ABSTRACT

This article offers a critical analysis of four contemporary sociological debates. Systematic analysis of the relevant literature suggests the existence of a project for reformulating the methodological foundations of empirical quantitative sociology. This reconstruction shows the emergence of the following idea: an alliance of variable analysis, mechanism methodology, and simulation techniques would be of great help in resolving some of the impasses that “standard” empirical quantitative sociology encounters. This thesis is then tested for quantitative sociology of social stratification: where there are “irrefutable signs” that the idea is becoming acceptable. Lastly, a link is established between these recent debates and older, similar proposals for resolving the problem, and the question of why the soundness of these proposals has only recently been recognized is examined.

Recent studies have discussed some of the “chronic” difficulties of our discipline. There have been critiques of the sometimes radical split between theory and research (Boudon, 1997; Cuin, 2000; Goldthorpe, 1997, 2000c; Hedstrom and Swedberg, 1996, 1998b); the weakness of sociological theory has been highlighted (Coleman, 1990; Van den Berg, 1998); questions have been raised about the foundations of empirical research (Ragin and Becker, 1992) and the scientific status of sociological analysis altogether (Cuin, 2000, 2004; Passeron, 1991; Raynaud, 2006). The exchange between Raymond Boudon (2002c) and John Goldthorpe (2004) around “sociology that matters” represents a meaningful synthesis of this “reflexive effervescence” in contemporary sociology.

This article originates from similar methodological concerns, though it bears not on sociology in general but a specific research tradition, “empirical quantitative sociology”.(1)

This particular current of sociological analysis has been the focus of lively critical debate aimed at getting beyond certain fundamental limitations. The fact that such important works as Stanley Lieberson’s (1985), Charles Ragin’s (1987) or Ray Pawson’s (1989) did not suffice to curb the potential “dangers” due to naive use of techniques that were becoming increasingly sophisticated is clearly visible in the virulent renewal of the debate in the 1990s (Clogg and Haritou, 1997; Esser, 1996; Sociological Methodology, 1991, 21, pp. 291-358; Sorenson, 1998). These critical discussions began generating precise methodological proposals for improving the situation. To increase the microfoundation of empirical quantitative sociology, Hans-Peter Blossfeld (1996, 1998) recommended extending the use and analysis of longitudinal quantitative data by means of “event history analysis” techniques; Andrew Abbott (1992a, 1992b, 1995, 2000; Abbott and Hrycak, 1990; Abbott and Tsay, 2000) proposed applying the notion of “narrative” using the technique of “optimal matching analysis” originating in biology; Peter Abell (1984, 1998, 2003) agreed with the idea of substituting the notion of “variable” (see also Abell, 2004) but in contrast to Abbott, proposed an algebraic rather than metric implementation method.²

This article analyzes the recent literature that has been fueling the debate on the limitations of empirical quantitative sociology, literature which offers clear indications of a way of reformulating that approach which would extend its explanatory power. The focus here is on four types of scientific contributions: 1) those that discuss problems specific to this approach; 2) those related to the notions of social action and rationality; 3) studies of the concept of generative mechanisms; 4) publications on applying simulation methods in sociology. It may be shown that these four types of sociological production refer back to each other. My research suggests that the intersection of the four constitutes a methodological project that reformulates empirical quantitative sociology.

To anticipate the conclusions reached, it can be said that the last three areas of the debate provide general solutions to the problems raised by the first. In other words, recent literature has begun to outline a type of empirical quantitative

---

² The potential of these proposals is currently being debated. On Abbott’s, see Halpin and Chan (1998); Levine (2000); Santoro (2003); Wu (2000). The Journal of Mathematical Sociology devoted a special issue to Abell’s work (Journal of Mathematical Sociology, 1993, 18, nos. 2-3). What seems certain at least for the time being is that improvement attempts have been quite limited, in that they have not been able to surmount the descriptivism specific to variable sociology. Furthermore, the very notion of “narrative” raises more problems than it solves, as Hempel had already pointed out (1965, pp. 447-453).
sociology wherein variable analysis describes, mechanism modeling (where mechanisms are built in non-reductionist methodological individualism terms) explains, and simulation activates and runs (as well as tests) the mechanisms assumed to be the basis for observed statistical relations. I am well aware that, thus formulated, this idea is supported by only a limited part of the literature; what’s more, certain aspects of these debates –above all the material on “social simulation”– are unfamiliar to most sociologists. The view advanced here is nonetheless supported by the few articles in which the connections discussed are explicitly brought together (see Edling, 2002; Fararo, 1997; Fararo and Butts, 1999; Goldthorpe, 1999; Hedstrom, 2005, ch. 6; Hummon and Fararo, 1995).

Clearly, then, this article is concerned with the debate “inside” the quantitative approach, fueled by researchers not hostile to that approach and who recommend amending rather than abandoning it. This choice seems justifiable. As Ray Pawson points out (1989, ch. 1), many of the attacks on this approach amount to no more than a means for the critics to forge their own sociological identity, and the criticisms have not been followed by constructive proposals. The clearest manifestation of this attitude is found among the so-called “interpretive” sociologists (see, for example, Berger, 2002 and Blumer, 1956).(3) However, the choice may produce an equivocation; i.e., it might give the reader a dichotomous vision of the contemporary epistemological and methodological debate, suggesting that there are “interpretive” or “constructivist” types of sociology on the one hand, on the other quantitative approaches with “nomothetic” aims. Though I am aware of this ambiguity, I hope the plurality of authors and contributions discussed in this reconstruction will suffice to dispel any sense of equivocation or sterile opposition, which is far from my purposes here.

The article is divided into three parts. The first concerns recent written discussion of the main limitations of variable analysis. Since this debate is more familiar than the others, only the essential components of it are presented. The second is concerned with how empirical quantitative sociology may benefit from being combined with an actionist perspective, an explanatory strategy based on generative mechanisms and simulation analysis. The third discusses a specific sociological area –the sociology of social mobility– where recent literature has been changing in major ways that suggest a reformulation and enriching of the “standard” quantitative approach. The analysis concludes by suggesting how close these contemporary debates are to older ideas and outlines a response to the question of why these ideas have only quite recently been fully received.

(3) Recognizing this does not mean denying the importance of “criticism from the outside”. Such attacks have in fact helped sensitise “quantitative sociologists” to the limitations of their analyses and thus to fuel debate “within” the quantitative approach.
Three major problems with “variable analysis”

Explanatory power has not kept pace with growing technical sophistication. The most attentive recent literature readily admits this major limitation in using multivariate statistical methods to analyze social phenomena (Freedman, 1991a, 1991b).

Three groups of problems may be cited to account for this “structural discrepancy” specific to empirical quantitative sociology: 1) its a-theoreticalness; 2) its reductive conception of causation; 3) its merely partial handling of the multi-level aspect of social phenomena.

“Quantitative sociology remains very theory-poor” (Sorensen, 1998, p. 238). Freedman (1991a) suggested that this feature affects several stages of quantitative research. The first is defining the model; i.e., the structure of relations among the variables to be tested (Goldthorpe, 1996a; Sorensen, 1998). The second concerns the crucial moment of selecting and accepting the model. Many researchers have remarked that the choice of the “best model” (supposing it can be made at all) cannot be made on a purely statistical basis (Aish-Van Vaerenbergh, 1994, p. 115; Bohnstedt and Knoke, 1998; Cobalti, 1992, p. 123; Cherkaoui, 2005, ch. 4; Wunsch, 1994, p. 37). Third, empirical quantitative sociology often underestimates the role of theory during analysis of “control variables”, either in choosing the variables to be inserted or in interpreting the effects of the interactions brought to light (Sorensen, 1998).

Lastly, the justification of conditions (forms of variable distribution, error structure and relations among variables) that any “statistical model” must establish in order to be reasonably applied and correctly estimated also lacks theory (Freedman, 1991a). The many statistical tests now available do not entirely resolve the problem: statistical justification and sociological justification do not necessarily coincide. This is particularly clear for one of the most generally assumed conditions of validity in empirical quantitative sociology, namely the linearity hypothesis (Clogg and Haritou, 1997, p. 88, p. 93; Abbott, 1992a, p. 433, p. 434), which is usually justified on the basis of the parsimony principle without specifying that that principle is often only computational, not sociological (Sorensen, 1998, p. 249). The a-theoreticalness (the first more than the second) are extremely hard to justify.

(4) “It is in turn generally agreed that, far from theory being output from causal path and suchlike analysis, it is, rather, necessary input to them”, writes Goldthorpe (1996a, p. 98). “Unfortunately, sociologists over the last decades have become less, rather than more, competent at translating theoretical ideas into models to be estimated by statistical techniques”, writes Sorensen (1998, p. 239).

(5) In this connection, Cherkaoui (2005, ch. 4) and Wunsch (1994, p. 30) observe that, without theoretical reasoning, both the “strong closure hypothesis” and the “weak closure” one (the first more than the second) are extremely hard to justify.

(6) Writes Sorensen: “The introduction of independent variables as controls in a multivariate statistical model is not usually seen as specifying a theory. [...] The result is a conceptually meaningless list of variables preventing any kind of substantive conclusion” (1998, pp. 243-244).

(7) “Typically, the assumptions in a statistical model are quite hard to prove or disprove, and little effort is spent in that direction”, writes David Freedman (1991a, p. 311).
of variable analysis also assumes another form. Ray Pawson (1989) stresses the lack of theoretical thinking in choosing and justifying measurement levels of inserted variables.

The second aspect likely to be called into question is the concept of causation specific to empirical quantitative sociology. The logical core of multivariate analysis as conceptualized by Paul Lazarsfeld in the 1950s consisted in studying variations in the intensity of the tie between two variables, X and Y, produced by consecutive insertion of a series of additional variables Wn. A relatively high degree of association intensity stability is considered a sign that the variables Wn are acting neither as “intermediary” or “parasitic” variables. In empirical quantitative sociology, qualifying the link between X and Y as causal depends on this “control process”. Given the effects of Wn, if X —> Y “resists”, the hypothesis of mere correlation may be ruled out in favor of a causation hypothesis (Becker, 1992, p. 206; Ragin, 1987, pp. 58-61). This idea of “causation as robust dependence”, to use John Goldthorpe’s expression (1999, pp. 138-142; see also Hedstrom, 2003), has one major limitation, however. In contrast to experimental control based on randomly assigning cases to groups, the “statistical control” procedure does not ensure that all variables susceptible of influencing dependent variable Y are taken into account (Lieberson, 1985(8), ch. 2 and 6; Ragin, 1987, pp. 61-67). Clogg and Haritou (1997) recently drew all the relevant conclusions from this observation. Because the hypothesis that relevant non-measured variables are not correlated with the independent variable cannot be tested using the data one is in the process of analyzing, no variant of the “generalized linear model” has anything to say about the causal aspect of the relations studied.(9)

Once this major problem has been acknowledged, the image of causation as “robust dependence” is seen to have three major limitations. First, it implies a “technicist” reduction of sociological explanation. The statistical control process may lead the researcher to identify the causal effects of a variable—and therefore its explanatory power— with the value of regression coefficients or more generally with one of the “goodness of fit” measures (Abell, 1984, p. 311; Clogg and Haritou, 1997, p. 92, pp. 93-94, p. 100; Freedman, 1991a; Sorensen, 1998, p. 241, p. 243). This same identification, probably coupled with a “determinist” conception of explanation, may then lead to the logical error of imputing causation to variables instead of actors. It is in this sense that Harmut Esser (1996, p. 160, p. 162, p. 164) qualifies “variable sociology” as “meaninglessness”: it lacks the dimension of the meaning of individual action and actors’ intentions.(10) Lastly, and contrary to what might be discussed in this remarkable work.

The authors’ main thesis: “We cannot know whether the causal effect is large or small, positive or negative, present or absent without additional knowledge that cannot be obtained from the data.” (Clogg and Haritou, 1997, pp. 105-106). Similar claims are made on pp. 94, 96, 100, 103, 104.

(8) See Vallet (2004) for a careful discussion of this remarkable work.
(9) The authors’ main thesis: “We cannot know whether the causal effect is large or small, positive or negative, present or absent without additional knowledge that cannot be obtained from the data.” (Clogg and Haritou, 1997, pp. 105-106). Similar claims are made on pp. 94, 96, 100, 103, 104.
expected, multivariate analysis favors a “mono-dimensional” vision of causation to the detriment of a “pluri-dimensional or configurational” one. Some of the conditions necessary for estimating any variant of the generalized linear model require that independent variable effects be uniform regardless of the levels of the other variables; this amounts to hypothesizing causal effect independence (Becker, 1992, p. 207; Ragin, 1987, pp. 63-64). Though this is not impossible from a purely technical point of view, empirical variable sociology tends to underestimate the complex nature of causation; i.e., the multiple interactions among explanatory factors (Abbott, 1992a). (11)

Empirical quantitative sociology’s handling of the micro-macro problem is the third main focus for criticism of this approach. There are two main difficulties to using multivariate statistical techniques in this connection. First, they may lead the researcher to identify the aggregate level of analysis with the macro level, which amounts to reducing transition modes from the individual level to the macro-social level to nothing more than the “simple process” of aggregating individual actions using a logical operation of the “juxtaposition of analytic units” sort. Empirical quantitative sociology thus tends to neglect another, sociologically essential means of composing individual actions, namely the “complex aggregation mechanisms” involving interdependence of analytic units (Abbott, 1992a, p. 431, p. 434; Cherkaoui, 2005, ch. 2, 3, 6; Cuin, 2002; Esser, 1996, pp. 160-162; Hedstrom, 2003; Hedstrom and Swedberg, 1996, p. 136). (12) The second difficulty of this approach with respect to the micro-macro link has to do the fact that there is no change in level of analysis when it comes to interpreting the variable relation structure brought to light. Arthur Stinchcombe (1991, pp. 370-371) insisted that the process of inserting control variables –the core of statistical explanation– remains at the aggregated level. This means that no true understanding of the micro-individual processes responsible for the emersion of this or that structure of relations is possible. (13)

A variety of ways of underestimating theory, a reductive conception of causation, incomplete handling of the micro-macro problem—these are the

(11) “Attributes determine each other principally as independent scales rather than as constellations of attributes; main effects are more important than interactions (main effects assumption).” (Abbott, 1992a, p. 433).
(12) My distinction between “simple aggregation process” and “complex aggregation mechanisms” corresponds to Cherkaoui’s between “resultant effects” and “emerging effects” (1998, ch. 1). Cherkaoui locates both the basis for this conceptual distinction and the definition of macro-social level as emerging effect in Durkheim (see also Cuin, 1997), Max Weber (1922, p. 40) and more recently James Coleman (1986a, p. 1321, 1990, p. 5, p. 12) proposed the same definition of the macro-social level.
(13) Diffusion of longitudinal data and methods for analyzing them may be of greater use to empirical quantitative sociology in resolving this second problem than the first (i.e., taking interdependence structures into account). This is one of the major implications of Tom Snijder’s work on statistical processing of longitudinal data on social networks (1996, 1997, 2001). Snijder shows that in order for these data, which are essential for the study of interdependence structures, to be analyzable and interpretable, available statistical techniques need to be combined with “actor-oriented” simulation methods. In fact, Hans-Peter Blossfeld’s proposal (1996, pp. 191-197, 1998), recalled in the introduction, deals with only one aspect of the micro-macro problem.
major problems that a significant segment of recent literature imputes to empirical quantitative sociology. What is to be done? For Blalock (1991, p. 333) the answer is to increase the complexity of models while continuously subjecting their conditions of validity and theoretical implications to critical examination and debate. Freedman’s answer to this is: “If I am right, playing the game harder will not help. It is the rules that we need to change” (1991b, p. 357).

The analysis of recent literature presented in the following section shows that the most careful empirical quantitative sociology is starting to move in the second of these directions.

Three possible types of integration

The “language of variables” and the “language of action”

One of the richest debates in contemporary sociology concerns the form and place that a theory of action should have in sociological analysis (Marini, 1992; Déchaux, 2002). Though rational choice theory appears to be at the core of these discussions, it is in fact only one aspect of the new interest in action, the actor and rationality. The thinking of Siegwart Lindenberg on the “method of decreasing abstraction” (Lindenberg, 1992, 2003, p. 362) and the “principle of sufficient complexity” (2002, 2003, p. 362) suggests in fact that rational choice theory can be no more than the point of departure for constructing a theory of action in sociology. The simplifications effected by this theory have non-negligeable consequences on our way of conceiving and constructing the phenomenon to be explained (Lindenberg, 1998). The aim of many of Raymond Boudon’s works (1989, 1993, 1995, 1996, 1998, 1999, 2001, 2002a, 2002b, 2003) has been to show that this conception of intentional rational action could be fitted into a more general analytic framework identifiable as a type of methodological individualism that does not reduce actors’ rationality to mere instrumental or consequentialist rationality. Lindenberg’s discussion of Boudon’s cognitive rationality model is evidence that the honing of a “social” understanding of rationality has been one of the richest developments in this debate (Lindenberg, 2000; see also the exchange between Cuin [2005] and Boudon [2005a]).

(14) The American economist Gary Becker’s aim to make this specific conception of the actor and action the unifying paradigm for the social sciences (1976, 1993, 1996, 2002) surely helps explain this focus. The fact that James Coleman (1990) then directly—and successfully—took up the challenge in sociology represented a further, probably decisive pull in that direction (Bouvier, 2000; Demeulenaere, 1994; Revue Française de Sociologie, 2003, 44, 2). This led to numerous discussions of rational choice theory; see among others Abell (1992, 1996, 2001); Archer and Tritter (2001); Bohman (1992); Coleman and Fararo (1992); Elster (1986); Hardin (2001); Scheff (1992); Sociologie et Société (2002, 24, 1).
Revue française de sociologie

The fact is that whatever variant of rational choice theory or methodological individualism one decides to adopt, the shared hypothesis of an ideal-typical intentional and rational actor has great methodological appeal.

In this connection two essential points should be noted. First, this hypothesis presents an “explanatory plus” in that no supplementary condition is required once it has been established that the phenomenon to be explained results from a composition of individual intentional, rational actions (Coleman, 1986b, p. 1). This frees explanation of all black boxes (Boudon, 1998). Indeed, we can only observe that there can be no source of causation other than individual action and the reasons that inspire it, norms being only a parameter of action, not its determinant (Hedstrom and Swedberg, 1998b, pp. 11-13). Second, the reference to rational individual action plays a major role in conceptualizing the micro-macro problem (Abell, 1992; Cherkaoui, 2005, ch. 2; Friedman and Hechter, 1988; Hedstrom and Swedberg, 1996). In this respect, the actor and his/her reasons may be understood to constitute the essential link in the following sequence: “structure” —> “interaction” —> “action” —> “interaction” —> “structure”.

**Figure I. – Coleman-Boudon Diagram**

Macro (structure) <--------> Macro (structure)

[Interaction] [Interaction]

Micro (action)

James Coleman explicitly formalized this schema (1986a, p. 1322, 1990, ch. 1). It is also present in many of Raymond Boudon’s works (1977b, 1984, 1986, 2002b). Peter Abell (2003) has recently taken it up, somewhat amended it and discussed it at length. This conception of sociological explanation has been termed “structural individualism” (Lindenberg, 1977; 42

(15) This heterogeneity has been thoroughly documented; see Blossfeld (1996); Goldthorpe (1998); Opp (2002); Udehn (2001, 2002).

(16) And two others mentioned. First, the hypothesis of rational action has a “logical privilege or priority” given that it is a required point of departure for analysis because it is a stable criterion for comparison (Goldthorpe, 1998, p. 134). James Coleman and Thomas Fararo (1992, pp. xiv-xv) speak of a “principle of order” to indicate that its absence would imply a situation of logical and empirical chaos that would block the analysis. But the idea is hardly a new one. Max Weber argued along the same lines when he affirmed that verstehende sociology was rationalist above all for reasons of heuristic utility (1903-1906, p. 69, 1913, p. 306, p. 303, 1917, pp. 426-427, 1922, p. 32). Second, the point is to favor what may be qualified as “normative priority” in that actors themselves mean to be rational and claim this character for their actions (Elster, 1986, p. 26, 2001, p. 12763).


The qualifier “structural” is essential. It means that reference to the actor does not lead to a mistaken identification between micro-foundation and micro-reduction. The presence of the term “interaction” suggests that this can be avoided by taking into account the different types of interdependence structures in which actors are involved and that shape their ways of thinking as well as how their actions are conceived (Barbera, 2004, pp. 8-11, ch. 5). It is indeed interdependence structures that constitute the core of the passage from micro to macro. A genuine multilevel vision of social reality thus seems increasingly acceptable to contemporary sociologists. The “neo-structural sociology” current derived from the innovative work of Harrison White (Edling, 2002, pp. 206-208; Fararo, 1997, p. 79) seems particularly well equipped for studying and modeling this meso level because of this tradition’s particular attentiveness to social networks (Lazega and Favereau, 2002, pp. 2-11; Lazega, 2003).(18)

What does this type of debate mean for critical analysis of empirical quantitative sociology?

Taking into account the above individualist (but non-reductionist) conception of explanation when doing variable analysis would have three beneficial effects in connection with the problems just brought to light. First, focusing on actors’ intentional and rational actions would be a powerful means of bringing theory back into the research process since it is clearly the researcher’s theoretical, hypothetical reasoning that would produce a representation of the link between structure, action and statistical regularities. Second, the reference to individual action means that researchers could not proceed by attributing causation to variables and their reciprocal effects. Causation would be formulated in terms of actors, intentional actions and the interdependence among those actions. Lastly, the multilevel nature of this form of actionism would force the researcher to move from the aggregated or macro level that the statistical analysis is situated at and pay attention to the way individual actions and combinations of them could have engendered the regularities brought to light by multivariate analysis.

It must be noted that empirical quantitative sociology has begun to accept these arguments. In 1996, John Goldthorpe devoted an entire article to discussing the following thesis: “QAD clearly does need to be informed by some explicit theory of action, at all events where it is used with more than purely descriptive ambitions; and RAT [...] would appear distinctively suited (18) The expression “micro-macro” frequently used here is only a linguistic formula designed to lighten the prose. It should not lead the reader to think that I favor a simplistic or dualist vision of the analytical levels sociologists can use. The constant attention given to interdependence structures throughout this article should suffice to eliminate any such ambiguity: these structures are what found the reality and analytic relevance of the meso-social level. See Hannan (1992) for a penetrating discussion in favor of a multi-level analytical framework for analyzing complex social systems.
Revue française de sociologie

to providing accounts of the generation of the probabilistic regularities, often extensive in time and space, that QAD has the capacity to reveal.” (1996a, p. 113). In 1998, Hans-Peter Blossfeld and Gerald Prein edited a collective work devoted to this same argument. Each of the contributors expressed awareness of the necessity to link empirical quantitative sociology and action theory.

Though these ideas are certainly not shared by all, as Goldthorpe himself realistically admits (2000a, p. 20), the first move has been made. Combining “variables language” and “action language” is one way of improving empirical quantitative sociology.

The “language of variables” and the “language of mechanisms”

The second contemporary sociology debate that must be taken into consideration bears on the notion of generative mechanism (see Cherkaoui, 2005). In particularly systematic works on this argument, the approach in which this notion is the essential methodological core of both theorizing and empirical research is called “analytic sociology” (Barbera, 2003, 2004; Hedstrom, 2005).

The concept of generative mechanism is based on the idea of “generativity” (Fararo 1989, pp. 39-43; Fararo and Butts, 1999, p. 60). Attention is focused on the emergence, engendering or genesis of what is observed. Attending to the mechanism means attending to the “mode of production of phenomena” (Cherkaoui, 1998, ch. 3, 2005, ch. 4). The postulate is the same regardless of


(20) The title of the work, Rational Choice Theory and Large-Scale Data Analysis, once again expresses the connection between literature on social action and the critical debate on variable sociology. Moreover, “rational choice theory” as used by Blossfeld and Prein has quite broad meaning (1998, p. 3).

(21) Much literature on this subject has been written since the early 1990s: see Blossfeld (1996); Boudon (1998); Bunge (1998, 2004); Cherkaoui (2005, ch. 4); Elster (1989, 1998, 2003); Erikson (1998); Fararo (1989); Hechter (1998); Hedstrom (2003); Pawson (1989); Rios (2004); Schelling (1998); Sorensen (1998); Stinchcombe (1991); Van den Berg (1998). The idea of generative mechanism has been the acknowledged core of such disciplines as physics, biology and physiology since they became modern sciences (Machamer, Darden and Craver, 2000). The notion has not been so readily accepted in the social sciences, however, specifically sociology. Among the classic authors, implicit use of the idea of mechanisms seems present in Tocqueville (Boudon, 2005b; Cherkaoui, 2005, ch. 1; Edling and Hedstrom, 2005; Elster, 2003, pp. 44-48), Durkheim (Cherkaoui, 1998, ch. 3, 2005, ch. 2; Collins, 1992, ch. 6 and 11; Fararo, 1989, pp. 134-137, p. 345, p. 346), Simmel (Bunge, 1997, p. 412), as well as Weber (Cherkaoui, 2005, ch. 3; Hedstrom and Swedberg, 1998b, p. 5). Though Merton (1949, 1967) helped introduce certain analytical aspects of the notion into modern sociology, the 1960s and 1970s were the real “cradle” of the concept both epistemologically (Harré, 1972; Harré and Secord, 1972; Bunge, 1973, 1983, 1997) and in terms of actual use (Boudon, 1973, 1976, 1979a; Davidovitch and Boudon, 1964; Fararo, 1969, Schelling, 1971). See the conclusion for further discussion of this aspect.
whether mechanisms are conceived of as real entities in the world (Harré, 1972; Bunge, 1997; Fararo, 1989) or as analytic constructs (Stinchcombe, 1991; Hedstrom and Swedberg, 1998b): what is observed at “level K” must be explicable as an effect of one or several mechanisms positioned at “level K minus N”. For Y and X phenomena, whatever they may be, Figure II specifies that a mechanism is active in the emergence process of the relation as it is—namely its shape and nature—rather than on the values or behavior of variables considered separately. In other words, it would be inaccurate to conceptualize a mechanism as the equivalent of an intermediary or parasitic variable (Pawson, 1989, pp. 130-131).

FIGURE II. – Generative mechanisms and levels of reality

(“Level K”) Y ←→ X (Level of observation, description)

“producing” or “engendering” relation

(“Level K-N”) MECHANISM (Level of modeling and explanation)

In functional terms, a mechanism—or a chain of mechanisms (Gambetta, 1998)—answers the need to know how and why a relation—or a structure of relations—was engendered (Harré, 1972, p. 6, p. 118; Hedstrom, 2003).

In substantive terms, on the other hand, a mechanism is nothing other than a theoretical model constructed in terms of individual actions and interactions (Cherkaoui, 1998, ch. 3, 2005, ch. 4; Gambetta, 1998, p. 105; Hedstrom and Swedberg, 1998b, pp. 24-25; Schelling, 1998). But this individualist foundation should not be interpreted in reductionist terms (Bunge, 1997, p. 440, p. 441, p. 448, pp. 454-455, p. 457). All explanations in terms of mechanisms must work by way of three types of interdependent mechanisms: 1) “situational mechanisms” or “macro-micro mechanisms”, which model the structural components of social action; 2) “action formation mechanisms” or “micro-micro mechanisms”, aimed at modeling individual actors’ beliefs and ends; 3) “transformational mechanisms” or “micro-macro mechanisms” concerned with the process of simple or complex combining of individual actions (Hedstrom and Swedberg, 1998b, pp. 21-24). These concepts make it possible to insert the Coleman-Boudon schema (Figure I) into a more general analytic framework. Explanation by way of mechanisms implies “structural individualism”(22) (see also Manzo, 2007a).

(22) Given that mechanisms are constructed in methodological individualism terms, the debates on action theory and generative mechanisms may be said to intersect and refer to each other. The reason for this was mentioned in the preceding section: because actors alone are capable of “connecting” and “transforming” (Abell, 2004, p. 293), the only units of analysis that can claim to have causal power and significance are situated at the level of individuals and their actions.
On these bases, advocates of the mechanism notion put forward the following major proposition: there can only be explanation, and therefore even more clearly causal explanation, on condition of modeling the mechanisms that underlie the observed relations (Blossfeld, 1996; Bunge, 1997, 1998; Elster, 1998, 2003; Fararo, 1989; Hedstrom and Swedberg, 1998b; Pawson, 1989; Sorensen, 1998; Stinchcombe, 1991). Neither X’s temporal (or logical) precedence to Y nor observation of a systematic and recurrent link between them can justify attributing causation to that connection. It is the presence of a mechanism that accounts for the production of Y on the basis of the existence of X. This is what Harré (1972, p. 116, p. 121, pp. 136-137) defined as a “generative theory of causality”, in diametrical opposition to a “successionist theory of causality”. (23)

Recent literature stresses that this kind of methodological position has obvious implications for empirical quantitative sociology.

Explanation by generative mechanisms disqualifies the causal claims of variable-centered statistical explanation in that the parameters of a “statistical model” only express the intensity and the sign of the tie between Y and X without saying anything about the mechanisms responsible for producing it (Bunge, 1997; Harré, 1972; Hedstrom and Swedberg, 1998b, pp. 9-10, pp. 15-17; Cherkaoui, 2005, ch. 2; Elster, 1998, 2003). The corollary cannot be neglected either: the explanatory ambition of variable analysis gets significantly downsized, and its role in the quantitative research process comes to be understood as no more than descriptive (Goldthorpe, 1999; Hedstrom, 2003; Hedstrom and Swedberg, 1998b). (24)

Does this disqualify empirical quantitative sociology? Hardly. The language of variables can synergize very well with reasoning in terms of mechanisms. First, this combination would bring theoretical thinking back to the fore because, given that mechanisms are not observable, they have to be theoretically modeled. Second, the partial, reductive conception of causation specific to variable analysis would be corrected in that the notion of

(23) Outside sociology, the notion of “generative causality” has been put forward by a number of eminent thinkers. David Cox (1992, p. 297) suggests limiting the notion of causality “to situations where some explanation in terms of a not totally hypothetical underlying process or mechanism is available” (see also Cox and Wermuth, 1993, p. 207); A. H. Simon and Y. Iwasaki (1988, p. 150) explicitly affirm that “causality arises when a mechanism links phenomena”.

(24) This reevaluation of the status of variable analysis is “radical” in Goldthorpe’s thinking. His judgment concerns both causal analysis techniques and the most recent and sophisticated statistical methods. Of the first he affirms: “Instead of being regarded as a means of inferring causation directly from data, its primary use should rather be seen as descriptive, involving the analysis of joint and conditional distribution in order to determine no more than patterns of association (or correlation).” (1999, p. 152); of the latter: “It is important that the use of rather advanced statistical techniques for these purposes of what might be called sophisticated description should be clearly distinguished from their use in attempts at deriving causal relations directly from data analysis.” (ibid., p. 153). Hedstrom and Swedberg, no less explicit, state that we have to give up our “faith in statistical analysis as a tool for generating theories”, as well as the “belief in an isomorphism between statistical and theoretical models”. However, statistical analysis remains essential “for descriptive purposes and for testing sociological theories” (1998b, p. 17).
mechanism prevents the researcher from identifying the “statistical significance” of parameters with sociologically significant relational causation. Lastly, modeling generative mechanisms would force researchers to position themselves in a multilevel analytic framework aimed at systematically relating the individual action (micro level), the interdependence structures existing between those actions (meso level) and the products emerging from the later in terms of institutions, norms and conventions (macro level).

This is precisely the type of virtuous combination of variables and mechanisms that is explicitly proposed and discussed at length by John Goldthorpe in an article that, with a few exceptions (Barbera, 2004, ch. 7; Grusky and Di Carlo, 2001), has not received the attention I believe it deserves. “Causation, statistics, and sociology” (1999) represents a genuine manifesto for reformulating empirical quantitative sociology. The English sociologist here sketches out an “alternative for sociology” in which the set of statistical tools would be used to bring to light empirical regularities while modeling of generative mechanisms would explain the emergence of those regularities. Goldthorpe explicitly endorses a “causation as generative process” view in which, once empirical regularities have been established, the point is to “hypothesiz[e] generative processes at the level of social action” (ibid., p. 151, pp. 154-155).

Goldthorpe thus explicitly relates three objects: critical discussion of empirical quantitative sociology, methodological individualism, and generative mechanisms strategy. His ideas leave no doubt about the possibility of enriching empirical quantitative sociology by means of 1) an actionist viewpoint, 2) reasoning in terms of mechanisms.

The “language of variables” and simulation methods

Since a phenomenon can be broken down into \( n \) variables, multivariate statistical techniques represent a powerful tool for formally describing the structure of relations existing between them. But how may the reasoning in terms of mechanisms required to explain the genesis of such a structure be made operative? The literature on social science applications of simulation methods provides possible answers to this question (see also Manzo, 2004a).

Though in sociology both the idea of simulating the mechanisms underlying social phenomena and the first real simulation attempts date from the 1960s and 70s, it was not until the late 1980s that simulation methods became a focus of the real methodological and technical debate that extended...
A general definition and concise description of a specific technique will surely give a clearer idea of the nature of this approach.

Above and beyond the specificities of the various techniques, we can begin to define “simulation” by stating that it is the execution of a program that translates a theoretical system (representing an object of analysis) into a set of algorithms written in a specialized computer language. By this means, the behavior of this system may be studied and observed as it evolves dynamically under different conditions (see for example Macy, 2001, p. 14439; Moretti, 2000, p. 137; Troitzsch, 1997, p. 46; Klein, 2002-2003, p. 7; Hanneman and Patrick, 1997; Hartmann, 1996, p. 83; Heise, 1995). The complex nature of this definition is due to the fact that simulation more closely resembles a family of techniques (clearly and pedagogically presented in Gilbert and Troitzsch, 1999) that share a common methodological spirit—this is what the foregoing definition tries to convey to the reader—rather than a single method with clearly standardized operating protocols. This is why certain authors readily use the artistic metaphor when defining the nature of the activity performed by practitioners of this approach (Axelrod, 1997a; Marney and Tarbert, 2000; Whicker and Sigelman, 1991, ch. 8). To illustrate,
we can consider a type of simulation known in the literature as the “multiagent system” (Davidsson, 2002), which is particularly promising for sociological analysis (Macy and Willer, 2002; Sawyer, 2003, 2004a, 2004b, 2005) but also for economics (Phan, 2004), political science (Axelrod, 1997b; Cederman, 2001; Johnson, 1999) and the social sciences in general (Amblard and Phan, 2006). This technique which is becoming increasingly widespread also complexifies another method, the “cellular automata” (Weisbuch 1992) well-known to biologists and physicists. As the name suggests, a multiagent system is made up of a set of \( n \) elementary units (named “automata” or “agents”). The researcher can program both the behavior of these units, either singly or grouped into subsets, and the way the units (or groups of units) interact in time. The aim of the technique is to observe how the system of interaction between agents evolves and its final “emerging” configuration. Being able to study the behavior of the system by way of different actor models and/or different interdependence structures and to model the recursiveness that gets established among the different levels of analysis (actions, interactions, aggregate products of interactions, etc.) is surely what makes this method so appealing (see also Manzo, 2007a).

But does it offer methodological advantages genuinely specific to simulation or should its recent accelerated diffusion in the social sciences be explained rather by “methodological mode” logic? In other words, why should we simulate?

I see three main reasons.

First, it can be claimed that simulation has a positive impact on sociological theory (Collins, 1992; Fararo, 1989, p. 158; Hanneman, 1995; Hanneman, Collins and Mordt, 1995; Hanneman and Patrick, 1997; Troitzsch, 1997, p. 48). As the above definition suggests, all simulations require preliminary modeling. This step brings the matter of theoretical thinking in the research process back to the fore: simulation methods give theory back its function of orienting sociological analysis. Second, simulation analysis increases the theory’s degree of formalization. The operation of translating the initial theoretical propositions into a series of algorithms that the computer can read requires the researcher to specify the structure and shape of relations among the model’s analytic components (Hanneman et al., 1995, p. 3; Jacobsen and Bronson, 1997, p. 98, p. 99; Kliemt, 1996, p. 20). Likewise, the fact that the computer has to be able to read the model requires constant control of its internal logical consistency (Collins, 1992, pp. 647-648). Lastly, simulation

(29) For methodological discussions and/or empirical applications of this technique, see, for example, Bainbridge (1995); Castelfranchi (1998); Conte et al. (1997b, p. 10); Conte et al. (1998); Doran (1998); Duong and Reilly (1995); Gilbert (1996b, pp. 4-5); Gilbert and Troitzsch (1999, ch. 8 and 9); Halpin (1999, pp. 1495-1496); Johnson (1999, pp. 1522-1524, p. 1525); Macy (2001); Minar et al. (1996); Moretti (2000, 2004); Moss (1998); Phan (2004); Sichman, Conte, and Gilbert (1998); Terna (1998).
complexifies the theory in that it allows for observing the behavior of a theoretical system under a variety of initial conditions (Bainbridge, 1995, p. 483, p. 484; Fararo, 1989, p. 238; Hanneman et al., 1995; Hanneman, 1995; Hanneman and Patrick, 1997; Hegselmann, 1996, pp. 222-230).

Second, simulation methods –some more than others, of course– represent a viable technical solution for making generative mechanism reasoning operative. One is struck in reading this literature by the recurrence of such terms as “mechanism”, “process”, “underlying process”, “causal processes”, “underlying generative mechanisms” (31) And the link between simulation methods and generative mechanism modeling is often explicitly asserted (Coleman, 1965, p. 95; Fararo, 1989, p. 139; Edling, 2002, p. 213; Gilbert, 1994, 1996a, 1999a; Kliemt, 1996, p. 14, p. 16, p. 19; Schelling, 1971). (32) Specifically, the biunivocal relation between a “generative epistemology” and “multiagent systems” is being increasingly recognized and accepted (Cederman, 2005; Epstein, 2006). This appreciable property of simulation analysis derives from the very core of the technique, namely the operation of writing a program that incorporates the theoretical model to be studied –this is termed “model translation” (Whicker and Sigelman, 1991, p. 37). Writing a series of algorithms that specify how and why the variables are linked to each other amounts in fact to postulating a series of generative mechanisms. (33) “To simulate” means to engender a structure of data from a set of theoretically significant rules that one assumes to be the foundation of the phenomenon under study (Halpin, 1999, p. 1500; Hanneman, Collins, and Mordt, 1995, p. 5). Direct modeling of the mechanism in this way is made possible by the use of a formal programming language. This allows us to “drive” the mechanism by dialoguing with the computer—to observe its dynamic form in action so to speak. In this sense—and without wishing to underestimate the difficulties involved in programming (Bruderer and Maiers, 1997; Gilbert, 1996c; Heise, 1995; Troitzsch, 1996)– the computer language can help overcome certain difficulties posed by both natural and mathematical languages, since the former may imply logically unmastearable complexity while the second can present analytically unmanageable complexity (Axtell, 2000; Coleman, 1965, p. 105; Collins, 1992, pp. 643-644; Epstein, 2006; Gilbert, 1996a, p. 449, 1999a, p. 1485; Gilbert and Troitzsch, 1999, p. 6; Hanneman et al., 1995, pp. 8-9; Johnson, 1999, p. 1511, p. 1514, p. 1518, p. 1526; Troitzsch, 1997, p. 47, p. 48).


(32) Nigel Gilbert sees simulation as the mechanism method: “It aims to explicate the mechanisms of social processes and so perhaps could be called ‘process-centred analysis’.” (1996a, p. 449); “One of the benefits of computational models is that they allow the social scientist to express ideas about process or mechanism in a flexible yet precise way.” (1999a, p. 1485).

(33) See in this connection the functional definition of mechanism given above.
The third methodological advantage that helps explain the appeal of simulation methods, particularly that of “agent-centered” ones, is their ability to powerfully, flexibly model the multi-level nature of social reality (Coleman, 1965, p. 91, p. 94, p. 96; Gilbert, 1996b, p. 1, pp. 6-7, 1999a, p. 1487, 1999b; Grémy, 1977, p. 40, p. 71; Hanneman, 1995, p. 461; Macy, 2002, p. 144, pp. 147-148, p. 161). Specifically, simulation allows for studying how complex, unexpected consequences are engendered at the macro level from relatively simple situations at the micro level (Gilbert, 1996a, p. 452; Gilbert and Troitzsch, 1999, pp. 9-12). This property of some simulation methods derives from their ability to handle the problem of interdependence; that is, the way a plurality of elementary units interact reciprocally and, quite often, non-linearly (Latané, 1996, p. 290, p. 291; Kliemt, 1996, p. 20; Nowak and Lewenstein, 1996, p. 255, pp. 256-257, p. 258, p. 259). In modeling the meso level, simulation analysis proves a good candidate for improving our understanding of the micro-macro problem. Once again it is the computer language that allows for obtaining such results. Certain sophisticated types of programming—namely the family of languages defined as “object-oriented”—facilitate the construction of complex models where the behavior of each basic unit, the interactions among such units and the reciprocal references between the “local” and the “global” can be modeled with great precision.

In what sense, then, can it be claimed that empirical quantitative sociology could benefit from a closer tie to simulation methods? If we recall the problems mentioned earlier, it can be claimed that these techniques reinforce that type of sociology precisely where it is weakest. Given that variable sociology tends to underestimate the role of theory, simulation works to strengthen theoretical models. Given that the language of variables underestimates the plurality of levels specific to sociological analysis and favors linear relations, simulation methods represent a powerful technical support for handling the micro-macro problem, and by directly modeling structures of interdependence among agents, they favor a “configurational”, non-linear view of causation. Given that variable analysis requires generative mechanism reasoning to explain the empirical regularities it brings to light, simulation constitutes a tool for formally studying the mechanisms of phenomena production (Gilbert and Troitzsch, 1999). (34)

The soundness of arguments in favor of a virtuous combining of the language of variables and simulation analysis has begun to be recognized in the literature. (35)

(34) “Simulation models are concerned with processes [...] We would expect a simulation model to include explicit representations of the processes which are thought to be at work in the social world. In contrast, a statistical model will reproduce the pattern of correlations among measured variables, but rarely will it be modeling the mechanisms which underlie these relationships.” (Gilbert and Troitzsch, 1999, p. 17).

(35) Statistical analysis and simulation methods have been brought together more frequently than I can discuss in detail here; see, among others, Boudon (1977, p. 18); Coleman (1965, p. 100); Collins (1992); Gilbert (1994, 1996a, pp. 448-449); Gilbert and Troitzsch (1999, pp. 14-17); Halpin (1999, pp. 1499-1501); Hanneman (1995, pp. 459-460); Whicker and Sigelman (1991, p. 69).
Michael Macy and Robert Willer (2002, p. 162) affirm that “agent-based models use simulation to search for causal mechanisms that may underlie statistical associations”. John Goldthorpe (1999, p. 158), discussing how it might be possible to test generative mechanisms constructed at the theoretical level, notes: “The simulation approach to hypothesis testing is not at a very advanced stage. Nonetheless, there are by now at least indications that its potential in helping to integrate theoretical and empirical quantitative work is becoming more fully appreciated.” Peter Hedstrom (2005, ch. 6) seems to be explicitly moving in this direction. Lastly, it is worthwhile quoting Brendan Halpin: “This interface between statistics, simulation, and sociological theory is critically important for the development of a sociology that is both theoretically sound and empirically founded, particularly when it comes to dealing with issues that are inherently complex.” (1999, p. 1501; see also p. 1503).

Recent developments in sociology of social stratification

One research area where empirical quantitative sociology has taken an extremely complex, sophisticated form is sociology of social stratification (Cobalti, 1995; Ballarino and Cobalti, 2003; Raftery, 2001). Not by chance was it systematically chosen as an exemplary case of the limitations of variable analysis (Cherkaoui, 2005, ch. 2; Esser, 1996; Sorensen, 1998). I make the same choice here, though with the opposite purpose, my aim being to show that social stratification sociology today represents a research area where an attempt is clearly being made to revisit variable sociology with the insights discussed in the preceding sections.

The first sign of change concerns awareness and explicit acceptance of the methodological limits of variable analysis, particularly its a-theoreticalness and explanatory insufficiencies. John Goldthorpe (2000b, p. 230) writes: “Sociologists engaging in quantitative analysis of social mobility, or indeed of other macrosociological phenomena, have, I believe, often shown an insufficient appreciation of the importance of theory; and, in particular, in failing to see that such analysis, no matter how sophisticated it may be, cannot itself substitute for theory in providing explanations of the empirical findings that it produces.” Though sociologists of stratification are not always ready to acknowledge all the effects of this “overestimation” of the power and explanatory value of the multivariate statistical tools (Goldthorpe, 2003b, p. 33n), the direction of the shift in perspective is clearly outlined. This is confirmed by another fact: since at least the mid-1990s, the literature shows a number of studies that seek to nourish stratification sociology with an explanatory theory constructed specifically in terms of rational action and generative mechanisms (Barbera, 2004, p. 99, p. 148; Breen and Jonsson, 2005; Goldthorpe, 2003b, pp. 19-25; Grusky, 2001, 25). (36)

(36) This view was already being clearly outlined in the early 1970s by Raymond Boudon (1973); see the conclusion of this article.
In this connection, studies bearing on the first phase of the social position transmission process, namely “origin —> educational attainment”, are the most advanced (see Barone, 2005; Manzo, 2004b(37)). In John Goldthorpe’s first version of an explanatory theory of educational inequalities based on the rational actor hypothesis (1996b, pp.167-178), the notion of generative mechanism is implicitly present but the term “generative processes” is used only once (ibid, p. 162). A year later, Goldthorpe and Breen published “Explaining Educational Differentials: Towards a Formal Rational Action Theory” (1997), and it is here that the theory takes it full, mathematicized form, the point being to show that the empirically observed regularities “reflect action on the part of children and their parents that can be understood as rational” (ibid, p. 184, pp. 202-203). Reasoning in terms of generative mechanisms is here explicitly adopted.(38) Jan Jonsson and Robert Erikson (2000) propose an explanatory model of educational inequalities inspired in large degree by the same methodological principles.(39) The rational actor hypothesis (ibid., p. 347, pp. 358-368) and the notion of generative mechanisms (ibid., p. 347, p. 349, p. 362, p. 373) constitute the basic components for constructing a theory to explain the genesis of the empirical regularities describing the phenomenon: “One way of using the individual-decision model presented above, and the proposed mechanisms connected to the family of origin and the structure of the school system, is for trying to understand relevant empirical regularities.” (ibid., p. 368).

These theoretical works have in fact already elicited a number of empirical developments. Antonio Schizzerotto (1997) takes off from rational individual strategies to explain the low level of education participation at the upper secondary and tertiary education levels of the Italian school system. Ralf Becker (2003) uses a complex model of rational school choice to analyze the increase in school participation in Germany since the 1950s, consistently linking the languages of mechanisms and variables linking (ibid., p. 2, p. 3, p. 4, p. 6, p. 13). Steffen Hillmert and Marita Jacob (2003) adopt a rational choice perspective enriched with human capital theory to study the effects of the presence of a vocational training track on scholastic inequalities at the higher levels of the German school system. The methodological principle guiding their analysis is the following: “We use a rational-choice approach to explain persisting differentials in educational decisions by looking at the

(37) In this study I provide a detailed reconstruction of the “rational educational choice approach”, comparing it to the “cultural capital” theory that developed out of Pierre Bourdieu and Jean-Claude Passeron’s analyses. I also develop a broader-scope analysis than the one presented here of empirical literature on inequality of educational opportunity (see also Manzo and Corposanto, 2003).

(38) The section presenting the core of the model is entitled “The generation of class differentials” (ibid., p. 188); the analytic structure of the model is constructed in terms of mechanisms – “We then propose three mechanisms through which class differentials in educational attainment may arise at the level of ‘secondary effects’. “ (ibid., p. 189, p. 192), the main point of the theory being to “grasp” the “key generative processes” (ibid, p. 203).

(39) Their article, entitled “Understanding Educational Inequality: the Swedish experience’, also further develops earlier work, e.g., Erikson and Jonsson (1996), particularly the introduction and ch. 1.
cause mechanisms and generative processes of the association of social origin and educational outcomes.” (ibid., p. 321). Richard Davies, Eskil Heinesen and Anders Holm (2002) have tried to empirically assess from Danish data the relevance of the essential mechanism in the Goldthorpe-Breen model compared to the hypotheses that can be derived from the human capital theory. Richard Breen and Meir Yaish (2006) have also tried to empirically test the “relative risk aversion” mechanism on English data; to date, this study is the most direct attempt at empirically estimating the crucial generative mechanism of the Goldthorpe-Breen model (see also Holm and Jaeger, 2006). That model is also at the center of Roy Nash’s study (2003) and that of Gösta Esping-Andersen and Josep Mestres (2003). Both of these articles attempt a critique of the model with the purpose of reevaluating the empirical effect of the socialization process on the emergence of educational inequalities.

Gabriele Ballarino and Fabrizio Bernardi (2001) have tried to empirically test the Erikson-Jonsson model: their analysis on the basis of “time budget” micro-data starts with the observation that the “mechanisms” responsible for the aggregate results have not been sufficiently analyzed by empirical social stratification research.

The penetration of the generative mechanism notion can also be detected in studies of the inequalities that structure access to the job market. Barbara Reskin’s article (2003) is a veritable manifesto in this area. The title itself –“Including Mechanisms in our Models of Ascriptive Inequality”– expresses the author’s programmatic intentions. Reskin defines mechanisms as “the processes that convert inputs (or independent variables) into outputs (or dependent variables)” (ibid., p. 7), declaring that only by adopting analysis in terms of mechanisms –i.e., constructing “how-explanations” (ibid., p. 1)– can the discrepancy between technical sophistication and explanatory power in empirical research into inequalities in job market access be reduced. She suggests the need to systematically model four types of “mechanisms”: “intrapsychic”, “interpersonal”, “societal” and “organizational” (ibid., pp. 8-14).

Lastly, rational action and mechanisms are acquiring a place in some analyses of the relative aspect of social mobility. John Allen Logan (1996) outlines what he defines as the “random matching model of opportunity” (ibid., p. 175, p. 180), a microsociological model of the interaction process between two groups of actors assumed to be rational: employers and job-seeking individuals. The model formally defines the rules (ibid., pp. 177-179)

(40) Both articles use recent OECD PISA survey data (Programme for International Student Assessment; see OECD, 2001). Both Nash and Esping-Andersen stress the importance of the socialization process, particularly during the early period of children’s scholastic careers. Nash’s criticism of the Goldthorpe-Breen model is harsher than Esping-Andersen’s.

(41) As early as the 1970s, some studies –rare but exemplary– sought to introduce or recommend modeling mechanisms and social processes in this area; see in particular the pioneering research of Harrison White (1970) and Aage Sorensen’s developments of that research (1977, 1979). See further discussion in the conclusion.
and processes (pp. 179-180) that allow these two social groups to meet up with each other. Maurizio Pisati (1997) places the notion of generative mechanism at the core of a comparative empirical analysis of Italian and American mobility regimes. He explicitly accepts a “generative modeling” strategy and his stated aim is to construct “a theoretical model for the explanation of mobility regimes in terms of underlying generative mechanisms” (ibid., p. 180). (42) John Goldthorpe has also worked on constructing a theory of relative mobility in terms of rational action. In his “Outline of a theory of social mobility” (2000b), he asserts the need for a micro-foundation to explain the “intrinsic” association between origin and destination (ibid., p. 237) and recognizes the notion of “mobility strategies” as a possible means of reaching this end (ibid., pp. 238-243). (43) The methodological principle driving the theory is defined as follows: “providing theoretical accounts or narratives that can show them to be capable of generating the empirical regularities in relative mobility rates that require explanation and at the same time to be rational, and thus intelligible, responses by individuals to the situations in which they find themselves” (ibid., p. 244). An article by Mohamed Cherkaoui (2005, ch. 6) completes this framework. Cherkaoui stresses the central role that the generative mechanism notion should play if the aim of mobility studies is to construct macro-social empirical generalizations. (44)

To demonstrate the existence in quantitative social stratification sociology of a process of methodological restructuration that is being guided by the ideas discussed in the second part of this article, we need to be able to identify studies where simulation methods are coupled with variable analysis with the purpose of applying mechanism reasoning. Though there are still very few such studies, two of those cited above do use a simulation procedure: Hillmert and Jacob (2003, pp. 326-332, p. 333) and Logan (1996, pp. 186-190). These two analyses use similar logic, in that simulation analysis is used to “drive” the mathematical equations that specify the generative mechanisms of the phenomenon, given that certain parameters cannot be estimated empirically. The types of numerical simulation used are relatively simple. This type of

(42) Pisati proposes four generative mechanisms which, by influencing the micro level of individual mobility propensity, structure the association between origin and destination observed at the aggregate level. The mechanisms are 1) availability of class resources for moving within social space; 2) availability of specific resources for privileged “buying” of certain social positions; 3) desirability differential for “arrival” social positions; 4) class preferences for one or another “arrival” position (ibid., pp. 181-182).

(43) The theory outlined by Goldthorpe is strictly linked to his model of educational inequalities (mentioned above), particularly to the mechanism of “relative risk aversion” (ibid., p. 242).

(44) Cherkaoui’s argument is based on a precise definition of the macro level; in his understanding, it is only legitimate to speak of “macro” if the interdependence structure of the individual actions potentially operative in the process of emergence of the phenomenon under study has been modeled (see n. 12 above). Since currently available empirical research studies in social mobility aggregate individual data simply (see comment at the end of part 1 of this article), Cherkaoui concludes that they are not yet in a position to construct macrosociological propositions. In this sense, then, he recommends further developed modeling of micro-macro mechanisms, to use the terms of Hedstrom and Swedberg’s typology discussed above.
simulation may also be identified in two other studies, though it is used only for methodological purposes and not substantive ones. Hellevik (1997) uses a sort of numerical simulation to demonstrate that the odds-ratio does not necessarily produce the most accurate image of inequality trends (ibid., p. 376, p. 378, pp. 383-389) and that their “insensitivity to the margins” is operative under more specific conditions than sociological social stratification literature might lead us to believe (ibid., p. 394, n. 6). Jones, Wilson and Pittelkow (1990, pp. 196-199, p. 203, pp. 208-209), on the other hand, use numerical simulation to choose between different (log-linear) specifications on a single mobility table, specifications that fit the data equally well. Though these studies clearly suggest there is a space for simulation methods within stratification sociology, the fact is that the types of simulation permitting particularly strong and flexible use of mechanism reasoning remain largely unknown in this area of research. A notable exception is Moretti’s attempt (2004, pp. 131-149) to outline a “multiagent” model using LISP, the purpose being to formalize certain generative mechanisms of the Italian intra-generational mobility regime. The work begun by Chattoe and Heath (2001) likewise deserves to be closely followed. There are also studies that, though remaining outside “classic” mobility sociology, show the possibility of effectively applying “object-oriented” programming languages to social stratification problems (Duong and Reilly, 1995; Fararo and Butts, 1999, pp. 48-64).

Attentive examination of stratification sociology literature thus shows that there are signs of change in the way empirical quantitative sociology in this area is practiced. The limitations of variable analysis are being recognized; there is increasing acceptance of rational action theory and generative mechanism-centered explanatory modes as essential complements to descriptive statistical analysis; simulation methods are beginning, timidly, to claim a certain legitimacy. As I see it, these “tensions” are fully inscribed in the analytic framework of “empirical quantitative sociology revisited” discernible in the recent methodological debates reconstructed here.

* * *

This analysis of certain currents of contemporary sociological literature purports to show that a growing number of authors are converging more or less directly toward a research program that may be summed up thus: describe by means of variables —> explain by means of mechanisms —> formalize by means of simulations. Reference is to a complex image of quantitative research involving strict integration of cognitive operations (description, explanation, formalization).

(45) Thesis also put forward by John Allen Logan in the article cited above (Logan, 1996, p. 175, p. 194, p. 197, p. 198).

(46) Nathalie Bulle’s sociology of education research (1996, 1999), in which certain aspects of the French and American school systems are studied by means of numerical simulation, does not entirely fit into the framework of social stratification analysis but it should nonetheless be included among these contributions.
explanation, modeling), languages (verbal, mathematical, computational) and technical tools (statistics and simulation) that may at first sight seem hard to combine. Among the sociologists considered here, Goldthorpe proposes a particularly synthetic, explicit version of this “alliance”. He advises limiting use of multivariate statistical analysis to description (1999, p. 152, p. 153, 2000c, p. 258) and combining such analysis with a theory of action (1996a; 2000c, p. 258) and with reasoning in terms of mechanisms (1999, pp. 151-154). Lastly, he acknowledges that simulation can provide a useful basis for such reasoning (1999, p. 158).

But it can be shown that this undeniable movement within a significant part of contemporary sociology makes use of ideas that have in fact already been present in the sociological community for quite some time. We may conclude with a brief exercise in sociology of knowledge, asking why such a research program has only recently come to be recognized as relevant and legitimate.

First, it is worthwhile recalling briefly some of the main protagonists in this longer and thus more distant history. There are of course the classical sociologists, particularly—and obviously—Max Weber. Several times in the theoretical and methodological essays (Weber, 1903-1906, pp. 69-70, 81, 1913, p. 316, 1917) as well as in the first chapter of Economy and Society (1922, p. 39), Weber explicitly claims that “rule and number knowledge” should be systematically combined with “knowledge reached through interpretation” since an empirical regularity only has meaning if it can be related to the subjective activity of the actors; conversely, an interpretation can only constitute a causally accurate statement if there are accompanying proofs of the empirical regularity engendered by such activity. As Weber explains in the first paragraph of The Protestant Ethic and the Spirit of Capitalism, the empirical association between worldly occupations and the Protestant denomination has already been fully brought to light; the problem “is to explain it” (1904, p. 83n).

The idea of linking the work of demonstrating empirical regularities with that of analyzing generative mechanisms was then explicitly put forward by certain epistemologists in the early 1970s. The works of Rom Harré (1972, p. 137, p. 179, p. 183; Harré and Secord, 1972, p. 66, p. 70, p. 125) and Mario Bunge (1973, 1983) are particularly important in this connection. And it was also during the 1970s that these ideas began to appear quite explicitly among sociologists. Raymond Boudon, for example, proposed...
systematically combining statistical analysis, explanation in terms of mechanisms, and simulation methods. Methodological texts (Boudon, 1965, 1967, 1977a, 1979a, pp. 51-52, p. 62, p. 63n, 1979b, ch. 7), empirical applications (Boudon, 1973; Davidovitch and Boudon, 1964), as well as exchanges with researchers using different methodological approaches from his own (Boudon, 1976) attest to this author’s effort to develop and defend these ideas.\(^{50}\)

Other highly innovative empirical research studies dating from the same period were done by Harrison White (1970) and Aage Sorensen (1977, 1979); in these studies the methodological objective of modeling the processes and generative mechanisms underlying the macro phenomena under study was explicitly pursued and fully attained.\(^{51}\)

It is therefore undeniable that the intellectual and cognitive resources for reformulating empirical quantitative sociology have long been available to sociologists. And yet these ideas only very recently began to be discussed explicitly and systematically. Why? What explains such slow reception and diffusion?

The first explanatory factor may have to do with the influence of positivist epistemology during the process of institutionalizing the sociology discipline, above all in developing protocols for quantitative empirical analysis (Barbera, 2004, p. 14; Bunge, 1997; Cherkaoui, 2005, ch. 4). From this perspective the resistance to accepting an explanatory methodology—generative mechanisms—centered on an analytical unit—the mechanism—that by definition cannot be directly observed, measured or “made operative” is understandable. Moreover, the epistemological nature of mechanisms is not comparable to that of “laws”: the scope of a mechanism is smaller, less general, than that of a nomological proposition (Elster, 1998, p. 49, pp. 51-52, p. 62, 2003, ch. 1).

The fact is, as Cuin remarks (2003), the current intellectual contingency tends to discredit naive positive attitudes as well as nomothetic impulses. As I see it, this may have helped make the idea of generative mechanisms more acceptable to empirical quantitative sociologists.

The second point is that despite the fact that the “generative mechanism” notion and the idea of coupling it with the “variable” concept are not new, the way they were presented may have diminished their visibility and thereby slowed reception. Reading the “precursors”, one realizes that one of the following three situations holds: 1) these ideas were often discussed in an analytic and argumentative framework where they were not the author’s main focus; 2) they were applied in concrete research studies that did not discuss...
them explicitly or in any detail; 3) they were presented in narrowly diffused
texts.\(^{52}\)

Third, the problems specifically linked to simulation methods have very
likely made those methods hard to accept as a source for enriching empirical
quantitative analysis. Among those problems should be mentioned an intrinsic
difficulty of the approach: writing the program that translates the theoretical
model into a set of instructions the computer can read (Bruderer and Maiers,
1997, p. 90; Whicker and Sigelman, 1991, ch. 5). This problem has not been
entirely resolved today (Gilbert, 1994, 1996c; Gulyas, 2003; Johnson, 2003),
and it is easy to imagine the resistances it provoked in decades when
interdisciplinarity was less widespread than it is today. It is worthwhile
recalling an oft-raised objection to simulation methods, namely their
irrealism, deriving from the simplifications required if a simulation is to run
correctly (Johnson, 1999, p. 1524; Macy, 2001, pp. 14441-14443; Nowak and
Lewenstein, 1996, p. 250). It is likely that due to the effect of theories of
complexity and chaos, we are more willing today to accept the idea that
extremely complex configurations can emerge from relatively simple initial
conditions (Hegselmann et al., 1996; Nowak and Lewenstein, 1996, pp. 256-
257). Third, simulation techniques (some more than others) encounter resis-
tance from quantitative sociologists because of their weakness in the applica-
tion stage (difficulty of empirically initializing many parameters, for
example) and model assessment. Simulation practitioners do not deny the
reality of these problems (Collins, 1992, p. 649; Hedstrom, 2005, ch. 6; Boero
and Squazzoni, 2005; Whicker and Sigelman, 1991, p. 67). A less naive,
technicist vision of statistical testing —see above § 1— as well as recent prog-
ress in comparing simulated and empirical data (Moss and Edmonds, 2005;
Fararo and Butts, 1999; Logan, 1996; Snijders, 1997, 2001) —may nonetheless
have nuanced the perception of this objective difficulty. The last factor that
has probably worked to inhibit acceptance of simulation methods in sociology
has to do with their being epistemologically and technically hybrid. As
Hanneman, Collins and Mordt pointed out (1995, pp. 5-6), the simulation
approach can be ranked with neither quantitative or qualitative methods. Its
components are methodologically heterogeneous: logical and/or mathematical
formalization, attention to types of logic used by actors, modeling of interde-
pendence structures, sensitivity to context variability, etc. This may be
displeasing to purists\(^{53}\) on the quantitativist and qualitativist side alike, the
first seeing it as insufficient rigor and formalization, the second viewing it as
the latest attempt at modeling what cannot be modeled.

\(\text{\footnotesize\(^{52}\)}\) The first situation is represented by the
above-cited contributions of Weber, Harre, and
Bunge, as well as certain important ideas
developed by Merton (1949, 1967); the second
by Boudon’s \textit{L’inégalité des chances} (1973)
and certain fundamental articles by Schelling
(1971); the third by Boudon’s 1979\textsuperscript{a} article as
well as an important article by Fararo (1969).

\(\text{\footnotesize\(^{53}\)}\) This occurred in the natural sciences:
when simulation methods were introduced,
many scientists found them an inelegant,
unsophisticated type of modeling (Troitzsch,
1996).
Though partial, this configuration of elements may help us understand why the sociological community began only belatedly to take into account methodological propositions for reformulating empirical quantitative sociology that have in fact been available since the 1960s and 70s. Now that this move has been made, it is important to confirm the usefulness of this program for quantitative sociological analysis by means of empirical research practice. I have begun working on this task (see Manzo, 2006a, 2006b, 2007b) but the road ahead is a long one.

Gianluca MANZO
Université de Paris IV-Sorbonne – France
Université de Trento – Italy
glmanzo@yahoo.fr
Translation: Amy Jacobs

Previously published: RFS, 2005, 46, 1

REFERENCES


Revue française de sociologie


Gianluca Manzo


Revue française de sociologie


— 2000. – Ce que (ne) font (pas) les sociologues. Petit essai d’épistémologie critique, Geneva, Droz.


Gianluca Manzo


Gianluca Manzo


67
Revue française de sociologie


Gianluca Manzo


OECD, 2001. – Knowledge and Skills for Life, Paris, OECD.


Revue française de sociologie


— 2004b. – “Social Explanation and Computational Simulation”, Philosophical Explorations, 7, 3.


Weber M., 1903-1906. – “Roscher et Knies e i problemi logici dell’economia politica di indirizzo storico” in P. Rossi, Saggi sul metodo delle scienze storico-sociali, Milano, Edizioni di comunità [Italian translation].


