

Cet article est disponible en ligne à l'adresse :

http://www.cairn.info/article.php?ID_REVUE=REL&ID_NUMPUBLIE=REL_732&ID_ARTICLE=REL_732_0173

Du paradoxe libéral-parétien à un concept de métaclassement des préférences

par Herrade IGER SHEIM

| De Boeck Université | Recherches économiques de Louvain

2007/2 - N° 73

ISSN 0770-4518 | ISBN 2-8041-5452-3 | pages 173 à 192

Pour citer cet article :

— Igersheim H., Du paradoxe libéral-parétien à un concept de métaclassement des préférences, Recherches économiques de Louvain 2007/2, N° 73, p. 173-192.

Distribution électronique Cairn pour De Boeck Université.

© De Boeck Université. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

Du paradoxe libéral-parétien à un concept de métaclassement des préférences

Herrade Igersheim *

1 Introduction

Nous assistons depuis ces trente dernières années à l'introduction de la liberté en économie normative. Depuis le début des années 1970, elle s'impose en effet comme un critère fondamental d'appréciation des différents états sociaux. D'après Fleurbaey, Gravel, Laslier et Trannoy (1998), on distingue actuellement trois courants de recherche traitant de la liberté en économie normative : la théorie des droits, la littérature dite du "*freedom of choice*" et un troisième courant issu de la philosophie davantage axé sur l'équité.

Notre étude a pour cadre le premier de ces courants de recherche: la théorie des droits. Sen est le premier à introduire la notion de droits et de libertés individuels dans la théorie du choix social et à en tirer les conséquences au sein de ce champ disciplinaire dont les cadres ont été dessinés par Arrow (1951). En effet, le paradoxe libéral-parétien de Sen (1970 a,b) illustré par un exemple désormais classique, celui de *L'amant de Lady Chatterley*, est le point de départ de pratiquement tous les débats concernant la façon de concevoir les droits individuels.

Pour Sen, la fonction essentielle du conflit Pareto-liberté est de dénoncer les abus de la Pareto-optimalité et non de souligner l'importance des droits et des libertés individuels. Pourtant, l'abondante littérature qui a suivi la parution du paradoxe a globalement axé sa réflexion sur les droits

* CNRS et CEPERC, Université de Provence, 29 avenue Robert Schuman, 13621 Aix-en-Provence Cedex 1, France (e-mail : igersheim@cournot.u-strasbg.fr).

Je remercie vivement Claude d'Aspremont, Ragip Ege, Isabelle Maret, Anne Rozan et deux rapporteurs anonymes pour leurs conseils et commentaires. Je reste naturellement responsable de toutes les erreurs et imprécisions qui subsistent.

Une partie de ce travail a été financée par le Ministère espagnol de l'Education et des Sciences via le projet SEJ2005-01481/ECON et FEDER.

et les libertés. Pour beaucoup, le résultat négatif de Sen n'est dû qu'à la définition de la liberté et des droits qu'il propose. Les travaux de Gibbard (1974) et de Nozick (1974) notamment, consacrés à la question des droits et développés peu de temps après ceux de Sen, ne sont pas étrangers à cette concentration. Ces derniers s'attaquent aux insuffisances de la définition suggérée par Sen : d'après Nozick (1974), le formalisme arrovien que Sen utilise est inapte à représenter les droits et libertés individuels. Gaertner, Pattanaik et Suzumura (1992), quant à eux, rejettent le formalisme arrovien et ses extensions et expriment les droits en termes de jeux. Cette approche, aujourd'hui prépondérante, a été peu à peu reconnue implicitement comme la plus apte à rendre compte des droits individuels, tandis que celle de Sen a été reléguée au second plan. La littérature la plus récente en témoigne clairement (voir, par exemple, Deb, Pattanaik et Razzolini (1997) ou encore Van Hees (1999)). Comment expliquer ce phénomène ? Nous croyons que le paradoxe libéral-parétien plaide en faveur, non pas d'une approche en termes de jeux, mais d'un nouveau formalisme en termes de métaclassement des préférences.

C'est dans cette optique que nous nous proposons de développer notre contribution : il s'agit de revenir sur cet aspect de l'histoire récente de la théorie du choix social et d'avancer la thèse selon laquelle le paradoxe libéral-parétien prêche pour l'élaboration d'un formalisme au sein duquel une place est accordée aux valeurs et aux motivations individuelles. Après avoir démontré ce point, nous suggérons une méthode de construction des métaclassements. Celle-ci, fondée sur la notion de cohérence entre les préférences et les valeurs individuelles, permet d'aboutir à un résultat de possibilité pour le paradoxe libéral-parétien.

Dans une deuxième section, après avoir exposé formellement le paradoxe libéral-parétien, nous montrons que les deux ensembles de solutions proposés tout d'abord à celui-ci et qui semblent *a priori* diamétralement opposés reposent finalement sur une volonté similaire : celle de considérer les motivations et les valeurs des individus au-delà des préférences qu'ils expriment. Dans une troisième section, nous expliquons en quoi le fait de tenir compte des aspirations et des valeurs individuelles justifie un recours au concept de métaclassement des préférences. Enfin, dans une quatrième section, nous développons une méthode de construction des métaclassements : celle-ci mène à une résolution du paradoxe libéral-parétien.

2 L'analyse des résolutions du paradoxe de Sen

Après avoir indiqué les notations utilisées, nous présentons le paradoxe libéral-parétien et deux de ses tentatives de résolution, l'une de Sen (1976), l'autre de Gibbard (1974). Sans apporter de solution réelle au paradoxe, nous montrons qu'elles révèlent le véritable enjeu du conflit Pareto-liberté : souligner

les insuffisances du formalisme arrovien quant à l'expression des aspirations individuelles. Le paradoxe de Sen incite, ainsi, à l'élaboration d'un nouveau formalisme qui prendrait en compte les motivations individuelles.

2.1 Le paradoxe libéral-parétien

Soit X , l'ensemble fini des états sociaux possibles. $|X|$ indique le cardinal de X et $|X| \geq 3$. On considère un ensemble fini d'individus de cardinal n : $N = \{1, 2, \dots, n\}$ avec $n \geq 2$. R_i correspond à la relation de préférence de l'individu $i \in N$ sur les états sociaux. Pour tout $x, y \in X$, $(x, y) \in R_i$ signifie "x est au moins aussi bon que y". $P(R_i)$ et $I(R_i)$ désignent respectivement les relations de préférence stricte [$(x, y) \in P(R_i) \Leftrightarrow (x, y) \in R_i$ et $(y, x) \notin R_i$] et d'indifférence [$(x, y) \in I(R_i) \Leftrightarrow (x, y) \in R_i$ et $(y, x) \in R_i$] de R_i . Nous supposons que R_i décrit un ordre sur X , c'est-à-dire que R_i est une relation complète [$\forall x, y \in X : (x, y) \in R_i \vee (y, x) \in R_i$]¹ et transitive [$\forall x, y, z \in X : (x, y) \in R_i$ et $(y, z) \in R_i$ implique $(x, z) \in R_i$]. Une n -liste d'ordres individuels de préférence $\{R_i\}_{i=1}^n$ est appelée un profil, tandis que \mathcal{R} représente l'ensemble de tous les profils possibles sur X . Une règle de choix collectif f spécifie une relation de préférence sociale R pour chaque profil : $R = f(\{R_i\}_{i=1}^n)$. Comme pour R_i , P et I sont les parties asymétrique et symétrique de R . Si R est toujours un ordre, f est nommée fonction de bien-être social ("social welfare function" ou SWF d'après la terminologie arrovienne). Si R est seulement complète et acyclique [$\forall x_1, \dots, x_j \in X : \{(x_1, x_2) \in P(R) \& (x_2, x_3) \in P(R) \& \dots \& (x_{j-1}, x_j) \in P(R)\} \Rightarrow (x_1, x_j) \in R$], alors f est appelée fonction de décision sociale ("social decision function" ou SDF selon Sen (1970a), p. 52).

Le paradoxe libéral-parétien repose sur l'incompatibilité de trois conditions imposées à la règle de choix collectif et qui, *a priori*, semblent raisonnables. Tout d'abord, on définit la condition "classique" du domaine non restreint :

Condition 1. (U) *Domaine non restreint.* Le domaine de f est l'ensemble \mathcal{R} .

La deuxième condition invoquée par Sen est la condition de Pareto faible utilisée par Arrow (1951) dans le théorème d'impossibilité. Critère d'unanimité par excellence, elle indique que si tous les individus préfèrent strictement un état social à un autre, alors la règle de choix collectif se doit de refléter cette préférence partagée par tous :

Condition 2. (P) *Pareto faible.* Pour tout $x, y \in X$, si $(x, y) \in \bigcap_{i \in N} P(R_i)$, alors $(x, y) \in P(R)$.

En outre, Sen définit une condition minimale de libéralisme qui s'inspire des théories de Mill (1859) et de Hayek (1960) faisant appel à la notion

¹ Le signe \vee est le symbole du "ou" inclusif.

de sphère protégée (“*protected sphere*”) d’un individu. Cette condition établit qu’au moins deux individus de la société doivent être décisifs – c’est-à-dire que la règle de choix collectif doit refléter leurs préférences – sur au moins une paire distincte d’états sociaux chacun. En d’autres termes, ces individus jouissent d’un droit à faire valoir leurs préférences sur ces paires. Mais, ainsi imposée à la règle de choix collectif et associée à la condition U, cette condition de libéralisme “naïf” est susceptible d’entraîner, lorsqu’elle implique plus de deux individus ou plus d’une paire d’états sociaux par individu, un second paradoxe, celui de l’incohérence interne des droits (Gibbard (1974)).

Dès lors, beaucoup d’auteurs, qui s’attachent moins à l’étude du paradoxe de Gibbard qu’à celle du paradoxe de Sen, se contentent de supposer une attribution cohérente des droits individuels. Il en va ainsi, par exemple, de Farrell (1976), de Sen (1976) ou encore de Suzumura (1978) :

Définition 1. Attribution cohérente des droits. *Au moins deux individus i et j de N ($i \neq j$) sont dotés d’ensembles non vides D_i et D_j de paires distinctes d’états sociaux ; une telle attribution est dite cohérente s’il n’existe pas de suite de paires d’états sociaux $\{(x^t, y^t)\}_{t=1}^t$ ($t \geq 3$) telle que :*

- (i) $(x^t, y^t) \in \bigcup_{i=1}^n D_i \cap P(R_i)$ pour tout $t = 1, 2, \dots, t$;
- (ii) $x^1 = y^t$ et $x^t = y^{t-1}$ pour tout $t = 2, \dots, t$.

A cette définition est adjointe la condition L+, considérée comme une condition de libéralisme “cohérent” :

Condition 3. (L+) Libéralisme cohérent. *Pour toute attribution cohérente des droits, pour tout $x, y \in X$ et pour tout $i \in N$, si $(x, y) \in D_i \cap P(R_i)$, alors $(x, y) \in P(R)$.*

Le paradoxe libéral-parétien repose sur le conflit des trois conditions précédentes, le domaine non restreint, le principe de Pareto faible et la condition de libéralisme “cohérent” : il n’existe pas de SDF satisfaisant aux conditions U, P et L+.

2.2 Deux résolutions du paradoxe de Sen

Nous examinons ici uniquement les tentatives de résolution du paradoxe qui restent globalement dans le cadre du formalisme arrovien² et tentons de montrer que les solutions proposées par Sen (1976) et Gibbard (1974), représentatives de l’ensemble des travaux de cette littérature, justifient, par leur incessant recours aux motivations supposées des individus, l’emploi d’une structure de préférences plus riche.

² En particulier, la critique libertarienne n’est pas considérée.

2.2.1 Affaiblir la condition de Pareto

Avec le paradoxe libéral-parétien, Sen désire démontrer les conséquences parfois peu souhaitables du concept élaboré par Pareto. Pour éviter tout malentendu, notons qu'il ne s'agit en aucun cas de remettre en cause le critère d'unanimité inhérent au principe parétien, mais de discuter les valeurs que ce dernier véhicule : "the Pareto principle (i) lists a set of virtues, and (ii) uses dominance of virtues as the criterion. What is in dispute here is the former, not the latter" (Sen (1979), p. 554).

L'hostilité de Sen à l'égard du principe de Pareto ne peut que l'amener à proposer une résolution du paradoxe exigeant l'affaiblissement de la condition P. Sen (1976) va donc s'employer non pas à attaquer frontalement la condition de Pareto, ce qui du reste serait malaisé, mais à introduire des considérations supplémentaires dont la règle de choix collectif va devoir tenir compte. Et c'est à l'appui d'un simple exemple qu'il tente de mener cette entreprise à bien.

"Lire ou ne pas lire *L'amant de Lady Chatterley*" : grâce à deux personnages (1 et 2) et un livre *L'amant de Lady Chatterley*³, Sen a su introduire en quelques lignes la question des droits et libertés individuels en théorie du choix social. Ainsi, cette société comporte trois états sociaux : l'individu 1 lisant ce livre (a), l'individu 2 le lisant (b) ou personne ne le lisant (c). Nous avons donc : $X = \{a, b, c\}$. 1, personnage prude, préfère c à a et a à b , 2 étant par trop influençable. 2, le lascif, aimerait par contre que 1 lise le livre afin qu'il soit choqué. Ainsi, il préfère a à b et b à c . De plus, le choix de lire le livre ou pas peut être considéré comme un choix personnel, donc il est vraisemblable de supposer que l'individu 1 est décisif sur la paire d'états sociaux $\{c, a\}$, tandis que l'individu 2 est décisif sur $\{b, c\}$. Par la condition L+, on obtient alors $(c, a) \in P(R)$ et $(b, c) \in P(R)$. En outre, la condition P aboutit à $(a, b) \in P(R)$. Dès lors, un cycle apparaît : il n'existe pas de SDF satisfaisant aux conditions U, P et L+.

Quel est l'état social qui paraît le plus acceptable au regard de cet exemple ? Quelle est la condition qu'il semble bon d'affaiblir ? On voit bien ici que privilégier un critère parétien et opter pour a ne paraît pas être la solution idéale. En effet, choisir a reviendrait à nier les droits individuels et à retenir la solution qui ne satisfait aucun des individus en leur faisant faire l'inverse de leur désir. La solution b respectant les droits individuels et allant à l'encontre du critère de Pareto semble être la plus souhaitable.

Malgré tout, supprimer purement et simplement le critère de Pareto est une démarche difficile à justifier. C'est la raison pour laquelle Sen (1976) va raisonner de façon plus subtile : selon lui, il existe une distinction entre

³ Roman de l'auteur anglais D. H. Lawrence (1885-1930), écrit au milieu des années 1920 et censuré en Grande-Bretagne jusque dans les années 1970 car considéré comme un ouvrage pornographique.

les préférences d'un individu sur une paire d'états sociaux et le fait que cette personne désire que ses préférences *soient prises en compte* dans la règle de choix collectif, en particulier si celles-ci concernent un autre individu. Mais Sen ne s'arrête pas en si bon chemin. En effet, si l'on suspend le raisonnement à ce niveau, si 1, dans l'exemple précédent, ne souhaite pas que sa préférence pour c sur b agisse, il n'en préférera pas moins c à a et a à b . Dès lors, puisque les préférences individuelles sont transitives, le conflit persistera malgré ce pseudo renoncement. Sen justifie la suite ainsi : "extending this reasoning, I may decide, *for the sake of consistency*, not to insist that my preferences be taken into account *even in choices over some pairs that are not exclusively your concern*" (Sen (1976), p. 236, nous soulignons). Dans l'exemple de *L'amant*, cela se traduit de la façon suivante : non seulement 1 renonce à sa préférence pour c sur b , car il s'agit du choix personnel de 2, mais en plus, *par souci de cohérence*, 1 doit renoncer également à sa préférence pour a sur b . La seule possibilité qu'il lui reste est d'exprimer sa préférence pour c sur a , ces deux états sociaux formant sa sphère privée. Dès lors, que 2 renonce ou pas à ses préférences pour a sur c et pour a sur b , l'état social préféré par la société sera b puisque, par la condition $L+$, $(c, a) \in P(R)$ et $(b, c) \in P(R)$, le principe de Pareto n'agissant plus sur aucune paire dans cet exemple. Notons que l'on obtient le même résultat si 1 et 2 renoncent tous deux à certaines de leurs préférences, ou si 2, seul, y renonce.

2.2.2 Affaiblir la condition de libéralisme

Gibbard (1974), de son côté, propose d'affaiblir la condition de libéralisme. Axant son analyse sur les droits et les libertés individuels, il constate que la condition de libéralisme pêche par l'aspect obligatoire qu'elle revêt. En effet, le fait d'être décisif sur une paire d'états sociaux implique que l'individu exerce son droit sur cette paire suivant sa préférence immédiate, quelles qu'en soient les conséquences. La résolution suggérée par Gibbard consiste à affirmer que, dans certains cas, il est plus profitable pour l'individu de ne pas exercer son droit. A l'instar de Sen, il fonde sa proposition de résolution sur un exemple : nous la présentons brièvement grâce à l'histoire d'Edwin et d'Angelina, deux jeunes gens dont les préférences en termes de mariage ne s'harmonisent guère⁴.

Soit $N = \{1, 2\}$, l'individu 1 se nomme Edwin et l'individu 2 Angelina. Les structures de préférences individuelles sont semblables à celles de l'exemple de *L'amant*, mais il s'agit maintenant de connaître les préférences de ces individus quant au mariage. Trois états sociaux sont possibles : $X = \{E, J, O\}$ où l'état social E correspond à la situation où Edwin et Angelina se marient ; dans l'état social J , Angelina épouse un second pré-

⁴ Cet exemple est inspiré d'une opérette anglaise datant de la fin du XIX^{ème} siècle.

tendant (le Juge), tandis qu'Edwin reste célibataire ; enfin, dans l'état social O , Angelina et Edwin restent tous deux célibataires. Les préférences des individus sont les suivantes : Angelina désire se marier quoi qu'il arrive mais Edwin remporte sa préférence. Elle préfère donc E à J et J à O . Edwin, de son côté, souhaite rester célibataire, mais préfère encore se marier avec Angelina plutôt que de la voir épouser quelqu'un d'autre. Il préfère ainsi O à E et E à J . Sachant qu'Angelina désire épouser Edwin quelles que soient les circonstances, il est vraisemblable de considérer qu'Edwin est décisif sur la paire d'états sociaux $\{O, E\}$: par la condition L_+ , $(O, E) \in P(R)$. Sachant que le Juge, le second prétendant d'Angelina, souhaite ce qu'elle souhaite, Angelina, quant à elle, est décisive sur la paire d'états sociaux $\{J, O\}$: on obtient alors $(J, O) \in P(R)$. De plus, par la condition P, on a $(E, J) \in P(R)$. La même conclusion s'impose : il n'existe pas de SDF satisfaisant aux conditions U, P et L_+ .

Quelle solution adopter alors ? "At what point should the cycle be broken ?", nous demande Gibbard ((1974), p. 398). Puisqu'en termes de préférences et de droits, cet exemple est identique à celui de Sen, nous pourrions appliquer la solution précédente, c'est-à-dire affaiblir la condition P et opter pour l'état social J dans lequel Angelina épouse son second prétendant. Mais, pour Edwin, cette solution est la moins désirable. Alors que, s'il renonçait à son droit sur la paire d'états sociaux $\{O, E\}$, il épouserait Angelina et obtiendrait un résultat plus satisfaisant. Le raisonnement qui s'appliquait pour l'exemple de *L'amant* n'est plus de mise ici. Là, c'est l'option E qui est la plus attractive, l'état social qui altère les droits au profit du critère parétien. La solution envisagée par Gibbard consiste ainsi à renoncer à son droit au profit du critère parétien si les conséquences de ce renoncement sont plus avantageuses.

2.3 La prise en compte des aspirations individuelles

Les solutions proposées par Sen (1976) et Gibbard (1974) illustrent les deux tendances de la plupart des contributions qui ont suivi immédiatement la parution du paradoxe libéral-parétien. Celles-ci résolvent le conflit Pareto-liberté en modifiant les unes, la condition de Pareto (Farrell (1976), Sen (1976), Suzumura (1978), Saari (1998)...), les autres, la condition de libéralisme (Gibbard (1974), Blau (1975), Gaertner et Krüger (1981)...) comme le font les solutions de Sen et Gibbard respectivement. Celles-ci sont-elles satisfaisantes ?

En réalité, les auteurs semblent faire preuve d'une préférence conceptuelle marquée pour telle ou telle condition, celle-ci entraînant l'affaiblissement de la condition concurrente. En effet, la proposition de Sen affaiblit systématiquement la condition P, celle de Gibbard systématiquement la condition L_+ . Pourtant, Sen indique que : "to discuss whether a person's

preference should count or not we may need to know more than what the preferences happen to be, e.g. the reasons for holding these preferences” (Sen (1976), p. 237). Il insiste sur le fait qu’il existe des cas où le rejet de la condition L+ est justifié comme des cas où c’est la condition P qui est peu souhaitable. Et c’est finalement par l’examen des motivations et des causes des préférences individuelles que nous serons en mesure de déterminer quelle condition il est nécessaire d’éliminer. Dès lors, Sen rejette la proposition de Gibbard qui affaiblit systématiquement la condition L+ et lui reproche de tenir compte des “*other-regarding preferences*”. Le résultat de Farrell (1976), qui revient précisément à conserver des préférences “*self-regarding*”, répond davantage aux idées de Sen. Pourtant, il ne trouve pas grâce à ses yeux et essuie la critique inverse : considérer les droits comme systématiquement prioritaires à la condition P⁵. En outre, Sen dénonce le défaut de justification de ces propositions (en particulier, celle de Gibbard, fondée sur un seul exemple, procédé d’ailleurs discutable comme nous l’avons indiqué précédemment).

A cet égard, la proposition de Sen est sujette aux mêmes objections. Comment Sen justifie-t-il sa propre suggestion ? Simplement par le souci de cohérence ! Certes, la notion de cohérence peut éventuellement être invoquée, mais elle doit être davantage explicitée et justifiée. En outre, la proposition de Sen répond en tous points au problème posé par l’exemple de *L’amant* et à lui seulement.

Ces approches et celles qui leur sont apparentées, par leur aspect systématique et *ad hoc*, ne sont pas crédibles. Il faut envisager des cas où les individus peuvent renoncer à leurs droits et leur préférer l’efficacité parétienne et vice versa. Nous soutenons que ces tentatives de résolution du paradoxe libéral-parétien révèlent les motivations et les aspirations prêtées aux individus en fonction du contexte donné. Ainsi, bien que dans les deux exemples mentionnés, les structures de préférences individuelles soient scrupuleusement les mêmes, les situations sont complètement différentes. Or, les motivations que l’on prête aux individus sont inhérentes au contexte. Ce sont elles qui, en fin de compte, nous incitent à affaiblir soit la condition P, soit la condition L+.

Il est désormais flagrant que les propositions de Sen et Gibbard visant à résoudre le conflit Pareto-liberté plaident en faveur d’une prise en compte véritable des motivations et des valeurs individuelles. Cela signifie qu’il faudrait aller au-delà des préférences individuelles pour déterminer pourquoi les individus expriment telle ou telle préférence. Ceci nous amène à reconsidérer l’ensemble du formalisme, et pas seulement l’une ou l’autre condition. C’est ce que nous envisageons de faire dans les parties suivantes.

⁵ “One problem with Farrell’s ingenious system of ‘amendments’ lies in the fact that it too – like (...) Gibbard’s ‘alienable rights’ – tries to make social judgments based on what individual preferences happen to be, without going into the motivation that lies behind these preferences” (Sen (1976), p. 232).

3 Emergence d'un concept

D'après la théorie néoclassique, chaque individu est représenté par un ordre unique de préférence : "this is supposed to reflect his interests, represent his welfare, summarize his idea of what should be done, and describe his actual choices and behaviour" (Sen (1977), p. 99). La distinction faite par Arrow (1951) entre les "*tastes*" et les "*values*" prend alors sens : les premiers correspondent aux préférences autocentrées des individus, aux préférences ordinaires des consommateurs, tandis que les seconds ont trait aux préférences globales des individus, à leurs valeurs. Cette distinction a également été exploitée par Harsanyi (1955, 1977). Celui-ci propose dans ses travaux de recourir à une notion de préférence élargie à une dimension morale ou sociale. Dans ce cadre, on distingue les préférences éthiques d'une personne de ses préférences subjectives : "the former must express what this individual prefers (or, rather would prefer), on the basis of impersonal social considerations alone, and the latter must express what he actually prefers, whether on the basis of his personal interests or on any other basis" (Harsanyi (1955), p. 315). Harsanyi adhère à la notion smithienne selon laquelle le point de vue moral est celui d'un Observateur Impartial (et bienveillant). Il revendique également pour sa théorie l'application de la règle d'universalisation kantienne, celle-ci déterminant une procédure capable de désigner les maximes qui devront être rejetées. Selon Mongin et d'Aspremont ((1998), p. 446), "the connecting link is as follows : the action that is declared to be best from the impartial observer's point of view can also be said to have successfully passed the universalization test". Le théorème d'Harsanyi suppose toutefois, pour atteindre dans tous les cas une solution collective, que les individus évaluent de la même manière la valeur d'un état social pour l'un d'entre eux, d'où le recours à un postulat de similarité, selon lequel les individus ont les mêmes préférences éthiques. Mais d'après Mongin et d'Aspremont ((1998), p. 449) qui rejoignent en cela Sen (1977), "Harsanyi's notion of ethics is restrictive in one sense, because he unexceptionally identifies the 'moral' and 'social' points of view, and thus appears to leave no room for private ethics. In another sense, his conception is an encompassing one, because it supposedly covers *all* ethical aspects of social relations". Si un individu ne souhaite pas – ou est incapable de – éprouver un sentiment impartial de bienveillance pour tous, s'il désire le faire seulement vis-à-vis d'un groupe particulier, son voisinage par exemple ou la classe sociale à laquelle il appartient, la structure proposée par Harsanyi ne permet pas de rendre compte de ceci : il est donc nécessaire d'introduire des étapes intermédiaires entre les préférences subjectives et les préférences éthiques d'un individu. Se développe alors l'idée selon laquelle différents degrés de préférences peuvent être envisagés.

En outre, d'après Harsanyi (1977), on ne peut accéder aux préférences éthiques qu'après la "correction" et la "censure" des préférences subjectives

des individus. L'Observateur Impartial peut y procéder dans la mesure où l'individu dont les préférences sont modifiées est susceptible d'approuver cette correction. Sur ce point, la proposition de Goodin (1986) constitue une alternative. En effet, grâce à son concept de "*self-laundering preferences*", il s'agit là non pas de corriger les préférences individuelles par le biais d'un Observateur Impartial qui tend à opérer aveuglément sans égard aucun pour les motivations individuelles, mais d'inciter les individus à modifier eux-mêmes leurs préférences, à les "nettoyer". On en revient ainsi à la tentative de résolution de Sen (1976) dans laquelle il affirme qu'il existe une différence entre un individu préférant un état social à un autre et un individu désirant que cette préférence soit prise en compte dans le processus de choix collectif.

Dès lors, deux notions se mettent en place : d'une part, il faut être à même d'envisager plusieurs degrés de réflexion sur les préférences des individus et, d'autre part, offrir la possibilité aux individus de corriger leurs préférences eux-mêmes. Ces deux idées se rejoignent grâce au concept de métaclassement des préférences, autrement dit un classement des ordres individuels de préférence effectué par l'individu. Ce concept mentionné par McPherson (1982) et Goodin (1986) vient notamment à la suite des travaux de Frankfurt (1971) et Sen (1974, 1977).

Ces auteurs décrivent différents aspects et avantages du métaclassement. Pour Frankfurt (1971), la capacité d'avoir un métaclassement des préférences est la qualité qui distingue de façon unique les êtres humains. En effet, d'après ce dernier, la personne humaine est caractérisée par le fait qu'elle est non seulement capable d'avoir des désirs de premier ordre, mais également de s'auto-évaluer, c'est-à-dire de former des désirs de second ordre : "besides wanting and choosing and being moved to do this or that, men may also want to have (or not to have) certain desires and motives. They are capable of wanting to be different, in their preferences and purposes, from what they are. Many animals appear to have the capacity for what I shall call 'first-order desires' or 'desires of the first order', which are simply desires to do or not to do one thing or another. No animal other than man, however, appears to have the capacity for reflective self-evaluation that is manifested in the formation of second-order desires" (Frankfurt (1971), p. 7). McPherson (1982) prétend pouvoir traduire en termes de métaclassement des préférences la notion développée par Mill (1863) selon laquelle les individus bénéficieraient d'une propension innée à développer des capacités propres à réaliser des plaisirs qualitatifs ("*higher pleasures*"). Goodin (1986), quant à lui, souligne que le processus consistant à épurer ses préférences premières est fondé en partie sur l'hypothèse d'existence d'un métaclassement des préférences : "laundering their preferences then simply amounts to respecting their own preferences for preferences" (Goodin (1986), p. 83). Pour Sen ((1974), p. 55), "morality would seem to require a judgement among preferences whereas rationality would not".

Par rapport au théorème d'Harsanyi, le concept de métaclassement permet de mettre au jour des étapes intermédiaires entre les préférences subjectives et les préférences éthiques. En outre, il est alors possible d'exprimer le phénomène d'*akrasia* ou faiblesse de la volonté, c'est-à-dire les préférences que l'on aurait aimé avoir, l'exemple le plus flagrant étant celui donné par Sen : "given my current tastes, I am better off with heroin, but having heroin leads me to addiction, and I would have preferred not to have these tastes" (Sen (1977), p. 101). Mais Sen (1974) ne semble tout d'abord pas s'apercevoir que l'avantage du métaclassement ne consiste pas uniquement à développer des valeurs morales et offre en réalité des possibilités très étendues : il revient donc sur sa première approche lorsque, en 1977, il indique que "a meta-ranking is a general technique usable under alternative interpretations of the meta-ranking relation. It can be used to describe a particular ideology or a set of political priorities or a system of class interests" (Sen (1977), p. 101). Il reconnaît donc finalement que la notion de métaclassement est une approche générale. Même si les métaclassements construits sur une base éthique forment une catégorie importante, ils ne restent qu'un sous-ensemble de l'ensemble des métaclassements : un individu peut en effet exhiber un métaclassement qui ne correspond pas forcément à son sens du juste.

Nous avons donc montré que les tentatives de résolution du paradoxe libéral-parétien plaident en faveur d'une prise en compte des motivations et aspirations individuelles, puis que le concept de métaclassement permet précisément d'exprimer celles-ci. Il paraît alors possible de tenter de résoudre le conflit Pareto-liberté dans un nouveau formalisme fondé sur la notion de métaclassement.

4 Construire les métaclassements : une méthode

Il nous faut maintenant traduire le concept de métaclassement en un formalisme au sein duquel une place est accordée aux motivations et aux valeurs individuelles. Les pistes de formalisation offertes par les auteurs évoqués sont fort minces, voire inexistantes. Pour donner un corps analytique à ce concept, divers points se doivent donc d'être abordés en sus : d'une part, pourquoi les individus éprouvent-ils le besoin de nettoyer par eux-mêmes leurs préférences ? Y a-t-il des circonstances particulières qui favorisent la constitution d'un métaclassement ou, du moins, le fait que les individus expriment des préférences différentes de leurs préférences premières ? D'autre part, comment les individus effectuent-ils le classement de leurs différents ordres individuels de préférence ? Existe-t-il des classements plus probables que d'autres ? Enfin, même s'il est possible d'aboutir à des métaclassements individuels convaincants, le processus de choix collectif doit être en mesure de les prendre en compte. De quelle manière ceci peut-il être envisagé ?

Notons tout d'abord que le concept de métaclassement entretient des liens étroits avec la notion de nettoyage "endogène" des préférences individuelles comme le suggère l'expression de Goodin (1986), "*self-laundering preferences*". Ainsi, les individus sont susceptibles de réviser leurs préférences premières et d'exprimer des préférences différentes. Mais pourquoi un tel comportement si les individus n'étaient pas également pourvus de valeurs qui ne s'accordent pas toujours avec leurs préférences premières ? A cette question, Festinger ((1957), cité par Bowles (1998), p. 81) offre une réponse fort éloquente : "the human organism tries to establish internal harmony, consistency or congruity among his opinions, attitudes, knowledge, and values...". Dès lors, ce serait en tentant de réduire des contradictions possibles entre ces différents éléments que les individus réviseraient leurs préférences et constitueraient de la sorte leur métaclassement.

Ces idées recourent celles avancées par Gravel, Laslier et Trannoy (2000) : ces derniers proposent en effet de tester la cohérence entre les préférences (nommées "*tastes*") et les valeurs d'un individu impliqué dans un processus de choix collectif. Précisons quelque peu cet apport : les individus ont donc des préférences, mais ils ont également des valeurs. Celles-ci sont scindées en deux sous-ensembles : les valeurs en termes de résultat ("*end-values*") qui limitent l'ensemble des relations de préférence sociale aux relations que l'individu juge acceptables par rapport à l'éthique ; et les valeurs en termes d'agrégation ("*aggregation-values*") qui spécifient comment la relation de préférence sociale dépend des préférences individuelles. Il s'agit dès lors d'évaluer la cohérence de ces trois éléments. Avant toute chose, estiment les auteurs, le test doit être indépendant des préférences d'autrui car les individus n'ont pas de pouvoir sur ces dernières. Deux types de cohérence peuvent donc être envisagés : soit les individus sont tenus d'exhiber des préférences compatibles avec l'ensemble de leurs valeurs, et ce, quelles que soient les préférences des autres individus de la société ; soit les individus sont censés avoir des préférences qui, lorsqu'elles sont identiques pour tous, sont conformes à leurs valeurs. Ce second test est qualifié de test d'universalisation ("*universalization test*"). Finalement, de tels procédés permettraient d'aboutir à des restrictions endogènes de la condition U, mais également, pour revenir à nos préoccupations, de justifier la notion de métaclassement en soi ainsi que la manière dont celui-ci peut se former.

Pouvons-nous alors emprunter ces idées pour fonder notre formalisation du métaclassement ? En ce qui nous concerne, nous sommes face à un processus de choix collectif où les valeurs sont fixées par le contexte du conflit Pareto-liberté. Ainsi, les valeurs en termes de résultat correspondent ici à l'obtention d'une relation de préférence sociale complète et acyclique, tandis que les valeurs en termes d'agrégation sont représentées par les conditions P et L+. Ces valeurs sont supposées être partagées par tous. Par la condition U, tous les profils de préférences individuelles sont possibles. Pour ce qui est du test de cohérence, nous pouvons retenir le premier test proposé

par Gravel, Laslier et Trannoy (2000), bien que ces derniers aient préféré le second. En effet, ce premier test qui préconise un ordre individuel de préférence conforme aux valeurs quelles que soient les préférences d'autrui nous semble plus adapté pour traiter le cas du paradoxe libéral-parétien car notre formalisme requiert des ordres individuels de préférence sur les états sociaux. Dès lors, si l'on considérait, dans le contexte du conflit Pareto-liberté, le test dit d'universalisation, on parviendrait systématiquement et trivialement à des préférences cohérentes pour tous les individus⁶.

Tentons à présent d'appliquer ce premier test au conflit Pareto-liberté. Il s'agit donc de déterminer si l'ordre de préférence d'un individu i est bien conforme à ses valeurs quelles que soient les préférences d'autrui. En termes analytiques, ceci pourrait être traduit de la manière suivante : soit \mathcal{R}_i , défini comme l'ensemble de tous les R_i possibles pour un individu i et \mathcal{R}_{-i} , l'ensemble de tous les R_{-i} possibles avec $R_{-i} = (R_1, R_2, \dots, R_{i-1}, R_{i+1}, \dots, R_n)$. Ainsi, \mathcal{R} , l'ensemble de tous les profils possibles, est égal à $\mathcal{R}_1 \times \mathcal{R}_2 \times \dots \times \mathcal{R}_n$. De plus, lorsque $R_i \in \mathcal{R}_i$ et $R_{-i} \in \mathcal{R}_{-i}$, on note $(R_i ; R_{-i}) = (R_1, \dots, R_{i-1}, R_i, R_{i+1}, \dots, R_n)$. On définit alors deux sous-ensembles de \mathcal{R}_i , \mathcal{R}_i^α et \mathcal{R}_i^β , et ce, en fonction de \mathcal{R}_{-i} :

$$\mathcal{R}_i^\alpha = \{R_i \in \mathcal{R}_i \mid f(R_i ; R_{-i}) = R \text{ pour tout } R_{-i} \in \mathcal{R}_{-i}\},$$

$$\mathcal{R}_i^\beta = \{R_i \in \mathcal{R}_i \mid f(R_i ; R_{-i}) \neq R \text{ pour au moins un } R_{-i} \in \mathcal{R}_{-i}\}.$$

Avant toute chose, il convient de vérifier que pour une telle société l'ensemble \mathcal{R}_i^α n'est pas vide. Ceci fait l'objet de la proposition suivante :

Proposition 1. *Pour une société qui souhaite appliquer les conditions U, P et L+ et parvenir à une relation de préférence sociale R complète et acyclique, l'ensemble \mathcal{R}_i^α tel que défini précédemment est non vide.*

Preuve. Il s'agit de prouver qu'il existe au moins un ordre individuel de préférence R_i dans \mathcal{R}_i tel qu'une relation de préférence sociale R complète et acyclique est obtenue lorsque $f(R_i ; R_{-i})$ satisfait aux conditions P et L+ pour tout R_{-i} dans \mathcal{R}_{-i} .

Il nous faut tout d'abord montrer comment un cycle se forme dans le cas du paradoxe libéral-parétien. Si R n'est pas acyclique pour un profil $\{R_i\}_{i=1}^n$, alors $\exists x_1, \dots, x_j \in X : [(x_1, x_2) \in P(R) \ \& \ (x_2, x_3) \in P(R) \ \& \ \dots \ \& \ (x_{j-1}, x_j) \in P(R) \ \& \ (x_j, x_1) \in P(R)]$. Chaque étape de ce cycle est due soit à l'application de la condition P, soit à l'application de la condition L+. Remarquons d'emblée que le cycle ne peut être provoqué par la seule application de la condition P puisque les préférences individuelles sont transitives, ni par la seule application de la condition L+ car celle-ci suppose une attribution cohérente des droits : les deux conditions sont donc forcément toutes deux pré-

⁶ Pour pouvoir appliquer le test d'universalisation au paradoxe libéral-parétien, il conviendrait de donner un autre sens aux termes "préférences identiques". On pourrait par exemple envisager des préférences identiques en termes de leur degré de "meddlesomeness", c'est-à-dire de leur degré d'empiètement sur la sphère privée d'autrui.

sentes dans le cycle. De plus, au moins deux étapes du cycle sont dues à l'application de la condition L+ pour deux individus différents. En effet, si une seule étape du cycle était due à l'application de la condition L+, alors l'individu responsable de celle-ci ferait preuve de préférences individuelles intransitives, puisque, par ailleurs, toutes les autres étapes du cycle seraient dues à l'application de la condition P. Si un seul individu était responsable de plusieurs étapes du cycle dues à la condition L+, la conclusion serait naturellement identique. Ainsi, un cycle de ce type fait intervenir au moins deux étapes dues à la condition L+ pour deux individus différents et une étape due à la condition de Pareto.

Considérons à présent l'ordre de préférence suivant pour l'individu i : $\forall x, y \in X, xI_i y$. Deux cas se présentent alors : 1) si i et un seul autre individu de la société sont pourvus d'une sphère privée non vide, alors $\forall R_{-i} \in \mathcal{R}_{-i}$ aucun cycle n'est possible dans la mesure où les préférences de i sur sa sphère privée n'engendrent aucune préférence sociale. 2) Si deux ou plusieurs autres individus de la société sont pourvus d'une sphère privée non vide, la même conclusion s'impose car les préférences de i empêchent toute préférence sociale respectant la condition P.

Remarquons que $\mathcal{R}_i^\alpha \cup \mathcal{R}_i^\beta = \mathcal{R}_i$ et $\mathcal{R}_i^\alpha \cap \mathcal{R}_i^\beta = \emptyset$. \mathcal{R}_i^α et \mathcal{R}_i^β forment donc une partition de \mathcal{R}_i . Ainsi, d'après le test de cohérence retenu et pour un individu i , si \mathcal{R}_i appartient à l'ensemble \mathcal{R}_i^α , alors on peut dire que les préférences de i sont cohérentes avec ses valeurs.

En se fondant sur ce test, il est possible de proposer un classement crédible des ordres individuels de préférence, c'est-à-dire un métaclassement. Pour ce faire, il nous faut introduire quelques notations supplémentaires. Soit M_i , le métaclassement de l'individu i . Pour tout $R_i, R'_i \in \mathcal{R}_i$, $(R_i, R'_i) \in M_i$ signifie que R_i est faiblement préféré à R'_i par l'individu i . De plus, nous supposons que M_i est complet et transitif, c'est-à-dire qu'il décrit un ordre sur \mathcal{R}_i . Comme pour R_i , $P(M_i)$ et $I(M_i)$ désignent respectivement les parties asymétrique et symétrique de M_i . En outre, $\{M_i\}_{i=1}^n$ est appelé un méta-profil.

Grâce aux ensembles \mathcal{R}_i^α et \mathcal{R}_i^β , nous considérons ici le métaclassement M_i le plus simple possible, composé de deux classes d'équivalence :

$$\forall R_i^\alpha \in \mathcal{R}_i^\alpha, \forall R_i^\beta \in \mathcal{R}_i^\beta, (R_i^\alpha, R_i^\beta) \in P(M_i),$$

et :

$$\forall R_i^\alpha, R_i \in \mathcal{R}_i^\alpha, (R_i^\alpha, R_i) \in I(M_i),$$

$$\forall R_i^\beta, R'_i \in \mathcal{R}_i^\beta, (R_i^\beta, R'_i) \in I(M_i)$$

Dans ce cas, M_i est bien complet et transitif. Notons qu'un métaclassement n'est pas nécessairement construit de la sorte, il s'agit ici d'une simple

hypothèse autorisée par le test de cohérence retenu. Mais que l'on demeure ou pas dans le cadre de ce test, il est bien sûr possible d'envisager plus de deux classes. Néanmoins, le métaclassement décrit ci-dessus nous semble convaincant : il pourrait bien correspondre à une application de la notion de nettoyage "endogène" des préférences dans le cas du paradoxe libéral-parétien.

Nous arrivons à présent à la dernière de nos interrogations : comment le processus de choix collectif peut-il prendre en compte de tels métaclassements ? Dans quelles circonstances les individus sont-ils prêts à les exprimer ? Une fois encore, nous revenons à Goodin (1986). Celui-ci affirme en effet que les individus, dans un contexte électoral, législatif ou administratif, sont enclins à n'exprimer que leurs préférences "éthiques" : "th[e] multiplicity of preference orderings matters because, in the context of collective decision-making, people will launders their own preferences. They will express only their public-oriented, ethical preferences, while suppressing their private-oriented, egoistic ones" (Goodin (1986), p. 88). Ainsi, les individus, face à un problème de décision publique, se reconnaissent en tant que "citoyens" et, de ce fait, mettent en sourdine leurs intérêts propres pour se soucier avant tout des intérêts de la communauté. Certes, mais le font-ils d'emblée, en toutes circonstances ? Ainsi que le soutient Bowles ((1998), p. 75, nous soulignons), "markets and other economic institutions do more than allocate goods and services : they also influence the *evolution* of values, tastes, and personalities". *Evolution*, ce mot fait bien référence à une démarche qui n'est pas instantanée. Il renvoie à une notion de modification progressive des préférences dès lors que les individus interagissent par le biais des institutions. Il convient donc de s'interroger sur les motivations d'une telle démarche : à quel moment un individu ferait-il appel à son métaclassement lorsque le processus de choix collectif est en marche ?

D'après les travaux récents de Poulsen et Svendsen (2005) qui corroborent un certain nombre d'études relevant de l'économie expérimentale, plus un individu dispose d'information sur les préférences d'autrui, plus ce dernier est incité à réviser ses préférences, à agir comme un "citoyen". En ce qui nous concerne, nous pourrions traduire cette idée de la manière suivante : lorsque les individus ont suffisamment d'information sur les préférences d'autrui – ils savent par exemple que tel profil de préférences individuelles ne permet pas d'aboutir à une relation de préférence sociale complète et acyclique –, ils tendent à invoquer leur métaclassement, construit grâce à un test de cohérence sur leurs propres préférences. Analytiquement, ceci pourrait être pris en compte grâce à une fonction φ agissant sur les profils de préférences individuelles dès lors que ceux-ci ne mènent pas à une relation de préférence sociale complète et acyclique. Dans ce cas seulement, cette fonction φ considérerait les métaclassements des individus. En effet, dans le cas où l'information dont disposent les individus rend compte de l'existence d'une relation de préférence sociale complète et acyclique, point n'est besoin pour eux de faire appel aux métaclassements puisque, d'emblée, il n'y a pas de contradiction entre leurs préférences et leurs valeurs. Ces considérations sont résumées par la définition suivante :

Définition 2. Soit une société qui souhaite appliquer les conditions U , P et $L+$ pour parvenir à une relation de préférence sociale R complète et acyclique et une fonction $\varphi : \mathcal{R} \rightarrow \mathcal{R}$ avec les caractéristiques suivantes :

- (i) $f \circ \varphi(\{R_i\}_{i=1}^n) = R$ pour tout $\{R_i\}_{i=1}^n \in \mathcal{R}$;
- (ii) $\varphi(\{R_i\}_{i=1}^n) = \{R_i\}_{i=1}^n$ pour tout $\{R_i\}_{i=1}^n \in \mathcal{R}$ tel que $f(\{R_i\}_{i=1}^n) = R$;
- (iii) pour tout individu i , la fonction φ est déterminée par un méta-classement M_i tel que lorsque R_i^φ , la i -ème composante de $\varphi(\{R_i\}_{i=1}^n)$, est différente de R_i , la i -ème composante de $\{R_i\}_{i=1}^n$, alors $(R_i^\varphi, R_i) \in P(M_i)$.

Une question reste alors en suspens à ce stade : une telle fonction φ existe-t-elle ? La proposition suivante énonce un résultat de possibilité :

Proposition 2. Pour une société qui souhaite appliquer les conditions U , P et $L+$, il est toujours possible de déterminer un méta-profil $\{M_i\}_{i=1}^n$ tel qu'une fonction φ existe.

Preuve. Nous considérons alors le métaclassement M_i défini précédemment à partir des ensembles \mathcal{R}_i^α et \mathcal{R}_i^β et supposons que tous les individus de la société ont un métaclassement tel. On aboutit de la sorte à un méta-profil $\{M_i\}_{i=1}^n$.

Soit une fonction $\gamma : \mathcal{R} \rightarrow \mathcal{R}$ définie à partir de $\{M_i\}_{i=1}^n$. Si $f(\{R_i\}_{i=1}^n) \neq R$ pour un $\{R_i\}_{i=1}^n$ dans \mathcal{R} , alors $\gamma(\{R_i\}_{i=1}^n) = (R_i^\alpha, R_{-i})$ pour au moins un i dans N , avec $R_i^\alpha \in \mathcal{R}_i^\alpha$. Dans tous les autres cas, $f(\{R_i\}_{i=1}^n) = R$ et $\gamma(\{R_i\}_{i=1}^n) = \{R_i\}_{i=1}^n$.

Vérifions que la fonction γ remplit toutes les conditions de la définition 2 : (i) On a $f \circ \varphi(\{R_i\}_{i=1}^n) = R$ pour tout $\{R_i\}_{i=1}^n \in \mathcal{R}$ car lorsque $f(\{R_i\}_{i=1}^n) \neq R$, alors $\gamma(\{R_i\}_{i=1}^n) = (R_i^\alpha, R_{-i})$ pour au moins un i dans N . Or, $f(R_i^\alpha, R_{-i}) = R$ pour tout R_i^α dans \mathcal{R}_i^α et tout R_{-i} dans \mathcal{R}_{-i} . (ii) Cette condition est satisfaite par construction. (iii) Si $f(\{R_i\}_{i=1}^n) \neq R$, alors $\gamma(\{R_i\}_{i=1}^n) = (R_i^\alpha, R_{-i})$ pour au moins un i dans N . Or, pour tout i dans N , si $f(R_i^\alpha, R_{-i}) \neq R$, alors $R_i \in \mathcal{R}_i^\beta$. Donc $(R_i^\alpha, R_i) \in P(M_i)$.

Ainsi, la fonction γ a toutes les caractéristiques d'une fonction φ et peut être considérée comme telle.

Finalement, définir une telle fonction φ à partir d'un méta-profil $\{M_i\}_{i=1}^n$ correspond à une restriction de la condition U . Ceci corrobore bien les remarques et résultats de Gravel, Laslier et Trannoy (2000) :

Condition 4. (U*) *Domaine non restreint en termes de méta-classement.* Le domaine de f est $\varphi(\{R_i\}_{i=1}^n)$, $\forall \{R_i\}_{i=1}^n \in \mathcal{R}$.

Ainsi, nous avons établi que la notion de métaclassement fondée sur un test de cohérence entre les préférences et les valeurs des individus per-

mettait de résorber le paradoxe libéral-parétien. Cependant, l'hypothèse selon laquelle tous les individus de la société pratiquent un tel test de cohérence et établissent un métaclassement en conséquence est très contraignante. Avant de conclure nos propos, il convient de déterminer une condition dite de cohérence minimale grâce à laquelle il existe, pour tout méta-profil $\{M_i\}_{i=1}^n$, une fonction φ qui conduit à un résultat de possibilité. Nous définissons cette condition à l'aide des ensembles \mathcal{R}_i^α et \mathcal{R}_i^β :

Condition 5. (CM) Cohérence minimale. *Au moins un individu j de N fait preuve d'un métaclassement M_j avec la caractéristique suivante : $\forall R_j^\beta \in \mathcal{R}_j^\beta, \exists R_j^\alpha \in \mathcal{R}_j^\alpha$ tel que $(R_j^\alpha, R_j^\beta) \in P(M_j)$.*

La condition CM signifie qu'au moins un individu de la société préfère strictement au moins un ordre appartenant à l'ensemble \mathcal{R}_j^α à n'importe quel ordre appartenant à \mathcal{R}_j^β dans son métaclassement. En d'autres termes, au moins un individu de la société doit pratiquer le test de cohérence que nous avons retenu et classer au moins un ordre individuel de préférence en fonction de celui-ci. Nous aboutissons alors à la proposition suivante :

Proposition 3. *Pour une société qui souhaite appliquer les conditions U, P et $L+$, il existe, pour tout méta-profil $\{M_i\}_{i=1}^n$ respectant la condition CM, une fonction φ qui mène à une SDF satisfaisant aux conditions U^*, P et $L+$.*

Preuve. Soit $J \subseteq N$, l'ensemble des individus tel que la condition CM est satisfaite avec $|J| \geq 1$. Pour tout méta-profil $\{M_i\}_{i=1}^n$ respectant la condition CM, définissons une fonction $\delta : \mathcal{R} \rightarrow \mathcal{R}$ telle que si $f(\{R_i\}_{i=1}^n) \neq R$ pour un $\{R_i\}_{i=1}^n$ dans \mathcal{R} , alors $\delta(\{R_i\}_{i=1}^n) = (R_j^\alpha, R_{-j})$ avec $j \in J$ et $\delta(\{R_i\}_{i=1}^n) = \{R_i\}_{i=1}^n$ dans tous les autres cas.

On peut vérifier aisément que la fonction δ correspond à une fonction φ : les étapes pour les conditions (i) et (ii) sont identiques à celles de la preuve de la proposition 2. Pour la condition (iii), si $f(R_j, R_{-j}) \neq R$, alors $R_j \in \mathcal{R}_j^\beta$. Dans ce cas, il est toujours possible d'avoir $\delta(\{R_i\}_{i=1}^n) = (R_j^\alpha, R_{-j})$ car, d'après la condition CM, $\forall R_j \in \mathcal{R}_j^\beta, \exists R_j^\alpha \in \mathcal{R}_j^\alpha$ tel que $(R_j^\alpha, R_j) \in P(M_j)$.

Dès lors, il existe une SDF satisfaisant aux conditions U^*, P et $L+$.

5 Conclusion

Ainsi, bien qu'à la suite de Nozick (1974) et Gaertner, Pattanaik et Suzumura (1992), un certain nombre d'auteurs s'attachent aujourd'hui à exprimer les droits individuels par le biais d'une approche en termes de jeux, les tentatives initiales de résolution du paradoxe libéral-parétien tendent à prouver qu'il est indispensable de tenir compte des motivations et valeurs individuelles. Pour nous, il s'agit du véritable enjeu du conflit élaboré par Sen (1970a,b).

Or, prendre en considération les aspirations individuelles revient en réalité à enrichir la structure de préférences individuelles. Il est possible de réaliser ceci grâce au concept de métaclassement des préférences. C'est ce que nous avons tenté de démontrer dans cette étude. Il nous semble que cette manière de concevoir les aspirations et valeurs individuelles permet de sortir de l'impasse structurelle engendrée par une résolution systématique. Le métaclassement donne lieu à un environnement théorique à l'intérieur duquel les individus ont l'opportunité de manifester leurs aspirations et valeurs sans intervention extérieure.

Nous avons proposé une méthode de construction des métaclassements dans le cas du paradoxe libéral-parétien. Celle-ci est fondée sur la cohérence des préférences et des valeurs individuelles et s'appuie en particulier sur un test de cohérence évoqué par Gravel, Laslier et Trannoy (2000). Cette méthode a finalement conduit à effectuer des restrictions endogènes de la condition U et nous a permis d'énoncer un résultat de possibilité pour le théorème de Sen.

En fonction des valeurs invoquées par les individus, on pourrait naturellement définir un certain nombre de manières de concevoir le métaclassement. En particulier, la mise au point de différentes méthodes permettant d'aboutir à des résolutions d'autres résultats d'impossibilité de la théorie du choix social constituerait une extension possible de ce travail. Ce que nous avons simplement souhaité démontrer ici est que, dans le contexte du conflit Pareto-liberté, un tel procédé est conceptuellement convaincant et formellement envisageable.

Références

- Arrow K. J., (1951), *Social choice and individual values*, 1^{ère} édition, New York, Wiley.
- Blau J. H., (1975), *Liberal values and independence*, *Review of Economic Studies*, 42, pp. 395-401.
- Bowles S., (1998), *Endogenous preferences : the cultural consequences of markets and other economic institutions*, *Journal of Economic Literature*, 36, pp. 75-111.
- Deb R., P. K. Pattanaik et L. Razzolini, (1997), *Game forms, rights, and the efficiency of social outcomes*, *Journal of Economic Theory*, 72, pp. 74-95.
- Farrell M. J., (1976), *Liberalism in the theory of social choice*, *Review of Economic Studies*, 43, pp. 3-10.

- Fleurbaey M., N. Gravel, J.-F. Laslier et A. Trannoy, (1998), Introduction and overview, in J.-F. Laslier, M. Fleurbaey, N. Gravel et A. Trannoy (éds), *Freedom in economics : new perspectives in normative analysis*, Londres, Routledge, pp. 1-13.
- Frankfurt H. G., (1971), Freedom of the will and the concept of a person, *Journal of Philosophy*, 68, pp. 5-20.
- Gaertner W. et L. Krüger, (1981), Self-supporting preferences and individual rights : the possibility of Paretian libertarianism, *Economica*, 48, pp. 17-28.
- Gaertner W., P. K. Pattanaik et K. Suzumura, (1992), Individual rights revisited, *Economica*, 59, pp. 161-177.
- Gibbard A., (1974), A Pareto-consistent libertarian claim, *Journal of Economic Theory*, 7, pp. 388-410.
- Goodin R. E., (1986), Laundering preferences, in J. Elster et A. Hylland (éds), *Foundations of social choice theory*, Cambridge, Cambridge University Press, pp. 75-101.
- Gravel N., J.-F. Laslier et A. Trannoy, (2000), Consistency between tastes and values : a universalization approach, *Social Choice and Welfare*, 17, pp. 293-320.
- Harsanyi J. C., (1955), Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility, *Journal of Political Economy*, 63, pp. 309-321.
- Harsanyi J. C., (1977), *Rational behavior and bargaining equilibrium in games and social situations*, Cambridge, Cambridge University Press.
- Hayek F. A., (1960), *The constitution of liberty*, Londres, Routledge.
- McPherson M. S., (1982), Mill's moral theory and the problem of preference change, *Ethics*, 92, pp. 252-273.
- Mill J. S., (1859), *On liberty*, New York, Appleton-Century-Crofts, 1947.
- Mill J. S., (1863), *Utilitarianism*, Indianapolis, Bobbs-Merrill, 1957.
- Mongin P. et C. d'Aspremont, (1998), Utility theory and ethics, in S. Barbera, P. J. Hammond et C. Seidl (éds), *Handbook of utility theory*, vol. 1 : Principles, Dordrecht, Kluwer Academic Publishers, pp. 371-481.
- Nozick R., (1974), *Anarchy, state and utopia*, Oxford, Blackwell.
- Poulsen A. U. et G. T. Svendsen, (2005), Social capital and endogenous preferences, *Public Choice*, 123, pp. 171-196.
- Saari D. G. (1998), Connecting and resolving Sen's and Arrow's theorems, *Social Choice and Welfare*, 15, pp. 239-261.
- Sen A. K., (1970a), *Collective choice and social welfare*, San Francisco, Holden-Day.
- Sen A. K., (1970b), The impossibility of a Paretian liberal, *Journal of Political Economy*, 78, pp. 152-157.

- Sen A. K., (1974), Choice, orderings and morality, in S. Körner (éd.), *Practical reason*, Oxford, Blackwell, pp. 54-67.
- Sen A. K., (1976), Liberty, unanimity and rights, *Economica*, 43, pp. 217-245.
- Sen A. K., (1977), Rational fools : critique of the behavioural foundations of economic theory, *Philosophy and Public Affairs*, 6, pp. 317-344.
- Sen A. K., (1979), Personal utilities and public judgements : or what's wrong with welfare economics, *Economic Journal*, 89, pp. 537-558.
- Suzumura K., (1978), On the consistency of libertarian claims, *Review of Economic Studies*, 45, pp. 329-342.
- Van Hees M., (1999), Liberalism, efficiency, and stability : some possibility results, *Journal of Economic Theory*, 88, pp. 294-309.