

# Could *homo æconomicus* become a revolutionary?

## On the need to teach and practice a different economics \*

Christian Arnsperger  
FNRS and Chaire Hoover (UCL)  
arnsperger@etes.ucl.ac.be

Philippe De Villé  
IRES (UCL)  
deville@ires.ucl.ac.be

November 2002

### 1. Introduction

Our paper does not aim at arguing in favor or against revolution. It is merely an attempt to investigate the *possibility* of revolutionary action within the particular set of assumptions about individuals and society made by mainstream economic theory.<sup>1</sup> Thus, we do not ask 'Should *homo æconomicus* become a revolutionary?' but 'Could *homo æconomicus* become a revolutionary?' Of course, even this more modest question would seem to presuppose that we do *not* believe the phenomenon of revolution to be *unacceptable* or *unwarranted*, always and everywhere. In fact, it only presupposes a concern for the *possibility* of explaining a historically observed phenomenon, namely the periodic uprooting of some fundamental features of the prevailing social arrangement. We shall see that the answer to our question is negative: man as represented by standard theory could never become a revolutionary. This will turn out to have important epistemological as well as political consequences. Indeed, even a moderate such as Amartya Sen will turn out to be an involuntary 'conservative', and the legitimacy of standard economics as a body theory that could be taught indiscriminately will turn out to vanish almost totally. A rather different economics will need to be taught if we want to make sense of the *practical*

---

\* A previous version of this paper was presented, in abridged form, at the IRES Workshop on 'Understanding power in economics', organized in Louvain-la-Neuve on April 19, 2002. Insightful comments by our colleague Enrico Minelli were instrumental in improving the argument and presentation, but he cannot be held responsible for any remaining errors or inconsistencies. This research is part of a program supported by the Belgian government (Poles d'Attraction inter-universitaires PAI P5/21).

<sup>1</sup> In the whole paper, we shall be avoiding the use of the word 'neoclassical', because this word seems to us rather misleading. It overemphasizes historical classification (i.e., neoclassicism as a heritage of, and differentiation from, classical economics) while understating the crucial aspects which we want to focus on: on first, the existence of a body of theorizing which forms a 'mainstream' in academic economics; second, the fact that this main stream has by and large become a 'standard' in approaching social problems, to the point that thinkers who want to push certain non-mainstream insights have to fight hard for legitimacy. Therefore, we will use mostly the expression 'standard economics', sometimes interchangeably for 'mainstream economics'. Clearly, then, the expression '*homo æconomicus*' refers to human agency as represented within this particular, so-called 'standard' body of economic theory.

*possibility* of economic agents to form a reflective judgment on society and, possibly, to contemplate uprooting the socio-economic system in which they live.

We will not engage analytical Marxism directly in this paper. The pioneering work of John Roemer [1982, 1988] has been an effort to reformulate Marxist perspectives and to adapt them to the prevailing theoretical vocabulary of standard economics. Roemer has succeeded in offering a ‘general’ theory of exploitation, identifying in each property-rights regime a certain form of exploitation (feudal, capitalist, or socialist exploitation) which is ‘characteristic’ of that regime. His key claim is then that ‘each revolution in an economic structure abolishes just that kind of property associated with its characteristic form of exploitation’ [1988, p. 137]. He adds, unsurprisingly given his Marxist background, that ‘class struggle is the midwife of social revolution, the instrument by which property relations are transformed in radical ways– from feudalism to capitalism, from capitalism to socialism’ [1988, p. 85]. However, nowhere in Roemer’s whole game-theoretic approach do we actually see individuals coalescing and revolutionizing the pattern of property rights. *The revolution, if it does take place, does not do so in the model, which is only able to portray and formalize successive ‘stages’ and to identify those forms of exploitation which will be superseded at each stage.* This is in itself an awesome accomplishment, to be sure, but it falls short of answering our question: where do we see, in the model’s structure and in its way of actually modelling individuals, that a revolution is possible? The models give no answer, and this seems to be because Roemer, despite his earnest commitment to Marxism and his explicit mentions of ‘revolution’ *outside* of his formal models, has adopted a standard economic vision of social interaction.

The standard economic vision of social interaction, as we will try to show, makes revolutionary phenomena largely a *non sequitur*. To show this, we need to go back to basics and to scan the most fundamental standard results for a clue to where a ‘revolutionary motivation’ might take root– if, indeed, it can at all. This is what we attempt to do in the next section.

## 2. The Second Theorem of Welfare Economics and the ‘revolutionary’s handbook’

### 2.1. Sen on the Second Theorem

Let us therefore comment in some detail Amartya Sen’s reflections on the Second Theorem of welfare economics, or what he calls the ‘second part’ of the ‘Fundamental Theorem’ of welfare economics. This Fundamental Theorem has two parts : the first part says that any competitive equilibrium is Pareto-optimal; the second part says that any Pareto-optimal allocation can be reached as a competitive equilibrium provided the initial endowments have been suitably redistributed prior to trading at Walrasian prices. The first part is labelled by Sen the ‘direct theorem’. It is usually taken by defenders of the ‘market mechanism’ as a powerful proof of the market’s ‘efficiency’: if all trades are carried out at equilibrium prices, all opportunities for mutually beneficial exchanges *at these prices* will be exhausted, so that the only way to make anyone even better off is to make at least one person worse off– and this, it is claimed, is in all cases objectionable. The second part of the Fundamental Theorem is labelled by Sen the ‘converse theorem’. It is usually taken to mean that the ‘market mechanism’ is ‘neutral’ in terms of distributive criteria: simply isolate, among all Pareto optima, those which in addition respect some

criterion of distributive justice, and reallocate the initial endowments in such a way that one of these just and Pareto-optimal allocations can be decentralized as a competitive equilibrium.<sup>2</sup>

We shall focus here on this converse theorem because this is where the issue of revolution is explicitly present. Sen [1993, pp. 521-522, italics added] says the following:

*To use the competitive market equilibrium to achieve any social optimum, we have to get the initial distribution of resources right, and depending on how equity-conscious our social objectives are, this could require a total reallocation of ownership patterns from whatever pattern we may have inherited historically. The 'converse theorem', thus, belongs to a 'revolutionary's handbook'. I am not bothered here by the sociological fact that enthusiastic advocates of the market mechanism are typically not particularly revolutionary in demanding radical redistributions of ownership. More immediately relevant is the recognition that if we are not able, for political, legal or any other reasons, to rearrange the resource distributions freely, the converse theorem does not guarantee even the limited achievement of Pareto efficiency for any given initial distribution of resources. In contrast, the 'direct theorem' does guarantee just that— it ensures something rather solid here and now, even if that achievement is far from adequate.*

This passage needs to be complemented with an earlier version taken from Sen [1987, pp. 37-38, italics added]:

*... the second part of the fundamental theorem' would be used only if it were politically possible to redistribute resources among the people in any way that would be required by considerations of social optimality. Even if the necessary lump-sum transfers were identifiable and also economically feasible, issues of political feasibility can be, obviously, extremely important when dealing with such fundamental matters as radical changes in ownership. Even though the invoking of the second part of the 'fundamental theorem' may often come from rather conservative quarters defending the market mechanism, the result can be of real use only as a part of some 'revolutionary's handbook', transforming the ownership of means of production before getting the market to do the rest. If radical redistributions of ownership were not possible, movements towards overall social optimality will require mixed mechanisms of a kind not covered by the 'fundamental theorem'.*

An immediate implication of these passages is that 'revolutionary' is opposed to 'politically feasible'. However, unless the underlying formal model is to be considered only as a fanciful metaphor, it is the *exact same* agents who

- (i) 'historically inherit' a given structure of initial endowments,
- (ii) either carry out or refuse a 'radical redistribution' of these initial endowments, and
- (iii) carry out Walrasian-equilibrium trades from these— either redistributed or unchanged— initial endowments.

---

<sup>2</sup> Obviously, while this way of using the 'converse theorem' to normatively justify the market mechanism is the most prevalent one, it is not the only possible one. We could, alternatively, interpret the converse theorem as saying the following: for each competitive equilibrium  $E$ , there exists a reallocation of resources such that  $E$  can be obtained through decentralized trades; therefore, if the 'social planner' required to carry out this reallocation is really all-powerful, she could simply reallocate all available resources directly to the competitive equilibrium itself, thus dispensing with the whole 'trading' process from 'initial' resources. This would then make the converse theorem a defence of a purely planned society with no decentralized economic institutions. While such an interpretation is coherent with the formal

There is therefore an issue of *anthropological and politological feasibility* which Sen neglects completely: we need to know whether the human makeup presupposed by step (iii), namely the Walrasian-market anthropology and politology, is compatible with the human makeup presupposed by step (i), namely the anthropology of socio-economic criticism and political deliberation. If it should turn out that there is no such compatibility,

?? either we would have to accept that step (ii) may induce an anthropology and a politology other than the Walrasian-market ones

?? or we would have to say that step (iii) makes step (ii) irrelevant– in which case all of Sen's hypothetical statements commencing with 'if' would have to be replaced by affirmative statements commencing with 'since', and the converse theorem would be part of a 'purely imaginary utopian's handbook'.<sup>3</sup>

## 2.2. The obsession with Pareto-improvement

That there is an often only implicit, but very definite, Walrasian-market anthropology and politology can best be seen through the fascination which the famous 'core-convergence theorem' has exercised over the decades since it was suggested by Edgeworth in the early 1880's. With a very large number of traders, each being virtually 'measureless' (see e.g. Debreu and Scarf [1963] and Aumann [1964]), the core of the economy– i.e., the set of allocations at which no individual or coalition of individuals does worse than in the initial situation– shrinks to the set of competitive equilibria: the *only* way to carry out mutually advantageous trades starting from any distribution of initial endowments is to trade to one of the Walrasian equilibria corresponding to these endowments. This result has been hallowed by mainstream economists because it epitomizes the *contractarian, mutual-advantage view of society* which was inherited from British liberalism and is deeply entrenched in standard economics through the all-pervading 'rationality' constraint: whether under perfect or imperfect competition, an individual or coalition of individuals will participate in trading only if it is individually or coalitionally 'rational' to do so, i.e., if the individual or coalition expects that it will do at least as well as if it refrained from trading at all. The core-convergence theorem says that in the limit, the only way to satisfy all (individual and coalitional) 'rationality' constraints simultaneously is to trade at Walrasian prices.

It is symptomatic of the ideology underlying standard theory that Sen [1987, p. 38] at the same time doubts the relevance of Pareto optimality and agrees fundamentally with the mutual-advantage view of society. The quote is symptomatic indeed: 'The idea that overall social optimality must *inter alia* require Pareto optimality is based on the notion that if a change is

---

content of the converse theorem, it is far from the dominant interpretation, which uses the theorem to advocate redistribution combined with 'free markets'. In what follows, our arguments can be applied to either of the two interpretations.

<sup>3</sup> In that case, putting the modified words between brackets, we would get a fully 'reactionary' version of Sen's second statement: '... the second part of the 'fundamental theorem' <will never> be used <since> it <is> politically <impossible> to redistribute resources among the people in any way that would be required by considerations of social optimality. Even if the necessary lump-sum transfers were identifiable and also economically feasible, issues of political feasibility can be, obviously, extremely important when dealing with such fundamental matters as radical changes in ownership. Even though the invoking of the second part of the 'fundamental theorem' may often come from rather conservative quarters defending the market mechanism, the result can <never> be of <any> real use since it will only be a part of <some 'purely imaginary utopian's handbook'>, with <no way> of <actually> transforming the ownership of means of production before getting the market to do the rest. <Since> radical redistributions of ownership are not possible, movements towards overall social optimality will always require mixed mechanisms ... '.

advantageous for each, then it must be a good change for the society. That notion must in some sense be correct, but to identify advantage with utility is far from obvious.' What Sen is talking about is in fact Pareto *improvement*, which focuses on mutually advantageous changes; he seems to neglect, however, the much less self-evident feature of Pareto *optimality*, namely that 'overall social optimality' requires the *refusal* of any allocative change that is to at least one individual's or coalition's *disadvantage*— regardless of whether that individual or coalition is extremely poor, in which case we might indeed legitimately reject the change, or is the sole owner of all the resources in the economy, in which case such a rejection will amount to nothing but the perpetuation of injustice.

The exclusive focus on the core of an economy as the area of legitimate reallocations is clearly linked to the idea that, at any given moment in historical time, the 'inherited' structure of initial endowments is the 'natural' starting point. How does this affect our discussion of the Second Theorem? If 'political feasibility' is equated with 'belonging to the core', *any rearrangement of the initial endowments will itself have to be in the core*. However, since from that rearranged profile the trading process will, in any case, lead to a competitive equilibrium belonging to the original core, the original rearrangement is irrelevant and might as well not have been operated. Indeed, the core-convergence theorem tells us that from the given initial endowments, the only reallocations which respect all (individual as well as coalitional) rationality constraint are *nothing but* the competitive equilibria; therefore, when the number of traders is very large, carrying out a Pareto-improving reallocation of initial endowments is tantamount to trading from the 'inherited' endowments directly to a Walrasian equilibrium... So it is the ideology underlying the First Theorem— 'all competitive equilibria are Pareto optima, and this is good because Pareto improvement is the first and foremost normative criterion for political acceptability'— which makes the Second Theorem effectively harmless, and eventually irrelevant.

### 2.3. Why self-enclosed ignorance might be theoretically necessary

The deep reason for this is that the First Theorem and its Walrasian-market anthropology and politology represent any individual or coalition being concerned with one thing and one thing only: that whatever allocation is proposed to them should be in the shaded area above the indifference curve(s) going through their initial-endowment bundle(s). The implementation of Pareto optimality and *a fortiori* of core allocations through a series of Pareto-improving reallocations therefore requires only a very restricted collective dimension of any individual's consciousness: the ability to envision self-interest-promoting coalitions. The Pareto-optimality and core-belonging criteria themselves, like all axioms in standard theory, are overarching criteria *decreed by the theorist*: they are never formulated consciously by any of the individuals themselves, or by any subset of individuals. Each of these individuals, whether standing alone or contemplating the formation of a coalition, literally *has no normative or reflective judgment* about the society in which he is located. Sure enough, he may know who some of the other traders are and in what situation they are relative to him, since this is the essence of coalition formation. However, neither any coalition nor any individual reflects on social justice or on social relations so as to form a normative judgment about society. 'Rationality'— whether individual or coalitional— is confined to evaluating the surplus to be gained, individually or as part of a coalition, from trading at given initial resources. Some theorists will euphemise this aspect by saying that such lack of reflective consciousness expresses only a 'sufficiency condition': for the theoretical results to go through, it is *sufficient* that each agent focus only on his own individual and coalitional 'rationality' constraints (since this will, for instance, minimize the size of the 'message space'; see Mount and Reiter [1974]), and hence the theory merely goes ahead without *needing* to assume more than that. These same theorists will assert that the assumed lack of reflexivity is not *necessary*: if the individuals in the model do, in addition, have an overall judgment about the social situation, this is not refused by the theorist but— so he will claim— it does not affect the results obtained under less demanding assumptions.

But is this so sure? We believe it is not. In fact, the constraint which the implicit politology of the First Theorem imposes on the Second Theorem is the following: since what counts ultimately is not Pareto optimality but *Pareto improvement* (of which Pareto optimality is merely an endpoint), the core of any economy need not be indexed by the particular vector of initial endowments of which it is a function. Take an elementary two-agent example that could be represented in a standard Edgeworth box. Let there be two agents *A* and *B*. The two initial endowments  $\omega^1 = \{(5.5, 3.5), (0.5, 0.5)\}$ , which is extremely inegalitarian, and  $\omega^2 = \{(3, 2), (3, 2)\}$ , which is completely egalitarian, generate two very different cores  $C(\omega^1)$  and  $C(\omega^2)$ , but the implicit assumptions underlying the Walrasian politology imply that neither individual *A* nor individual *B* has any 'preference' or 'judgement' which would allow him to *rank* these two cores— i.e., to rank the two corresponding initial-endowment vectors. Suppose the society is currently located in point  $\omega^1$ . If both *A* and *B* were able to 'look up' from their single indifference curve and the shaded area above it, they would *necessarily* realize that  $\omega^2$  is also a possibility. Individual *B* would perhaps on impulse rank  $\omega^2$  first and  $\omega^1$  second because he benefits more from trading at  $\omega^2$ ; if individual *A* has no concern whatsoever for social justice, he is likely to impulsively form the opposite ranking. Since *A* is very rich while *B* is very poor, both individual rankings might not be equally legitimate from the point of view of social justice. But alternatively, *A* might care for justice and may in fact come to militate for a shift from  $\omega^1$  to  $\omega^2$ . Therefore, this mere fact of 'looking up' and coming to know certain neglected social possibilities creates a situation which, potentially at least, could jeopardize the glib distinction between sufficiency and necessity outlined just above: while the unawareness of other social possibilities might indeed be sufficient for the attainment of a competitive equilibrium inside  $C(\omega^1)$ , this same unawareness might turn out to be *necessary* in order to avoid that at least one of the two individuals (and perhaps even both) express a preference for a competitive equilibrium attained inside  $C(\omega^2)$ — i.e., in order to avoid that prior to Walrasian trading the initial endowments be (in Sen's own phrasing) 'radically redistributed' from  $\omega^1$  to  $\omega^2$ .

We may want to ask immediately, Why would anyone want to 'avoid' this radical redistribution? Certainly, a theorist such as Sen is not a defender of injustice? Of course not. While he dismisses the 'converse theorem' by writing that 'despite its apparently greater relevance, [it] has, in many ways, less interest for practical economic policy than the direct theorem' [1993, p. 535], this is essentially because he claims to have a strong concern for 'political feasibility'. However, the model within which he discusses the possibility of a move towards radical redistribution is in fact *structurally unable* to make sense of such a move anyway, because it is part of a larger and long-forgotten heritage which can be called somewhat inelegantly '*empiricist monadology*' (or 'E-M' for short). As we shall try to demonstrate in the next section, the standard economic model is built on this long-forgotten heritage, and this makes it impossible to think of conflict and redistribution in any other terms but those of what we shall call '*empiricist-monadological coordination*'. Thus, while Sen does not purposefully avoid the possibility of revolution, he discusses it— and then dismisses it— within a model which *cannot but* avoid it; as a result, despite all his best intentions, he might turn out to be an unwitting accomplice of reactionary, or at least merely reformist, politics.

### 3. The empiricist-monadological heritage: Mainstream economics as a particular form of metaphysics

Let us now venture for a while into the philosophical underpinnings of the standard economic model. We will claim that standard economics, as a discipline concerned with the analysis of the interactions between individual decisions and the collective effects of these interactions, is a particular form of metaphysics, based on *a very particular articulation between*

*self-centeredness and other-centeredness*. More specifically, economics will be claimed to be a specific metaphysical discipline based on the *monadological* worldview inherited from Leibniz and re-arranged to fit the *empiricist* bias of 18th-century thinkers. This is a rather general claim, and not a completely new one, but it has received precious little attention among standard economists so far, with the result that economics has become increasingly blind to the philosophical roots of its most fundamental views on individual decision and on social order.

### 3.1. From atomism to monadology

Although it is always difficult to say what old thinkers might actually have meant, it seems plausible to read atomism nowadays as a materialist theory of the composition of reality, that is, a theory which postulated (i) that we are able to know the reality that is 'out there' and (ii) that this reality (including invisible things like thought or consciousness) is composed exclusively of and/or induced exclusively by material 'atoms.' Change, that is, de-composition followed by re-composition, took place within the non-being of space and was made possible by inner tendencies (which are assimilated to 'necessity') inscribed within the atoms. The best-known metaphor is that of Lucretius' *clinamen*, or 'swerve,' which allowed free-falling atoms in space to huddle haphazardly into composites, themselves moving to encounter other composites, and so on. Without this swerve inherent in atoms, but fundamentally unpredictable in its effect on the composition of matter, there would be no ordered reality but only eternally separate particles following one another in free fall through space.

This was a form of open determinism which did not appear to exclude that some atoms or composites of atoms exercise what *looks like* a free will (but is really only a built-in mechanism). Scharfstein [1998, p. 177] paraphrases Lucretius' position as follows: 'if motion followed motion in exactly fixed order, without any swerve that "burst the bonds of fate," what would be the source of the "will torn from the fates," the will by which each man moves to where his pleasure— not merely the mechanism of motion— leads him?' In other words, the *clinamen* was the factor which explained the coherence of reality and the order that arises from free individual decisions guided by the search for 'pleasure.' We ought to note carefully, however, that this was a *non-autonomous* form of freedom, i.e., one inscribed within the atoms' ontological structure as a feature *retroactively* explaining the *observed* cohesion of reality.

While the aim here is not to reconstruct a history of atomism, nor to suggest that all forms of atomism flow from a single intuition, one cannot avoid a parallel between ancient atomism and Leibniz's theory of monads. We must simply note that, while atomists tended to reduce spirit to matter (thought was the material reaction of a bunch of atoms being 'hit' by another bunch of atoms), Leibniz tended to reduce matter to spirit; this distinction, however, fundamental as it may be for some philosophical issues, makes little difference for our present discussion of the roots of standard social science. What *is* crucial, on the contrary, is the fact that here again we had a theory where individuals (*a-tomos* and *in-dividuum* are synonyms in Greek and Latin) unwittingly created a global order by following *built-in tendencies which are nothing but an internalized reflection of this pre-existing order*.

The Leibnizian monads, while being simple and indivisible entities quite like the ancient atoms, are nevertheless somewhat more complex than their predecessors because they are capable of what Leibniz called 'small perceptions': all monads, and hence also the composed monads that are human beings, are endowed with 'perception' and 'appetition' as inner tendencies which made it possible to imagine movement and change within a universe where, following Leibniz's famous phrase, 'monads have neither windows nor doors.' This is the atomistic heritage of monadology which will later flow into economics: atoms have only limited contact with each other but they obey inner tendencies which, when combined according to the requirements of 'nature' (which for Leibniz implied no materialism whatsoever or even naturalism in the

usual sense), generate *the global order whose attainment was already inscribed within these individual tendencies* (see De Villé 1988, Dupuy 2002).

As Renaut [1987] has very clearly shown, the passage from 17th-century theism, which still made it possible for Leibniz in his *Theodicy* to postulate the existence of a God or ‘monad of monads’ who imprinted upon all of natural and social reality a ‘pre-established harmony,’ to 18th-century rationalism and empiricism effected a radical change within the monadological tradition itself. Mandeville and Smith were to ‘crack open’ the monads to a significant extent *while nevertheless remaining within the monadological worldview*. Mandeville’s cynically vain and self-interested individuals, as well as Smith’s more subtly imitative and other-regarding individual, look in retrospect very much like a minimal concession to empiricism within a fundamentally atomistic and monadological framework. Clearly, Smith’s idea that collective order (rather than complete and violent chaos) emerges ‘as if’ individual actions were ‘guided by an invisible hand’ was little more than a secularised, empiricist re-casting of the Leibnizian theodicy.

### 3.2. The censored empiricist project: Other-regardingness and the fear of social chaos

Dupuy [1992] has argued convincingly that Smith in fact had a more ambitious project, which was to explain the emergence of social order on the basis of *nothing but* generalized mimetic behavior on the part of all individuals, but that he was deterred from this radically ‘emergentist’ project by his and his fellow philosophers’ inability to conceive adequate feedback mechanisms which would make the fixed point resulting from *other-regarding* individual actions *both* collectively endogenous *and* collectively sustainable. The Hobbesian threat was still very much active: too much looking at each other and trying to imitate and emulate one another created the permanent risk of degeneration into generalized brutality.

Therefore, as confirmed by Renaut, the prospect of generalized chaos and violence led Smith, Hume, Berkeley and all subsequent empiricist-liberal thinkers to revert to a strongly monadological view of social cohesion: individual optimising decisions were coordinated by a *deus ex machina* called the invisible hand of the market. This required, of course, some ‘framing’ of the way decisions are made: agents refrain from various ‘side activities’ and accept some basic variables (such as prices) as parameters in their search for the maximal fulfilment of their preferences. Such framing, which to a significant extent injected the requirement of social order and cohesion *into the individual tendencies*, ‘was a close kin to the atomistic *clinamen* and to the monadological combination of appetite and perception. It transformed the invisible hand into a true *deus in* (rather than *ex*) *machina*, working from within the system through the individual ‘tendencies’.

And this is the background story which, although by now in an almost nonexistent and ‘implicit’ form (because the process view of Smith, Hayek and others has yielded to the equilibrium-state view of Arrow, Debreu and others), still pervades standard economics. Whether it is perfect competition, imperfect competition or noncooperative game theory, the basic logic is always the same: individual monads, possibly somewhat ‘cracked open’ in order to have enough ‘windows and doors’ to carry out certain strategic interactions, take individual decisions whose interaction is mysteriously and more or less invisibly (there might be some learning) coordinated into an ‘equilibrium.’ Let us call this the *empiricist-monadological* (henceforth E-M) framework. We must be quite careful in our understanding of its basic logic: contrary to a truly ‘emergentist’ logic in which order could– but might not– emerge from uncoordinated individual decisions, the E-M framework postulates the reverse: *since* the only way to think about society is as a strictly ‘coordinated’ set of decisions, this imperative of *ex post* coordination has to be, to a minimal extent at least, injected back into the *ex ante* structure of individual decisions and hence individual motivations. Let us put it succinctly but, we believe, adequately as follows: empiricist monadology cannot postulate

fully self-enclosed monads; they need to have 'windows and doors,' but only just enough so that the strategic interactions of their decisions yield a pre-defined social order— pre-defined not in the sense that the content of the decisions is prescribed to the agents, but pre-defined in the sense that, to ward off the threat of social chaos and deadly mimetism, many forms of other-regardingness are excluded by assumption *because they are detrimental to a vision of social order that wants to hold on to two fundamental precepts: (1) agents, even when endowed with strategic other-regardingness, are largely self-enclosed monads and (2) social theory has to study the conditions of global social order.*

This last point is absolutely crucial: carrying the legacy of atomism and monadology right into modern-day social analysis, standard economics— including many approaches other than Walrasian competition, such as all non-Walrasian and most game-theoretical equilibrium models— has remained 'stuck' within a metaphysical framework which seeks to explain the harmony of the social whole on the basis of its *self-enclosed* parts. There is a deep ethical reason for this, which pervades Greek thought and has stayed with us all the way into contemporary analytical philosophy: all these ways of thinking seek to obey 'the Socratic exigency of a mind nothing can force, an exigency Leibniz ... answers to in refusing the monads windows' (Levinas [1969], p. 219).

### 3.3. Is 'mimetic violence' relevant to the question of revolution?

The shift from Greek and German idealism to British empiricism did not fundamentally change the problem, but simply made it all more difficult because, once 'human nature' is conceived not only methodologically but factually as self-reflecting 'intentionality,' the answers to the problem of peaceful coordination become even more elusive than if humans are merely passive atoms.

In fact, as already indicated, standard economics has largely abandoned the Mandeville-Smith project of founding social coexistence on other-regardingness driven by vanity and recognition, and has reverted with the E-M project to a more aseptified monadological outlook, where the potentially violent other-regarding passions are replaced by strategic but still self-enclosed, optimizing calculation. Some theorists (see e.g. Dupuy [1992]) believe that economics should pay more attention to the Mandeville-Smith project because, they claim, this project at least had the merit of combining the concern for coordination with the fact of social violence. This violence, however, is predominantly of the 'mimetic' kind in Mandeville-Smith: envying what others have, individuals engage in attempts at violent reallocations of resources, and this generates both social order and social change. The well-known work of Girard [1972, 1982] has developed this hypothesis to its fullest (see Dumouchel and Dupuy [1979] for a specific application to economics). While this is certainly not the place to go into the intricacies of this fascinating model, we feel it has little or no kinship with the question of revolution which preoccupies us at present. Envy of the poor for the rich may be one possible interpretation of certain revolutionary phenomena and the violence accompanying them; however, as Hannah Arendt has noted perspicaciously, if envy was at all present during the American or the French revolutions, it could almost certainly not explain the way the overall events occurred: 'The realistic views of the Founding Fathers with regard to the shortcomings of human nature are notorious, but the new assumptions of [19th-century] social scientists that those who belong to the lower classes of society have, as it were, a right to burst with resentment, greed, and envy would have astonished them, not only because they would have held that envy and greed are vices no matter where we find them, but perhaps also because their very realism might have told them that such vices are much more frequent in the upper than in the lower social strata' (Arendt [1963], p. 73).

Thus, moving back and forth between atomism and mimetic violence does not seem to give us a plausible clue for our present question, i.e., the possibility of revolution in the standard economic framework. True enough, the main element of envy, namely that agent *B*'s welfare is diminished by agent *A*'s having more than agent *B* (see Kolm [1995]), is absent from

our earlier example of the Edgeworth box; however, Sen's own definition of what a revolution would be, i.e., a radical reshuffling of the initial endowments, does not require that  $B$  should in any relevant sense of the word want to 'imitate' agent  $A$ .

This clarification notwithstanding, there are indeed some basic requirements, as we shall see in the next section, and we will show that they are, very clearly, excluded by the E-M heritage of standard economic theory.

## 4. The prerequisites of revolution

We have mainly shown up to now that despite what Sen says rhetorically, in actual fact the anthropology and politology which he accepts for step (iii) invalidates the anthropology and politology of socio-economic criticism and political deliberation which would be necessary in step (ii), i.e., in what he himself calls a 'revolutionary' phenomenon— namely, the 'radical redistribution' of initial endowments through extra-political means. Thus, in a nutshell, if the Walrasian market mechanism is taken seriously as the representation of a desirable way to organize interaction, then there can be no revolution and there never will be— simply because, in the Walrasian mindset, this phenomenon has no meaning.

But in describing more precisely what the anthropology and politology required by step (ii) would look like, we will discover a deep reason why it is not 'pushed' by even such justice-conscious authors as Sen : indeed, it is likely to imply that step (iii), in turn, is no longer relevant because the individuals who have engaged in socio-economic criticism and political deliberation of step (ii) are no longer 'suited' for the Walrasian framework— or, for that matter, for any self-enclosed, strategic interaction. This jeopardizes the whole standard theoretical setup. So let us ask what anthropology and politology would be required if the individuals are to either carry out or refuse a 'radical redistribution' of the initial endowments.

### 4.1. Going beyond the individual and collective 'rationality' constraints

The first and minimal element, as already indicated, is the mere capacity to look beyond individual or coalitional 'rationality' at given initial endowments. One way to formalize this in a manner that perhaps still remains broadly congenial to the mindset of standard theory is to assume that each individual has the ability to rank initial-endowment vectors. If there are  $n$  individuals and  $m$  commodities, this means a ranking covering the whole  $n \times m$ -dimensional Edgeworth 'hyperbox', since any vector in that space could be a possible initial-endowment vector instead of the 'historically inherited' one.

Of crucial importance, of course, is the criterion according to which this ranking would be constructed. A simple criterion would be for each individual  $i$  to identify with each endowment vector  $\omega$  the bundle  $x^*(\omega)$  he would obtain in the competitive equilibrium (if it exists) from  $\omega$ ;  $i$ 's preferred initial-endowment vector would then be the one yielding his preferred competitive-equilibrium bundle. A slightly more general criterion would be for each  $i$  to associate with each  $\omega$  his preferred bundle inside the corresponding core  $C(\omega)$ . Both these criteria seem to be minimal departures from the attitude of individual rationality merely at given endowments; they allow for a variation of endowment vectors but stay extremely close to the individual-'rationality' and coalitional-'rationality' constraints once the vectors have been varied. A more elaborate criterion would to apply one of the above individually rational criteria only to the subset of all allocations which, in each core  $C(\omega)$ , satisfy a distributive norm such as no-envy, egalitarian-equivalence, or any other such norm. This would require the internalisation by each individual of a distributive norm which would have lexicographic priority over individual preference in the ranking of initial-endowment vectors.

One could spend time analysing and axiomatising numerous such mechanisms for the formation of individual rankings; this would have some use in itself, but it would not ward off the basic and unsurprising problem which all these rankings

would have in common— namely, that *unless all individual rankings turn out to attribute the top rank to the same initial-endowment vector*  $\omega$ , i.e., *unless there is unanimity concerning the 'best' initial endowment vector, there will be a conflict of interests between the individuals.*

## 4.2. Mechanisms for solving inevitable conflicts of interests

Therefore, the second element of the anthropology and politology of socio-economic criticism and political deliberation is the mechanism— whether peaceful or violent— through which the individuals will 'solve' this very likely conflict of interests. There are essentially three 'canonical' ways in which this can be done, as we know: through voting, through bargaining (which may or may not be unanimity-inducing), or through armed struggle. Sen is correct in saying that a revolution will ensue only in the third case, i.e., if there is no way to either vote or bargain. We should be careful, though: the absence of peaceful means is a necessary, but not a sufficient condition for there to be a revolution— what is certain is that there will be some kind of 'civil conflict', but whether the outcome will be a radical change in the endowment vector or only a return to the *status quo* is entirely uncertain.

What Sen calls 'political feasibility' lies in the realm of voting or bargaining, but upon reflection the three canonical modalities are not so easy to keep apart. In fact, modifying our previous example, suppose there are 10 percent of *A*-individuals and 90 percent of *B*'s. The competitive equilibria and cores will not be the same as in the two-agent case, but we can still define an initial vector  $\omega^1$  such that the sum of all (identical) *A*-agent owns most of each resource and an initial vector  $\omega^2$  which is the 'center' of our modified Edgeworth box. Given the great discrepancy between  $\omega^1$  and  $\omega^2$ , majority voting or even a standard qualified-majority system such as 'two-thirds' majority will obviously tilt the society in the direction of an egalitarian redistribution; only unanimity, basically, can block the move. So unless the well-to-do *A*'s have been able previously to inscribe the unanimity rule in the Constitution, *or* unless for some reason they are able to muster enough energy and resources— to be deduced from their  $\omega^1$  vectors as already a form of disguised redistribution— in order to convince a sufficient number of *B*'s that they should not vote for  $\omega^2$ , the radical redistribution will be 'politically feasible'. Apart from campaigning and perhaps threatening with social unrest (all the way from looting to denouncing the Constitution), both sides are then likely, *already within* the process leading up to the vote, to engage in bargaining about how much each *A* has to pay each of the *B*'s who accept to bargain so that  $\omega^1$ — or what is left over from it, presumably still sufficiently distant from  $\omega^2$ — will win the vote under whatever voting rule is inscribed in the Constitution.

Part of the process may also be one of economic persuasion : perhaps the *A*'s will attempt to convey to the *B*'s an argument according to which too much redistribution at once is actually detrimental to the *B*'s themselves, analogously to the arguments sometimes advanced for the existence of 'socially necessary exploitation'.<sup>4</sup> However, such arguments rely on the notion of *incentives* in production and in technological change; it is hard to see what the analogue would be in the Walrasian model we are discussing now: initial resources here are only means of exchanging in a Pareto-improving manner, and

---

<sup>4</sup> Roemer [1988, p. 145] puts this argument as follows: 'There are two principal ways in which the redistribution of property might render those who are putatively exploited worse off. First, this might occur as a result of the incentive effects of the new, egalitarian distribution of property. Conservatives argue that an egalitarian distribution of the capital stock deprives potential entrepreneurs of incentive, and perhaps deprives the proletarians of the whip they require to be productive. ... Second, a redistribution of property, while perhaps not affecting incentives immediately, might lead to a slower level of technological change; if so, the exploitation can be viewed as socially necessary from a dynamic point of view. In either case, if exploitation is socially necessary, then there is an argument against eliminating it.'

therefore it matters little who holds them. Economic persuasion on the part of the wealthy may therefore not be possible as a tool in the bargaining process. Could a reverse persuasion on the part of the indigent be envisaged? This would take the form of persuading the minority of *A*'s that the only legitimate way to rank the  $\varphi$ -vectors is according to a norm of distributive justice, so that all could eventually agree that  $\varphi^2$  is to be preferred to  $\varphi^1$ . Unanimity would then make possible the radical redistribution towards  $\varphi^2$  without the minority of *A*'s feeling that they have lost substantially compared to  $\varphi^1$ .

However, if this complicated and multifaceted bargaining does not suffice in persuading either a sufficient number of the majority of *B*'s that  $\varphi^1$  should be upheld or the minority of *A*'s that  $\varphi^2$  should be instituted, and if either the *A*'s had previously inscribed unanimity as a constraint in the Constitution or they refuse *ex post* the result of the majority vote in favor of  $\varphi^2$ , then armed conflict is the only way out left. If the minority of *A*'s wins the conflict through massive purchase of arms, the end result may be an  $\varphi$ -vector which— due to the expenditure in means of violence— is closer to  $\varphi^2$  than what each surviving *A* would want ; if the majority of *B*'s win the conflict thanks to massive conversion of resources into weapons, and carry out the radical redistribution, conversely the resulting (post-revolutionary)  $\varphi$ -vector may be more distant from  $\varphi^2$  than what each surviving *B* would desire.

### 4.3. Putting into question atomistic coordination

Whatever the case may be, one thing seems to us absolutely certain: there is no way in which this complicated episode can simply be 'squeezed in' between two periods of innocuous Walrasian trading— or, for that matter, between two periods of Nash-equilibrium or other atomistic interaction. Whatever the outcome of the process of criticism, bargaining, voting, and possibly even armed conflict, the agents will *emerge* from this process profoundly different from the standard monads portrayed earlier— and in fact, in order to even have engaged in such a complicated process, they must *already* have been profoundly different from the standard monads *before*, i.e., after the previous trading period which supposedly determined the initial-resource vector for the next period. Therefore, by a regressive reasoning, they can never have been such monads. In other words, bringing in the complicated and multifaceted modalities of socio-economic criticism and political deliberation at any point invalidates the whole Walrasian-market anthropology and politology— before and after the episode. *To be potential revolutionaries, the agents can never have been, and can never be in the future, standard monads*; this has far-reaching consequences.

Indeed, the third element of the anthropology and politology of socio-economic criticism and political deliberation is that the 'market mechanism' itself, which has been presupposed in all our arguments up to now, comes to be considered as a non-self-evident way of organizing interactions between the individuals. To put it differently, agents who have the individual capacity to 'look up' from their 'rationality' constraints at given initial endowments, who have the individual capacity to rank various initial endowments, and who have the collective capacity to bargain, vote or fight about what the appropriate initial endowments ought to be, are also likely to have the collective capacity to construct their basic rules of social and economic interaction. If so, then the focus of standard theory on the Walrasian or, for that matter, any other form of atomistic coordination mechanism has to be *itself justified* as the possible outcome of a procedure of collective deliberation, reciprocal persuasion, and social conflict. By making the whole model of society into an aseptic monadology, standard economists have difficulties even making sense of such questions. For them, the 'market mechanism' acts once and for all as the harmonizing principle pre-inscribed in the agents, who are quasi-empty monads devoid of any 'dispositions' and 'tendencies' except those required by... the harmonizing principle itself. The market mechanism, or any other E-M coordination

mechanism, is presented as a natural, metaphysically justified environment for social interaction— certainly not a mechanism whose rules have emerged through a process of interaction in which the agents themselves have participated.

This is a vision which can no longer be upheld after our discussion here. Take the competitive-market mechanism as a particular case of an E-M coordination mechanism. Once each individual is assumed to be able to express a preference between a competitive equilibrium *attained within C(?<sup>1</sup>)* and a competitive equilibrium *attained within C(?<sup>2</sup>)*, he must also *in principle* be assumed to be able to express a preference between a *competitive equilibrium* attained within *C(?<sup>2</sup>)* and any *alternative pattern of distribution of resources* attained within *C(?<sup>2</sup>)*.<sup>5</sup> Such agents will therefore be interested in knowing alternative, non-market mechanisms, such as generalized self-management or planning with absence of private property of endowments, in order to be able to form their judgment (whether favourable or unfavourable) about the market mechanism— one of the implications of which is that, if they do end up preferring the market mechanism, they will then *voluntarily and knowingly* act as if they were empty monads, simply because this is in their eyes the best way to organize their social interactions. This, we believe, has some unexpected but crucial implications for the teaching of economics and for the assumptions we make about the reflective consciousness of the individuals *within* the model. Once these implications are really taken into account *at the level of model-building methodology*, the fundamental epistemological stance of standard empiricist-monadological economics collapses.

#### **4.4. The teaching of economics and of ‘comparative economic systems’ as a crucial part of the economic model itself**

To grasp these crucial implications, let us go back for a moment to a distinction which Sen [1987, pp. 3-7] has suggested— namely, the distinction between economics as an ‘ethical’ approach to social questions, and economics as an ‘engineering’ approach to social problems. The engineering approach, Sen tells us, consists in ‘being concerned with primarily logistic issues rather than with ultimate ends ... the ends [being] taken as fairly straightforwardly given’ [1987, p. 4]. Through the ethical approach, in contrast, ‘economics relates ultimately to the study of ethics and that of politics, ... [so that] the study of economics, though related immediately to the pursuit of wealth, is at a deeper level linked up with other studies, involving the assessment and enhancement of more basic goals’ [1987, p. 3]. The ethical approach, Sen tells us, relates both to the individual existential level (‘How should I live?’) and to the collective political level (‘What is the common good?’).

Despite these promising distinctions, Sen’s openness collapses on itself once we realize that, on the one hand, he agrees (classically) that the ‘direct theorem’ ‘is really supremely unconcerned with distributional issues [so that] a Pareto efficient outcome may well be thoroughly unequal and nasty’ [1993, p. 536] but that, on the other hand (as already quoted), ‘the converse theorem, despite its apparently greater relevance, has, in many ways, less interest for practical economic policy than the direct theorem’ [1993, p. 535]— in fact, the direct theorem ‘ensures something rather solid here and now, even if that achievement is far from adequate’ [1993, p. 522]. When we ask ourselves what, in Sen’s mind, would relieve this ‘inadequacy’, all we get is some allusion to ‘mixed mechanisms of a kind not covered by the “fundamental theorem”’ [1987, p. 38]— presumably, the search for a second- or third-best reallocation of bundles after the market mechanism has operated,

---

<sup>5</sup> Indeed, in coherence with what we said earlier about the abusive focus of standard economics on the mutual-advantage view of society and its correlative obsession with Pareto-improvement, even the restriction to the core would have to be questioned. But although this is a generalization of our argument which would certainly be worthwhile and ultimately necessary, it is *not necessary, at this stage*, for us to drive the point we want to make.

an operation which the theory presumes to be constrained by more or less stringent 'incentive imperatives' due to the fact that too much redistribution at once is likely to cause disincentives or, at least, to encourage some wealthy agents to distort the preference information they provide to the social planner who is to construct an optimal tax scheme. This, of course, is exactly the argument with which, as we argued earlier, the minority of *A*-individuals could try to persuade the majority of *B*-individuals that revolutionizing the pattern of initial endowments may in the end go against their own interests. Which goes to show that, *in actual fact*, Sen is participating in the (often unconscious) endeavour to construct an intellectual model of society in which the very notion of revolution is simply senseless.

This is rather striking, because after all the concept of revolution which Sen invokes has to do exclusively with the distribution of initial endowments, not with the uprooting of the whole socio-economic system itself. It is thus a rather 'minimal' concept of revolution, if we compare it to what Arendt [1963, pp. 28 and 34] takes as the essence of the revolutionary phenomenon: 'The modern concept of revolution [is] inextricably bound up with the notion that the course of history suddenly begins anew, that an entirely new story, a story never known or told before, is about to unfold... . Only where this pathos of novelty is present and where novelty is connected with the idea of freedom are we entitled to speak of revolution.' But it may be precisely because his concept of revolution is so narrow that Sen succeeds in stripping it of meaning. Standard theory *de facto* has constructed and proclaims in its teachings an image of the human subject as a monad which does not for one moment desire *to found the socio-economic system in which it lives on its own freely willed agreement*. The existence of other-regardingness in the form of strategic behavior and coalition formation as it is currently conceptualized in game theory remains fully compatible with such an empiricist monadology. If we were such monads groping around with only very limited, strategic concern for what is outside of ourselves, we would indeed have no socially driven concern at all for the system in which we 'operate', since this 'operation' would be completely orchestrated and mechanically conducted by a 'monad of monads' which, in standard economics, is an uneasy mix of a *deus in machina* (not 'ex' machina) called 'the market' and a strangely all-knowing theorist who, *qua* theorist, seems to ignore that for all other purposes he is also one of these monads.

However, the individuals endowed with the three anthropological and politological properties outlined in sections 4.1, 4.2 and 4.3 are simply not such monads— and they cannot ever have been, as we have shown. Consequently, the correct model of an economy is not the Walrasian monadology, but a model in which the agents' tendency and desire to compare alternative economic systems (*i*) is an *explicit part of their being economic agents* and (*ii*) is *not reducible to the very peculiar mix of self-regard and other-regardingness* which underlies even sophisticated, game-theoretic monadologies.

This presents the following challenge: part of the agent's resources— including time, in a more elaborate model with production and, hence, 'leisure' time— may be *devoted to some form of 'permanent education'*, whether in the guise of schooling, university education, or any other arrangement— in much the same way as the theory of rational choice represents other choices such as consumption, leisure time, working time, number of children, or even suicide (see Becker [1976]). Such education, however, is not aimed at acquiring new production skills: its explicit aim is to acquire tools of judgment (ethical theories, philosophical images of Man, factual knowledge) in order to be able to rank various mechanisms for distributing resources (of which the 'market mechanism' is only one) as well as to rank various initial-endowment vectors for a given resource distribution mechanism. Let us call '*critical economic agents*' the individuals who have acquired such tools of judgment.

The crucial question now becomes the following: how likely is it that critical economic agents will recognize themselves in a teaching of economics which represents them as the contrary of what they are— namely, as self-enclosed monads obeying to an un-reflected, pre-inscribed law of coordination? In other words, how much appeal would standard economics and its E-M vision of agency have on an 'educational market' populated by critical agents seeking to rank initial-endowment bundles

and alternative socio-economic systems? This question might seem outlandish at first, but in fact it flows directly from the logic of the standard economic model itself: *a theory of rational choice must be able to go as far as including its own basic assumptions and postulated mechanisms as objects of choice on the part of the agents which this theory itself posits as rational*. If this is so, the teaching of economics and especially of 'comparative economic systems' now becomes *an endogenous part of the model itself*: standard economic theory with its E-M presuppositions has to view itself— because it will be viewed by the agents in the model— not as a 'self-evident' representation to be taught to everyone, but rather as a particular approach taught by certain agents (the academic economists professing these so-called 'standard' presuppositions) who *de facto* suggest a particular way of ranking socio-economic systems and initial-endowment vectors. The anthropology and politology underlying the First and Second Theorems of Welfare Economics turn out to be merely *one of many* ways of ranking systems and endowments. How, then, should economists treat the question of critically reflective agency, of which revolutionary agency is a special case? In the next section, we critically investigate a simple formal representation of this more encompassing framework, and we use this representation to highlight again the strong limitations of standard economic modeling and the strong demands of an alternative approach.

## 5. Revolution as the exercise of 'constituent power'

### 5.1. Can revolution be a limit case in social choice theory?

So to begin with, let us suppose that there is some list  $S = \{s_1, s_2, \dots, s_p\}$  of possible systems  $s_p$  ( $p=1, \dots, P$ ) for organizing interactions and exchanges starting from any given vector  $?$  of initial endowments. Suppose furthermore that each individual has, through various forms of 'permanent education' as outlined above, formed a judgment about the relative merits of different possible arrangements. Individual  $i$ 's *social-system meta-ranking* would then be a ranking of pairs  $(s_p, ?)$ , expressing the individual's— possibly complex and incomplete— judgments about which socio-economic system should 'best' be combined with which vector of initial endowments.

One particularly simple way (though perhaps not one which is always technically possible) of constructing this meta-ranking would be to separate the distributive judgment from the system-related judgment. In that case, individual  $i$  would have on one side a system-independent ranking  $R^2_i$  of initial-endowment vectors and on the other side an endowment-independent ranking  $R^s_i$  of socio-economic systems. The two rankings would then be applied one after the other, with the order of application depending on the individual's basic social philosophy: a *systems absolutist* will apply  $R^s_i$  before  $R^2_i$ , thus determining first his preferred socio-economic system and then determining the initial-endowment vector that will be preferable; an *endowments absolutist* will, on the contrary, apply  $R^2_i$  before  $R^s_i$ . In either case, of course, the independence of the two rankings implies that the top-ranked pair  $(s^*_i, ?^*_i)$  will be the same. More generally, we can denote by  $?(s, ?)$  any pair of a system and an endowment vector, and we can call  $?^*_i$  the top-ranked pair of individual  $i$  within his ranking  $R^2_i$  of all possible pairs.

Given this simple notation, let us see how far we can go in terms of conceptualizing revolutionary motivation and agency, and we shall also be able to see how the shortcomings of the standard economic approach become manifest. A *pre-revolutionary situation* would be the case where either (a) the prevailing voting rule does not allow for a unique collective choice  $?^*$  (the extreme case, as already discussed earlier, being the case where the voting rule is the unanimity rule and where at least two individuals  $i$  and  $j$  have different top-ranked vectors  $?^*_i$  and  $?^*_j$ ); or (b) the prevailing bargaining rule (which, as already discussed earlier, possibly includes a transfer scheme modifying the rankings of some individuals) does

not allow for an enforceable compromise between all individual rankings  $R^i$ . A *revolutionary situation* would be a situation (a) in which a coalition  $K$  of individuals agree on a pair  $z^{**}=(s^{**}, z^{**})$  which is distinct, in one of its components at least, from the currently prevailing pair  $z^*=(s^*, z^*)$  and (b) in which the currently available resources within that coalition, namely the sum of all components  $z_k^*$  for all  $k \in K$ , are large enough so that armed conflict is feasible. Finally, there would be a *revolutionary action* on the part of coalition  $K$  if (a) the *expected* post-revolution pair  $z^+=(s^+, z^+)$  is 'sufficiently close' to the *desired* post-revolution pair  $z^{**}=(s^{**}, z^{**})$  and (b) the well-known 'collective-action problem' highlighted by Olson [1965, pp. 105-110] can be solved within the coalition.<sup>6</sup>

This is essentially how a simple formal model inspired by standard economic methodology would go about conceptualizing revolutionary motivation and agency. However, how limited this first step of modeling really is can be seen when we address questions such as: what is the individuals' criterion in the construction of their ranking– is it really a broad view of social justice or absence of social oppression, or is it rather the personal gain obtainable in each pair  $z$ ? Consequently, what is the criterion used by the coalition  $K$  to 'agree' on the pair  $z^{**}$ , and more importantly still, what is the *discussion procedure* within coalition  $K$  which allows to reach this pair? How is the collective-action problem to be solved within coalition  $K$ , except by postulating a mode of coordination which is not E-M, i.e., which relies fully and explicitly on the non-selfish and non-strategic, justice- and/or oppression-oriented motivations of the individuals in  $K$ ? Answering any of these questions requires us to step outside of the basic anthropology and politicalology adopted by the standard economic model. (See, e.g., Arnsperger 1998 and Arnsperger and Varoufakis 1999.) Indeed, in order to make sense of revolutionary action, a model has to be able to present such action essentially as a non-E-M-coordinated, not exclusively self-interested action aimed at implementing through force a  $z^{**}$  which no social-choice procedure (voting, bargaining) is able to implement.

In that sense, it is indeed an action inaugurating a 'radical beginning' or a 'new origin' in the sense described by Arendt. It also shows the essentially *extra-political* character of revolution, since such an action remains defined with reference to 'the conviction that political relations in their normal course do not fall under the sway of violence, and this conviction we find for the first time in Greek antiquity, in so far as the Greek *polis*, the city-state, defined itself explicitly as a way of life that was based exclusively upon persuasion and not upon violence' (Arendt [1963, p. 12]). Thus, 'insofar as violence plays a predominant role in wars and revolutions, both occur outside the political realm, strictly speaking... . There exist events which, though they may occur in a strictly historical context, are not really political and perhaps not even connected with politics' (Arendt [1963], p. 19).

The difficulty of standard economics in even conveying the *possibility* of violent social conflict may well be due, deep down, to the fundamental avoidance of 'Hobbesian' violence and chaos to which we alluded earlier. This avoidance may have been desirable in an intellectual climate where one felt the only alternative was either social order or 'pre-social' chaos, i.e., the blind violence of wretched and lonely creatures groping in the dark and lashing out at whomever comes close to

---

<sup>6</sup> Discussing Marxian class conflict, Olson presents the collective-action problem related to revolutions as a standard *free-riding* problem: '... the absence of the sort of class action Marx predicted is due in part to the predominance of rational utilitarian behaviour. *For class-oriented action will not occur if the individuals that make up a class act rationally.* ... in both cases [for a worker as well as for a bourgeois] the individual would find that he would get the benefits of the class action whether he participated or not. ... Marxian class action then takes on the character of any endeavour to achieve the collective goals of a large, latent group. A class in Marxist terms consists of a large group of individuals who have a common interest arising from the fact that they do or do not own productive property or capital. As in any large, latent group, each individual in the class will find it to his advantage if all of the costs or sacrifices necessary to achieve the common goal are borne by others' [1965, pp. 105-106]. This free-riding argument has been used extensively by Elster [1985], while an attempt at refuting it by recourse to ethical arguments had earlier been offered by Buchanan [1979].

them, for fear of being lashed out at. But as Macpherson [1962] has remarkably demonstrated, even in this state of nature Hobbes actually had in mind a group of individuals already socialized into a 'possessive market society': '... Hobbes moved from his original physiological postulates to the conclusion that all men necessarily seek ever more power over others, by introducing assumptions which are valid only for possessive market societies. And it was only after he had established his conclusion that all men in society necessarily seek ever more power over others that he introduced his hypothetical state of nature ... . Hobbes's social assumptions were required ... for his deduction of the behaviour of man in the state of nature' [1962, pp. 68-69].

Therefore, claiming as Arendt does that it was 'the experience of wars and revolutions' which, in the 17th century, led to 'the assumption of a pre-political state, called "state of nature"' [1963, p. 19] seems unwarranted. In fact, revolutionary action as we have been analysing it here is something which occurs within a society of individuals who have *at the same time* already been socialized into a certain socio-economic system *and* who have been endowed (through the unfolding of their reflective capacities) with a capacity to judge (and perhaps rank) various alternative socio-economic systems. A revolution as we understand it is therefore not– in the way standard theorists such as Sen would want to have it, so that it is easy for them to reject the notion of revolution– a return to a 'state of nature' but *the instatement of a 'new beginning' which is seen not to be feasible through voting or bargaining within the current socio-economic system and given the current vector of initial endowments.*

## 5.2. 'Constituent power' and the all too limited relevance of standard economics

It should be clear by now that the answer to the question in our paper's title is negative: no, *homo œ conomicus* could never become a revolutionary. Individuals capable of revolutionary action can never have been *homini œ conomici*, since if they were then becoming revolutionaries would mean changing their own subjectivity in a way that requires reflective resources excluded by the standard economic model itself. As a result, this so-called 'standard' model must from now on be viewed as an extremely 'local', rudimentary and– as we shall be claiming– *degenerate* model of society: a model that might be willingly created by individuals who would be convinced (the question being how they could ever have become convinced) that acting 'as if' they were monads with narrowly circumscribed interests and 'as if' there were nothing but an invisible hand coordinating their narrowly motivated actions. The set of assumptions and results which currently present themselves as 'economics' must from now on accept to present themselves for the very specific, narrow, empiricist-monadological body of theory which they actually are.

Revolution, if and when it occurs (and, to repeat, we have not been defending it or arguing that it is always and everywhere called for), is fundamentally about what a substantial body of juridical theory has called *constituent power*, i.e., the root-source power which lies at the basis of all Constitutions and forms, so to speak, the necessary condition to even make sense of the notion of a 'Constitution'. In his discussion of constituent power, Antonio Negri [1999] has this remarkable passage, which deserves to be quoted here in full length because it illustrates very clearly what we said earlier about the necessity to include *within* the model itself the energy-consuming and time-consuming activities of the individuals who seek to critically decide which socio-economic system they find preferable. In the course of the French Revolution, Negri writes, 'the constituent power of the masses encounters the time of the bourgeoisie– that is, the organization of the time of the working day– as its obstacle. It is on this point, that the productivity of power, its economic organization, and its social power manifest themselves, for one and the other, for the bourgeoisie and the proletariat– because both rely on this awareness of

temporality (during the revolution and the further the conflict goes) to build up antagonistic class consciousness. Thus time confronts the Parisian masses as limit. Time is finite. Time must be that of the repetition of the working day. It is this obstacle of time that, on the contrary, increases the consciousness of the masses: it leads them from politics to society, from the critique of power to the critique of labour. The masses respond to the obstacle of time with sudden and formidable accelerations, accelerations that each time reach beyond the obstacle and push the limit forward. The rupture of time touches and covers more and more the social space: it breaks it and tries to invert it. Situated in temporality these accelerations of the revolutionary movement of the masses reveal the will to rupture social time, working time' [1999, pp. 197-198]. This says approximately the following: for the Parisian masses, to conduct a revolution in order to constitute a new (s, ?)-pair meant breaking away from the time constraints imposed by the dreary working rhythm prevailing under the old (s, ?)-pair constituted by the bourgeoisie, in order to be able both to formulate a 'critique of labour' and to derive from this critique the strength to 'accelerate' time in order to overrule the prevailing social order.

In this paper we have not been able, obviously, to actually suggest a formal model of revolutionary action. Our aim was more modest. We wanted to show, and we believe we have shown, that once the true and deep prerequisites for revolutionary action in an economic model have been described, the standard representation of an economy suddenly appears under a very pale light. In fact, given what we have just said about revolution as the exercise of constituent power, the only possible rationale for the standard economic model is the following: a successful revolution has been conducted by *highly reflective* individuals who, having studied numerous alternative socio-economic systems and numerous alternative forms of rational action, have concluded from this extensive study that it is best for everyone *to consciously transform themselves* into a group of unreflective monads and to bet somehow on the possibility for their actions to be coordinated through a harmonizing mechanism which they have not created, and whose commands will somehow be 'injected' into them the moment they decide to become monads. Doing this, these individuals moreover *voluntarily relinquish* all future possibilities to revolutionize their social world– it is something like a voluntary, collective lobotomy. It is indeed the case that in such a model, as Sen implicitly argues, any further idea of 'revolution' becomes senseless.

No doubt, most standard theorists will recoil in horror in front of such an account of how their favourite models might become adequate representations of social reality. No wonder, indeed, that they do not make such a background story explicit in front of their students. Neoclassical theory would, to say the least, begin to look only very dimly relevant even as an 'allegory', and the students would probably begin to ask vehemently for a methodological... revolution.

## References

- Arendt, H. [1963], *On Revolution*, New York, Viking.
- Arnsperger, Ch. [1998], 'Optimisation individuelle et engagement moral', in F.-R. Mahieu and H. Rapoport (eds), *Altruisme : Analyses économiques*, Paris, Economica, 191-214.
- Arnsperger, Ch. and Y. Varoufakis [1999], 'Beyond Altruism, Duty or Collusion: Introducing Solidarity Into Economics', IRES Discussion Paper 9930.
- Aumann, R. [1964], 'Markets with a Continuum of Traders', *Econometrica* **32**, 39-50.
- Becker, G. [1976], *The Economic Approach to Human Behavior*, Chicago, Chicago University Press.
- Buchanan, A. [1979], 'Revolutionary Motivation and Rationality', *Philosophy and Public Affairs* **9**, 59-82.
- Debreu, G. and H. Scarf [1963], 'A Limit Theorem on the Core of an Economy', *International Economic Review* **4**, 235-246.
- De Villé, Ph. [1988], 'Marché et concurrence comme fondements de l'ordre social', *Revue Européenne des Sciences Sociales* **82**, 121-30.
- Dumouchel, P. and J.-P. Dupuy [1979], *L'Enfer des choses. René Girard et la logique de l'économie*, Paris, Seuil.
- Dupuy, J.-P. [1992], *Le Sacrifice et l'envie. Libéralisme et justice sociale*, Paris, Calmann-Lévy.
- Dupuy, J.-P. [2002], 'La Cité-Machine', *Le Banquet*, no. 17, 107-119.
- Elster, J. [1985], *Making Sense of Marx*, Cambridge, Cambridge University Press.
- Girard, R. [1972], *La Violence et le sacré*, Paris, Grasset.
- Girard, R. [1982], *Le Bouc émissaire*, Paris, Grasset.
- Kolm, S.-Ch. [1995], 'The Economics of Social Sentiments. The Case of Envy', *The Japanese Economic Review* **46**, 63-87.
- Levinas, E. [1969], *Totality and Infinity. An Essay on Exteriority*, Pittsburgh, Duquesne University Press.
- Macpherson, C.B. [1962], *The Political Theory of Possessive Individualism. Hobbes to Locke*, Oxford, Oxford University Press.
- Mount, K. and S. Reiter [1974], 'The Informational Size of Message Spaces', *Journal of Economic Theory* **8**, 161-192.
- Negri, A. [1999], *Insurgencies. Constituent Power and the Modern State*, Minneapolis, University of Minnesota Press.
- Olson, M. [1965], *The Logic of Collective Action. Public Goods and the Theory of Groups*, Cambridge (Mass.), Harvard University Press.
- Renaut, A. [1987], *L'Ère de l'individu. Contribution à une histoire de la subjectivité*, Paris, Gallimard.
- Roemer, J. [1982], *A General Theory of Exploitation and Class*, Cambridge (Mass.), Harvard University Press.
- Roemer, J. [1988], *Free to Lose. An Introduction to Marxist Economic Philosophy*, Cambridge (Mass.), Harvard University Press.
- Scharfstein, B.-A. [1998], *From the Upanishads to Kant. A Comparative History of World Philosophy*, Albany, SUNY Press.
- Sen, A. [1987], *On Ethics and Economics*, Oxford, Blackwell.
- Sen, A. [1993], 'Markets and Freedoms: Achievements and Limitations of the Market Mechanism in Promoting Individual Freedoms', *Oxford Economic Papers* **45**, 519-541.