

Developpement méthodologique et analyse du vote

Autor(en): **Joye, Dominique**

Objekttyp: **Article**

Zeitschrift: **SVPW-Jahrbuch = Annuaire ASSP**

Band (Jahr): **27 (1987)**

PDF erstellt am: **16.08.2024**

Persistenter Link: <https://doi.org/10.5169/seals-172574>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Dominique Joye

Developpement méthodologique et analyse du vote

Cet article présente l'intérêt de développements méthodologiques récents pour l'analyse du vote, tout particulièrement en Suisse. Cette réflexion amène à montrer qu'il n'existe pas de bonne méthodologie en soi mais qu'il faut plutôt considérer un triplet «hypothèses-données-méthodes» dont aucun des termes n'est pertinent sans tenir compte des autres.

L'analyse des données agrégatives et l'analyse des données individuelles ayant des propriétés particulières, elles justifient, souvent, d'une méthodologie spécifique. Des exemples d'utilisation dans le contexte suisse peuvent alors être proposés.

In diesem Artikel werden die jüngsten methodologischen Entwicklungen zur Analyse des Wahlverhaltens am Beispiel der Schweiz dargestellt. Es zeigt sich dabei, dass es *die* gute Methode an sich nicht gibt, sondern keine der drei Schritte «Hypothese-Daten-Methode» ohne Berücksichtigung der anderen aussagekräftig sein kann.

Sowohl die Analyse der aggregierten Daten wie auch diejenige der Individualdaten haben Eigenschaften, welche die Anwendung spezifischer Methoden verlangen. Einige Beispiele aus dem schweizerischen Kontext werden im Artikel beschrieben.

1. Introduction ¹

«Le vote est un phénomène complexe. Il est à la fois individuel et collectif; il peut être abordé comme processus ou comme résultat, dans sa distribution spatiale ou sa répartition sociale; sous l'angle de ses motivations ou de ses effets. Il existe plusieurs façons de l'étudier ou de l'interpréter; chacune révèle certaines dimensions du phénomène en cachant ou en minimisant les autres.» ²

L'enjeu de la méthode n'est donc pas mince puisqu'elle est liée à une approche, à des hypothèses, à des données même. En d'autres termes, suivant les hypothèses de recherche et la méthodologie retenue, ce sont certains aspects du vote qui vont être analysés. Il m'a dès lors semblé intéressant, dans cet article, d'axer davantage la discussion sur la logique de la démarche, sur la stratégie de recherche plus que sur le détail des procédures statistiques utilisées. Une telle approche implique certes de présenter certaines recherches mais ne prétend en aucun cas couvrir l'ensemble des recherches qui se sont faites en Suisse. De même, je me bornerai ici à parler des méthodes quantitatives, sans aborder d'autres formes de productions de connaissance tout aussi pertinentes: certains articles présentés dans cet annuaire sont bien là pour en témoigner.

Généralement deux approches sont distinguées: l'analyse des données agrégées, comme les résultats par canton ou par commune; et l'analyse des données individuelles telle l'exploitation des données de sondage.

Ce sont les propriétés de ces deux approches que je vais explorer dans cet article. Mais, pour les mieux situer, il importe de revenir au préalable sur quelques caractéristiques du vote.

2. Les dimensions du vote

Le vote est souvent considéré comme l'acte de base du citoyen dans un système démocratique. Pour en parler, il est parfois utile de repartir d'une définition élémentaire pour montrer, sur une telle base, le consensus d'auteurs venus d'horizons parfois très différents.

Le petit Robert définit ainsi l'action de voter: «*Voter* . . . contribuer à faire adopter par son vote, décider par un vote majoritaire.»

Dans une telle définition, la référence au thème de la décision est manifeste. Ce concept a été largement étudié en science politique. De nombreux auteurs se sont penchés sur un modèle prévoyant un acteur parfaitement informé, prenant des décisions rationnelles.

1 Je tiens à remercier ici Dominique Auderset, Frédéric Bon, Claude Longchamp, Ioannis Papadopoulos et Martin Schuler pour leurs amicales critiques d'une version préliminaire de ce texte.

2 F. Bon, «Qu'est-ce qu'un vote», in *Histoire* No 2 1979, p. 106.

Cette hypothèse de rationalité et d'information peut sembler difficile à soutenir quand on pense à la diversité des sujets soumis à votation populaire: arrêtés fédéraux urgents destinés à combattre la surchauffe, modification du tarif des douanes, arrêté sur l'économie laitière et j'en passe. Effectivement les analyses VOX semblent confirmer ce diagnostic: des projets comme l'article conjoncturel ou le contre-projet à l'initiative sur la protection des locataires n'ont été vraiment compris que par une minorité des votants.³

Plusieurs auteurs sont conscients de ce problème de limites dans la compréhension des électeurs et renvoient alors l'explication à des facteurs sociaux. Simon écrit par exemple: «Le fait que ces limites ne sont pas physiologiques et fixés mais au contraire largement déterminées par des forces sociales voire organisationnelles crée des problèmes de construction théorique d'une grande subtilité.»⁴ De même Olson mentionne que «il est bien certain que les motivations idéologiques existent dans toutes les sociétés, jusque dans les groupes les plus stables et les mieux définis. Aux Etats-Unis on trouve ce type de comportement dès qu'on approche des partis politiques».⁵

En bref, ces considérations me semblent avoir démontré l'impossibilité d'omettre l'analyse des valeurs⁶ et de leurs liens avec l'insertion sociale quand on veut étudier le vote. Une part des déterminants est d'ordre sociologique, et la méthodologie se doit d'en tenir compte. De manière intéressante, ce débat entre niveau individuel et social est à la base d'une grande discussion méthodologique sur l'usage des données agrégées mais n'est pas sans intérêt non plus pour l'analyse de données individuelles. Comme ces deux types d'information sont très souvent utilisés en comportement électoral, il est intéressant d'y consacrer un examen attentif.

3. L'usage des données agrégées

L'usage des données agrégées a été marqué en sciences sociales par une fameuse controverse qui n'est d'ailleurs pas prête de finir. Fondamentalement, le point de départ de cette discussion se trouve dans un article célèbre de Robinson⁷ qui montrait que la corrélation entre race et analphabétisme variait suivant le type de données: 0,23 en considérant des données individuelles, 0,773 en prenant pour base l'état américain. Comment expliquer ce résultat et quelle est la vraie valeur de la corrélation?

3 E. Gruner et H. P. Hertig, *Der Stimmbürger und die «neue Politik»*, Haupt, Berne, 1983, p. 411.

4 H. A. Simon, *Models of Man*, Wiley, New York, 1957, p. 199, traduction personnelle.

5 M. Olson, *Logique de l'action collective*, PUF, Paris, 1978, p. 186.

6 Notons d'ailleurs que l'analyse des valeurs a été considérablement développée ces dernières années.

7 cf W. S. Robinson, «Ecological correlation and the behavior of individuals», *American Journal of Sociology*, vol. XV, 1950.

Le problème de base est que l'on ne dispose pas de la même information dans les deux cas. Prenons l'exemple d'un chercheur s'efforçant d'étudier la relation entre classe ouvrière et vote socialiste. Au niveau agrégé il obtient par exemple:

Tableau 1: Résultat agrégé

Circonscription	% ouvriers	% socialistes
1	20	20
2	50	50
3	80	80

Soit une corrélation parfaite; mais peut-on en déduire que ce sont les ouvriers qui votent socialistes? Certainement pas car une telle situation peut correspondre aux tableaux suivants:

Tableau 2: Résultats individuels

Circonscription 1 (ou 3)

Situation 1	Situation 1			Situation 2		
	Socialiste	Non socialiste	Total	Socialiste	Non socialiste	Total
Ouvrier	20	0	20	0	20	20
Non ouvrier	0	80	80	20	60	80
Total	20	80		20	80	

Circonscription 2

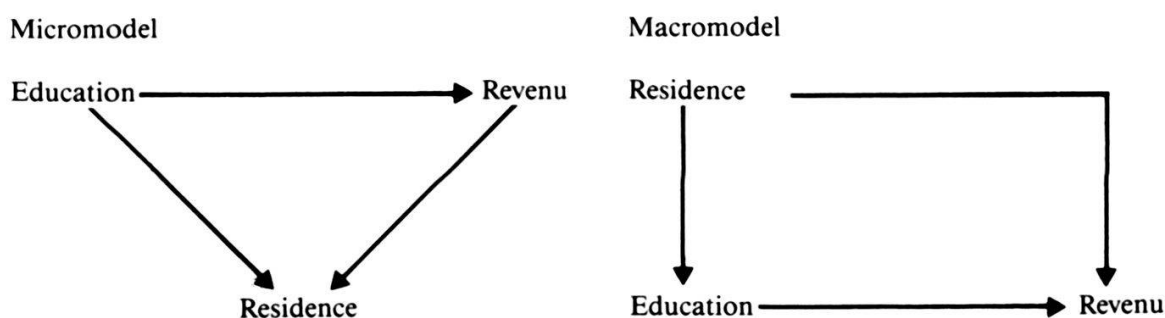
	Situation 1			Situation 2		
	Socialiste	Non socialiste	Total	Socialiste	Non socialiste	Total
Ouvrier	50	0	50	0	50	50
Non ouvrier	0	50	50	50	0	50
Total	50	50		50	50	

Le problème de l'inférence écologique s'explique ainsi: avec des données agrégées on ne connaît jamais que les marginales des tableaux que l'on peut construire pour les unités d'analyse. Les chercheurs se sont dorénavant efforcés de trouver une méthodologie permettant d'estimer les valeurs individuelles.

Plusieurs solutions dans ce sens ont été tentées: l'idée générale a été soit de poser certaines hypothèses supplémentaires,⁸ soit d'introduire d'autres varia-

⁸ Par exemple *Flanigan* et *Zingale* montrent l'utilisation d'hypothèses pour réduire les valeurs limites de variation de la corrélation individuelle: *W. H. Flanigan* et *N. H. Zingale*, *Alchemist's gold: inferring individual relationship from aggregate data*, ECPR Joint Sessions, Lancaster, 1981. Pour poursuivre une illustration fictive, dans le cas du tableau 2 on peut supposer que la moitié au moins des ouvriers vote socialiste. Une autre méthode très répandue est celle de la régression écologique de *Goodmann* qui fait l'hypothèse que le contexte n'intervient pas dans la relation. Autrement dit, dans cet exemple, que la proportion d'ouvrier qui vote socia-

bles dans l'analyse, qui prennent en compte implicitement l'effet d'agrégation,⁹ soit encore de modifier le schéma d'analyse sous-jacent. Langbein et Lichtman posent l'hypothèse que le problème est en fait une question de spécification du modèle d'analyse. Dans l'exemple de la relation entre éducation et revenu, le lieu de résidence serait une conséquence du revenu et de l'éducation au niveau individuel; dans un modèle agrégé une cause (graphique 1). De telles solutions ne me semblent pas toujours les plus pertinentes car il s'agit de constructions spécifiques à chaque niveau. Il est pourtant intéressant de voir qu'elles renvoient de manière implicite ou explicite à l'unité d'analyse et à ses caractéristiques. C'est cette dernière perspective qu'il convient de développer rapidement.



Graphique 1¹⁰: Exemple de modèle différent suivant le niveau

Alker (1973) montre que le coefficient de corrélation individuel est en fait la somme de deux composantes: la variation interrégionale et intrarégionale.¹¹ Une telle proposition amène à reprendre plus globalement le type d'interprétation que l'on peut faire de ces données.

On voit ainsi se dessiner une tendance qui, en privilégiant l'unité d'analyse, permet d'éviter le problème de l'inférence écologique. Boudon distingue trois types d'interprétation pour les données agrégées:

1. En termes de phénomènes purement individuels. Le groupement ne joue aucun rôle comme variable sociologique. Le seul but de l'analyse est de

liste est la même dans toutes les circonscriptions: L. A. Goodman, «Ecological regression and behaviour of individuals», *American Sociological Review*, vol. XVIII, 1953. Une extension de ce modèle peut se trouver par exemple dans G. R. Iversen «Recovering individual data in the presence of group and individual effects», *American Journal of Sociology*, 1973.

9 Cf notamment les travaux de L. I. Langbein et A. J. Lichtmann, *Ecological Inference*, Sage, Beverly Hills, 1978 ou W. P. Shively «Ecological inference: the use of aggregate data to study individuals», *American Political science review*, 1969.

10 Source: L. I. Langbein et A. J. Lichtmann, op. cit., p. 20.

11 Alker va en fait plus loin et sa démonstration mathématique permet de déboucher sur une véritable typologie des fausses inférences: cf H. R. Alker, *Typologie des fausses inférences écologiques*, Dept de science politique, Université de Genève. Malheureusement la place manque ici pour développer ce raisonnement.

retrouver les corrélations individuelles. On a vu que cette attitude sous-tendait plusieurs approches.

2. En termes de phénomènes individuels et d'influences sociales. Cette démarche correspond bien à une sociologie qui suppose, d'une part un déterminisme social mais aussi une certaine liberté de l'acteur. Cette approche pose souvent des problèmes d'estimation des paramètres.
3. En termes stricts de corrélation collective. A propos de l'article de Robinson, *Boudon* écrit: «des circonstances historiques *identiques* et certains aspects de leur développement historique ont conduit certains états à:
a) importer et maintenir une masse importante de main d'oeuvre noire et
b) négliger le développement de leur système scolaire.»¹²

Comme il le signale lui-même, les approches de ce type sont plus simples, une variable globale satisfaisant à l'explication. L'utilisation de ce type d'interprétation implique pourtant certaines précautions. Principalement, il faut que l'unité d'analyse soit pertinente.

Scheuch écrit à ce propos: «Traiter un agrégat comme un système, un contexte dans lequel les individus évolueraient, implique bien sûr que cette unité tienne lieu de système social.»¹³ Par ailleurs il mentionne: «la question reste ouverte de savoir si une nation est réellement une société, et si une société est bien un contexte englobant, pertinent pour la plupart des comportements.»¹⁴

En Suisse, l'analyse agrégée est le plus souvent basée sur des unités fortement structurées: la commune et le canton ne sont pas de simples définitions administratives mais des niveaux du fédéralisme, avec un pouvoir politique propre. De plus ce sont des espaces géographiques bien reconnus par leurs habitants qui s'y identifient assez souvent.¹⁵ Dans ce sens, en répondant à la question de Scheuch, on peut parler de sociétés, contexte pertinent pour une analyse écologique du vote.¹⁶

12 R. Boudon, *L'analyse mathématique des faits sociaux*, Plon, Paris, 1967, pp. 162, 3.

13 E. K. Scheuch, «Social context and individual behavior», in M. Dogan et S. Rokkan, *Quantitative ecological analysis in the social sciences*, MIT Presse, Cambridge, 1969, p. 166, traduction personnelle.

14 idem, p. 164, traduction personnelle.

15 De nombreuses enquêtes ont posé la question de l'espace de référence. Cf en particulier M. Bassand, M. C. Brulhardt, F. Hainard et M. Schuler, *Les Suisses entre la mobilité et la sédentarité*, PPR, Lausanne, 1985 ou D. Sidjanski, C. Roig, H. Kerr, R. Inglehart et J. Nicola, *Les suisses et la politique*, Lang, Berne, 1975.

16 Ce point est beaucoup plus longuement développé dans D. Joye, *Structures sociales et structure politique*, Thèse de doctorat, université de Genève, 1984.

4. L'usage des données individuelles

Les données individuelles paraissent, a priori, contenir une information plus riche que les données agrégées: pour reprendre l'exemple du tableau 1, on est en possession de l'ensemble de l'information et plus seulement des «marges». De plus de telles données semblent particulièrement pertinentes si l'on postule un modèle démocratique «One man one vote» où il s'agit de recueillir les explications du votant sur ces propres motivations, expériences et préférences.

Toutefois, le raisonnement mené jusqu'ici montre, en fait, que l'on se borne, souvent, à interpréter de telles données en termes individuels, en oubliant la part d'explication due au contexte et aux valeurs sociales. Ceci est d'autant plus patent quand on pense à l'usage le plus répandu des tableaux croisés: on n'étudie pas l'attitude d'un groupe social en tant que tel mais bien si le fait d'avoir un attribut change l'opinion. Par exemple, on ne cherche que rarement à construire le groupe social des électeurs socialistes mais on observe plutôt la différence induite dans la probabilité de voter socialiste par le fait d'être ouvrier.

De tels problèmes arrivent souvent à cause de considérations pratiques: avec les effectifs disponibles dans les enquêtes, typiquement de 1000 à 2000 personnes, il est quasiment impossible, pour des raisons de fiabilité, de multiplier les croisements et les contrôles entre les variables. De plus, il est presque exclu de régionaliser les données, de les analyser suivant le contexte géographique ou institutionnel.¹⁷ Le type d'interrogation et d'interprétation sera donc forcément différent.

Il faut mentionner aussi que la variabilité des réponses individuelles est considérable ce qui explique que les proportions de variance expliquée soient relativement faibles. Il n'est pas rare que des modèles publiés dans les revues spécialisée n'expriment en fait que le tiers de la variance! Ce résultat est parfois un peu étonnant si l'on pense aux régularités que l'on peut retrouver lors de l'analyse agrégative.¹⁸ En fait l'agrégation, comme d'autres techniques statistiques, induit un effet de lissage sur la gamme des variations individuelles.

Enfin, autre caractéristique des données individuelles: si l'on met à part les panels, études remarquables tant par leur qualité que par leur rareté,¹⁹ elles ne constituent qu'une photographie de l'opinion à un moment donné mais ne permettent guère d'analyses diachroniques. Les études s'efforçant de tirer des tendances, des évolutions temporelles à partir de questions identiques posées dans

17 Il suffit de penser à l'effectif qu'il reste si l'on répartit 1000 personnes dans 26 cantons et demi-cantons!

18 Pour une intéressante discussion des régularités que l'on peut retrouver au niveau individuel, cf *R. Inglehart*, «Aggregate stability and individual-level flux in mass belief systems: the level of analysis paradox», *American political science review*, 1984.

19 On peut mentionner, par exemple, l'étude de *H. T. Himmelwilt*, *P. Humphreys*, *M. Jäger* et *M. Katz*, *How voters decide*, Academic Press, London, 1981, qui ont suivi des électeurs pendant plus de vingt ans.

différents sondage, telles les études de popularité, ne sont plus à proprement parler des analyses individuelles mais bel et bien de données agrégées: l'unité est ici un échantillon national à un temps donné. Ce type de données est parfois utilisé pour étudier les effets de l'activité du gouvernement sur le comportement politique ultérieur.²⁰

La perspective de recherche basée sur des données individuelles est donc bien différente de celle basée sur les données agrégées. A ces optiques différentes correspondent des méthodologies spécifiques.

5. Les modèles méthodologiques

Après ces considérations, on est en mesure d'examiner plus exactement les grands modèles méthodologiques qui correspondent à ces approches. Il est toutefois intéressant de signaler qu'ils peuvent être considérés comme des cas particuliers de stratégies de recherche beaucoup plus larges, inspirées aussi bien de l'analyse de système que du développement récent de la statistique.

5.1 Modèles ouverts et fermés

L'analyse systémiste a depuis longtemps posé la question des limites d'un système, de son caractère ouvert ou fermé.²¹ A ces deux approches correspondent deux démarches. A propos de l'économie, Godelier distingue l'approche en système fermé, où l'on considère un domaine comme «champ particulier d'activité» ou en système ouvert qui «interdit de décrire l'économie sans montrer en même temps sa relation avec les autres éléments du système social».²²

Dans la statistique moderne, on retrouve une opposition analogue entre les méthodes exploratoires, où l'on s'efforce de découvrir des structures en faisant un minimum d'hypothèses sur les données, n'excluant pas a priori l'inclusion de connaissances extérieures et les techniques confirmatoires où l'on tâche de vérifier des relations ou des modèles parfaitement spécifiés. Bref, pour reprendre une analogie de *Tukey*, d'un côté la démarche d'un détective qui cherche des indices, de l'autre celle d'un juge qui tranche une question.²³

20 Pour une discussion, critiquable par certains aspects, de ce type d'analyse, cf par exemple *G. H. Kramer* «The ecological fallacy revisited: aggregate versus individual level findings on economics and elections, and sociotropic voting» in *American political science review*, 1983.

21 Même si peu de chercheurs posent encore un système a priori fermé, la démarche réalisée pour la mesure des variables s'apparente bel et bien à une opération de fermeture du système.

22 *M. Godelier, Rationalité et irrationalité en économie*, Maspéro, Paris, 1969, t II, p. 140. Cette discussion rejoint en fait la distinction entre modèles «soft» et «hard» opérée par *A. Rapoport*, «La théorie moderne des systèmes, un guide pour faire face au changement, *Revue française de sociologie*, 1970, ou *H. Ozbekhan*, «Esquisse d'une théorie de la planification», in *Prospective et politique*, OCDE, 1969.

23 Cf *J. Tukey, Exploratory data analysis*, Addison Wesley, 1977.

Les deux grandes démarches que j'ai présentées correspondent, le plus souvent car des exemples a contrario sont aussi intéressants, chacune à un de ces types de modèle: approche ouverte, modèle exploratoire pour les données agrégées; approche fermée et modèle confirmatoire pour les données individuelles.

5.2 L'analyse des données agrégées

L'analyse des données agrégées, telle que je l'ai présentée, offre souvent les caractéristiques suivantes:

1. La démarche est basée sur une unité d'analyse sur laquelle on possède plus d'information que les seules données à disposition. Par exemple le fait de trouver une participation élevée à Schaffhouse n'a pas le même sens que de voir l'individu 845 d'un sondage avoir un comportement particulier. Le programme d'analyse utilisé devrait pouvoir tenir compte d'une telle propriété.²⁴
2. La quantité d'information à disposition est considérable: il suffit de penser, dans le domaine qui m'occupe ici à l'ensemble des données de votations, d'élections et du recensement pour s'en convaincre!²⁵ Ces données sont souvent disponibles sur une longue durée.
3. La base d'analyse est un système social mais non un individu avec ces caractéristiques psychologiques, sur lequel l'inférence n'est pas légitime.

L'ensemble de ces caractéristiques amène une approche ouverte: on dispose de plus d'information que les seules variables quantitatives.

On peut trouver un exemple typique de ce genre d'approche dans des travaux récents de géographie électorale. L'analyse d'un grand nombre d'élections (ou de votations) permet de retrouver les grands clivages qui structurent la vie politique d'un pays, rejoignant ainsi une préoccupation chère, notamment, à un Rokkan.²⁶ En outre ce type d'analyse permet une interprétation duale des données: non seulement on retrouve des grands clivages mais on dégage aussi des régularités régionales.

D'un point de vue méthodologique, les outils qui correspondent à une telle démarche sont de l'ordre de la statistique multivariée. Si l'analyse factorielle et ses variantes sont actuellement bien connues, notamment depuis les travaux de Rummel,²⁷ il faut signaler, sans être en aucune façon exhaustif, quelques développements importants.

24 C'est un particulier le cas d'un programme comme EDA élaboré par E. Horber de l'Université de Genève. C'est, à ma connaissance, le seul programme élaboré dans cette optique.

25 Notons au passage que seul un recensement exhaustif peut être correctement régionalisé.

26 Cf par exemple S. Rokkan, *Une famille de modèles pour l'histoire comparée de l'Europe occidentale*, Journée d'étude de l'association française de science politique, Paris, 1976.

27 Cf R. J. Rummel, *Applied factor analysis*, Northwestern University Press, Evanston, 1969.

Les développements à partir de l'analyse en composantes principales:²⁸ on peut mentionner les comparaisons de facteur, le calcul de variables supplémentaires, les analyses en composantes principales de scores factoriels, etc. De tels modèles ont permis de montrer les grands clivages qui structurent la vie politique suisse mais aussi une remarquable stabilité dans la distribution spatiale du vote. Ces travaux mériteraient un développement systématique mais ils ont déjà été bien utilisés pour décrire la vie politique de certains pays, de la Suisse en particulier.²⁹ Dans le même registre citons les applications particulières, notamment de l'analyse en corrélation canonique, que réalise W. Miller³⁰ pour étudier la stabilité et le changement du vote.

Les techniques de classifications automatiques ou analyse typologique ont connu un développement important dans les années septante, avec un accès facilité à des ordinateurs rapides. Pour s'en convaincre, il suffit de consulter la littérature consacrée à ce domaine.³¹ En cherchant à créer des groupes homogènes, on peut faire apparaître des régularités: par exemple, alors que l'analyse factorielle des votations en Suisse permet de décrire des clivages comme progressisme-traditionnalisme ou centre-périphérie, l'analyse typologique montre des groupes de cantons au comportement bien défini: «nein-sager» d'une part, comme on pouvait s'y attendre, mais aussi un groupe de «ja-sager» qui tend plus facilement que la moyenne suisse à accepter les votations.³²

L'analyse des échelles multi-dimensionnelles représente une famille de méthodes, appliquées dans de nombreux domaines. L'avantage de l'une d'elle, *Indscal*,³³ est de permettre de résoudre un problème courant mais très délicat: le «cube de données». Les techniques classiques s'appliquent à une matrice de données composée de cas en lignes et de variables en colonnes. Il arrive pourtant que l'on dispose de structure de données plus compliquées: par exemple du même ensemble de variables à différents moments dans le temps. Un exemple typique en serait les résultats des élections au Conseil National par cantons depuis 1919. On obtiendrait alors les dimensions communes qui sous-tendent l'ensemble de ces élections, mais aussi une pondération permettant de voir l'importance d'une dimension particulière à un moment donné.

28 R. Nef montre d'excellents exemples de l'utilisation de l'analyse factorielle sur les votations. Cf notamment R. Nef, «Struktur, Kultur und Abstimmungsverhalten», *Revue suisse de sociologie*, No 2, 1980. Notons que l'esprit de ce type de recherche se trouve déjà chez P. Gilg et E. Frischknecht, «Regionales Verhalten in eidgenössischen Volksabstimmungen», *Annuaire suisse de science politique*, 1976.

29 Pour la France, cf F. Bon, *Les élections en France*, Seuil, Paris, 1978 et pour la Suisse, R. Nef, op. cit. De nombreuses propriétés de ces méthodes multivariées sont également utilisées pour décrire la situation suisse dans D. Joye op. cit.

30 Cf W. Miller, *Electoral dynamics*, Mac Millan, London, 1977.

31 Citons comme livre de base B. Everitt, *Cluster Analysis*, Heinemann, London, 1974 ou M. R. Anderberg, *Cluster analysis for applications*, Academic Press, New York, 1973.

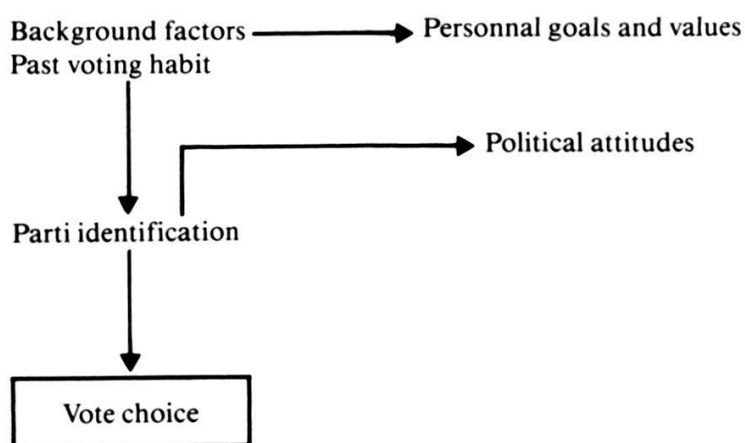
32 Pour plus de détail cf Joye, op. cit.

33 Cf notamment A. P. M. Coxon, *The User's guide to multidimensional scaling*, Heinemann, London, 1982, ou J. B. Kruskal et M. Wish, *Multidimensional scaling*, Sage, Beverly Hills, 1978.

Enfin, même si les arguments présentés ci-dessus tendent à justifier plutôt d'une méthodologie ouverte lorsque l'on utilise les données agrégées, on peut mentionner certaines applications de stratégies confirmatoires, souvent proches de modèles économétriques.³⁴

5.3 L'analyse des données individuelles

En matière de sondage, ce sont beaucoup plus des modèles fermés qui sont appliqués. La méthodologie qui y correspond est beaucoup plus d'ordre confirmatoire. On peut s'en rendre compte en présentant rapidement un modèle particulier: le modèle de Michigan. C'est en effet le modèle qui a été le plus souvent utilisé pour étudier le comportement politique et la plupart des auteurs y font référence, au moins pour le critiquer, le commenter ou le compléter. Fondamentalement on peut le résumer par le graphique 2.³⁵



Graphique 2: The strong version of vote choice

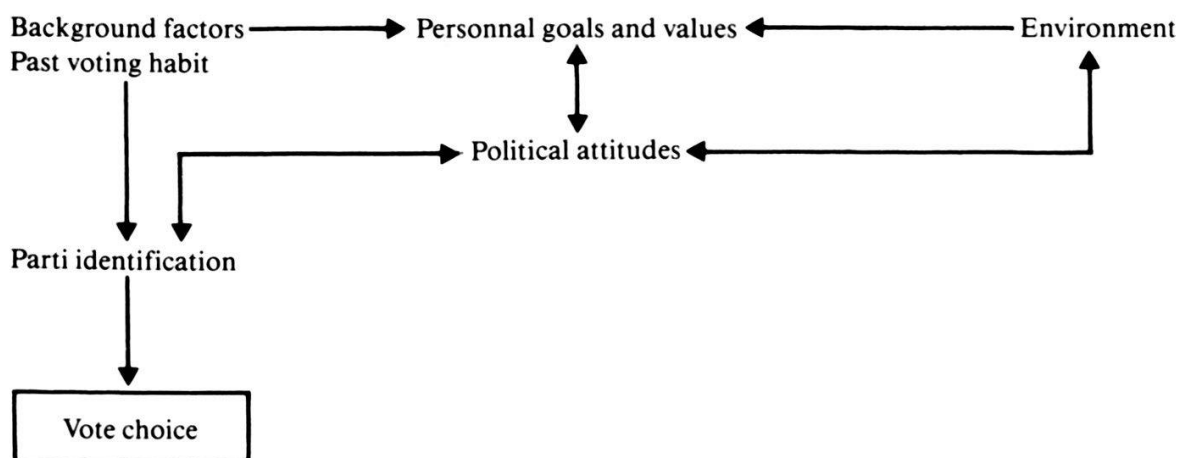
Les caractéristiques essentielles de ce modèle peuvent être résumées dans les traits suivants:

1. Un grand nombre de variables décrivant la situation personnelle du répondant: âge, éducation, revenu, etc.
2. Un rôle clé attribué à une variable particulière, l'identification partisane. Certains auteurs jugent d'ailleurs un peu tautologique de prédire le vote par l'appartenance partisane mais il ne faut pas oublier qu'en Suisse par exemple, tous les électeurs sont loin d'être capables de s'identifier à un parti. Dans ce sens cette variable est loin d'être négligeable.

34 Pour la Suisse, on peut mentionner, par exemple, des travaux comme *F. Schneider et W. Pommerhne, Ideologie versus Eigennutz, eine empirische Untersuchung der AH-Vorlage von 1972*, présenté au Congrès suisse de science politique, Jongny, 1983.

35 Tel qu'il est présenté par *Himmelwilt et al.*, op. cit., p. 9.

Ce modèle a subi de nombreuses critiques et extensions. Par exemple *Himmelweit* et al parlent d'un modèle II (graphique 3) qui accorde une place beaucoup plus grande aux valeurs et aux autres facteurs de l'environnement, notamment les media.³⁶



Graphique 3: The weak version of vote choice

On pourrait multiplier les exemples de ce type. Par exemple *Himmelweit* et al. développent leur propre modèle, qu'ils qualifient de «consommériste»,³⁷ ce qui renvoie de nouveau à certains aspects de l'économie. Ce qui m'intéresse pourtant ici c'est que l'on se situe ici dans une toute autre optique que celle de l'analyse des données agrégées telle que je l'ai présentée: la référence au caractère social ou sociologique ne se fait plus en considérant des caractéristiques de l'unité d'analyse mais par l'intermédiaire des valeurs, du passé et de l'insertion sociale du répondant qui déterminent les attitudes et le vote. De plus, il s'agit réellement de modèles, où les variables à utiliser sont strictement déterminées.

Classiquement, ces recherches utilisent des techniques d'analyse causale, montrant aussi bien l'influence d'une variable, indépendamment de celle des autres, que le pouvoir explicatif du modèle. Malheureusement cette technique exige des variables mesurées sur un échelle à intervalles, que l'on obtient rarement dans des sondages. A la suite des travaux de *Goodmann* notamment,³⁸ une technique a été développée qui permet quasiment la même opération sur des données nominales: les modèles log-linéaires. Basés sur une décomposition du X^2 ,³⁹ l'idée fondamentale est de trouver un modèle aussi parcimonieux que possible qui permette de décrire les relations entre les variables. Dans le même

36 idem.

37 Cf *Himmelwilt* et al., *ibid*, p. 13.

38 Pour une excellente présentation, cf *P. Whiteley*, «The analysis of contingency tables» in *D. McKay, N. Schofield et P. Whiteley, Data analysis and the social sciences*, Frances Pinter, London 1983 ou *G. Upton, The analysis of cross-tabulated data*, Wiley, New York, 1978.

39 Le chi carré est une mesure statistique de l'écart à l'indépendance entre deux variables.

sens, il faut mentionner des techniques comme l'analyse logit ou probit ⁴⁰ qui permettent de relaxer certaines hypothèses statistiques des modèles complexes.

Un autre modèle est conceptuellement important pour l'analyse du comportement politique: celui de Downs. Il part d'une optique très semblable à celle de Simon: celle d'un électeur surchargé d'informations, le plus souvent incomplète, qu'il ne peut contrôler mais qui sont déterminantes pour le vote: «L'incertitude altère cette situation en supprimant la compétence du votant de lier une décision partisane et sa propre idéologie. Les électeurs ne connaissent pas les détails des décisions du gouvernement, et ne peuvent d'ailleurs pas les connaître sans un coût significatif. Et même s'ils les connaissaient, ils ne pourraient pas toujours prédire ce qu'une décision donnée va entraîner. Ils sont dès lors incapables de suivre les conséquences de chaque décision et de les relier à leur propre idéologie. De plus, ils ne savent pas à quels problèmes le gouvernement va devoir faire face dans la prochaine période. Dans ces conditions, plus d'un électeur trouve les idéologies utiles car elles permettent de ne plus relier chaque décision à sa propre philosophie. Les idéologies l'aident à mettre en évidence les différences entre les partis; elles peuvent être utilisées comme un échantillon des moyens de différenciation. Par là même, un électeur peut faire l'économie du coût de l'information sur une large gamme de décisions.» ⁴¹

Cette perspective est importante à divers égards. Sur un plan conceptuel, elle introduit la notion d'utilité du vote, non seulement dans son expression mais aussi dans le fait d'aller ou non voter, ce qui introduit des perspectives intéressantes pour l'étude de la participation. Sur un plan plus méthodologique, elle introduit l'idée que l'un des éléments déterminants du vote est la situation de l'électeur dans un système de référence marqué par des grandes dimensions idéologiques. A partir de là, certains auteurs ont défini ce que l'on peut appeler une «théorie spatiale du vote». ⁴² Par exemple, si un électeur se définit comme de gauche, il doit, dans un système bipartisan, voter pour le parti qui se situe le plus proche de lui.

Un des développements de cette construction est de chercher à construire la position des candidats dans un tel système, de pouvoir les distinguer de la situation de leur parti. L'analyse des échelles multidimensionnelles a été utilisée dans ce sens mais il faut mentionner une méthode, bien plus générale par ailleurs, qui permet de construire des représentations spatiales du système de représentation des électeurs: l'analyse factorielle des correspondances. Très utilisée en France, elle travaille également sur une décomposition du X^2 . Souvent utilisée suivant une optique exploratoire, elle montre une fois de plus que logiques exploratoires et confirmatoires ne sont pas seulement fonction des données analysées

40 Cf J. H. Aldrich et F. D. Nelson, *Linear probability, logit and probit models*, Sage, Beverly Hills, 1984.

41 A. Downs, *An economic theory of democracy*, Harper and Row, New York, 1957, p. 98, traduction personnelle.

42 Pour un développement récent, cf J. M. Enelow et M. J. Hinich, *The spatial theory of voting*, Cambridge university Press, Cambridge, 1984.

mais dépendent aussi de la problématique de recherche. Fondamentalement le but d'une telle analyse est d'obtenir une représentation simplifiée des données, un peu à l'instar de l'analyse factorielle, en reprenant le maximum d'information dans un espace de dimension moindre. Le chercheur peut ainsi disposer d'une «carte» de proximité entre les modalités de réponse, le plus souvent dans un plan bidimensionnel.⁴³

Si l'analyse factorielle des correspondances est aujourd'hui bien connue et ses propriétés de réduction d'information largement utilisées, il faut en mentionner un développement intéressant: la méthode TRI2 de Philippe Cibois.⁴⁴ Un des atouts de cette approche, c'est de se placer dans la situation d'un utilisateur qui interprète un résultat d'analyse factorielle des correspondances par la simple proximité des indicateurs. Si l'on inscrit sur ce plan les relations statistiquement significatives, on peut alors écarter les effets de projection des espaces d'ordre supérieur et voir se dégager des ensembles de modalités pertinents pour l'explication.

6. Conclusion et développements

Si l'on considère, en conclusion, les applications possibles de ces méthodes, il faut reconnaître que le système politique suisse jouit d'une qualité estimable pour le politologue: en matière de comportements politiques, les référendums et les initiatives produisent une quantité extraordinaire de données, en sus des élections que connaissent la majorité des pays à tradition démocratique. Reprenons rapidement quelques domaines potentiels d'application comme illustrations de difficultés qui peuvent se présenter ou thème de réflexion sur des problèmes à résoudre.

En données agrégées, il faut mentionner, en se limitant au seul comportement politique conventionnel,⁴⁵ les diverses votations aussi bien fédérales que cantonales ou communales. De sérieux problèmes méthodologiques restent pourtant à résoudre mais il s'agit autant de problèmes substantiels que de traitement statistique.

En particulier, alors que j'ai insisté sur l'importance de l'unité d'analyse, une question de niveaux reste posée. Certains travaux montrent que si l'on raisonne sur les grands clivages culturels, une analyse basée sur le canton est pertinente. Par contre, pour d'autres clivages, il est important de descendre au niveau de la commune tant des variables comme les modes de vivre et d'habiter deviennent

43 Dans cet annuaire, l'article de R. Knuesel et D. Joye montre l'utilisation de cette technique: on distingue le paysage politique vaudois tel qu'il se dégage des élections de mars 1986.

44 Cf Ph. Cibois, *L'analyse des données en sociologie*, Paris, PUF, 1984.

45 Les remarquables analyses de l'activisme politique de H. P. Kriesi, R. Levy, G. Ganguillet et H. Zwicky, *Politische Aktivierung in der Schweiz*, Rüegger, Diessenhofen, 1981 sont en fait une analyse agrégée, où l'unité d'agrégation est «L'évènement». Mais on sort en l'occurrence du domaine de l'analyse du vote.

importantes. Par exemple, une distinction entre progressisme et traditionalisme peut se retrouver entre des cantons comme Genève d'une part, Uri de l'autre. Pourtant, une semblable distinction joue entre Genève-ville et des communes comme Gy ou Vandoeuvres. Le problème se complique du fait que pour rendre compte de ce dernier effet, il importe de tenir compte du premier: généralement, Vandoeuvres peut être qualifiée de plus progressiste que Uri!⁴⁶ Il s'agit de trouver un modèle global qui rende compte de ces interrelations en fonction des dimensions en jeu.

Un autre problème est redoutable et, à ma connaissance, n'a pas encore trouvé de solution satisfaisante: l'effet multipack ou le fait que plusieurs votations fédérales soient en même temps soumises au vote, ceci sans exclure d'autres enjeux dans certains cantons ou certaines communes. Si l'on peut légitimement supposer, au vu de la régularité des résultats, que cet effet n'est pas déterminant pour le résultat, il est loin d'en être de même pour la participation: on sait que les motivations au fait d'aller voter sont différentes suivant qu'il s'agit d'enjeux fédéraux, cantonaux ou communaux.⁴⁷ Même à un niveau donné un enjeu peut être prépondérant même s'il est extrêmement difficile de savoir a priori quel est le déterminant.

La réponse à une telle question se trouve d'ailleurs peut-être dans des données individuelles qui devraient permettre de raisonner plus correctement sur ce qui motive l'électeur. Dans le même esprit, on peut mentionner les études qui restent à réaliser, en Suisse, sur la «volatilité électorale», tendance de l'électeur à changer de parti mais peut être aussi à quitter le système politique pour se réfugier dans l'aliénation politique ou se confier à d'autres formes d'agrégation des intérêts. Une telle préoccupation me paraît d'autant plus intéressante alors que les signes de désintérêt du citoyen pour le comportement politique conventionnel sont loin de disparaître⁴⁸ et que, analyse VOX mises à part, on souffre cruellement en Suisse d'un manque de grandes enquêtes de comportement politique.

Dans le même sens, le développement des liens entre analyse de sondage et analyse agrégée serait à développer. L'analyse agrégée a en effet aidé à dégager de grands clivages importants pour la vie politique. Dans un autre sens, et là on peut réellement parler d'interprétation duale à son propos, elle a permis de mettre en lumière des persistances de différences régionales dans le comportement politique montrant de véritables façons différentes de vivre et de participer.⁴⁹ De son côté, l'analyse des données individuelles a sans conteste

46 Pour une première tentative dans ce sens, on peut consulter notamment D. Joye et M. Schuler, «Agglomérations et comportement politique» in M. Bassand et J. Csillaghy, *Politiques locales et gestion urbaine*, IREC-EPFL, Lausanne, 1986.

47 Cf par exemple, U. Windisch, *Luttes de clans, luttes de classe. Chermignon, la politique au village*, Lausanne, 1976.

48 Cf D. Joye et I. Papadopoulos, «La mobilisation partisane est-elle en crise?» in *Annuaire suisse de science politique*, Haupt, Berne, 1986.

49 Cette diversité est bien décrite dans M. Schuler et al., *Atlas structurel de la Suisse*, Ex Libris, Zurich, 1985.

permis de mieux comprendre les motivations de l'électeur, aussi bien à participer au jeu politique que les modalités de son choix. Pourtant, ces deux domaines restent, notamment à cause des raisons que j'ai évoquées, bien trop séparés, et il faudrait se donner les moyens conceptuels d'une synthèse.

Mais, au delà de ces exemples qui n'offrent pas un portrait exhaustif du problème, il faut insister sur la profonde imbrication entre hypothèses, données et méthodes. Il n'existe pas de méthode toute faite qui soit la panacée en matière de comportement politique. Il n'est pas non plus de méthode qui soit a priori mauvaise. Il est par contre essentiel de réfléchir à l'adéquation de ses propres données et de la méthodologie à une perspective de recherche. Dans ce sens, et les quelques remarques formulées ici y sont une invite, il reste une large place à «l'imagination sociologique» pour trouver, ou adapter, une méthodologie à ses préoccupations scientifiques.