



**Mathématiques
et sciences humaines**
Mathematics and social sciences

175 | 2006
175, Varia

Actions, interactions et structure dans l'émergence de la stratification sociale des diplômes : un modèle de choix discrets avec externalités

*Actions, interactions and structure in the emergence of educational
stratification: A discrete choices model with externalities*

Gianluca Manzo



Édition électronique

URL : <http://msh.revues.org/3549>
DOI : 10.4000/msh.3549
ISSN : 1950-6821

Éditeur

Centre d'analyse et de mathématique
sociales de l'EHESS

Édition imprimée

Date de publication : 1 décembre 2006
Pagination : 53-99
ISSN : 0987-6936

Référence électronique

Gianluca Manzo, « Actions, interactions et structure dans l'émergence de la stratification sociale des diplômes : un modèle de choix discrets avec externalités », *Mathématiques et sciences humaines* [En ligne], 175 | Automne 2006, mis en ligne le 11 décembre 2006, consulté le 07 octobre 2016. URL : <http://msh.revues.org/3549> ; DOI : 10.4000/msh.3549

Ce document est un fac-similé de l'édition imprimée.

© École des hautes études en sciences sociales

ACTIONS, INTERACTIONS ET STRUCTURE DANS L'ÉMERGENCE DE LA STRATIFICATION SOCIALE DES DIPLÔMES : UN MODÈLE DE CHOIX DISCRETS AVEC EXTERNALITÉS

Gianluca MANZO¹

RÉSUMÉ – *L'article présente un modèle générateur de la stratification sociale des diplômes qui tâche de décomposer analytiquement l'émergence de celle-ci en termes de choix individuels, d'interactions directes et indirectes entre eux et de contraintes structurelles qui pèsent sur les acteurs. Le modèle proposé s'inspire de la classe de modèles dits « des choix discrets avec externalités » et se propose de raffiner par là l'une des théorisations les plus couramment adoptées dans les études quantitatives de la stratification éducative, à savoir l'approche dite « du choix éducatif rationnel ». L'article guide le lecteur à travers les étapes de construction et d'étude du modèle. Dans un premiers temps, les idées sociologiques qui fondent ce modèle sont exposées. Dans un deuxième temps, leur formalisation mathématique, d'abord, leur traduction informatique (dans une « société artificielle »), ensuite, sont disséquées. Dans un troisième temps, les résultats du traitement déductif par simulation du modèle théorique sont élaborés inductivement. D'une part, les données « simulées » sont confrontées à des données empiriques françaises et italiennes pour établir le degré de conformité entre la stratification éducative théorique et celle empirique. D'autre part, on étudie les données produites par le modèle sous des conditions expérimentales différentes de celles adoptées pour reproduire les données empiriques : le rôle de la variabilité interindividuelle dans la genèse de la stratification éducative est exploré. L'article se conclut par une discussion des limites de la modélisation et de la formalisation proposées.*

MOTS-CLÉS – Analyse statistique, Choix scolaires, Émergence, Hétérogénéité, Interdépendance, Modèles mathématiques, Systèmes multi-agents

SUMMARY – Actions, interactions and structure in the emergence of educational stratification: A discrete choices model with externalities

The article presents a generating model of educational stratification which aims analytically decompose the emergence of this one in terms of individual choices, of interactions direct and indirect between them and of structural constraints which weigh on the actors. The model suggested takes as a starting point the framework of "discrete choices models with externalities". By this one, it proposes to refine the most usually adopted theorizations in the quantitative studies of the educational stratification, namely the "rational educational choice approach". The article guides the reader through the stages of construction and study of the model. In the first time, the sociological ideas which melt this model are exposed. In the second time, their mathematical formalization, initially, their computational translation (in an agents-based system), then, are dissected. In the third time, the results of the deductive treatment by simulation of the theoretical model are elaborate inductively. On the one hand, the "simulated" data are confronted with French and Italian empirical data to establish the degree of conformity between the theoretical educational stratification and empirical one. In addition, one studies the data produced by the model under experimental conditions different from those adopted to reproduce the empirical data: namely, the role of the interindividual variability in the genesis of the educational stratification is explored. The article is concluded by a discussion about the limits of our formal model.

KEYWORDS – Agents-based models, Educational choices, Emergence, Heterogeneity, Interdependency, Mathematical models, Statistical analysis.

¹ Institut des Sciences Humaines Appliquées (ISHA), Université Paris IV-Sorbonne, 28 rue Serpente 75006 Paris, et Université de Trento, département de sociologie et recherche sociale, glmanzo@yahoo.fr

« Les grandes différences qui existent entre la méthode des sciences de la nature et celle des sciences sociales expliquent pourquoi le physicien qui se tourne vers les travaux des sociologues ressent si souvent l'impression de se trouver en compagnie de gens qui commettent habituellement tous les péchés mortels qu'il est soucieux d'éviter et qu'il n'existe pas encore de science de la société conforme à ses exigences ».

F. Hayek [1953, p. 65]

« [...] entre le régime sec, proche d'une certaine forme d'obscurantisme, d'une certaine pratique de l'« analyse des données » et l'ivresse d'une mathématique coupée de l'empirisme, il faut tenter de raison garder : entre la théorie et sa mathématisation, et l'empirisme appuyé de bonnes méthodes statistiques, les allers et retours doivent être permanents ».

M. Barbut [1994, p. 12]

1. INTRODUCTION

Dans la lignée de travaux récents en sociologie de la stratification éducative [Breen, Goldthorpe, 1997], cet article présente un modèle mathématique du choix scolaire qui s'inspire de la classe de modèles dite « modèles de choix discrets avec externalités » (cf. [Schelling, 1973 ; et, plus récemment, Nadal, Gordon, 2005 ; Phan, Payot, 2006]).

Cependant, si ce modèle mathématique constitue le point focal de l'article, il ne saurait être dissocié ni, en amont, de ce qu'il le fonde ni, en aval, de ce qu'il le justifie. La construction d'un tel objet mathématique puise en effet, d'une part, dans un problème sociologique précis (comment la stratification éducative émerge-t-elle ?), d'autre part, dans une théorie sociologique existante (l'« approche du choix éducatif rationnel ») ayant pour but de résoudre ce problème sociologique (partie 2). Notre modèle mathématique formalise donc un ensemble d'idées théoriques sur la genèse d'un phénomène social délimité, certaines d'entre elles étant disponibles en littérature, d'autres étant proposées par nous-mêmes (partie 3). Convaincus qu'un modèle n'est ni vrai ni faux, mais simplement utile, notre but principal est cependant de démontrer la pertinence du modèle construit en étudiant les conséquences. En aval, notre travail de formalisation est donc intimement lié au traitement informatique du modèle (partie 4) ainsi *in fine* qu'à l'étude statistique inductive des données qu'il produit. Ces données « simulées » seront systématiquement confrontées à des données empiriques (françaises et italiennes) dans le but de déterminer, d'une part, si le modèle est à même d'en reproduire la structure (partie 5), d'autre part, s'il est en mesure de nous apprendre des éléments nouveaux quant à certains aspects de leurs sources de production (partie 6).

Ainsi, l'article défend une conception spécifique de l'analyse sociologique selon laquelle l'on part de la construction d'un modèle théorique, on en traduit la structure dans un modèle mathématique, on met en œuvre celui-ci par des techniques de simulation appropriées pour revenir enfin aux observations à l'aide de la statistique classique. Bien qu'elle soit prônée de l'extérieur [Barbut, 1994] tout comme de l'intérieur [Halpin, 1999] de la discipline, cette intégration à l'allure circulaire entre théorie sociologique, mathématisation et traitement informatique et statistique nous paraît encore largement ignorée en sociologie.

Ce ne sont en effet que des fragments de cette chaîne dont on peut repérer la présence.

La sociologie dite « mathématique » – depuis les travaux fondateurs de Coleman [1964], de White [1963] ou de Fararo [1973] jusqu'aux bilans les plus récents [Fararo, 1997 ; Scott, 1997 ; *The journal of mathematical sociology*, 1984 ; *Sociological forum*,

1997 ; *Sociological theory*, 2000] – a endossé une démarche analytico-déductive de modélisation théorique et a fait du langage mathématique son outil privilégié d'expression [Edling, 2002 ; Fararo, 2005(a) ; Leik, 1992]². En dépit des ambitions originaires de ce courant, force est de reconnaître néanmoins que la présence la plus capillaire des mathématiques en sociologie passe plutôt par l'application des multiples méthodes statistiques qu'elles supportent [Raftery, 2001]. La notion même de modèle est d'ailleurs bien davantage associée en sociologie à la statistique multivariée qu'à la formalisation mathématique d'idées théoriques précises [Fararo, 2005(b)]. Au sein de la sociologie faisant grand usage des méthodes statistiques – parfois dite « sociologie des variables » [Esser, 1996] – les objectifs descriptifs et de quantification priment toutefois sur les ambitions de construction théorique propre à la « sociologie mathématique » [Hayes, 1984]. C'est d'ailleurs cet écart qui a conduit certains auteurs à conseiller une utilisation des méthodes statistiques en tant que moyen d'estimation des paramètres de modèles mathématiques qui, eux, formaliseraient des idées substantielles [Aberg, 2004 ; Backman, Edling, 1999 ; Hedstrom, 1992, 1994, 2000 ; Sorensen, 1976, 1998].

Quant enfin à l'univers, à présent fort hétérogène, des méthodes de simulation informatique (cf. pour une revue, [Gilbert, Troitzsch, 1999]), celles-ci n'ont jamais joué un rôle de premier plan en sociologie. Bien que leur présence puisse être repérée dès les années soixante [*Archives européennes de sociologie*, 1965 ; Coleman, 1962, 1965 ; Boudon, 1965, 1967, 1977(a), 1979(a) ; Davidovitch, Boudon, 1964 ; Grémy, 1971(a), 1971(b), 1977], la visibilité de la simulation n'a commencé à augmenter, du moins dans la sociologie anglo-saxonne, que depuis le début des années quatre-vingt-dix [*The Journal of Mathematical Sociology*, 1990 ; *American Journal of Sociology*, 2005 ; Bainbridge *et al.*, 1994 ; Carley, 1996]. Il importe de noter que la « sociologie mathématique » est au centre de ces évolutions récentes : c'est en son sein en effet que la sociologie dite « computationnelle » a vu le jour [Hummon, Fararo, 1995 ; cf. aussi, Carley, 1994, 2001 ; Fararo, Butts, 1999 ; Heise, 1995 ; Macy, 2001 ; Macy, Willer, 2002]. Elle prône la combinaison du langage mathématique avec les méthodes de simulation en raison de la grande souplesse de certains langages de programmation [Collins, 1988, annexe]. Une fois encore, cependant, le parcours circulaire qui va de la théorie aux observations n'est parcouru que partiellement. Comme on l'a justement fait remarquer [Boero, Squazzoni, 2005 ; Hedstrom, 2005, chap. 6], la « sociologie computationnelle » a pour l'heure encore tendance à privilégier le versant de l'exploration théorique au détriment du contrôle empirique des modèles. Ce qui est très clairement visible dans les écrits de Carley [1999] ou de Hanneman [1995 ; Hanneman, Patrick, 1997 ; Hanneman, Collins, Mordt, 1995].

Ainsi, au plan méthodologique, notre exercice de modélisation des choix scolaires se veut une tentative d'application plus organique que cela ne semble habituellement se faire en sociologie de la chaîne d'opérations et de langages qui va de la modélisation théorique aux observations empiriques en passant successivement par la formalisation mathématique, la simulation informatique et l'analyse statistique (cf. [Manzo, 2005]). Car, si c'est toujours une théorie que l'on mathématise [Barbut, 1994, p. 8-10 ; 2000, p. 206-207 ; Bunge, 1973, p. 131], ce sont toujours des données empiriques qu'il nous faut expliquer. Dans le cas de modèles particulièrement élaborés, leur traitement

² Précisons que l'application des mathématiques à l'étude des phénomènes sociaux est bien plus ancienne que la période que nous considérons implicitement ici – la sociologie moderne dès les années cinquante – et elle intéresse les sciences sociales plus en général. Pour un aperçu plus large, (cf. [Barbut, 2000 ; Martin, 2000, 2002 ; Martin, Barbut, 2005] ainsi que *Mathématiques & Sciences humaines* [2005, n°172]).

déductif par simulation et, ensuite, l'analyse statistique inductive des résultats auxquels ils conduisent (cf. [Moss, Edmonds, 2005]) nous semble pouvoir grandement aider à relier ces deux versants de la production des connaissances en sociologie (cf. [Manzo, 2006(a)]).

2. LE PROBLÈME ET LES HYPOTHÈSES THÉORIQUES

Le phénomène des inégalités sociales d'opportunités éducatives se présente avant tout comme un état systémique. Par là, il faut entendre que le lien, qualifiable *lato sensu* de proportionnalité inverse, entre le groupe social d'origine des individus et le diplôme qu'ils obtiennent constitue une régularité que l'on détecte au niveau de la société dans son ensemble. D'où cette régularité sociétale, fort résistante au changement, émerge-t-elle ?

Nous proposons de considérer la stratification sociale des diplômes comme l'attracteur de la dynamique d'un « système complexe adaptatif ». En dépit de l'hétérogénéité du concept de « complexité (cf. [Atlan, 1991 ; Benkirane, 2002 ; Sawyer, 2005, chap. 2 ; Simon, 1996, chap. 4 ; 1999 ; Valence, 2004, chap. 0]), ce dernier peut être défini comme un système constitué d'un grand nombre de composantes élémentaires qui interagissent de manière continue s'influencent réciproquement (cf. par exemple, [Axelrod, Cohen, 2000 ; Morgan, 2005 ; Weisbuch, 1989, 2003]). Ainsi conceptualisé, un « système complexe adaptatif » présente deux propriétés fondamentales. D'une part, aussi simple que le comportement local de ses composantes soit-il, l'issue de leurs interactions au niveau systémique peut être difficile à prédire. On dira alors que les états systémiques sont originaux ou « émergents ». D'autre part, le produit des interactions continues entre les unités élémentaires revient sur celles-ci en affectant le comportement et les interactions. C'est ce qu'on qualifie de « récursivité » ou de « rétroaction » (cf. [Morin, 1984, 1990 ; Morin, Le Moigne, 1999]).

Ainsi, penser la stratification sociale des diplômes comme étant générée par un « système complexe adaptatif » revient à poser le problème de son enracinement microsociologique. C'est là une idée ancienne en sociologie de l'éducation : « [...] *les phénomènes de stratification doivent être considérés comme la résultante non désirée de structures d'interdépendance* [...] » lit-on, dès les années soixante-dix, chez Boudon [1977(b), p. 206-207].

Dans ce cas, les composantes élémentaires du système sont les individus (et leur famille) et le comportement de ces composantes tient à la décision de participer ou non à tel ou tel niveau scolaire. Ces choix discrets binaires constituent vraisemblablement l'origine microsociologique de ce que l'on observe au niveau sociétal. Ces choix ont par ailleurs un trait spécifique : leur caractère itéré. Comme le suggère la Figure 1, les individus doivent en effet émettre un choix à tout moment où un barrage est prévu, plus ou moins explicitement, par un système scolaire donné.

Si déjà Boalt [1947] (cité dans [Jonsson, Erikson, 2000]) avait inscrit l'étude empirique des inégalités éducatives dans le cadre de cette suite de bifurcations successives (ce qui deviendra la « règle » après les travaux de Mare [1980, 1981]), c'est en revanche Boudon [1973(a)] qui – en s'inspirant d'idées plus anciennes [Keller, Zavalloni, 1964] – proposa la première théorisation des choix sous-jacents. À présent, une véritable « approche du choix éducatif rationnel » s'est constituée [Becker, 2003 ;

Breen, 2001 ; Breen, Goldthorpe, 1997 ; Breen, Yaish, 2006 ; Davies, Heinesen, Holm, 2002 ; Gambetta, 1987 ; Goldthorpe, 1996 ; Hillmert, Jacob, 2003 ; Jonsson, Erikson, 2000 ; Raftery, Hout, 1993 ; Schizzerotto, 1997] et, bien que des voix critiques se lèvent régulièrement [Esping-Andersen, Mestres, 2003 ; Nash, 2003], cette perspective explicative semble désormais constituer le cadre de référence de la plupart des études quantitatives de la stratification éducative [Barone, 2005, 2006 ; Breen, Jonsson, 2005 ; Manzo 2004].

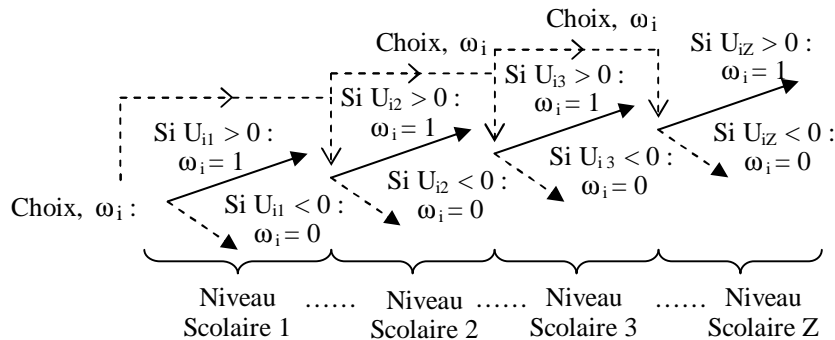


Figure 1. Itération de la prise de décision scolaire

Note de lecture. (Le symbolisme renvoie au formalisme introduit dans la section suivante, cf. § 3) ω_i = choix de l'acteur i correspondant à la décision de participer ($\omega_i = 1$) ou non ($\omega_i = 0$) au niveau scolaire D (D allant de 1 à Z) ; U_{iD} (fonction qui sera supposée numérique) = préférence de l'acteur i pour le niveau scolaire D .

Le cœur de cette approche tient à trois hypothèses essentielles.

H_1 : les acteurs évaluent subjectivement les bénéfices d'un diplôme du point de vue de leur position sociale d'origine, l'objectif de chacun d'entre eux étant d'atteindre un diplôme utile pour éviter toute expérience de mobilité sociale intergénérationnelle descendante. Il en découle l'effet inégalitaire suivant : plus l'on se situe aux rangs supérieurs de la structure sociale, plus l'on nécessitera d'instruction pour y rester, et donc plus l'on évaluera positivement tous les diplômes, notamment ceux de niveau élevé.

H_2 : les acteurs évaluent subjectivement les coûts d'un diplôme du point de vue de leur groupe social d'origine. Il en dérive l'effet inégalitaire suivant : plus l'on se situe aux rangs supérieurs de la structure sociale, plus l'on dispose de ressources pour soutenir les différents types de coûts (matériels, immatériels, cognitifs, relationnels) impliqués par le cursus scolaire, et donc les diplômes apparaîtront d'autant moins coûteux.

H_3 : les acteurs utilisent leur réussite scolaire passée pour évaluer subjectivement leurs chances de réussite future. Ce qui implique l'effet inégalitaire suivant : si la réussite scolaire passée est influencée par certaines ressources inégalement distribuées au départ, ceux qui en sont davantage dotés (notamment les individus issus des groupes supérieurs) auront une meilleure réussite et, donc, une propension plus favorable à la poursuite des études.

La perspective du « choix éducatif rationnel » soutient que l'agrégation de décisions scolaires répondant à ces trois facteurs conduit, au niveau systémique, à une stratification sociale des diplômes. Si l'on suppose des acteurs rationnels, en effet, plus

l'on monte dans la hiérarchie des groupes sociaux, plus les individus disposeront de bilans « bénéfiques/coûts » positifs ainsi que d'une réussite scolaire favorable : ils arrêteront alors plus tardivement leur carrière scolaire par rapport aux individus issus de milieux sociaux moins bien lotis.

Pour se rapprocher du concept de « système adaptatif complexe » évoqué plus haut, ces trois hypothèses paraissent cependant insuffisantes. Les acteurs de l'« approche du choix éducatif rationnel » sont en effet des acteurs isolés : les structures d'interdépendance qui les relient n'apparaissent qu'en arrière-plan. Le rôle que les choix éducatifs d'autrui joue dans la prise de décision scolaire nécessite en revanche d'être explicitement thématiqué.

Nous introduisons ainsi trois hypothèses additionnelles.

H₄ : les acteurs considèrent les choix favorables à tel ou tel diplôme émis par les individus faisant partie de leur entourage immédiat (cf. pour des preuves empiriques, [Beaud, 2003 ; Maurin, 2004]). En tant que source d'information (cognitive et/ou normative) ainsi qu'en vue de la constitution de micro-économie d'échelle, les acteurs peuvent imiter le choix éducatifs de leurs contacts. Nous mobilisons donc une conception rationnelle du mimétisme (cf. [Cherkaoui, 2006 ; Hedstrom, 1998] ; plus anciennement, [Tarde, 1890] ; en économie, [Orléans, 1999, chap. 2 ; 2002]). Selon que le contact soit socialement homophile (interaction dyadique intragroupe) ou hétérophile (interaction dyadique intergroupe), les effets inégalitaires entraînés par le choix éducatif d'autrui sont cependant différents. Dans le premier cas, les choix éducatifs ont tendance à se renforcer mutuellement tandis que, dans le second, des phénomènes d'« innovation éducative » peuvent se produire (songeons, à cet égard, à la notion de « liens faibles » originairement proposée par Granovetter [1973]).

H₅ : la prise de décision scolaire est affectée par les choix éducatifs d'autrui également de façon indirecte par l'intermédiaire des cumuls agrégés de choix qui se constituent au fil du temps. Ces cumuls ont un effet sur la valeur des diplômes : à mesure qu'un diplôme se diffuse, sa valeur décroît. C'est là une application qu'on qualifie en théorie des jeux d'interdépendance de « paramétrique » [Abell, 1996].

H₆ : les acteurs prennent en compte ce processus d'inflation des diplômes et développent une stratégie défensive qui les amène à réévaluer à la hausse les diplômes en cours de dévaluation car leur obtention est de plus en plus nécessaire. C'est là une application qu'on qualifie en théorie des jeux d'interdépendance de « stratégique » [Abell, 1996].

Il est aisé de voir alors quels sont les effets inégalitaires de ces deux formes indirectes d'interdépendance : puisque leur combinaison entraîne un allongement de la durée moyenne de la scolarité, elles permettent à tous les autres mécanismes inégalitaires de déployer leurs effets pendant un laps de temps plus étendu (cf. [Duru-Bellat, 2006]).

Ces six hypothèses nous paraissent en mesure de fonder un « système complexe adaptatif ». En effet, des milliers d'acteurs agissent selon une logique qui leur est propre (**H₁**, **H₂** et **H₃**) tout en prenant en compte les choix de leur voisinage local (**H₄**), étant affecté par l'agrégation systémique continue de ces choix locaux (**H₅**) et essayant de réagir à ceux-ci (**H₆**). Ce sera là le « Modèle du Choix Éducatif Interdépendant », MCEI dans la suite (cf. aussi, [Manzo, 2007]).

3. LA FORMALISATION MATHÉMATIQUE

Les sociologues quantitativistes de la stratification (éducative) mobilisent de plus en plus fréquemment H_1 , H_2 et H_3 pour interpréter *a posteriori* les structures de données agrégées qu'ils décrivent au préalable. Notre objectif est en revanche de faire de ces structures statistiques la conséquence de H_1 , H_2 et H_3 – complétées par H_4 , H_5 et H_6 – posées *a priori*. Dans un tel but, il est certainement opportun d'en proposer une formalisation mathématique.

Considérons alors une population P composée de N acteurs ($i = 1, 2, \dots, N$). Considérons ensuite un nombre w de groupes G dans lesquels P est divisée : chaque acteur i appartient à un et un seul de ces w groupes (les groupes G constituent donc une « partition » de P). Considérons enfin un nombre z de niveaux scolaires D qui, conformément à la Figure 1, sont séquentiellement organisés : on n'a pas accès au niveau 2, par exemple, sans avoir d'abord passé le niveau 1³.

Sous de telles conditions, admettons alors que, face à chaque niveau scolaire D qui se présente à lui, chaque acteur i doit faire le choix ω_i correspondant à la décision d'y participer ($\omega_i = 1$) ou non ($\omega_i = 0$). Ce choix sera fonction de la préférence, notée U , de l'acteur i pour le niveau scolaire D : si U_{iD} est positive, $\omega_i = 1$, sinon $\omega_i = 0$.

Le problème revient donc à modéliser cette quantité U_{iD} . Comme l'indique la formule (1), la structure générale du modèle que nous proposons fait de cette quantité une fonction de quatre termes (RO, RS, C et I) dont trois sont pesés par des grandeurs (λ , ϕ et ψ) qui sont à leur tour fonction de certains paramètres (α , β , γ et δ étant les poids – dont les valeurs seront fixés à 1 dans toutes les analyses reportées dans la suite – de la combinaison de ces quatre termes).

$$U_{iD} : f[\alpha(\lambda RO_D), \beta(\phi RS_{iGiD}), \gamma(\psi C_{iGiD}), \delta(I_{iD})] \quad (1)$$

Procédons par ordre et définissons chacun de ces termes – lesquels seront, tous, *in fine* des variables numériques – à la lumière des hypothèses exposées au paragraphe précédent⁴.

- 1) H_1 affirme que les acteurs évaluent les diplômes en vue d'en faire un instrument pour atteindre un groupe social au moins similaire à celui de leur famille d'origine. Nous distinguons alors, d'une part, RO_D – la valeur d'un niveau scolaire D sur le marché du travail (la « Rentabilité Objective », pour ainsi dire) – d'autre part, RS_{iGiD} – l'appréciation personnelle que l'acteur i donne de ce niveau scolaire D du point de vue de son groupe social d'origine G (la « Rentabilité Subjective », pour ainsi dire).

Comme indicateur de RO_D , nous retiendrons le pourcentage (empirique) d'individus qui atteignent le groupe socioprofessionnel « cadres, professions libérales, chefs d'entreprise, etc. » à partir de tel ou tel niveau scolaire D (cf. Annexe A, Tableau A). Quant à RS_{iGiD} , en l'absence de données empiriques adéquates, ses valeurs suivront une distribution de probabilité de forme log-normale et ont été construites à partir

³ Comme cela pouvait déjà se déduire de la formulation verbale de H_1 et de H_2 , nous posons ici que G tout comme D constituent un ordre total. Nous avons cependant conscience que, du point de vue sociologique, c'est là une simplification (cf. pour un choix similaire, [Breen, Goldthorpe, 1997, p. 184]).

⁴ Bien que nous ne le fassions pas pour alléger la présentation, il importe de préciser que chaque terme de la formule (1) devrait être indicé par t, le temps. Comme cela apparaîtra plus clairement dans la suite, le MCEI implique en effet un *processus*.

d'une famille de distributions de probabilité (de forme normale) dont les moyennes respectent les contraintes suivantes : 1) chaque diplôme est d'autant mieux évalué que le groupe social de l'acteur i est élevé ; 2) chaque groupe social G a une évaluation prioritaire pour un segment spécifique de la hiérarchie des diplômes (cf. Annexe B, Tableau B).

- 2) H_2 affirme qu'un acteur i évalue subjectivement les coûts d'un niveau scolaire D du point de vue des ressources auxquelles lui donne accès son groupe social d'origine G : C_{iGiD} représente cette évaluation.

Suit le même principe d'initialisation que RS_{iGiD} , mais les contraintes imposées sur les moyennes des distributions sont bien évidemment différentes. Notamment : 1) un diplôme est d'autant plus coûteux qu'il se situe aux rangs supérieurs de la hiérarchie des diplômes ; 2) ces coûts sont d'autant plus soutenables que le groupe d'où l'agent est issu se situe aux rangs supérieurs de la hiérarchie des positions sociales (cf. Annexe B, Tableau B)⁵.

- 3) H_3 affirme que les acteurs prennent en compte leur réussite scolaire passée pour estimer leurs chances de réussite future. En adoptant le nombre de redoublements, noté Nr , comme indicateur de réussite, nous imaginons un acteur qui est de plus en plus sensible à leur accumulation jusqu'à un certain seuil, au-delà duquel il l'est de moins en moins. Nous avons d'ailleurs conçu la réussite comme un facteur de pondération des coûts perçus de l'investissement scolaire : ceux-ci seront ressentis comme étant d'autant moins lourds que les résultats négatifs sont peu nombreux. En raison de nos connaissances et de sa simplicité, une fonction logistique a été choisie pour représenter ces idées⁶ :

$$\Psi = \frac{e^{\theta(Nr)i}}{1 + e^{\theta(Nr)i}} \quad (2)$$

- 4) H_4 affirme que, par delà les bénéfices (en termes de position sociale atteignable), les coûts et les risques de l'investissement scolaire, les acteurs prennent en compte également les choix scolaires des individus avec lesquels ils entretiennent des liens directs. Nous supposons alors que chaque acteur i est en contact avec un nombre k d'autres acteurs : l'ensemble de ces contacts constitue le voisinage de l'acteur i , noté V_i . Parmi ces contacts, l'acteur i comptabilisera ceux qui ont émis un choix favorable vers le niveau scolaire D : notons V_i^{FD} ce sous-ensemble de V_i . Comme l'indique la formule (3), l'influence sociale I_{iD} dont l'acteur i fera l'expérience en correspondance du niveau scolaire D sera alors calculée par « proportionnalité » :

⁵ Le choix de la distribution log-normale pour RS et C tient à la conviction que les évaluations individuelles sont plus hétérogènes que le choix de la plus habituelle distribution normale n'aurait permis de le représenter. Somme toute, il semble en effet peu réaliste de penser que les évaluations individuelles se distribuent symétriquement autour d'une valeur moyenne et que les évaluations s'éloignant de cette valeur seraient, par construction, peu probables. La distribution de Laplace-Gauss est certainement fort répandue en sociologie et, plus généralement, en sciences sociales. Historiquement, cette diffusion est probablement liée à l'influence que les idées de Quételet eurent aux débuts des applications massives de la statistique à la description des phénomènes sociaux. Une histoire fouillée de la notion de probabilité montre pourtant l'existence d'une tradition de recherche, encore peu connue, qui s'est intéressée aux phénomènes rares qui exigent des distributions de probabilités spécifiques (cf. [Crepel, Hertz, 2003]). Barbut [2003] suggère explicitement que d'autres distributions de probabilité que la « normale » pourraient utilement, peut-être plus utilement, être adoptées dans la modélisation mathématique des phénomènes sociaux. Walter [2003], par exemple, développe cette idée au sein de l'économie financière.

⁶ Le paramètre θ figurant dans la formule (2) a été inséré pour laisser la possibilité d'augmenter ou de diminuer le poids assumé par le « nombre de redoublements » dans la construction des préférences des acteurs. Dans le cadre des analyses reportées dans cet article, sa valeur sera cependant toujours égale à 1.

$$I_{iD} = \frac{V_i^{FD}}{V_i} \quad (3)$$

- 5) **H₅** affirme que les choix éducatifs des acteurs s'influencent les uns les autres également par une première voie indirecte : à travers les effets de dépréciation que la diffusion d'un niveau scolaire donné exerce sur la valeur « marchande » de ce même niveau scolaire. C'est donc RO_D (la « rentabilité objective ») qui est ici mis en cause. Soit alors NP_D le nombre d'acteurs qui possèdent, à un instant donné, le diplôme D : nous avons imaginé que la valeur de celui-ci est une fonction inverse de celui-là. Pour que cette décroissance ne soit pas cependant trop rapide, nous avons choisi, en raison de nos connaissances et de sa simplicité, de prendre le logarithme (népérien) comme forme fonctionnelle de ce processus d'inflation des diplômes scolaires. Le terme λ , figurant dans la formule (1), qui pèse RO_D peut donc s'exprimer selon la formule (4) :

$$\lambda = \frac{1}{\text{Ln}(NP_D)} \quad (4)$$

- 6) **H₆** affirme que les acteurs prennent activement en compte les effets inflationnistes dus aux choix éducatifs d'autrui. Pour les contrer, ils réévaluent à la hausse leurs appréciations du niveau scolaire en cours de dévaluation car son obtention est d'autant plus nécessaire (en vue de l'obtention d'autres diplômes qui, eux, seraient encore à l'abri d'une telle dépréciation). C'est donc RS_{iGiD} (la « rentabilité subjective ») qui est ici mise en cause. Nous postulons alors que cette réévaluation à la hausse de l'intérêt que l'on porte vers un diplôme en cours de dévaluation est une fonction croissante du nombre d'acteurs qui possèdent ce diplôme (noté plus haut NP_D). Cependant, nous supposons que cette croissance n'est ni linéaire ni continue : elle augmenterait de manière progressivement plus lente et, surtout, elle procéderait en « marche d'escalier » (ce qui traduit l'idée d'effets dépréciateurs qui se déploient dans le temps de manière différée ou retardée dès que leurs tailles est suffisamment grande pour être visible⁷). Une manière d'exprimer le terme ψ , figurant dans la formule (1), qui pèse RS_{iGiD} peut alors être la suivante⁸ :

$$\varphi = \text{Ln} \left(e + \text{ENT} \left(\frac{NP_D}{1000} \right) \right) \quad (5)$$

Ainsi, en rappelant la règle de décision et en substituant, dans l'ordre, la (4), la (5), la (2) et la (3) dans la (1), le MCEI peut s'exprimer *in fine* sous la forme suivante (6) :

$$\omega_i = 1 \mid U_{iD} > 0, \quad \text{avec} \quad (6)$$

$$U_{iD} = \alpha \left(\frac{1}{\text{Ln}(NP_D)} RO_D \right) + \beta \left(\text{Ln} \left(e + \text{ENT} \left(\frac{NP_D}{1000} \right) \right) RS_{iGiD} \right) - \gamma \left(\frac{e^{\theta(Nr)i}}{1 + e^{\theta(Nr)i}} C_{iGiD} \right) + \delta \frac{V_i^{FD}}{V_i}$$

Le MCEI fait donc de la préférence individuelle pour tel ou tel diplôme une combinaison linéaire de quatre termes dont la configuration globale rappelle un modèle classique « coûts (C_{iGiD}) / bénéfices (RS_{iGiD} , RO_D) ». Cependant, suivant certains

⁷ Cf. sur ce point, [Chauvel, 1998(a), p. 113-125 ; 1998(b)].

⁸ La valeur 1000 comme base fixe de la « division entière » a été choisie en raison de la taille des populations auxquelles nous appliquerons le modèle. La présence de la constante « e » tient à la nécessité d'avoir une valeur de RS_{iGiD} qui ne soit pas affectée tant que la première marche de la fonction n'est pas atteinte (pour $NP_D > 1000$).

modèles économiques hétérodoxes récents [Durlauf, 2001 ; Phan, Pajot, 2006 ; Steyer, Zimmerman, 2004] dont Schelling [1971, 1973] a été le pionner, la formule (6) introduit un quatrième terme (V_{iFD}/V_i) prenant en compte l'interaction directe entre les choix éducatifs des acteurs. Par rapport au cadre strictement individualiste de l'approche sociologique du choix éducatif rationnel (défini par H_1 , H_2 et H_3), c'est là une tentative d'enrichissement de la théorie. Par ailleurs, par rapport à ces modèles – dits « de choix discrets avec externalités » [Nadal, Gordon, 2005] – qui nous inspirent, nous prenons en compte aussi de façon explicite les effets de l'influence sociale indirecte, à savoir ces effets-là qui passent par la constitution d'agrégats sociaux. Selon la formule (6), cette seconde forme d'interdépendance modifie dynamiquement les bénéfices que les acteurs peuvent tirer des diplômes. Dans le cas de la composante individuelle de ceux-ci (la « Rentabilité Subjective », RS), en outre, l'effet indirect des choix éducatifs d'autrui ne s'active que par paliers successifs, ce qui constitue une application particulière de l'idée plus générale de « seuil d'activation » (cf. [Schelling, 1971, 1978 ; Granovetter, 1978, 1983, 1988]), la spécificité étant que tous les acteurs sont soumis aux mêmes paliers.

4. LA MISE EN ŒUVRE INFORMATIQUE

Le problème qui se pose à présent est celui de l'étude du MCEI. Par delà la complexité de sa formalisation, ce sont essentiellement nos objectifs théoriques qui nous éloignent d'une tentative de résolution par voie analytique. Nous souhaitons en effet établir si une population composée de dizaines de milliers d'acteurs qui agissent et interagissent selon les règles du MCEI dans le cadre d'un système scolaire séquentiellement organisé (cf. Figure 1) conduit à une stratification sociale des diplômes. Nous sommes donc moins intéressé à déterminer l'espace des solutions d'équilibre du modèle qu'à activer la dynamique contenue dans les mécanismes qui le fondent. En empruntant le concept à Fararo [1969] et à Boudon [1979(a)], nous voulons faire du MCEI un « modèle générateur », c'est-à-dire un ensemble d'hypothèses capable de générer déductivement des structures de données.

Dans un tel but, une des solutions possibles est représentée par les « systèmes multi-agents », une technique de simulation informatique qui permet de dessiner concrètement et de modéliser individuellement des entités (les agents), de les insérer dans des structures réticulaires, d'étudier leur évolution dynamique et, par là, d'établir des liens (éventuellement circulaires) entre le comportement de ces entités, leurs interactions et les résultats systémiques de celles-ci. De ce point de vue, dans le choix entre résolution analytique et simulation informatique de type « multi-agents », nous souscrivons à la remarque faite par Epstein [2005], selon qui « [...] *the issue is not whether equivalent equations exist, but which representation (equation or program) is most illuminating* » (cf. aussi [Axtell, 2000]). La diffusion croissante de cette technique en sociologie [Gilbert, 1996 ; Hedstrom, 2005, chap. 6 ; Macy, Willer, 2002 ; Moretti, 2004 ; Sawyer, 2003, 2005], en économie [Phan, 2004], en science politique [Axelrod, 1997 ; Cederman, 2001 ; Johnson, 1999] et, plus généralement, dans les sciences sociales [Amblard, Phan, 2006] semblent ainsi se répandre en raison des voies qu'elle ouvre en termes de construction de véritables « sociétés artificielles ».

Pour comprendre la nature du « système multi-agents » (stylisé en Figure 2) que nous utilisons pour « animer » le MCEI, trois éléments principaux méritent d'être brièvement discutés : les « agents », le réseau et le temps.

Les agents constituent les acteurs de notre société artificielle. Comme le précise Amblard [2003, p. 19-20, 32], un « agent » est un vecteur d'états (numériques, logiques ou les deux à la fois). Dans notre cas, ces états sont définis par la formule (6) : cette fonction qui formalise H_1 , H_2 , H_3 , H_4 , H_5 et H_6 gère le comportement de chacune des entités de notre système artificiel face à la séquence de diplômes à franchir (cf. Figure 1). Cette uniformité de fonction s'accompagne d'ailleurs d'une très grande hétérogénéité dans la valeur de ses arguments. Nos agents appartiennent en effet à des groupes G différents (cf. Figure 2) : en raison de cette appartenance groupale, ils ont des évaluations subjectives hétérogènes des bénéfices (RS dans la formule (6)) et des coûts (C dans la formule (6)) des différents diplômes. En outre, au sein même de chaque groupe, nous admettons une certaine variabilité : nos agents ne sont donc pas « aplatis » sur les valeurs moyennes de leur groupe d'origine (cf. plus haut, § 3, point 1 et 2).

Parmi les états des agents, il y en a d'ailleurs certains qui ont une importance particulière : les choix éducatifs émis par les agents avoisinants. Un système multi-agents présente en effet le grand intérêt de rendre possible la construction concrète de tels voisinages. Dans notre cas, la structure résiliaire qui supporte les interactions entre les agents se configure comme un réseau intermédiaire entre un réseau régulier et un réseau aléatoire (cf. [Watts 1999]). Notamment, pour représenter la notion théorique d'« interaction dyadique intragroupe », les agents sont disposés au sein de chaque groupe sur un réseau régulier : chacun d'entre eux est en contact direct avec un nombre moyen K d'autres agents. Ensuite, puisque le MCEI postule également l'existence d'un versant « intergroupe » de l'interaction dyadique (cf. plus haut, H_4), une certaine proportion B de liens intragroupe est débranché (aléatoirement) pour en faire des liens reliant des agents qui font partie de groupes sociaux différents⁹. Enfin, pour représenter la prédominance dans les interactions sociales de l'homophilie sur l'hétérophilie, ce rebranchement vers l'extérieur est fait, grâce à une procédure stochastique, de telle sorte que les liens intergroupe soient une fonction inversement proportionnelle de la « distance sociale » qui sépare les groupes.

À cet égard, la Figure 2, loin de donner une représentation exacte d'un tel réseau, a comme seul but de suggérer la nature socialement hétérogène de celui-ci ainsi que de visualiser les interactions entre les agents sous forme de « flux » de choix (ω_i).

Ce sont en effet ces choix qui traversent le système au cours du temps, le dernier élément essentiel constitutif de notre système multi-agents. Chez les informaticiens, on lit souvent que *la simulation consiste à faire évoluer dans le temps l'abstraction du système : le modèle* [Amblard, 2003, p. 5], (cf. aussi, [Amblard, Dumoulin, 2004 ; Ramat, 2006]). Techniquement, la simulation en tant que telle ne démarre en effet que lorsque le programme informatique traduisant le modèle est exécuté : c'est seulement à ce moment-là que le *processus* contenu dans une concaténation de mécanismes se déploie [Simon, 1996, p. 170] ; (cf. aussi, [Hartmann, 1996]). Simuler le MCEI signifie alors laisser évoluer dans le temps « informatique », c'est-à-dire au fil de plusieurs centaines d'itérations, l'ensemble d'agents dont le comportement et les interactions dépendent de la formule (6). Ce n'est qu'au fil de ces itérations que les choix éducatifs des uns affectent ceux des autres (à chaque itération 20 % des agents tirés au hasard sont mis à jour) et que ces choix sont transformés en diplômes obtenus (toutes les 52 itérations) et, donc, en cumuls agrégés de choix affectant indirectement les choix des agents qui agissent à des instants temporels successifs. Comme la Figure 2 tâche de le

⁹ Par rapport à l'algorithmique proposée par Watts (notamment, l'« algorithme β »), la différence tient ici à ce que le rebranchement aléatoire s'opère vers l'extérieur du groupe, et non en son sein.

suggérer, la mise en œuvre informatique du MCEI permet ainsi de représenter concrètement l'enchevêtrement des formes d'interdépendance impliquées dans la prise de décision scolaire : un agent peut en effet être affecté directement par le choix de son voisins et, dans le même temps, subir les effets indirects des agents ayant agi antérieurement par l'intermédiaire des cumuls de choix que ceux-ci ont contribué à générer.

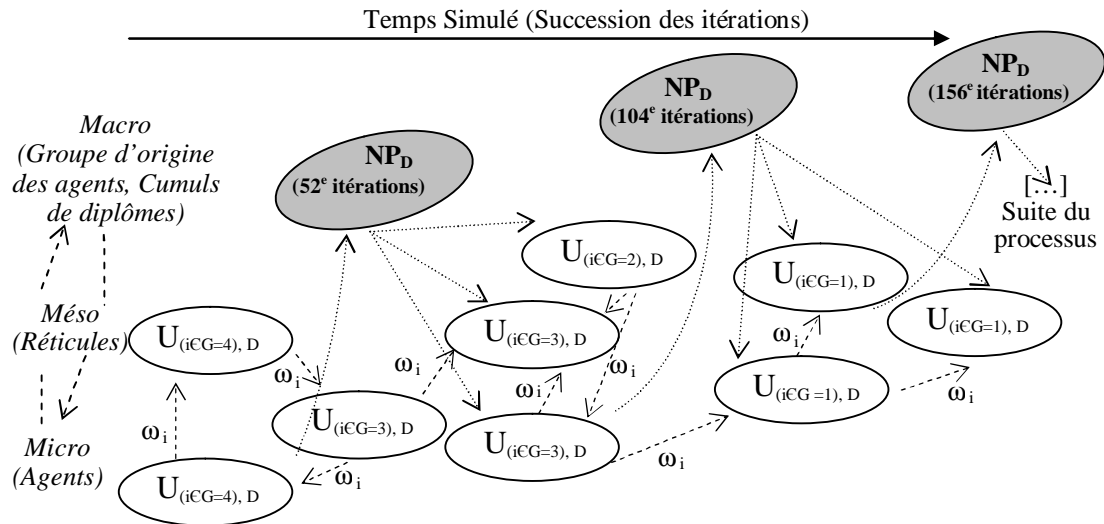


Figure 2. Stylisation du Système Multi-Agents qui met en œuvre le MCEI

Note de lecture. a) Les ellipses blanches représentent les agents issus de différents groupes G : chacun d'entre eux est caractérisé par une valeur de U_{iD} géré par la formule (6) ; b) les flèches qui relient ces ellipses représentent des interactions dyadiques entre les agents ; c) les choix de ces derniers (ω_i) se transmettent à travers ces interactions dyadiques ; d) les flèches ascendantes (en pointillé) représentent l'agrégation de ces choix en cumul d'obtentions éducatives (NP_D dans la formule (6), ellipses grises ici) tandis que les flèches descendantes (en pointillé) représentent les effets de rétroaction de ces cumuls de choix sur la fonction d'utilité des agents (notamment, sur la composante « rentabilité objective » et « rentabilité subjective » des diplômes (cf. la formule (6)) ; e) ces boucles se déploient au cours du temps informatique, c'est-à-dire grâce à une suite d'itérations).

Conseil de lecture. On consultera cette figure en s'accompagnant de la Figure 1 afin de percevoir la profondeur du processus par rapport à la suite de transitions conditionnelles dans lesquelles s'articule le système des diplômes sous-jacents au système multi-agents.

Ainsi, la décomposition analytique opérée par H_1 , H_2 , H_3 , H_4 , H_5 et H_6 de la genèse de la stratification sociale des diplômes en termes d'actions, de réseaux et de structure trouve un support concret dans l'infrastructure informatique de notre système multi-agents. D'un point de vue sociologique, celui-ci rend opérationnel une forme complexe d'individualisme méthodologique [Dupuy, 1992, 1997] qu'on a parfois qualifiée de « structurelle » [Barbera, 2002 ; Boudon, 1983 ; Boudon, Bourricaud, 1982(a) ; Lindenberg, 1977 ; Raub, 1982 ; Udehn, 2001, 2002 ; Wippler, 1978, 1985]. L'évolution dynamique de ce système, rendue possible par la simulation (cf. Figure 2), permet d'ailleurs de démultiplier dans le temps la boucle « Macro \rightarrow Micro \rightarrow Macro » sous-jacente à cette forme complexe d'individualisme [Coleman, 1986, 1990] faisant *in fine* du MCEI une application de ce qu'on pourrait appeler l'« extension temporelle » du « Coleman Boat » [Abell, 1996, 2003 ; Coleman, 1993].

5. LES DONNÉES SIMULÉES ET LES DONNÉES EMPIRIQUES

En introduisant cet article, nous avons avancé l'idée selon laquelle la formalisation d'un modèle ne devrait pas être une fin en soi, mais plutôt un moyen voué à la compréhension de structures spécifiques de données issues de l'observation.

C'est donc aux conséquences du MCEI que nous nous tournons maintenant dans le but de déterminer si ses hypothèses constitutives peuvent prétendre à une « présomption de validité et de réalisme » dans l'explication de la stratification sociale des diplômés existant en France et en Italie dans les années 90¹⁰.

Le Tableau 1 reporte les données générées par le système multi-agents animé par H_1 , H_2 , H_3 , H_4 , H_5 et H_6 et les compare aux données empiriques françaises et italiennes¹¹.

Une première indication de la proximité globale existant entre les données « simulées » et les données empiriques vient de l'« écart absolu moyen ». Cette statistique (calculée ici comme $\sum_{ij} |f_{ij} - F_{ij}| / 2N$, f étant la fréquence empirique et F celle théorique) vaut 2.54 pour la France et 1.76 pour l'Italie. De manière très générale, donc, il semblerait que le « système complexe adaptatif » où chaque agent se comporte selon la formule (6) produit une structuration sociétale des diplômés largement superposable à celle que l'observait en France et en Italie dans les années quatre-vingt-dix.

¹⁰ Les données françaises sont issues de l'*Enquête Formation-Qualification professionnelle* (FQP), réalisée en 1993 par l'INSEE sur un échantillon (aléatoire) représentatif de la population française (cf. site du Centre Maurice Halbwachs). En raison des particularités de la procédure d'échantillonnage adoptée, nous avons élaboré les données après les avoir pondérées, notamment en faisant comme s'il s'agissait d'un échantillon aléatoire simple (cf. [Vallet, 2004]). Les données italiennes proviennent en revanche de la deuxième vague de collecte, réalisée en 1999, de l'enquête (réalisée également par échantillonnage aléatoire) dite ILFI (*Indagine longitudinale sulle famiglie italiane*, cf. <http://www.sociologiadip.unimib.it/ilfi/>). Quant aux nomenclatures des groupes socioprofessionnels et des diplômes, nous avons jugé opportun de recoder les nomenclatures française et italienne dans une classification commune, notamment les classifications dites CASMIN (cf. respectivement, d'une part, [Goldthorpe, Llewellyn, Payne, 1980, p. 39-42 ; Erikson, Goldthorpe, 1992, p. 35-47 et Breen, 2004, p. 9-14], d'autre part, [Koning, Luttinger, Muller, 1988 ; Muller, Luttinger, Konig, Karle, 1989 et Muller, Karle, 1993]). Précisons enfin que, pour ne pas démultiplier excessivement le nombre de paramètres de nos simulations, nous nous sommes contentés d'une recodification en quatre postes des nomenclatures CASMIN (cf. du moins pour les groupes socioprofessionnels d'origine, [Erikson, Goldthorpe, 1992, p. 39 et Breen, 2005, p. 41]).

¹¹ Précisons que la « société artificielle » ayant produit ces données a les caractéristiques suivante : a) une densité de liens relativement limitée (chaque agent a, en moyenne, 4 voisins) ; b) une probabilité d'établir des liens intergroupe relativement basse, mais non nulle (10 % environ des liens intragroupe sont déconnectés et rebranchés vers l'extérieur) ; c) une probabilité bien plus grande que ces liens relient des individus socialement proches que socialement éloignés. Quant aux valeurs de RO, RS et C, le lecteur peut en revanche se reporter aux Tableaux A et B en annexe. Pour des preuves empiriques du bien-fondé du point c), on pourra d'ailleurs consulter : 1) pour ce qui est de l'homogamie sociale, en Italie, Schizzerotto [1990], Cobalti, Schizzerotto [1994, chap. 5] et Bernardi [2002] et, en France, Thélot [1982, chap. 8] et Forsé, Chauvel [1995] ; 2) au plan international, pour ce qui est de l'homogamie éducative, Smith *et al.* [1998] ; 3) au sujet de l'homophilie sociale, Cousteaux, Lemel [2004] et Godechot [2000] et, plus en général, Mercklé [2004, chap. 3] et Degenne, Forsé [2004, chap. 2].

	1a	1b+1c	2a+2b	2c	3a+3b	Total
Distribution simulée (France)						
	188.1	426.4	520.6	495.35	1290.55	2921
	730	1253.75	837.9	422.2	429.15	3673
	1210.5	1596	662.35	307.65	332.5	4109
	1578.1	2236.5	787.15	264.85	206.4	5073
	3706.7	5512.65	2808	1490.05	2258.6	15776
Distribution empirique (France, Enquête FQP 1993)						
I + II (1)	196	402	557	492	1274	2921
IIIa + IIIb (2)	747	1331	710	455	430	3673
IV (3)	1121	1661	632	304	391	4109
V + VI + VIIa,b (4)	1576	2174	746	316	261	5073
Total	3640	5568	2645	1567	2356	15776
Distribution simulée (Italie)						
I + II (1)	4.8	35.4	101.15	397.3	333.35	872
IIIa + IIIb (2)	10.45	113.65	215.65	455.6	194.65	990
IV (3)	318.4	860.4	1058.45	590.7	152.05	2980
V + VI + VIIa,b (4)	393.35	1128.05	1508.05	618.65	75.9	3724
Total	727	2137.5	2883.3	2062.25	755.95	8566
Distribution empirique (Italie, Enquête ILFI 1999)						
I + II (1)	3	26	125	393	325	872
IIIa + IIIb (2)	7	80	251	455	197	990
IV (3)	321	874	967	612	206	2980
V + VI + VIIa,b (4)	365	1104	1436	677	142	3724
Total	696	2084	2779	2137	870	8566

Tableau 1. Distribution du diplôme le plus haut obtenu (en colonne) selon le groupe social d'origine (en ligne). *Fréquence absolue* produites par les systèmes multi-agents mettant en œuvre le MCEI (moyenne de 20 réplifications) et fréquences absolues observées

Note de lecture. Groupes sociaux : I+II = « cadres supérieurs, professions libérales, chefs d'entreprise » (1) ; IIIa+IIIb = « professions intermédiaires » (2) ; IV = « autonomes » (3) ; V+VI+VIIa,b = « ouvriers » (4). *Diplômes* : 1a = sans diplôme ; 1b+1c = diplômes de type CEP ou CAP ; 2a+2b = diplômes de type BEPC ou BEP ; 2c = diplômes secondaires longs ; 3a+3b = diplômes de l'enseignement supérieur

Pour mieux en apprécier la configuration, il paraît indispensable d'élaborer de façon inductive les données produites par nos sociétés artificielles et de les confronter, point par point, aux mêmes élaborations opérées sur les données empiriques. Selon une distinction classique en sociologie de la mobilité et, omniprésente dans les études quantitatives des inégalités scolaires, cette analyse inductive des données simulées sera ici opérée dans une direction double. Dans un premier temps, nous étudierons les données d'un point de vue « absolu » ; successivement, nous les décrirons sous leur profil « relatif »¹².

¹² Pour ce qui est des couples conceptuels « mobilité structurelle/mobilité nette » et, plus récemment, « mobilité absolue/mobilité relative (ou fluidité sociale) », l'on consultera, entre autres, [Boudon, 1973(a), p. 263-264 ; 1973(b) ; Boudon, Bourricaud, 1982(b) ; Cherkaoui, 1989, 1992 ; Cobalti, 1989(b), (c) ; 1995, chap. 3 ; Cuin, 1993, chap. 3 ; Goldthorpe, Llewellyn, Payne, 1980, chap. 3 ; Erikson, Goldthorpe, 1992, chap. 2 ; 2002 ; Lemel, 1991, p. 156-157 ; Merllié, 1994, chap. V ; Merllié, Prévot, 1997, p. 24-27 ; Thélot, 2003, p. 25 ; Vallet, 1989]. Quant aux études quantitatives des inégalités

5.1 LES MULTIPLES FACETTES DE LA DISTRIBUTION DES DIPLÔMES

Commençons alors par regarder le volume des différentes catégories de diplômes qu'on pourrait qualifier de « morphologie de l'éducation ». La Figure 3 montre que, conformément à la distribution observée des diplômes en France, la « société artificielle française » a engendré une distribution « hypertrophique » aux niveaux inférieurs de la hiérarchie des diplômes. Le système multi-agents italien a en revanche correctement généré une sorte d'« hypertrophie des niveaux scolaires intermédiaires » : cette configuration était observée dans la société italienne à la fin des années quatre-vingt-dix.

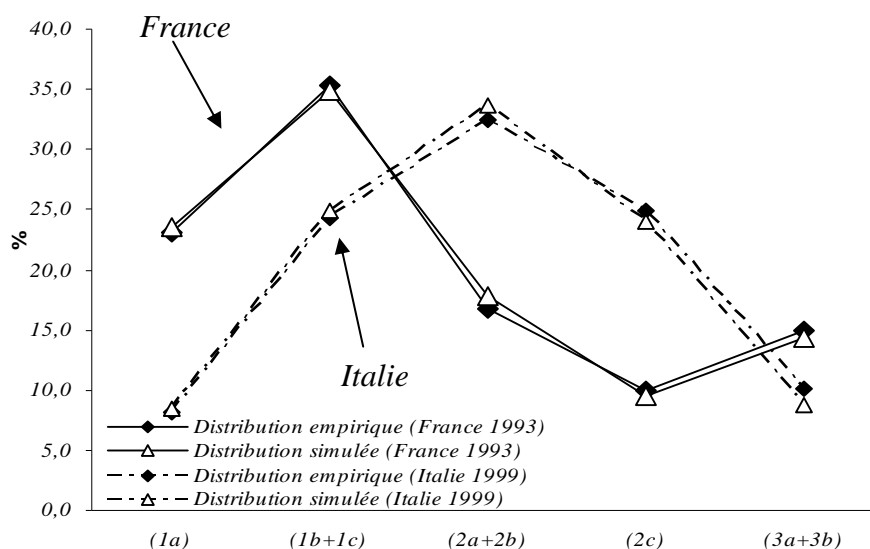


Figure 3. Distribution des diplômes.

Systèmes multi-agents mettant en œuvre le MCEI et sociétés réelles

Note de lecture. Diplômes : 1a = sans diplôme ; 1b + c = diplômes de type CEP ou CAP ; 2a + 2b = diplômes de type BEPC ou BEP ; 2c = diplômes secondaires longs ; 3a + 3b = diplômes de l'enseignement supérieur

Pour décrire les modulations de telles « morphologies » des diplômes à l'intérieur de chaque groupe social, il est ensuite possible d'étudier les pourcentages de ligne calculables sur les données du Tableau 1. On obtient ainsi les destinations éducatives les plus fréquemment atteintes au sein des différents groupes, ce qu'on peut qualifier de « flux scolaires ».

Comme le montre la Figure 4, le MCEI génère une structuration fine de ces flux qui suit de près la relation empirique de « proportionnalité inverse » qui relie le groupe social d'origine et la destination éducative la plus fréquemment atteinte par les individus des sociétés réelles. Au-delà des différences qui existent entre les deux pays, ce qu'il nous importe de noter ici c'est la capacité du modèle à reproduire aussi bien les

scolaires, l'affrontement entre « mesures absolues » et « mesures relatives » a généré des débats considérables : en France [Barbut, 1984, 1985 ; Combessie, 1984, 1985 ; Florens, 1984 ; Grémy, 1984 ; Merllié, 1985 ; Prévot, 1985 ; Vallet, 1988 ; Morel, 2002 ; Merle, 2000 ; Duru-Bellat, 2002 ; Duru-Bellat, Kieffer, 2000 ; Bulle, 2005], en Italie [Cobalti, 1989(a)], aux Etats-Unis [Mare, 1980, 1981], dans les pays scandinaves (cf. [Hellevik, 1997, 2000, 2002 ; Kivinen *et al.* 2001, 2002 ; Marshall, Swift, 1999, 2000 ; Marks, 2004 ; Ringen, 2000, 2005] ; au plan international [Blossfeld, Shavit, 1993]).

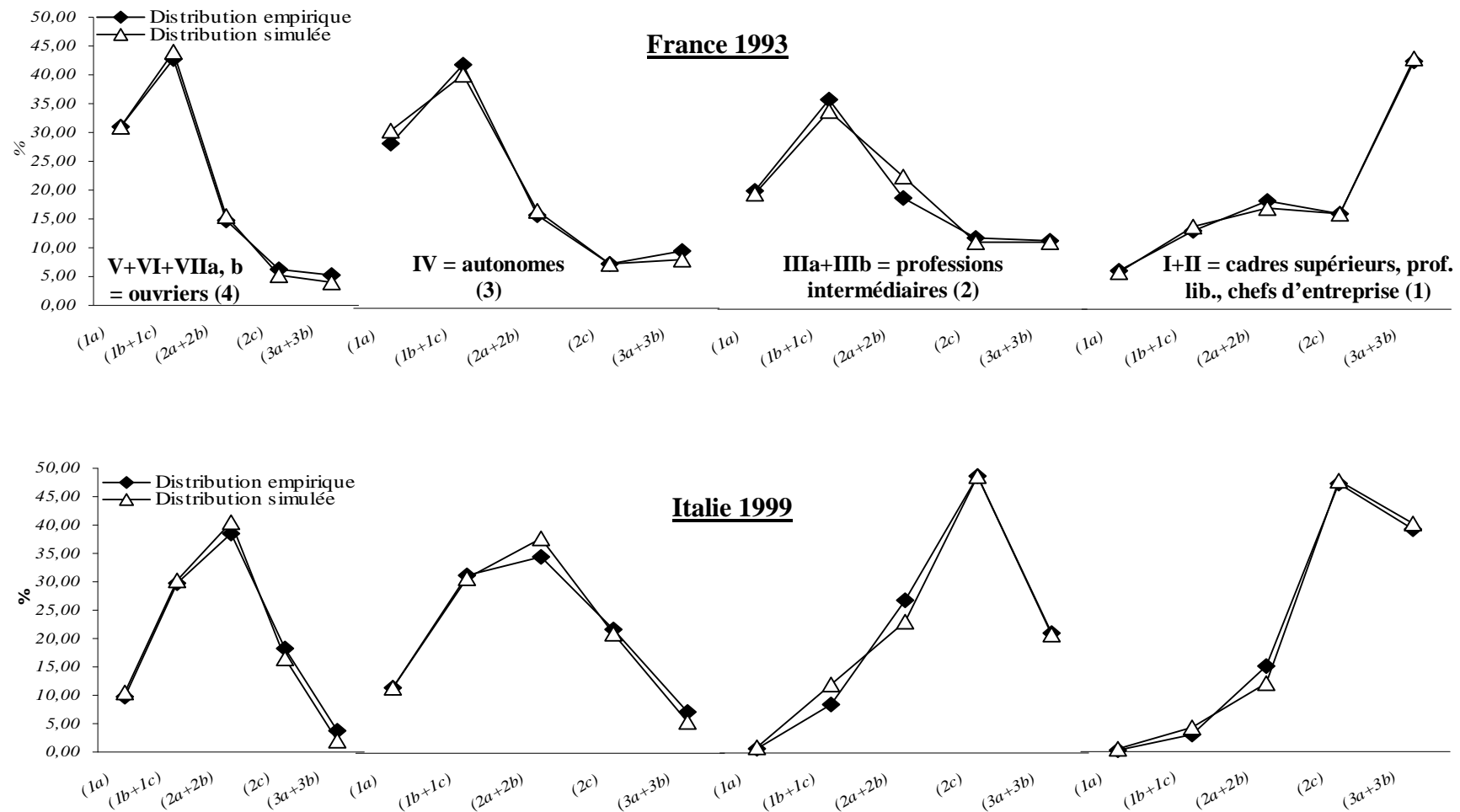


Figure 4. Pourcentages d'individus (« agents ») issus des différents groupes sociaux qui atteignent tel ou tel diplôme en tant que diplôme le plus haut obtenu.

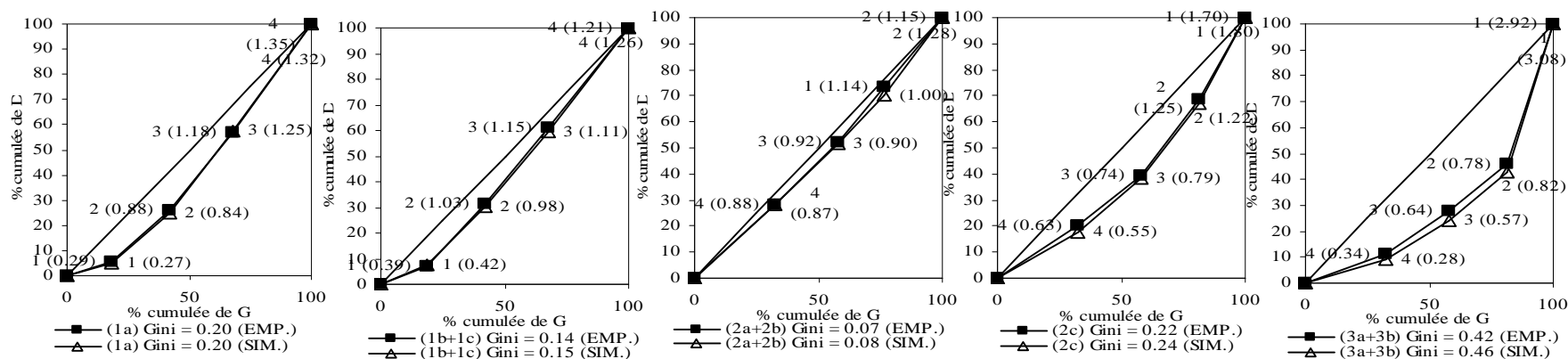
Systèmes multi-agents mettant en œuvre le MCEI (triangles blancs) et sociétés réelles (carrés noirs).

(Pourcentages de ligne calculés sur les données reportées dans le Tableau 1)

Diplômes : 1a = sans diplôme ; 1b+1c = diplômes de type CEP ou CAP ; 2a+2b = diplômes de type BEPC ou BEP ;

2c = diplômes secondaires longs ; 3a+3b = diplômes de l'enseignement supérieur

France 1993 (Indice de GINI moyen : valeur empirique = 0.21 ; valeur simulée = 0.23)



Italie 1999 (Indice de GINI moyen : valeur empirique = 0.23 ; valeur simulée = 0.25)

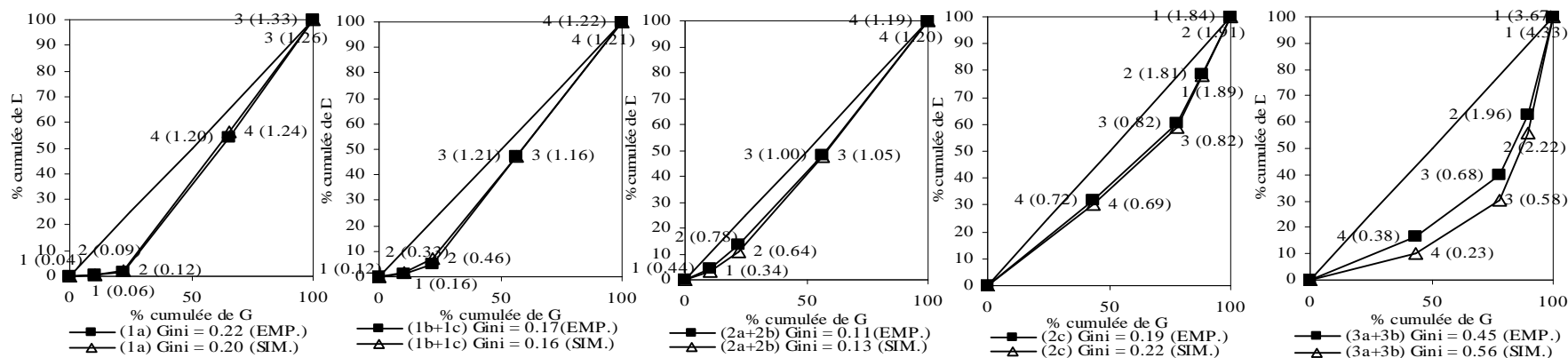


Figure 5. Courbes de Lorenz, indices de GINI et ordre entre les groupes (selon le « coefficient distributif », entre parenthèses).

Systèmes multi-agents mettant en œuvre le MCEI (SIM. = données simulées) et sociétés réelles (EMP. = données empiriques).

(Construction à partir des pourcentages de colonne calculés sur les données reportées dans le Tableau 1 ; se reporter également à ce tableau pour la signification des postes)

situations éducatives les plus fréquentes chez les différents groupes que les obtentions les plus improbables. Ainsi, par exemple, les deux systèmes multi-agents produisent certes la suprématie « statistique » chez les ouvriers des diplômes inférieurs en France et des diplômes secondaires inférieurs en Italie, mais ces systèmes engendrent aussi le petit groupe de « dissidents » qui, parmi les individus issus de ce groupe, réussissent à décrocher un diplôme de l'enseignement supérieur.

Ce résultat peut être imputé à l'hétérogénéité des préférences et des ressources de nos agents. Nous avons déjà rappelé que, si ces derniers ont des valeurs groupales typiques sur RS et sur C, notre mise en œuvre informatique du MCEI admet un certain degré de variabilité intragroupe de ces deux grandeurs. Si l'on considère les voies de diffusion ouvertes par les structures de liens existant entre les agents, l'on entrevoit la possibilité que des choix éducatifs hétérodoxes, par rapport aux choix moyens, se consolident à la fin du processus¹³.

Afin de caractériser la stratification sociale des diplômes de nos sociétés artificielles, une dernière classe d'informations peut être extraite des données simulées. Au lieu d'observer où se dirigent majoritairement les individus issus des différents groupes (les « flux scolaires »), nous pouvons nous intéresser à la composition sociale de chaque diplôme. Il s'agit de ce qu'on qualifie parfois d'« afflux ». Pour les quantifier de façon précise et synthétique, les courbes de Lorenz [1905] et l'indice de Gini [1912] s'avèrent des outils attrayants¹⁴. Les graphiques de la Figure 5 tracent ces courbes pour chaque diplôme et affichent les « coefficients distributifs » qui permettent d'explicitier l'ordre entre les groupes sous-jacents aux courbes¹⁵.

La similarité qui existe entre les courbes de Lorenz simulées et celle observées ressort très nettement de ces graphiques. Le partage sociétal de l'éducation qui caractérise nos deux sociétés artificielles présente une structuration statistique étonnement proche de celle observée empiriquement. À cet égard, la valeur empirique de l'indice de Gini moyen vaut 0.21 pour la France et 0.23 pour l'Italie : les valeurs calculées sur les données simulées valent respectivement 0.23 et 0.25. Si l'on examine patiemment les « coefficients distributifs », ensuite, on constate que la hiérarchie entre les groupes propres aux données simulées suit de près celle décelable dans les données empiriques. Les systèmes multi-agents reproduisent parfaitement l'inversion dans le degré de concentration que l'on observe entre les groupes sociaux dès lors que l'on monte dans la hiérarchie des diplômes. Conformément aux courbes de Lorenz empiriques, l'on retrouve, aussi bien dans la société virtuelle française qu'italienne, les groupes inférieurs et moyens concentrés (au sens de Gini) dans les diplômes inférieurs et intermédiaires tandis que la situation exactement inverse se vérifie en correspondance des diplômes supérieurs. Dans le contexte d'une telle régularité, il importe d'insister sur le fait que les systèmes artificiels différencient finement les deux pays quant au contenu

¹³ Dans la partie 6, nous reviendrons sur la « variabilité intragroupe » de RS et de C pour y intervenir de façon expérimentale. On tâchera ainsi de convaincre le lecteur du bien-fondé de l'intuition ici avancée.

¹⁴ Par ordre chronologique, Blau [1977], Allison [1978], Chauvel [1995], Lampard [2000] et Pedersen [2004] permettent de suivre les débats générés en sociologie par ces outils. Parmi les auteurs qui leur ont reconnu un intérêt spécifique pour l'analyse empirique des inégalités éducatives, citons Barbut [1984, 1985(a), (b), 1990] et Hellevik [1997, 2000, 2002] ; d'autres l'ont nié [Combessie, 1985 ; Marshall, Swift, 1999, p. 247 ; Kivinen *et al.*, 2001, p. 173].

¹⁵ Rappelons que ces coefficients représentent le rapport entre la proportion d'individus issus d'un groupe social donné qui sont présents à un niveau scolaire donné et le poids démographique de ce groupe : la valeur de ce coefficient exprime ainsi le degré de sur-représentation (si supérieure à 1) ou de sous-représentation (si inférieure à 1) de ce groupe en correspondance de tel ou tel diplôme [Barbut, 1984 ; Hellevik, 1997].

spécifique de cet ordre. Notamment, en correspondance des trois situations éducatives « $1a$ », « $1b+1b$ » et « $2a+2b$ », la société virtuelle italienne est caractérisée par une sorte de « monopole éducatif » de la part du groupe ouvrier (groupe 4) et du groupe intermédiaire (groupe 3) : en France, cela n'est pas le cas. Compte tenu de ce fait et de l'existence dans la société artificielle française d'au moins un diplôme caractérisé par un degré d'inégalité distributive particulièrement contenu (les diplômes secondaires inférieurs, « $2a+2b$ »), on peut dire que les systèmes multi-agents génèrent un profil du partage sociétal de l'éducation conforme aux données empiriques que l'on pourrait qualifier d'« égalité distributive localisée » pour la France et d'« inégalité distributive diffuse » pour l'Italie.

Ainsi, qu'il s'agisse de la distribution des diplômes, des diplômes les plus fréquemment atteints au sein de chaque groupe (les « flux scolaires ») ou de la composition sociale des différents diplômes (les « afflux scolaires »), le « système complexe adaptatif » (du type ébauché en Figure 2) activé dans la mémoire vive de notre ordinateur grâce à la simulation multi-agents produit des données structurellement conformes aux données empiriques françaises et italiennes. Les hypothèses à l'origine de ces sociétés artificielles – le MCEI (cf. la formule (6)) – semblent donc en mesure d'engendrer une stratification sociale des diplômes.

5.2 LA CONFIGURATION DE LA « FLUIDITÉ ÉDUCATIVE »

Pour nous en assurer, interrogeons à présent les données simulées contenues dans le Tableau 1 d'un second point de vue – dit « relatif », comme nous l'avons anticipé au début de cette section 5 – c'est-à-dire eu égard à la proportion d'individus qui obtiennent tel diplôme plutôt que tel autre au sein d'un groupe donné par rapport à la situation, sur cette même alternative, d'un autre groupe social. Pour dresser un bilan à la fois synthétique et général des multiples comparaisons possibles – qu'on pourrait qualifier de « fluidité éducative » –, nous choisissons d'utiliser à la mesure d'association dite, dans la littérature anglo-saxonne, « odds ratio généralisé » (cf. [Kaufman, Schervish, 1986, 1987]) qu'en raison de sa structure mathématique, nous proposons de qualifier ici de « birapport généralisé ». Cette statistique exprime le degré de sur-représentation (si supérieure à 1) ou de sous-représentation (si comprise entre 0 et 1) moyenne propre à un groupe social donné sur un diplôme donné par rapport, non pas à tel ou tel autre diplôme, mais à n'importe quel autre diplôme en comparaison, non pas à tel ou tel groupe, mais à n'importe quel autre groupe¹⁶.

¹⁶ Formellement, si L_i et C_j (avec $i = 1, \dots, G$ et $j = 1, \dots, D$) indiquent respectivement la ligne et la colonne générique d'un tableau « groupe social d'origine/diplôme » de dimension (G, D) et $L_i C_j$ l'effectif qui se trouve au croisement de la ligne i et de la colonne j , l'« odds ratio » (le birapport) pour quatre cases contiguës peut s'écrire, indifféremment, des deux façons suivantes (cf. par exemple, [Bohrstedt, Knoke, 1998, p. 162-166 ; Hout, 1983, p. 16 ; Powers, Xie, 2000, p. 95-97 ; Wong, 2003(a)]) :

$$\frac{(L_i C_j)}{(L_i C_{j+1})} \setminus \frac{(L_{i+1} C_j)}{(L_{i+1} C_{j+1})} = \frac{(L_i C_j)}{(L_{i+1} C_j)} * \frac{(L_{i+1} C_{j+1})}{(L_i C_{j+1})}$$

Or, chaque case du tableau sera caractérisée par $(G-1)*(D-1)$ odds ratios (ce qu'on qualifie d'« ensemble de base », cf. [Goodman, 1969]), chacun d'entre eux exprimant la situation d'avantage ou de désavantage d'un groupe donné par rapport à un autre groupe sur une alternative éducative spécifique : la *moyenne géométrique* de ces coefficients locaux donne naissance à un *odds ratio généralisé*. Cette mesure a parfois été proposée pour décrire la « mobilité sociale relative » (cf. [Cobalti, 1989(b), (c) ; 1995 ; Cobalti, Schizzerotto, 1994 ; Ballarino, Cobalti, 2003 ; et Pisati, Schizzerotto, 1998]). Beaucoup plus rarement, en revanche, on l'a proposée pour décrire l'aspect « relatif » de la stratification éducative (cf. [Cobalti, 1992, p. 142]).

La Figure 6 représente ainsi graphiquement l'ensemble de ces coefficients calculés sur les données simulées ainsi que sur les données empiriques : la hauteur des cônes équivaut à la valeur des « birapports généralisés » correspondant à un groupe donné pour une situation éducative donnée.

Empiriquement, une configuration « polarisée » des opportunités éducatives apparaît¹⁷ : les sociétés artificielles gérées par le MCEI sont également caractérisées par une telle structure. Les cônes les plus hauts se situent en effet aux deux coins opposés du plan : d'une part, en correspondance des situations éducatives les moins désirables, d'autre part, en correspondance des diplômes secondaires longs et, plus encore, de l'enseignement supérieur. Ici et là, les groupes qui disposent des avantages concurrentiels les plus forts ne sont évidemment pas les mêmes : les « ouvriers » et les « autonomes » monopolisent les opportunités d'obtention des situations éducatives inférieures tandis que les « intermédiaires » et les « cadres supérieurs » ont des avantages compétitifs particulièrement forts en correspondance des strates supérieures de la hiérarchie des diplômes.

Ces deux aires d'« attraction forte » s'expriment d'ailleurs bien plus intensément en Italie qu'en France. Elles sont aussi moins étendues ici que là : les avantages du groupe des « cadres supérieurs » en correspondance des diplômes de l'enseignement supérieur (et, bien que de manière atténuée, du « secondaire long », 2c) sont en quelque sorte uniques en France tandis qu'en Italie, le groupe « intermédiaire » s'associe à cette position d'avantage relatif. Comme en témoigne la valeur moyenne des « birapports généralisés », on observe pour l'Italie, des résultats moins fluides que pour la France : le MCEI conduit ponctuellement à cette différence structurelle.

Afin d'évaluer le bien-fondé de cette différence importante, nous avons également tenté d'appréhender globalement la configuration des « birapports généralisés » de la Figure 6 en appliquant aux données empiriques et aux données simulées deux modèles log-linéaires parmi les plus fréquemment utilisés dans les études quantitatives de la mobilité sociale (cf. par exemple, [Breen, 2004, chap. 2]). Il s'agit, d'une part, du modèle dit d'« association constante »¹⁸, d'autre part, du modèle dit « UNIDIFF »,

¹⁷ Dans les études sur la mobilité sociale, cette configuration apparaît d'ailleurs souvent (cf. [Grusky, Hauser, 1984, « Within all sixteen countries, the picture that emerges is one of severe immobility at the two extremes of the occupational hierarchy and considerable fluidity in the middle »]), p. 35.

¹⁸ Formellement, le « modèle d'association constante » exprime les effectifs de chaque celle d'un tableau de contingence à trois variables (le Tableau 1, dans notre cas) de la manière suivante (cf. par exemple [Powers, Xie, 2000, chap. 4]) :

$$\ln(F_{ijw}) = \ln(u) + \ln(u_i) + \ln(u_j) + \ln(u_w) + \ln(u_{iw}) + \ln(u_{jw}) + \ln(u_{ij})$$

u étant une constante, u_i le terme expressif des effets marginaux de la variable de ligne (dans notre cas, le groupe social d'origine des individus), u_j le terme expressif des effets marginaux de la variable de colonne (dans notre cas, le diplôme le plus haut obtenu), u_w le terme expressif des effets d'une variable tierce (dans notre cas, le pays) dont les catégories constituent les niveaux de contrôle de la relation de premier ordre entre i et j , u_{iw} et u_{jw} étant les termes expressifs des variations des distributions marginales des deux variables principales selon la catégorie de la variable tierce, u_{ij} le terme expressif de l'association entre les deux variables principales. Comme on le voit, aucun terme expressif des variations de cette association au sein des catégories de la variable tierce n'est présent (ce terme serait u_{ijw}). Le modèle de « différence uniforme » tâche précisément de modéliser cette interaction à trois variables. Cela est fait de la manière suivante :

$$\ln(F_{ijw}) = \ln(u) + \ln(u_i) + \ln(u_j) + \ln(u_w) + \ln(u_{iw}) + \ln(u_{jw}) + (\beta_w) \ln(\theta_{ij})$$

Le modèle suppose ainsi une structure des « birapports » commune aux différentes sous-tables (ce qui est représenté par θ_{ij}), mais il admet que leurs valeurs se modifient uniformément selon un facteur β_w . Les deux modèles en question ont été ici estimés sous des contraintes de paramétrisation de type ANOVA (égalisation à 1 du produit des paramètres) et au moyen d'un algorithme de « maximum de vraisemblance » connu sous le nom de « uni-dimensional Newton algorithm » (cf. [Vermunt, 1996, p. 27-28, 317-321 ; 1997, p. 5, 71, 89]). Les estimations ont été faites grâce au logiciel « Lem » (« Log-linear and event history analysis with missing data using EM algorithm ») mis au point par Jeroen Vermunt de l'Université de Tilburg (cf. [Vermunt, 1997]).

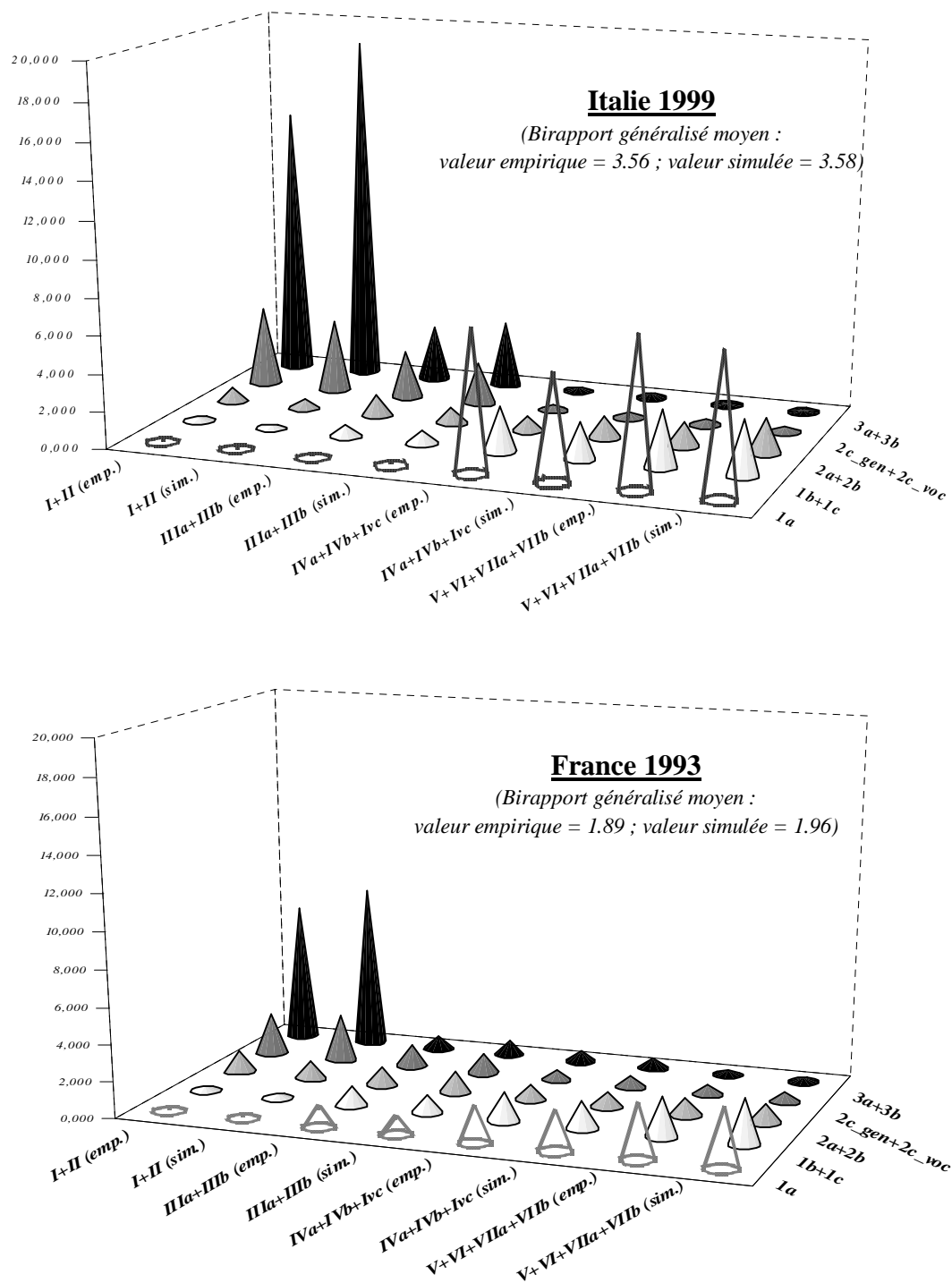


Figure 6. Configuration de la « fluidité éducative » (« Birapports généralisés »).

Systèmes multi-agents mettant en œuvre le MCEI et sociétés réelles.

(« Emp. » = valeurs des birapports généralisés calculées sur les données empiriques

« Sim. » = valeurs des birapports généralisés calculées sur les données simulées)

Légende. Groupe sociaux : I+II = « cadres, professions libérales, chefs d'entreprise » ; IIIa+III = « professions intermédiaires » ; IV = « autonomes » ; V+VI+VIIa, b = « ouvriers » ; **Diplômes :** 1a = sans diplôme ; 1b+1c = diplômes de type CEP ou CAP ; 2a+2b = diplômes de type BEPC ou BEP ; 2c_gen+2c_voc = diplômes secondaires longs ; 3a+3b = diplômes de l'enseignement supérieur

acronyme qui est pour « différence uniforme » [Erikson, Goldthorpe, 1992, p. 91-92] ; (cf. aussi [Xie, 1992]).

Dans notre cas, la signification substantielle de ces deux modèles log-linéaires est la suivante : le premier suppose qu'aucune différence dans la « fluidité éducative » n'existe entre la France et l'Italie tandis que le second suppose qu'une différence existe, mais de type particulier (notamment, tous les « birapports généralisés » seraient plus proches, ou plus éloignés, de 1 selon un facteur uniforme β , dit paramètre UNIDIFF).

Le Tableau 2 reporte les résultats de ces estimations. On constate ainsi que, quel que soit l'indice d'ajustement considéré, le modèle de « différence uniforme » peut être préféré au modèle d'« association constante », aussi bien pour les données empiriques que pour les données simulées. La structure de la fluidité éducative des deux pays paraît ainsi effectivement différer. La Figure 7 reporte, ensuite, la valeur du paramètre UNIDIFF β qui synthétise le sens et l'intensité de ces différences. Sur la base des valeurs de ce paramètre, on peut effectivement conclure que la structure des opportunités éducatives italiennes est largement plus rigide que celle observée en France. L'estimation de β sur les données simulées montre d'ailleurs qu'une telle différence apparaît ponctuellement aussi au sein de nos systèmes multi-agents mettant en œuvre le MCEI.

	DI	L ²	Bic	ID	rL ²
France (1993) et Italie (1999) réelles					
(1) « Association constante »	12	378.68 (0.00)	257.48	0.039	92.56
(2) « Différences uniformes »	11	171.34 (0.00)	60.24	0.026	96.63
Systèmes multi-agents français et italien					
(3) « Association constante »	12	401.00 (0.00)	279.80	0.041	92.93
(4) « Différence uniforme »	11	190.13 (0.00)	79.04	0.027	96.65
Significativité des différences selon test classique : (1) - (2) = 207.34 significatif au seuil de <0.001 (3) - (4) = 210.87 significatif au seuil de <0.001					

Tableau 2. Estimation du modèle log-linéaire d'« association constante » et du modèle log-multiplicatif de « différence uniforme » (dit UNIDIFF) sur les données empiriques et sur les données simulées (présentées dans le Tableau 1)

Ainsi, à la lumière des résultats jusqu'ici présentés, la simulation de ce modèle engendre des données qui, aussi bien sous le profil « absolu » que « relatif », sont conformes aux multiples facettes des stratifications éducatives française et italienne des années quatre-vingt-dix. Mais, au-delà de sa capacité à reproduire la configuration des observations transversales, l'utilité heuristique du MCEI doit être évaluée pour ce qu'il nous apprend au sujet de leur genèse. Comme la Figure 2 le laissait entendre, les hypothèses constitutives de ce modèle théorique donnent en effet naissance à un « système complexe adaptatif » dont la dynamique permet de mieux comprendre le processus d'émergence de la stratification éducative.

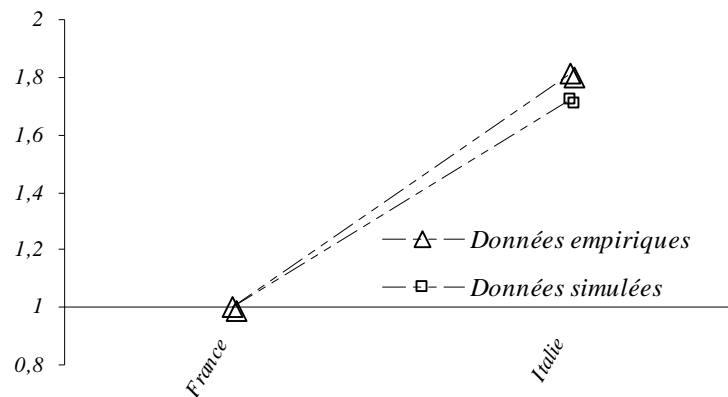


Figure 7. Valeurs assumées par le paramètre UNIDIFF β selon les estimations des modèles (2) et (4) (cf. Tableau 2)

Note. Une valeur de $\beta > 1$ par rapport à la France indique un lien plus fort en Italie qu'en France entre groupe socioprofessionnel d'origine et diplôme

Indépendamment de la compréhension progressivement plus fine que l'on pourrait acquérir du fonctionnement du MCEI, celui-ci ne contient pas d'éléments ayant la capacité de conduire *ipso facto* à une stratification aussi structurée des diplômes des agents. Le code du programme informatique qui supporte la configuration et la « vie » de nos sociétés artificielles ne prévoit *a priori* aucune règle qui ait l'effet de forcer l'issue systémique du processus microsociologique dans un tel sens. Ce sont bien les interactions locales entre des milliers d'agents, qui se déterminent selon les principes du MCEI, qui conduisent à ce résultat global.

Certes, la population d'agents est différenciée au départ en groupes en termes d'objectifs éducatifs (RS dans la formule (6)) et de ressources (C dans la formule (6)). Mais, et c'est là un point essentiel, il existe une certaine variabilité au sein de chaque groupe car les agents sont initialisés sur ces grandeurs par tirage aléatoire à partir de distributions probabilistes. Les frontières entre les groupes sont en réalité, au départ, particulièrement floues : les distributions d'où l'on extrait les valeurs d'initialisation des agents ont des aires fort étendues de chevauchements. Nos agents, surtout aux premiers niveaux scolaires à franchir, ont donc des appréciations subjectives des diplômes et des ressources qui ne sont pas aussi différenciées qu'on pourrait le croire à la lumière de la répartition sociétale des diplômes repérables à la fin du processus.

Ce décalage a son origine dans les interactions dyadiques continuellement à l'œuvre entre les agents : ces interactions creusent progressivement les différences de départ. Les interactions dyadiques locales entre les agents sont en effet le moteur du modèle à double titre. D'une part, elles permettent que les choix éducatifs se diffusent au sein des différents groupes ; d'autre part, elles alimentent un processus lent de cristallisation autour des choix éducatifs typiques, c'est-à-dire les plus fréquemment émis en raison des évaluations éducatives subjectives et des ressources majoritairement répandues au sein d'un groupe donné. Puisque le réseau global qui relie les agents privilégie les liens intragroupes et, parmi les liens intergroupes, les liens entre des groupes socialement proches (cf. plus haut, note 11), les interactions dyadiques locales conduisent à un écartèlement progressif des choix éducatifs des agents selon le groupe.

Ce qui était relativement peu différencié au départ le devient de plus en plus à mesure que la dynamique du modèle progresse¹⁹.

Ainsi, similairement à ce que Schelling [1971] a fait de la ségrégation spatiale, le MCEI animé par simulation multi-agents fait de la stratification éducative le résultat émergent non voulu (ce qui ne veut pas forcément dire « surprenant » du point de vue de l'observateur) d'un processus parfois qualifié d'« amplification » (cf. [Boudon, 1979(b), p. 156-157]), c'est-à-dire une dynamique qui accentue itérativement les différences de départ en termes de préférences et de ressources individuelles (cf. néanmoins [Forsé, Parodi, 2006], pour une analyse critique pénétrante de certaines interprétations abusives du travail de Schelling).

6. DONNÉES SIMULÉES, DONNÉES EMPIRIQUES ET « FAITS STYLISÉS »

La capacité de H_1 , H_2 , H_3 , H_4 , H_5 et H_6 (formalisées dans la formule (6) et mises en œuvre par un système semblable à celui esquissé en Figure 2) à reproduire des structures spécifiques de données empiriques ayant été attestée, une dernière question mérite d'être posée. Quelles conséquences logiques, autres que celles directement testables empiriquement, pourrait-on tirer de ce modèle théorique ?

L'utilisation de la simulation informatique nous permet de répondre à une telle question par une opération simple : le changement de certains aspects de la structure paramétrique du modèle et l'étude de son comportement faisant suite à ces modifications. C'est là ce qu'on qualifie d'« expérimentation » – « [...] un processus par lequel on récolte des données sur un système en agissant sur ces entrées » [Ramat, 2006] – ou bien, parfois, de « quasi-expérimentation » [Boudon, 1973(a), p. 172], ou encore, surtout en informatique, d'« analyse de sensibilité » [Ginot, Monod, 2006]. Dès lors, ce sont des « faits stylisés » [Sanders, 2006] ou, si l'on préfère, des phénomènes purement computationnels (cf. [Livet, 2006]) que l'on produit en ce sens que les structures de données engendrées par le modèle n'ont pas un corrélat empirique direct et immédiat. Dans notre cas, cependant, cette mise entre parenthèses des données empiriques n'est que partielle. Notre objectif est en effet d'évaluer le type de stratification sociale des diplômes produit par le modèle sous des conditions différentes par rapport à celles qui nous ont permis de reproduire correctement les données empiriques transversales françaises et italiennes.

Parmi les nombreuses modifications paramétriques possibles (cf. [Manzo, 2006(b), chap. 9]), nous nous intéressons tout particulièrement ici à celles qui concernent deux termes centraux du MCEI : les « évaluations subjectives » que les acteurs donnent des diplômes et leur perception subjective des coûts de ces derniers

¹⁹ Lors de l'analyse expérimentale du MCEI (dont certains aspects sont discutés dans la section 6), deux régularités virtuelles témoignant de l'existence d'un tel processus ont été attestées (cf. [Manzo, 2006(b), chap. 9]). Premièrement, nous avons observé que, si l'on efface complètement les interactions dyadiques entre les acteurs, aucune forme de stratification des diplômes n'apparaît au sein de nos sociétés artificielles. La dynamique du modèle est brisée. Presque tous les agents, quel que soit leur groupe social, s'arrêtent aux situations éducatives de plus bas niveau. Deuxièmement, nous avons constaté que, si l'on élimine le versant intergroupe de l'interaction dyadique, on obtient une stratification des diplômes encore plus ségréguée que celle que l'on observe en présence d'un nombre limité de contacts intergroupes. Les interactions dyadiques intergroupes alimentent donc bien des choix hétérodoxes au sein des différents groupes (cf. plus haut, H_4). Or, puisque, dans nos sociétés artificielles, ce type de liens est (par construction) rare, c'est bien l'interaction dyadique intragroupe qui prime en retranchant progressivement les groupes les uns par rapport aux autres.

(respectivement, RS et C dans la formule (6)). En effet, comme la description de la dynamique du modèle proposée à l'instant le suggère, si les interactions entre les acteurs constituent le « moteur » du modèle, leurs évaluations et leurs ressources en constituent le « carburant ». Ce n'est cependant pas la structure de ces paramètres entre les groupes qui sera changée, mais plus spécifiquement leur variabilité intragroupe. Ainsi, par rapport à la configuration paramétrique de RS et de C adoptée pour reproduire les données empiriques françaises et italiennes (cf. Annexe B, Tableau B), nous agissons à présent sur la variance des distributions de probabilité à partir desquelles l'on initialise les agents des sociétés artificielles permettant de mettre en œuvre le MCEI²⁰.

Le Tableau 3 reporte les résultats d'un premier groupe d'expériences. La variance des distributions de RS et de C est ici modifiée de manière contenue (sa valeur passe de 0.25 à 0.50) et localisée (elle ne touche que les agents « ouvriers » et uniquement pour les diplômes inférieurs et intermédiaires inférieurs, d'abord, pour ceux de niveau intermédiaire supérieur et supérieur, ensuite). Deux faits stylisés principaux apparaissent :

- 1) Si l'on compare la sous-table (a) à la sous-table (b), on constate que cette modification modeste de la variance des distributions de RS et C circonscrite aux agents « ouvriers » suffit à engendrer une élévation non négligeable de la proportion d'agents qui obtiennent des diplômes secondaires inférieurs et supérieurs non seulement parmi les « ouvriers » mais également chez les autres groupes d'agents dont les préférences et les ressources n'ont été affectées par aucune modification exogène. C'est dire l'importance des processus de diffusion rendus possibles par le réseau d'interactions entre les agents postulé par le MCEI : des modifications locales se sont propagées et progressivement amplifiées²¹.
- 2) Si l'on compare ensuite la sous-table (c) à la sous-table (b), on constate que ce processus vertueux de propagation est néanmoins largement moins opérant si l'augmentation de la variabilité des choix des agents « ouvriers » concerne non pas les diplômes inférieurs et intermédiaires inférieurs mais les diplômes de niveau plus élevé. C'est là l'effet de la structure séquentielle du système de diplômes auquel le MCEI s'applique (cf. Figure 1). Puisque le choix des diplômes supérieurs ne se pose que conditionnellement à l'obtention des diplômes inférieurs, des préférences et des ressources plus favorables au choix des premiers n'entraînent pas d'effets conséquents car les difficultés à franchir les seconds n'ont pas été affectées.

²⁰ Pour des raisons d'espace, nous nous limiterons à étudier les données simulées sous le profil des « flux scolaires », c'est-à-dire des diplômes les plus fréquemment obtenus au sein des différents groupes d'agents (techniquement, il s'agit des « pourcentages de ligne » calculables sur les données simulées). Pour ne pas alourdir le texte, en outre, nous reporterons uniquement les résultats qui concernent les données françaises : chacune des régularités dont il sera question apparaît cependant ponctuellement aussi dans l'application du modèle aux données italiennes (cf. [Manzo, 2006(b), chap. 9]).

²¹ L'analyse de modèles formels contenant des structures d'interaction conduit souvent au résultat selon lequel « there is no necessary proportionality between the size of a cause and the size of its effect » [Hedstrom, 2005, p. 144]. Des études récentes par simulation multi-agents du rôle que la structure des interactions entre les acteurs joue dans les phénomènes de diffusion des opinions et des innovations sont [Amblard, Deffuant, 2004] et [Deffuant, Huet et Amblard, 2005]. Steyer et Zimmermann [2004] ont récemment abordé de manière formelle, claire et synthétique le lien qui existe entre processus de diffusion et structures résiliantes. À ce sujet, l'étude sociologique pionnière reste celle de Coleman, Katz et Menzel [1966].

	1a	1b+1c	2a+2b	2c	3a+3b	TOTAL
(a) Variance = 0.50 sur (4) pour '1b+1c' et '2a+2b'						
I+II (1)	2.12	3.42	13.43	26.34	54.68	2921
IIIa+IIIb (2)	8.71	17.39	40.02	22.55	54.68	3673
IV (3)	12.76	26.94	37.28	14.44	8.57	4109
V+VI+VIIa,b (4)	2.85	6.89	55.15	26.87	8.24	5073
(b) Simulation de référence : Variance = 0.25 (Valeur qui permet de reproduire les données empiriques)						
I+II (1)	6.44	14.60	17.82	16.96	44.18	2921
IIIa+IIIb (2)	19.87	34.13	22.81	11.49	11.68	3673
IV (3)	29.46	38.84	16.12	7.49	8.09	4109
V+VI+VIIa,b (4)	31.11	44.09	15.52	5.22	4.07	5073
(c) Variance = 0.50 sur (4) pour '2c' et '3a+3b'						
I+II (1)	7.46	16.01	17.92	15.13	43.49	2921
IIIa+IIIb (2)	22.37	36.28	19.84	10.24	11.27	3673
IV (3)	30.24	40.18	14.27	7.35	7.96	4109
V+VI+VIIa,b (4)	32.29	43.02	10.98	7.92	5.80	5073

Tableau 3. Pourcentage d'agents ayant obtenu tel ou tel diplôme (en colonne) au sein des différents groupes (en ligne) sous une modification de la variance des distributions de RS et de C localisée à *un seul groupe* (« ouvriers ») pour deux seules classes de diplômes.

Légende. Groupes sociaux : I+II = « cadres, profession libérales, chefs d'entreprise » (1) ; IIIa+IIIb = « professions intermédiaires » (2) ; IV = « autonomes » (3) ; V+VI+VIIa,b = « ouvriers » (4). *Diplômes* : 1a = sans diplôme ; 1b+1c = diplômes de type CEP ou CAP ; 2a+2b = diplômes de type BEPC ou BEP ; 2c = diplômes secondaires longs ; 3a+3b = diplômes de l'enseignement supérieur

Le Tableau 4 reporte ensuite les résultats d'un deuxième groupe d'expériences dans lesquelles les modifications de la variance des distributions de RS et de C ont assumé un caractère plus généralisé : tous les diplômes que les agents doivent évaluer sont touchés, mais la modification intéresse un seul groupe d'agents à la fois.

Là aussi, deux faits principaux peuvent être détectés :

- 1) Si l'on compare la sous-table (e) à la sous-table (d), on constate qu'un accroissement de la variabilité interindividuelle qui touche tous les diplômes auxquels les agents sont confrontés entraîne des effets d'ampleur plus étendue en termes d'élévation des flux scolaires que cela n'était le cas pour des modifications touchant uniquement les niveaux scolaires inférieurs (cf. Tableau 3, sous-table (a)/(b)). Dans ce dernier cas, par exemple, les agents « ouvriers » qui obtenaient un diplôme de l'enseignement supérieur (« 3a+3b ») passaient de près de 4 % à près de 8 % tandis que ce pourcentage passe de près de 4 % à près de 14 % dans le cas d'une augmentation qui intéresse tous les niveaux scolaires.
- 2) Ces effets de propagation semblent cependant être d'autant plus réduits que le groupe social intéressé par l'augmentation de la variance de RS et de C se situe au rangs supérieurs de la hiérarchie des positions. Si l'on compare, par exemple, la sous-table

(h) à la sous-table (e), on constate en effet qu'une telle augmentation conduit près de 14 % d'agents « ouvriers » à un diplôme de l'enseignement supérieur (« 3a+3b ») si elle concerne les agents « ouvriers », mais seulement 4.16 % d'entre eux si elle concerne les agents « cadres ». Plus généralement, si l'augmentation de la variance de RS et de C ne touche que ces derniers, aucun autre groupe ne bénéficie de ces choix plus ambitieux.

	1a	1b+1c	2a+2b	2c	3a+3b	TOTAL
(d) Simulation de référence : Variance = 0.25 (Valeur qui permet de reproduire les données empiriques)						
I+II (1)	6.44	14.60	17.82	16.96	44.18	2921
IIIa+IIIb (2)	19.87	34.13	22.81	11.49	11.68	3673
IV (3)	29.46	38.84	16.12	7.49	8.09	4109
V+VI+VIIa,b (4)	31.11	44.09	15.52	5.22	4.07	5073
(e) Variance = 0.50 que pour 'V+VI+VIIa,b' (4) sur tous les diplômes						
I+II (1)	2.33	3.02	7.10	27.07	60.48	2921
IIIa+IIIb (2)	7.38	17.02	35.78	27.88	11.94	3673
IV (3)	12.49	25.55	32.71	19.59	9.66	4109
V+VI+VIIa,b (4)	2.54	7.04	18.68	56.96	14.79	5073
(f) Variance = 0.50 que pour 'IV' (3) sur tous les diplômes						
I+II (1)	4.50	9.03	17.60	21.45	47.41	2921
IIIa+IIIb (2)	13.33	27.99	31.86	14.72	12.09	3673
IV (3)	6.50	16.17	27.44	26.27	23.62	4109
V+VI+VIIa,b (4)	15.09	38.81	31.17	10.55	4.38	5073
(g) Variance = 0.50 que pour 'IIIa+IIIb' (2) sur tous les diplômes						
I+II (1)	4.85	9.41	16.20	18.86	50.68	2921
IIIa+IIIb (2)	5.23	9.83	23.17	29.06	32.70	3673
IV (3)	21.15	36.16	24.38	9.46	8.86	4109
V+VI+VIIa,b (4)	19.00	39.38	27.16	10.26	4.20	5073
(h) Variance = 0.50 que pour 'I+II' (1) sur tous les diplômes						
I+II (1)	3.27	4.74	7.21	9.68	75.09	2921
IIIa+IIIb (2)	19.10	32.67	25.78	11.97	10.48	3673
IV (3)	27.44	37.82	18.68	8.30	7.77	4109
V+VI+VIIa,b (4)	28.02	43.38	18.02	6.42	4.16	5073

Tableau 4. Pourcentage d'agents ayant obtenu tel ou tel diplôme (en colonne) au sein des différents groupes (en ligne) sous une modification de la variance des distributions de RS et de C localisée à *un seul groupe*, mais pour tous les diplômes.

Légende. Groupes sociaux : I+II = « cadres, professions libérales, chefs d'entreprise » (1) ; IIIa+IIIb = « professions intermédiaires » (2) ; IV = « autonomes » (3) ; V+VI+VIIa,b = « ouvriers » (4). *Diplômes* : 1a = sans diplôme ; 1b+1c = diplômes de type CEP ou CAP ; 2a+2b = diplômes de type BEPC ou BEP ; 2c = diplômes secondaires longs ; 3a+3b = diplômes de l'enseignement supérieur

À la lumière de la structure mathématique du MCEI (cf. la formule (6)), cette asymétrie dans la propagation des effets locaux s'explique aisément. La présence de choix plus favorables au sein des groupes supérieurs a peu de chances d'entraîner une diffusion généralisée de ces choix car les agents des autres groupes ne disposent pas au départ de valeurs sur RS et sur C qui rendent opérant la valeur quelque peu plus positive du terme d'influence I. Pour le dire avec un exemple, si un agent « ouvrier » a un bilan « rentabilité subjective (RS)/coûts (C) » négatif, voire très négatif, en correspondance

d'un diplôme secondaire supérieur, le fait que son voisin « cadre » choisisse plus fréquemment ce diplôme ne suffira pas à faire basculer la valeur de sa fonction U_{ID} vers une valeur positive. Pour que cela se vérifie, il faudrait que cet agent « ouvrier » ait des nombreux voisins « cadres », ce qui est très peu probable car la configuration du réseau fait (par construction) que ses voisins sont plus fréquemment issus d'un groupe socialement proche.

Ainsi, le MCEI suggère qu'en raison de cette « impossibilité à suivre » de la part des groupes sociaux inférieurs, il est peu probable qu'une course généralisée à l'éducation puisse être déclenchée par des investissements scolaires accrus de la part des groupes supérieurs. À cet égard, notre étude par simulation multi-agents d'un modèle théorique précis semble donc clore une querelle ancienne qui opposa Cherkaoui [1982, p. 91-93] à Bourdieu [1979, p. 147] quant à l'hypothèse, soutenue par ce dernier, selon laquelle l'augmentation de la participation scolaire au niveau secondaire supérieur durant les années soixante et soixante-dix se fonda sur l'antériorité de l'accroissement de la demande d'éducation de la part des groupes supérieurs.

Le Tableau 5 reporte enfin les résultats d'un dernier groupe d'expériences sur la variance des distributions de RS et de C : des modifications de taille différente ont été opérées pour tous les groupes d'agents et pour tous les diplômes à la fois. Deux résultats généraux apparaissent :

- 1) Si l'on compare la sous-table (m) à la sous-table (l), on constate une élévation de très grande ampleur des flux éducatifs. Près de 96 % des agents « cadres », près de 50 % des agents « intermédiaires », près de 37 % des agents « autonomes » et près de 30 % des agents « ouvriers » atteignent un diplôme de l'enseignement supérieur (« $3a+3b$ »). Ce qui va bien au-delà des modifications générées par une modification de la variance de RS et de C limitée à un seul groupe d'agents (cf. Tableau 4) ou, plus encore, à un seul groupe d'agents sur quelques catégories de diplômes seulement (cf. Tableau 3). En parcourant successivement les sous-tables (n), (o) et (p), on constate aisément d'ailleurs que l'ampleur d'une telle élévation des flux scolaires s'amplifie à mesure que la variabilité intragroupe croît. Ainsi, si tout le monde peut se distancier positivement des choix typiques de son groupe, les conditions sont posées pour qu'une cumulation en chaînes d'effets positifs se déploie à travers le réseau d'interactions qui existe entre les agents.
- 2) Si l'on compare ensuite la sous-table (i) à la sous-table (l), on constate qu'exactement le contraire se vérifie dans le cas où la variance intragroupe se réduit. Une diminution de la variance de RS et C conduit en effet à un rabaissement des flux éducatifs vers les situations éducatives de niveau inférieur.

Ainsi, plusieurs « faits stylisés » sont réunis pour croire au rôle que la variabilité interindividuelle joue dans le processus de sédimentation de la stratification sociale des diplômes. Quoique partiel, cet ensemble d'expérimentations sur la variance des distributions de RS et de C nous paraît en effet confirmer la place centrale occupée par la variabilité intragroupe dans la dynamique sous-jacente au MCEI. Son accroissement implique que les agents seront plus fréquemment initialisés avec des valeurs qui s'éloignent positivement, dans le cas de RS (c'est-à-dire, des évaluations subjectives plus favorables), et négativement, dans cas de C (c'est-à-dire, des coûts perçus comme étant moins lourds), des valeurs moyennes propres à leur groupe social d'origine. Ce type de modifications engendre le processus suivant :

- 1) la présence d'un nombre accru de valeurs rares au sein du groupe permet, à travers les interactions entre les agents, la cristallisation en son sein de hiérarchies de choix plus ambitieux ;
- 2) dès lors, les agents des différents groupes entrent en interaction intergroupes avec des choix plus ambitieux ;
- 3) ce qui augmente la probabilité que ces choix se répandent ;
- 4) il s'en suit un nombre agrégé de diplômes plus élevé et, donc, une perte de valeur plus rapide et, donc aussi, une réévaluation des évaluations subjectives des agents plus consistante ;
- 5) celles-ci étant dès le départ plus favorables, il en dérive que le cycle suivant engendre des effets d'autant plus intenses ; et ainsi de suite.

	1a	1b+1c	2a+2b	2c	3a+3b	TOTAL
(i) Variance = 0.125						
I+II (1)	37,29	23,13	11,13	7,43	21,01	2921
IIIa+IIIb (2)	70,41	15,29	5,64	3,65	5,01	3673
IV (3)	78,58	13,11	3,46	1,57	3,29	4109
V+VI+VIIa,b (4)	92,87	5,06	0,51	0,30	1,27	5073
(l) Simulation de référence : Variance = 0.25 (Valeur qui permet de reproduire les données empiriques)						
I+II (1)	6,44	14,60	17,82	16,96	44,18	2921
IIIa+IIIb (2)	19,87	34,13	22,81	11,49	11,68	3673
IV (3)	29,46	38,84	16,12	7,49	8,09	4109
V+VI+VIIa,b (4)	31,11	44,09	15,52	5,22	4,07	5073
(m) Variance = 0.50						
I+II (1)	1,15	0,82	0,52	1,81	95,70	2921
IIIa+IIIb (2)	2,91	2,87	6,98	35,70	51,55	3673
IV (3)	4,13	5,61	11,34	41,64	37,28	4109
V+VI+VIIa,b (4)	1,96	2,44	4,23	61,91	29,46	5073
(n) Variance = 0.75						
I+II (1)	0,47	0,33	0,38	0,29	98,53	2921
IIIa+IIIb (2)	1,50	1,26	2,30	10,56	84,38	3673
IV (3)	1,50	1,93	3,51	16,25	76,81	4109
V+VI+VIIa,b (4)	0,74	0,72	0,84	19,31	78,38	5073
(o) Variance = 1						
I+II (1)	0,47	0,22	0,22	0,32	98,77	2921
IIIa+IIIb (2)	1,04	0,69	1,24	3,65	93,38	3673
IV (3)	1,31	1,24	1,76	5,77	89,91	4109
V+VI+VIIa,b (4)	0,48	0,43	0,37	3,96	94,75	5073
(p) Variance = 1.25						
I+II (1)	0,49	0,21	0,21	0,19	98,91	2921
IIIa+IIIb (2)	0,88	0,78	0,71	1,74	95,89	3673
IV (3)	0,77	1,06	1,23	2,67	94,27	4109
V+VI+VIIa,b (4)	0,27	0,30	0,19	1,64	97,61	5073

Tableau 5. Pourcentage d'agents ayant obtenu tel ou tel diplôme (en colonne)
au sein des différents groupes (en ligne) sous des valeurs différentes
de la variance des distributions de RS et de C
pour tous les groupes et pour tous les diplômes

Légende. Groupes sociaux : I+II = « cadres, profession libérales, chefs d'entreprise » (1) ; IIIa+IIIb = « professions intermédiaires » (2) ; IV = « autonomes » (3) ; V+VI+VIIa,b = « ouvriers » (4). *Diplômes :* 1a = sans diplôme ; 1b+1c = diplômes de type CEP ou CAP ; 2a+2b = diplômes de type BEPC ou BEP ; 2c = diplômes secondaires longs ; 3a+3b = diplômes de l'enseignement supérieur

7. CONCLUSION

Cet article a présenté un modèle de la formation des choix scolaires et a tenté de montrer que les structures d'interaction directe et indirecte existant entre elles permettent de remonter de ces choix à la stratification sociale des diplômes se situant au niveau systémique. Pour ce faire, nous avons étudié, de façon inductive, les données produites de façon déductive, par un système multi-agents dont le fonctionnement était régi par notre modèle théorique et sa formalisation mathématique. En raison de l'homologie de structure existante entre ces données « simulées » et les données empiriques de la France et de l'Italie, nous avons alors avancé une *présomption de validité et de réalisme* au sujet des hypothèses constitutives du modèle. Seulement ensuite, nous nous sommes engagés dans une exploration expérimentale de certains des paramètres permettant leur mise en œuvre.

Il importe cependant de noter que c'est là une forme bien spécifique de validation. Le contrôle empirique auquel nous avons soumis le MCEI repose en effet sur les conséquences macro découlant de ses hypothèses micro et mésosociologiques. En dépit de la pertinence d'une telle procédure de validation (reconnue tant en informatique, Amblard [2003, 2006], qu'en sociologie, Boudon [1984, p. 49]), force est de reconnaître qu'une validation plus directe de ces hypothèses fait pour l'heure défaut à notre travail. Ainsi, au-delà de la possibilité de les justifier sur la base d'éléments théoriques et de données empiriques éparpillés dans la littérature sociologique, nous voudrions, en guise de conclusion, revenir plus spécifiquement sur certains aspects de l'infrastructure opérationnelle du MCEI, c'est-à-dire sur les éléments qui ont permis concrètement de le traiter. Car, s'il est indéniable que formaliser signifie choisir et, obligatoirement, appauvrir, il est essentiel d'avoir conscience de l'« arbitraire » existant dans un modèle afin d'évaluer correctement la portée des résultats qu'il produit.

À cet égard, un premier élément qui mérite d'être signalé tient à la forme fonctionnelle (notamment, logarithmique et logistique) adoptée pour représenter l'interdépendance paramétrique, l'interdépendance stratégique et le rôle de la réussite scolaire passée (cf. les termes qui pèsent RO, RS et C dans la formule (6)). En l'absence d'éléments empiriques solides pour justifier le choix de telle ou telle fonction, il serait certainement pertinent d'étudier le comportement du modèle sous des formes fonctionnelles différentes ayant des propriétés mathématiques qualitatives équivalentes (monotonie, croissance, convexité, etc.). L'objectif d'une telle analyse serait de déterminer le degré de stabilité du comportement du modèle ainsi que de sa capacité à reproduire correctement les données empiriques.

Il importe cependant de noter que, même en présence d'une telle preuve de stabilité, notre modèle pourrait s'exposer à la critique d'une excessive complication de sa formulation mathématique. Parmi ceux qui pratiquent le type de modélisation proposée dans ce travail, certains auteurs influents soutiennent en effet que « *the simpler the model, the easier it may be to discovery and understand the subtle effects of its hypothesized mechanisms* » (cf. [Axelrod, 2005, p. 4] ; cf. aussi [Axelrod, 1997, introduction]). Au prix de quelques difficultés additionnelles dans la phase de compréhension de la dynamique du modèle, nous croyons que l'éloignement de cette règle d'or se justifie à la lumière de l'articulation du phénomène que nous nous proposons d'expliquer. Dans l'optique, qui est la nôtre, de percer un domaine d'étude bien consolidé et présentant déjà une théorie qui focalise les débats, les mécanismes générateurs de la stratification éducative n'auraient pas pu être simplifiés davantage que nous ne l'avons déjà fait. C'est donc la conjugaison des principes « d'abstraction

décroissante » et de « complexité suffisante » énoncés par Lindenberg [1992, 1998, 2002, 2003, p. 362] qui justifie, à nos yeux, une telle complication du modèle (ainsi que de sa formalisation et de sa mise en œuvre).

Un deuxième élément qui doit être mentionné tient à la nature des acteurs de nos sociétés artificielles, les « agents ». De ce point de vue, contrairement à ce qu'on vient de dire, on pourrait objecter à notre modèle d'être encore trop simple. Suivant une distinction courante, les agents animés par le MCEI ne sont en effet que de type « réactif » : ils ne sont dotés ni de croyances, ni d'intentions ou de mémoire, ce qui aurait fait basculer notre modèle vers un modèle à « agents cognitifs » [Castelfranchi, 1998]. Le MCEI suppose, par exemple, que les agents évaluent différemment les bénéfices et les coûts des diplômes selon leur groupe social d'origine, mais ces évaluations ne sont pas représentées par un système de croyances explicitement modélisé et programmé. Le MCEI admet en outre que ces évaluations se modifient dynamiquement sous la pression des cumuls agrégés de choix alimentés par les acteurs agissant à des instants temporels antérieurs, mais aucun mécanisme cognitif explicite n'est ni prévu ni programmé à cet effet. De ce point de vue, on pourrait nous objecter de ne pas être en droit d'attribuer à notre modèle la capacité de prendre en compte des véritables boucles de rétroaction du « macro » sur le « micro » car les agents ne sont pas caractérisés par une véritable « conscience » des traits macro qui affectent leur comportement (cf. sur ce point [Boero, Castellani, Squazzoni, 2006 ; Dessalles, Phan, 2005]). Certains auteurs se demandent même s'il est sensé de parler de « modèles à agents » en présence de simples « agents réactifs » [Squazzoni, 2006].

Un troisième élément qui mérite une brève discussion tient d'ailleurs précisément aux modalités de représentation des évaluations subjectives (RS) que les acteurs donnent des diplômes ainsi que de leurs perceptions subjectives des coûts (C). Comme le lecteur le sait désormais, les agents de nos sociétés artificielles ont été initialisés sur ces deux grandeurs par tirage aléatoire de distributions de probabilité dont les moyennes respectent certaines contraintes d'ordre (cf. § 3, point 1 et 2 ; Annexe B, Tableau B). En raison du manque de données, la valeur de ces paramètres n'a donc pas été estimée empiriquement, comme le voudrait, par exemple, Hedstrom [2005, p. 167] en proposant la notion de « empirically calibrated agent-based models » pour indiquer des modèles où les éléments ayant trait au lien entre le contexte (le groupe social d'origine, dans notre cas) et le comportement des acteurs seraient initialisés sur base empirique.

De ce point de vue, on pourrait alors nous objecter que les valeurs auxquelles nous sommes arrivées sur RS et C sont arbitraires. Cet argument consiste à dire que, puisqu'avec des valeurs numériques différentes l'on pourrait atteindre des résultats similaires, la valeur des simulations s'écroule. L'argument a probablement moins de poids qu'il ne semblerait au premier abord car la valeur numérique *stricto sensu* de RS et de C n'est que le moyen pour faire « tourner » le modèle. Ce qui compte ce sont en revanche les relations que ces paramètres expriment dans leur ensemble ainsi que le respect des contraintes posées *a priori* sur leur structure. Certes, on aurait pu choisir des valeurs différentes. Mais, encore faudrait-il démontrer que le modèle aurait conduit à une reproduction aussi adéquate des données empiriques que celle atteinte ici. Même toutefois, dans ce cas, de deux choses l'une : ou bien ces valeurs alternatives expriment des relations d'ordre différentes entre les groupes, auquel cas il s'agirait d'en éclairer la signification théorique ; ou bien ces valeurs nouvelles violent les relations d'ordre imposées, auquel cas elles ne seraient pas acceptables. Au cas où, enfin, des structures de paramètres différentes exprimeraient des hypothèses théoriques également justifiables (à la vue du bon sens et de la littérature existante), l'on se trouverait

simplement dans une situation classique dans laquelle plusieurs modèles théoriques s'affrontent.

Un quatrième élément qui ne doit pas être négligé a trait ensuite à la manière dont les acteurs de notre modèle théorique opèrent leur choix. Pour l'heure, le MCEI se fonde sur une règle de décision binaire selon laquelle le diplôme est choisi si l'utilité que l'individu associe au diplôme est supérieure à zéro. Une règle de décision probabiliste (de forme logistique, cf. par exemple [Nadal, Gordon, 2005]) pourrait certainement substituer ce critère dichotomique. Une règle de ce type s'avérerait également fort utile pour affiner la différence entre « décision » et « obtention » d'un diplôme, distinction qui reste, en l'état actuel de son développement, peu élaborée au sein du MCEI.

Un dernier élément qu'il importe de ne pas oublier concerne les structures réticulaires dans lesquelles les acteurs sont encastrés. Le MCEI postule en effet que les interactions dyadiques entre les acteurs ont lieu au sein de liens ayant tous le même poids. Dans les termes de la théorie des graphes, cela veut dire que nos systèmes multi-agents mettant en œuvre ce modèle se basent sur des graphes non valués. Les agents ne « pèsent » pas les choix de leurs voisins en fonction du groupe social de ces derniers : le choix d'un agent « ouvriers », par exemple, a pour un agent « cadre » exactement la même importance que le choix de ce dernier a pour le premier. C'est donc là une simplification par rapport à la classe de modèles qui a pourtant inspiré ce travail – les « modèles de choix discrets avec externalité » – dans lesquels on fait habituellement l'hypothèse de l'existence de liens valués entre les acteurs (cf. [Nadal, Gordon, 2005]).

Les cinq points que nous venons de soulever constituent ainsi autant de limites du modèle présenté dans cet article que de pistes de recherche à poursuivre activement pour l'améliorer. Car, s'il est incontestablement vrai, comme le remarque Paul Valéry [1942, p. 864], que « ce qui est simple est toujours faux. Ce qui n'est l'est pas est inutilisable », on ne saurait trouver le juste milieu entre ces deux extrêmes que par un processus de modélisation progressif et incrémental.

Remerciements. L'auteur tient à remercier chaleureusement Frédéric Amblard, maître de conférences en informatique à l'Université de Toulouse I, pour avoir bien voulu programmer le modèle théorique présenté et formalisé dans cet article. Sans son aide précieuse la conception interdisciplinaire de l'analyse sociologique qui anime ce travail aurait manqué d'un maillon essentiel.

BIBLIOGRAPHIE

- ABELL P., "Sociological theory and rational choice theory", Turner B. S. (ed.), *The Blackwell companion to social theory*, Oxford, Blackwell, 1996.
- ABELL P., "The role of rational choice and narrative action theories in sociological theory", *Revue française de sociologie*, 44(2), 2003.
- ABERG Y., "Social action and macro level dynamics: a formal theoretical model", *Acta Sociologica*, 43, 2004.
- ALLISON P. D., "Measures of inequality", *American Sociological Review*, 43, 1978.
- AMBLARD F., *Comprendre le fonctionnement des simulations sociales individus-centrées. Application à des modèles de dynamique d'opinions*, Thèse doctorat en informatique, Université Blaise Pascal - Clermont II, 2003.
- AMBLARD F., « Évaluation et validation dans les modèles multi-agents », Amblard F., Phan D., (eds.), *Modélisation et simulation multi-agents : Applications pour les Sciences de l'Homme et de la Société*, Paris, Hermès (chap. 4), 2006.
- AMBLARD F., DEFFUANT G., "The role of network topology on extremism propagation with the relative agreement opinion dynamics", *Physica A*, vol. 343, 2004.
- AMBLARD F., DUMOULIN N., « Mieux prendre en compte le temps dans les simulations individus-centrées », *Actes des 11^e Journées de Rochebrune 2004* (Rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels), Rochebrune, 2004.
- AMBLARD F., PHAN D. (eds.), *Modélisation et simulation multi-agents : Applications pour les Sciences de l'Homme et de la Société*, Paris, Hermès, 2006.
- American journal of sociology*, special issue: "Computation" 10(4), 2005.
- Archives européennes de sociologie*, numéro spécial : "Simulation in sociology", VI(1), 1965.
- ATLAN H., « L'intuition du complexe et ses théorisations », Fogelman Soulié F. (éd.), *Les théories de la complexité. Autour de l'œuvre d'Henri Atlan*, Paris, Seuil, 1991.
- AXELROD R., *The Complexity of cooperation: agent-based models of competition and collaboration*, Princeton (NJ), Princeton University Press, 1997.
- AXELROD R., COHEN M.D., *Harnessing complexity: organizational implications of a scientific frontier*, New-York, Free Press, 2000.
- AXELROD R., "Advancing the art of simulation in the social sciences", Rennard J.-Ph. (ed.), *Handbook of research on nature inspired computing for economy and management*, Hersey (PA), Idea Group, 2005.
- AXTELL R., "Why agents? On the varied motivations for agent computing in the social sciences", *Proceedings of the workshop on agent simulation: applications, models and tools*, Argonne National Laboratory (IL), 2000.
- BACKMAN O., EDLING C R., "Mathematics matters: on the absence of mathematical models in quantitative sociology", *Acta Sociologica*, 42(1), 1999.
- BAINBRIDGE W., BRENT E., CARLEY K., HEISE D., "Artificial social intelligence", *Annual Review of Sociology*, 20, 1994.
- BALLARINO G., COBALTI A., *Mobilità sociale*, Roma, Carocci, 2003.
- BARBERA F., "Quattro secoli di individualismo metodologico: bilancio e prospettive", *Rassegna italiana di sociologia*, 43(2), 2002.

- BARBUT M., « Note sur quelques indicateurs globaux de l'inégalité », *Revue française de sociologie*, XXV(4), 1984.
- BARBUT M., « Commentaire sur la note de J-C. Combessie », *Revue française de sociologie*, XXVI(1), 1985(a).
- BARBUT M., « Sur les indicateurs de l'inégalité : croissance logistique et mesure de l'inégalité et de quelques effets 'paradoxaux' dans la comparaison des inégalités », *Mathématiques et Sciences humaines*, 90(2), 1985(b).
- BARBUT M., « Introduction à l'analyse mathématique des inégalités », *Archives de psychologie*, 58, 1990.
- BARBUT M., « Sur la formalisation dans les sciences sociales », *Histoire & Mesure*, IX(1-2), 1994.
- BARBUT M., « Les mathématiques et les sciences humaines. Esquisse d'un bilan », Baechler J., Chazel F., Kamrane R. (éds.), *L'acteur et ses raisons. Mélanges en l'honneur de Raymond Boudon*, Paris, Presses Universitaires de France, 2000.
- BARBUT M., « Homme moyen ou homme extrême ? De Vilfredo Pareto (1896) à Paul Lévy (1936) en passant par Maurice Fréchet et quelques autres », *Journal de la Société Française de Statistique*, 144(1-2), 2003.
- BARONE C., "La teoria della scelta razionale e la ricerca empirica. Il caso delle disuguaglianze educative", *Rassegna Italiana di Sociologia*, 3, 2005.
- BARONE C., "A new look at schooling inequalities in Italy and their trends over time", article présenté au RC28 Spring Meeting, Nijmegen, 11-14 Mai 2006.
- BEAUD S., *80 % au bac... et après? Les enfants de la démocratisation scolaire*, Paris, La Découverte, 2003.
- BECKER R., "Educational expansion and persistent inequality of education. Utilizing subjective expected utility theory to explain increasing participation rates in upper secondary school in the Federal republic of Germany", *European sociological review*, 19(1), 2003.
- BENKIRANE R., *La complexité, vertiges et promesses*, Paris, Le Pommier, 2002.
- BERNARDI F., "Chi sposa chi? Omogamia educativa e occupazionale in Italia", Schizzerotto A., (ed.), *Vite Ineguali. Disuguaglianze e corsi di vita nell'Italia contemporanea*, Bologna, Il Mulino, 2002.
- BLAU P., "A macrosociological theory of social structure", *American Journal of Sociology*, 83, 1977.
- BLOSSFELD H.-P., SHAVIT Y., "Persisting barriers. Changes in educational opportunities in thirteen countries", Shavit Y., Blossfeld H.-P. (eds.), *Persistent inequality: changes educational attainment in thirteen countries*, Westview press, Boulder, 1993.
- BOERO R., CASTELLANI M., SQUAZZONI F., "Processi cognitivi e studio delle proprietà emergenti nei modelli ad agenti", Terna P., Boero R., Morini M., Sonnessa M. (eds.), *Modelli per la complessità - La simulazione ad agenti in economia*, Il Mulino, Bologna, 2006.
- BOERO R., SQUAZZONI F., "Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science", *Journal of Artificial Societies and Social Simulation* vol. 8, n° 4, 2005.
- BOHRNSTEDT G. W., KNOKE D., *Statistica per le scienze sociali*, Bologna, Il Mulino, 1998. [Or. *Statistics for social data analysis*, Itasca, Peacock publishers, 1994].
- BOUDON R., « Réflexion sur la logique des modèles simulés », *Archives européennes de sociologie*, VI(1), 1965.

- BOUDON R., « Simulation et analyse des processus », Boudon R. (éd.), *L'analyse mathématique des faits sociaux*, Paris, Plon, chap. IX, 1967.
- BOUDON R., *L'inégalité des chances. La mobilité sociale dans les sociétés industrielles*, Paris, Colin, 1973(a).
- BOUDON R., *Mathematical structures of social mobility*, Elsevier, London, 1973(b).
- BOUDON R., « Les modèles expérimentaux », Boudon R., Grémy J.-P. (éds.), *Les modèles en sociologie*, Paris, Lemtas, 1977(a).
- BOUDON R., *Effets pervers et ordre social*, Paris, Presses Universitaires de France, 1977(b).
- BOUDON R., "Generating models as a research strategy", Merton R. K., Coleman J. S., Rossi P. H. (eds.), *Qualitative and quantitative social research*, New-York, The free press, 1979(a).
- BOUDON R., *La logique du social*, Paris, Presses Universitaires de France, 1979(b).
- BOUDON R., « Préface » [à la seconde édition de] *La logique du social*, Paris, Presses Universitaires de France, 1983.
- BOUDON R., *La place du désordre. Critique des théories du changement social*, Paris, Presses Universitaires de France, 1984.
- BOUDON R., BOURRICAUD F., « Individualisme », Boudon R., Bourricaud F. (éds.), *Dictionnaire critique de la sociologie*, Paris, Presses Universitaires de France, 1982(a).
- BOUDON R., BOURRICAUD F., « Mobilité sociale », Boudon R., Bourricaud F. (éds.), *Dictionnaire critique de la sociologie*, Paris, Presses Universitaires de France, 1982(b).
- BOURDIEU P., *La distinction. Critique sociale du jugement*, Paris, Minuit, 1979.
- Breen R., "A rational choice model of educational inequality", *Working Paper* 2001/166, Madrid, Instituto Juan March, 2001.
- BREEN R. (ed.), *Social mobility in Europe*, Oxford, Oxford University Press, 2004.
- BREEN R., "Foundations of a neo-weberian class analysis", Wright E. O. (ed.), *Approaches to Class Analysis*, Cambridge, Cambridge University Press, 2005.
- BREEN R., GOLDTHORPE J., "Explaining educational differentials: towards a formal rational choice theory", Goldthorpe J. (ed.), *On sociology. Numbers, Narratives, and the integration of research and theory*, Oxford, Oxford University Press, 1997.
- BREEN R., JONSSON J. O., "Inequality of opportunity in comparative perspective: recent research on educational attainment and social mobility", *Annual Review of Sociology*, 31, 2005.
- BREEN R., YAISH M., "Testing the Breen-Goldthorpe model of educational decision making", Stephen L., Morgan D. B., Grusky D., Fields G. S. (eds.), *Frontiers in Social and Economic Mobility*, Stanford, Stanford University Press, 2006.
- BULLE N., "Trend and comparative analysis of inequality of social opportunity", *Quality and Quantity*, 40(1), 2005.
- BUNGE M., "Mathematical modeling in social science", Bunge M. (ed.), *Method, model and matter*, Dordrecht, Reidel, 1973.
- CARLEY K., "Sociology: computational organization theory", *Social Science Computer Review*, 12(4), 1994.
- CARLEY K., "Artificial intelligence within sociology", *Sociological Methods & Research*, 25(1), 1996.

- CARLEY K., "On generating hypotheses using computer simulations", *Proceedings of the 1999 international symposium on command and control research and technology*, Newport, 1999.
- CARLEY K., "Computational approaches to sociological theorizing", Turner J. (ed.), *Handbook of Sociological Theory*, New York, Kluwer Academic Publishers, 2001.
- CASTELFRANCHI C., "Simulating with cognitive agents : the importance of cognitive emergence", Sichman J., Conte R., Gilbert N. (eds.), *Multi-agent systems and agent based simulation*, Berlin, Springer, 1998.
- CEDERMAN L.-E., "Agent-based modelling in the political science", *Political Methodology*, 10(1), 2002.
- CHAUVEL L., « Inégalités singulières et plurielles : l'évolution de la courbe de répartition des revenus », *Revue de l'OFCE*, 5, 1995.
- CHAUVEL L., *Le destin des générations, structure sociale et cohortes en France au XX^e siècle*, Paris, Presses Universitaires de France, 1998(a).
- CHAUVEL L., « La seconde explosion scolaire : diffusion des diplômes, structure sociale et valeur des titres », *Revue de l'OFCE*, 66, 1998(b).
- CHERKAOUI M., *Les changements du système éducatif en France 1950-1980*, Paris, Presses Universitaires de France, 1982.
- CHERKAOUI M., « Mobilité sociale », Cherkaoui M., Boudon R., Besnard Ph., Lécuyer B. (éds), *Dictionnaire de Sociologie*, Paris, Larousse, 1989, [nouvelle édition 2005].
- CHERKAOUI M., « Mobilité », Boudon R. (éd.), *Traité de Sociologie*, Paris, Presses Universitaires de France, 1992.
- CHERKAOUI M., *Le paradoxe des conséquences*, Genève-Paris, Droz, 2006.
- COBALTI A., "La diseguaglianza delle opportunità educative: problemi di misura del cambiamento nel tempo", Moscati R., (ed.), *La sociologia dell'educazione in Italia. Centralità e marginalità della scuola*, Bologna, Zanichelli, 1989(a).
- COBALTI A., "A relative mobility table. A modest proposal", *Quality and Quantity*, 23, 1989(b).
- COBALTI A., "L'analyse dei dati di mobilità. La tavola di mobilità relativa", *Quaderni di sociologia*, 12, 1989(c).
- COBALTI A., "Origine sociale e livello di istruzione: un modello", *Polis*, 6(1), 1992.
- COBALTI A., *Lo studio della mobilità sociale. Metodi e prospettive dell'indagine sociologica*, Roma, Nis, 1995.
- COBALTI A., SCHIZZEROTTO A., *La mobilità sociale in Italia. L'influenza dei fattori di diseguaglianza sul destino educativo, professionale e sociale dei singoli nel nostro paese*, Il Mulino, Bologna, 1994.
- COLEMAN J. S., "Analysis of social structures and simulation of social processes with electronic computers", Guetzkow H. S. (ed.), *Simulation in social sciences: readings*, Englewood Cliffs, Prentice-Hall, 1962.
- COLEMAN J., *Introduction to mathematical sociology*, New-York, Free Press, 1964.
- COLEMAN J. S., "The use of electronic computers in the study of social organization", *Archives européennes de sociologie*, VI(1), 1965.
- COLEMAN J. S., "Social theory, social research and a theory of action", *American Journal of Sociology*, 96(6), 1986(a).
- COLEMAN J. S., *The foundations of social theory*, Cambridge, Harvard University Press, 1990.

- COLEMAN J. S., "Reply to Blau, Tuomela, Diekman and Baurmann", *Analyse und Kritik*, 15, 1993.
- COLEMAN J. S., KATS E., MENZEL H., *Medical innovation: a diffusion study*, New-York, Bobbs Merrill, 1966.
- COLLINS R., *Theoretical sociological*, Orlando, Harcourt Brace Jovanovich Inc., 1988.
- COMBESSIE J.-C., « L'évolution comparée des inégalités : problèmes statistiques », *Revue française de sociologie*, XXV, 1984.
- COMBESSIE J.-C., « Paradoxe des fonctions de concentration », *Revue française de sociologie*, XXVI, 1985.
- COUSTEAUX A.-S., LEMEL Y., « Étude de l'homophilie socioprofessionnelle à travers l'enquête Contacts », Document de travail CREST, 10, 2004.
- CREPEL P., HERTZ S., « Les petits nombres ont-ils leur loi? », *La recherche*, 13, 2003.
- CUIN C.-H., *Les sociologues et la mobilité sociale*, Paris, Presses Universitaires de France, 1993.
- DAVIDOVITCH A., BOUDON R., « Les mécanismes sociaux des abandons de poursuite judiciaire. Analyse expérimentale par simulation », *L'année sociologique*, troisième série, 1964, p. 111-244.
- DAVIES R., HEINESEN E., HOLM A., "The relative risk aversion hypothesis of educational choice", *Journal of Population Economics*, 15(4), 2002.
- DEFFUANT G., HUET S., AMBLARD F., "An individual-based model of innovation diffusion mixing social value and individual benefit", *American Journal of Sociology*, 110(4), 2005.
- DEGENNE A., FORSÉ M., *Les réseaux sociaux*, Paris, Armand Colin, 2004.
- DESSALLES J.-L., PHAN D., "Emergence in multi-agent systems: cognitive hierarchy, detection, and complexity reduction part I: methodological issues", Mathieu F., Beaufils B., Brandouy O. (eds.), *Agent-based methods in finance, game theory and their applications*, Series: Lecture Notes in Economics and Mathematical Systems, Springer, Vol. 564, 2005.
- DUPUY J.-P., *Introduction aux sciences sociales. Logiques des phénomènes collectifs*, Paris, Ellipses, 1992.
- DUPUY J.-P., « Introduction », Dupuy J.-P., Livet P., (éds.), *Les limites de la rationalité. Rationalité, éthique et cognition*, I, Paris, La Découverte, 1997.
- DURLAUF S. N., "A framework for the study of individual behaviour and social interactions", *Sociological Methodology*, 31, 2001.
- DURU-BELLAT M., *Les inégalités sociales à l'école. Genèse et mythes*, Paris, Presses Universitaires de France, 2002.
- DURU-BELLAT M., *L'inflation scolaire. Les désillusions de la méritocratie*, Paris, Seuil, 2006.
- DURU-BELLAT M., KIEFFER A., « La démocratisation de l'enseignement en France : polémiques autour d'une question d'actualité », *Population*, 55(1), 2000.
- EDLING C. R., "Mathematics in sociology", *Annual Review of Sociology*, 28, 2002.
- EPSTEIN J., "Remarks on the foundations of agent-based generative social science", *CSED Working Paper*, 41, 2005.
- ERIKSON R., GOLDTHORPE J., *The constant flux : a study of class mobility in industrial societies*, Oxford, Clarendon Press, 1992.

- ERIKSON R., GOLDTHORPE J. H., "Intergenerational Inequality: a sociological perspective", *Journal of Economic Perspectives*, 16(3), 2002.
- ESPING-ANDERSEN G., MESTRES J., "Ineguaglianza delle opportunità ed eredità sociale", *Stato e Mercato*, 67, 2003.
- ESSER H., "What is wrong with 'variable sociology'?", *European Sociological Review*, 12(2), 1996.
- FARARO T. J., "Stochastic processes", Borgatta E. F. (ed.), *Sociological Methodology*, San Francisco, Jossey-Bass, 1969.
- FARARO T. J., *Mathematical Sociology*, New York, Wiley, 1973.
- FARARO T. J., "Reflections on mathematical sociology", *Sociological Forum*, 12(1), 1997.
- FARARO T. J., « Mathématiques et sociologie », Cherkaoui M., Boudon R., Borlandi M., Valade B. (éds.), *Dictionnaire historique de la pensée sociologique*, Paris, Presses Universitaires de France, 2005(a).
- FARARO T. J., « Modèle / Modélisation », Cherkaoui M., Boudon R., Borlandi M., Valade B. (éds.), *Dictionnaire historique de la pensée sociologique*, Paris, Presses Universitaires de France, 2005(b).
- FARARO T. J., BUTTS C. T., "Advance in generative structuralism: structured agency and multilevel dynamics", *Journal of Mathematical Sociology*, 24(1) 1999.
- FLORENS J.-P., « Inégalité et dépendance statistique », *Revue française de sociologie*, XXV, 1984.
- FORSÉ M., CHAUVEL L., « L'évolution de l'homogamie en France », *Revue française de sociologie*, XXXVI, 1995.
- FORSÉ M., PARODI M., « La ségrégation spatiale selon Schelling : la perversité est ailleurs », document de travail de l'OFCE, Paris, Mars 2006.
- GAMBETTA D., *Where they pushed or did they jump? Individual decision mechanisms in education*, Cambridge, Cambridge University Press, 1987.
- GILBERT N. G., « Holism, individualism and emergent properties. An approach from the perspective of simulation », Hegselmann R., Mueller U., Troitzsch G. K., (eds), *Modelling and simulation in the social sciences from the philosophy of science point of view*, Dordrecht, Kluwer, 1996.
- GILBERT N., TROITZSCH K., *Simulation for the social scientist*, Philadelphia, Open University Press, 1999.
- GINI C., "Variabilità e mutabilità, 1912", Pizetti E., Salvemini T. (eds), *Memorie di metodologica statistica*, Roma, Libreria Eredi Virgilio Veschi, 1955.
- GINOT V., MONOD H., « Explorer les modèles par simulation : application aux analyses de sensibilité », Amblard F., Phan D., (éds), *Modélisation et simulation multi-agents : Applications pour les Sciences de l'Homme et de la Société*, Paris, Hermès (chap. 3), 2006.
- GODECHOT O., *Plus d'amis plus proches ? Comment comparer deux enquêtes peu comparables ?*, Document de travail, Insee, 2000.
- GOLDTHORPE J. H., "Class analysis and the reorientation of class theory: the case of persisting differentials in education attainment", 1996(c), Goldthorpe J., *On sociology. Numbers, narratives, and the integration of research and theory*, Oxford, Oxford University Press, 2000.
- GOLDTHORPE J. H., LLEWELLYN C., PAYNE C., *Social mobility and class structure in modern Britain*, Oxford, Clarendon Press, 1980.

- GOODMAN L., "How to 'ransack' social mobility tables and other kinds of crossclassification tables", *American Journal of Sociology*, 75, 1969.
- GRANOVETTER M., "The strength of weak tie", *American Journal of Sociology*, 78, 1973.
- GRANOVETTER M., "Threshold models of collective behavior", *American Journal of Sociology*, 83(6), 1978.
- GRANOVETTER M., SOONG R., "Threshold models of diffusion and collective behavior", *Journal of Mathematical Sociology*, 9, 1983.
- GRANOVETTER M., SOONG R., "Threshold models of diversity: chinese restaurants, residential segregation and the spiral of silence", *Sociological Methodology*, 18, 1988.
- GRÉMY J.-P., "Use of computer simulation techniques in sociology", *International Social Science Journal*, 23(2), 1971(a).
- GRÉMY J.-P., « Les techniques de simulation », Boudon R. (éd.), *Les mathématiques en sociologie*, Paris, Presses Universitaires de France, 1971(b).
- GRÉMY J.-P., « Les modèles simulables », Boudon R., Grémy J.-P. (éds), *Les modèles en sociologie*, Paris, Lemtas, 1977.
- GRÉMY J.-P., « Sur les différences entre pourcentages et leur interprétation », *Revue française de sociologie*, XXV, 1984.
- GRUSKY D., HAUSER R., "Comparative social mobility revisited: models of convergence and divergence in 16 countries", *American Sociological Review*, 49, 1984.
- HALPIN B., "Simulation in sociology", *American Behavioral Scientist*, 42(10), 1999.
- HANNEMAN R. A., "Simulation modelling and theoretical analysis in sociology", *Sociological Perspectives*, 38(4), 1995.
- HANNEMAN R. A., COLLINS R., MORDT G., "Discovering theory dynamics by computer simulation: experiments on state legitimacy and imperialist capitalism", *Sociological Methodology*, 25(1), 1995.
- HANNEMAN R. A., PATRICK S., "On the uses of computer-assisted simulation modelling in the social sciences", *Sociological Research on Line*, 2(2), 1997.
- HARTMANN S., "The world as a process. Simulation in the natural and social sciences", Hegselmann R., Mueller U., Troitzsch G. K. (eds), *Modelling and simulation in the social sciences from the philosophy of science point of view*, Dordrecht, Kluwer, 1996.
- HAYEK F., *Scientisme et sciences sociales. Essai sur le mauvais usage de la raison*, Paris, Plon, 1953. [Traduction française partielle de Hayek F., *The Counter-Revolution of Science : Studies on the Abuse of Reason*, Glencoe, Free Press, 1952].
- HAYES A. C., "Formal model building and theoretical interests in sociology", *Journal of Mathematical Sociology*, 10, 1984.
- HEDSTROM P., "Organizational vacancy chains and the attainment process", *Journal of Mathematical Sociology*, 17(1), 1992.
- HEDSTROM P., "Contagious collectivities: On the spatial diffusion of swedish trade unions, 1890-1940", *American Journal of Sociology*, 99(5), 1994.
- HEDSTROM P., "Mesolevel networks and the diffusion of social movements: the case of the swedish social democratic party", *American Journal of Sociology*, 106(1), 2000.
- HEDSTROM P., "Rational imitation", Hedstrom P., Swedberg R. (eds), *Social Mechanisms. An analytical approach to social theory*, Cambridge, Cambridge University Press, 1998.

- HEDSTROM P., *Dissecting the social: on the principles of analytical sociology*, Cambridge, Cambridge University Press, 2005.
- HEISE D., "Sociological algorithms: preface", *Journal of Mathematical Sociology*, 20(2-3), 1995.
- HELLEVIK O., "Class inequality and egalitarian reform", *Acta Sociologica*, 40(4), 1997.
- HELLEVIK O., "A less biased allocation mechanism", *Acta Sociologica* 43(1), 2000.
- HELLEVIK O., "Inequality versus association in educational attainment research: Comment on Kivinen, Ahola and Hedman", *Acta Sociologica*, 45(2), 2002.
- HILLMERT S., JACOB M., "Social inequality in higher education. Is vocational training a pathway leading to or away from university?", *European Sociological Review*, 19(3), 2003.
- HOUT M., *Mobility tables*, London, Sage, 1983.
- HUMMON N. P., FARARO T. J., "The emergence of computational sociology", *Journal of Mathematical Sociology*, 20(2-3), 1995.
- JOHNSON P. E., "Simulation modelling in political science", *American Behavioral Scientist*, 42(10), 1999.
- JONSSON J. O., ERIKSON R., "Understanding educational inequality: the Swedish experience", *L'année sociologique*, 50(2), 2000.
- KAUFMAN R. L., SCHERVISH P. G., "Using adjusted crosstabulations to interpret log-linear relationship", *American Sociological Review*, 51, October 1986.
- KAUFMAN R. L., SCHERVISH P. G., "Variations on a theme. More uses of odds ratios to interpret log-linear parameters", *Sociological Methods and Research*, 16(2), 1987.
- KELLER S., ZAVALLONI M., "Ambition and social class: a respecification", *Social Forces*, 43, 1964.
- KIVINEN O., HEDMAN J., AHOLA S., "Expanding education and improving odds. Participation in higher education in Finland in the 1980s and 1990s", *Acta Sociologica*, 44(4), 2001.
- KIVINEN O., HEDMAN J., AHOLA S., "Changes to differences in expanding higher education: reply to Hellevik", *Acta Sociologica*, 45(2), 2002.
- KÖNIG W., LÜTTINGER P., MÜLLER W., "A Comparative analysis of the development and structure of educational systems. Methodological foundations and the construction of a comparative educational scale", *CASMIN working paper*, 12, Mannheim, Universität de Mannheim, 1988.
- LAMPARD R., "Measuring inequality in a cross-tabulation with ordered categories: from the Gini coefficient to the Tog coefficient", *International Journal of Social Research Methodology*, 3(1), 2000.
- LEIK R., "Mathematical sociology", Borgatta E. F., Borgatta M. L. (eds), *Encyclopaedia of Sociology*, vol. 3, New-York, Macmillan, 1992.
- LEMEL Y., *Stratification et mobilité sociale*, Paris, Armand Colin, 1991.
- LINDENBERG S., "Individuelle Effekte, kollektive Phänomene und das Problem der Transformation", Eichner K., Habermehl W. (eds), *Probleme der Erklärung sozialen Verhaltens*, Meisenheim, Anton Hain, 1977, p. 46-84.
- LINDENBERG S., "The method of decreasing abstraction", Coleman J. S., Fararo T. (eds), *Rational choice theory. Advocacy and critique*, Newbury Park (CA), Sage, 1992.

LINDENBERG S., "The influence of simplification on explananda : phenomenon-centered versus choice-centered theories in the social sciences", Blossfeld H-P., Prein G. (eds), *Rational choice theory and large-scale data analysis*, Boulder Co., Westview Press, 1998.

LINDENBERG S., "The extension of rationality: framing versus cognitive rationality", dans Baechler J., Chazel F., Kamrane R. (eds), *L'acteur et ses raisons. Mélanges en l'honneur de Raymond Boudon*, Paris, Presses Universitaires de France, 2000.

LINDENBERG S., "Social rationality versus rational egoism", Turner J. H. (ed.), *Handbook of sociological theory*, New-York, Kluwer Academic publisher, 2002.

LINDENBERG S., « Coleman et la construction des institutions : peut-on négliger la rationalité sociale ? », *Revue française de sociologie*, 44(2), 2003.

LIVET P., « Statut épistémologique de la simulation en sciences de l'homme et de la société », Amblard F., Phan D. (éds), *Modélisation et simulation multi-agents : Applications pour les Sciences de l'Homme et de la Société*, Paris, Hermès (chap. 8), 2006.

LORENZ M. O., "Methods of measuring the concentration of wealth", *Journal of the American Statistical Association*, 9, 1905.

MACY W. M., "Social simulation: computational approaches", *International encyclopedia of the social and behavioral sciences*, Elsevier, vol. 21, 2001.

MACY W. M., WILLER R., "From factors to actors: computational sociology and agent-based modeling", *Annual Review of Sociology*, 28, 2002.

MANZO G., "Verso una teoria delle diseguaglianze di opportunità educative", *Studi di Sociologia*, 42(1), 2004, p. 79-113

MANZO G., Variables, mécanismes et simulations : une synthèse des trois méthodes est-elle possible ? Une analyse critique de la littérature », *Revue française de sociologie*, 46(1), 2005, p. 37-74.

MANZO G., "Generative mechanisms and multivariate statistical analysis. Modeling educational opportunity inequality by multi-matrix log-linear topological model: contributions and limits", *Quality and Quantity*, 40(5), 2006(a), p. 721-758.

MANZO G., *Analyse comparée de la stratification éducative en France et en Italie dans les années quatre-vingt-dix et au cours du XX^e siècle. De la description statistique à l'explication par simulation « individus-centrés »*, Thèse de doctorat de sociologie, Université Paris IV – Sorbonne, 2006(b), 681 pages.

MANZO G., « Le modèle du choix éducatif interdépendant. Des mécanismes théoriques aux données empiriques », *Archives Européennes de Sociologie*, 48(1), 2007.

MARE R., "Social background and school continuation decisions", *Journal of the American Statistical Association*, 75, 1980.

MARE R., "Change and stability in educational stratification", *American Sociological Review*, 46, 1981.

MARKS G. N., "The Measurement of socio-economic inequalities in education: a further comment", *Acta Sociologica*, 47, 2004.

MARSHALL G., SWIFT A., "On the meaning and measurement of inequality ", *Acta Sociologica*, 43(1), 1999.

MARSHALL G., SWIFT A., "Reply to Ringen and Hellevik", *Acta Sociologica*, 43(1), 2000.

MARTIN O., « Mathématiques et sciences sociales au XX^e siècle ", *Revue d'histoire des sciences humaines* », 6, 2002.

- MARTIN O., BARBUT M., « Mathématique sociale », Cherkaoui M., Boudon R., Borlandi M., Valade B. (éds), *Dictionnaire historique de la pensée sociologique*, Paris, Presses Universitaires de France, 2005.
- Mathématiques et Sciences humaines*, Numéro spécial sur « Modèles et méthodes mathématiques dans les sciences sociales : apports et limites », 172(4), 2005.
- MAURIN E., *Le ghetto français. Enquête sur le séparatisme social*, Paris, Seuil, 2004.
- MERCKLÉ P., *Sociologie des réseaux sociaux*, Paris, La Découverte, 2004.
- MERLE P., « Le concept de démocratisation de l'institution scolaire : une typologie et sa mise à l'épreuve », *Population*, 55(1), 2000.
- MERLLIÉ D., « Analyse de l'interaction entre variables. Problème statistique ou sociologique », *Revue française de sociologie*, XXVI, 1985.
- MERLLIÉ D., *Les enquêtes de mobilité sociale*, Paris, Presses Universitaires de France, 1994.
- MERLLIÉ D., PRÉVOT J., *La mobilité sociale*, Paris, La découverte, 1997.
- MOREL G., « Modéliser l'évolution de l'inégalité des taux de scolarisation », *Mathématiques et Sciences humaines*, 164(4), 2003.
- MORETTI S., *Modelli e conoscenza scientifica. Problemi di formalizzazione nella ricerca sociologica*, Milano, Guerini scientifica, 2004.
- MORGAN M., « Calcul et dynamique : de l'informatique aux systèmes complexes », [communication au] *Premier colloque d'histoire et épistémologie des mathématiques discrètes*, CAMS, Paris, 11-12 Mars 2005.
- MORIN E., *Sociologie*, Paris, Seuil, 1984. [Nouvelle édition, 1994, cité dans cette édition].
- MORIN E., *Introduction à la complexité*, Paris, Seuil, 1990. [Nouvelle édition, 2005, cité dans cette édition].
- MORIN E., LE MOIGNE J.-L., *L'intelligence de la complexité*, Paris, l'Harmattan, 1999.
- MOSS S., EDMONDS B., "Sociology and simulation: statistical and qualitative cross-validation", *American Journal of Sociology*, 110(4), 2005.
- MÜLLER W., KARLE W., "Social selection in educational systems in Europe", *European Sociological Review*, 9, 1993.
- MÜLLER W., LÜTTINGER P., KÖNIG W., KARLE W., "Class and education in industrial nations", *International Journal of Sociology*, 19, 1989.
- NADAL J.-P., GORDON M., « Physique statistique de phénomènes collectifs en sciences économiques et sociales », *Mathématiques et Sciences humaines*, 172(2), 2005.
- NASH R., "Inequality/difference in education: is a real explanation of primary and secondary effects possible?", *British Journal of Sociology*, 54(4), 2003.
- ORLÉANS A., *Le pouvoir de la finance*, Paris, Odile Jacob, 1999.
- ORLÉANS A., « Les interactions mimétiques », Lesourne J., Orléans A., Walliser B., (éds), *Leçons de microéconomie évolutionniste*, Paris, Odile Jacob, 2002.
- PEDERSEN A. W., "Inequality as relative deprivation: a sociological approach to inequality measurement", *Acta Sociologica*, 47, 2004.
- PHAN D., "From agent-based computational economics towards cognitive economics", Bourguine P., Nadal J.-P. (éds), *Cognitive Economics : An Interdisciplinary Approach*, Berlin, Springer, 2004.

- PHAN D., PAJOT S., "Complex behaviours in binary choice models with global or local social influence", Bruun C. (ed.), *Advances in artificial economics. The economy as a complex dynamic system*, Series: *Lecture Notes in Economics and Mathematical Systems*, Springer, Vol. 584, 2006.
- PISATI M., SCHIZZEROTTO A., « La mobilité sociale en Italie », *La Revue Tocqueville / the Tocqueville Review*, 19(1), 1998.
- POWERS D. A., XIE Y., *Statistical methods for categorical data analysis*, San Diego (CA), Academic Press, 2000.
- WONG R. S.-K., *Introduction to log-linear modeling*, University of California-Santa Barbara, 2003(a). [Disponible à : www.socwkp.sinica.edu.tw].
- PRÉVOT J., « À propos d'indices et de comparaisons de proportions », *Revue française de sociologie*, XXVI, 1985.
- RAFTERY A. E., "Statistics in sociology, 1950-2000: a selective review", *Sociological Methodology*, 31, 2001.
- RAFTERY A. E., HOUT M., "Maximally maintained inequality: expansion, reform and opportunity in Irish education, 1921-1975", *Sociology of Education*, 66(41-62), 1993.
- RAMAT E., « Introduction à la modélisation et à la simulation à événements discrets », Amblard F., Phan D. (éds), *Modélisation et simulation multi-agents : applications pour les sciences de l'homme et de la société*, Paris, Hermès, (chap. 2), 2006.
- RAUB W., "The Structural-Individualist approach: towards an explanatory sociology", Raub W. (ed.), *Theoretical models and empirical analyses*, Utrecht, E. S. Publications, 1982.
- RINGEN S., "Inequality and its Measurement", *Acta Sociologica*, 43(1), 2000.
- RINGEN S., "The Truth about class inequality", 2005. [Disponible à : <http://users.ox.ac.uk/~gree0074/documents%20en/truthclass.doc>].
- SANDERS L., « Les modèles agents en géographie urbaine », Amblard F., Phan D. (éds), *Modélisation et simulation multi-agents : applications pour les sciences de l'homme et de la société*, Paris, Hermès, (chap. 7), 2006.
- SAWYER R. K., "Artificial societies. Multiagent systems and the micro-macro link in sociological theory", *Sociological Methods and Research*, 31(3), 2003.
- SAWYER R. K., "Social explanation et computational simulation", *Philosophical Explorations*, 7(3), 2004(a).
- SAWYER R. K., "The mechanisms of emergence", *Philosophy of the Social Sciences*, 34(2), 2004(b).
- SAWYER R. K., *Social emergence. Societies as complex systems*, Cambridge, Cambridge University Press, 2005.
- SCHADEE H. M. A., SCHIZZEROTTO A., "Processi di mobilità maschili e femminili nell'Italia contemporanea", *Polis*, 1, 1990.
- SCHELLING T. C., "Dynamic models of segregation", *Journal of Mathematical Sociology*, 1, 1971.
- SCHELLING T. C., "Hockey helmets, concealed weapons and daylight saving. A study of binary choices with externalities", *Journal of Conflict Resolution*, 17, 1973.
- SCHELLING T. C., *Micromotives and macrobehaviour*, New-York, Norton & Co., 1978. [Traduction française, *La tyrannie des petites décisions*, Paris, Presses Universitaires de France, 1980).

SCHIZZEROTTO A., "Perché in Italia ci sono pochi diplomati e pochi laureati ? Vincoli strutturali e decisioni razionali degli attori come cause della contenuta espansione della scolarità superiore", *Polis*, 11(3), 1997.

SCOTT L. F., "Mathematics in thinking about sociology", *Sociological Forum*, 12(1), 1997.

SIMON A. H. *The sciences of the artificial*, Cambridge, MIT Press, 1996. [Traduction française, *Les sciences de l'artificiel*, Paris, Gallimard, 2004, cité dans l'édition française].

SIMON A. H., "Coping with complexity", GRASCE (éd.), *Entre systémique et complexité, chemin faisant... Mélanges en l'honneur du Professeur Jean-Louis Le Moigne*, Paris, Presses Universitaires de France, 1999.

SMITS J., ULTEE W., LAMMERS J., "Educational homogamy in 65 countries: an explanation of differences in openness using country-level explanatory variables", *American Sociological Review*, 63, 1998.

Sociological Forum, [Special issue]: "Mathematics in thinking about sociology", 12(1), 1997.

Sociological theory, "A Symposium on Formal Theory", 18(3), 2000.

SORENSEN A.; "Models and strategies in research on attainment and opportunity", *Social Science Information*, 15(1), 1976.

SORENSEN A.; "Theoretical mechanisms and the empirical study of social processes", Hedstrom P., Swedberg R. (eds), *Social mechanisms. An analytical approach to social theory*, Cambridge, Cambridge University Press, 1998.

SQUAZZONI F., "Review of, Hedstrom P., dissecting the social: on the principles of analytical sociology", *Journal of Artificial Societies and Social Simulation*, 9(2), 2006.

STEYER A., ZIMMERMANN J.-B., « Influence sociale et diffusion de l'innovation », *Mathématiques et Sciences humaines*, 168(4), 2004.

TARDE G., *Les lois de l'imitation*, Paris, Seuil, coll. Les empêcheurs de penser en rond, 2001.

The Journal of Mathematical Sociology, [special issue]: "Mathematical ideas and sociological theory", 10(3-4), 1984.

The Journal of Mathematical Sociology, [special issue]: "Computer simulation in sociology", 15(2), 1990.

THÉLOT C., *Tel père, tel fils ? Position sociale et origine sociale*, Paris, Dunod, 1982. [Nouvelle édition, 2003, cité dans cette édition].

THÉLOT C., « Préface », Thélot C. (éd.), *Tel père, tel fils ? Position sociale et origine sociale*, Paris, Hachette, 2003.

UDEHN L., *Methodological individualism. Background, history and meaning*, London, Routledge, 2001.

UDEHN L., "The changing face of methodological individualism", *Annual Review of Sociology*, 28, 2002.

VALENCE A., *Changement structurel endogène et complexité. Propriétés émergentes des dynamiques du travail, du capital et des salaires*, thèse de doctorat, EHES, 17 décembre 2004.

VALÉRY P., « Mauvaises pensées et autres », *Œuvres*, tome II, Paris, Gallimard, Bibliothèque de La Pléiade, 1960.

VALLET L.-A., « L'inégalité des chances : un modèle d'évolution », *Revue française de sociologie*, XXIX(3), 1988.

- VALLET L.-A., « Fluidité », 1989, dans Cherkaoui M., Boudon R., Besnard Ph., Lécuyer B. (éds) , *Dictionnaire de sociologie*, Paris, Larousse [nouvelle édition, 2005].
- VALLET L.-A., « Modélisation log-linéaire et log-multiplicative des tableaux de contingence - Éléments de Cours », *Point-Recherche CREST*, 20, 2004.
- WALTER C., « La dictature des valeurs extrêmes », *La Recherche*, 13, 2003.
- VERMUNT J. K., *Log-linear event history analysis : a general approach with missing data, unobserved heterogeneity and latent variables*, Tilburg, Tilburg University Press, 1996.
- VERMUNT J. K., *Lem : A general program for the analysis of categorical data*, Tilburg University, 1997. [Disponible à : <http://spitswww.uvt.nl/~vermunt/>].
- WATTS D. J., *Small worlds: the dynamics of networks between order and randomness*, Princeton, Princeton University Press, 1999.
- WEISBUCH G., *Complex systems dynamics*, Redwood City, Addison Wesley, 1989.
- WEISBUCH G., « Systèmes complexes en sciences humaines », 2003. [Disponible à : <http://www.lps.ens.fr/~weisbuch/>].
- WHITE H. C., *An Anatomy of Kinship*, Prentice-Hall, 1963.
- WIPPLER R., “The structural-individualist approach in Dutch sociology”, *The Netherlands Journal of Sociology*, 14, 1978.
- WIPPLER R., “Explanatory sociology: the development of a theoretically oriented research program”, *The Netherlands Journal of Sociology*, 21, 1985, p. 63-74.
- XIE Y., “The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables”, *American Sociological Review*, 57, 1992.

ANNEXE A

	1b+1c	2a+2b	2c	3a+3b
France	0.10	0.25	0.44	0.79
Italie	0.01	0.06	0.26	0.66

Tableau A. Valeurs d'initialisation des « agents » sur RO_D
(« Diplôme » en colonne) (cf. § 3, point 1)

Légende. Diplômes : 1a = absence de diplôme ; 1b+c = diplômes de type CEP ou CAP ; 2a+2b = diplômes de type BEPC ou BEP ; 2c = diplômes secondaires longs ; 3a+3b = diplômes de l'enseignement supérieur.

Note de lecture. Les valeurs numériques représentent la proportion (empirique) d'individus qui atteignent le groupe socioprofessionnel « cadres, professions libérales, chef d'entreprise » à partir de tel ou tel diplôme.

ANNEXE B

France, RS _{iGD}				
	1b+1c	2a+2b	2c	3a+3b
I+II (1)	0.70 (0.25)	0.90 (0.25)	1.10 (0.25)	1.50 (0.25)
IIIa+IIIb (2)	0.62 (0.25)	0.65 (0.25)	0.45 (0.25)	0.40 (0.25)
IV (3)	0.57 (0.25)	0.50 (0.25)	0.35 (0.25)	0.30 (0.25)
V+VI+VIIa,b (4)	0.45 (0.25)	0.35 (0.25)	0.30 (0.25)	0.05 (0.25)
Italie, RS _{iGD}				
	1b+1c	2a+2b	2c	3a+3b
I+II (1)	0.75 (0.25)	0.95 (0.25)	1.00 (0.25)	0.65 (0.25)
IIIa+IIIb (2)	0.68 (0.25)	0.85 (0.25)	0.88 (0.25)	0.45 (0.25)
IV (3)	0.55 (0.25)	0.70 (0.25)	0.65 (0.25)	0.35 (0.25)
V+VI+VIIa,b (4)	0.50 (0.25)	0.64 (0.25)	0.55 (0.25)	0.20 (0.25)

France, C _{iGD}				
	1b+1c	2a+2b	2c	3a+3b
I+II (1)	1.58 (0.25)	1.74 (0.25)	1.92 (0.25)	2.60 (0.25)
IIIa+IIIb (2)	1.71 (0.25)	1.83 (0.25)	2.00 (0.25)	2.80 (0.25)
IV (3)	1.75 (0.25)	1.86 (0.25)	2.05 (0.25)	2.93 (0.25)
V+VI+VIIa,b (4)	1.87 (0.25)	1.92 (0.25)	2.08 (0.25)	3.45 (0.25)
Italie, C _{iGD}				
	1b+1c	2a+2b	2c	3a+3b
I+II (1)	1.38 (0.25)	1.55 (0.25)	1.68 (0.25)	2.32 (0.25)
IIIa+IIIb (2)	1.41 (0.25)	1.59 (0.25)	1.69 (0.25)	2.33 (0.25)
IV (3)	1.55 (0.25)	1.66 (0.25)	1.72 (0.25)	2.34 (0.25)
V+VI+VIIa,b (4)	1.70 (0.25)	1.75 (0.25)	1.83 (0.25)	2.36 (0.25)

Tableau B. Moyenne et variance (entre parenthèses) des distributions de probabilité (de forme normale) à partir desquelles on a construit les valeurs d'initialisation (qui suivent une distribution de forme log-normale) des « agents » sur RS et C

Légende. Groupe sociaux : I+II = « cadres supérieurs, profession libérales, chef d'entreprise » ; IIIa+IIIb = « professions intermédiaires » ; IV = « autonomes » ; V+VI+VIIa,b = « ouvriers ». *Diplômes* : 1a = « absence de diplôme » ; 1b+1c = diplômes de type CEP ou CAP ; 2a+2b = diplômes de type BEPC ou BEP ; 2c = diplômes secondaires longs ; 3a+3b = diplômes de l'enseignement supérieur.

Note I : le lecteur constatera que, quant à RS_{iGD}, les moyennes des distributions respectent les contraintes suivantes : 1) chaque diplôme est d'autant mieux évalué que le groupe social de l'acteur i est élevé ; 2) chaque groupe social G a une évaluation prioritaire pour un segment spécifique de la hiérarchie des diplômes (valeur en gras). En revanche, pour C_{iGD} les contraintes imposées sont les suivantes : 1) un diplôme est d'autant plus coûteux qu'il se situe aux rangs supérieurs de la hiérarchie des diplômes ; 2) ces coûts sont d'autant plus soutenables que le groupe d'où l'agent est issu se situe aux rangs supérieurs de la hiérarchie des positions sociales

Note II : les valeurs figurant dans ce tableau sont le résultat final d'un travail extrêmement long d'ajustements progressifs et incrémentaux. Comme dans le cas de tout « modèle statistique », les valeurs des paramètres ne sont que le résultat d'un processus itératif de rapprochements progressifs aux données empiriques. En informatique, on qualifie cette phase d'« identification » (des paramètres) ou de « calibration » (du modèle) (cf. Amblard, 2006 ; Ginot, Monod, 2006). Dans notre cas, ce processus n'a pas été automatisé mais, bien au contraire, le fruit d'un travail patient, par « essais et erreurs », de recherche des « bonnes valeurs ». C'est d'ailleurs l'absence de procédures automatiques standardisées d'estimation des paramètres d'un modèle traité par simulation qui rend certains auteurs sceptiques quant à la viabilité de cette technique (cf. par exemple, [Raftery, 2001, p. 27]). Il est certainement vrai qu'en absence d'algorithmes d'estimation statistiquement fondés, il est plus difficile d'évaluer l'ajustement du modèle et de le comparer à des modèles alternatifs. Il est vrai aussi que le type d'estimation manuelle à laquelle nous avons été obligé nous a contraint à rester toujours près de la théorie et à réfléchir à la signification des valeurs à attribuer aux paramètres.