

# LA COMMISSION DES SONDAGES, GARDIEN DU TEMPLE DES SONDAGES ÉLECTORAUX EN FRANCE

Marc Christine<sup>1</sup> & alii<sup>2</sup>

<sup>1</sup> *Ancien Inspecteur Général de l'Insee, France, mchristine7577@gmail.com*

## Résumé.

La Commission des sondages a été mise en place par une loi de 1977, modifiée en 2002 et 2016. Cette loi a fixé des mécanismes de nature à garantir la qualité et l'objectivité des sondages d'opinion publiés. La Commission des sondages est chargée d'assurer le respect de ces règles et de réguler l'activité des instituts de sondage dans le domaine politique.

Selon l'article 5 de la loi, « la Commission des sondages est chargée d'étudier et de proposer des règles tendant à assurer dans le domaine de la prévision électorale l'objectivité et la qualité des sondages publiés ou diffusés tels que définis à l'article 1er.

La Commission a tout pouvoir pour vérifier que [ces] sondages [...] ont été commandés, réalisés, publiés ou diffusés conformément à la présente loi et aux textes réglementaires applicables ».

L'objet de ce papier est de décrire les missions et attributions de la Commission, de présenter son fonctionnement récent dans le contexte des élections européennes de juin 2024 et d'en dégager une réflexion d'ensemble sur la pratique des enquêtes d'intentions de vote en France.

L'auteur principal de ce papier – ancien statisticien public - est membre de la Commission mais s'exprime à titre personnel dans le cadre d'une démarche objective et si possible scientifique, sous l'aval cependant du Président de la Commission. Le but est notamment, au travers de la présentation au Colloque Sondages, de mieux faire connaître la Commission et de lui assurer une publicité au-delà des cercles habituels qui suivent ses travaux.

À travers l'action de la Commission, l'objectif est aussi de mieux connaître les pratiques des instituts qui réalisent ces enquêtes, dégager leurs points communs et leurs différences.

Enfin, une réflexion plus méthodologique conclura ce papier, ouvrant des voies pour d'éventuels futurs travaux de recherche ou d'amélioration ou de normalisation des techniques.

**Mots-clés.** Commission des sondages, loi, sondage électoral, institut de sondage, access panel

## Abstract.

The Commission des sondages was set up by a 1977 law, amended in 2002 and 2016.

This law established mechanisms to guarantee the quality and objectivity of published opinion polls. The Commission des sondages is responsible for ensuring compliance with these rules and regulating the activities of polling organizations in the political arena.

According to article 5 of the law, "the Commission des sondages is responsible for studying and proposing rules designed to ensure the objectivity and quality of opinion polls published or broadcast as defined in article 1, in the field of electoral forecasting.

The Commission is empowered to verify that the surveys defined in Article 1 have been commissioned, carried out, published or distributed in accordance with the present law and the applicable regulatory texts".

The purpose of this paper is to describe the Commission's missions and remit, to present its recent operations in the context of the June 2024 European elections, and to offer an overall reflection on the practice of polling for voting intentions in France.

The main author of this paper - a former public statistician - is a member of the Commission, but expresses his views in a personal capacity as part of an objective and, if possible, scientific approach, albeit with the endorsement of the Commission's Chairman. The aim of presenting this paper at the Colloque Sondages is to raise the Commission's profile and publicize it beyond the usual circles that follow its work.

Through the Commission's action, the aim is also to gain a better understanding of the practices of the institutes that carry out these surveys, and to identify their commonalities and differences.

Finally, a more methodological reflection will conclude this paper, opening up avenues for possible future research, improvement or standardization of techniques.

**Keywords.** Commission des sondages, law, electoral poll, polling institute, access panel

## 1 Introduction

La Commission des Sondages a été mise en place par une loi de 1977, modifiée en 2002 et 2016.

L'objet de ce papier est de décrire les missions et attributions de la Commission, de présenter son fonctionnement récent dans le contexte des élections européennes de juin 2024 et d'avoir une réflexion d'ensemble sur la pratique des enquêtes d'intentions de vote en France.

L'auteur principal de ce papier est membre de la Commission mais s'exprime à titre personnel dans le cadre d'une démarche objective et si possible scientifique, sous l'aval cependant du Président de la Commission. Le but est notamment, au travers de la présentation au Colloque Sondages, de mieux faire connaître la Commission et de lui assurer une publicité au-delà des cercles habituels qui suivent ses travaux.

Cela permettra, à travers l'action de la Commission, de mieux connaître les pratiques des instituts qui pratiquent ces enquêtes, dégager leurs points communs et leurs différences. Mais il ne s'agit en aucun cas de juger de la qualité du travail ni de la Commission ni des instituts qui procèdent aux enquêtes, ni d'analyser la qualité des sondages et leur pouvoir prédictif, au moins pour le dernier scrutin que la Commission a eu à connaître : les élections européennes de juin 2024.

## 2 Attributions légales, champs de compétence et pouvoirs normatifs

La Commission des Sondages a été instituée par la loi n°77-708 du 19 juillet 1977 relative à la publication et à la diffusion de certains sondages d'opinion. Cette loi, modifiée en 2002 puis en 2016, a pour objet d'empêcher que la publication de sondages électoraux vienne influencer ou perturber la libre détermination du corps électoral.

### 2.1 Une définition des sondages

Dans sa version initiale, la loi de 1977 ne donnait pas de définition des sondages d'opinion auxquels elle s'applique. Dans sa version issue de la loi du 25 avril 2016, elle comporte désormais, dans son article 1, une définition du sondage :

« Un sondage est, quelle que soit sa dénomination, une enquête statistique visant à donner une indication quantitative, à une date déterminée, des opinions, souhaits, attitudes ou comportements d'une population par l'interrogation d'un échantillon. [...]. Les personnes interrogées sont choisies par l'organisme réalisant le sondage de manière à obtenir un **échantillon représentatif**<sup>1</sup> de la population concernée ».

La Commission a retenu un champ d'application qui peut être modulé au cours du temps. À toute époque, les sondages sur les intentions de vote ou sur l'opinion des électeurs à l'égard des candidats relèvent de la loi. En revanche, un sondage sur une question de société peut ou non entrer dans le champ de sa compétence selon qu'au moment du sondage, cette question est ou n'est pas au centre du débat électoral.

**Un cas particulier concerne les** simulations de vote opérées sur des panels d'internautes. Dans la mesure où ces enquêtes ne seraient pas menées auprès d'échantillons représentatifs de la population, elles n'entreraient pas dans le champ de la loi. Il est donc impératif qu'elles soient accompagnées de précautions de présentation, indiquant qu'il ne s'agit pas de sondages au sens de la loi et appelant l'attention des lecteurs sur la prudence avec laquelle il convient d'en interpréter les résultats.

### 2.2 L'obligation de fournir des notices descriptives

Une autre évolution importante de la loi a été que la publication ou la diffusion des sondages électoraux s'accompagne d'indications précises. À cette fin, elle instaure l'obligation pour l'organisme ayant réalisé le sondage, de **déposer auprès de la Commission avant toute publication d'un sondage, une notice** décrivant les conditions d'élaboration des sondages et différentes caractéristiques techniques. Les articles 2 et 3 les listent :

- Les noms de l'organisme ayant réalisé le sondage et de son ou ses commanditaires ;
- Le nombre de personnes interrogées ;
- La ou les dates auxquelles il a été procédé aux interrogations ;
- Le texte intégral de la ou des questions posées ;
- Une mention précisant que tout sondage est affecté de marges d'erreur ;

---

<sup>1</sup> C'est l'auteur qui souligne.

- Les marges d'erreur des résultats publiés ou diffusés, le cas échéant par référence à la méthode aléatoire.

Et plus précisément, pour expertise détaillée :

- L'objet du sondage ;
- La méthode selon laquelle les personnes interrogées ont été choisies, le choix et la composition de l'échantillon ;
- Les conditions dans lesquelles il a été procédé aux interrogations ;
- La proportion des personnes n'ayant pas répondu à l'ensemble du sondage et à chacune des questions ;
- S'il y a lieu, la nature et la valeur de la gratification perçue par les personnes interrogées ;
- S'il y a lieu, les critères de redressement des résultats bruts du sondage.

Ces notices sont rendues publiques par la Commission et toute personne peut y accéder.

► Deux points importants doivent être soulignés ici :

- Le champ de la loi concerne des sondages prévus pour **être publiés ou diffusés**. Certains sondages peuvent ne pas être destinés à la publication mais ils tombent sous le coup de la loi dès lors que tout ou partie de leurs résultats sont divulgués.
- L'envoi des notices à la Commission doit être **préalable à la publication du sondage**.

## 2.3 Pouvoir de sanction

De manière générale, la Commission des sondages est une autorité chargée de veiller au respect de cette réglementation.

Elle peut, en cas de non-respect de la loi, publier ou diffuser une **mise au point** ou, le cas échéant, mentionner les indications qui n'auraient pas été publiées ou diffusées.

Cette mise au point est diffusée dans des conditions qui lui assurent une audience équivalente à celle du sondage incriminé, notamment dans les media qui l'auraient publié.

# 3 Fonctionnement

## 3.1 Un renouvellement récent

La Commission a été entièrement renouvelée à la fin de l'année 2023. Le texte officiel a été publié au JO du 28 décembre 2023 et a arrêté la liste de ses membres. Ceux-ci - comme pour les précédentes Commissions - sont issus du Conseil d'État, de la Cour de Cassation et de la Cour des Comptes, soit six personnes titulaires plus six suppléants, auxquels s'adjoignent trois personnalités qualifiées désignées respectivement par le Président de la République, le Président du Sénat et la Présidente de l'Assemblée Nationale. Le Président de la Commission est un président de section honoraire du Conseil d'État.

Les services permanents de la Commission sont constitués d'un Secrétaire général, d'un secrétaire permanent et de deux experts indépendants.

Enfin, la Commission dispose d'un site internet qui assure sa publicité et sa transparence. Le site publie, outre les textes législatifs et réglementaires, les notices ainsi que les avis, mises au point etc...

### 3.2 Un travail régulier, voire quotidien

On peut résumer simplement le travail de la Commission (de manière très réductrice..) en disant qu'en l'absence de problème, il n'y a pas d'action particulière ! En effet, la Commission des sondages **ne délivre pas explicitement de label**. Elle intervient lorsqu'un sondage pose problème au regard des règles fixées par le législateur.

Cependant, les experts font un travail quotidien (surtout en période pré-électorale) d'examen et de vérification des notices. Ce n'est que s'il y a problème que la Commission débat en séance plénière en vue d'arrêter une position. Dans ce cas, elle peut émettre une mise au point qui doit être publiée dans le média qui a commandité le sondage et qui est portée à la connaissance du public via le site internet de la Commission.

Notamment : des notices insuffisantes, des erreurs manifestes dans les chiffres techniques fournis, des modifications de méthode non justifiées et surtout non explicitées peuvent conduire à des mises au point.

Un exemple de mise au point récente :

La Commission des sondages a décidé d'ordonner, sur le fondement de l'article 9 de la loi du 19 juillet 1977, la mise au point suivante qui sera publiée par le site en ligne de [media], partenaire média de XXX :

*« Après une instruction approfondie et de nombreux échanges avec la société concernée, la Commission des sondages a pu constater plusieurs séries d'insuffisances de la part de la société XXX dans l'élaboration de certains sondages réalisés à l'occasion de l'élection au Parlement européen en 2024.*

*[...]*

*Il est rappelé que l'écart relativement faible des estimations publiées par cet institut avec les estimations publiées par d'autres instituts au cours de la même période ne saurait en soi l'exonérer de l'obligation d'élaborer des sondages relatifs au débat électoral dans des conditions qui en assurent l'objectivité et la traçabilité. ».*

En période électorale, comme lors des élections européennes, l'activité de la Commission et, notamment, des experts, est évidemment intense. Elle devait ralentir après le 9 juin 2024 mais la vie politique en a décidé autrement.

### 3.3 Des communiqués de rappel aux obligations légales

La Commission veille au respect des textes législatifs en vigueur. Outre son travail quotidien de vérification des notices envoyées par les instituts, la Commission a rappelé dans deux communiqués récents, avant les élections européennes :

- d'une part l'interdiction de la publication, la diffusion ou le commentaire de tout sondage électoral la veille et le jour de chaque tour de scrutin

- et, d'autre part, l'obligation de faire figurer l'intégralité des listes dans les choix d'intentions de vote proposés aux répondants. Ceci prend effet à partir du moment où l'ensemble des listes ont été arrêtées par le Ministère de l'Intérieur. Dans la pratique, avant le début de la campagne officielle, seules les listes connues ou anticipées, sous une appellation provisoire ou une tête de liste non encore désignée, étaient proposées.

En particulier, le 24 mai, en relation avec un institut posant problème :

*« Comme indiqué plus haut, ne pas tenir compte de l'intégralité de l'offre politique, désormais connue, n'est plus possible depuis le 17 mai.*

*Au surplus, une telle omission est de nature à introduire un biais dans l'administration de l'enquête. La commission des sondages appelle donc à regarder ce sondage avec réserve ».*

En ce qui concerne la tenue des élections législatives, la Commission a rappelé qu'en dehors du cas où serait effectué un sondage pour l'une des 577 circonscriptions, les instituts réalisent des sondages à l'échelle du territoire métropolitain, donnant une indication générale et nationale, à un moment donné, de l'intention de vote à ces élections. La commission assure le contrôle de tels sondages.

Mais les exercices de projection en nombre de sièges sont considérés comme ne présentant pas les mêmes encadrements méthodologiques que les sondages et demeurent tributaires d'un grand nombre de paramètres locaux (offre politique de chaque élection locale, personnalité des candidats, taux de participation...). De ce fait, la Commission des sondages précise qu'elle n'exerce aucun contrôle sur ces estimations en sièges et souligne que de tels projections doivent être interprétées avec prudence et distinguées clairement, dans les publications, des sondages eux-mêmes.

## **4 Compréhension des pratiques des instituts**

La nomination récente de la Commission a conduit ses membres, au cours des six premiers mois de son fonctionnement en 2024 - en sus de son fonctionnement courant -, à auditionner la plupart des grands instituts de sondage réalisant des enquêtes d'intentions de vote. Le but était de comprendre leurs pratiques, de mettre en évidence les lignes communes de leurs actions, d'identifier leurs différences, dans le but d'acquérir une connaissance sur leur fonctionnement, sans jugement de valeur, à ce stade, sur leur conduite.

De ces auditions, on peut retenir un tableau général des pratiques de ces instituts, des points qui font l'objet des préoccupations de la Commission et des constantes dans les questions que s'est posées la Commission ou les informations qu'elle a cherché à obtenir auprès des instituts de sondage. On notera ici les éléments les plus saillants issus de ces auditions.

### **4.1 Contexte général**

#### **4.1.1 L'opinion des instituts de sondage vis-à-vis de la Commission**

La plupart « jouent totalement le jeu » et apprécient d'avoir des relations régulières avec celle-ci et l'opportunité de pouvoir présenter devant elle leurs entreprises et leurs travaux. Il peut s'établir ainsi des relations de confiance. Inversement, la Commission se montrera

particulièrement interrogative et sourcilleuse face à des instituts faisant une apparition à l'occasion d'une élection, peu ou pas connus, voire pas implantés en France. Naturellement, les mêmes règles et obligations s'imposent à tous, instituts traditionnels comme nouveaux venus, mais la vigilance de la Commission sera plus aiguisée pour les seconds.

#### **4.1.2 Le diagnostic général posé par les instituts dans l'exercice de leur métier**

Tous soulignent des constantes dans le paysage politique actuel : volatilité de l'opinion, stabilisation et cristallisation tardive, difficulté à être certains de la stabilité d'un choix, populations difficiles à joindre (jeunes) et fragmentation de l'offre politique.

Plusieurs difficultés sont par ailleurs mises en avant : difficulté à prévoir l'abstention. Difficulté également dans les élections locales (municipales, départementales, voire législatives) dans la mesure où interviennent des considérations locales et où le passage d'un % de votants en faveur de tel parti à un nombre de sièges n'est pas simple et est très tributaire du mode de scrutin. Il peut arriver que des instituts locaux interviennent dans ces situations.

#### **4.1.3 Le poids de la fonction « sondages politiques » dans le modèle économique de l'institut**

Il est très variable. Pour certains instituts, il est très faible (quelques % de leur chiffre d'affaires) par rapport à l'ensemble constitué principalement d'enquêtes marketing. Pour d'autres, il est important, voire ne constitue que la seule activité de l'institut. Ceci conditionne évidemment le modèle économique des instituts, les derniers cités ayant en général un effectif employé très faible.

#### **4.1.4 Les trois acteurs dans la démarche du sondage électoral**

- Les media qui commanditent les enquêtes. Ce sont eux in fine qui ont la responsabilité du commentaire et de l'interprétation des résultats, en relation avec leur ligne éditoriale.
- Les instituts de sondage qui réalisent la collecte des informations et produisent et fournissent de premières analyses.
- Dans certains cas, le fournisseur des panels, s'il est distinct de l'institut.

Il faut noter que la présentation des résultats par le media commanditaire ne relève pas de la responsabilité de la Commission, qui ne juge et ne valide que la façon dont a été collectée l'information. Ainsi, en avril 2024, un député, mécontent du fait que des résultats publiés dans un media ne présentaient que ceux de certaines listes, a saisi la Commission des sondages et l'Arcom. En la matière, la Commission était juridiquement incompétente.

## **4.2 Les conditions d'organisation de la collecte**

### **4.2.1 Le mode de collecte**

Au cours des 40 dernières années, le mode de collecte a évolué depuis le face à face au

téléphone et est presque exclusivement constitué aujourd'hui de l'internet. La méthode standardisée est celle de « *l'access panel* », vivier de répondants potentiels sous forme d'adresses électroniques recueillies par tous moyens de contacts à l'occasion notamment d'achats en ligne, sur les réseaux sociaux ou parfois ciblées sur certaines sous-populations plus difficiles à joindre ou de comportement de réponse plus fragile, comme les jeunes.

Pour une enquête donnée, une sélection d'adresses est constituée (au moyen d'un procédé par quota, de manière à ce qu'elles soient « représentatives ») et les personnes sont sollicitées pour répondre, en général dans un délai très bref. La déperdition est évidemment très importante (ratio de l'ordre de 10 entre les sollicités et les répondants).

#### **4.2.2 La question de la propriété des access panels**

Selon les cas, les instituts sont propriétaires ou utilisent les services d'un gestionnaire de panels. Les arguments en faveur du panel propriétaire sont une meilleure garantie de la qualité du travail et une meilleure maîtrise du processus de production. Dans l'autre cas, la sous-traitance, même encadrée par des spécifications précises, ne permet pas toujours de vérifier que les sélections d'échantillons ont été faites au plus près des spécifications de l'institut.

À l'inverse, les instituts qui recourent à des gestionnaires de panels estiment que cela leur permet de se centrer sur leur cœur de métier : la réalisation d'enquêtes.

#### **4.2.3 La taille des panels**

Elle est difficile à cerner, dans la mesure où il y a soit comptabilisation du noyau des membres actifs, soit un élargissement à un cercle plus large, composé d'un grand nombre de « dormants ». D'où des chiffres allant de quelques centaines de milliers à plusieurs millions.

#### **4.2.4 L'ergonomie du questionnaire**

La tenue des élections européennes a posé la question du très grand nombre de listes candidates. D'une part, il est obligatoire qu'à partir du moment où celles-ci sont arrêtées officiellement, toutes figurent parmi les choix offerts aux répondants. D'autre part, se pose la question de l'ordre dans lequel les listes sont présentées : fixe, arbitraire, tournant aléatoirement ?

Enfin, de plus en plus de répondants, notamment les jeunes, peuvent être amenés à utiliser un support autre que l'ordinateur (smartphone, tablette..), ce qui pose d'évidents problèmes de lisibilité de listes longues .

### **4.3 Les traitements**

#### **4.3.1 Les filtres de traitement**

Il convient de ne retenir que les répondants qui ont une certitude assez forte d'aller voter.

Celle-ci peut être mesurée par des questions fermées ou par une échelle de notation sur l'intensité de l'intention. Il est souhaitable que ces règles d'exclusion soient fixes dans le temps pour un institut donné, voire d'un institut à l'autre, faute de quoi l'analyse diachronique risquerait d'être faussée. Un institut de sondage ayant modifié ces règles au cours de la période de pré-campagne, au motif qu'il procédait ainsi pour des enquêtes effectuées dans d'autres pays européens, se l'est vu reprocher par la Commission des sondages.

### **4.3.2 Les conditions d'exclusion ou de mise en sommeil**

Les instituts procèdent en général à une détection et à l'exclusion de répondants robotisés ou manifestement insincères ; cocher toujours la même case par exemple.

Par ailleurs, une période de latence de quelques mois est en général mise en place avant une sollicitation ultérieure.

Enfin, des unités dormantes depuis un temps trop long sont éliminées des panels. Compte tenu des flux d'entrée-sortie, les instituts estiment que les effectifs globaux restent stables.

### **4.3.3 Le redressement**

C'est une phase essentielle du traitement et les notices techniques examinées par la Commission doivent mettre en miroir les résultats bruts et les résultats redressés.

Les variables servant au redressement se composent tout d'abord des variables socio-démographiques standards et d'autre part des variables politiques, à savoir **la déclinaison dans le questionnaire des souvenirs de vote à une élection antérieure**, qui donne lieu à la comparaison avec les vrais résultats connus de cette ou ces élections.

Ce redressement politique est essentiel. Il permet à la fois de corriger les biais de sous-déclaration (non déclaration de votes en faveur de partis extrêmes) et aussi de donner une image plus fidèle du corps électoral au moment de l'interrogation. De surcroît, les questions sur le vote passé permettent des analyses fines produites en aval par les media ou les politologues, sur la question : comment l'électorat de X il y a 5 ans s'est-il reporté vers les partis Y en 2024 ?

La question se pose alors du choix des élections de référence. Dans le cas des élections européennes de 2024, ont été en général choisies l'élection européenne précédente (2019) – mais avec des risques d'oubli du vote réel ou une offre politique structurée différemment à l'époque - et la dernière élection présidentielle, élection nationale majeure et en l'occurrence plus récente.

On peut objecter que les modes de scrutins de ces deux élections sont très différents. Aussi, il semble préférable – et la plupart des instituts le font – d'utiliser les résultats du premier tour des présidentielles, reflétant mieux la diversité des opinions, et éventuellement, d'y rajouter ceux du second tour. Certains instituts ont cependant déclaré qu'ils n'utilisaient que le second tour, sans d'ailleurs justifier cette stratégie, ce qui est certainement très réducteur.

À la date de rédaction de ce papier, il n'est pas encore possible, pour les sondages relatifs aux élections législatives consécutives à la dissolution de l'Assemblée Nationale prononcée par le Président de la République le 10 juin 2024, de savoir quelle sera la stratégie des instituts en la matière.

#### **4.3.4 La rémunération des panélistes**

De l'avis général, celle-ci est infime et ne permettrait en aucun cas à un panéliste « multirécidiviste » d'en tirer un revenu appréciable. Dans certains cas, la possibilité est offerte d'accéder à des loteries.

#### **4.4 Le rythme des réinterrogations**

Il dépend évidemment de la commande des media. Certains instituts pratiquent le « *rolling* » : interrogations quotidiennes par tranches de 500 ou 1000 répondants par exemple et agrégation de trois vagues quotidiennes pour fournir une estimation plus robuste. D'un jour à l'autre, un noyau constant de 2/3 des répondants contribue aux résultats. Ceci assure un lissage des résultats.

En revanche, la réinterrogation des mêmes individus est peu pratiquée. Dans la mesure où les adresses électroniques sont conservées et stockées par le gestionnaire du panel, des réinterrogations pourraient cependant être mises en œuvre : l'intérêt serait de vérifier la stabilité des opinions, voire de comparer le vote réel d'un répondant au vote déclaré lors d'une enquête. De telles informations seraient précieuses dans une optique de prédiction du vote réel au moyen de modèles économétriques par exemple.

#### **4.5 Les risques**

Plusieurs risques sont pointés : cyber attaques, intrusion partisane, ingérence extérieure..

Les premiers risques sont appréciés et traités à l'instar de ceux qui peuvent affecter les systèmes d'information existant dans les entreprises de toute nature.

Quant aux seconds, ils semblent peu probables du fait que, même en cas d'inscription massive de catégories partisans, la probabilité qu'elles soient sollicitées simultanément dans une enquête donnée doit rester faible.

### **5 Réflexions statistiques**

La mécanique statistique (constitution des échantillons à partir des access panels, suivi des réponses, procédures d'arrêt de la collecte ..) relève des secrets commerciaux des instituts. En ce domaine, la Commission des sondages, mis à part des principes très généraux, n'a pas vocation à donner des prescriptions ni les moyens de contrôler dans le détail les techniques mises en œuvre. Néanmoins, on peut – sans engagement de la Commission sur ce point – dessiner des pistes méthodologiques permettant de formaliser une mécanique « souhaitable » ou pouvant servir de modèle de référence.

## 5.1 Les quatre paradigmes

- Dans ce cadre, il faut clairement admettre le **premier paradigme de l'échantillonnage non probabiliste**.

La statistique publique peut être amenée à considérer comme seule vertueuse celle qui s'appuie sur des échantillons probabilistes. C'est effectivement la technique qu'elle emploie. Aussi bien s'est-elle interdite par ailleurs de réaliser toute enquête sur les intentions de vote, qui sont donc dévolues à des instituts ad hoc, spécialisés ou non.

La littérature a souvent, et dans de nombreux articles, analysé, modélisé les méthodes de quota et tenté de mettre en exergue leurs propriétés et leurs faiblesses, mises en regard de celles des sondages probabilistes. Tel n'est pas le propos dans cet article : on admettra sans jugement le recours aux méthodes « empiriques ».

- **2<sup>ème</sup> paradigme**, lié au précédent, voire cause de ce dernier : **l'absence de base de sondage**. Celle-ci est remplacée par un vivier d'adresses électroniques recueillies par tous moyens, peu importe lesquels.
- **3<sup>ème</sup> paradigme : admettre les biais** liés au fait que seuls ceux qui disposent d'internet et ont les capacités cognitives et techniques suffisantes pour le faire sont répondants. Et admettre que restreindre le champ effectif de l'enquête à cette catégorie produit un biais réductible ou corrigible par les méthodes de redressement.
- **4<sup>ème</sup> paradigme : admettre que seuls les volontaires répondent** sans qu'il y ait véritablement de mécanisme explicite de connaissance, de modélisation et de correction de la non-réponse (a minima, cependant, via les procédures de redressement).

## 5.2 Une procédure idéalisée

Sous réserve d'admettre ces paradigmes, on pourrait néanmoins imaginer une formalisation souhaitable de la démarche statistique comme suit :

- 1<sup>ère</sup> étape : la base de sondage.

Le vivier d'adresses est considéré comme une base de sondage utile dans laquelle on puisera toutes les enquêtes mais qui ne représente qu'une fraction de la population de référence. Cette base peut être déformée en termes de structures socio-démographiques par rapport à celles de la référence. Il conviendrait donc de réattribuer des poids à chaque unité du vivier au moyen d'un *calage préalable sur des structures de référence connues* pour la rendre « représentative ».

Faute d'information sur la constitution de ce vivier – et sous réserve d'avoir éliminé en amont les doublons potentiels -, on ferait l'hypothèse que les poids initiaux sont identiques (de valeur 1) pour toutes les unités. Cette procédure de calage conduirait ainsi à des nouveaux poids  $\omega_i$  qui assureraient la conformité de la base redressée à la population de référence.

À noter qu'idéalement, les structures (données par l'Insee) doivent être celles du *corps électoral*, c'est-à-dire des inscrits sur les listes électorales. Même si cette inscription est obligatoire – mais non sanctionnée en pratique –, on estime qu'environ 10 % de la population française en âge de voter n'est pas inscrite. À cela s'ajoutent les jeunes dont l'inscription n'a pas encore été réalisée pour diverses raisons, dont, par exemple, ceux atteignant la majorité entre la date de la dissolution et celle des élections législatives de juin-juillet 2024.

- 2<sup>ème</sup> étape : la sélection d'un échantillon.

Les instituts procèdent à un échantillonnage par quota mais on en ignore le mécanisme exact ; on ne sait pas non plus si l'échantillon est réajusté au cours de la collecte (qui cependant est en général très brève ; une ou deux journées). On pourrait alors remplacer cette technique par un *tirage équilibré* selon un jeu de caractéristiques socio-démographiques connues sur les unités de la base, en partant des probabilités d'inclusion fictives constituées par les inverses  $1/\omega_i$  des poids calculés à la première étape.

On pourrait même raffiner la méthode en prenant en compte des probabilités de réponses a priori  $r_i$ , issues de l'analyse statistique des enquêtes passées et d'une modélisation des probabilités de réponse par catégories, selon un modèle logistique par exemple. Ainsi, les probabilités d'inclusion fictives utilisées dans ce tirage équilibré seraient le produit  $r_i/\omega_i$  de l'inverse des poids initiaux de calage par ces probabilités de réponse a priori.

- 3<sup>ème</sup> étape : redressement d'une enquête.

Là, à nouveau, on met en jeu une procédure de calage standard, qui corrigerait implicitement les biais liés à la non-réponse, à défaut de la modéliser (et sous réserve d'admettre tous les autres biais), en calant les répondants sur des structures socio-démographiques, mais aussi sur les caractéristiques de votes aux précédentes élections, données devant être évidemment être connues sur l'ensemble des répondants.

Les poids en entrée de la procédure seraient eux issus de la 2<sup>ème</sup> étape ; les poids finaux en résultant seraient les poids définitifs utilisés pour la production des résultats.

► On voit donc qu'on pourrait ramener des procédures empiriques telles que pratiquées par les instituts à trois phases de procédures contrôlées statistiquement selon des méthodes satisfaisant aux règles de l'art.

Il ne s'agit là que de propositions méthodologiques qui pourraient avoir pour intérêt d'unifier les pratiques des instituts et d'assurer une meilleure transparence des méthodes. En l'état actuel, les instituts restent évidemment maîtres de leurs techniques et la Commission des sondages n'a pas vocation à leur imposer des méthodes ou outils particuliers.

Aussi les idées développées ci-dessus sont-elles seulement des pistes pour la réflexion méthodologique qui pourraient être mises en œuvre à titre expérimental par ceux qui le

souhaiteraient, ne serait-ce que pour comparer les résultats ainsi obtenus à ceux issus de leurs pratiques habituelles.

## **Bibliographie**

Ardilly P., Peut-on se fier aux sondages empiriques ? *In* Courrier des statistiques N10, Novembre 2023 - Insee 113.

Ardilly P., Haziza D., Lavallée P. et Tillé Y, Les contributions de Jean-Claude Deville à la théorie des sondages et à la statistique officielle, incluant un article : « L'échantillonnage par quotas », Techniques d'enquête, Statistique Canada, décembre 2023.

Deville J-C, Särndal K-E, Une théorie des enquêtes par quotas, Techniques d'enquête Statistique Canada, 17, 177-195, 1991.

Site internet de la Commission des sondages, <https://www.commission-des-sondages.fr>