Davide Secchi

Altruism and Selfish Behavior. The Docility Model Revisited

2005/08

UNIVERSITÀ DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

http://eco.uninsubria.it

In questi quaderni vengono pubblicati i lavori dei docenti della Facoltà di Economia dell'Università dell'Insubria. La pubblicazione di contributi di altri studiosi, che abbiano un rapporto didattico o scientifico stabile con la Facoltà, può essere proposta da un professore della Facoltà, dopo che il contributo sia stato discusso pubblicamente. Il nome del proponente è riportato in nota all'articolo. I punti di vista espressi nei quaderni della Facoltà di Economia riflettono unicamente le opinioni degli autori, e non rispecchiano necessariamente quelli della Facoltà di Economia dell'Università dell'Insubria.

These Working papers collect the work of the Faculty of Economics of the University of Insubria. The publication of work by other Authors can be proposed by a member of the Faculty, provided that the paper has been presented in public. The name of the proposer is reported in a footnote. The views expressed in the Working papers reflect the opinions of the Authors only, and not necessarily the ones of the Economics Faculty of the University of Insubria.

Printed in Italy in September 2005
Università degli Studi dell'Insubria
Via Ravasi 2, 21100 Varese, Italy

Altruism and Selfish Behavior. The Docility Model Revisited[*]

Davide Secchi[§]

September 2005

Abstract

Herbert A. Simon is widely known for his studies on rationality, artificial intelligence and for his pioneering approach to organizational studies. In one of his latest works, he presented a theory of human interaction, focused on the conflict between the selfish and the altruistic that can be seen as the essence of human relationships. The model is quite ambiguous: (1) it follows a kind of social Darwinism that (2) postulates selfish individuals' extinction. Taking up Simon's hypotheses on altruism, docility, and selfish behavior, we develop an alternative model of human interaction. The main objective of the paper is to show that rejecting neo-Darwinism and assuming slight complications in the model can explain more in terms of social system interactions. We assume that docility and then altruism, in a technical sense, is the basis of social interaction as it shapes the whole system. It is worth noting that, in the model, selfish individuals do not disappear.

Key words: docility, altruism, social system, bounded rationality, social interactions, social Darwinism

1. The problem of altruism

This paper analyzes a model of social systems that is developed on the basis of one of Herbert A. Simon's later ideas (1990; 1993).[1] The model here outlined tries to explain the interactions between selfish and altruist individuals.

The word "selfish" here refers to the paradigmatic *homo economicus* of neo-classical studies, based on the work of utilitarian theorists such as Bentham and John Stuart Mill. Selfish behavior is clearly defined by the fundamental works of the fathers of the economic discipline, starting with Edgeworth (1881), and continuing with Friedman's (1953; Friedman and Savage, 1948) and Becker's (1974; 1976) contributions, just to mention a few of the most important Authors. The point is that the selfish individual

[*] *Working Paper 2005/08*, Dipartimento di Economia, Università dell'Insubria.

[§] Lecturer in Management, Dipartimento di Economia, Università dell'Insubria. Via Monte Generoso, 71 - 21100, Varese, Italy; e-mail: dsecchi@eco.uninsubria.it.

[1] It is worth noting that in *Administrative Behavior* (1947) Simon mentioned the issue of docility, although he didn't developed it further in an analytical way. This is the reason why we refer in the paper to the later works.

behaves in a fully self-interested and rational way, has an ordered set of preferences, and does not care about the other members of the economic system. This approach has never left the economic basis for explaining human behavior, and strong traces still remain in recent microeconomic work (see, for example, Mas-Colell, Whinston, and Green, 1995).

In broad terms, altruism refers to someone who gives something, but gains no return from the beneficiary of his/her altruistic action. Despite this first and basic assumption, the altruist needs someone to reciprocate the action, when the interaction is reiterated (Axelrod, 1984). Reciprocity refers to something that the receiver will do, in order to praise the altruistic party's behavior; for example, it could even be another act of altruism. Theoretical and empirical experiments seem to confirm reciprocation as common behavior within social systems (Axelrod, 1984; see also the recent Yung-An and Day-Yang, 2003).

The international debate on the meaning of altruism and its implications is growing (for a review, see Khalil, 2004). In particular, Khalil (2004) defines altruism qua charity, stressing that efforts should be directed to promoting an interdisciplinary matrix in order to study altruism, since its links with a large number of scientific domains becomes particularly evident (Wilkinson, 2004; Lunt, 2004).

However, if not altruism, social interactions have been addressed by different Authors as an important argument for re-defining the basic postulates of microeconomics. In particular, it is worth noting that Etzioni's "I&We" paradigm (1988) is very similar to Simon's concept of docility (as detailed below), when he presented it as the "properly socialized" version of the under-socialized neo-classical economic man (1988: 13ff). His work *The Moral Dimension* attempts to found economics on different postulates, trying to include the other's interests in the self. Etzioni criticises the fully-rational paradigm and suggests the introduction of non-rational variables, like emotions and values, into human behavior modeling. One of the main implications is that "[i]ndividuals and community are both completely essential, and hence have the same fundamental standing" (1988: 9). This approach leads to the fusion of social and economic domains into a new branch that he called "socioeconomics" (1988).

Another very interesting view is that of Amartya Sen. He defines the two concepts, that of "sympathy" and that of "commitment" as necessary tools for understanding human

behavior (Sen, 1977: 236). Sen specifies that "[t]he former corresponds to the case in which the concern for others directly affects one's own welfare. If the knowledge of torture of others makes you sick, it is a case of sympathy; if it does not make you feel personally worse off, but you think it is wrong and you are ready to do something to stop it, it is a case of commitment" (1977: 236).

In other words, "sympathy" arises when someone is interested in improving somebody's conditions, and this will also enhance those of the one who acts. It is a self-interested behavior.

On the contrary, when "commitment" occurs then someone *feels* that it should be morally correct to act in order to improve somebody else's condition. Hence, it is altruistic behavior since the one who acts does not have any (direct or economic) improvement of his/her personal condition.

Commitment belongs to the so called non-rational domain of human cognitive capabilities, and it modifies the traditional economic assumptions on human behavior. In the sense that commitment is derived from one's values, it reflects the idea of the social system that the individual develops in his personal experience. The analogy with Simon's concept of docility is clearly defined.

Robert Frank is another pioneer of "heterodox" theories concerning the basic economic postulates. In his analytical works (1994) and specifically in his *Passions within Reason* (1988), Frank tries to focus on the role of emotions in human reasoning and in social and economic equilibriums. Once again, similarities between Frank's and Simon's work are wide and interesting (as sustained also by Khalil, 2004). The two provide models in order to find the economic nature of altruism, and define the interaction between altruistic and selfish individuals, reaching quite different results, however (discussed below).

All these approaches lead to or imply a concept of rationality that breaks with the fully-rational individual. The concept of bounded rationality (Simon, 1947) is the one that, in recent years, has successfully provided the alternative way of thinking about the human decision making processes. Our cognitive capacity is limited by external and internal elements: (a) the environment plays a crucial role in terms of constraints, and (b) our brain cannot perceive and control all the variables because of structural cognitive limits. Recent works show difficulties, challenges and, above all, results of the bounded

rationality approach (Conlisk, 1996; Foss, 2003).

Finally, here we follow Simon's idea on altruism (1993). It is individual tendency to interact having both the idea that the other will reciprocate (in some way, in the future) and to be open-minded, i.e. to learn from the others (docility). Within these hypotheses, altruism is a "technical" word that indicates human thinking in terms of the social system. Selfish individuals limit their thoughts only towards themselves, no matter how the system is organized.

The second section of the paper is dedicated to synthesizing Simon's theory of altruism. In the third section we outline some implications of the model, arguing that its strengths are based on strict hypotheses. Further, Simon's model is modified, using a game theory approach, into a game where selfish individuals face altruists. The objective is to define a model that can capture some real features, defining a stable relationship between selfish and altruistic individuals. In the last section we try to extend some results of the models to other disciplinary domains, other than sociology and economics.

2. Simon's model of altruism

Simon discusses his theory of altruism in two papers (1990; 1993). The core ideas are traced by the bounded rationality hypothesis and the satisficing approach (Simon, 1947; 1955; 1956; 1979) is used as the basis for explaining human behavior.

The basic concept is that of *fitness*. Following a Darwinian approach, the model is based on the "survival of the fittest" postulate, that is to say that "[i]f several species inhabit the same niche, the fittest will eliminate the others; and species will evolve with the appearance of mutants having greater fitness than the original genotype" (Simon, 1993: 126). In evolutionary biology, fitness relates the altruism hypothesis to the "behavior that reduces the actor's fitness while enhancing the fitness of others" (Simon, 1993: 126). This is the technical meaning of "altruism", and it significantly differs from the common use of the same term (Knudsen, 2003).

Sometimes altruistic behavior emerges and contributes decisively to the species' survival, which means, in biological terms, growth in fitness. In some other cases, altruism can be limited to a very narrow influence or might eventually disappear. Simon suggests that the former case is that of human beings.

In order to survive, humans are "docile," in the sense that our fitness is enhanced by

"the tendency to depend on suggestions, recommendations, persuasion, and information obtained through social channels as a major basis for choice" (Simon, 1993: 156). In other words, we support our limited decision-making capabilities by receiving inputs, perceptions, data, and so on, from the social environment. The social context gives us the main data filter,[2] available to increase individual fitness.

To some extent, the concept of "docility" embodies that of altruism, in the sense that one cannot be altruistic if they are not docile. That is to say that one cannot be altruistic without any impetus toward others in a broad sense (i.e. the social system of interactions). This implies that totally selfish individuals are not docile.

Docility is the core of Simon's approach, as it postulates that this is the way individuals behave in modern human societies, and altruism is likely to develop in docile contexts. Since the general gain (fitness) deriving from being docile overcomes the loss of fitness related to altruistic actions, altruism will evolve in our society.

Following the hypotheses of bounded rationality and docility, Simon determines the fitness of the selfish (*fS*) and of the altruistic individual, dividing the intelligent from the unintelligent[3] (*fI* and *fU* respectively).

$$fS = fn + faI \cdot qI \cdot cI + faU \cdot qU \cdot cU \quad (1)$$

$$fI = fn + fd \cdot dI + faI \cdot qI \cdot cI + faU \cdot qU \cdot cU - c \cdot cI \quad (2)$$

$$fU = fn + fd \cdot dU + faI \cdot qI \cdot cI + faU \cdot qU \cdot cU - c \cdot cU \quad (3)$$

He specifies that "*fn* is normal fitness; $fd \cdot dI$ and $fd \cdot dU$ are increments of fitness for docility; *faI* and *faU* are increments in fitness from others' altruism; *cI* and *cU* represent the extent to which *I* and *U* are altruistic; *dI* and *dU* denote the abilities of *I* and *U* to benefit from docility; *qI* and *qU* are percentages of *I* and *U* in the population. The cost of altruism is *c*" (Simon, 1993: 157-158).

Evidence from the model shows that the selfish individual does not benefit from docility fitness "surpluses," and has less chances of survival. This chance of survival derives from the costs of altruistic actions that, as stated in the "technical" definition, must be taken as deteriorations in fitness. If these costs of the intelligent and unintelligent

[2] I do not use, voluntarily, the term "information" that relates to the cerebral activity of data decrypting and interpreting.

[3] The difference between the intelligent and the unintelligent altruistic individual is related to their degree of altruistic behaviour. The intelligent are less altruistic, in the sense that they can differentiate among different actions, while the unintelligent behaves altruistically without any distinction. Not being able to discriminate between different alternatives classifies unintelligent behaviour.

altruists are compensated by the docility effect and the fitness surpluses derived from the effects of altruistic behavior of the others, the altruists will be the fittest individuals. In the opposite case, their proportion in the total population will soon decline.

However, human beings are social individuals where being social means to be docile. If docility will appear to foster altruism as a way to enhance the general social fitness, then altruists (the intelligent ones) will be fitter than the selfish individuals.

Equations (1), (2), and (3) define the framework through which the social system evolves. Selfish individuals are defined as the stereotype of the economic actor, as in neoclassical economics. This *homo economicus* is not docile, by definition. He/she cannot even imagine behaving without a direct personal return, so that altruism is not contemplated. Also docility is not part of his/her world because it is grounded on a bias toward the others, and a selfish individual does not depend on others, *strictu sensu*. This independence (only possible in experimental thoughts) leads him/her to found reasoning only on his/her thoughts and personal cognitive achievement. This is a typical competitive disadvantage in a world where the altruists also learn through the social system.

Simon assigns values to parameters[4] and assumes that the population is equally distributed between the three individuals, in the period zero (the first period). After 30 generations, the fitness-related interactions lead to a strong prevalence of intelligent altruists (53%), to a sensible decrease in the selfish (18%), while the unintelligent altruists manifest a slower disappearance (29%).

The intelligent altruist is the fittest. However, the most important element seems to be docility more than altruism. This can also be derived from the initial hypotheses, as docility is a necessary condition in order to make altruism emerge. However, values attributed to parameters in the model show that the main contribution is supported by the ability of the intelligent altruist to gain advantage from docile behavior. Of course, the costs of altruism impact only on *I* and *U*. Thus, docility makes *I* fitter than the others, considering altruism as a kind of side-effect of it.

The "technical" definition of altruism that Simon gives in the first part of his paper becomes complete when docility elements are inserted. Within this framework, altruism integrates a sort of social identity, manifested at the single individual level. The

[4] The values are attributed as follows: "$fn = 1.01$; $fd = 0.02$; $c = 0.005$; $faI = 0.01$; $faU = 0.005$; $q_0I = 1/3$; $q_0S = 1/3$; $p_0U = 1$; $dI = 2$; $cI = 0.8$; $cU = 1$; $dU = 1$." (Simon, 1993: 158).

individual behaves altruistically because it perceives that being part of a social group involves taking direct actions toward an increase in the total fitness (i.e. the whole group fitness).

In other terms, this is not a simple theory of the "altruism paradox". On the contrary, this is a theory of docile society where altruism is more diffused than previously thought in standard economic theories. Docility fosters altruism, and this is what makes social interaction possible.

3. Challenges to the basic model

The concept of docility as an explanatory bias for social systems. Belonging to Khalil (2004), Simon's theory of altruism can be defined as 'altercentric,' within the framework of interactional rationalistic theories of altruism. This group of theories is called "interactional because, despite their differences, they model action on the standard economic approach, viz., as the outcome of optimization" (Khalil, 2004: 99). In particular, the altercentric perspective "can account for resource sharing where the agent does not stand to collects a benefit in egoistic form or in the egocentric form. It maintains that the agent, at least in some occasions, may share income because he is built with a *pro-social trait*. However, such a trait is not modeled as the desire to enhance the welfare of recipient, but rather modeled as springing form, what one may call, a *moral gene*" (102; emphasis added).

Simon's theory is typically biological in his basic elements (Sesardic, 1995), as it refers to conditions for the evolution of the population (species). Sesardic distinguishes between psychological altruism and evolutionary altruism, even if they are related (1995: 130). From this point of view, Simon refers to evolutionary biological altruism as it is not related to the contingency of cultural and traditional elements. However, giving altruism a social bias means transferring the evolutionary discourse into the everyday interaction between individuals. Social interactions imply cultural, political, economic, moral, and psychological elements that intervene in determining the way people behave. Thus docility goes ahead in describing evolutionary altruism, and this is one of the points addressed in the discussion. In other terms, we praise Khalil's emphasis on the need for interdisciplinary efforts in order to analyze the altruism phenomenon, and human behavior in general.

Simon's simple model opens the borders of narrow neoclassical economics. Docility (or "socializability," as in 1990) is the basis of human acquisition of knowledge, because it basically consists of cognitive delegations to other society members. In other terms, from our birth we operate this kind of delegation, first to our parents, and then to other people. After that we begin to select between people from which to learn something important or insignificant, and the level of a personal role in acquiring information becomes ever higher.

The item here is twofold. In the first instance, people delegate data acquisition to their experience and to the external resources and individuals (Magnani, 2001). On the second, this implies that people do trust others (Magnani, 2005). Moreover, from a neurological point of view, we can infer that "a big cortex can provide an evolutionary advantage only in presence of a massive storage of meaningful information and knowledge on external supports that only an already developed small community of human beings can possess. […] If we consider high-level consciousness as related to a high-level organization […] of human cortex, its origins can be related to the active role of environmental, social, linguistic, and cultural aspects" (14). Thus, "docile" interaction lays on the very basis of our social (and neurological) development.

These two conditions imply that without a significant level of trust this "docile" process is not possible at all. It is very high during the first phases of our lives, then decreases or increases depending on the group in which we operate (e.g. family, firm, association, political party, church, etc.). It is clear that the "docile" individual is, in the technical meaning here used, the same as the properly socialized individual (Etzioni, 1988), or the individual displaying "pro-social" behavior (Khalil, 2004).

> **H1.** *People that behave altruistically pass the docility effect to the overall members of society.*

It is worth noting that social behavior seems to be much more fundamental than usually thought, both for humans and animals (especially for the other primates; Humphrey, p. 307). Following Humphrey, we can affirm that docility and altruism make the "transmission of information" and "individual learning" possible. Then, the first hypothesis is concerned with underlining this essential role of docility for the overall members of society.

Neo-Darwinism in Evolutionary Theories and in Social Sciences. Another point to be discussed in Simon's model is that of applying Darwinism, in its pure form, to economics and, in general, to the social sciences. Darwinism is eminently a biological concept, and its automatic application to the characters of social evolution may lead to distortions (Sesardic, 1995).

The first reason relates to the concept of natural selection, in the formulation accepted in Simon's view. It is mechanical in its essence, as it refers to the organism's adaptation to a specific environment (niche). The hypothesis is that when different organisms strive for survival, only the fittest will survive while the others will disappear. This approach encounters numerous difficulties if transposed to social and psychological behavior, where mechanical relations very often do not apply (Simon, 1979; Frank, 1987). Thus, social evolution is modified by biological bias (i.e. genes), but only in the long run, and the mechanical application of Darwinism to the social sciences (economics in particular) needs at least to be clarified.

> **H2.** *The process of "social selection" of behaviors that fit is not mechanical, nor deterministic in its essence.*

The second reason to question the neo-Darwinian approach to social systems refers to recent developments in the field that challenge the "survival of the fittest" concept. Some Authors suggest that evolution does not work exactly as Darwin thought. The evolutionary process is more dynamic than mechanical, so that we find many options, even if they are not actual options (Maturana and Varela, 1984).[5] The basic concept here is that of co-evolution (1984: 92ff), that is the way in which the organism and the environment find a mutual variation. The organism modifies its character in order to reach better fitness; however, the environment equally, is not insensible to modifications. Within this complex system of changes, many organisms might fit the same environment (niche), which is neither stable nor insensible with regard to the organisms which live in it.

The number of evolutionary possibilities relates to environmental change and the adaptation process that founds co-evolution. It means that evolution is not "deterministic" (in classic terms), but dynamic and adaptive. Nonetheless, this concept

[5] Going back to Mendel, sometimes we find "recessive genes" or "characters" in the group or species. That is to say that the individual with lower fitness does not entirely disappear.

can be found, somehow, in Darwin's works, as Maturana and Varela pointed out, implying that the major responsibility lays on neo-Darwinism and on the scientific *vulgate* of the Master's work. Hence, absorbing these concepts into the model may lead to distortions in the understanding of social interactions.

The other point is more disruptive (to some extent, it breaks with orthodox Darwinian tradition). Maturana and Varela found that, if evolution follows a dynamic and adaptive approach, the rule of the "survival of the fittest" does not entirely apply (1984: 103). In fact, the individual that is nearest to the environmental conditions (and even its changes) shall survive. It is not the "fittest" in absolute terms, but the one that shows its conditions to be more suited to that specific environment (niche), as specified by the concept of co-evolution. This is a "relative" and dynamic approach to evolutionary biology that can be successfully adapted to theories of social and economic behavior. Moreover, it seems to be very similar to the "satisficing approach," in its biological application. In fact, relativity means that the biological organism does not have a clear perception of the overall natural and social variables, and it adapts itself in order to be part of the natural and social system. A model of docile behavior needs to embed these concepts of dynamic adaptations, and to open towards the possible survival of different organisms in the same niche.

> **H3.** *The social context is based on the survival of adapted individuals, hence we do not find the one best survivor, but we might have the contemporary presence of different individuals in the same environment (niche).*

One of the implications of Simon's assumptions on Darwinism can be seen directly from the model. If we do not limit the calculus on population growth and ratios, and extend it indefinitely, we find that the fittest will survive, and that the others with lower fitness will gradually disappear. After 60 interactions the selfish decrease to 8.7%, while the unintelligent altruists stand on 21%, and the intelligent reach 70.3%. The other results are shown in Table 1 (the basis for calculations are those written in the previous pages).

Table 1 – Population movements of *I*, *U*, and *S* in *n* interactions[6]

Periods	Intelligent altruist	Selfish	Unintelligent altruist
1	33.3%	33.3%	33.3%
30	52.7%	18.5%	28.8%
60	70.3%	8.7%	21.0%
100	86.0%	2.6%	11.4%
199	98.2%	0.0%	1.7%
344	99.9%	0.0%	0.0%
721	100.0%	0.0%	0.0%

Source: Our data on Simon's (1993) hypotheses.

The table clearly shows that the intelligent altruist, after 721 generations, will be the only individual in society. The result should be the same even if starting from other ratios in the population. The first, and obvious, result is that the model does not reach any equilibrium between the three (in Pareto-efficiency terms), since the intelligent altruist only will survive. The second result is that the one who survives is the fittest individual, as Darwin's approach can confirm.

However, we cannot deny that selfish (i.e. totally selfish, but intelligent) individuals do exist in our societies. From this point of view, Simon's model is a sort of approximation that fails to give a coherent impression of the society where we all live. Furthermore, if the "survival of the fittest" should be substituted by the "survival of the more adaptable" view, then we cannot exclude selfish behavior as it can be the "more adaptable" in relation to eventual future social dynamics.[7]

If these arguments are consistent, the model fails to address three important points: (1) it fails to record that docility leads to important fitness effects on everybody (and not only for altruists); (2) the model needs to go beyond the mere application of neo-Darwinism, i.e. "mechanical" determinism; (3) it also needs to encounter the potential contemporary survival of diverse organisms (or individuals) within the same environment (niche).

For these and other reasons that will be explained below, we decided to modify Simon's model. The objective is that of reaching a model of real interplay between actors.

[6] The sum of percentages shown in lines five and six (periods 199 and 344) do not reach 100%, leaving a marginal 0.1% error. The reason is that values 0.0% are approximated to zero, but they are not so far from it. On the contrary, the last line shows 100% because we are approximating values with ten and six zeroes after comma.

[7] Of course, we cannot imagine that fully-selfish individuals exist, just as we cannot presume the existence of fully-altruistic individuals. The model approximates human behaviors into variables, as common in economic modeling.

4. A model of docile society

Since Simon began working on bounded rationality (1947), economic modeling has improved at a very high level, leading to the improvement of the discipline. However, we want to remember Hayek's advice on analyzing social and economic variables. With reference to the engineers of economics, he stated that exchanges in the economy are basically the opinions of the interacting individuals. Market prices, for example, express the opinion of the seller on the fair value that a good or service has, and might not match the opinion of the buyer. Therefore, we cannot analyze the price volatility just as if it were a gas molecule (Hayek 1975). Economics is fundamentally a study of human nature (Marshall 1921).

This premise leads directly to the sense we attribute to our model. It is a means (or a tool) that provides us with useful representation of the interactions we suppose individuals have. In no case do we assume that the model exactly mirrors actual human interactions; moreover, the main variables assume significance only if reported to the general theoretical background here exposed. In other terms, we suppose that the fully-selfish or the fully-altruist individuals do not exist, so that players in the model reflect the trend to behave either mainly altruistically or selfishly.

The first move from Simon's model to the one here presented is that of changing actors and the scheme of interactions. If we assume that the fully-selfish and the fully-altruist (or docile) individual does not exist, then the two actors are just approximations, and "altruist" or "selfish" elements have to be intended as the dominance of the first or the second character within the individual. Therefore, there is no need to assume an unintelligent altruist, because the altruistic action might be rational or not (in the sense that goal-attainment is, more or less, the matter of evaluation) depending on the situation, i.e. on environmental variables and personal resources. Capabilities vary from individual to individual and even the "unintelligent" can behave in an intelligent way, depending on the circumstances.

Another reason not to consider the unintelligent individual is the ordinary rule followed by economists: we assume that if the model is consistent for two actors, it should be so also for more than two of them. In the end, the unintelligent altruist seems to be both insignificant, to the extent that it is implied in the behavior of the other two, and unnecessary, if starting with a simpler model.

The structural ideas developed in the model are taken from the work of Robert Frank (1994: 256ff); the attempt is to match his "Hawks and Doves" model to Simon's (Frank, 1994: 256-257).

We switch Simon's docility model of evolutionary change into an interactive game between two players, the selfish (*S*) and the altruist (*A*).[8] The payoff (fitness, *f*) of the two individuals are described as follows:

$$fS = qS \cdot \left(\frac{(PS_w + PS_l)}{2} \right) + qA \cdot xPA \tag{4}$$

$$fA = qS \cdot yPS_w + qA \cdot PA + dA \tag{5}$$

The fitness level of *S* and *A* depends on the probability for *A* to meet *A* or *S*, and vice-versa. Then, the equations express probability to meet an altruist or a selfish individual; this incidence will depend on the quantity of altruists or selfish individuals in the population.

The parameters are defined as follows: *qS* and *qA* refer to the percentages of population of *A* and *S* in a given period; PS_w and PS_l stand for the payoff of a selfish individual meeting another selfish individual, and respectively winning or losing; *PA* stays for the altruist meeting another altruist payoff; *dA* expresses the docility effect.

Moreover, when *S* meets *A*, suppose that *S* is really capable of having a great advantage from *A*'s behavior. Therefore, the effect of altruism is enhanced and expressed by the *PA* multiplied by the *x*. In addiction, consider that the effect of being docile applies also to the selfish. Docility is typically a character that pertains to the individual. However, if it can be defined as the social "vision" of the altruistic individual, then it should be possible to think about the effects of docile behavior on the society overall. This implies a multiplier effect on the altruist-meeting-a-selfish individual payoff, i.e. the selfish one is a free rider.

On the contrary, we suppose that when *A* meets *S* it is possible to recognize a little altruism. This is a sort of side-effect of taking always self-oriented actions, in the sense that even a selfish individual cannot take into account all of the impacts deriving from its behavior. The *y* has a discount effect on the value of the winning selfish payoff (PS_w). In other words, we suppose that a little portion of selfish actions have an impact

[8] To some extent, the original model can also be viewed as an interactive game where each player is associated to a probability of meeting the others in relation to their success, i.e. fitness and population growth or decrease.

on the other members of society, despite the intentions of the player. This is a realistic assumption since, for example, a shopkeeper can clean the sidewalk near her/his shop in order to give customers a better welcome; while in any case, most of the citizens passing along the sidewalk receive benefit from this self-directed action. It is a kind of positive externality effect (Coase, 1960) or positive external economy (Pigou, 1950).

The docility effect does apply only to altruists. In this case, the fitness of *A* is enhanced by a relative competitive advantage for survival because *S* doesn't have it at all. The hypothesis is that *S* tries to improve his/her presence in society through the optimizing process. That is to say that *S* gives a value to every alternative in order to choose the best way to optimize the variables. *A*'s behavior differs from this because his/her goal is twofold, as he/she (1) tries to find a way for survival, and (2) gives a major relevance to variables of the social context (i.e. the others).[9]

This view suggests that, in the struggle for survival, the altruist always gives something, no matter if she/he comes into contact with another altruist or with a selfish individual. On the contrary, the selfish individual doesn't give anything to anyone (except for the above mentioned side-effect) and, uses optimizing mechanisms, trying to reach very high benefits from the altruists. Moreover, the interaction between selfish individuals can be positive or negative, i.e. it can be costly or provide a high return in terms of fitness. Since the selfish entity does not pay for altruist actions, and he/she is not docile (or does not display pro-social behavior), he/she has the same probability of gaining or loosing positions when meeting another selfish person (PS_w or PS_l).

We decided not to highlight the costs of altruism, as they appear in Simon's expressions. The reason is very simple, and refers to the fact that they can be integrated in the return parameters. Therefore, we consider net-revenues (or payoffs), i.e. netted by costs. Thus, the net costs of the game can be expressed as follows:

$$PA = GPA - cA \tag{6}$$

$$PS_w = GPS_w - cSy \tag{7}$$

$$PS_l = GPS_w - cS \tag{8}$$

It is clear from the equation that the net payoff for one altruist derives from the gross payoff for another altruist (*GPA*) to which the costs of the action are subtracted (*cA*). We also suppose that even selfish individuals support, to some extent, costs due to their

[9] The latter is similar to Sen's concept of "sympathy" while the former is nearer to "commitment."

actions, so that the winning selfish payoff (PS_w) is netted by the costs of the action, discounted at y ratio. On the contrary, the losing selfish payoff (PS_w) is simply netted by the costs of the selfish action. Finally, we follow Simon and the common hypotheses on the costs supposing that the selfish action is cheaper than the altruistic one ($cS > cA$).

The first equation shows what should be the average level of fitness (or the average payoff) that selfish and altruistic individuals gain from living in a context that follows the above mentioned hypotheses. These results, however, have to be related to the population growth, since we want to know if docility causes altruists to survive or die. Additional hypotheses have to be done.

Suppose that the average payoff, i.e. the level of fitness (fA and fS), leads to better conditions of survival. This means that if the value of fS is higher than fA, then the selfish percentage in society will increase ($q_1S > q_0S$) at a given rate of growth (g); vice-versa, if $fA > fS$, then altruists will increase, at the same rate (g). Since this is an interactive game, where players meet following simple behavioral hypotheses, we suppose the rate of growth to be the same for the two types of players. This assumption does not affect the evolutionary significance of the model, as one can presume that the reproductive "mechanisms" are the same for every woman and man. The reproductive system works effectively if social conditions permit it, otherwise the other type of individual will advance.

However, it is also possible to think in terms of possible changes, i.e. the possibility that a selfish individual will become an altruist and the other way round. Some terminological modifications could be needed, in that case.

In formal terms, the S and the A population in period 1 (q_1) are defined in the following expressions (9 and 10), under the constraints that the first (9) is valid only if $fS > fA$, and the second (10) only if $fA > fS$:

$$q_1S = qS \cdot gS \tag{9}$$

$$q_1A = qA \cdot gA \tag{10}$$

As the hypotheses are different, we cannot apply the values Simon's gave to his model. Hence, for example, suppose that: $qS = 0.5$; $PS_w = 0.08$; $PS_l = -0.1$; $PA = 0.008$; $dA = 0.04$; $x = 6.3$[10]; $y = 0.01$; $g = 0.02$. Then, after the first 46 interactions, the social system finds the equilibrium with $qS = 3.5\%$. The model shows how such a system will

[10] This value is very high in order to integrate S's ability to keep advantage on A, and also to take into account the free riding effect, that means general costs reduction.

improve living conditions for altruists, while leaving selfish individuals a reduced role. Moreover, it supports the hypothesis that, even in the social systems dominated by altruism, the selfish do not disappear, even if they cover a very limited role in the game. However, the latter might increase in number when the system changes its conditions, registered in parameter modifications.

5. The end of selfish economy

The new model obtained addresses the major points discussed in the previous section. In particular, it integrates the effect of docility on every society member, as selfish members can also take advantage of the altruist's behavior. This point is, to some extent, also present in the original model. However, we tried to give it a specific emphasis in the discussion and in the main variables of the model.

Our model doesn't strictly follow the neo-Darwinism theory. It is not "mechanical" or "determinist" in a classical meaning; on the contrary, the attempt to base it on game theory gives it a probabilistic background. This means that the model follows social modifications, as it is open to modify the results if environmental and individual conditions change.

Furthermore, the "survival of the fittest" thesis is abandoned, as the results show that different individuals can survive in the same environment (niche). Different conditions in environment and in individual behavior lead to different payoffs that open the way to other survival strategies and fitness equilibriums between the players.

Problems arising from Simon's theory of docility were fundamentally those of (1) explaining the permanence of selfish individuals in docile societies, and (2) avoiding the hypothesis of social Darwinism. The new model encounters these problems as part of its hypotheses, and can be interpreted as a partial solution of them.

The model reaches its equilibrium, leaving little space to selfish individuals, so that self-interested people are set aside. This element does not imply that self-interest will be relegated to a marginal role in the future but that, under the conditions that we find at present in our society, it is. In other terms, the model is not a foreteller; it is behavioral-descriptive. In fact, as Simon's model, it has to be modified in case of environmental conditions changes. It provides a good estimation only if we suppose that the interplay between the actors will continue on the basis of those parameters'. However, if we are

confident about the parameters' consistency, then we have a fair description of our society.

To be more precise on this point, we also suggested, in order to test H1, that "one of the chief functions of society is to act as it were as a 'polytechnic school' for the teaching of subsistence technology. The social system serves the purpose in two ways: (i) by allowing a period of prolonged dependence during which young animals, spared the need to fend for themselves, are free to experiment and explore; and (ii) by bringing the young into contact with older, more experienced members of the community from whom they can learn by imitation (and perhaps, in some cases, from more 'formal lessons')" (Humphrey, 1976: 310). This appears very clearly in the model, as "the resulting mix of old and young, caretakers and dependents, sisters, cousins, aunts and grandparents not only calls for considerable social responsibility but also has potentially disruptive social consequences. The presence of dependents (young, injured or infirm) clearly calls at all times for a measure of tolerance and unselfish sharing" (Humphrey, 1976: 310). We can argue that this "measure of tolerance and unselfish" behavior must be incredibly widened in our societies since it is fundamental for their survival. The "measure" might increase or decrease in relation to the environment, time and other variables.[11]

The hypotheses here tested seem to have implications on three levels, at least, so that this model, or a more developed one, can be applied to specific fields of social sciences. The most direct level to which we can refer to is the *system*, though it is also the one we tried to describe in the paper. If the docile-altruist character is dominant, then selfish and fully-rational economics faces a serious challenge. The interaction between two dominant individual types leaves neo-classical economics a marginal role in normative and prescriptive terms. However, this is a highly challenging domain where new models and approaches need to be developed. Time has passed since the first Authors denounced the neo-classical theoretical derive, and modeling and approaches have been developed (see, for example, Frank, 1988; Simon, 1958). The new institutionalism in economics, as in sociology and politics (for a review, see Scott, 2001), provides us with very useful concepts for "complicating" traditional models (DiMaggio and Powell, 1991; Williamson, 1975).

[11] From this point of view, we don't see "moral education" as a "form of coercion" (Yung-An and Day-Yang, 2003: 683).

The docile approach here described has its most significant implications mainly at the *organization* level, as it can be used, for example, to explain why different types of organizations survive within the same environment (niche). If we define "selfish" and "altruist" as a way in which to define organizational external behavior, we find that the whole theory of externalities, both positive and negative, can be understood as prevalence of the first or of the second character. Finally, the organization needs to be docile in order to fit the environment. In other terms, its behavior has to be "socially oriented," in the sense that we explained in the paper, i.e. "the tendency to depend on suggestions, recommendations, persuasion, and information obtained through social channels as a major basis for choice" (Simon, 1993: 156). This approach can also give a very significant role to organizational sociology and corporate social responsibility (CSR) studies. From this viewpoint, CSR becomes the study of docility as it applies to corporations (Secchi, forthcoming).

The third level to which the docility approach and the model can be applied is that of the *individual*. As mentioned in the paper, selfishness and altruism might be addressed to the same individual, as his/her behavior may follow the former or the latter inclination. The individual cannot be thought of as being the same over time and giving the same answers even to the same environmental stimulus. This is further complicated by environmental and individual changes (the concept of co-evolution is worth noting here). In order to integrate these hypotheses in explaining individual behavior, the model (and the approach here stressed) can be useful in deriving which behavior will prevail, under certain external conditions. This is a challenge to neo-classical microeconomics too, as shown by Foss (2003). A new model of human cognition seems to be needed in order to integrate non-rational and only apparently non-rational behavior (Etzioni, 1988; Frank, 1988). However, this is not sufficient as rationality doesn't appear to be consistent with docility and altruistic behavior; hence, recent advancements in cognitive science can be useful for defining a new model of rational choice (Margolis, 1981; Bardone and Secchi, 2005). Moreover, trustworthiness and docility effects are enhanced in restricted contexts, such as within organizations, so further studies need to be directed there too.

Finally, altruism and docility seem to be two important concepts for social and economic reasoning. Many advances appear possible if we think of applying them to

psychological, social, political and, mostly, economic domains.

In order to be more precise, "[e]conomic theory has treated economic gain as the primary human motive. An empirically grounded theory would assign comparable weight to other motives, including altruism and the organizational identifications associated with it. The theory would recognize that human motives change over time, responding to experience and surprises of history. Individuals do not form their preferences in isolation from other individuals, but in response to both public events and information that is widely broadcast. Theory must make room for tulip crazes, responses to oil shocks, or the unexpected rise of ethnicity. Instead of political science or history as derivative from economics analysis, there is a need for economics based upon the facts of history and political social life. [...] Altruism, especially altruism derived from group and organizational loyalties, will play a major role in it" (Simon, 1993: 160).

This is a very stimulating research program to which we will dedicate our future interests, starting with an empirical evaluation of the model.

References

Axelrod R. 1984. *The Evolution of Cooperation*. New York: Basic Books, Inc.

Bardone E. and Secchi D. 2005. "Extending the Bounded Rationality Model. The Distributed Cognition Approach." *Working Paper*, 9. Varese: Department of Economics, University of Insubria.

Becker G.S. 1974. "A Theory of Social Interactions." *Journal of Political Economy*, 82(6): 1063-1093.

_____ 1976. "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology." *Journal of Economic Literature*, 4(3): 817-826.

Coase R.H. 1960 (1988). "The Problem of Social Cost." In *The Firm, the Market, and the Law* (1988), The University of Chicago Press, Chicago and London.

Conlisk J. 1996. "Why Bounded Rationality?" *Journal of Economic Literature*, 34(2): 669-700.

DiMaggio P. and Powell W.W. (eds.) 1991. *The New Institutionalism in Organizational Analysis*. London: University of Chicago Press.

Edgeworth F.Y. 1881. *Mathematical Psychics. An Essay on the Application of Mathematics to the Moral Sciences*. London: C. Kegan Paul & Co.

Etzioni A. 1988. *The Moral Dimension. Toward A New Economics*. New York: The Free Press.1988.

Foss N.J. 2003. "Bounded Rationality in the Economics of Organization: 'Much Cited and Little Used.'" *Journal of Economic Psychology*, 24: 245-264.

Frank R.H. 1987. "Shrewdly Irrational." *Sociological Forum*, 2(1): 21-41.

_____ 1988. *Passions Within Reason*. New York: W.W. Norton.

_____ 1994. *Microeconomics and Behavior*, Second Edition. New York: McGraw-Hill.

Friedman M. 1953. *Essays in Positive Economics*. Chicago: University of Chicago Press.

Friedman M. and Savage L.J. 1948. "The Utility Analysis of Choices Involving Risk." *Journal of Political Economy*, 56(4): 279-304.

Hayek F.A. von 1975. "The Pretence of Knowledge." In *New Studies in Philosophy, Politics, Economics and the History of Ideas*, pp. 23-34. London and Henley: Routledge.

Humphrey N.K. 1976. "The Social Function of Intellect." In Bateson P.P.G., and Hinde R.A., *Growing Points in Ethology*, pp. 303-317. Cambridge: Cambridge University Press.

Khalil E.L. 2004. "What is Altruism?" *Journal of Economic Psychology*, 25: 97-123.

Knudsen T. 2003. "Simon's Selection Theory: Why Docility Evolves to Breed Successful Altruism." *Journal of Economic Psychology*, 24: 229-244.

Lunt P. 2004. "Questions of Definition Scope in Economic Theories of Altruism: A Commentary on 'What is Altruism?' by Elias Khalil." *Journal of Economic Psychology*, 25: 135-139.

Magnani L. 2001. *Abduction, Reason, and Science. Processes of Discovery and Explanation*. New York: Kluwer Academic/Plenum Publishers.

_____ 2005. "Mimetic Minds. Meaning Formation through Epistemic Mediators and External Representations." *Draft paper* (early consultation allowed by Author's courtesy).

Marshall A. 1920. *Principles of Economics*, Eight edition. London: MacMillan and Co.1921.

Mas-Colell A., Whinston M.D., and Green J.R. 1995. *Microeconomic Theory*. Oxford:

Oxford University Press.

Maturana H. and Varela F.J. 1984. *El àrbol del conocimiento*. Santiago de Chile: Editorial Universitaria (read in its Italian translation, 1987).

Pigou A.C., (1920) 1950. *The Economics of Welfare*, Fourth edition. London: MacMillan and Co., London.

Scott W.R. 2001. *Institutions and Organizations*, Second Edition. Thousand Oaks (Calif.): Sage Publications.

Secchi D. forthcoming. *The Docile Corporation. Utilitarian, Managerial, and Relational Theories of Corporate Social Responsibility*.

Sen A. 1977. "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy and Public Affairs*, 6(4): 317-344.

Sesardic N. 1995. "Recent Work on Human Altruism and Evolution." *Ethics*, 106(1): 128-157.

Simon H.A. 1947. *Administrative Behavior*. New York: The Free Press.

_____ 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, 69(1): 99-118.

_____ 1956. "Rational Choice and the Structure of the Environment." *Psychological Review*, 63(3): 129-138.;

_____ 1979. "Rational Decision Making in Business Organizations." *The American Economic Review*, 69(4): 493-513.

_____ 1990. "A Mechanism for Social Selection and Successful Altruism." *Science*, 250(4988): 1665-1668.

_____ 1993. "Altruism and Economics." *The American Economic Review*, 83(2): 156-161.

Simon H.A. and Newell A. 1958. "Heuristics Problem Solving: The Next Advance in Operations Research." 6(1): 1-10.

Yung-An, H., and Day-Yang, L. (2003). "Altruism versus Egoism in Human Behavior of Mixed Motives. An Experimental Study." *The American Journal of Economics and Sociology*, 62(4): 677-705.

Wilkinson G.S. 2004. "A Reply to Elias Khalil's 'What is Altruism?'" *Journal of Economic Psychology*, 25: 125-127.

Williamson O.E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications. A Study in the Economics of Internal Organization*. New York: The Free Press.