

Congrès de l'AFSP – 22-24 juin 2015 – Aix-en-Provence
Section thématique 7 : Les appropriations méthodologiques d'internet dans la recherche sur des objets politiques

Anne Jadot, anne.jadot@univ-lorraine.fr
Université de Lorraine / IRENEE / CEVIPOF
Pierre Lefébure, pierre.lefebure@univ-paris13.fr
Université Paris 13 – Sorbonne Paris Cité / Laboratoire Communication et Politique-IRISSO

**Recruter des répondants pour des enquêtes en ligne
grâce à une « Voting Advice Application ».
Les enjeux méthodologiques de la 'panélisation' d'utilisateurs de la
Boussole Présidentielle 2012.**

Cette contribution propose une double réflexion sur l'intérêt et les difficultés du recrutement de répondants par un dispositif internet. Plus précisément, nous explorons les possibilités d'Internet pour recruter des répondants pour un panel électoral en ligne. Le point de départ est une application en ligne de type « Voting Advice Application » (VAA) avec l'objectif de constituer un panel, c'est-à-dire un vivier de répondants qui puissent être interrogés à plusieurs reprises dans le temps. Nous nous fondons sur l'expérience menée en 2012 et 2014. L'usage d'internet est, dans notre expérience, triple :

- il permet de récolter, dans un premier temps, des données et des métadonnées auprès d'une population très vaste d'internautes volontaires, qui utilisent un outil pour leur propre intérêt de s'informer sur une élection en cours ; ces données pourront d'ailleurs être appariées aux réponses ultérieures aux différentes vagues du panel
- il permet de tenter de convaincre la population des utilisateurs de la VAA de participer au panel, en accomplissant la démarche proactive de fournir une adresse email, grâce à une fenêtre « pop up » recueillant un consentement éclairé et express
- il permet, enfin, d'administrer les différentes vagues du panel, avec de nombreux avantages de souplesse et de réactivité pour l'équipe scientifique.

L'application internet développée par l'équipe de recherche et la satisfaction d'usage qu'en tirent les utilisateurs constitue la meilleure motivation d'entrer dans le panel, en saisissant des internautes volontaires au moment d'une campagne, recueillant des données au moment où leur décision électorale est encore en train de se faire, et avec la possibilité d'entretenir un panel sur plusieurs années. Hormis la possibilité de s'exprimer sur des sujets qui leur tiennent à cœur (et nous verrons qu'il s'agit probablement pour eux d'une dimension importante, parfois de manière négative / critique), il n'y a donc pas de récompense ultérieure matérielle à intégrer le panel tiré d'une Boussole électorale. L'entrée dans le panel est en quelque sorte une manière pour les utilisateurs de l'application Boussole de récompenser l'équipe de recherche pour un dispositif qui leur a plu et qui leur a fourni une information utile, au point souvent qu'ils en aient parlé autour d'eux, qu'ils l'aient recommandée à des proches via les réseaux sociaux, alimentant un « bouche à oreille » et, partant, le vivier de recrutement. Avant de présenter notre démarche, nous présentons dans une première section un état des démarches en cours à l'international dans le domaine des panels en ligne, qui peuvent être en population générale à large vocation de recherche pour les sciences sociales ou plus étroitement électoraux liés à une campagne précise. Cela nous permet de montrer les

avantages et les inconvénients du web pour l'administration des questionnaires, le problème d'inégalité d'accès à Internet, pas toujours compensé par des dispositifs visant à amoindrir les biais et, enfin, l'enjeu crucial du moment du recrutement, quand les futurs panélistes sont sélectionnés ou s'auto-sélectionnent.

1) L'intérêt des panels en ligne pour la sociologie, et plus spécifiquement la sociologie électorale

a) Administrer des enquêtes en ligne : de nombreux avantages, renforcés en matière de recherche électorale

Les avantages d'internet pour l'administration d'une enquête (qu'elle soit *ad hoc* ou longitudinale) sont importants pour le domaine électoral. Il y a à la fois des aspects logistiques d'organisation du terrain et des aspects substantifs par rapport aux questionnaires qu'il est possible d'administrer en ligne.

Du point de vue opérationnel, la diffusion en ligne permet une grande réactivité, une véritable souplesse pour l'envoi simultané en masse de questionnaires à la date - et même à l'heure - précisément choisie pour le lancement du terrain. L'envoi d'un lien vers l'enquête en ligne peut ainsi avoir lieu juste après un événement précis de la campagne, par exemple juste après l'annonce des résultats d'un premier tour, ou après une émission télévisée marquante. Cela se fait selon le choix de l'équipe, sans aucune autre contrainte que sa propre organisation interne. Ladite équipe acquiert d'ailleurs aussi plus d'autonomie par rapport aux terrains pour lesquels elle doit s'en remettre à l'institut de sondage et au travail des enquêteurs. On « internalise » non seulement la rédaction mais aussi le montage concret du questionnaire en ligne, éventuellement en laissant certains choix pour la toute dernière minute, par exemple, une fois connue l'identité des qualifiés pour un second tour. Il est tout de même possible de pré-tester voire de 'piloter' plusieurs fois le questionnaire. En outre, le taux de participation peut être stimulé par la possibilité pour les répondants de remplir le questionnaire à l'heure qu'ils souhaitent, à leur convenance, et parfois par le canal qu'ils souhaitent (ordinateur, tablette voire smartphone). Eventuellement, si l'équipe de recherche le décide et paramètre ainsi l'enquête (via des logiciels tels que Sphinx ou Qualtrics), les répondants ont même la souplesse d'enregistrer le début de leurs réponses et de reprendre le questionnaire plus tard. Cela augmente considérablement l'amplitude horaire des réponses par rapport à un terrain où il y a des enquêteurs, quel que soit le mode d'administration. La déconnexion entre le contact initial et le remplissage du questionnaire peut donc s'étaler dans le temps, avec en outre la possibilité de gérer des relances, en 'double aveugle' : on ne sait pas qui a répondu quoi mais on peut relancer le listing mails de ceux qui n'ont pas du tout commencé à répondre et/ou de ceux qui ont commencé le questionnaire mais qui ne l'ont pas terminé. Ce 'rappel' (au sens de *reminder*) est une occasion de plus de convaincre de participer alors que, via la méthode des quotas la plus habituelle en France, chaque répondant potentiel qui n'est pas immédiatement disponible quand il est contacté est définitivement perdu. Il est d'ailleurs possible d'enregistrer des précisions sur le comportement de participation (réponse intégrale en une seule fois ou en plusieurs connexions) au titre des « métadonnées » collectées en ligne, comme par exemple le temps de réponse total, voire item par item, selon la programmation effectuée par l'équipe.

Au titre des aspects substantifs, la nature des questions posées est impactée positivement par l'auto-administration en ligne. On peut inclure des questions dites 'sensibles', avec moins de biais de désirabilité sociale. L'absence d'enquêteur, en particulier pour ce qui touche au politique et au vote, serait en quelque sorte similaire au dispositif de la « fausse urne »

longtemps mis en place dans les enquêtes électorales du CEVIPOF quand il y avait un terrain en face à face, ce qui évitait les effets structurels de sous-déclaration de certains électors (notamment l'extrême droite) et conjoncturels connus sous le nom de « spirale du silence ». En outre, comme dans les enquêtes administrées en face à face, il y a la possibilité de soumettre aux répondants des visuels à l'appui des questions, ce qui peut être important quand il y a des listes d'items un peu longues, qui sont plus difficiles à administrer par téléphone. Cela minimise les biais de récession et de primauté des derniers et premiers items. On peut même être encore plus créatifs grâce au support web, en incluant si on le souhaite des stimuli auditifs, voire vidéos : par exemple, soumettre une photo ou une affiche de campagne, un extrait de discours, d'émissions TV ou des clips de campagne. C'est certes également possible avec des interviews CAPI mais cela suppose que l'enquêteur laisse pendant un temps l'ordinateur portable à l'enquêté, éventuellement avec écoute au casque, ce qui rallonge la durée moyenne de passation. Un autre aspect positif est la possibilité plus grande d'incorporer des questions ouvertes, pour laisser s'exprimer les répondants avec leurs propres mots, même si une relance éventuelle (en fonction du paramétrage du questionnaire) pour inciter à saisir quelque chose en cas de question ouverte laissée vide a sans doute moins de poids qu'avec un enquêteur qui interroge. D'expérience, c'est un type de questions auquel collectivement les équipes de recherche rechignent parfois avec d'autres modes d'administration, non seulement parce que, au stade des analyses, le traitement du verbatim est plus long et compliqué que pour les questions fermées, mais aussi parce que la saisie par les enquêteurs de réponses en clair allonge la durée de passation et, partant, le coût de l'enquête. Car c'est un autre avantage, peut-être même un des principaux, des enquêtes en ligne : administrer ainsi un questionnaire coûte beaucoup moins cher qu'avec les autres modes, et l'on peut ainsi viser des échantillons de plus grande taille pour un coût très raisonnable, en économisant sur le poste budgétaire des enquêteurs. Cela ne doit évidemment pas être le seul argument en faveur des enquêtes électorales en ligne, mais c'est un aspect important. D'autant plus qu'on peut concentrer le terrain sur quelques jours seulement (par exemple entre le débat télévisé et le jour du second tour) et tout de même vite atteindre plusieurs milliers de questionnaires remplis, sans pour autant déployer de manière intensive sur le terrain un grand nombre d'enquêteurs travaillant en parallèle, ce qui serait d'un coût prohibitif.

Bien sûr, ce mode d'administration en ligne présente aussi des inconvénients, le principal étant qu'une partie de la population n'a pas d'accès à Internet (cela a été estimé à 20% selon Gombault 2013), même si c'est régulièrement en baisse. Certains chercheurs restent par ailleurs sceptiques sur l'identité de qui répond vraiment et la véracité des réponses collectées en ligne, puisqu'on n'est pas sûr de qui ouvre le courriel reçu avec le lien vers l'enquête et parce qu'il pourrait y avoir une certaine 'licence' à la fantaisie quand il n'y a pas d'enquêteur pour entendre et saisir les réponses. Mais pourquoi les sondages par téléphone seraient-ils exempts d'un soupçon similaire ? Et, de toute façon, les chercheurs travaillent toujours *in fine* avec du déclaratif dans des enquêtes d'opinion, sauf quand il y a une « enquête validation de la participation électorale ». Il n'y a certes pas en ligne de lien de confiance tel que celui qui peut être bâti dans le cadre d'une interaction enquêteur / enquêtés, mais il n'y a pas non plus de biais de désirabilité sociale. Au delà de ces bémols, l'inconvénient principal des enquêtes en ligne est qu'il faut trouver un moyen de toucher la population cible, de diffuser le questionnaire : soit disposer d'une base d'adresses à qui envoyer un message avec un lien vers l'enquête, soit toucher la population des internautes d'une autre manière, en mettant directement l'enquête (ou, à tout le moins, la tentative de recrutement de répondants) à disposition sur internet. Différentes pistes ont été explorées ces dernières années par la recherche publique et universitaire pour administrer des enquêtes en ligne qui sont de surcroît longitudinales, pour constituer de véritables panels d'internautes, lesquels peuvent avoir

l'ambition forte de représenter la population générale, ou bien viser plus spécifiquement les seuls électeurs inscrits.

b) Des recherches en plein essor : des panels en « population générale » qui cherchent à réduire les biais d'accès à internet...

A travers le monde la recherche, soit dans des centres de recherche incorporés au sein d'une université, soit en lien avec des organismes de statistique publique, a entrepris ces dernières années de constituer des panels d'internautes qui sont voulus représentatifs de la population générale. C'est un champ en plein développement, avec une actualité passionnante.

Au titre des expériences étrangères pionnières, on peut citer le *Knowledge Panel* de Knowledge Networks aux Etats-Unis, créé en 1999 par deux universitaires américains. Ce dispositif repose sur un échantillonnage aléatoire et sur l'installation d'une connexion à internet chez les répondants qui n'en disposent pas à leur domicile. L'échantillon, tiré aléatoirement à partir d'une base d'adresses de logements, est composé de 55 000 personnes de 18 ans et plus. Il est à noter que ce dispositif est également ouvert aux études commerciales. Il y a d'ailleurs eu un tournant important en 2011 : le Knowledge Panel a été racheté par l'institut de sondage GfK. Dans le cadre du programme TESS, l'accès à ce panel reste toutefois gratuit pour les chercheurs en sciences sociales, le programme prenant en charge le coût financier pour les projets d'enquête qui sont sélectionnés. Dans cette lignée, l'enquête *Longitudinal Internet Studies for the Social Sciences* (LISS Panel) a été mise en place par l'institut de recherche néerlandais CentERdata de l'Université de Tilburg. Elle est représentative des ménages néerlandais et le panel a été constitué en 2007 selon un plan de sondage probabiliste réalisé en collaboration avec le Centraal Bureau voor de Statistiek (CBS), l'Institut national de statistique des Pays-Bas. Il repose là aussi sur la mise à disposition d'un ordinateur simplifié et d'une connexion à internet pour les ménages qui en sont dépourvus. Le LISS Panel est constitué de 5 000 ménages, soit 8 000 individus de 16 ans et plus. Ce dispositif est gratuit et exclusivement dédié à des opérations de recherche. Le *German Internet Panel* (GIP) de l'Université de Mannheim est, lui, représentatif de la population âgée de 16 à 75 ans résidant en Allemagne. Il repose sur un échantillonnage aléatoire de 1 500 personnes et les membres du panel sont interrogés tous les 2 mois. Comme dans le cas du LISS Panel, un ordinateur simplifié et une connexion internet sont mis à disposition des personnes qui en sont dépourvues. Ce dispositif est réservé aux seuls chercheurs de l'Université de Mannheim et, thématiquement, les questionnaires sont très centrés sur les réformes politiques. Le *GESIS Panel* du GESIS Leibniz Institute for the Social Sciences de Mannheim se caractérise, par contraste avec les exemples précédemment cités, par la coexistence de deux modes possibles d'interrogation, les deux étant toutefois auto-administrés. En effet, les personnes disposant d'un accès à internet depuis leur domicile peuvent répondre sur le web ; les personnes ne disposant pas d'un accès à internet ou qui ne souhaitent pas répondre en ligne peuvent répondre sur papier aux questionnaires envoyés par voie postale. L'échantillon a été tiré aléatoirement à partir des registres municipaux, il est constitué de 4 000 individus âgés de 18 à 70 ans et résidant en Allemagne. Les enquêtes bimestrielles sont proposées par des équipes de recherche en sciences sociales dans le cadre d'appels à projets et ne peuvent en aucun cas servir un intérêt commercial. Dernier exemple présenté ici, le *Citizen Panel* a été lancé en Suède par le Laboratory of Opinion Research (LORE), un centre académique au sein du département de Science Politique de l'Université de Göteborg précisément créé en 2010 pour explorer l'opinion publique et la démocratie en développant des initiatives méthodologiques. En novembre 2012, 29 000 individus ont été tirés au sort dans un registre de la population où sont disponibles, outre l'adresse postale, des

informations sur le genre, la date de naissance et la composition des ménages. Ils ont reçu une carte postale les invitant à participer au panel, avec la mention d'un identifiant et d'un mot de passe pour qu'ils aillent s'identifier sur un site dédié. De très gros efforts d'expérimentations ont été mis en place dans ce cadre, par exemple pour comprendre l'efficacité de différentes rédactions - plus ou moins personnalisées - de l'invitation à participer, ou encore de l'inclusion, pour certains sous-échantillons tirés au sort, d'une motivation monétaire peu coûteuse pour l'équipe (de l'ordre de 0,30€). Il s'agissait en l'occurrence d'un n° de loterie attribué par avance à chaque personne contactée mais ledit numéro était validé pour pouvoir participer au tirage de la loterie nationale uniquement en cas d'inscription en ligne pour le panel. En 2014, les chercheurs du LORE ont mis en place avec les chercheurs néerlandais de Kieskompas une « Voting Advice Application », en collaboration avec le grand quotidien Aftonbladet, ce qui leur a permis de recueillir l'adresse mail de quelques milliers d'utilisateurs. Nous y reviendrons car cela permet de comparer leur approche avec celui de notre équipe pour les Boussoles électorales développées au CEVIPOF.

En France, en s'inspirant d'ailleurs de certains de ces panels et en travaillant en collaboration notamment avec le panel néerlandais LISS, le dispositif d'*Enquête Longitudinale par Internet pour les Sciences Sociales* (ELIPSS) a été mis en place par le Centre de Données de Sciences Po avec un consortium d'universités et en partenariat avec des institutions de la statistique publique tels que l'INED, et l'INSEE pour l'échantillonnage. Il se distingue des dispositifs similaires étrangers par le choix de l'internet mobile comme mode de collecte principal. Une tablette tactile et un abonnement 3G sont en effet fournis à l'ensemble des panélistes en échange de leur participation. Cela représentait une « incentive » intéressante au moment du lancement, puisque 9% des ménages étaient équipés d'une tablette tactile début 2012, même si le taux d'équipement a rapidement progressé et qu'on l'estimait à un ménage sur trois fin 2013. Ainsi, les panélistes recrutés peuvent répondre aux questionnaires même s'ils ne disposent pas de connexion internet avant leur participation au dispositif. La phase de pilote s'achèvera à l'hiver 2015-2016 mais, d'ores et déjà, des enseignements sur le recrutement peuvent être tirés d'ELIPSS¹ (cf. Cornilleau et al, rapport d'ELIPSS à l'INPES, juin 2014). La base de sondage était constituée des logements qui avaient été recensés en 2011 et, en son sein, un échantillon de 4 500 logements a été tiré au sort par l'INSEE, avec des stratifications dont le détail technique n'est pas utile ici. Les 4 500 adresses ont été divisées en trois sous-échantillons : un échantillon principal de 3 500 adresses et deux échantillons de réserve de 300 et 700 adresses, à utiliser au cas où l'objectif initial des 1 500 panélistes ne serait pas atteint. Au sein de chaque logement sélectionné par l'INSEE, une fois l'accord de principe obtenu au niveau du ménage, une seule personne était ensuite tirée au sort parmi les éligibles (les 18-75 ans) et cette personne doit accepter de participer au panel, c'est-à-dire s'engager à répondre à une dizaine d'enquêtes par an via la tablette tactile (laquelle peut évidemment être utilisée en dehors des enquêtes, connexion 3G incluse, par tous les membres du ménage). Les répondants devaient signer une convention, se faire envoyer la tablette, bénéficier si besoin d'un tutoriel d'apprentissage, lequel était l'occasion de poser une « enquête zéro » sur les pratiques numériques, afin de mesurer ensuite les effets en ce domaine de la participation au panel. L'exploitation de l'échantillon principal a eu lieu de juin 2012 à février 2013. Trois modes de contact ont été mis en œuvre. Une invitation à participer a été envoyée par courrier postal. Ce premier contact a fait l'objet d'une expérimentation : la moitié des courriers avait été envoyée avec un bon cadeau de 10 euros (acquis quelle que soit la décision du ménage, c'est-à-dire non conditionné à l'acceptation du panel) et l'autre moitié sans bon cadeau. Des

¹ Certaines réflexions de la présente communication sont d'ailleurs nourries par le fait que le premier co-auteur est membre depuis fin 2012 du Comité Scientifique et Technique d'ELIPSS, qui est le volet quanti de l'Equipex DIME-SHS, lequel fut lauréat en 2011 des « investissements d'avenir ».

relances postales et téléphoniques ont ensuite été réalisées. Mais l'échantillon principal avait été épuisé sans que les objectifs de recrutement aient été atteints. A l'issue d'un appel d'offres, l'institut de sondage TNS Sofres a été retenu pour exploiter en face-à-face les adresses pour lesquelles il n'y avait pas eu de contact et pour tenter de convaincre certains des ménages qui avaient refusé de participer. Par ailleurs, de janvier à avril 2013, les 1 000 adresses de réserve ont été exploitées par téléphone (du moins, pour les ménages pour lesquels un n° avait pu être trouvé) et en face à face par TNS Sofres. Au terme de cette double procédure, 1 026 personnes ont accepté de participer au panel, plus de la moitié (51%) ayant été recrutée en face à face, 31% par courrier et 18% par téléphone. Il est d'ailleurs intéressant de constater que ce dispositif, précisément conçu pour limiter l'impact de la fracture numérique sur les biais de représentativité a, dans les faits, conduit à recruter un panel où la proportion des répondants qui étaient équipés d'un accès internet *préalablement* à leur entrée dans le panel ELIPSS est *supérieure* à la moyenne des ménages français. Plus précisément, parmi les panélistes recrutés, seuls 9% n'étaient *pas* dotés d'un accès à internet à domicile avant ELIPSS, alors qu'une enquête sur les technologies de l'information et de la communication et le commerce électronique réalisée en 2012 par l'INSEE chiffre, sur la population cible des 18-75 ans, cette proportion de *offliners* à 17%. La mise à disposition d'équipement ET d'abonnement ne contribue donc pas à l'objectif de réduire les biais d'accès à internet. De tels décalages avaient d'ailleurs également été pointés dans le cas des dispositifs LISS et GIP. On peut se poser la question de savoir si l'acceptation de ce type de dispositif longitudinal serait facilitée en amont par une familiarité préalable avec l'outil numérique, ce qui tempère un peu la volonté de tirage au sort probabiliste. Un mécanisme d'auto-sélection semble tout de même à l'œuvre (sans mentionner ici les effets d'âge et de diplôme).

Au delà de ces équipements et abonnements, il y a eu des expérimentations sur d'autres formes de rétributions matérielles de la participation, qui peuvent être plus ou moins symboliques, avec l'idée de maximiser l'acceptation du panel sans trop renchérir le coût du terrain. Il y a par ailleurs des passerelles entre les projets rapidement présentés ici, avec des coopérations entre les équipes citées : par exemple, LISS, GIP, GESIS et ELIPSS ont administré de façon simultanée la même enquête au printemps 2014, avec un questionnaire qui allait pour partie être posé également en ligne mais à un échantillon *ad hoc* 'frais', en relation avec les enquêtes sur les élections européennes. Cela permet de faire office de 'contrepoint' et de mesurer, selon les pays, les effets du mode d'administration en plus du recrutement ponctuel ou dans le cadre d'un panel. Pour généraliser, on peut souligner que ces différentes enquêtes, conçues pour être représentatives de la population générale, ont en tout cas vocation à durer plusieurs années et à être utiles à plusieurs domaines de sciences sociales. C'est ce qui justifie le coût du recrutement initial, les investissements étant parfois conséquents. Cela pose aussi des questions autour de l'attrition, c'est-à-dire la « mortalité » du panel, mais aussi d'un possible « conditionnement », pour la partie des questionnaires qui est régulièrement répétée d'une vague à l'autre pour mesurer la stabilité des attitudes et le changement et qui peut, à tout le moins, entraîner une certaine lassitude. D'autres stratégies de plus court terme, ciblées sur le comportement politique, cherchent également à constituer des panels d'internautes, autour d'une campagne électorale.

c) ... et des enquêtes auprès d'électeurs recrutés en ligne, avec l'enjeu crucial du mode d'(auto) sélection

Par contraste avec les études sociologiques citées ci-dessus, des enquêtes en ligne peuvent être mises en place de façon plus ponctuelle par rapport à une élection donnée, soit avec un seul questionnaire (pré- ou post-électoral), soit avec un nombre plus ou moins élevé de vagues

de panel pour pouvoir suivre un échantillon avant et après la décision électorale. Le coût réduit des enquêtes en ligne, déjà mis en avant, plaide d'ailleurs assez logiquement pour mener plutôt de telles enquêtes 'panélistant' l'échantillon, en commençant plus ou moins en amont du jour de l'élection, précisément pour tenter de saisir de potentiels effets de campagne sur les choix électoraux (renouant en quelque sorte avec l'ambition initiale de l'école de Columbia et l'équipe de P. Lazarsfeld). Ceci n'a pas forcément vocation à remplacer mais à compléter (du moins, quand les équipes de recherche disposent d'un budget très conséquent !) le standard de l'enquête électorale qui était – et qui reste aux yeux de certains chercheurs – l'entretien en face à face au domicile des électeurs, comme cela fut initié dans la série des *American National Election Studies* (ANES), dont s'était ensuite largement inspirée la série des *British Election Studies* (BES). Mais, même dans ces deux équipes à la longue tradition d'interviews en face à face maintenues malgré les coûts induits, il y a eu l'introduction plus récente d'enquêtes en ligne conçues comme complémentaires, n'ayant pas vocation à remplacer le noyau traditionnel. Ce qui cristallise le débat scientifique, c'est le mode de recrutement de ces internautes pour une enquête électorale : invite-t-on à un panel en ligne grâce à un échantillonnage probabiliste ou en faisant fonctionner le volontariat ? Ce qui revient dans le second cas à laisser les répondants s'auto-sélectionner *via* ce qu'on appelle des « *opt in surveys* ». Autrement dit, comment choisir (ou atteindre) les électeurs à qui l'on souhaite soumettre le questionnaire en ligne (ou dont on souhaite qu'ils s'en saisissent d'eux-mêmes) ? Il n'existe en effet pas d'annuaire de l'ensemble des adresses emails de la population des électeurs inscrits - sans parler des doublons qu'un tel listing contiendrait ! Il ne peut donc pas y avoir de tirage au sort directement parmi un hypothétique fichier d'adresses mail pour désigner au hasard les répondants à qui envoyer un lien unique vers un questionnaire en ligne. Il faut donc disposer d'un autre type de base de sondage et d'une autre stratégie si l'on veut maintenir une volonté d'échantillonnage probabiliste, comme cela devrait être la 'norme' selon l'Association Américaine pour la Recherche sur l'Opinion Publique (AAPOR). Une croyance partagée par un article séminale de 2011 par Yaeger et al². Du coup, à partir de 2006, les ANES ont interrogé une partie du *Knowledge Panel* déjà présenté ci-dessus, pour conduire des enquêtes mensuelles pendant la période pré-électorale 2007-2008.

Dans les BES, il y a eu un mode de recrutement très différent³ pour la partie panel en ligne introduite dans les années 2000. Par exemple, pour l'élection générale de 2010, ils se sont appuyés sur un extrait du panel « You Gov », qui lui-même avait contacté plus de 360 000 adultes via des 'bannières' et d'autres moyens de contact. La structure des personnes à sélectionner en leur sein étaient dérivées de quotas (âge, genre, type de quotidiens lu

² Même si, selon le directeur de LORE et responsable scientifique du Citizen Panel suédois, Johan Martinsson, les données prises en compte dans cet article sont un peu anciennes (dans un domaine qui progresse rapidement et a vu le développement de fortes innovations en quelques années) et se sont limitées aux enquêtes probabilistes qui avaient un taux de réponse assez élevé, alors qu'on sait que c'est en fort et rapide déclin. Autrement dit, une enquête probabiliste avec un taux de participation faible n'est pas forcément intrinsèquement supérieure, cf. l'argument convergent de Rivers et Bailly, 2009.

³ Ce qui permet d'ailleurs des comparaisons intéressantes, pas seulement sur la qualité respective des données et les significations statistiques de tests et marges d'erreur, mais aussi pour explorer des résultats empiriques substantifs. En particulier, il peut y avoir d'éventuelles différences dans le pouvoir prédictif des variables selon les différents modes d'enquêtes. C'est le but d'un article de recherche par J. Karp et M. Luhiste, à paraître dans *Public Opinion Quarterly* (manuscrit communiqué par les auteurs) pour montrer qu'estimer la force du lien entre, d'une part, l'âge et, d'autre part, la participation politique et électorale varie beaucoup selon la source ANES / BES utilisée pour 2008 et 2010 respectivement car, pour chaque année, on dispose à la fois d'une composante traditionnelle en face à face et d'un panel en ligne.

régulièrement) et avec des pondérations fondées à la fois sur des données de recensement et sur les statistiques de lectorat, ainsi que sur la proximité partisane telle que mesurée régulièrement par You Gov auprès de 80 000 autres internautes.

Cela se rapproche en cela d'autres enquêtes menées aux Etats-Unis : en 2006 le *Cooperative Congressional Election Study* et en 2008 le *Cooperative Campaign Analysis Project* (cf. Vavreck et Rivers, 2008). Ce sont de bons exemples de « opt-in » : le plus souvent, il y a une bannière (qui ressemble à une publicité) qui est présente sur une page internet, via des pages d'actualité, des médias etc., sans qu'il y ait de partenariat particulier ni de synergie pour inciter les visiteurs du site où la bannière est 'embarquée' à cliquer sur celle-ci pour ensuite répondre à l'enquête. Outre le biais de couverture déjà mentionné (selon l'accès ou pas à Internet), il y a avec ce mode d'auto-sélection un biais d'exposition (selon les sites où la bannière de pub était présente) et un biais d'auto-sélection (car parmi tous les internautes visitant les pages où la bannière était embarquée, seuls un petit nombre ont eu la curiosité de cliquer et d'aller vers un autre site). Dans cette optique, les internautes entrant en contact avec l'invitation à participer ne sont pas sélectionnés par un institut de sondage, ils choisissent eux-mêmes de rentrer dans le panel.

Cette caractéristique est rédhibitoire aux yeux de certains chercheurs, tandis que d'autres insistent sur la possibilité de tout de même appliqué un plan d'enquête mais en deux temps. Il s'agit d'abord de collecter un très grand nombre de réponses volontaires, mais ensuite de chercher à sélectionner parmi elles, ou à « post stratifier » la population ainsi constituée. Il y a un pan de recherche stimulant (Couper 2000, Rivers 2006) avec des développements statistiques importants en cours. Grâce à des données fines et fiables (telles qu'un recensement de la population) sur la population 'mère' ou 'cible', on construit un « sample framing » et on sélectionne ensuite parmi la masse des répondants volontaires un nombre plus restreint qui correspond à ce profil de population, un peu comme une application de quotas pour sélectionner parmi les centaines de milliers de questionnaires recueillis. Alternativement, outre les opérations classiques de pondérations, on peut aussi envisager le « sample matching » : on administre en parallèle le même questionnaire à un échantillon beaucoup plus petit, mais construit selon des critères de représentativité reconnus assez largement, et le pattern des réponses à l'ensemble du questionnaire va servir de 'benchmark', pas seulement pour du classique calage aux marges.

Il est intéressant d'ailleurs que, dans le monde scientifique anglo-saxon, quand il est question de plan d'enquête ou de son absence, certains conçoivent l'échantillonnage probabiliste comme la méthode la plus 'pure', la post-stratification étant une manière de corriger des biais de volontariat. En France, on peut noter que l'échantillonnage par quotas a été directement utilisé pour recruter et maintenir un panel en ligne pendant quelques mois. Il s'agit de l'enquête Présidoscopie de novembre 2011 à juin 2012, organisée par le CEVIPOF avec le soutien du Monde, de la Fondapol et de la Fondation Jean Jaurès, administré par IPSOS avec Logica Business Consulting. Le recrutement initial de 6 000 individus a été fait au sein de « l'access panel » (ou « panel propriétaire ») de l'institut mais, ensuite, ils n'étaient plus interrogés pour aucune autre enquête électorale pendant le temps de la campagne. Les quotas étaient basés sur l'âge croisé avec le genre, le niveau de diplôme et la profession de la personne de référence, après stratification par région et catégorie d'agglomération. Il y avait des points de fidélité en cadeau pour tous, un bon d'achat électronique pour ceux faisant l'objet d'un entretien individuel semi-directif par téléphone (une dizaine par vague) et un tirage au sort final d'une tombola, prévu pour récompenser 10 participants. 3 309 individus ont participé à l'ensemble des 10 vagues prévues d'emblée et cette opération, perçue comme une démarche de R&D par l'institut, avait été prolongée de 2 vagues, post-présidentielle et post-législative (cf. Chanvril et al, 2013). D'autres développements sont en cours d'élaboration pour un futur panel en ligne 2016-2017 au CEVIPOF, avec une dizaine de

vagues entre décembre 2015 et le printemps 2017, visant d'avoir 10 000 répondants à chacune, mais les choix méthodologiques sont encore en cours, en fonction des réponses des instituts à l'appel d'offre. Il est intéressant en tout cas que, pour mettre en place un panel de moyenne durée (plus court qu'un *British Election Panel Study* sur 5 ans ou la *German Longitudinal Electoral Study* sur 4 ans, mais plus long que les Panels Electoraux Français de 2002 et 2007 et la Présidoscopie 2012), ce soit de nouveau une enquête en ligne qui soit possible.

Quelle que soit la méthode retenue parmi toutes celles présentées ci-dessus, il est important de souligner une chose. Quand il y a recrutement d'internautes volontaires, parfois via le site internet de médias où sont intégrés des bannières qui ne sont pas des publicités mais des appels à participer à une recherche, on pourrait croire – à tort – qu'on en reviendrait à un élément existant dans la 'préhistoire' des enquêtes d'opinion, les fameux « votes de paille ». Or il ne s'agit pas du tout d'une forme moderne qui serait réinventé au moyen d'internet, par rapport aux coupons réponses que les journaux américains diffusaient à la fin du 19^{ème} siècle / début 20^{ème} siècle. Car il ne s'agit en aucun cas de produire des intentions de vote ni de prédire les résultats de l'élection à venir, cela ne sert pas à des publications de tendances pendant le temps de la campagne mais bien à une manière de collecter des informations utiles. Il s'agit d'utiliser le medium internet comme moyen d'entrer en contact avec des répondants potentiels et de les solliciter pour répondre à une (ou plusieurs) enquêtes en ligne sur les élections. Dans la suite de ce papier, nous présenter la spécificité d'opérer cette démarche grâce à une VAA.

2) Une démarche alternative : développer une « Voting Advice Application » comme mode de recrutement de panélistes

a) Présentation rapide des principes d'une VAA

Une VAA consiste à proposer à l'internaute qui l'utilise un outil de repérage de ce qui devrait constituer (dans une logique spatiale du vote) son choix préférentiel au sein de l'offre électorale (partis ou candidats) à un scrutin donné. L'internaute est amené à utiliser la VAA de sa propre initiative et n'est donc pas initialement sollicité par l'équipe de recherche qui développe le dispositif. A partir d'une série de question sur les options qui ont sa préférence au sujet de thèmes traités habituellement par les forces politiques en compétition ou, mieux encore, des thèmes explicitement traités par ces forces dans leur programme pour l'élection en cours, l'internaute répondant est catégorisé en termes de plus ou moins proximité à chacune de ces forces. Divers types de visualisation de la proximité ou de la distance du répondant aux forces politiques peuvent être élaborés. Par ailleurs, l'équipe de recherche qui développe la VAA peut y intégrer divers modules complémentaires permettant de recueillir des données utiles à ses analyses mais qui n'interviennent pas dans le bénéfice promis à l'utilisateur.

Développées de manière relativement récente, les VAA ont suscité une attention soutenue depuis quelques années. Une littérature désormais bien établie leur est consacrée, de même qu'un groupe permanent de l'ECPR. Les difficultés que les chercheurs rencontrent pour recueillir des données grâce aux VAA, pour les exploiter et les interpréter sont ainsi régulièrement discutées. Nous nous inscrivons dans ce contexte en détaillant plus spécifiquement le dispositif de VAA que nous avons développé sous le nom de « La Boussole Présidentielle 2012 » avec des collègues néerlandais (Vrije Universiteit, Amsterdam) concepteurs d'une application de ce type (Kieskompas) en animant une équipe implantée au CEVIPOF lors de la dernière campagne présidentielle et qui a ensuite prolongé ses travaux à

l'occasion d'autres scrutins.

Un premier type de réflexion concerne la stratégie de déploiement du dispositif de VAA pour obtenir le recrutement le plus satisfaisant possible. C'est en quelque sorte la partie visible du dispositif. Un second type de réflexion concerne les enjeux méthodologiques du recueil et de l'exploitation des données recueillies auprès d'une population ainsi recrutée, cette population de répondants ayant notamment pour caractéristique de ne pas constituer un échantillon représentatif d'une population mère.

Le principe d'une VAA repose sur sa capacité à fournir un bénéfice à ses utilisateurs qu'on peut donc espérer recruter gratuitement et en très grand nombre. Le dispositif doit donc, d'une part, être identifié par ses utilisateurs potentiels comme susceptible de leur apporter un bénéfice (promesse) pour déclencher l'usage effectif et, d'autre part, être d'un usage suffisamment satisfaisant pour ne pas provoquer de défection en cours d'utilisation et bien entraîner les utilisateurs jusqu'aux dernières fonctionnalités intéressant le chercheur qui a élaboré le dispositif. Une stratégie assez fine et adaptée aux dispositions et pratiques des internautes doit donc être mise en œuvre.

b) Les enjeux de notoriété et d'attractivité du dispositif pour maximiser la population d'utilisateurs

La taille et la composition du groupe des utilisateurs de la VAA est un enjeu méthodologique qui engage la nature et la fiabilité des analyses qui pourront être menées. Il importe donc de développer une stratégie appropriée pour former un groupe conforme aux exigences de recherche.

A cet égard, il faut assumer que, les utilisateurs de la VAA s'auto-recrutant volontairement, il est quasiment impossible qu'ils forment (comme par miracle statistique) un groupe représentatif de l'ensemble de la population (le corps électoral). En revanche, on peut en diversifier le recrutement pour amoindrir les biais statistiques par rapport aux caractéristiques sociodémographiques vis-à-vis desquelles il est assez habituel dans les analyses politiques de rechercher une représentativité de l'échantillon étudié. Puis, si la diversification est suffisamment assurée, on peut mettre en œuvre différents techniques de redressement pour envisager de mener quelques analyses avec une perspective de généralisation des résultats. Cependant, sans visée de généralisation par représentativité, il restera plus recommandé d'envisager, d'une part, des analyses de « carottage » consistant à observer spécifiquement une catégorie de population pour y déceler telle ou telle disposition et, d'autre part, des analyses par la recherche de contrastes entre catégories spécifiques.

En outre, par un processus de recommandation d'individu à individu, l'auto-sélection peut donner lieu à des effets de grappes, c'est-à-dire des afflux d'utilisateurs qui présentent les mêmes caractéristiques. Cette endogamie préjudiciable à l'objectif de diversification peut involontairement relever de l'appartenance à un même milieu socio-professionnel ou à une même catégorie d'âge mais on peut également redouter qu'elle résulte d'une stratégie volontaire d'investissement de la VAA par un groupe militant qui souhaiterait peser sur la moyenne des résultats.

Face à ces difficultés, l'objectif de diversification aussi bien que la possibilité de traiter statistiquement des sous-catégories spécifiques et, éventuellement, de mener des opérations de redressement, sont d'autant plus envisageables que les utilisateurs sont très nombreux. La volonté de servir le plus grand nombre est également consubstantielle à un dispositif qui se présente comme pourvoyeur d'informations utiles à l'exercice informé du droit de vote. L'utilité sociale de la VAA et la fiabilité des analyses qu'elle doit permettre font donc

pareillement primer l'objectif d'un « très grand nombre » d'utilisateurs. De fait, « La Boussole Présidentielle 2012 » a plutôt bien atteint cet objectif en obtenant près de 1,6 millions de visiteurs s'étant connectés à sa « version 1 » disponible durant les quatre semaines précédant le premier tour et 200 000 à la « version 2 » aménagée pour les deux semaines de campagne du second tour. Par ailleurs, parmi les internautes qui se sont connectés à la version 1, la proportion de ceux qui l'ont utilisée jusqu'à la séquence de publication de leur résultat est assez importantes : ce que nous appelons le taux de conversion (d'utilisateur potentiel à utilisateur effectif) est de 49% (790 000). Après avoir établi divers critères de sélection (usage de la version en Français, usage supérieur à deux minutes, etc.), le nombre d'utilisateurs « utiles » et donc d'unités d'observation est de plus de 600 000 (soit, en moyenne, 20 700 individus par jour). Quelques exigences supplémentaires ramènent finalement les utilisateurs étudiés à 554 000 (19 100 par jour).

Comment en arrive-t-on à une telle performance ? Essentiellement en construisant des partenariats payants avec des médias qui hébergent sur leur site internet un accès à l'application et, selon leur intérêt bien compris, s'efforcent ensuite de générer du trafic vers leur page web dédiée à l'application. Nous avons ainsi noué un partenariat payant avec un quotidien national gratuit, un grand quotidien régional et un regroupement de trois médias nationaux (une chaîne de télévision / une chaîne de radio / un portail internet) associés pour couvrir la campagne présidentielle 2012 avec un site internet commun. Cette diversité de médias a été recherchée pour atteindre des audiences différentes en termes de répartition géographique et de profils sociodémographiques. S'y ajoutent des accès par recommandation sur les réseaux sociaux (notamment Facebook) et plus marginalement des accès par des sites internet des institutions des membres scientifiques du projet. Le volontarisme éditorial des médias partenaires pour promouvoir la VAA ayant été assez contrasté, cela explique des différences importantes dans l'origine des connections des utilisateurs « utiles ». A lui seul, le média le plus impliqué génère ainsi plus de la moitié des connections. Ces contrastes manifestent l'importance crucial du rôle joué par les partenaires pour recruter des utilisateurs. La puissance du média télévision reste très largement avérée car c'est lui qui a généré les records de connections par heure.

A l'aune de ces éléments, il importe de sélectionner des partenaires relativement généralistes et dont la ligne éditoriale est politiquement neutre de telle sorte qu'un éventuel afflux de connections ne soit pas associé à un public ayant une préférence politique spécifique. De manière complémentaire, il apparaîtrait stratégiquement approprié d'inclure à l'avenir un partenaire dont l'audience couvre une ou des catégories de population dont on a pu constater avec « La Boussole présidentielle 2012 » qu'elle était significativement sous-représentée (personnes âgées, ruraux).

b) La nécessité de produire une satisfaction d'usage de l'application en vue du recrutement pour le panel

Une fois connecté, l'internaute doit devenir un « utilisateur », c'est-à-dire ne pas faire défection avant d'avoir pu visualiser le résultat lui fournissant le bénéfice qu'il est censé rechercher (37e page de l'application). Il faut donc que l'ergonomie de l'application soit satisfaisante à l'usage (composition des pages, lettrage, boutons à cliquer) et que les éléments de résultats apparaissent suffisamment conformes aux attentes des utilisateurs. Des indicateurs quantitatifs semblent indiquer un bon niveau de performance en ces domaines. Par exemple, le temps passé sur l'application par l'utilisateur médian est de 7,5 minutes, ce qui est relativement prolongé et donc plutôt indicateur d'une attention soutenue liée à l'utilisation de diverses fonctionnalités optionnelles au-delà des réponses apportées aux 30 questions permettant de situer le répondant par rapport aux candidats en compétition. Un autre

indicateur satisfaisant est le taux de réponse à chacune de ces 30 questions dont certaines sont objectivement plus exigeantes que d'autres : d'une question à l'autre, ce taux oscille entre 97% et 99% des utilisateurs « utiles ».

En revanche, après visualisation du résultat apportant le bénéfice principal attendu par l'utilisateur, seulement 21% (116 000) des 553 400 utilisateurs étudiés ont rempli le questionnaire supplémentaire permettant de compléter la gamme des données dont le chercheur peut faire l'analyse. Cette proportion d'un sur cinq constitue une performance relativement élevée par comparaison avec les résultats obtenus par le même type de VAA dans d'autres pays. La déperdition est encore plus forte pour la fonctionnalité qui conditionne directement la capacité à constituer un panel, c'est-à-dire un vivier de répondants pour de futures enquêtes : seuls 22 400 personnes ont répondu positivement à la demande de laisser leur e-mail pour être recontacté. Là encore, la performance de l'ordre de 4% n'est pas mauvaise par rapport à d'autres VAA équivalentes. Il ne s'agit donc pas d'être défaitiste au sujet des questionnaires supplémentaires et du recrutement des potentiels panélistes mais de bien assumer qu'il y a une forte déperdition des répondants pour les fonctionnalités qui ne relèvent pas du bénéfice qui a pu initialement motiver leur usage de l'application.

Ceci étant dit, plusieurs éléments nous renseignent sur l'hétérogénéité des bénéfices réels que les utilisateurs peuvent rechercher. Alors que les sondages d'opinion indiquent habituellement que la proportion des électeurs indécis diminue au fur et à mesure qu'avance la campagne, nous avons constaté inversement que la proportion des utilisateurs de « La Boussole présidentielle 2012 » déclarant sur la première page de l'application n'avoir pas encore décidé augmentait à l'approche du scrutin : 46% à trois semaines du premier tour, 52% à une semaine du premier tour. D'une certaine manière, compte tenu de l'information que notre VAA apporte, cela correspond à sa vocation première : littéralement, fournir une « aide sur le vote ». On peut donc imaginer que les recommandations d'utilisateurs précédents satisfaits de l'usage qu'ils ont fait de l'application l'ont particulièrement promue auprès de leurs connaissances éprouvant le besoin de stabiliser leur préférence de vote. Ce profil d'utilisateur aurait ainsi constitué progressivement un afflux de plus en plus important. Par ailleurs, les messages de promotion de l'application par les partenaires médias en fin de période ont concentré l'attention sur l'utilité qu'elle avait tout particulièrement pour les électeurs encore indécis, ce qui a très probablement stimulé chez les individus de ce type un intérêt renforcé pour y recourir.

Mais cela manifeste que, de toute façon, une part très importante des utilisateurs ne recherche pas prioritairement un « conseil » concernant lequel des candidats devrait recevoir leur soutien. L'aspect ludique de la comparaison avec les résultats des amis, la curiosité pour la visualisation de son propre positionnement, un certain attrait pour les informations associées à l'application sur les programmes des candidats peuvent aussi entrer en ligne de compte. Tout cela renforce la nécessité de soigner la présentation de toutes les fonctionnalités de l'application. Tout particulièrement, si la fonctionnalité première de l'application qui consiste à fournir une visualisation du positionnement politique de l'utilisateur permet de recueillir ses réponses à 30 enjeux d'opinion, le chercheur a cependant besoin d'autres informations pour pouvoir en développer l'analyse. Il s'agit notamment d'informations signalétiques (conditions de vie, caractéristiques sociodémographiques) pour compléter celles recueillies au tout début de l'usage de la VAA (sexe, âge, niveau de diplôme, commune de résidence). Ce sont aussi des questions d'opinion et de rapport au politique. Selon les modalités habituellement retenues par les développeurs de l'application, nous avons prévu deux questionnaires complémentaires.

Le premier questionnaire supplémentaire était dit « pop up » car surgissant sous forme de fenêtre surimposée à la page en cours de consultation quelques instants après la publication de

la page de visualisation du résultat politique de l'utilisateur. Ce pop up a donné lieu à un taux de retour satisfaisant (les 21% évoqués ci-dessus) mais il n'est pas le questionnaire le plus étendu précisément parce que le choix a été fait de limiter le document à l'espace immédiatement visible dans le cadre de la fenêtre surgissant de façon surimposée à la page en cours de consultation (« pop up »). Le second questionnaire supplémentaire était dit « extra » car il était consultable sur une page à laquelle l'utilisateur n'accédait qu'en cliquant volontairement sur un bouton intitulé « bonus ». Ce questionnaire « extra » était beaucoup plus développé que celui du pop up, ce qui devait fournir un riche ensemble de données, mais il n'a été rempli que par une très faible proportion d'utilisateurs parce qu'il était en fait confiné dans une rubrique mal identifiée et finalement très rarement ouverte. Ainsi, même parmi les 22 400 utilisateurs qui nous ont donné leur e-mail en exprimant leur accord express pour être recontacté à l'occasion de futures enquêtes, c'est-à-dire a priori les personnes les plus enclines à répondre à nos questionnaires, 93% n'ont pas rempli cet « extra ». Il est donc clairement avéré que, si l'équipe de recherche tient à pouvoir traiter les données recueillies via ce type de questionnaire supplémentaire, il faut impérativement l'intégrer de manière plus efficiente dans l'ergonomie de l'application.

c) Entretenir la disponibilité des panélistes volontaires

Dès lors qu'un dispositif de VAA est, d'une part, relayé de manière suffisamment puissante par ses partenaires médias et par les recommandations de ses premiers utilisateurs de bouche à oreille et sur les réseaux sociaux numériques et, d'autre part, doté d'une ergonomie suffisamment incitative pour que les utilisateurs fassent usage des fonctionnalités allant au-delà du bénéfice initialement promis et tout particulièrement acceptent de donner leur e-mail pour être recontactés, alors la stratégie de développement du dispositif en ligne permet de recruter un panel. Ce recrutement impose évidemment un certain nombre de précautions techniques pour s'assurer de l'obligation légale de confidentialité selon les normes établies par la CNIL.

Dans le cas de « La Boussole Présidentielle 2012 » (Vague 1 d'enquête : V1), la première sollicitation suivante auprès de notre panel (V2 : deuxième enquête pour les panélistes qui y répondent) a été menée très rapidement. Elle a consisté à interroger les panélistes recrutés jusqu'au soir du débat télévisé d'entre deux tours dès le lendemain pour traiter notamment de leur rapport à cet événement médiatique de la campagne et dans la perspective du vote du second tour. Disposant d'un vivier de 22 406 individus, nous avons fait le choix de les interroger en deux vagues pour traiter l'hypothèse d'une réception différenciée de ce qui avait pu marquer les publics selon qu'ils répondraient dès le lendemain matin ou le lendemain en soirée après que les commentaires sur le débat se seraient cristallisés et largement diffusés. Nous avons ainsi effectué la première sollicitation à répondre à un questionnaire en ligne auprès 10 867 panélistes à 7h et la seconde sollicitation auprès de 11 539 panélistes à 18h.

Les taux de contact effectif étaient respectivement de 97,5% et de 96,6%, ce qui confirme la qualité du recueil des adresses e-mail même si le léger taux de pertes incite à améliorer cette procédure de recueil en recourant à une « confirmation d'inscription » par retour de mail automatisé comme cela est généralement fait pour la création d'un compte ou la signatures d'une pétition en ligne. Dans un délai de 48h après clôture de la seconde sollicitation, le taux de réponse de chaque vague a été respectivement de 31,4% (3 321 questionnaires individuels remplis) et de 23% (2 649) soit un taux moyen de 27% (5 970). Ce taux moyen est assez satisfaisant mais le différentiel assez marqué entre la population ciblée dans la première sollicitation et la population ciblée dans la seconde sollicitation mérite un commentaire. Une première hypothèse tient à ce que la réception du mail de sollicitation dès la matinée a suscité des réponses en cours de journée selon la disponibilité respective des panélistes, ce qui serait

garant d'une meilleure faisabilité par rapport au mail reçu en soirée quand les activités personnelles et familiales peuvent laisser moins de disponibilité jusqu'à une « péremption » ou faible rémanence de l'effet incitatif du mail à partir du lendemain de sa réception. Ces hypothèses quant au cycle de vie de la sollicitation par mail n'ont pas été testées à partir des variables sociodémographiques susceptibles de l'éclairer et reste à discuter par rapport à la littérature sur les modalités de sollicitation pour des questionnaires auto-administrés, notamment en ligne. Cependant, un facteur inattendu est venu perturber la seconde vague de sollicitation et pourrait aussi contribuer à expliquer son plus faible taux de retour : une erreur d'écriture dans le lien inséré dans le texte du mail et permettant de se connecter au questionnaire en ligne nous a obligé à opérer un second envoi pour cette seconde vague de sollicitation, ce qui fut fait un peu avant 19h après que de multiples panélistes nous aient alertés par retour de mail. Malgré un intitulé explicite (« erratum ») du second mail correctif, il est donc probable qu'un effet de découragement se soit manifesté chez les panélistes de la seconde vague qui auraient tenté de se connecter au questionnaire en ligne en utilisant le lien erroné.

La sollicitation suivante (V3 : troisième enquête pour les panélistes qui y répondent) a été effectuée après les élections législatives de juin 2012 auprès de tous les panélistes qui s'étaient connecté au questionnaire de la V2 et y avaient donné au moins une réponse (6 298 individus dont les 5 970 qui avaient répondu à tout le questionnaire). Auprès de cette population dont nous étions donc assurés par son comportement préalable qu'elle était motivée à se connecter à nos questionnaires, le taux de réponse a été de 42,8% (2 682 questionnaires complets). C'est une performance comparativement élevée par rapport au taux de réponse moyen de V2 (+ 15,8 points) mais qui peut aussi être interprétée comme relativement faible compte tenu de la propension préalablement observée à nous répondre. On peut ici faire l'hypothèse que, sans VAA dédiée à ce scrutin telle qu'elle serait susceptible de relancer l'intérêt pour notre relation avec les panélistes et, d'autre part, s'agissant d'un scrutin qui mobilise moins les électeurs que l'élection présidentielle comme en témoigne le différentiel d'abstention (42,8%, soit 22,3 points de plus), la propension à nous répondre est sensiblement affaiblie.

Par comparaison, les taux de retour auprès de l'ensemble des panélistes recrutés en 2012 et sollicités pour répondre à des questionnaires après les élections municipales de mars 2014 (V4) et après les élections européennes de mai 2014 (V5) sont respectivement de 11,7% et de 14% donc significativement les plus faibles de toutes nos vagues d'enquête. La V4 ayant été conduite deux semaines après le scrutin, il est fort probable que ce délai a été défavorable, surtout en l'absence de VAA dédiée. La V5 a été conduite dans les trois jours suivant le scrutin et a disposé d'une VAA dédiée (« La Boussole européenne 2014 ») mais dont la visibilité et donc l'attractivité sont restées faibles (un peu moins de 90 000 utilisateurs). En revanche, parmi les utilisateurs de « La Boussole européenne 2014 », 2 300 (soit 5,5%) ont donné leur e-mail et ainsi intégré notre panel et, parmi eux, le taux de réponse à la V5 est de 37% (852 questionnaires individuels finalisés) soit 23 points de plus que parmi les panélistes recrutés en 2012. Tout en admettant que les élections européennes sont peu mobilisatrices et que les utilisateurs de notre VAA sont donc inversement sur-mobilisés par rapport au reste de la population, cela ne suffit pas à expliquer le différentiel de taux de réponse avec les panélistes recrutés en 2012, d'autant que ces derniers étaient plutôt plus intéressés par la politique que la population générale telle qu'elle est régulièrement interrogée à ce sujet dans les enquêtes par sondage sur des échantillons représentatifs.

Une hypothèse complémentaire et qui compte probablement pour une large part de l'explication tient à la « fraîcheur » du rapport à notre démarche d'enquête : celle-ci tendrait à s'atténuer avec le temps et ne maintiendrait donc pas un taux de réponse aussi élevé d'une vague d'interrogation du panel à l'autre. Par contraste, les panélistes recrutés le plus

récemment éprouveraient un intérêt et une motivation pour notre démarche qui est avivée par un sens de la nouveauté. Si nous avons l'avantage de recruter notre panel et de le maintenir sans contrepartie financière ni fourniture d'avantages commercial, ce qui permet de développer un appareil de recueil de données avec une grande parcimonie de moyens, un enjeu important reste la capacité à entretenir la motivation et la disponibilité des panélistes.