as distinguished from that of *implementation*. Justification is a context of rational acceptance from the point of view of universalizable judgements. Justification is neutral and impartial. From the contractarian standpoint justification coincides with the request that rational autonomous agents choose by a general agreement whether to enter or not any given arrangement of the society. It is a test of general ex ante acceptance that has to be carried out, *before* considering *ex post* stability of the arrangement, i.e. whether it is realistic to think that the agreement will be *complied* with. The content of this test is the *internal rationality* of an agreement, i.e. the rationality or not of choosing any given agreement, taking for granted that, if it will be chosen, it will be also put in practice (Gauthier 1986).

A theory of norms implementation takes a quite different point of view. In this perspective the problem is ex post compliance, or conformity, and the question concerns whether the agent is driven by any kind of motive to act in such a way that she/he will carry out a given justified principle (or the social contract). Reasons to act in this case are of many kind, not just only impartial and univesalisable judgments, expressed form a neutral point of view. As Nagel suggested, agent relative reasons to act are appropriate in this context (which may include both personal preferences and the sense of obligation for duties that the agent is asked to accomplish just because he/she is that particular decision maker in that particular position). In contractarian terms this is the problem of *external* rationality, not the rationality of choosing some contact amongst many other, but the *ex post* rationality to comply with it, allowing also that are possible diverging

incentive, like in a DP game (Gauthier 1986). Solving the implementation problem hence means to consider the full set of the players' reasons and motivations to act according to a norm.

According to the contractarian approach we may define CSR as an extended model of corporate governance such that who runs a firm (entrepreneurs, directors, managers) have responsibilities that range from the fulfilment of fiduciary duties towards the owners to the fulfilment of analogous fiduciary duties towards all the firm's stakeholders. Hence the scope of CSR is to extend the concept of fiduciary duty *from a mono-stakeholder* perspective (where the sole relevant stakeholder is the owner of the firm) *to a multi-stakeholder one* in which the firm owes fiduciary duties to *all* its stakeholders (the owners included).

However stakeholders interests may diverge, and also be at least partly contradictory, and hence the company's fiduciary duties may become inconsistent. The main question therefore is whether we are able to work out a consistent and acceptable principle of fair balancing of stakeholders legitimate claims.

Since fiduciary duties are based on fiduciary relations between those who run the firm and those (stakeholders) who delegate authority by accepting to submit to the governance structure of the firm, the best way to understand CSR as a set of enlarged fiduciary duties is to model it as based on a social contract amongst all the stakeholders by which they decide to institute that particular governance structure that we call the firm. In fact, the social contract typically is a model of rational acceptance which explains any delegation of authority as the rational decision on the part of those who submit to an authority to *trust* those who are in the position of authority. Notice that this is not a social contract *between* any company and the *society at large*, but *every* special, or local, social contracts amongst all the stakeholders - both because involved in transactions or touched by externalities - of *each* particular company.

Many will say that these social contracts are not real, but just hypothetical, and of course this is true for every social contract model. Its paradigm is an hypothetical contract from which we may deduce how the company would have been instituted by rational stakeholders, and at the same time it is a test able to check whether that arrangement of the firm could have been agreed upon, and hence justified, by all its stakeholders rationally. However, there is a factual basis for understanding relations amongst the firm and its stakeholders as a fiduciary (which refers to a sort of implicit promise or tacit agreement that the productive organisation will be conducive to the stakeholders best interest, and will restrain from damaging them). Since lot of contracts are incomplete, they leave when renegotiation occurs many *ex post* decision variables to the discretion of the strongest parties. In some cases contracts are explicitly "completed" by a delegation of authority that takes place by allocating to a party the residual right of control over every *ex ante* non contractible decision variables (economics says that this occurs by giving a party property right over the physical asset of the firm). Members of the organisation

(mainly employees but also non controlling shareholders or minority partners) explicitly accept an authority relation with those who are in a position of authority in the organization. Nevertheless, also those contractual parties who remain formally independent in a market contract, but are unprotected due to the incompleteness of contacts, and hence undergo the risk of opportunistic renegotiation, as far as they accept to stay the relation, they also show to trust the company who has the strongest position in the contact. They believe that it will not substantially abuse its *de facto* discretion over the *ex ante* un-contractible decision variables.

Let me summarize how the company would have been instituted by means a social contract, which serves also as a test for its moral justification. I have a two stage game account for that (see also Sacconi 2006)

The *first* stage is a collective choice over firm's constitutions modelled by a bargaining game amongst all the cooperating stakeholder. The second stage is a coalition game that the stakeholders play within the rules of the game (the constituted firm) selected at the first stage. This second step generates a final allocation of payoffs.

Being linked in a sequential way, the two games can be solved reasoning backward. Constitutions are seen a restrictions over the strategies available to players at the second stage game. Each second stage game has a solution in terms of payoff allocation, so that at the first stage a constitution may be selected according to the final allocation of payoffs associated to it. Being a bargaining game, the first stage game is solved by the most accredited solution concept for such games, i.e. *Nash bargaining solution*, which prescribes to maximize the product of the players' payoffs for agreements net of the status quo payoffs. Formally this coincides to a distribution of the surplus proportional to relative marginal variation of the players' utility. Under the additional assumption of interpersonal utility comparability this may be interpreted as a distribution proportional to a rough measure of their *relative needs* (how much *my* satisfaction positively changes at margin in relation to a negative marginal change of *your* utility due to an infinitesimal transfer of surplus from you to me).

The second stage game is a coalitional game played within a given institutional framework (the model of firm governance) that assigns each player certain rights or responsibilities. It must be solved in terms of a solution concept that allocates to each player an amount of utility related to his importance for each possible coalition. This carries us to a distributive principle based on proportionality to relative contribution. But notice that the institutional arrangement - a structure of rights and duties that influence the level of contribution of each player to each coalition - is chosen at 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.

Assume that the alternative arrangements of property rights institutionally feasible are only those that will generate a disproportionate imbalance of power and hence in the final distribution of payoffs amongst the players (stakeholders). Therefore, the two stage structure of our constitutional and post constitutional choice problem implies that from the constitutional perspective an institutional framework will be chosen only if under such arrangement those who are entitled with authority in the firm are constrained to carry out a utility transfer to the non controlling stakeholders such that the final distribution of surplus will approximate as much as possible the Nash bargaining solution. This solution roughly split the surplus in equal parts as far as the utility space, and the underlying space of joint actions, is symmetrical.

CSR follows from this model in a quite natural way: as an “extended” structure of fiduciary duties, CSR amounts the obligation to carry out these utility transfers in favour of the un-controlling stakeholders, in order to make sure that, whatever allocation of property rights is selected, the principle for a constitutional agreement will be respected. Hence, the constitution implies that stakeholders will gain from the company payoffs proportional to their own relative contributions/deserts. But what they are able to contribute, and hence the payoffs that they will deserve, is strictly related to their relative needs (i.e. property rights over assets, and responsibilities toward non controlling stakeholders are arranged so that stakeholders may deserve by contribution what they really need).

This theorizing reaches some results that completely reverse the evaluation of contractarianism given by Helen Alford about the distributive justice problem faced in balancing different stakeholders interests (for more details see Sacconi 2004). By the way, it also answers professor Jensen’s criticism that a multistakeholder approach to strategic management would make too complex and unmanageable the objective function of the firm.

- a) the objective function of the firm is univocally defined not as the maximisation of shareholder value, but as the maximisation of the Nash bargaining product of the stakeholders utilities, after having put at a minimum the negative externality over other non cooperating stakeholders;
- b) this objective function is explained as the one that stakeholders would have agreed in the position of an hypothetical contract in which they could have decided to start the firm as a cooperative venture to their mutual advantage;
- c) the objective function is genuinely normative and can be translated in a set of practical prescriptions, concerning the goals of the firm (see Sacconi 2004):
 - *First step*: minimize the negative externalities affecting stakeholders in the broad sense (perhaps by paying suitable compensation);

- *Second step*: identify the agreements compatible with the maximization of the joint surplus and its simultaneous fair distribution, as established by the impartial cooperative agreement among the stakeholders in the strict sense;
 - *Third step*: if more than one option is available in the above defined feasible set, choose the one that maximizes the *residual* allocated to the owner (for example, the shareholder).
- d) even if it is quite abstract (no more however than any other the piece of theoretical economics and ethics), it can be made operational in two ways: (i) first, under certain constraints stakeholders through a decentralised bargaining process would reach exactly an agreement on this specification of the objective function, given that *it is* the solution of a bargaining problem; (ii) second, it could be also operationalised through an explicitly ethical management procedure (the procedure for reaching a social contract among stakeholders by means of an ethical management or board – see Sacconi 2004), which can be specified as follow:
- *Force, fraud and manipulation must be set aside.*
 - *Each party comes to the bargaining table with only its capacity to contribute and its assessment of the utility of each agreement or non-agreement proposed (dispensing with any form of threat other than its possible refusal to agree).*
 - *The bargaining status quo must be set at a level such that each stakeholder results immune against the cost of its specific investments – that is, each stakeholder must obtain from the social contract at least reimbursement of the cost of the specific investment with which it has contributed to the surplus (otherwise the bargaining process would permit opportunistic exploitation of the counterparty's lock-in situation). The distribution of the surplus is regulated by the social contract – and by the corresponding deliberative procedure – on the basis of 'initial endowments' thus defined.*
 - *Each party in turn puts itself in the position of all the others, and in the position of each of them he can accept or reject the contractual alternatives proposed.*
 - *If solutions are found which are acceptable to some stakeholders but not to others, these solutions must be discarded and the procedure repeated (which reflects the assumption that cooperation by all stakeholders is recognized as necessary).*
 - *The terms of the agreement reached are therefore those that each stakeholder is willing to accept from its particular point of view: that is, the non-empty intersection of the joint strategies and relative distributions acceptable to each of them.*

4. The “relational” person within the social contract view on CSR: : ideal conformity and reciprocity

Emphasizing these achievements of the contractarian approach to CSR I do not want to hide under the carpet the main allegation raised by Helen Alford. Her main criticism points to the compliance problem and asks why stakeholders (and mainly shareholders) should conform by means of their concrete behaviour to a simply hypothetical, even though normative, social contract (see also Zamagni 2005) . I do not pretend to have solved this point so far, for the bargaining model is addressed to a different goal: to solve just the problem of internal rationality of the

contract, i.e. why in an ideal bargaining situation we should agree on such a governance structure of the firm that provides for extended fiduciary duties.

This question has to be squarely faced, for, as I already said, it is correct to distinguish between the contexts of justification and that of implementation, so that to implement a normative justified CSR principle or norm, we may need a set of reasons to act or motives for choice that goes further the basically neutral and impartial reasons that justify a given course of action or an institution. In this I follow Thomas Nagel distinction between the neutral point of view of justification and the agent-relative point of view that becomes relevant in the implementation context, when we ask ourselves about the complex set of personal reasons that we have for carrying out our personal actions. Notice that according Nagel, agent-relative reasons to act are not only reasons based on the agent's self-interest or individual well-being, but also, and perhaps mainly, reasons of deontology. In fact we realize to have a certain duty when we consider a decision under the perspective of being properly the actor who is specifically in charge, or in the position, to make that particular decision, whereas when we consider the consequences of a decision we may detach ourselves from the perspective of the specific agent and assess it according to the consequences to whatever person, also moral patients who are very far from being influent on the taking on that specific decision. Therefore, whereas deontological reasons are agent-relative, symmetrically, utilitarianism is a typical perspective generating neutral reasons to act.

My answer to the compliance problem is strictly consistent with this point of view, as it is based on the idea that compliance with the social contract follows from the complexity of human motivation represented by agents' overall utility function (the stakeholders but also managers, board of directors, entrepreneurs etc. who have direct influence over the company strategic conduct). This idea allows room for both material utility and ideal utility derived from conformity to an ethical ideal of justice, conditional on reciprocal and reciprocally expected conformity to the same ideal on the part of the other players. In this sense the second component is exactly a sort of *utility from deontology*, where deontological behaviour is seen as conformity to a duty or acting consistently with a given ideal.

In my view spontaneous adhesion to a CSR norm results basically from the *reputation mechanism* in a repeated trust game. But I also stress that reputation mechanism is weakened by many *fragilities*, both *cognitive* and *motivational* (see Sacconi 2004, 2006b) In particular, we should realize that a reputation of being partly unethical can also be supportive of an equilibrium behaviour (that is a combination of reciprocal best responses). For example, the reputation of abusing stakeholders' trust not all the time, but exactly up to the level at which stakeholders are indifferent between staying the relationship or exiting from it. In this sense, if there are only self-interested stakeholders and companies (even in enlightened sense) companies will abuse a significant portion of the time and stakeholders will surrender to such nearly opportunistic companies.

Therefore, the answer must be found in the complementarity between reputation games and a different social phenomena that I call “ideal-conformist preferences”.

The idea is that stakeholders (but also entrepreneurs) have *two* kind of preferences both able to motivate their action. On one hand (more basic), they describe the outcomes of their interaction as consequences and over consequences preferences are defined as *consequentialist*. These may be not only the typical self-interested preferences, but also altruistic preferences and also preference understood in less subjectivist terms than in standard economic theory. For example, inter-subjective preferences for “primary goods” that are necessary means for the achievement of any personal plan of life pursued by stakeholders, minimally significant to them.

But this part of the argument is not new at all. The new part of the argument concerns *conformist preferences*. Stakeholders (and also entrepreneurs) have also preferences defined over states of the world resulting from their interaction, which are described in terms of their consistency to an ideal - that is considering how far the agents actions (a state) is far from the set of actions that would put in practice an ethical ideal. By ethical ideal I mean a principle of justice for the distribution of material utilities.

Let assume that stakeholders have agreed upon a social contract concerning the principle of justice that should govern the distribution of the social surplus produced by the firm. Then conformist preferences may come into the picture. Intuitively speaking, a stakeholder will get intrinsic utility from the simple fact of complying with the principle, if he expects that this way he will be able of contributing in carrying out the distributive ideal, admitted that he expects that the other stakeholders will also contribute in carrying out the same principle, given their expectations.

A measure of ideal-conformist preferences is composed by three elements (see Grimalda and Sacconi 2005) :

First, a index of how much, given the other agents actions the action, the agent contributes to a fair distribution of material payoffs. This may also be put in terms of how much the agent is responsible for a fair distribution given what (he expects that) other players will do.

Second, a measure of how much the other stakeholders (or the company) is expected to contribute to a fair distribution, given what they expect form the first agent’s behaviour. This may also be put in terms of the (expected) responsibility of other players with respect to the generation of a fair allocation of the surplus given what they believe.

Third, a weight representing the agent’s psychological disposition to act on the motives of reciprocal conformity to an ideal.

Summing up, if a stakeholder expects that others are responsible for a fair distribution, given what they expect about his behaviour, and he also is responsible for a fair distribution, given the others' (expected) behaviour, then the motivational weight of conformity to an ideal (deontology) will enter the utility function. That is will enter effectively his system of preference, so that complying with the principle may gain additional utility (in psychological sense) with respect to the basic material payoff given by the same strategy.

This can change dramatically the calculation of the overall utility that induces the stakeholders to surrender or not in the face of a company's sophisticated opportunism that manage to keep compliance with the social contract at its minimum. In other words a stakeholder will refuse to surrender and will prefer a more resolute, hard-nose approach by quitting the relationship. At the same time if the company complies with the social contract, the stakeholder will gain additional utility by also complying with it, and this reinforces the overall preference for compliance.

On the other hand, a company with conformist preferences will add to its overall utility a component conditional on mutual compliance with the ideal. On all the occasions stakeholders also complies by entering the relationship, they both gain additional utility due to the fact that they are both responsible for achieving an higher level of ideal implementation. Instead, as far as most stakeholders are conformists, the typical self-interested far-sighted company will not gain substantially by adopting a sophisticated strategy of opportunism, for stakeholders do not accept to be exploited by it.

Notice that ideal-conformism does not equates itself with conformity to whatever behaviour may be socially shared by a group. Ideal-conformity here is seen as the desire to comply with an *agreed norm of justice*, given the expected reciprocal conformity by others to the *same* norm. If others' behaviour does not conform to the principle, the player will not gain any utility by replicating the others' behaviour. Hence it is essential for this definition of conformity the origin and the nature of the norm. It must not be whatever convention may emerge from social evolution, but exactly the norm that has moral content and is mutually believed to be complied with.

Social contract gives a basis for these norms of justice. The idea is that stakeholders and companies participate to a social dialogue by which they reach agreements on the principles of distributive justice lying behind the claims of CSR. Moreover by dialogue they set standards of behaviour and voluntary norms reflecting that principles. Social dialogue, under certain constraints of impartiality and honesty, simulates the hypothetical situation of agreement in which stakeholders (and entrepreneurs) may agree over the constitution of the firm. Since this gives some basis to the belief that they have *agreed*, and supports the shared ideology that companies have *subscribed* to the constitutional agreement concerning principles of justice regulating their relations to the stakeholders, then there is some basis to

emergence of conformist preferences (of course, a second condition is still needed: mutual beliefs of reciprocal conformity must be formed) .

Notice that I'm not saying that companies conform to the ideal (or to the ideological) agreement for it is their best material interest to do that. Instead, I'm taking for granted that it may be in the best interest of those who run the firm (at least in the short run) not to comply, or that there may be preferred equilibria in which the company incompletely complies with the norm. Nevertheless, the simple fact of an *ex ante* constitutional agreement, or moreover the simple fact of an *ex ante* ideology supporting the myth of an implicit social contract of the firm, allows the formation of conformist preferences, admitted that the beliefs that stakeholders and companies will conform also are formed. *Ex ante* norms and beliefs over reciprocal conformity enter the utility function as constituents of its ideal part. Therefore they create new motivations that may give additional support to the decision to comply with a CSR social norm.

In a sense, the model is quite akin to the Rawlsian idea of a "sense of justice": a *desire to conform* with those principles of justice that would have been accepted rationally in an hypothetical agreement. But it adds to that the more realistic hypothesis that the sense of justice (or the desire to be just) is not absolute and unconditional, but based upon an expectation of reciprocity in conformity. To comply "alone" does not command a practical motivation drive, but being aware to be able to get near to the ideal given the others' behaviour and believing that others are also doing their best to get near to the ideal (given what they believe about ourselves) introduces an additional motivational component within our preference system that can push us to comply with the hypothetical agreement. Therefore, it may be considered as a moderately Kantian (liberal) account for conformity, sensitive to the subtleties of the Humean understanding of the nature and formation of social norms.

Nevertheless, this theory accounts also for many suggestions that can be derived from the personalist thought of a Catholic thinker like Jaques Maritain. Maritain stresses that human being may be seen at the same time as an individual, and as a person. As an individual he/she entertains material needs and a merely instrumental relations with society, which is seen as a mean for obtaining relief from material needs. But as a person she/he is capable of sociality, relationality and spirituality. The person aims to higher order (ideal) goods and he wants to create social relationships with other people seen as members of a community similarly aimed to such spiritual higher order ends.

The model satisfies in a quite natural way both the part of Maritain's view of humanity. In correspondence to what he calls the "individual" I have material utility or consequentialist preferences, but correspondent to the "person's" higher order aims I introduce *ideal utility* derived from conformist preferences based on mutually expected reciprocity in complying with the ideal. Relationality is here introduced not in the sense that persons simply like relations, to stay together *per se*,

but in the sense of a desire to share with other agents the same behaviour of consistency with an ideal of justice. Hence they are pleased of expecting that other agents will reciprocate their consistency to the same ideal principle. I guess that a personalist would appreciate this aspect of the theory, since Maritain did not confine the person's desire of sociality to reciprocity within the limits of any concrete social group, but saw the person as aiming to an ideal community.

But there is nevertheless a big difference. In this approach conformist preferences, that is higher order desires of the "person", depend on ideal principles that are completely human and grounded on human autonomy - i.e. on the ideal of an impartial agreement. Therefore, I do not need any idea of a transcendent common good to ground conformist preferences for reciprocity, since on the contrary they are grounded in an ideal of justice, which is a human artefact. In fact, I only need a set of principles of justice that the individuals would have agreed in an hypothetical choice situation, which satisfy universalisability of the terms of agreement and their invariance with respect to the permutation of the individuals' point of view.

This perspective answers the question about the origin of principles - from where does they come from? – that we would like to see carried out by reciprocal conformity and would make us happy. They come from our ex ante moral choice. Whatever its weakness, the theory is therefore *self-contained*.

Moreover, it is able to account for many social behaviours and intuitions so relevant for the subject of CSR, but such that the theory of enlightened self-interest would disregard as purely irrational.

- a) First, it accounts for stakeholders' activism, like as socially responsible consumers or investors who give up some material advantage in order to punish opportunistic behaviour by companies that do not damage their material interests, and explain how this is connected to the observation that a company purports, may be deceitfully but nevertheless explicitly, its adherence to ethical principles (for example by a code of ethics, social reports etc.) ;
- b) Second, it accounts for the possibility of *social enterprises* ad non profit organisation, involved not just in advocacy of rights or social causes, but in the delivery of welfare goods; and explain how the internal members of these enterprises escape from the egoistic incentive to appropriate the entire surplus, not so much for they are legally submitted to the "non distribution constraint", but because of the ideology of employees and entrepreneurs, with or without religious backgrounds.
- c) Third, it moreover accounts for the intuition, quite shared, that companies effectively engaged in CSR are those where top managers in some sense seriously believe it, and not those who comply just because of an instrumental strategy of reputation gaining given the traditional objective function (maximizing the shareholder value).

In fact, what this model allows us to predict is that from an evolutionary point of view this kind of enlightened self-interested companies will be superseded from both the sides (compliance and non compliance) by two alternative types of firms. On the one hand, by conformist companies, who gain additional utility from satisfaction of their conformist preferences by interacting positively with conformist stakeholders. On the other hand, by sophisticated opportunist firms, who will take advantage of the presence of non conformist and simply consequentialist stakeholders, ready to surrender to sophisticated opportunism on the part of the firms.

This is not a consolatory view point, however. It may result clear by considering what will happen if in the population of companies and stakeholders, interacting one another by random matching, it will be observed that most stakeholders have not strong enough conformist motivations. This can occur for in the population the basic disposition to be motivated by reciprocal conformity to an ideal is too weak, or because its members do not trust each other enough about their reciprocal compliance with a given principle (even if they recognize it as rational and justified). Therefore the evolutionary equilibrium where to be a conformist company pays more in terms of overall utility (ideal utility included) will not emerge, and we may predict that the population will evolve toward an equilibrium state where most companies are soft opportunist and most stakeholders surrender to them.

This possibility however suggests very strong policy implications from the perspective of an institution design specifically addressed to support spontaneous compliance with CSR social norms. Assumed a normal level of psychological disposition to conform, the basic priority for a CSR *policy* should be to favour the institutionalisation of social dialog (mimicking the social contract) over CSR social norms ad standard of behaviour. And, on the other side, to support the building up of middle level social bodies, institutions and social agencies with the primary goals to verify and spread amongst stakeholders information and knowledge about companies' compliance with CSR principles and standards. These two roles are crucial to improve conformist preferences within firms and stakeholders. By social dialog, first, they generate an approximation to the constitutional agreement and give the basis for sharing an ideology. Second, by promoting verification and spreading information they generate the informational basis for reciprocal beliefs that are a necessary condition for activating conformist preferences.

5. Is the Thomistic alternative really a better foundation for CSR? The inescapable need for too demanding metaphysical assumptions

It is quite obvious that the version of motivational complexity given before may seem not completely satisfying to the true follower of Maritain's philosophy. Would the neo-Thomist alternative work better?

Let me remind that one basic tenet of Helen Alford's proposal is the distinction between the individual and the person as both necessary aspects of the human being, who at the same time has both material and spiritual aspirations and needs. The person tends to sociality, relationality and communication as inherent parts of the "common good", which must be disentangled from the instrumental need for social relations, as means for gaining material goods and services that characterizes the individual. Now ask: what guarantees that the search for relationality and sociality will not be confined to closed and narrow communities, possibly focused on maximising their internal friendship and bonding social capital, but inimical to strangers (imagine a strong company culture entirely built around the myth of its "supermen and superwomen" people, but adversarial to any other company, external stakeholders like clients and supplier, and the institutions representing public interests)? More generally what will guarantee that the good of the particular community (for example a specific company) will be really connected to the *common* good?

This problem is clearly seen by Maritain in his book "the man and the common good" (Maritain 1983), where he advocated the constraints of justice over the pursuing of the common purposes of particular communities, like social groups, States, nations (we may add also companies). But this problem also infects most of the neo-Aristotelian and communitarian virtues ethics theories, especially when applied to the realm of business - as in the case of Solomon's application of MacIntyre's virtue ethics to business ethics. In this application the company is an identifying community based on its typical productive goal. Virtues within a company are equated to conformity to role-requirements, which are associated to the positions that each person occupies with the division of labour. These roles are in fact functional to the pursuing of excellence in that given goal. This implies that virtues coincide with the proper accomplishment of the demands of each role within the hierarchical structure of the firm. The individual is functionalized to the internal finalism of the firm, seen as a community with its own understanding of its excellence standard from where virtues descend.

In this context happiness (as eudemonia) depends on the conformity of personal character to virtues, which in turn is intrinsically related to the inherent standard of excellence of the practice as understood within its communitarian tradition. In fact the tradition of that activity sets its standard of excellence, i.e. the end at which the practice is ordered and necessarily tends to achieve as part of the constitutive definition of the practice itself. You cannot say to practise medicine – so to say – without acting according to the ideal of excellence defining the good physician and his discipline according the tradition of medicine. Therefore the company is a community, or perhaps an industry is a community, endowed with an intrinsic goal and an ideal of excellence, set by its own tradition for the given productive practice. Any activity functional to excellence in gaining that end defines a virtue, and in a system of rational division of labour it is captured by a role. Hence,

everybody will be happy, in eudemonian sense, as far as they discharge the demand of their roles.

But what is the end of the company in our traditional understanding? Someone could say that, at least in the Anglo-Saxon part of western tradition, the end of the corporation is to make profit and its excellence is maximizing shareholder value. Hence, managers and employees will be happy to discharge the demands of their roles functional to this end. As you can see, it is not so difficult to construe within virtues ethics an apologetical argument for the traditional view of capitalist enterprise. Virtues ethics complements Milton Friedman, *pace* our concern for CSR. (see also Miller, this conference).

At the end, it seems not so strange that Solomon himself says that in front of virtues ethics the same (individualistic) concept of stakeholders results outdated. In fact the company doesn't need to undertake fiduciary duties owed to its stakeholders. The stakeholders, instead, will reach the bliss of happiness by discharging their role-relative obligations toward the company, as this is the way to shape their character consistently to virtues. Hence they may give up their stakes against the firm, their real, intrinsic satisfaction does not depend on that.

Beyond this aspect, virtues ethics faces the well known problem of dissenting members of the community, or the strangers. How to manage those members of a community (or stranger) that start to not recognize as really worthwhile the standard of excellence received from the tradition? Must they be ostracized?

I have to admit that Helen Alford's proposal, as it is based on personalism and neo-Thomism, and not on simple virtue ethics as such, is able to escape this difficulty by resorting to an idea of common good that is transcendent with respect to the good of any particular community or particular social group – the firm or any industry included. The second basic tenet of this position, together with the duality individual-person put forward by Maritain, is that common good is the inherent end to which human nature (and the society) is ordered, i.e. necessarily aims to.

To make sense of this idea let me try to reconstruct how the argument consistently works. It is based on two sets of assumptions.

The major premise is a metaphysical hypothesis, concerning the “common good” as inherent end to which the human being is necessarily ordered, which derived from S. Thomas' synthesis between Aristotle's theory of causality and Christian theology. All probably remember how S. Thomas reconnected God to each form of Aristotle's causality, but let me summarize.

a) At the beginning of the chain connecting each “movements in act” to each “movements in power”, that explains every movement, cannot be but an “immovable mover”, and it cannot be but God;

b) similarly, God must be the first efficient uncaused cause, i.e. the “stopping rule” of the infinite regress from one efficient cause to another;

c) given that qualities have different degrees of perfection, there must be a maximal perfection to each quality from which any other degree descends by corruption, and this perfection cannot be but God;

d) causality has finalistic structure: causes are ends inherent to each nature (they are “final causes”), but the primitive “final cause”, the only that must be uncaused to stop the regress, cannot be but an intelligent designer endowed with a will capable of assigning the end to each nature, and he cannot be but God. However, God cannot have but an absolutely good will. Hence, he cannot want but the good, so that the inherent end, or final cause of every nature, to which whatever is ordered, must be the universal good. It follows that human nature aims necessarily to the good, and is ordered to it in so far as the good is the inherent end and final cause of humanity. Notice that this is a metaphysical concept of good: it is a transcendent good identifiable with the reign of God, or, as also Maritain says, *beatitude*.

Also any single person is therefore ordered to the good, since he cannot aspire but to the good. This is simply analytical, since any will wants what it understands as a good. Even the modern economist, sitting on the side of a subjectivist theory of value, could subscribe this statement. In fact preferences underlie decisions, and they can be reduced to a formal relation of *betterness* that may include any reason that the decision maker may have for preferring a given course of action over another. *Betterness* is just a formal binary relation, without any specific content (Broome 2000). Therefore to say that the human being aims to the good is simply tautological, as far as “good” is simply what he *subjectively* assesses as *better* than the alternative. However according to S. Thomas the man may be mistaken about the *real* good: only the good to which he is *objectively* ordered is the true good, and this exists independently on him being aware of it.

Second, the anthropological premise. The person is the relational and spiritual part of human being, and as such she aims to establish rich human relations with other humans, through communication. As far as such relations are based on love, they are not instrumental but of intrinsic value *per se*, and therefore this relationality induces human beings to enter communities aimed to the good of their members.

But what guarantees that these communities have moral character? What makes us sure that they will not be condemned to parochialism? It is the assumption that the human being is naturally ordered to a *transcendent* good, and therefore any community created by persons must be in a sense a “good” community. The metaphysical hypothesis here encounters and supports the anthropological one. This is very clear in Maritain when he says that the society to which the person basically is ordered is the perfect community of the Saints and the Blessed, which may take place only in the Reign of Heavens. Hence the good to which humans are ordered is a transcendent common good which is identical with a perfect society,

where he will get to *beatitude*. In our mundane life we may reach only intermediate degrees of approximation to this perfect community, so that our aiming to the good cannot ever be satisfied into mundane, social or business life. This sets the basis for the relationship between the human being and society, and between human freedom and the human's obligations to any concrete form of human society within which the individual and the person are parts (Maritain talks about political society, but we may extend the reasoning to the smaller scale society of the company or business community).

The individual aims to society (and the firm) due to his material needs, but in order to receive the benefit of society he must accept his subordination to the social whole. Individuals are functionalised to the community, and therefore far from being free, they are dependent on the social organisation (as in the firm hierarchy). There is not freedom or independence in the way the individual is related to society. The correct image must be taken from biology: a beehive or an ant-heap.

But, on the other side, the person remains free under any kind of temporal and historical society. The reason is that, since he/she aims to a transcendent good, he/she cannot be satisfied by any concrete temporal community, and hence exceeds any temporal social "whole", like as the family, the company, the state, the nation etc. In fact he/she is ordered to a wider and more perfect community, which allows also the plain development of his spiritual powers. Any concrete particular community hence is no more than a mean by which the person tries to satisfy these higher aspirations. To tell it in Rawlsian terms, concrete communities are *primary good* required to the pursuing of such an ideal good of the person. These concrete community are subordinated to the person as he/she uses them as a means for his/her higher aspirations, and admitted that this ideal community cannot be put in practice within our temporal society, the person retains always his freedom and remains unquiet at a certain extent under each these concrete societies. This is a basis for a concept of *negative* liberty.

This tendency of the person to go beyond any concrete society or group and to enlarge and enrich the sphere of her/his relationships toward the model of a perfect society, produces in practice effects quite similar to the requirements of the contractarian model. To stay at the company level, stakeholders are not tools for the company, employed in order to reach some goal or excellence, but on the contrary it is the company that serves as an instrument the aspirations of stakeholders, seen as persons, to help them reaching an ideal that goes further the company itself (in this sense I understand Helen Alford's statement that the company is a community of work in which each member should be helped to pursue the full development of her/his human potential).

The difficult point of course is about where this independence of persons is grounded. As it should be clear, all rests on the condition that an *end* or *final cause* of human nature and society does exist, to which persons are *necessarily ordered*. But this is a *metaphysical* premise to which we must be committed from outside any

social, economic or philosophical theory, so that these cannot be self-contained. Assumptions like as “final cause”, “inherent end”, and “immovable mover” (not properly accepted by contemporary natural sciences), and the assumption that an intelligent designer puts into the world, as its necessary finalism, the tendency to a transcendent common good, are presuppositions that in all honesty one could not ask his interlocutor to take for granted *tacitly*.

Dispensing with these metaphysical presuppositions (that I deem too demanding), personalist anthropology - as much insightful as it may be by emphasizing relationality, sociality and communication - however does not escape the risk of confining the person within narrow communities. Here I see a serious risk. If the direct link between the anthropological assumption and the metaphysical one is severed, then the personalist analysis of the company would be left with just the “biological” analogy of society such that individuals are functionalised to the need of the “whole” – an idea that Maritain himself finds a bit humiliating.

This cannot happen within a social contract view of the firm, since according to Kantians (see Bowie 1999) the firm is seen as a mean functional to the “reign of ends” constituted by all the firm’s stakeholders: “You (the manager) shall never treat a neighbour (a stakeholder) as just a mere mean (for the company), but always also as an end in herself”. Contractarianism ensures that the individual is able to rationally agree upon a constitution of the firm that reflects the value of autonomy and rationality of choice. It also allows stakeholders to develop conformist ideal preferences that represents the immaterial part of their system of motivations. This component of the system of preference reflects the desire for relationality and reciprocity and the ambition to an ideal, admitted it has been at least hypothetically agreed.

In short, it is from “mere” individuality, shaped by the moral exercise of unanimous rational choice under the “veil of ignorance” or any other conditions of impartiality, that emerges the “person” seen as the human capability to carry out moral reasoning and to hold immaterial desires for reciprocity, relations and conformity to an ideal of justice.

I have already said that my worry is that the neo-Thomist perspective, once the direct connection between the *transcendent* “common good” and the *concrete particular* good is cut out, may open the route to a reduction of the company to a community endowed with a finalism on its own, but without any substantial respect for stakeholders’ rights and claims of justice, as they are seen as “extrinsic” to the proper finalism of the firm. Hence, I perfectly understand the catholic believer who wants to foster the idea of person by maintaining the metaphysical premise that allows him to see the person as naturally ordered to a transcendent universal common good. Simply I think that he should also recognize that this intellectual strategy presumes an act of faith. Neither it can be deduced by purely rational steps of reasoning, nor be based on a stringent logic of the “excluded third”, or by

any kind of “reduction ad absurdum” that would demonstrate that in absence of that presumption we could not predict what is obvious to our experience.

As far as the intellectual requirements of simplicity and general intelligibility are accepted as rules of the game, we may indeed fare better by dispensing ourselves with such hypotheses. In fact, alternatives are available metaphysically much less demanding (especially in a dialog over a penultimate matter as, admittedly, CSR continues to be), which enable us to explain the same phenomena of human motivation, like as the search for relationality, reciprocity and the ambition to ideal values. You can just start from the simple exercise of human impartial choice and rational agreement on the part of the company’s stakeholders, within which any belief and also any view of the “common good” may find appropriate room as the view of one particular stakeholder. Consequently we don’t need that internal link between anthropology and metaphysics (even if we may like it as consistent with our faith) in order to explain the causal force of motivations that push individuals to act according an ideal of justice or to advocate an ideal of just company.

Of course it is true that all behavioural economics models of motivational complexity make some important assumptions about the fabric of human psychology. When the strategic conditions for a positive conformist preference, or social preference, or a preference for reciprocal kindness are satisfied (often assuming that the appropriate beliefs have formed), then these models predict that some basic disposition to act (or to prefer) is activated, which is represented by a motivational weight attached to these variables.

Is this the place where the natural disposition to the common good is represented? I would say no. There is no reason to say that this dispositional weight is an “universal constant” across any individual and any culture, and I take that it cannot be activated without the reaching of an agreement and the formation of the appropriate beliefs, which are both social artefact. Moreover, rather than a natural tendency to the common good, it may be interpreted as the weight of the psychological disposition to be just (according to the Kantian primacy of the “right” over the “good”), which in fact cannot be activated without an *ex ante* fair agreement over a principle of justice and the corresponding system of mutual expectations.

Summing up, I see the possibility of a convergence of our dialog toward a view of CSR inspired by how, at the end of his philosophical career, John Rawls saw justice as “political not metaphysical”. I guess that, for the conduction of a rational dialog about CSR, the idea of a general metaphysical final cause identified with the transcendent common good, should be put within parenthesis (as one “complete theory of the good” which is valid for those who accept it). Then the search for a “good” which is “common” because it is acceptable by all, and hence not metaphysical but practical, becomes substantially identical to the effort to discover the *overlapping consensus* amongst all the interests, the identities and the complete

views of good upheld by different persons or (in the case of a company) by all the different stakeholders.

By submitting hypotheses about this practical “common good” to the exercise of universalisation of our judgment and by putting one into the shoes of the others, we (the stakeholders) will be able to find out a set of basic goods produced and allocated by companies, such that the claims (interests, needs and aspirations) for them are invariant with respect to the permutation of every personal stakeholder standpoint within the company and around it. Then, perhaps, by the same invariance-seeking device, we will be able to *agree* over a set of distributive principles for these goods. And finally will be also able to develop our ideal motivation supporting our reciprocal conformity to the agreed ideal principles for the company conduct.

5. Some practical limitations of the idea of “company as a community of work”.

Let me conclude with a remark on the company as an institution, between economics and the law. We have seen that without the effectiveness of a metaphysical theory of the common good, the same personalist approach is not able to prevent the risk of parochialism, narrow communities with narrow, maybe collusive, interests amongst the members. How this risk may affect the good company? I see a possible, even if quite mild, effect also on the idea of the company as a “community of work” upheld by Helen Halford.

The economic theory of *team production*, and the model of a super-additive characteristic function of the coalition formed by productive members of the team, is certainly a natural counterpart of the idea of a community of work. Moreover there is not antagonism between social contract approach and the community of works approach about this aspect. The surplus is the fruit of genuine cooperation and individual contributions are not separable (to be fair it should be also recognized the contribution of the *nexus of contract* theorist to this idea – see Alchian and Demsetz 1972 -, even though their model of the firm clashes with CSR; however more recent team production theorist elaborate a model of corporate governance that is very akin to my definition of CSR, see Blair and Stout 1999). Therefore distribution may be only the result of a cooperative agreement on the *common* surplus .

This implies - contractarians and personalists do agree on that - rejection of libertarian models, upholding that every allocation of values amongst the company’s members can be made on the basis of just negative rights, property right in particular. And at most on the basis of the voluntary exchange between property rights over assets, input etc., so that the negative rights of any party are not violated. But keeping to this view we simply cannot account for the distribution of a cooperative *surplus*, which typically is not covered by negative rights that at most may fix the base-line *form where* the purely cooperative *surplus* (the

“common good”) is measured, so that no one can be entitled to part of it just because of a negative right (for a stakeholder libertarian theory however see Freeman E. and R. Phillips 2002).

Consider nevertheless that also “communities of work” (in empirical sense) may suffer from opportunistic behaviour and collusion. This is true for production Co-ops (that may act opportunistically toward the consumers), and also to the productive non-profits. Therefore, also to these organisations, that already are legally recognized as “community of work”, CSR is recommended as a kind of “external mutuality”, extending the cooperative treatment to all the stakeholders of the firm.

The social contract approach provides a natural collective decision procedure capable to include, albeit hypothetically, all the stakeholders who are in a potential relation of mutual advantage with the firm, even if they are just external stakeholders, like as consumers and clients, suppliers, the local communities hosting the plants, till future generations interested in the positive but not the negative (environmental) externality of the company activity, etc.

Moreover the social contract approach complemented by the idea of conformist ideal preference is also able to explain how are possible *social enterprises* that follow apparently altruistic goals. They restrain them-selves from appropriating all the surplus, even though they could in principle resort to this form of opportunist behaviour. The universalistic approach implicit in the social contract allows to include in the ex ante contract both the internal members of the organisation, entitled in practice to decide, and the external stakeholders, possibly weak decision makers, or beneficiaries. Conformist preferences developed around a principle of justice, make possible equilibrium solutions where internal members do not abuse and leave substantial part of the surplus to be used for improving the beneficiaries’ well-being.

Otherwise, my worry is that the communitarian approach, and also the relational-personalist approach, once the link to a transcendent common good is severed, may carry us to a view of the “community of work” seen as closed corporatist institution, entrenched into the defence of just the interests and the “bonding ties” amongst the producers (entrepreneurs, managers, employees, partners, maybe the owner if he is a natural person, but not when ownership is dispersed throughout a large number of anonymous individuals having impersonal relationship to the firm by means of financial market institutions), and giving few or no weight to the claims from external stakeholders, like the consumers - who cannot be immediately reduced to internal members of a community of work.

This is not at all a necessary conclusion of the idea, but it is a possible risk if the common good of the members of a community of work are not submitted to the exercise of universalisation that enlarges the view beyond the company walls toward external stakeholders. Exactly as Maritain suggested, also in our little scale

society we need to “transcend” the small community of work to look for ideal “relations” with external stakeholders, who are eligible only for sporadic, impersonal, not so warm concrete relationships. Hence, my invitation to discover the “overlapping consensus” amongst all the company stakeholders apply here all the more so.

References

- Alchian A. & Demsetz H. (1972) "Production, Information costs and economic Organization", 62, *Amer. Economic Review*, pp.777-795.
- Alford H. (2006), *Stakeholder theory*, presented at “The Good Company” conference , mimeo
- Alford H. and Shcherbinina Y. (2006) , *Towards a Thomistic personalist Foundation for CSR*, mimeo
- Alford H. Sena, B. and Shcherbinina Y. (2005) *Strengthening the ethical basis of Csr: a personalist input form Catholic Social Thought (CST)*, mimeo
- Binmore K. (2005), *Natural Justice*, Oxford UP, Oxford
- Blair. M and L. Stout (1999) “A Team Production Theory of Corporate Law”, *Virginia Law Review*, Vol. 85, No. 2
- Bowie M. (1999), *Business Ethics: a Kantian Perspective*, Oxford, Basil Blackwell.
- Broome J. (1999), *Ethics out of Economics*, Cambridge, Cambridge UP
- Freeman R.E., R. Phillips (2002), “Stakeholder Theory; A Libertarian Defense”, *Business Ethics quarterly*, vol. 13, 3, pp.331-349
- Gauthier D. (1986), *Morals by Agreement*, Oxford, Clarendon Press.
- Grimalda G. and Sacconi L. (2005) “The Constitution of the Not-For-Profit Organisation: Reciprocal Conformity to Morality” *Constitutional Political Economy* Vol. 16(3), September, 2005, pp.249-276
- Jensen M.C. (2001) “Value Maximization, Stakeholder Theory, and the Corporate Objective Function” *Journal of Applied Corporate Finance*, Vol. 14, No 3, Fall.
- Keeley M. (1988), *A Social-Contract Theory of Organization*, University of Notre Dame, Indiana.
- Mariatin J. (1983), *La persona e il bene comune*, 7a ed., Morcelliana, Brescia.
- Maritain J.(1947), *Umanesimo Integrale*, 2a ed., Editrice Studium, Roma.
- McIntyre, 1981, *After Virtues*, Univ. of Notre Dame, Indiana.
- Miller R. T. (2006), *Maximizing Shareholder Value in the Good Company*, mimeo (this conference)
- Nagel T. (1986), *A view from nowhere*, Oxford University press, Oxford
- Rabin, M. (1993): “Incorporating Fairness into Game Theory”, *American Economic Review*, Vol. 83, N. 5, pp. 1281-1302
- Rawls J. (1993), *Political Liberalism*, Columbia U.P. New York.
- Sacconi L. (2000): *The Social Contract of the Firm*, Springer Verlag, Berlin.
- Sacconi L. (2006), “A Social Contract Account For CSR as Extended Model of Corporate Governance (I): Rational Bargaining and Justification” in *Journal of Business Ethics*, Special Issue on “Social Contract Theories in Business Ethics” (in print)

- Sacconi L. (2006), “A Social Contract Account for CSR as Extended Model of Corporate Governance (II): Compliance, Reputation and Reciprocity” in *Journal of Business Ethics*, (forthcoming)
- Sacconi L. (2004) “CSR as a model of extended corporate governance, an explanation based on the economic theory of social contract, reputation and reciprocal conformism”, Liuc paper .142, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=514522, now in F. Cafaggi (ed.), *Profiles of self-regulation*, Kluwer Law, 2006 (in print)
- Solomon R. (1991) *Ethics and Excellence, Cooperation and Integrity in Business*, Oxford University Press,
- Sugden R. (2005), “Correspondence of sentiment: An Explanation of the pleasure of social interaction”, in L. Bruni and P.L. Porta (eds.) *Economics & Happiness*, Oxford, Oxford U.P.
- Tommaso d’Aquino, *De Veritate* (trad. it. testo latino a fronte, Bompiani, 2005)
- Zamagni S. (2005), “Critica delle critiche alla CSR ed il suo ancoraggio etico” in L. Sacconi (a cura di), *Guida critica alla responsabilità sociale e al governo di impresa*, pp.319-336, Bancaria editrice, Roma