

Commission des sondages

Rapport d'activité 2022

Élection présidentielle et élections législatives

Mme Aubin, Présidente

M. Hoynck, Secrétaire général

Au cours de la période qui a précédé les élections présidentielle et législatives de 2022, l'activité de la commission s'est partagée entre :

- le contrôle systématique des conditions de publication des sondages ;
- l'examen systématique de la fiabilité de tous les sondages relevant de son champ de compétences ;
- l'examen quasi exhaustif de la qualité du travail de gestion des panels.

1 – Le contrôle systématique de la publication des sondages

- 1– 1 Une baisse apparente du nombre de sondages
- 1– 2 L'article 3 de la loi de 1977
- 1– 3 Les sondages de second tour
- 1– 4 La présence des médias étrangers
- 1– 5 Les publications la veille et le jour du scrutin dans les médias français
- 1– 6 Publication des estimations en sièges.
- 1– 7 Publication d'estimations locales tirées d'estimations nationales
- 1– 8 Publications diverses.
- 1– 9 Publications de la commission

2 – L'examen systématique de la fiabilité et de la sincérité de tous les sondages relevant de son champ de compétences

- 2– 1 Les moyens de la commission
- 2– 2 Les règles imposées par la commission
 - 2– 2.1 L'objectivité et la sincérité
 - 2– 2.2 La réclamation de « La France insoumise »
 - 2– 2.3 Cluster 17

3 – L'examen quasi exhaustif du travail des panélistes.

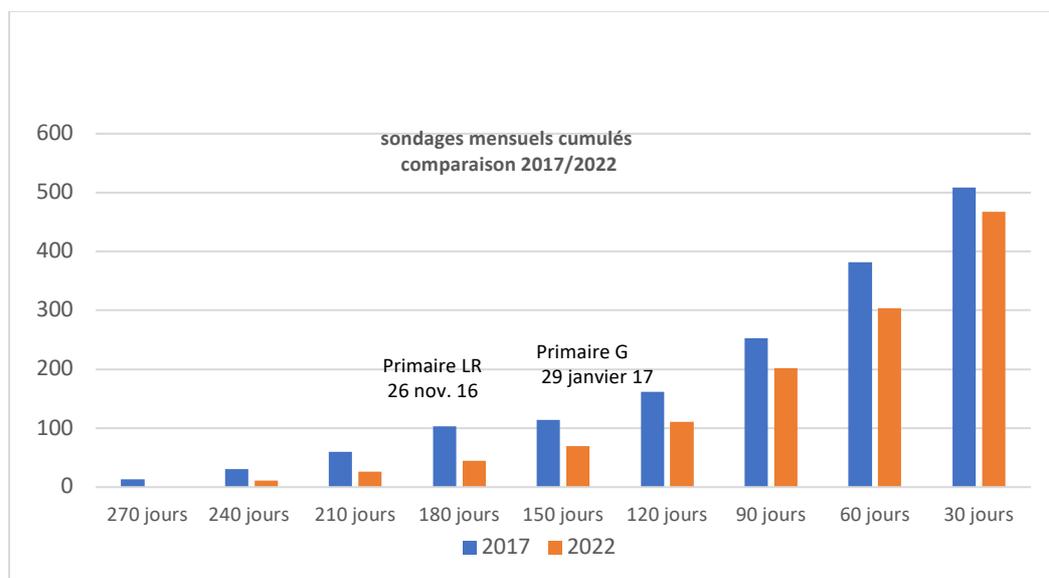
- 3– 1 Le recrutement des panels
 - 3– 1.1 Les sources de recrutement des panélistes
 - 3– 1.2 Les modalités de recrutement
 - 3– 1.3 Contrôle de la motivation des panélistes
 - 3– 1.4 Contrôle de l'ancienneté moyenne des panélistes.
- 3– 2 Le contrôle de la participation aux échantillons
 - 3– 2.1 Fréquences des enquêtes
 - 3– 2.2 Des vérifications complémentaires sur les membres de l'échantillon
- 3– 3 Les défauts des panels

Conclusion

1 – Le contrôle systématique des conditions de publication des sondages

• 1 – 1 Une baisse apparente du nombre de sondages

De 2017 à 2022, le nombre de sondages relatifs à l'élection présidentielle est passé de 560 à 467. La diminution est forte, mais cette constatation doit être nuancée. La centaine de sondages en moins s'explique par le fait que la commission a choisi, en 2007, d'attirer les sondages relatifs aux élections dites « primaires » dans le champ de son contrôle ; or, deux primaires ouvertes ont eu lieu, en 2016 (pour 117 sondages), une à gauche et une droite, alors qu'en vue du scrutin de 2022, les formations de droite comme de gauche s'en sont remises, pour le choix de leur candidat à la présidentielle, au vote de leurs militants, ce qui, eu égard au nombre restreint de ceux-ci, ne permet pas de procéder à des sondages. Il n'en demeure pas moins que l'évolution à la hausse du nombre de sondages pour la présidentielle, très nette depuis 2012, connaît une pause.



Avant que soient désignés les candidats *Les Républicains* ou socialistes, les premiers sondages comprenaient pour chacun de ces camps plusieurs hypothèses en vue de donner à l'instant choisi une mesure du potentiel relatif de chacun de leurs candidats possibles. À neuf mois ou encore à quatre mois du scrutin, à défaut de primaires ouvertes, les comparaisons entre candidats d'un même parti non encore départagés par

le vote des militants se font donc tout naturellement par la comparaison de leur putative performance telle que mesurée par des sondages portant notamment sur le second tour du scrutin. Nous verrons plus avant ce que cela modifie des recommandations de la commission.

Au cours des cent derniers jours précédant le scrutin, on a dénombré environ 300 sondages, soit trois par jour. Notons qu'une part significative des commanditaires de ces sondages appartenait à la presse économique (*Les Échos*, *Challenges*) et que la presse généraliste est apparue moins intéressée par la commande de sondages.

Le champ médiatique était saturé en particulier par trois « rollings » quotidiens (OW, Ipsos, Ifop) pendant les derniers mois de campagne¹.

Sur le nombre de sondages, encore faut-il préciser que, dans un climat lié à la crise du Covid non refermée et à la guerre ouverte en Ukraine, les commentateurs de la vie publique, dans un avis partagé, ont considéré que la campagne électorale était restée assez atone, ce qui, tout à la fois, donnait une certaine importance aux sondages d'intentions de vote et rendait difficile de faire vivre la disposition législative encore récente qui prévoit que la commission doit se saisir de tout sondage portant sur une question d'opinion entrant dans "le débat électoral".

Les **élections législatives** se sont présentées dans un cadre inédit dans la mesure où leur premier tour était organisé sept semaines après l'élection du Président de la République, soit deux semaines plus tard qu'en 2017 ; pour autant, l'atonie de cette seconde campagne électorale, qui devait déboucher, croyait-on alors par une habitude trop vite prise, sur un scrutin de confirmation, a peu intéressé les médias et s'est donc traduite par une rétraction du nombre de sondages : 47 contre 58 en 2017.

• 1 – 2 L'article 3 de la loi de 1977 dans sa rédaction issue de la loi du 25 avril 2016, précise que la première publication ou diffusion d'un sondage doit être accompagnée « d'une mention précisant que tout sondage est affecté de marges d'erreur » et indiquer « les marges d'erreur des résultats publiés ou diffusés, le cas échéant par référence à la méthode aléatoire. » Une telle mention présente deux défauts, théorique et pratique, et une qualité. La qualité est celle de relativiser les estimations des

¹ Un rolling est une série de sondages quotidiens qui réunissent trois échantillons : celui du jour, celui de la veille et celui de l'avant-veille. Ce dernier est quotidiennement abandonné au profit de celui du nouveau jour. Dans la mesure où il existe un nouvel échantillon chaque jour, la commission des sondages compte pour une enquête chaque publication quotidienne.

sondages, ce qui est toujours nécessaire. Un défaut théorique est souligné par la rédaction du texte : « le cas échéant par référence à la méthode aléatoire ». En effet, le propre de la méthode par quotas est de se fonder sur des échantillons dont la constitution ne s'en remet pas intégralement à l'aléa et donc de permettre de se contenter d'échantillons de dimension réduite. Les marges d'erreurs statistiques, qui se calculent aisément pour un sondage par méthode aléatoire, ne sont pas calculables pour des sondages par la méthode des quotas. Un défaut pratique est mis en lumière par l'enquête que le CEVIPOF, le centre de recherche de Sciences-Po, organise avec Ipsos à partir des sondages réalisés selon la méthode des quotas mais en utilisant des échantillons très importants qui sont suivis dans le temps. Au début des enquêtes, ces échantillons comptent plus de 20 000 personnes. Si donc l'on devait en croire les calculs, les marges d'incertitude seraient très faibles, de l'ordre de 0,5 %. Les initiateurs même de ces enquêtes ne revendiquent pas des marges aussi faibles et soulignent le caractère artificiel de ce calcul qui oublie les biais sur lesquels, depuis qu'elle existe, la commission des sondages porte son attention.

L'article 6 de la loi organique n° 2021-335 du 29 mars 2021 portant diverses mesures relatives à l'élection du Président de la République dispose que : *« Pour la prochaine élection du Président de la République suivant la publication de la présente loi organique, toute publication ou diffusion de sondage, au sens de l'article 1er de la loi n° 77-808 du 19 juillet 1977 relative à la publication et à la diffusion de certains sondages d'opinion, est accompagnée des marges d'erreur des résultats publiés ou diffusés, le cas échéant par référence à la méthode aléatoire, établies sous la responsabilité de l'organisme qui l'a réalisé. »*

La commission des sondages s'est efforcée de veiller à l'application de cette nouvelle disposition législative spécifique à l'élection présidentielle de 2022. Elle a alerté les médias sur son existence et demandé aux instituts de sondages de mettre à la disposition des destinataires de leurs sondages les messages appropriés pour assurer le respect de la loi. En effet, ces marges d'erreur sont déjà disponibles, conformément à la loi, dans les notices publiées sur le site internet de la commission et sur celui des instituts. Elles prennent parfois la forme de tableaux difficilement reproductibles à chaque diffusion d'un sondage, en particulier si celui-ci est commenté à la radio ou à la télévision : la marge d'erreur par référence à la méthode aléatoire n'est en effet pas unique pour toutes les intentions de vote d'un sondage, elle dépend notamment du pourcentage d'intentions attribué à chaque candidat. En toute rigueur, il faudrait pour chaque mention d'intention de vote pour un candidat accompagner le pourcentage de l'intervalle d'erreur qui lui correspond. La commission estime que les instituts ont

globalement fait cet effort de pédagogie, mais que tous les médias n'ont pas nécessairement respecté cette obligation avec constance.

- 1 – 3 Les sondages de second tour

La mesure de la performance possible de tout candidat d'un parti face à celui d'un autre parti passe préférentiellement par des sondages sur des hypothèses de second tour, de sorte que l'ancienne règle prudentielle qui interdisait la publication de sondages de second tour avant l'acquisition des résultats du premier tour, a été abandonnée. Mais, en 2017, les recommandations de la commission à ce sujet sont restées strictes :

S'ils ne publient qu'une hypothèse de second tour, les instituts ne doivent publier que celle qui oppose les deux candidats qui arrivent en tête du sondage de "premier tour". Lorsque les résultats établis du premier tour sont suffisamment proches pour que, compte tenu des marges d'incertitude qui les affectent, l'identité des candidats qualifiés pour le second tour soit incertaine, les instituts sont recevables à tester et publier plusieurs hypothèses de second tour. Cette pluralité des hypothèses envisagées est, en effet, de nature à relativiser la portée des résultats de "deuxième tour" publiés et à inciter le public à les interpréter avec toute la prudence nécessaire.

En vue de l'élection de 2022, la commission a été conduite à appliquer cette recommandation avec une certaine souplesse. Comme cela a été déjà noté dans le précédent rapport de la commission (2018-2021), si l'obligation de tester l'hypothèse la plus probable de second tour issue du sondage de premier tour est fortement maintenue, la publication d'estimations de second tour fondées sur des candidats qui ne dépassaient guère les 10 % au premier tour a été acceptée. Cela a été contesté par des observateurs des sondages. Il n'en reste pas moins que deux raisons fortes militent en faveur d'une liberté accrue laissée aux instituts et aux commanditaires pour tester des hypothèses moins probables.

La première s'exprime ainsi : alors que l'offre politique finale est encore loin d'être "cristallisée", de tels sondages de second tour peuvent avoir pour certains partis politiques une fonction de "sondage de primaires". Il s'agit, indépendamment de leur estimation de premier tour, de connaître les performances putatives au second tour de tels ou tels candidats d'un camp face à ceux d'un autre. Certains partis ont même étudié une procédure, finalement abandonnée, de départage de leurs champions, non pas en recourant à des primaires mais en se fondant sur une série de sondages. En tout état de cause, les partis comme les commentateurs ont besoin d'une mesure de l'efficacité au

second tour « du troisième homme », voire de tout candidat nouveau. Aucune raison impérative ne s'oppose à ce que les acteurs de la vie politique bénéficient de telles indications, et renvoyer de tels sondages à la clandestinité n'apparaît pas comme un progrès du contrôle, ni comme un gage de sérénité de la campagne.

La seconde raison de cette souplesse est consubstantielle à la réserve qu'il convient d'avoir quant au caractère prédictif des sondages, particulièrement lorsqu'ils sont administrés longtemps avant l'échéance électorale. Des candidats qui, à neuf mois d'une élection, ne sont pas « donnés d'avance » au second tour par les sondages peuvent en pratique y parvenir. L'exemple de l'actuel Président de la République et des estimations de vote dans les sondages un an avant sa première élection en sont une bonne illustration.

• 1 – 4 La présence des médias étrangers le jour du vote

Certes, on peut mesurer l'importance d'un pays à l'intérêt qu'il suscite à l'étranger ; néanmoins, en matière de sondages, certaines prévenances s'avèrent parfois indélicates. Des médias belges, en effet, semblent prêts à bien des approximations pour obtenir un surcroît d'audience sur internet venant de France et les recettes publicitaires afférentes un dimanche tous les cinq ans. Cette année, une sorte de palme peut être décernée à *La Libre Belgique*. Jusqu'alors *La Libre* se contentait de publier des « fuites » venant d'instituts de sondages français, sans autres précisions. Et cela suffisait à provoquer, au cours de l'après-midi du dimanche électoral, un grand nombre de connexions sur son site et, ce faisant, un grand bonheur de sa régie publicitaire.

Cette année le média belge a monté d'un cran dans le sens de l'imagination en précisant dès 17 heures 30 : « *Selon un premier sondage sortie des urnes (soit réalisé auprès d'électeurs sortant des bureaux de vote), Emmanuel Macron et Marine Le Pen seraient en tête à égalité. Ils récolteraient tous deux 24 % des votes, a appris La Libre à bonne source... il s'agit là de résultats issus d'un premier sondage sortie des urnes, réalisé par l'un des grands instituts. Il faut donc être très prudent avec ces résultats, qui ne sont qu'indicatifs et qui peuvent encore fortement évoluer. Précisons d'ailleurs que ces résultats sont des chiffres bruts non corrigés (sic). Plusieurs sources bien informées nous signalent qu'il faudra attendre 19 heures 45 pour départager les deux premiers candidats, tant l'écart est faible.* ».

Interrogé par *Libération* quelques jours plus tard, M. Dorian de Meeûs, le rédacteur en chef de *la Libre Belgique*, change sa version et précise avoir eu « *les chiffres de trois instituts différents* ». Et d'affirmer que « *si nous n'avions pas la fiche technique*

des études, il s'agissait bien d'enquêtes « sortie des urnes », portant sur plusieurs milliers de personnes. »

L'appel à la prudence du rédacteur en chef semble constituer un hommage du vice à la vertu, dans la mesure où *La Libre Belgique*, qui se bornait habituellement à reprendre avec un certain bonheur la moyenne des sondages publiés le dernier vendredi de la campagne, a souhaité, cette fois, choisir une voie plus risquée, celle de la prolongation des tendances tracées par les dernières estimations alors qu'il eût été plus avisé de choisir un léger retournement de celles-ci.

Quoi qu'il en soit, aucun des instituts n'a réalisé de sondages "sortie des urnes", pour la raison tirée de l'imprécision de ces derniers. Sans doute agacé par la méthode, l'électeur qui sort de l'isoloir rechigne à voter une seconde fois et quand il le faisait par le passé, jusqu'en 2002, c'était avec un singulier mauvais souvenir de vote. Le prétendu recours aux sondages "sortie des urnes" pour établir les estimations de 20 heures est donc un pur fantasme.

La réalité est que les instituts qui présentent des tendances dès 20 heures dans les principaux médias utilisent deux sources principales : leurs sondages du vendredi et des estimations tirées de l'analyse de bureaux tests dont on connaît l'écart à la moyenne nationale mesuré lors de précédents scrutins et qui ferment dès 19 heures. Ainsi les instituts, et seulement leurs clients, et peut-être ceux des commentateurs qui ont été informés par ces derniers, connaissent-ils vers 19 h 45 les tendances du dépouillement partiel.

Par ailleurs, dans la journée du samedi et parfois aussi du dimanche, des instituts procèdent à des sondages pour publier en fin de soirée des analyses du vote par sexe, par âge et pour toutes sortes de catégories géographiques ou socioprofessionnelles : qui a voté quoi ? Ces derniers sondages font l'objet d'embargos stricts. Ils ne peuvent en rien servir à l'établissement d'estimations et certainement pas à la fameuse rumeur belge de cette année dite "24-24-19" avec dans l'ordre, Mme le Pen, M. Macron, M. Mélenchon. Rappelons pour relativiser l'intérêt de ces « fuites » du dimanche après-midi que les estimations à 20h étaient assez éloignées de ces chiffres, elles donnaient 28,6% à M. Macron (résultat officiel et définitif de l'élection : 27,85%), 24,4% pour Mme Le Pen (résultat officiel 23,16%) et 20,2% pour M. Mélenchon (résultat officiel 21,35%).

Une nouvelle fois, un certain appât du gain a incité la presse étrangère à agir de façon déraisonnable. La commission des sondages note que c'est avec raison que la France se dote à chaque élection de dispositifs de veille pour éviter l'influence délétère qu'une puissance étrangère hostile pourrait exercer sur nos élections, mais c'est avec amertume que l'on note que certains médias de pays voisins interviennent de façon répétée, en contrariété avec l'un des principes majeurs des élections : le calme, le silence dans lequel le corps électoral doit forger sa volonté.

• 1 – 5 Les publications la veille et le jour du scrutin dans les médias français

Le calme dans lequel on souhaite, en France que le corps électoral forge son sentiment la veille et le jour du scrutin fonde deux règles principales : la fin de la campagne le samedi qui précède le scrutin à zéro heure et l'article 11 de la loi de 1977 qui précise que : « En cas d'élections générales et de référendum et le jour de chaque scrutin, aucun sondage électoral ne peut faire l'objet, par quelque moyen que ce soit d'une publication, d'une diffusion ou d'un commentaire. (...) Cette interdiction ne fait obstacle ni à la poursuite de la diffusion de sondages publiés avant la veille de chaque scrutin ni au commentaire de ces sondages (...). »

L'article 11 contient donc un principe fort assorti d'une exception claire. Toutefois, cette année, la possibilité de seconde publication ou de reprise a donné lieu dans plusieurs grands médias, (*Le Figaro* la veille du premier tour des présidentielles comme des législatives, le *Journal du Dimanche* et *Le Parisien* le jour même des scrutins) à d'amples doubles pages de rétrospectives de sondages dans lesquelles, en somme, les dernières estimations des instituts prenaient, le jour même de l'abstinance voulue par le législateur, davantage de place que la veille, au moment même où de telles publications étaient pleinement licites.

On peut comprendre que le législateur n'ait pas voulu qu'un média ayant publié un sondage dans une édition en ligne dans la période autorisée soit obligé de rendre les pages internet correspondantes inaccessibles pendant la période d'interdiction pour respecter la lettre de l'ancienne version du texte. Mais on peut se demander si les pratiques décrites au paragraphe précédent, si elles respectent la lettre de la loi, sont bien conformes à son esprit. Aussi appartiendra-t-il au législateur de dire si une telle situation lui convient ou s'il serait préférable de revenir au principe d'abstinance sondagière si près du scrutin en interdisant les pages de reprise ou de synthèse des sondages antérieurs. Les journaux du dimanche en souffriraient spécialement sans doute de même

que les éditions du samedi de certains quotidiens, mais le silence avant le vote tel que souhaité au premier alinéa de l'article 11 ne mérite-t-il pas un effort de la part des médias et de leurs lecteurs ?

- 1 – 6 Publication des estimations en sièges.

Les sondages qui entrent dans le champ de contrôle de la commission sont définis à l'article 1^{er} de la loi du 19 juillet 1977. La commission a ajouté à cette liste les sondages relatifs aux primaires. En revanche, la commission n'a jamais contrôlé les estimations qui, après un sondage relatif aux élections législatives mesurant les intentions de vote au niveau national, tentent d'établir des projections en sièges de la nouvelle assemblée. Ce type de sondage national couplé à une projection en sièges est différent d'un sondage d'intentions de vote pour une circonscription déterminée, qui appelle le contrôle habituel de la commission.

Ainsi certaines études législatives sont-elles contrôlées pour partie seulement par la commission et le caractère partiel de ce contrôle est parfois critiqué ; il reste fondé aux yeux de la commission.

En effet, si les sondages bénéficient d'un double système de référence, théorique d'une part et tiré de près de 10 000 sondages expertisés d'autre part, les projections en sièges, pour sophistiquées qu'elles soient parfois, restent insaisissables au contrôle. En effet, la commission manque de références scientifiques, de labélisation ISO et surtout d'expériences documentaires pour déterminer avec certitude ce qui correspond aux règles de l'art. Cette étape de projection en sièges ne constitue pas au demeurant à proprement parler un sondage. Par ailleurs, de telles estimations sont très sensibles aux évolutions des estimations de vote. Un mouvement de toute fin de campagne qui affecte les estimations en voix de 2 % ou 3 %, soit une proportion inférieure à la marge d'incertitude, peut provoquer des évolutions en sièges considérables.

Ce défaut de contrôle par la commission des estimations en sièges incite certains à en prôner l'interdiction ; ce n'est pas l'avis de la commission qui estime que généralement les publications de ces enquêtes ont su faire le départ entre la partie « sondages » proprement dite et la partie « estimation en sièges », de sorte que le lecteur de bonne foi et normalement attentif peut aisément distinguer les deux exercices.

- 1 – 7 Publication d'estimations locales tirées d'estimations nationales

De la même façon que l'on peut estimer un résultat en sièges en se fondant sur un sondage national, on peut anticiper sur la même base le résultat d'une circonscription déterminée. Mais, pour les mêmes raisons, la publication de telles données chiffrées est sans rapport avec celles d'un sondage. Ces publications qui ont été faites sur les réseaux sociaux par certains candidats ou leurs soutiens ne sont pas interdites, mais elles ne sont pas pour autant des « sondages » au sens de la loi et la commission est intervenue chaque fois qu'elle a pu dans l'heure où elle a eu connaissance de ces errements pour préciser que l'attribution du terme « sondage » à de tels exercices ne correspondait pas à la définition législative du mot « sondage ». (2^{ème} de Loir-et-Cher, 4^{ème} de Seine Maritime, 1^{ère} du Territoire de Belfort, 5^{ème} du Pas-de-Calais.)

- 1 – 8 Publications diverses

Pour qu'on ne puisse relever une utilisation abusive du terme “sondage”, il ne suffit pas que le cœur de la publication précise d'où provient l'échantillon des personnes ayant répondu, par exemple du lectorat du média, mais il faut aussi que le terme de « sondage » ne figure pas dans l'article publié. (Communiqué du 30 mai : Les Français.press). De même, donnent lieu à communiqué de la commission les initiatives de candidats qui, par voie de tract, ont posé des questions à la population et qui se vantent, en fin de campagne, d'estimations défavorables à leur adversaire. Dans un tel cas, il ne suffit pas que le mot « sondage » n'ait pas été utilisé, mais encore que les données récoltées ne soient pas présentées comme le seraient les estimations d'un sondage (cf. communiqué du 26 mai 2022 : 7^{ème} circonscription du Rhône)

- 1 – 9 Publications de la commission

L'article 3 de la loi prévoit que la commission doit publier sur son « service de communication au public en ligne », pour chaque sondage publié, les notices dont il précise le contenu. Cette obligation trouve sa raison d'être dans la volonté du législateur de rendre plus transparente l'activité des instituts de sondage.

Pourtant, par la publication même de leurs notices sur le site de la commission, certains instituts souhaitent obtenir ce qu'ils considèrent comme une sorte de labélisation ou du moins de certificat de qualification dans ce qui serait à leurs yeux le premier cercle des instituts français et ils le font savoir.

Dès lors, la commission se trouve en porte à faux avec une obligation de publier d'un côté et une retenue à le faire d'un autre pour des études qui ne sont pas du tout ou pas tout à fait au standard des études de qualité. Au cours des campagnes de sondages de 2022, la commission a été confrontée aux situations les plus diverses.

La commission estime qu'il serait préférable à l'avenir de ne pas publier les notices de sondages sur son site dans deux hypothèses.

En premier lieu, dans le cas d'instituts qui n'ont pas suivi la procédure de déclaration préalable et qui ont transmis à la commission des éléments lacunaires ne permettant pas aux experts de la commission un contrôle complet des études publiées.

En second lieu, dans l'hypothèse d'instituts qui, ayant fait l'objet d'une mise au point pour un défaut général de méthode ou pour une absence de culture de contrôle interne et dont les productions ultérieures à la mise au point de principe n'ont pas encore apporté la preuve solide d'un amendement de leur méthode et qui restent donc "couvertes" par la mise au point initiale qui n'a pas encore épuisé ses effets.

Dans ces deux cas la publication des notices sur le site de la commission est davantage de nature à alimenter la confusion qu'à clarifier les choses.

2 – L'examen systématique de la fiabilité de tous les sondages relevant de son champ de compétences

• 2 – 1 Les moyens de la commission

À son apogée, qui est aussi son terme, la campagne électorale en vue de l'élection présidentielle a compté plus d'une vingtaine de sondages par semaine.

Il faut ici féliciter les deux experts de la commission qui s'organisent par un travail constant afin d'éviter tout retard et assurer promptement un contrôle exhaustif.

Deux conditions expliquent cette réussite. D'abord, la présence de rollings qui, par eux-mêmes, simplifient un peu la réalisation du contrôle et ensuite et surtout la confiance qui marque généralement les rapports entre les instituts et la commission des sondages². En effet, le contrôle de la commission repose d'abord sur les déclarations des instituts. Cela signifie que si la commission réussit à obtenir des instituts l'intégralité des documents préalables à l'élaboration des estimations publiées et de refaire étape par étape tout le travail qui a conduit auxdites estimations, sa présidente, son secrétaire général et ses experts travaillent avec les instituts dans un climat de confiance qui n'oblige pas de refaire à chaque sondage un examen exhaustif.

Cette campagne a été marquée par une crise de confiance majeure avec un institut dont on a suspecté, avec quelque raison, qu'il utilisait des outils informatiques inadaptés qu'il ne maîtrisait pas, ce qui a conduit la commission à demander l'intégralité des fichiers de base ; mais, en général, à propos de bien des éléments, nous croyons les instituts capables d'un autocontrôle serré.

Dès lors, on peut s'interroger : le règne de la confiance constitue-t-elle une barrière à l'apparition de nouveaux instituts ? Afin de pallier ce risque, le secrétariat de la commission a mis en place un dispositif fondé sur l'article 7³ de la loi du 19 juillet 1977 qui impose aux nouveaux entrants une déclaration préalable. Dès qu'un institut déclare sa volonté d'effectuer des sondages électoraux, le secrétariat transmet au demandeur une liste des documents nécessaires aux experts. Si elle n'est pas exhaustive, cette liste est cependant suffisamment complète. Elle est établie de sorte que chacun puisse comprendre en quoi consiste le contrôle de la commission et, ce faisant, quelles sont les spécificités d'un sondage électoral. Le travail de la commission est aussi, autant que faire se peut, d'éviter, par un échange en amont, des difficultés en aval.

De nouveaux instituts, dont de grandes sociétés étrangères, rechignent ou négligent ce moment d'échange qui leur est ouvert, ce qui a pu entraîner des mises au point dues à une mauvaise compréhension des choses. En sens inverse, un exemple venu d'un petit institut grenoblois créé par un professeur d'université, nous a permis de

² Sur la notion de confiance opposée à la notion de connivence cf. le rapport annuel de la commission 2007

³ **Article 7** modifié par Loi n° 2017-55 du 20 janvier 2017 - art.24 : « Nul ne peut réaliser des sondages, tels que définis à l'article 1er et destinés à être publiés ou diffusés, s'il ne s'est engagé, par une déclaration préalablement adressée à la commission des sondages, à appliquer les dispositions de la présente loi et les textes réglementaires applicables. Nul ne peut publier ou diffuser les résultats d'un sondage, tel que défini à l'article 1er, s'il a été réalisé sans que la déclaration prévue à l'alinéa qui précède n'ait été préalablement souscrite. »

constater que le dispositif mis en place permettait à une structure qui s'apparente plus à une start-up qu'à un groupe multinational d'être dans une juste intelligence des exigences de l'art, de s'assurer d'un contrôle interne serré, d'utiliser un outil éprouvé, de recourir à une panélisation sérieuse, de respecter des règles de la commission en matière de représentativité et de redressement, d'échanger avec franchise avec les services de la commission et ainsi d'obtenir d'un seul coup la confiance nécessaire.

Cette réussite est l'occasion de préciser pour l'avenir à tout institut qui souhaiterait investir le champ des sondages électoraux qu'il convient de ne pas traiter comme une simple démarche administrative cette déclaration préalable auprès de la commission.

• 2 – 2 Les règles de contrôle statistique

Les campagnes de sondages de 2022, tant présidentielle que législative, n'ont fait l'objet d'aucune difficulté de méthode avec les principaux instituts. La règle stricte d'une colonne de référence unique, c'est-à-dire d'un même type de redressement maintenu de sondage en sondage tout au long de la campagne, assortie de la possibilité de déplacer, une seule fois, le curseur de la certitude du choix, n'est plus contestée dans son principe. Elle a simplement donné l'occasion d'un débat avec un institut qui, à l'approche du premier tour, a souhaité modifier une seconde fois son option d'intensité de la certitude d'aller voter. Une telle demande a fait l'objet d'une réunion qui s'est tenue entre les responsables de l'institut, la présidente, le secrétaire général et les experts de la commission, au terme de laquelle l'institut a décidé de s'en tenir à la règle stricte d'un seul changement de référence.

Ce rapport étant le dernier de l'actuelle commission des sondages, il n'est pas inutile, avant d'aborder les problèmes et réclamation de 2022, de revenir sur une quinzaine d'années de contrôle statistique.

2 – 2.1 L'objectivité et la sincérité

La commission est chargée de l'objectivité et de la qualité des sondages publiés, ce qui recouvre bien des aspects techniques : le choix d'une base de sondages, l'élaboration d'un plan de sondage, la neutralité des questions posées et l'ordre de celles-ci. Parmi ces aspects, figurent les règles de redressement statistique.

En 2007, les sondages en vue d'élections nationales, qu'ils portent sur une seule circonscription (présidentielle) ou sur 577 (législatives), étaient effectués par téléphone, conduisant à une triple difficulté :

- Un “sinistrisme des échantillons”. Les personnes interrogées déclaraient une intention de vote plus à gauche que leur souvenir de vote qui lui aussi se situait plus à gauche que les résultats effectifs.

- Une mauvaise représentativité de tous les échantillons marqués par un vote d’extrême droite profondément sous-évalué, le comblement entre le vote passé, par exemple celui de M. Jean-Marie Le Pen au premier tour de 2002 (18 %) et le nombre de personne se souvenant d’avoir voté Le Pen à cette même élection (circa 6 %), conduisait à des coefficients de pondération qui ont pu dépasser 3.

- des artéfacts que les sondeurs appelaient « accidents d’échantillon » qui pouvaient donner lieu dans des séries hebdomadaires à des évolutions fortes non justifiées par la vie de la campagne et qu’il convenait d’amortir.

Le cumul de ces difficultés explique que la commission autorisait alors les instituts à choisir, pour chaque candidat, un redressement parmi un grand nombre d’hypothèses. En effet, après le redressement socio-démographique, les redressements s’opéraient, a minima, à partir d’une élection de référence, en général entre la dernière élection générale (européenne) ou la dernière élection présidentielle ou encore une combinaison des deux, soit trois hypothèses, puis pour chacune de ces hypothèses en fonction de la certitude d’aller voter (6 hypothèses) puis de la sûreté du choix (12 hypothèses), etc. Au maximum, certains instituts ont parfois présenté des tableaux à 48 colonnes, soit 48 hypothèses ! Ces colonnes multiples donnaient pour chaque candidat une « fourchette » de résultats entre lesquels les instituts étaient libres de fixer l’estimation la plus haute et la plus basse. Le tableau résumé suivant donne une idée des écarts de fourchettes des principaux candidats en février 2007 pour un sondage téléphonique (*sondage 6180*).

	mini	maxi
Ségolène Royal	24,2	30,8
François Bayrou	10,8	18
Nicolas Sarkozy	30,5	36,1
Jean-Marie Le Pen	8,8	14,2

Au surplus, pour amortir les artéfacts, les instituts souhaitaient parfois sortir de la fourchette⁴. Pour toutes ces raisons, le temps des sondages nationaux par téléphone,

⁴ Sur les progrès du contrôle statistique, lire avec un intérêt historique les développements sur les « sorties de fourchette dans le rapport annuel 2007 de la commission. »

que la doctrine⁵ encense parfois, pouvait apparaître comme une période où l'enquête statistique ne limitait pas de façon déterminante la fixation des estimations. Dans une telle situation, où, *in fine*, c'est à partir des considérations politologiques qu'étaient fixées les estimations, le contrôle proprement statistique des sondages d'intentions de vote était alors réduit.

Le contraste est fort avec la situation actuelle. Aujourd'hui, les instituts doivent se limiter à une seule colonne de référence, c'est-à-dire à un seul mode de redressement tout au long de la campagne de sondages, moyennant la possibilité de varier, une fois, l'intensité de la certitude d'aller voter ou de la sûreté du choix.

Ceci s'explique par le fait que nous sommes passés aux sondages par internet, fondés sur une base solide donnée par "panélisation". Ce passage a été généralisé après le cycle électoral de 2012.⁶ Parallèlement, la réserve à déclarer un souvenir de vote d'extrême droite s'est estompée, ce qui a également facilité le retour des sondages dans un espace statistique plus conforme aux principes.

Ainsi, la commission, sur une base stabilisée, a-t-elle pu imposer aux instituts d'abord de s'en tenir aux limites de la fourchette, puis de se fonder sur une seule colonne de référence, enfin, d'exclure toute possibilité d'exception à la colonne de référence. Longtemps tenus dans les limbes de la statistique, les sondages en lien avec le débat électoral se sont progressivement pliés à ses lois.

En d'autres termes, les règles du contrôle et le niveau d'exigence ne sont pas seulement fonction de la volonté de l'organe de contrôle – celle-ci n'a jamais manqué – mais des modalités pratiques d'élaboration des sondages.

Pour autant, la commission des sondages, toujours soucieuse de laisser ouvert le champ des possibles, n'a jamais rendu obligatoire ni le recours à internet ni la

⁵ Une autre partie de la doctrine, qui souvent est la même, porte également aux nues le sondage aléatoire.

⁶ Même si l'efficacité prédictive des ultimes sondages du vendredi avant le scrutin n'est pas pour la commission un critère déterminant ni même majeur, chacun a pu constater que trois instituts travaillaient par internet et trois par téléphone. La différence de performance fut telle entre ces sous-groupes que le téléphone s'en trouva condamné. L'ensemble des deux premiers candidats de gauche (Mélenchon, Hollande) a recueilli 39,7 % des voix ; il était estimé à 40,2 % par la moyenne des estimations des instituts internet et à 42,5 par la moyenne des instituts par téléphone. De l'autre côté de l'échiquier politique, l'ensemble Le Pen, Sarkozy a recueilli 45,1 % des voix ; il a été estimé à 43 % par internet et à 41,5 % des voix par téléphone.

panélisation⁷. Cependant, actuellement sur une base de près de 10 000 sondages dont près de 2 000 administrés par internet, la commission note que si la méthode qui combine l'internet et la panélisation est parfois présentée comme une technique *low cost*, elle permet, pour l'heure encore, comme aucune autre technique, tout à la fois d'éviter les artéfacts en dessinant des courbes d'opinion sans à-coups inexplicables au cours d'une campagne électorale et d'aboutir à des performances prédictives des ultimes sondages plus que satisfaisantes, les écarts aux résultats du vote étant inférieurs aux marges d'incertitude. Cette dernière performance est toutefois réservée aux élections dont le taux de participation avoisine 70 % des votants. Une abstention massive reste un écueil qui ne peut être évité par la sûreté d'un échantillon de bonne qualité.

2 – 2.2 La réclamation de « La France insoumise »

Des représentants de « La France Insoumise » ont souhaité en début de campagne être entendus par la Commission des sondages. Ils l'ont été au cours d'une réunion du 14 septembre 2021.

Devant la commission, M. Manuel Bompard a développé le raisonnement suivant.

La démocratie n'existe que si elle se donne les moyens de libérer l'acte de vote des influences extérieures. Or, les résultats des dernières élections régionales ont très largement contredit les estimations des sondages et ces « erreurs » ont eu un impact évident sur le déroulement de la campagne.

Dès lors, pour celui qui était alors député européen, certains instituts devaient être plus transparents sur le « niveau estimé » de participation de leurs enquêtes qui devraient être toujours affichées.

Inversement, les instituts qui rendaient publiques leurs méthodes de calcul révélaient, ce faisant, une pratique trop restrictive⁸.

En retenant une plus large participation, les sondeurs donneraient voix à des populations qui votent en effet moins mais qui pourraient faire varier les estimations.

⁷ Notez que le redressement par souvenir de vote n'a jamais été rendu obligatoire. Il existe d'autres façons de qualifier politiquement un échantillon. Au surplus, la base de données de la commission révèle que, dans certains départements où la fluidité des étiquettes politiques est forte, le redressement par souvenir de vote est loin d'être utile.

⁸ Les instituts demandent aux personnes interrogées qui se disent « sûres d'aller voter » de noter leur propre certitude de 1 à 10 et ne retiennent que ceux qui se sont attribuées la note 10 ou les notes 9 et 10.

Le débat qui succéda à ces interrogations fut de qualité. Les éléments portés en réponse à la réclamation furent principalement les suivants. Le législateur est revenu deux fois déjà, en 2002, puis en 2016, pour préciser le contenu de l'article 3 de la loi n° 77-808 relative à la publication et à la diffusion de certains sondages d'opinion. En cette matière, qui touche au secret de fabrication et donc au secret des affaires, il faut relever que certains instituts ont délibérément choisi de donner une information plus large alors que d'autres sont plus discrets.

Cependant, il n'appartient pas à la commission des sondages d'imposer de son propre chef la publicité d'informations s'ajoutant à celles dont la liste a été élaborée par le législateur. La commission s'efforce de faire en sorte que la loi soit respectée ; elle ne peut, en la matière, demander aux instituts d'aller au-delà des exigences législatives. L'avenir nous dira si l'actuel législateur souhaitera revenir sur la liste des documents figurant à l'article 3 de la loi ou s'il précisera que, par principe de prudence, il ne convient pas, à neuf mois d'une élection, de présenter une estimation de participation.

Par ailleurs, 9 ou 6 mois avant une élection, le problème des sondeurs n'est pas de publier un « taux de participation » estimé, mais d'opter pour une colonne de référence, un redressement de référence qui devra être tenu de sondage en sondage, tout au long de la campagne⁹. Or ce problème se pose de la façon suivante : quel que soit le mode d'interrogation, les personnes interrogées par sondage sont davantage promptes à répondre qu'à réellement se rendre aux urnes le jour du scrutin. C'est une constante, indépendante du mode d'interrogation. Dès lors, il faut trouver les moyens de concentrer l'exercice sur celles des personnes ayant répondu et dont on pourrait penser qu'elles présentent le plus de garanties effectives d'aller voter. Quand, par exemple, un institut estime devoir ne retenir dans ses estimations que ceux des répondants qui, à distance du scrutin, notent leur certitude d'aller voter 10/10, le résultat qui importe à l'issue de cette sélection n'est pas un taux de participation apparent, mais la base de calcul fixe qui servira tout au long de la campagne aux estimations de vote d'un institut donné. La nécessaire stabilité des méthodes tenue de vague en vague est la seule garantie offerte aux lecteurs attentifs de sondages leur permettant de lire une analyse dynamique des évolutions dans le temps qui importe bien plus qu'un seul niveau estimé à un instant déterminé des différents candidats.

La commission se félicite en tout cas d'échanges documentés avec des représentants de formations politiques, qui poussent le débat et donc la réflexion sur les

⁹ Modulo (cf. infra) la possibilité unique à la fin de la campagne de jouer sur la focale afin de retenir une population plus étroite ou éventuellement plus large.

sondages plus avant. De tels échanges offrent une compréhension plus fine des souhaits des acteurs de la vie publique, une meilleure connaissance du rôle de la commission, des limites de son action, voire une orientation d'éventuels débats législatifs.

2 – 2.3 Contentieux Cluster 17

À l'automne 2021, un nouvel institut, dénommé Cluster 17, a commencé à publier sur son site des estimations relatives aux élections présidentielles. Les difficultés de compréhension furent immédiates avec cet institut. Faute de réponses claires, la commission invita les représentants de Cluster 17 à répondre oralement devant elle le 24 novembre 2022.

La réunion a d'abord porté sur la bonne compréhension de la méthode Cluster 17 qui repose sur un ensemble disparate.

- D'une part, un fichier acheté de 6,5 millions d'adresses auprès des sociétés *Fichiers-emails* dont les fichiers, selon les informations disponibles, ne comportent que des adresses de particuliers âgés de 20 à 60 ans et de *Base-emails.com* qui ne traite que des professionnels, donc pauvre en retraités et personnes âgées de 65 ans et plus.

- D'autre part, un panel qualifié de 100 000 personnes qui ont accepté d'être recontactées à la suite d'une première réponse à un sondage de l'institut. Cette seconde base est mobilisée à chaque sondage selon une proportion variable, qui tourne autour de dix pour cent.

Dans un second temps, la réunion a permis de faire le point sur les doublons qui affectaient l'échantillon de Cluster 17 pour une vague donnée. Plus de 500 réponses étaient « doublonnées », c'est-à-dire que des réponses conservées dans l'échantillon du sondage étaient identiques sur un grand nombre de variables (âge, sexe, souvenir de vote, région de résidence etc...), qui ne correspondaient pas à des personnes sondées distinctes. L'institut qui n'avait manifestement mis en place aucune mesure de contrôle de la qualité de ses opérations de recueil, a semblé découvrir l'existence de cette anomalie.

D'une façon étrange, ces doublons affectaient prioritairement, et donc amplifiaient, la catégorie des plus de 65 ans, celle-là même qui faisait défaut aux fichiers utilisés pour l'administration du sondage.

On pouvait être alors enclin à acquiescer à la conviction, sans disposer de moyen de le prouver, que nous avions affaire pour la vague expertisée à un « sondage guidé », ce guidage ayant pour fonction de corriger les déficiences de l'échantillon initial. En d'autres termes, pour cette vague, on pouvait suspecter que le choix des personnes faisant partie de

l'échantillon et le nombre de réponses données par chacune d'elles étaient déterminés par les données que l'on souhaitait publier.

À l'issue de la réunion, Cluster 17 s'engagea à retirer ses études de son site internet et de travailler différemment à l'avenir.

Cela ne s'est pas déroulé ainsi, malgré de très nombreux échanges téléphoniques et par mails. Face aux imprécisions, aux contradictions et aux doutes qui s'accumulaient, la commission a convoqué les représentants de Cluster 17 une seconde fois le 1^{er} février 2022 dans l'éventualité d'une mise au point. Cet entretien était précédé d'un long questionnaire écrit.

L'insuffisance des réponses apportées par écrit et oralement par C17 et le manque de contrôle de l'institut conduisit la commission à publier à l'encontre de Marianne, magazine qui avait publié les derniers sondages de Cluster 17 une mise au point dans les termes suivants.

Mise au point de la commission des sondages du 4 février 2022

Après avoir entendu les représentants de la structure dénommée « Cluster 17 » au cours de ses séances des 24 novembre 2021 et 1^{er} février 2022, 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 magazine papier de « Marianne », partenaire média de Cluster 17 ainsi que sur le site en ligne du magazine.

Les méthodes de travail de Cluster 17 et la faiblesse des moyens mis en œuvre font douter de la qualité des sondages produits. Est particulièrement en cause la représentativité d'échantillons provenant non d'un panel constamment entretenu et renouvelé de répondants qualifiés mais d'un stock d'adresses électroniques acquises à bas coût, sans aucune garantie sur l'origine de ces adresses, sur la qualification des personnes qu'elles représentent et sans aucun contrôle de Cluster 17 sur cette base. Par ailleurs, le mode de questionnement des titulaires de ces adresses, invités expressément à répondre à une enquête en vue de l'élection présidentielle, est susceptible de créer des biais dans la constitution des échantillons alors qu'aucune procédure de contrôle n'est prévue pour détecter ces biais et les corriger.

En raison de ces insuffisances méthodologiques et de la faiblesse des contrôles dans la réalisation des opérations de recueil, la commission des sondages invite les lecteurs des sondages réalisés par Cluster 17 à une prudence particulière dans l'appréciation de leur pertinence.

Cette décision a fait l'objet d'un recours pour excès de pouvoir devant le Conseil d'État.

Par une décision du 20 décembre 2022, celui-ci a entièrement confirmé la position de la commission des sondages. D'une façon générale, le Conseil d'État confirme sa

jurisprudence constante (notamment une décision Société Indice du 13 décembre 1985), selon laquelle il appartient à la commission des sondages, chaque fois que la qualité ou l'objectivité d'un sondage lui paraît en question, de demander la publication d'une mise au point appropriée, quels que soient les résultats du sondage et cela même lorsqu'elle constate que le sens général du résultat ne lui apparaît pas susceptible d'être mis en cause. Il précise que la commission des sondages est fondée à ordonner la publication d'une mise au point, dans des termes appropriés, dès lors que, au vu des éléments d'information dont elle dispose à la date de sa décision et en tenant compte, le cas échéant, de l'urgence qui s'attache à son intervention eu égard en particulier à la proximité du scrutin auquel un sondage se rapporte, elle n'est pas en mesure de s'assurer de la conformité d'un tel sondage aux exigences de la loi du 19 juillet 1977 et de ses décrets d'application.

Ce point est important pour conforter l'office de la commission : les sondages électoraux sont réalisés sous la responsabilité des instituts, c'est à eux d'être en mesure de justifier du sérieux de leur méthode, la commission se prononçant au vu des éléments qui lui sont fournis en réponse aux interrogations qu'elle formule et compte tenu de l'effet utile de son contrôle : une mise au point intervenant à une date trop éloignée des faits de nature à la justifier perdant tout intérêt pratique pour ses destinataires, c'est-à-dire les futurs électeurs.

S'agissant du cas spécifique de Cluster 17 et de la mise au point attaquée, le Conseil d'Etat a rejeté l'argumentation de la société selon laquelle elle aurait recours à une méthode aléatoire conforme à la théorie des sondages, méthode qui la dispenserait de s'assurer du caractère représentatif de ses échantillons. Il a relevé que la base « aléatoire » de l'institut « *ne saurait être assimilé à la mise en œuvre de la « méthode aléatoire » de sondage consistant à tirer au sort les personnes sondées parmi une base contenant la totalité, ou une partie très substantielle, de la population pertinente, qui comptait, en l'espèce, plus de 48 millions d'électeurs* ». Il juge surtout que le mode de constitution de cette base « aléatoire », « *eu égard aux conditions dans lesquelles ces adresses ont été acquises et à l'absence de dispositifs de vérification* », ne permet pas de prévenir les biais susceptibles d'affecter cette base, dont la société ignore comment les adresses emails ont été obtenues. La haute juridiction a confirmé que la circonstance que des contrôles aient été mis en place pour répondre aux

demandes de la commission à la suite de la première audition n'est pas de nature à corriger les biais pouvant exister en amont, touchant à la base elle-même. Le Conseil d'État écarte enfin l'argument tiré de ce que les répondants ne bénéficient d'aucune rémunération ou récompense en contrepartie de leur participation au sondage, dès lors que cette circonstance n'est pas « de nature à prévenir les biais susceptibles d'affecter la base « aléatoire » ».

3 – L'examen quasi exhaustif du travail des panélistes

La Commission des sondages s'est intéressée cette année au travail des sociétés d'access-panels, qu'elles soient liées structurellement aux instituts de sondages ou sous-traitantes.

Nous avons vu ci-dessus que la base documentaire de la commission portant sur près de 10 000 sondages révèle de façon empirique que seuls les sondages réalisés par interrogation en ligne et reposant sur des panels offraient des garanties d'absence d'artéfact et permettaient de fonder un véritable contrôle statistique.

Cette réalité est loin de suffire au rassemblement d'un consensus. Des chercheurs en sciences sociales contestent cette réalité empirique, préférant, par exemple, s'adosser au mythe de la pureté des sondages strictement aléatoires ou porter aux nues les sondages administrés par téléphone. Sans doute ont-ils raison s'agissant de l'avantage théorique de véritables sondages aléatoires par rapport à des sondages par quota, mais on ne peut aisément les suivre dans la mesure où la base de données de la commission ne contient aucune étude qui s'approche réellement du prétendu modèle pur et parfait. À l'opposé, ce furent des résultats concrets, l'expérience documentée de la campagne de 2012, qui expliquent le recours aux panels dès que cela est possible.

En 2021, la contestation des access panels a été reprise dans la grande presse. Un article de Luc Bronner du service politique du *Monde* paru le 5 novembre 2021, a rassemblé, à l'encontre des panels, des avis et opinions assez forts : des panélistes seraient principalement motivés par l'argent, ils répondent n'importe comment, peuvent mentir sur leurs identité et profil, sont inscrits sur plusieurs panels et répondent sans cesse, ils ne représentent pas la France dans sa diversité, etc. Tous ces défauts étaient étayés par l'expérience du journaliste qui avait consacré du temps à s'inscrire sur plusieurs panels, à mentir sur son identité, à répondre n'importe comment à un grand nombre de questions. Cette expérience mettait donc à jour une question simple : si un individu peut faire n'importe quoi, un groupe d'individus concertés et animés par un mobile cohérent ne peut-il pas, à des fins partisans par exemple, faire de "l'entrisme" et, au final, altérer la sincérité du sondage ?

M. Bronner a, par ailleurs, rassemblé l'avis de chercheurs et de consultants qui critiquent le caractère déclaratif des réponses, regrettent certains avantages du face-à-face ou du téléphone, ou estiment que les panels ne maîtriseraient pas le processus de recrutement, le caractère « volontaire » de la participation aux sondages constituant un biais. Bref, les panels seraient « un déni de la notion de sondages ». Autant d'avis qui ne sont pas confirmés par l'expérience de la commission fondée, quant à elle, sur l'examen systématique des sondages panélisés au cours de la dernière décennie.

Un tel écart de perception appelait donc un examen quasi systématique des pratiques des instituts que la commission n'avait pas réalisé depuis longtemps. Il le fut au cours de plusieurs semaines, pour l'essentiel en séance plénière.

Les conclusions de la commission s'établissent comme suit :

Le travail des sociétés d'access-panel, qu'elles soient juridiquement indépendantes ou non des instituts de sondages¹⁰, est complexe, méticuleusement mené, empreint d'une culture de l'autocontrôle. On appellera ci-dessous sociétés d'access-panel toutes les sociétés qui assurent cette activité, quelle que soit la nature juridique du lien qui les rapproche des instituts de sondage. Notons par ailleurs que les sondeurs, quel que soit ce lien, développent, en sus des contrôles de l'access-panel, leur propre contrôle de cohérence, s'agissant notamment des intentions de vote en vue de se prémunir d'évolutions non crédibles. La recherche de résultats cohérents dans le temps est fondée

¹⁰ Les instituts se sont organisés pour que, fonctionnellement, l'activité d'access-panel soit toujours séparée de l'activité de sondeurs.

sur l'idée qu'avec l'internet a disparu désormais tout risque d'évolution erratique de l'opinion ou plutôt de sa mesure que l'on appelle ici artéfacts.¹¹

Aucun acteur du secteur ne prend les avertissements avancés par l'article du *Monde* à la légère mais tous relèvent que recourir à l'internet revient à retenir le mode d'interrogation qui a le taux d'équipement le plus large, que le choix du *on-line* correspond au souci de disposer d'échantillons bénéficiant d'une forte représentativité, c'est-à-dire d'une qualité de souvenir de vote bien meilleure qu'avec le téléphone et donc offre l'avantage de minimiser les coefficients du redressement politique.

Certes, les « panélistes » sont plus connectés que la population moyenne, mais tous les autres modes d'interrogation – face-à-face ou téléphone – souffraient de biais autrement plus problématiques.

En cas d'abstention massive et différenciée (régionales 2021 par exemple), celle des jeunes et des ouvriers en particulier, l'incapacité des panels à justement apprécier le phénomène n'est pas de nature à mettre en cause leur performance enregistrée en 2017 et 2022 : fluidité des évolutions et, in fine, proximité des estimations et des résultats.

L'article de presse laisse entendre que les instituts ou les access-panels n'auraient à aucun moment réfléchi aux questions qu'ils posent. Cela n'est pas tout à fait vrai.

Au-delà de ces généralités, les principaux enseignements qui peuvent être rendus publics et qui ont trait aux moyens mis en œuvre pour parer au risque « d'entrisme » que des pratiques du type de celles décrites et mises en œuvre par le journaliste du *Monde*, qui seraient appliquées à plus grande échelle, pourraient entraîner sont les suivants :

- 3 – 1 Le recrutement des panels

Le contrôle du recrutement des panélistes passe d'abord par une vérification automatique en continu de l'importance des afflux.

Il existe parallèlement un contrôle sur l'utilisation d'une même IP et sur l'existence des personnes qui ne doivent pas être des robots.

¹¹ Si la campagne 2022 a vu les intentions de vote de certains candidats progresser rapidement, puis stagner avant de subir une progressive attrition, les sous-ensembles gauche + extrême gauche et droite + extrême droite n'ont évolué que très lentement.

Rappelons ici que le but n'est pas d'avoir un panel représentatif, mais un panel volumineux, renouvelé, sain et qui permet de constituer des échantillons représentatifs.

3 – 1.1 Les sources de recrutement des panélistes sont toujours multiples, de cinq à dix et variées. L'accord est unanime sur le fait que la grande diversité de recrutements est gage de qualité.

- le recrutement aléatoire, proposition faite à la fin d'enquêtes téléphoniques
- le recrutement naturel via un site dédié au recrutement dont le nom n'évoque pas celui de l'institut en général et alimenté en particulier par les réseaux sociaux
- les campagnes de « display »¹²
- les campagnes publicitaires sur réseaux sociaux
- le recrutement par sites relais
- le recrutement par accords avec d'autres acteurs de l'internet qui sert à couvrir des cibles particulières (jeunes, personnes plus âgées, ouvriers)
- le recrutement direct par support papier ou à l'occasion d'opérations spéciales
- le recrutement par campagne d'emailing
- le parrainage

3 – 1.2 Les modalités de recrutement

Du fait qu'un individu se fasse passer pour ce qu'il n'est pas, on ne peut déduire que les access-panel ne peuvent ni identifier, ni bloquer des entrées massives.

Au surplus, le fait même qu'un répondant soit arrivé au bout des enquêtes ne signifie pas que ses réponses ont finalement été prises en compte.

Les sociétés d'access panel « monitorent » le nombre de personnes inscrites. Elles disposent de ratios d'entrants que toute opération de recrutement doit apporter. Si les inscriptions se font plus massives ou plus localisées sur un site d'origine, elles sont alertées des écarts à la moyenne. Les sociétés d'access-panels ont pu craindre que l'article du *Monde* soit considéré par certains comme un mode d'emploi du *trolling* ou de la manipulation des sondages ; elles ont donc été particulièrement vigilantes¹³ dans les semaines qui ont suivi sa publication sans relever de mouvements anormaux.

Pour l'heure, le risque de *trolling* ne s'est pas manifesté.

L'inscription s'opère en général à double *opt-in*

¹² Annonces sur des sites web ou sur des applications

¹³ Il n'est pas difficile aux acces-panel de mettre en quarantaine les nouveaux inscrits ou de comparer leurs intentions avec celles des anciens. Ils l'ont fait sans déceler de distorsion. Il a été procédé de même à l'automne 2016 pour les sondages relatifs à la primaire de la droite et du centre en raison de la poussée du nombre de supporters de François Fillon à la suite du premier débat télévisuel.

Les panélistes ne recherchent pas d’hyperqualification du panel, l’objectif n’étant pas de tout savoir sur les personnes qui s’inscrivent, mais de connaître essentiellement leurs variables socio-démographiques qui suffisent aux besoins des études, notamment au *screening*¹⁴ et évitent un questionnaire trop long qui pourrait rebuter certains et donc d’engendrer un biais de motivation.

Le recrutement participe de la recherche d’un équilibre entre ce qui permet de soutenir la participation de panélistes sans créer d’aspects communautaires, ce qui tendrait vers le professionnalisme des panélistes.

3 – 1.3 Contrôle de la motivation des panélistes

Une remarque de M. Bronner est ici confirmée par les sociétés d’access-panel : la “rémunération” est trop faible pour constituer une réelle motivation. Cette rémunération n’en est au demeurant pas une : le fait de compléter la réponse à un sondage n’entraîne pas le versement d’une somme, mais permet de cumuler des points, dont le total, au prix d’une certaine fidélité, peut se matérialiser dans l’obtention d’un bien dont la valeur est estimée en fonction du nombre de points acquis.

Au reste, la question de la motivation de la présence dans un panel n’est pas abordée par l’article du *Monde*. Elle n’est pourtant pas négligeable. Les panélistes doivent être intéressés pour donner, un temps, leur avis sur de grandes décisions des entreprises ou des pouvoirs publics, découvrir en exclusivité de nouveaux services, produits ou mesures, exprimer des idées ou préférences politiques.

Un intérêt exclusif pour les sondages politiques est une cause d’exclusion des panels.

Notons que la motivation varie avec l’outil de recueil. Les jeunes, par exemple, ne veulent pas s’engager dans la durée, ils nécessitent des procédures de recrutement coûteuses et spécifiques via les réseaux sociaux. Au surplus, les jeunes ne répondent pas par ordinateur comme leurs parents, mais sur leur téléphone portable.

3 – 1.4 Contrôle de l’ancienneté moyenne des panélistes.

Les sociétés recherchent généralement une durée moyenne de présence de leurs panelistes, entre 3 ans et 5 ans. Au-delà des exclusions, il faut laisser se produire un “départ naturel” des panélistes.

¹⁴ Dégager une population particulière pour une enquête rapide portant sur un nombre très limité de questions.

- 3 – 2 Le contrôle de la participation aux échantillons

3 – 2.1 Fréquences des enquêtes

sont contrôlés :

- le nombre de sollicitations annuelles
- le nombre d'enquêtes réalisées en général avec un maximum par mois
- le nombre d'enquêtes politiques par mois
- le délai de viduité entre deux sollicitations.
- le délai de viduité sur le même sujet.

Aucune réponse ne doit être obligatoire. La non-réponse est toujours un choix.

Parfois, la sélection des panélistes se fait en fonction d'un algorithme qui, à chaque type d'étude, affecte les panélistes d'un poids en fonction de son historique de participation.

3 – 2.2 Des vérifications sur les membres de l'échantillon viennent s'ajouter :

- vérification du n° de téléphone via SMS
- contrôle de l'intégrité des réponses ;
- *Sanity shake*
 - contrôle des profils strictement identiques sur un même questionnaire
 - contrôle des profils enregistrés sur le compte et déclarés dans le questionnaire
 - contrôle de cohérence interne par questions similaires
 - contrôle du temps de réponse
 - contrôle d'une même réponse à toutes les questions. (*Straight liner*)
 - contrôle de l'horaire d'interrogation

Un répondant ne peut participer deux fois à la même enquête. Le lien envoyé à chaque panéliste est personnalisé et ne peut servir qu'à une seule réponse. S'agissant des access-panels utilisés par plusieurs instituts, l'absence de participation d'un même panéliste à plusieurs enquêtes électorales concomitantes est assurée.

D'une façon plus large, il convient de ne pas perdre de vue que, pour la constitution de chaque échantillon, c'est la société qui "choisit" les panélistes, ce qui supprime tout biais de motivation de la part des panélistes sur-intéressés par la politique.

Les sociétés d'access panel disent n'avoir jamais observé d'inscriptions en masse. Sans doute l'investissement serait-il trop lourd pour un résultat trop aléatoire. Même 1 000 militants bien organisés qui seraient entrés très progressivement dans les panels ne seraient pas assurés d'être interrogés en nombre suffisant pour une enquête donnée afin d'en altérer la sincérité. Le risque d'entrisme provoquant un biais est plus théorique que pratique.

Notons enfin que bien des instituts croisent leur panel de sociétés sœurs avec des panels de sociétés d'access-panel sous-traitantes soit pour des besoins spécifiques, soit pour des enquêtes lourdes.

Tout mode de recueil a ses défauts. Il n'y a pas de moyen idéal de recueillir l'opinion. Qu'il soit mené plus ou moins indépendamment du métier de sondeur, le métier de panéliste cherche le moins mauvais mode de recueil par arbitrage entre des solutions limitées pour des raisons techniques ou économiques.

• 3 – 3 Les défauts décelés par les sondeurs à propos des panels ont trait par exemple à la possibilité de réaliser des études sur des territoires restreints, des enquêtes portant exclusivement sur des populations de plus de 75 ans, ou les enquêtes sur l'appétence envers les nouvelles technologies. Ces défauts concernent marginalement les sondages électoraux en vue de scrutins nationaux. À l'issue d'un tel tour d'horizon des pratiques des panélistes, qu'ils soient plus ou moins indépendants des instituts de sondages, la conviction de la commission est que la culture du contrôle domine ce métier. Peut-être n'est-elle pas sans faille, le suivi de l'opinion est un art difficile, mais rien n'est plus éloigné des procédures *low cost* que le métier de panéliste tel qu'il est, pour l'heure, pratiqué en France¹⁵.

Ainsi, le travail de constitution des échantillons apparaît-il intelligemment conçu, sérieusement et méticuleusement mené. Cela explique possiblement l'absence d'artéfacts constatés dans les sondages électoraux tout au long de la dernière décennie.

¹⁵ Une nouvelle norme ISO 20252 2019 intègre la gestion d'un panel. Mais l'Afnor n'audite pas encore sur cette version. Ses règles, toutefois, semblent moins contraignantes que ce que la profession s'impose spontanément à l'heure actuelle.

L'échec des tentatives de se libérer de la discipline des panels, tentatives qui se sont traduites sans délais par le retour d'errements anciens¹⁶ établissent la fiabilité du recours aux panels notamment pour les grandes enquêtes nationales. Pour autant la commission des sondages ne considère pas devoir imposer, une méthode, une pratique, un contrôle de qualité qui, en l'état du savoir, s'impose de lui-même. Laissons, élection après élection, la possibilité de découvrir d'autres outils de suivi de l'opinion, d'autres procédures, d'autres règles de l'art. Peut-être que les progrès de l'intelligence artificielle sauront demain mieux gérer le "monde data" dans lequel nous entrons.

Relevons que les actuels responsables d'instituts appartiennent tous à une génération qui a commencé son activité avec l'interrogation exclusivement par téléphone pour tous les types de sondages et qui recourent encore au téléphone pour les estimations liées aux scrutins locaux. Quand l'interrogation *on line* est devenue possible, ils ont tous choisi ce mode de constitution des échantillons méthodologiquement plus contraignant ; aujourd'hui aucun ne voudrait revenir en arrière.

¹⁶ Cf. Supra 2– 2.3

Conclusion

Au-delà du rapport de l'année électorale 2022, l'actuelle commission des sondages, qui ne siègera plus à compter du printemps 2023, souhaite évoquer l'effet que provoquera alors la première application de la nouvelle disposition (née de la loi n° 2017-55 du 20 janvier 2017), de l'article 6 de la loi n° 77-808 du 19 juillet 1977 relative à la publication et à la diffusion de certains sondages d'opinion. Cette disposition qui prévoit le renouvellement complet de la commission peut apparaître comme une règle inutilement rigide dans une matière aussi technique que les sondages électoraux et pour une activité aussi intermittente

Par ailleurs, la commission pense que le législateur pourrait réfléchir à une éventuelle interdiction la veille et le jour du scrutin de la publication de synthèses et commentaires des derniers sondages.

Pour l'essentiel, le bilan de cette année électorale montre que le texte et l'esprit initial de la loi restent adaptés aux évolutions considérables de la pratique des sondages électoraux depuis près d'un demi-siècle.