Chapitre IV – ANALYSE AXIOMATIQUE DU CHOIX COLLECTIF

Après avoir étudié en détail deux méthodes de choix collectifs, la majorité simple et le décompte de Borda, on s'aperçoit que chaque méthode a ses propres défauts, et qu'aucune ne parvient à vérifier un ensemble de propriétés désirables :

- la règle à la majorité simple ne remplit les conditions requises que s'il y a deux candidats à départager ; dans le cas de trois candidats ou plus, toute méthode de vote généralisant la règle majoritaire pose des difficultés : soit elle ne sera pas toujours décisive (absence de vainqueur de Condorcet), soit elle ne sera pas efficace (elle ne choisit pas le vainqueur de Condorcet lorsqu'il existe ou elle est manipulable) ;
- les méthodes à scores présentent le même type de défaut.

On en vient alors à questionner la possibilité logique de la construction d'une règle compatible avec ces règles.

1) Qualités et défauts de base

Deux qualités des méthodes de choix collectif sont mises en exergue et ont déjà été présentées. L'une, le caractère de procédure ordonnante à propos de la méthode de Borda ; l'autre, l'indépendance par rapport aux informations extérieures, à propos du critère de Condorcet.

« Rappelons ce qui est en jeu.

Chaque électeur est supposé capable d'ordonner tous les candidats (éventuellement avec des ex-aequo). De ces ordres individuels, une procédure ordonnante déduit un classement collectif. C'était le cas de la procédure de Borda, ce n'était pas le cas de la méthode des duels (Condorcet) puisque qu'une majorité d'électeurs peut se dégager pour :

- préférer un candidat *a* à un candidat *b* ;
- préférer un candidat *b* à un candidat *c* ;
- préférer le candidat *c* au candidat *a*.

D'autre part, dans la méthode des duels, l'ordre de deux candidats au classement général ne dépendant que de la façon dont ces deux candidats étaient placés dans l'esprit de chaque électeur. La présence éventuelle d'autres candidats ne modifie en rien ce classement. Nous avons vu que cette indépendance par rapport aux informations extérieures n'appartient pas à la procédure de Borda. Observons déjà que se trouvent exclues des informations prises en compte toute notion d'intensité des préférences individuelles et toute comparaison interpersonnelle entre ces intensités. » (Boursin)

« A côté des qualités de base sur lesquelles se concentre l'attention, le défaut majeur des méthodes de majorité consiste en l'existence de cycles. Une majorité peut se dégager pour préférer *a* à *b*, *b* à *c*, et *c* à *a*. Un cycle peut d'ailleurs comporter plus de trois éléments et il n'est pas plus satisfaisant qu'une méthode oblige à préférer :

$$a_1 \text{ à } a_2,\ a_2 \text{ à } a_3,\ a_3 \text{ à } a_4,\ \ldots \text{ et } a_{n-1} \text{ à } a_n, \text{ et pourtant } a_n \text{ à } a_1.$$

Cette notion de cycle est liée à celle de vainqueur de Condorcet. Dans les situations que nous étudions, un vainqueur de Condorcet est un candidat qui, opposé séparément à chacun des autres, est choisi par la majorité des électeurs. On peut alors démontrer que :

S'il n'existe pas de vainqueur de Condorcet, alors il existe au moins trois candidats formant un cycle.

La démonstration est un bon exemple de cette sorte de raisonnement.

S'il existe un candidat a_i qui est battu (au vote majoritaire) par tous les autres, éliminons-le.

S'il existe un candidat a_j qui est battu (au vote majoritaire) par tous les autres sauf a_i, éliminons-le à son tour, etc.

Au bout de cet exercice, il reste au moins trois candidats.

En effet, s'il n'en restait que deux, celui des deux que la majorité préfère serait un vainqueur de Condorcet.

Notons a_x, a_y, a_z, a_t, … les candidats restants. Chacun d'eux bat au moins un des autres (sinon, il aurait lui aussi été éliminé), et est battu par au moins un des autres (sinon, il serait vainqueur de Condorcet).

Prenons l'un d'eux, par exemple a_x, et écrivons à sa suite l'un des candidats qu'il bat, par exemple a_y, et ainsi de suite. Comme tout candidat bat au moins un des autres, il n'y a pas de raison que cette chaîne s'arrête, ou plus exactement, un moment viendra où l'on aura à écrire un candidat figurant déjà dans la chaîne. Nous avons ainsi démontré l'existence d'un cycle. » (Boursin)

2) Théorème de Hansson (1959)

Hansson cherche une méthode de choix collectif vérifiant les quatre propriétés suivantes :

- elle est ordonnante ;
- elle est anonyme ;
- elle est neutre ;
- elle est indépendante par rapport aux informations extérieures (IIE).

Il démontre que la seule règle de choix collectif compatible est l'indifférence généralisée.

Théorème (Hansson, 1959) : Si une règle de choix collectif vérifie les quatre propriétés ci-dessus, elle est telle que, pour toutes options x et y, x et y sont équivalentes.

Démonstration :

On montre d'abord qu'en cas de partage égal des voix entre deux candidats, une telle règle les classe ex-aequo.

Notons :

- $x\ A_i\ y$, lorsqu'un électeur i classe x strictement avant y,
- $x\ D_i\ y$, lorsqu'un électeur i classe x strictement derrière y,
- $x\ E_i\ y$, lorsqu'un électeur i classe x équivalent à y.

On utilise les mêmes notations, sans l'indice i, pour le classement social résultant. La règle de dépouillement doit donner une relation $x\ R\ y$. Nous ignorons pour l'instant s'il s'agit de A, D ou E.

Imaginons que le corps électoral est composé de 2 n membres. Les électeurs numérotés 1 à n ont pour classements $x\ A_i\ y$; les électeurs numérotés $n + 1$ à 2 n ont pour classements $x\ D_i\ y$.

Si chaque d'électeur change d'avis, à cause de la neutralité de la règle, le comité doit inverser son classement. Mais comme il y avait autant d'électeurs d'avis A que d'avis D, la composition de l'urne est inchangée et d'après la règle d'anonymat, le résultat du dépouillement doit être inchangé.

Ceci exclut que R soit A, car, en inversant le classement, A devient D. Cela exclut pareillement que R soit D. En définitive, $R = E$. Le classement général donne donc une équivalence entre x et y.

On prouve maintenant qu'en cas de partage presque égal des voix entre les deux candidats ($n + 1$ électeurs dans un ordre, n dans l'autre), une telle règle les classe ex-aequo.

Imaginons les opinions suivantes dans un comité de 2 n + 1 membres :

votant n° 1 :	z	x	y
votant n° 2 :	y	z	x
votant n° 3 à $n + 2$:	x	y	z
votant n° $n + 3$ à 2 $n + 1$:	z	y	x

La règle de dépouillement doit donner une relation $x\ R\ y$, nous ignorons si R est A, D ou E.

Limitons nos observations aux candidats x et y. Parmi les électeurs, $n + 1$ préfèrent x à y et les autres préfèrent y à x. La relation $x\ R\ y$ résulte de ce partage, et seulement de ce partage, à cause des propriétés d'anonymat et d'IIE.

Limitons notre observation aux candidats z et x. Parmi les électeurs, $n + 1$ préfèrent z à x et les autres préfèrent x à z. Donc $z\ R\ x$ (de même que $x\ R\ y$, par la propriété de neutralité).

Limitons enfin notre observation aux candidats y et z. Parmi les électeurs, $n + 1$ préfèrent y à z et les autres préfèrent z à y. Donc $y\ R\ z$ (de même que $x\ R\ y$, par la propriété de neutralité).

Nous avons donc :

- d'une part $x\ R\ y$;
- d'autre part, $y\ R\ z$ et $z\ R\ x$, ce qui entraîne, la relation résultante étant transitive (propriété de règle ordonnante), que $y\ R\ x$.

Comme dans le cas précédent, cela exclut que la relation soit A, car en inversant le classement, A devient D (propriétés d'anonymat et de neutralité). Cela exclut pareillement que R soit D. En définitive $R = E$.

Comme le résultat ne dépend pas de la présence de z (IIE), nous en concluons que, chaque fois que $n + 1$ électeurs préfèrent x à y et n préfèrent y à x, le dépouillement aboutit à classer ces candidats ex-aequo.

Nous pouvons maintenant montrer le résultat :

Chaque fois que l'unanimité des électeurs classe x avant y, le classement général place des deux candidats ex-aequo.

- Si le collège électoral est pair : 2 n, soit les opinions :

votant n°1 à n :	x	y	z
votant n° $n + 1$ à 2 n :	z	x	y

D'après l'étude préliminaire, le partage égal entre x et z entraîne $x\ E\ z$; de même, $y\ E\ z$. Par transitivité, $x\ E\ y$. Du fait de la propriété d'IIE, cela ne dépend pas de la présence du candidat z : l'unanimité préfère x à y et le vote aboutit à l'équivalence de x et y.

- Si le collège électoral est impair : 2 $n + 1$, soit les opinions :

votant n°1 à $n + 1$:	x	y	z
votant n° $n + 2$ à 2 n :	z	x	y

On sait que si $n + 1$ votants sont d'un avis et n d'avis contraire, le dépouillement aboutit à l'équivalence. Donc :

$$x\ E\ z \text{ et } y\ E\ z,$$

et, par transitivité,

$$x\ E\ y.$$

Du fait de la propriété d'IIE, cela ne dépend pas de z : l'unanimité préfère x à y et le classement collectif les considère équivalents.
Le paradoxe se prolonge. Pensons à un collège avec des opinions arbitraires sur x et y. Demandons lui de voter au cas où un troisième candidat z totalement nul serait ajouté en queue de liste des préférences de chaque électeur. L'unanimité préfère x à z et y à z. Donc, le dépouillement donne

$$x\ E\ z \text{ et } y\ E\ z,$$

et, par transitivité,

$$x\ E\ y.$$

On a donc une indifférence généralisée.

3) Théorème d'impossibilité de Arrow (1951)

i) *Enoncé*

Arrow pose les conditions suivantes sur la règle de choix social :

- elle est ordonnante (elle associe à tout profil de préférences individuelles un (pré-) ordre de préférence social transitif) ;
- elle vérifie la propriété de Bentham (elle est définie pour tout profil d'opinions) (parfois appelée condition de domaine universelle) ;

- elle respecte le critère de Pareto : si tous les individus classent une option x devant y, alors le groupe doit classer x devant y ;
- elle vérifie l'indépendance par rapport aux informations extérieures ;
- elle n'est pas dictatoriale : il n'existe pas d'individu tel que la méthode de choix collectif classe toute paire d'issues x devant y quelconque dans le même sens que lui, quelles que soient les opinions des autres.

Arrow prouve le théorème :

Théorème d'impossibilité : Arrow (1951)

Si le choix porte sur plus de deux candidats, il n'existe pas de fonction de choix social qui vérifie les cinq axiomes. En fait, on montre que toute fonction de choix social remplissant les quatre premiers axiomes est dictatoriale.

ii) *Démonstration*

Elle repose sur la notion de partie décisive dans l'ensemble des votants. Une partie Q de l'ensemble des votants est dite décisive pour x contre y s'il suffit que tous les électeurs de Q classe x devant y (on notera cela $x > y$) pour qu'il en soit de même au résultat du vote, quelles que soient les préférences des autres. Une partie de l'ensemble des électeurs est dite simplement décisive si elle est décisive pour n'importe quel x contre n'importe quel y.

Une première étape de la démonstration est : avec les propriétés exigées, une partie Q décisive pour x contre y est tout simplement décisive.

Démonstration : Soit donc une partie Q décisive pour x contre y. Montrons qu'elle est décisive pour n'importe quel candidat u.
Imaginons d'abord un candidat u tel que

pour les électeurs de Q : $x > y > u$
pour les autres électeurs : $y > u > x$

Alors, au classement résultant :
- x bat y, puisque Q est décisive pour cette paire ;
- y bat u, puisque l'unanimité classe y devant u.

On en conclut, puisque la procédure est supposée ordonnante, que x bat u, selon le vœu des électeurs de Q et contre le vœu des autres électeurs. De plus, à cause de la propriété d'i.i.e., ce résultat est acquis en tenant compte seulement des classements des électeurs entre x et u, et pas du tout en fonction de la présence, de l'absence ou du classement de y.
Alors, effaçons y de nos hypothèses. Lorsque

pour les électeurs de Q : $x > u$
pour les autres électeurs : $u > x$

x bat u ; autrement dit, la partie Q est décisive pour x contre tout autre candidat u.

Imaginons maintenant un autre candidat z tel que :

pour les électeurs de Q : $z > x > u$

pour les autres électeurs : $u > z > x$

Alors, au classement résultant :

- x bat u, puisque Q est décisive pour cette paire ;
- z bat x, puisque l'unanimité classe z devant x.

On en conclut, puisque la procédure doit être ordonnante, que z bat u, selon le vœu de Q et contre le vœu des autres. De plus, à cause de la propriété d'i.i.e., ce résultat est acquis en tenant compte seulement des classements des électeurs entre z et u, indépendamment de la présence, de l'absence ou du classement de x.

Effaçons donc x de nos hypothèses. Lorsque

pour les électeurs de Q : $z > u$

pour les autres électeurs : $u > z$

z bat u ; autrement dit, Q est décisive pour tout candidat z contre tout autre candidat u.

En un mot, Q est simplement décisive.

La seconde étape de la démonstration consiste à partir d'une partie décisive (il en existe au moins une, c'est l'ensemble des votants, du fait du critère de Pareto) et à déduire séquentiellement l'existence d'un dictateur.

Soit Q une partie décisive. Divisons-la en deux ensembles disjoints A et B. Imaginons des candidats tels que :

pour les électeurs de A : $x > y > u$

pour les électeurs de B : $y > u > x$

pour les autres électeurs : $u > x > y$

Dans ces conditions, y bat u, puisque c'est le vœu des électeurs de Q, qu'ils appartiennent à A ou à B. Alors, de deux choses l'une :

- ou bien y bat x, selon le vœu des électeurs de B et contre celui de tous les autres. B est alors décisif ;
- ou bien x bat y, selon le vœu des électeurs de A et contre celui de tous les autres. A est alors décisif.

Ainsi, dans toute partition d'une partie décisive en deux sous-ensembles propres, l'un est décisif. Il ne reste, puisque l'ensemble des électeurs est fini, qu'à recommencer cette dichotomie jusqu'à trouver une partie décisive contenant un seul électeur, c'est-à-dire un dictateur.

Conclusion : les propriétés exigées de notre règle de décision ne sont remplies que pour une seule de ces règles, la dictature.

iii) *Affaiblissement des axiomes*

Le résultat obtenu par Arrow est frappant, d’autant plus que les postulats posés semblent raisonnables, d’une part, et que d’autres conditions pourraient légitimement être demandées à une règle de choix collectif, d’autre part (ce qui ne ferait que « renforcer » le résultat).

- <u>Abandon de l’hypothèse de transitivité</u> : Arrow introduit cette condition pour deux raisons (procédure ordonnante) : un premier motif est que la règle choisie désigne une (meilleure) issue pour tout état de l’opinion (c’est-à-dire évite des configurations du type du paradoxe de Condorcet) ; un second motif est que la règle de choix collectif doit être telle que le choix final ne dépende pas du cheminement qui y conduit (appelé l’agenda électoral).

Ces deux justifications sont en fait très différentes.

Concernant la première justification, les conditions posées par Arrow sont plus strictes que nécessaires. Il impose que la relation de préférence faible définisse un (pré-) ordre complet et transitif. Donc, la règle de choix collectif organise l’ensemble des candidats suivant un schéma du type :

Sens de la préférence stricte

		a			g	
Classes d’ équivalence	…	b	d	e	h	… →
		c		f	i	

Certains auteurs ont remarqué que deux conditions de rechange pouvaient garantir l’existence d’un (meilleur) choix pour tout état de l’opinion, dans la mesure où il n’est pas nécessaire de définir un (pré-) ordre, mais seulement une issue :

- la quasi-transitivité, définie comme la transitivité des préférences strictes uniquement :

$$R \text{ est transitif} \Leftrightarrow P \text{ est transitif et } I \text{ est transitif}$$

Avec la quasi-transitivité, on pose que seul P doit être transitif (des études expérimentales montrent que des individus peuvent être indifférents entre x et y, entre y et z, et malgré tout préféré x à z).

- l’acyclicité : $x_1 \, I \, x_n$ et $x_1 \, P \, x_2 \, P \, x_3 \dots x_{n-1} \, P \, x_n$ est possible.

Ces conditions suffisent pour désigner un ensemble des meilleurs choix collectifs pour tout état de l’opinion. Il a été démontré que ces conditions sont compatibles avec l’existence de fonctions de choix collectives non dictatoriales (Sen, 1969 ; Gibbard, 1969). On montre cependant que le pouvoir du dictateur ne disparaît pas complètement ; il se trouve dilué dans un groupe plus grand (Gibbard), c’est-à-dire une oligarchie.

Concernant l'autre justification de la transitivité de R, il a été montré que la transitivité de P seulement était une condition nécessaire et suffisante de l'indépendance par rapport à l'ordre des choix (i.e. l'agenda électoral). Donc, la condition d'invariabilité du choix peut être obtenue à une condition plus faible que la transitivité de R (qui impose qu'à la fois P et I soient transitifs).

- <u>Abandon de la propriété de Bentham</u> (appelée aussi domaine universel) : La justification de cette condition est : la liberté d'opinion et d'expression ; la nécessité du choix. On sait qu'il existe des théorèmes d'existence si cette condition n'est pas vérifiée : Cf. le théorème de l'électeur médian (Black, 1948).

- <u>Abandon de l'i.i.e.</u> : Arrow avance deux justifications de cet axiome : l'économie de la collecte d'information ; la non-manipulabilité du choix collectif.

Cette hypothèse revient à interdire que la règle de choix collectif soit construite sur une information cardinale (Sen). Ceci était précisément dans les intentions de Arrow, qui affirme que la mesure de l'utilité est difficile et arbitraire, manipulable aussi bien par le sondeur que par le sondé.

4) Théorème d'impossibilité de Sen (1970)

Sen (1970) pose une méthode de choix social telle que :

- elle vérifie la propriété de Bentham ;
- elle associe à tout profil de préférences individuelles une préférence sociale acyclique ;
- elle est libérale : tout individu est décisif sur une paire d'options au moins ; autrement dit, pour tout individu, il existe une paire x et y que la règle de choix collectif classe dans le même sens que lui, quelles que soient les opinions des autres ;
- elle respecte le critère de Pareto : si tous les individus classent une option x devant y, alors le groupe doit classer x devant y.

Sen (1970) donne alors la démonstration du théorème suivant.

Théorème : Sen (1970)

Si le choix porte sur au moins deux issues, il n'existe pas de méthode d'agrégation des préférences individuelles en un ordre de préférence social acyclique, qui soit à la fois libérale et Parétienne.

Intuitivement, ce théorème montre une contradiction entre le fait de doter chaque individu d'une sphère privée, à l'intérieur de laquelle il est décisif (condition de liberté), et celui de se plier à la règle d'unanimité (critère de Pareto).

Démonstration : Nous proposons ici plutôt une illustration.
Considérons le cas où il y a deux individus, notés a et b, et trois options, notées x, y et z.
Pour fixer les choses, admettons que a est décisif sur $\{x, y\}$ et qu'il préfère x à y. Ses préférences sont donc :

$$x > y > z \text{ ou } x > z > y \text{ ou } z > x > y.$$

Alors, peu importe les préférences de b, du fait que la règle de choix social est supposée libérale, le classement social doit mettre x devant y :

$$x \, P \, y.$$

Il s'agit maintenant de montrer que, quelle que soit la paire d'options sur laquelle b est censé être décisif, il existe des préférences de a (appartenant aux trois cas prévus ci-dessus) et de b induisant un cycle dans le classement social.
Si b est aussi décisif sur $\{x, y\}$, il suffit qu'il préfère y à x pour avoir un classement social cyclique, puisque si la règle de choix est supposée libérale, on a :

$$y \, P \, x.$$

Si b est décisif sur $\{x, z\}$, supposons que :

$$b : y > z > x.$$

Alors, peu importe les préférences de a, le classement social doit mettre z devant x :

$$z \, P \, x,$$

du fait que la règle de choix social est libérale.
Supposons alors que :

$$a : x > y > z.$$

Du fait que a et b classent tous les deux y devant z, le critère de Pareto fait qu'il doit en être de même dans le classement social :

$$y \, P \, z.$$

En rassemblant les résultats précédents, il vient finalement le cycle :

$$x \, P \, y \, P \, z \, P \, x.$$

Si b est décisif sur $\{y, z\}$, supposons à nouveau que :

$$b : y > z > x.$$

Alors, peu importe les préférences de a, le classement social doit mettre y devant z :

$$y \, P \, z,$$

du fait que la règle de choix social est libérale.
Supposons alors que :

$$a : z > x > y.$$

Du fait que a et b classent tous les deux z devant x, il en va de même dans le classement social :

$$z \, P \, x,$$

en vertu du principe Parétien.
En rassemblant les résultats précédents, il vient finalement le cycle :

$$x \, P \, y \, P \, z \, P \, x.$$

5) Théorème de Gibbard (1973) et Satterthwaite (1975)

Définition : Une méthode de choix social est dite non manipulable en stratégie dominante si elle incite les votants à déclarer honnêtement leur opinion, quel que soit le profil d'opinion des autres.

Gibbard (1973) et Satterthwaite (1975) s'intéressent aux méthodes de choix social non manipulables, sous les conditions supplémentaires suivantes :

- propriété de Bentham (domaine universelle) : la règle de choix social est définie pour n'importe quel profil d'opinions de l'ensemble des électeurs ;
- en parcourant l'ensemble des profils d'opinions possibles, la fonction de choix social désigne au moins trois vainqueurs.

Gibbard (1973) et Satterthwaite (1975) établissent le théorème suivant.

Théorème : Gibbard (1973) et Satterthwaite (1975)

Si une méthode de choix collectif vérifiant ces deux propriétés est non manipulable, alors elle est dictatoriale.