

Contractarianism and Personalisms in a dialogue on the ethical foundation of CSR - comment on Helen Alford's "Stakeholder theory"

Lorenzo Sacconi

Introduction

In her insightful paper Helen Alford makes several interesting points concerning the current state of stakeholder theory and the contribution of a neo-Thomist and personalist perspective to the foundation of CSR. Let me divide some of them between a *pars destruens* and a *pars construens*, not to give a complete reconstruction of her argument but just to focus on what I shall discuss in this comment.

Pars destruens

I set aside Alford's criticism of the business case for CSR, on which we completely agree, and also her criticism of the enlightened self-interest and utilitarian accounts, on which we almost agree, so that I can concentrate on her criticisms on what is of most concern to me: individualistic ethics and the social contract approach to CSR.

1. according to Helen, because all individualistic ethics, contractarian included, are based on an inappropriate understanding of the "nature" of the company, they are unable to account for shared goods and common interests within it. The contractual analogy, equated to Alchian-Demestz's idea of a *nexus of contracts*, is an inappropriate representation of corporate reality and fails to recognise its genuine nature as a "community".

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

2.1. Insofar 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;

2.2. The social contract sees CSR (and also moral responsibility) only as *instrumental*, that is, as a *means* to achieve 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 *means* to achieve further ends, not as an intrinsic moral value *per se*. In other words, the social contract sees *morality* as *instrumental*, not as based on an *intrinsic* value – i.e. an idea of the good;

2.3. Since the social contract is merely a hypothetical and abstract model of potential choice, it is unable to act as 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 to in practice?

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

Pars construens

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

I shall argue that the foregoing charges against individualism in general, and the social contract approach to CSR in particular, are unwarranted

1. 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 at, sociality and relationality.

2. As a specialisation of the Thomist idea of “common good”, the firm may be seen as a “community of work”, and hence as providing shared values and goods which underlie the constellation of stakeholders’ different and possibly conflicting interests.

3. The two last ideas are connected by how persons (as opposed to individuals) are related to the social “whole” - and the “small society” constituted by a company - in 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 to do so (this is implicit, but I want to make it explicit here).

4. This would be not only conceptually more satisfying, but also practically able to provide a better explanation of the concrete motivational drives that induce a company’s members to carry out CSR practices and organisational forms.

In what follows I shall argue that the foregoing charges against individualism in general, and the social contract approach to CSR in particular, are unwarranted. Secondly, even if I admit to having been fascinated by Maritain’s personalist account of what could be called the complexity of the Self’s motivations to act (which has so many prophetic analogies to contemporary models in behavioural economics – mine included), I shall question whether personalist anthropology can by itself keep the promise of giving an ethical grounding to CSR, without assuming the much more questionable Thomist metaphysical view of the “common good”, which in fact is inherent to Maritain’s view of person freedom (with respect to any temporal societal organisation). At the same time, endorsing these assumptions – with their theological implications – seems unnecessary, given what we can reach through a proper contractarian understanding of the complexity of the Self’s motivation to act, which is also relevant to CSR implementation.

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

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

I begin with methodological individualism. Let me ask: is this really unable to account for shared interests, common ends and interactions? If we recognise that game theory is the basic theoretical approach whereby the main individualist methodological explanations of social facts have been put forward over the last three decades, we must also recognize that even if we restrict 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 a number of models in which conflicts 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 is in the common interest of the players, but at the same time they may disagree on the particular coordination mode to be chosen, and hence may have conflicting preferences about what mode of coordination should be selected among the many possible.

Cooperative games, on the other hand, give the players opportunities to join coalitions (even the total coalition of all of them) with a super-additive characteristic function. That is to say, 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, the players are involved in a bargaining game in which they agree on the coalition structure in light of how much each of them will 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 whether or not to enter coalitions according to their personal objectives and their individual preferences.

No doubt this is an instrumental explanation of cooperation (and I will go beyond it by considering *ethical* individualism as well). 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 indeed want to reach agreement on a joint action plan, for they recognise that only by joint action will a surplus be attainable. Moreover, in the *ex ante* bargaining session they outguess each other's decisions, and they reciprocally adapt their decisions in order to reach agreement on the coalition formation. They form reciprocal expectations on their relationships, so that they not only *interact* with each another but are also *aware* of their interaction, and some degree of *de facto* rational symmetry or equality arises among the interacting parties.

Institutions, coalitions, organisations, joint action plans are worthwhile because they are unique sources of value. But ultimately they are *means* to maximize individual utility. This may have a desirable ethical implication, however, for there is no room here for *idolatry* of organisations, teams, companies or institutions. They can be "good" only in so far as they are conducive to individual well being or goal-attainment or preferences satisfaction, so that individuals are still free centres of decision. (I presume that personalists would be attracted by this depiction of mundane organisations as functional to the retention of freedom).

Not only is a coalition (and its super-additive value) a value *for* the individual members (hence the coalition members are not mere means *for* the coalitional value), but consideration is also made of the coalition's self-enforceability. This aspect takes us back to the non-cooperative branch of game theory, where simple endorsement of an agreement is not sufficient to predict the players' actual behaviour, i.e. their compliance with a joint action plan. Also necessary is that the agreement be *ex post* self-enforceable, and this is not the same as saying that there are reasons for agreeing on it *ex*

ante. It instead means that a coalition needs incentives, motivations or reasons to act that render the agreement self-supporting, which in turn implies that it cannot be simply enforced by an external authority. As a consequence, some equilibrium concepts (i.e. mainly Nash equilibrium) are used, which means that players are free to choose whether or not to comply with a given coalition structure or agreement. It is therefore possible to predict that the agreement will be implemented only when the players' reciprocal, and reciprocally expected, actions supporting it adapt to each other as reciprocal best responses. Again, society cannot entirely overrule the individual, and the latter retains 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 epistemic support for an individual's choice in that they provide representations of the other players' choices, which 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 the direct subject matter of our preferences.

However, in recent behavioural economic models, expectations also enter the domain of preference functions. Reciprocal expectations about behaviour enter the utility function because they are an intrinsic source of preference and utility for the Self (this approach is typically formalised in the branch of game theory called "psychological games" - see e.g. Rabin 1993). Since expectations concern reciprocity in the players' behaviours in terms of mutual compliance with some norm or principle of payoff distribution (such as minimizing direct inequality or maximizing kindness), they represent relationships amongst players (the extent to which one reciprocates the other's 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 preferences are today seen as more complex than they were in the

Classical game
theoretical models
view decisions
according to the
typical instrumental
model of economic
rationality

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

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

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

1. *other* individual characteristics of *moral value* that are irreducible to an idea of the "good", such as autonomy, rights, duties, as well as pleasure or pain, and wellbeing;

2. the moral treatment of these characteristics (whoever possesses them) through some impartial procedure of collective decision-making is recognized as having moral value.

Generally speaking, in order to generate any ethical model, we must answer two questions: *first*, which *characteristics* affected by or affecting a given practice, decision or institution *have moral value*? *Second*, what is the *moral treatment* of the

characteristics that we have recognized to have value, whoever may possess them?

The individualist will answer the first question by citing characteristics *possessed 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 *individuals' rational adherence* to it", "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 a second stage, however, qualification is made concerning the *mode of treatment* to which characteristics of value are subjected, *whoever* may possess them. This treatment may be generalised and universalised with reference to all the individuals possessing (to the same extent) the characteristic of value. What is needed is specification of the general idea of *equality of treatment*, or impartiality of treatment, with reference to a particular value, whoever may possess or be affected by it. A contractarian, who would answer the first question by saying "autonomy" or "rationality", would answer the second by formulating a procedure of impartial treatment, or equal treatment, which fitted moral characteristics like "rationality" or "autonomy". Therefore, the appropriate (moral) treatment in this case is an *impartial agreement* amongst all the rational agents involved. In fact, because a "social contract" requires unanimous (i.e. equal) rational acceptance by all the autonomous agents, it *equally* respects the autonomy of *all* of them.

Hence, it may be correct to say that a contractarian will see society and institutions as means to achieve the individual ends of those who agree to the contract. But this would not be an appropriate account of what really carries normative force in the contractarian argument. Institutions or societies are not important *per se*, because what *has moral value* is individuals' rationality and autonomy, so that the universalistic treatment of this value consists of a rule of impartial treatment for rationality (or autonomy) — that is, a general, unanimous agreement among rational choosers. Institutions, or society itself, is "ethical" because it is rationally agreed upon by all

the members through impartial exercise of their rational decision and autonomy.

Hence, the social contract (and its implications: for example, CSR social norms), has intrinsic moral value in so far as it is an impartial exercise of rationality by all the individuals involved. This is not a teleological account of ethics, of course, but teleology is not the only form of ethical justification. If consequences (in terms of pleasure or pain) were the characteristics of value involved by a practice, decision or institution, the answer to the second question would be to treat them in terms of the sum of the *equally weighted* pleasures and pains of all the parties involved – that is, utilitarianism.

As Thomas Nagel suggested some years ago, the requirement 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 that individuals possess. If liberty were the main substantive value (characteristic of value), liberty should be treated equally with respect to however individuals 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 elaborate 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 “universalisability” of moral statements, because it is a general requirement for moral reasoning independently of the substantive characteristics of value that identify a particular ethical theory.

Summing up, to return to the instrumentality of the social contract, it is not to pursue our private ends that we recommend a social contract. It is on the contrary because we grant moral value to the equal treatment of whichever individuals’ autonomy/rationality.

It is appropriate here to recall the distinction between what I call the “context of justification” and the “context of implementation”

of any ethical principle. In so far as we are interested in justification, we play the game of moral reasoning and moral language. We are therefore committed to respecting their rules. One of these rules is that whatever the substantive value we recognize as involved in a given practice or decision, we must treat it equally whoever possesses that characteristic. In the absence of this condition, we cannot be said to have given justification for the practice or decision at hand. This amounts to making judgements able to retain their moral significance amid the permutation of different concrete situations (where different individuals are involved) in so far as the individual’s characteristic of value (considered anonymously) remains unchanged.

A completely different matter is how we become able to perform this moral reasoning and to speak with moral language

A completely different matter is how we become able to perform this moral reasoning and to speak with moral language. The answer is probably connected to the millenary evolution that has given us the cognitive ability to put ourselves in the shoes of others, and to grasp symmetries and invariances amid the permutation of individual positions, points of view or roles. To be sure, this is a matter concerning the emergence of a capacity for “empathy” rather than “sympathy”, i.e. the cognitive ability to take the point of view of others and understand their evaluations in their own terms, more than the ability to identify with their emotions (which probably presupposes a certain commonality of feeling between the two parties – see Sugden 2005).

This ability is probably an evolutionary trait that has benefited the species in a millenary evolutionary process, for it has enabled members to win high payoffs if able to do so in cooperation or coordination situations where they have been selected at random to occupy different positions or roles over time (Binmore 2005). However, the explanation of *why* we possess these capabilities does not answer the question of *why* equal treatment justifies a given practice or institution. I would not say that an institution is “good” or “fair” because the rule for making this judgment shows high evolutionary fitness (unless I do not grant 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*

I will now defend a particular version of ethical individualism in business ethics. I refer to 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 resolve the *vexata quaestio* of how stakeholders' possibly (at least partly) conflicting interests can be balanced. Consequently, it also amounts to rejecting the allegation that, lacking a concept of common good, contractarianism has nothing to say about the "stakeholders balancing problem".

First, let me state the domain in which this discussion is located: we are in the domain of *justification* as distinct from *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 requirement that rational autonomous agents choose by general agreement whether or not to enter any given societal arrangement. It is a test of general *ex ante* acceptance that must be performed *before* considering the *ex post* stability of the arrangement, i.e. whether it is realistic to believe that the agreement will be *complied* with. The content of this *ex ante* test is the *internal rationality* of an agreement, i.e. the rationality or otherwise of choosing any given agreement, taking for granted that, if it is chosen, it will also be put in practice (Gauthier 1986).

A theory of norms implementation takes a quite different point of view. From this perspective, the problem is *ex post* compliance, or conformity, and the question is whether the agent is driven by any kind of motive to act in such a way that s/he will carry out a given justified principle (or the social contract). Reasons to act in this case are of many kinds, not just impartial and universalisable judgments expressed from 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 concerning duties that the agent is required to accomplish just because s/he 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 contract amongst many others, but the *ex post* rationality of complying with it, also given possibly diverging incentives, as in a DP game (Gauthier 1986). Solving the implementation problem, therefore, means considering the full set of the players' reasons and motivations to act in accordance to a norm.

According to the contractarian approach we may define CSR as an extended model of corporate governance such that those who run 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 purpose 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 they may also be at least partly contradictory, so that the company's fiduciary duties may become inconsistent. The main question therefore is whether we can work out a consistent and acceptable principle for the 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 agreeing to submit to the firm's governance structure, the best way to understand CSR as a set of enlarged fiduciary duties is to model it as being based on a social contract amongst all the stakeholders whereby they decide to institute that particular governance structure which we call the firm. In fact, the social contract is typically 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 in the position of authority. Note that what I'm elaborating on is not a social contract *between* a company and *society at large*, but *every* special, or local, social contract amongst all the stakeholders - both those involved in transactions and those affected by externalities - of *each* particular company.

Many will say that these social contracts are not real but only 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. 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 between 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 interests, and will refrain from damaging them). Since many contracts are incomplete, they leave many *ex post* decision variables to the discretion of the strongest parties when renegotiation occurs. 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 variable (economics says that this occurs by giving a party property rights over the physical assets of the firm). Members of the organisation (mainly employees but also non controlling shareholders or minority partners) explicitly accept an authority relation with those in a position of authority in the organization. Nevertheless, also those contractual parties who remain formally independent in a market contract, but are unprotected owing to the incompleteness of contacts, and hence are at risk of opportunistic renegotiation because they accept the relation, also trust the party with the strongest position in the contract. They believe that the latter will not substantially abuse its *de facto* discretion over the *ex ante* un-contractible decision variables. Hence in order to understand the hierarchical structure of the firm, we need to see it as based on a network of fiduciary relationships, and to account for these relationships in terms of some set of reasons that all the non controlling stakeholders must have in order to accept the authority of those who run the firm. My claim is that the better account for these trust-based authority relationships is the social contract amongst the stakeholders of the firm.

Let me summarize how the company would

have been instituted by means of a social contract, which serves also as a test for its moral justification. I give a two stage game account of it (see also Sacconi 2006). The *first* stage is a collective choice on the firm's constitution modelled by a bargaining game amongst all the cooperating stakeholders. 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.

The *first* stage is a collective choice on the firm's constitution modelled by a bargaining game amongst all the cooperating stakeholders

Because they are linked sequentially, the two games can be solved by reasoning backwards. Constitutions are seen as restrictions on the strategies available to the players in the second stage game. Each second stage game has a solution in terms of payoff allocation, so that in the first stage a constitution may be selected according to the final allocation of payoffs associated with it. Because it is a bargaining game, the first stage game is solved by the most accredited solution concept for such games, i.e. the *Nash bargaining solution*, which prescribes maximizing the product of the players' payoffs for agreements net of the 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 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 brings us to a distributive principle based on proportionality to relative contribution. But note that the institutional arrangement - a structure of

rights and duties that influences the level of each player's contribution 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.

Let me now introduce the legally understandable that what as to be chosen at the first step is an arrangements of the ownership structure over the physical assets of the firm. Assume that the legally feasible alternative arrangements of ownership are only exclusive private property rights, that is those ownership arrangements 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 this arrangement those with ownership-based authority in the firm are constrained to carry out a utility transfer to the non controlling non-owner stakeholders such that the final distribution of surplus will approximate the Nash bargaining solution as closely as possible. This solution roughly splits the surplus into equal parts insofar as the utility space, and the underlying space of joint actions, are symmetrical.

CSR follows from this model quite naturally: as an "extended" structure of fiduciary duties, CSR amounts to the obligation to perform these utility transfers in favour of the non-controlling stakeholders in order to ensure 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 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 towards non controlling stakeholders are arranged so that stakeholders may deserve by contribution what they really need).

This theorizing yields results that completely reverse the evaluation of contractarianism given by Helen Alford in regard to 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 multi stakeholder approach to strategic management would make the objective function of the firm too complex and unmanageable.

1. 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 set the negative externality on other non cooperating stakeholders at a minimum;

2. This objective function is the one to which stakeholders would have agreed in the case of a hypothetical contract in which they could have decided to start up the firm as a cooperative venture to their mutual advantage;

3. The objective function is genuinely normative and can be translated into a set of practical prescriptions concerning the firm's goals (see Sacconi 2004):

3.1. First step: minimize the negative externalities affecting stakeholders in the broad sense (perhaps by paying suitable compensation);

3.2. 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;

3.3. 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).

4. Even if this is rather abstract (but no more than any other piece of theoretical economics and ethics), it can be made operational in two ways: (I) first, under certain constraints, stakeholders would reach through a decentralised bargaining process exactly an agreement on this specification of the objective function, given that it is the solution of a bargaining problem; (II) second, it could also be 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 follows:

4.1. Force, fraud and manipulation must be

set aside.

4.2. 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).

4.3. The bargaining status quo must be set at a level such that each stakeholder is immune to 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.

4.4 Each party in turn puts itself in the position of all the others, and in the position of each of them it can accept or reject the contractual alternatives proposed.

4.5. 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).

4.6. 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

On emphasizing these achievements of the contractarian approach to CSR, I do not want to brush under the carpet the main allegation raised by Helen Alford. Her principal criticism concerns 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 claim to have faced this point thus far, for the bargaining

model was addressed to a different goal: solving only the problem of the contract's internal rationality, i.e. why in an ideal bargaining situation we should agree on a governance structure of the firm that provides for extended fiduciary duties.

This question has nevertheless to be squarely faced, for, as I have already said, it is correct to distinguish between the contexts of justification and that of implementation, so that in order to implement a normative justified CSR principle or norm, we may need a set of reasons to act, or motives for choice that go further than the basically neutral and impartial reasons that justify a given course of action or an institution. In this I follow Thomas Nagel's 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 performing our personal actions. Note 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 that we have a certain duty when we consider a decision from the perspective of being the actor 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 specific agent's perspective and assess it according to its consequences for

whatever person, also moral patients very far from influencing the taking of that specific decision. Therefore, whereas deontological reasons are agent-relative, utilitarianism symmetrically amounts undertaking detached perspectives generating neutral reasons to act.

My answer to the compliance problem is strictly consistent with this point of view, because 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 on the company strategic conduct). This idea allows room for both material utility and ideal

**In my view,
spontaneous
compliance with a
CSR norm results
basically
from the *reputation
mechanism* in a
repeated trust game**

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 (ideal) component is a sort of *utility from deontology*, where deontological behaviour is seen as fulfilling a duty or acting consistently with a given ideal.

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

Therefore, the answer must lie in the complementarity between reputation games and different social phenomena that I call "ideal-conformist preferences". The idea is that stakeholders (but also entrepreneurs) have *two* kinds of preferences, both able to motivate their action. On one hand (more basic), they describe the outcomes of their interaction as consequences, and their preferences regarding consequences are defined as *consequentialist*. These may be not only the typical self-interested preferences, but also altruistic preferences, and also preferences understood in less subjectivist terms than in standard economic theory: for example, inter-subjective preferences for "primary goods" that – as Rawls taught us – are necessary means to achieve any personal life plan minimally significant to the stakeholders pursuing them.

But this part of the argument is not new at all. The new part concerns *conformist preferences*. Stakeholders (and also entrepreneurs) also have preferences defined over states of the world resulting from their interaction, which are described in terms of their consistency with an ideal – that is, considering how far the agents' actions (a state) is

from the set of actions that would fulfil an ethical ideal. By ethical ideal I mean a principle of justice for the distribution of material utilities.

Let us 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. Conformist preferences may now enter the picture. Intuitively speaking, a stakeholder will gain intrinsic utility from the simple fact of complying with the principle, if the same stakeholder expects that in this way she/he will be able to contribute to fulfilling the distributive ideal, admitted that she/he expects the other stakeholders also contribute to fulfilling the same principle, given their expectations.

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

First, an index of how much, given the other agents' expected actions, the agent by her/his action 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) are expected to contribute to a fair distribution, given what they (are expected to) expect from 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 (are believed to) 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/her behaviour, and he/she also is responsible for a fair distribution, given the others' (expected) behaviour, then the motivational weight of conformity with an ideal (deontology) will enter the utility function. That is, it will effectively enter the stakeholder's system of preference, so that complying with the principle may yield additional utility (in psychological sense) with respect to the basic material payoff given by the same strategy.

This may dramatically change the calculation of the overall utility that induces the stakeholders to surrender or otherwise to a company's sophisticated opportunism which manages to keep compliance with the social contract to its minimum. In other words, a stakeholder will refuse to surrender and will prefer a more resolute, hard-nosed 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.

So far for the ideal-conformist stakeholder, but what about the possibility of an ideal-conformist enterprise? A company with conformist preferences will add to its overall utility a component conditional on mutual compliance with the ideal. On all the occasions that stakeholders also comply by entering the relationship, they both gain additional utility due to the fact that they are both responsible for achieving a higher level of ideal implementation. Instead, because most stakeholders are conformists, the typical self-interested far-sighted company will not substantially gain by adopting a sophisticated opportunism strategy, for stakeholders will refuse to be exploited by it.

Note that ideal-conformist does not equate with conformism to whatever behaviour may be socially shared by a group. Ideal-conformity is seen here 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 from replicating the others' behaviour. Hence, essential for this definition of conformity is the origin and the nature of the norm. It cannot be whatever convention may emerge from social evolution, but exactly the norm that has moral content and is mutually believed to be complied with.

The social contract provides a basis for these norms of justice. The idea is that stakeholders and companies participate in a social dialogue by which they reach agreements on the principles of distributive justice underpinning the claims of CSR. Moreover, by means of dialogue they set standards of behaviour and voluntary norms reflecting those

principles. Social dialogue, under certain constraints of impartiality and honesty, simulates the hypothetical situation of an agreement whereby stakeholders (and entrepreneurs) may agree on the firm's constitution. Since this gives some basis for the belief that they have *agreed on the company constitution*, and supports a shared ideology such that the company has *subscribed* to a constitutional agreement on principles of justice regulating its relations with stakeholders, there is some basis for the emergence of conformist preferences. (Of course, a second condition is still necessary - mutual beliefs of reciprocal conformity must be formed.)

Note that I am not saying that companies conform to the ideal (or ideological) agreement because it is in their best material interest to do so. Instead, I am taking it for granted that it may be in

the best material 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, provided that the beliefs that stakeholders and companies will conform are formed as well. *Ex ante* norms and beliefs concerning 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 akin to the Rawlsian idea of a "sense of justice": a *desire to conform* to those principles of justice that would have been accepted rationally in a hypothetical agreement. But it adds 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 generate a practical motivation drive, but being aware that one can approach the ideal given the others' behaviour, and believing that others are also doing their best to approach the ideal (given what they believe about ourselves) introduces an additional motivational component

The social contract provides a basis for these norms of justice

into our preference system that may induce us to comply with the hypothetical agreement. Therefore, it may be considered as a moderately Kantian (liberal) account of conformity, sensitive to the subtleties of the Humean notion of the nature and formation of social norms.

Though genuinely contractarian, this theory also accounts for some suggestions that can be derived from the personalist thought of a Catholic thinker like Jaques Maritain. Maritain (1983) stresses that human beings may be seen at both individuals and persons. As individuals they have material needs and a merely instrumental relation with society, which is seen as a means to obtain relief from material needs. But as persons they are capable of sociality, relationality and spirituality. Persons aim at higher order (ideal) goods and they want to create social relationships with other persons seen as members of a community similarly seeking to achieve such spiritual higher order ends.

My model satisfies quite naturally both parts of Maritain's view of humanity. In correspondence to what he calls the "individual" I envisage material utility or consequentialist preferences, but in correspondence to the "person's" higher order aims I introduce an *ideal utility* derived from conformist preferences based on mutually expected reciprocity in complying with the ideal. Relationality is introduced here 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 complying with an ideal of justice. Hence they are likely to expect that other agents will reciprocate their compliance with the same ideal principle. I presume that a personalist would appreciate this aspect of the theory, since Maritain did not confine the person's desire for sociality to reciprocity within the bounds of any concrete social group; rather, he saw the person as aspiring to an ideal community.

But there is nevertheless a major difference. On this approach, conformist preferences – that is, the higher order desires of the "person" – depend on ideal principles that are entirely 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 in order to ground conformist preferences for reciprocity, because 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 to which the individuals would have agreed in a hypothetical

choice situation and which satisfy universalisability of the terms of agreement and their invariance with respect to the permutation of the individuals' points of view.

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

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

1. First, it accounts for stakeholders' activism, such as that by socially responsible consumers or investors who forgo some material advantage in order to punish opportunistic behaviour by companies that do not damage their material interests, and it explains how this is connected to the observation that a company claims, maybe deceitfully but nevertheless explicitly, its adherence to ethical principles (for example by a code of ethics, social reports etc.);

2. Second, it accounts for the possibility of *social enterprises* and non profit organisations involved not just in the advocacy of rights or social causes but also in the delivery of welfare goods; and it explains how the internal members of these enterprises are immune to the egoistic incentive to appropriate the entire surplus, not so much because they are legally subject to the "non distribution constraint" but because of the ideology of employees and entrepreneurs, with or without religious beliefs.

3. Third, it accounts for the widely shared intuition that companies effectively engaged in CSR are those whose top managers in some sense seriously believe in it, and not those who comply merely because of an instrumental strategy of reputation gaining, given the traditional objective function (maximizing 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 company will be superseded on both sides (compliance and non compliance) by two alternative types of firm. On the one hand, by conformist companies, which gain additional utility from satisfaction of their

conformist preferences by interacting positively with conformist stakeholders. On the other hand, by sophisticated opportunist firms, which 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 can be made clear by considering what could happen in the population of companies and stakeholders interacting with each other through random matching in the case it was observed that most stakeholders do not have sufficiently strong conformist motivations. This may occur because the basic disposition in the population to be motivated by reciprocal conformity to an ideal is too weak, or because its members do not place sufficient trust in each other concerning their reciprocal compliance with a given principle (even if they recognize it as rational and justified). In case these are the initial conditions in the evolutionary game under consideration, the evolutionary equilibrium whereby being a conformist company pays more in terms of overall utility (ideal utility included) will not arise, and we may predict that the population will evolve towards an equilibrium state where most companies are soft opportunists and most stakeholders surrender to them. In other words, there are many possible evolutionary equilibria, and which of them will emerge depends on initial conditions.

This possibility, however, has very strong policy implications from the perspective of an institution design specifically intended to support spontaneous compliance with CSR social norms. Assuming a normal level of psychological disposition to conform, the main priority for a CSR *policy* should be to favour the institutionalisation of social dialogue (mimicking the social contract) on CSR social norms and standards of behaviour. And, on the other side, to support the building of middle level social bodies, institutions and social agencies with the primary goal of verifying and spreading amongst stakeholders information and knowledge about companies' compliance with CSR principles and standards. These two tasks are crucial for improving conformist preferences within firms and stakeholders. Through social dialogue,

first, they generate an approximation to the constitutional agreement and lay the bases for sharing an ideology. Second, by promoting verification and spreading information they generate the informational bases for the reciprocal beliefs that are a necessary condition for activating conformist preferences.

5. Is the Thomist 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 set out above may seem not entirely satisfactory to the true follower of Maritain's philosophy. Would the neo-Thomist alternative work better?

Let me recall that a basic tenet of Helen Alford's proposal is Maritain's 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 to gain material goods and services that characterize the individual. Now we may ask: what is it 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", but adversarial to any other company, external stakeholders like clients and suppliers, 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 was clearly seen by Maritain in his book *The Man and the Common Good* (Maritain 1983), where he advocated the constraints of justice over the pursuit of the common purposes of particular communities, like social groups, states,

There are many possible evolutionary equilibria, and which of them will emerge depends on initial conditions

nations (we may also add companies). But this problem also affects 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 with role-requirements, which are associated with the positions that each person occupies within the division of labour. These roles are in fact functional to the pursuit 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 whence virtues descend.

In this context, happiness (as eudemonia) depends on the conformity of personal character with 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 to which the practice is directed and necessarily seeks to achieve as part of the constitutive definition of the practice itself. One cannot claim to practise medicine - so to speak - without acting in accordance with the ideal of excellence defining the good physician and his/her 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 achieving that end defines a virtue, and in a system with a rational division of labour it is captured by a role. Hence, everybody will be happy, in the eudemonian sense, in so far as they discharge the demands of their roles.

But what is the end of the company in our traditional understanding? Some might say that, at least in the English-speaking part of the Western tradition, the end of the corporation is to make profits, 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 one can see, it is not difficult to detect 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).

In the end, it seems not so strange that Solomon himself says that in face of virtues ethics the (individualistic) concept itself of stakeholder seems outdated. In fact, the company does not need to perform the fiduciary duties owed to its stakeholders. The stakeholders, instead, will achieve happiness by discharging their role-relative obligations to the company, as this is the way to shape their character consistently with virtues. Hence they may relinquish their stakes in (or claims against) the firm, for their real, intrinsic satisfaction does not depend on these.

Besides this aspect, virtues ethics faces the well known problem of dissenting members of the community, or the strangers. How should a community manage those of its members (or strangers) that begin not to recognize as really worthwhile the standard of excellence received from the tradition? Must they be ostracized? The contractarian would answer to renegotiate the term of the social contract looking for the mutual advantage of both the standard members and the dissenting or the new comer, or to seek the overlapping consensus of their views of the good. But what the answer of the virtue-ethics communitarian?

I must admit that Helen Alford's proposal, because it is based on personalism and neo-Thomism, and not on simple virtue ethics as such,

is able to evade this difficulty by resorting to an idea of a 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 individual-person duality put forward by Maritain, is that the common good is the inherent end to which human nature (and the society) is "ordered", i.e. necessarily seeks to achieve.

To clarify this idea let me reconstruct how the argument works. It is based on two sets of assumptions.

First, the major premise is a metaphysical hypothesis concerning the "common good" as an inherent end to which the human being is necessarily directed and which derives from St

Thomas' synthesis between Aristotle's theory of causality and Christian theology. It will probably be recalled that St Thomas linked God to each form of Aristotle's causality. But let me summarize.

1. At the beginning of the chain connecting each "movement in act" to each "movement in power" that explains every movement cannot but be an "immovable mover", and this cannot but be God;

2. Likewise, God must be the first efficient uncaused cause, i.e. the "stopping rule" of the infinite regress from one efficient cause to another;

3. Given that qualities have different degrees of perfection, there must be a maximum perfection of each quality from which any other degree descends by corruption; and this perfection cannot but be God;

4. Causality has finalistic structure: causes are ends inherent to each nature (they are "final causes"), but the primitive "final cause", the only one that must be uncaused to halt the regress, cannot but be an intelligent designer endowed with a will capable of assigning the end to each nature, and this cannot but be God. However, God cannot but have 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 directed, must be the universal good. It follows that human nature necessarily aims at the good, and is "ordered" to it (that is in plain English "directed" to it) because the good is the inherent end and final cause of humanity. Note that this is a metaphysical concept of good: it is a transcendent good identifiable with the Kingdom of God, or, as also Maritain says, *beatitude*.

Also any single person is therefore "ordered" to the good, since s/he cannot but aspire to the good. This is simply analytical, since any will wants what it understands as a good. Even the modern economist, taking the side of a subjectivist theory of value, would subscribe to 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, saying that the human being aims at the good is simply tautological, because "good" is simply what s/he *subjectively* assesses as *better* than the alternative. However, according to St Thomas man may be mistaken about the *real* good: only the good to which he is *objectively* "ordered" (directed) is the true good, and this exists independently of his being aware of it.

Second, the anthropological premise. The person is the relational and spiritual part of the human being, and as such it seeks to establish rich human relations with other humans through communication. Because such relations are based on love, they are not instrumental, but have intrinsic value *per se*. This relationality therefore induces human beings to join communities directed to the good of their members.

**But what guarantees
that these
communities have
moral character?**

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, so that any community created by persons must in a sense be a "good" community. Here the metaphysical hypothesis encounters and supports the anthropological one. This is very clear in Maritain when he says that the society to which the person is basically "ordered" is the perfect community of the Saints and the Blessed, which can only exist in the Kingdom of Heaven. Hence, the good to which humans are "ordered" is a transcendent common good which is identical with a perfect society where they will achieve *beatitude*. In our mundane lives we may reach only intermediate degrees of approximation to this perfect community, so that our pursuit of the good can never be satisfied in mundane social or business life. This sets the basis for the relationship between the human being and society, and between human freedom and human obligations to any concrete form of human society of 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).

Individuals aim at society (and the firm) because of their material needs, but in order to

receive the benefits of society they must accept 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 no freedom or independence in the way that the individual is related to society. The correct image must be taken from biology: a beehive or an ant-heap.

But, on the other hand, “persons” remain free under any kind of temporal and historical society. The reason is that, since they aim at a transcendent good, they cannot be satisfied by any concrete temporal community, and hence exceed any temporal social “whole”, like the family, the company, the state, the nation, etc. In fact, they are “ordered” to a wider and more perfect community, which also allows for the plain development of their spiritual powers. Any particular concrete community is therefore no more than a means by which the person seeks to satisfy these higher aspirations. Put in Rawlsian terms, concrete communities are *primary goods* required for the pursuit of such an ideal good of the person. These concrete communities are subordinated to the person because he/she uses them as a means to achieve his/her higher aspirations, and given that this ideal community cannot be put in practice within our temporal society, the person always retains his/her freedom and is restless to a certain extent in each of 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 towards the model of a perfect society, in practice produces effects quite similar to the requirements of the contractarian model. At the company level, stakeholders are not tools for the company, employed in order to achieve some goal or excellence; on the contrary, it is the company that serves as an instrument for the aspirations of stakeholders, seen as persons, and helps them reach an ideal that extends beyond the company itself (it is in this sense that 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 where this independence of persons is grounded. As should be clear, everything rests on the condition that an *end*

or *final cause* of human nature and society does exist, to which persons are *necessarily directed*. But this is a *metaphysical* premise to which we must be committed regardless of any evidence reported by social, economic or political theory, so that these cannot be self-contained. Assumptions like “final cause”, “inherent end”, and “immovable mover” (not properly accepted by the contemporary natural sciences), and the assumption that an intelligent designer puts the tendency to a transcendent common good into the fabric of world as its necessary finalism, are presuppositions that, in all honesty, one could not ask one’s interlocutor to take for granted *tacitly*.

However, dispensing with these metaphysical presuppositions (that I deem too demanding), personalist anthropology- as insightful as it may be by emphasizing relationality, sociality and communication - does not avert 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 will be left only with the “biological” analogy of society, so that individuals are functionalised to the need of the “whole”- an idea that Maritain himself finds somewhat humiliating.

This cannot happen on a social contract view of the firm, since, according to Kantians (see Bowie 1999), the firm is a means 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 means (for the company), but always also as an end in itself”. Contractarianism ensures that the individual is able to agree rationally upon a constitution of the firm that reflects the values of autonomy and rationality of choice. It also allows stakeholders to develop conformist ideal preferences that represent the immaterial part of their system of motivations. This component of the preferences system reflects the desire for relationality and reciprocity, and the ambition to an ideal, provided 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 condition of impartiality, that emerges the “person” seen as the human capability to conduct 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 severed, may open the way 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, because they are seen as “extrinsic” to the proper finalism of the firm. Hence, I perfectly understand the Catholic believers who want to foster the idea of person by maintaining the metaphysical premise that allows them to see the person as naturally directed to a transcendent universal common good. Simply, I think that they should also recognize that this intellectual strategy presumes an act of faith. Neither can it 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.

In so far as the intellectual requirements of simplicity and general intelligibility are accepted as rules of the game, we may fare better by dispensing with such hypotheses. In fact, metaphysically much less demanding alternatives are available (especially in a dialogue on a penultimate matter as, admittedly, CSR continues to be), which enable us to explain the phenomena itself of human motivation as the search for relationality, reciprocity and the ambition to ideal values. We may 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 do not need that internal link between anthropology and metaphysics (even if may be desirable as far as consistency with one’s faith is concerned) in order to explain the causal force of motivations that induce individuals to act according an ideal of justice or to advocate an ideal of the just company.

Of course, it is true that all behavioural economic models of motivational complexity make 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 thus activated is some basic disposition to act (or to prefer) which is represented by a motivational weight attached to these variables.

Is this the place where a natural disposition to seek the common good should be presumed? I would say not. There is no reason to say that this dispositional weight is a “universal constant” across any individual and any culture, and I assume that it cannot be activated without the reaching of an agreement and the formation of the appropriate beliefs, which are both social artefacts. 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 on a principle of justice and the corresponding system of mutual expectations.

**Is this the place
where a natural
disposition
to seek the common
good should be
presumed?
I would say not**

Summing up, I see the possibility of a convergence of our dialogue on 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 conduct of a rational dialogue on CSR, the idea of a general metaphysical final cause identified with the transcendent common good should be placed in brackets (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 to all, and hence not metaphysical but practical, becomes substantially identical to the effort to discover the *overlapping consensus* amongst all the interests, identities and complete views of the good held 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 placing ourselves in the shoes of the others, we (the stakeholders) will be able to identify 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* on a set of distributive principles for these goods. And finally we will also be able to develop our ideal motivation supporting our reciprocal conformity with the agreed ideal principles for the company's conduct.

6. *Some practical shortcomings to the idea of the "company as a community of work".*

Let me conclude with a remark between economics and the law on the company as an institution. We have seen that, without the effectiveness of a metaphysical theory of the common good, the personalist approach is unable to prevent the risk of parochialism, narrow communities with narrow, maybe collusive, interests amongst the members. How may this risk affect the good company? I see a possible, albeit 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 to the idea of a community of work. Moreover, there is no antagonism on this aspect between social contract approach and the community of work approach. The surplus is the fruit of genuine cooperation, and individual contributions are not separable. Therefore distribution may be only the result of a cooperative agreement on the *common* surplus. (to be fair, one should also recognize 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 theorists elaborate a model of corporate governance that is closely akin to my definition of CSR, see Blair and Stout (1999).)

This implies – and contractarians and personalists agree - rejection of libertarian models claiming that every allocation of values amongst the company's members can be made on the basis

of only negative rights, property rights in particular. And at most on the basis of a voluntary exchange between property rights on assets, inputs, etc., so that the negative rights of any party are not violated. In fact, if we adhere to this view, we simply cannot account for the distribution of a cooperative *surplus*, which typically is not covered by negative rights which at most can fix the baseline *form* whereby the purely cooperative *surplus* (the "common good") is measured, so that no one is 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 the empirical sense) may suffer from opportunistic behaviour and collusion. This is true of production cooperatives (which may act opportunistically toward the consumers), and also of productive non-profits. Therefore, also to these organisations, which are already legally recognized as "communities 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 able to include, albeit hypothetically, all the stakeholders in a potential relation of mutual advantage with the firm, even if they are only external stakeholders, like consumers and clients, suppliers, the local communities hosting the plants, 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 *social enterprises* pursuing apparently altruistic goals are possible. They refrain 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 inclusion in the ex ante contract both the internal members of the organisation, entitled in practice to decide, and the external stakeholders, possibly very weak third parties or beneficiaries. Conformist preferences developed around a principle of justice enable equilibrium solutions where internal members do not abuse and leave a substantial part of the surplus to be used to improve the beneficiaries' well-being (see Grimalda and Sacconi 2005).

Otherwise, my worry is that, once the link to a transcendent common good has been severed, the communitarian approach, and also the relational-personalist one, may bring us to a view of the “community of work” as a closed corporatist institution, entrenched in defence of only the interests and the “bonding ties” amongst the producers (entrepreneurs, managers, employees, partners, maybe the owner if s/he is a natural person, but not when ownership is dispersed among a large number of anonymous individuals impersonally linked to the firm by financial market institutions), and giving little or no weight to the claims of external stakeholders like consumers - who cannot be immediately reduced to internal members of a community of work.

This is not at all a conclusion that derives necessarily from the idea, but it is a possible risk if the common good of the members of a community of work is not submitted to the exercise of universalisation that enlarges the view beyond the company to external stakeholders. Exactly as Maritain suggested, also in our small-scale society we need to “transcend” the small community of work to look for ideal “relations” with external stakeholders eligible only for sporadic, impersonal, cooler concrete relationships. Hence, my invitation to uncover the “overlapping consensus” amongst all the company stakeholders applies all the more so here.

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” Costitutional 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.3, N. 5, pp. 1281-1302
- Rawls J. (1971) A Theory of Justice, Oxford U.P.
- 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

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.