

**CORRECTION DES EFFETS DE MODE ET BIAIS DE SÉLECTION :
LES APPORTS D'UNE EXPERIMENTATION DE L'ENQUETE TIC (TECHNOLOGIES DE L'INFORMATION ET DE LA
COMMUNICATION) EN FACE-A-FACE**

Stéphane LEGLEYE^{1,2}, Louise VIARD-GUILLOT¹ & Amandine NOUGARET¹

1 : Insee, Département des ressources et conditions de vie des ménages

2 : CESP, Inserm

stephane.legleye@insee.fr
louise.viard-guillot@insee.fr
amandine.nougaret@insee.fr

Mots-clés : Échantillonnage, collecte, multimode, effet de mode

Domaine concerné : Collecte, théorie des sondages aval

Résumé

L'enquête TIC auprès des ménages interroge chaque année environ 30 000 individus sur leur équipement et leurs usages en matière d'Internet, d'informatique et de téléphonie fixe et mobile. Son protocole repose sur une collecte séquentielle et stratifiée par téléphone et par Internet/papier.

Son questionnaire ne comporte aucune question véritablement sensible. En revanche, Internet étant à la fois un des thèmes et un des modes de collecte de l'enquête classique, on peut suspecter un biais de sélection. Certains individus sélectionnés pour répondre par Internet pourraient en effet avoir tendance à ignorer l'enquête lorsqu'ils sont peu à l'aise avec cet outil (bien qu'une solution de réponse par papier leur soit proposée) ou lorsque le sujet de l'enquête ne les mobilise pas. En effet, la précédente édition avait permis de montrer qu'environ 15 % de la population de 15 ans et plus n'utilisait pas ou très peu Internet et que 38 % des usagers dans l'année avait un manque de compétence dans le domaine, 17 % de la population pouvant ainsi être considérée en état d'illectronisme¹ (Legleye and Rolland 2019).

Or, si 45 % du sous-échantillon interrogé par Internet/papier dispose d'une adresse courriel, signe d'une potentielle utilisation d'Internet, ce n'est pas le cas des 55 % restants, soit environ 13 % de la population générale, a priori très peu utilisatrice d'Internet. Même si la possibilité de répondre par papier est mentionnée dès la première lettre-avis, remplir le questionnaire papier exige davantage d'efforts (les filtres doivent être appliqués par le répondant, les instructions doivent être lues avec attention, le questionnaire doit être posté...) : cela concourt à diminuer le taux de réponse. De plus, les ménages du sous-échantillon Internet/papier sont plus pauvres que les autres alors que les taux de collecte sont très bas dans ce sous-échantillon : 39,3 % en 2021, dont 46 % pour les logements avec une adresse courriel et 34 % pour ceux qui en sont dépourvus.

Enfin, en plus du biais de non-réponse évoqué ci-dessus, la passation via le papier ou Internet ne permet pas de s'assurer de l'identité du répondant. Or l'enquête TIC repose sur le tirage d'un kish – individus de 15 ans ou plus du ménage dont la date de naissance est la plus proche du 1^{er} juin. On suspecte qu'au sein des ménages participant en auto-administré, la personne la plus intéressée par le sujet ou la plus à l'aise sur Internet, répondrait

¹ L'illectronisme désigne le fait de ne posséder aucune des quatre compétences numériques de base (savoir rechercher des informations, communiquer en ligne, résoudre des problèmes informatiques et utiliser des logiciels) ou de ne pas se servir d'Internet (incapacité ou impossibilité matérielle).

en lieu et place de la personne Kish, ce qui constitue un second biais de sélection. Ce dernier fait l'objet d'un traitement en standard à l'aval. Toutefois, il s'appuie sur une comparaison des sous-échantillons méthodologique et téléphonique, donc sur les sous-échantillons de ménages ayant un numéro de téléphone disponible et dont les caractéristiques sociodémographiques sont dissemblables de celles des individus de l'échantillon Internet sans numéro de téléphone.

*Au quatrième trimestre 2021 a eu lieu un **test de l'enquête TIC en face-à-face**, auprès d'un échantillon de 5 000 individus. L'objectif de cette expérimentation était de tester la présence de biais dans l'enquête TIC classique. La collecte face-à-face bénéficie en effet d'un meilleur taux de collecte, en particulier pour les ménages sans téléphone, et offre échantillon de contrôle monomode pour tous les sous-échantillons qui n'interfère pas avec la thématique de l'enquête et serait donc a priori moins biaisé que l'échantillon classique.*

Les différents échantillons de répondants à l'enquête classique ont été comparés à chaque étape du redressement avec l'échantillon de l'enquête face-à-face, afin de mettre en évidence le mécanisme global de réponse à l'enquête. Puis la comparaison des indicateurs d'intérêt obtenus dans l'enquête classique avec ceux obtenus en face-à-face (après les redressements des deux enquêtes) renseigne sur la qualité de l'enquête classique et de son redressement.

Abstract en anglais

The ICT household survey interviews approximately 30,000 individuals each year by telephone and Internet/paper. The ICT survey is a reference in France for the monitoring of the household Internet access rate, the proportion of Internet users and of people with digital illiteracy. Because it uses the Internet and digital communication tools to ask about their use, it is natural to suspect the existence of self-selection bias among respondents in the survey. In the fourth quarter of 2021, a face-to-face test of the ICT has been conducted, with a sample of 5,000 individuals. The objective of this experiment was to test the presence of bias in the classic ICT survey. The different samples of respondents to the conventional survey were compared at each stage of adjustment with the sample from the face-to-face survey, in order to highlight the overall response mechanism of the survey. Then the comparison of the indicators of interest obtained in the conventional survey with those obtained in the face-to-face survey (after the adjustments of both surveys) had provided information on the quality of the conventional survey and its adjustment.

1. Introduction

1.1. L'enquête TIC classique

L'enquête TIC auprès des ménages est une des plus anciennes enquêtes ménages multimode mise en place à l'Insee. Elle interroge chaque année environ 30 000 individus sur leur équipement et leurs usages en matière d'Internet, d'informatique et de téléphonie fixe et mobile. Son protocole repose sur une collecte par téléphone et par Internet/papier. Le champ de l'enquête est composé des personnes de 15 ans ou plus vivant dans un logement ordinaire, en France métropolitaine. Une personne par ménage est interrogée ; elle répond aux questions de niveau ménage et individu. L'objectif du dispositif TIC est de collecter des indicateurs décrivant l'équipement des ménages et les usages des individus dans le domaine des technologies de l'information (informatique, Internet, téléphone fixe et mobile), de façon à satisfaire aux demandes d'Eurostat. Dans TIC classique, on trouve un sous-échantillon de ménages dont on connaît un numéro de téléphone (parce que les personnes en ont communiqué un à l'administration fiscale ou parce qu'il figure dans l'annuaire) et un sous-échantillon de ménages sans numéro connu. Le premier est divisé en deux : un échantillon téléphonique qui doit répondre au téléphone (et qui peut répondre par Internet en cas de non-réponse, suivant un protocole séquentiel) ; un échantillon méthodologique qui doit répondre par Internet (et n'est jamais relancé par téléphone, suivant un protocole monomode). Enfin les ménages sans numéro connu doivent répondre par Internet ou Papier. Tous les contacts initiaux se font par courrier postal et envoi de courriels d'information doublant les lettres (d'invitation ou de rappel) : 80 % de la base de sondage dispose d'une adresse courriel.

L'échantillon téléphonique dispose du meilleur taux de collecte (voir infra) ce qui lui permet d'être considéré comme la référence pour les estimations d'effets de mode que l'on suspecte pour les échantillons interrogés par Internet (échantillon méthodologique) et par Internet / papier.

Les échantillons méthodologiques et Internet/papier sont en effet comparés à l'échantillon téléphone, après calage des premiers sur la structure socio-démographique du dernier, pour les estimations des taux de ménages équipés d'Internet et d'usagers d'Internet au cours des trois derniers mois. La comparaison est faite par tranche d'âge : les proportions de ménages équipés et d'usagers dans les trois derniers mois observées dans l'échantillon téléphone sont alors utilisées comme nouvelles marges pour redresser les autres échantillons par calage.

1.2. L'enquête TIC face-à-face

L'enquête TIC face-à-face a eu lieu du 11 octobre au 11 décembre 2021 auprès de 5 000 individus tirés en juin 2021, dans les Fichiers démographiques sur les logements et les individus 2020 (Fideli 2020). Le sondage est stratifié et sur-représente les ménages sans numéro de téléphone dans la base de sondage : 34 % des unités tirées doivent n'avoir aucun numéro connu dans Fideli ; les 66 % restant doivent avoir un numéro. Le champ de l'enquête et le questionnaire du test méthodologique en face-à-face étaient les mêmes que ceux utilisés pour l'enquête TIC 2021 "classique". Une seule question a été rajoutée, dans le premier module sur l'équipement du ménage en Internet : il s'agissait de savoir si le ménage avait accès à Internet à son domicile en avril 2021. Les réponses à cette question permettent de réaliser de meilleures comparaisons avec les résultats obtenus par la collecte multimode effectuée entre avril et juin 2021, dans un contexte de potentielle croissance du taux d'équipement durant la pandémie.

1.3. Biais suspectés dans l'enquête classique

La possibilité de répondre au téléphone n'existe pas dans le sous-échantillon Internet/Papier (sans numéro connu). Si 45 % de ce sous-échantillon dispose d'une adresse courriel, signe d'une potentielle utilisation d'Internet, ce n'est pas le cas des 55 % restants, soit environ 13 % de la population générale, qui est donc a priori très peu utilisatrice d'Internet. Dans l'enquête classique, nous avons pris la précaution d'annoncer dès la lettre-annonce qu'à la suite de la sollicitation initiale de réponse via Internet, le questionnaire papier serait envoyé en cas de non-réponse. Toutefois, remplir le questionnaire papier exige davantage d'efforts (les filtres doivent être appliqués par le répondant, il n'y a pas de contrôle de cohérence et les instructions doivent être lues avec attention, etc.) ; puis le questionnaire doit être posté, ce qui requiert une action supplémentaire. Ces deux aspects concourent à diminuer le taux de réponse. De plus, les ménages du sous-échantillon Internet-papier sont plus pauvres que les autres : 40 % des ménages de ce sous-échantillon gagnent moins de 20 000 euros par an, contre respectivement 25 % et 31 % pour les échantillons téléphonique et méthodologique. Leur capacité à se mobiliser pour l'enquête pourrait être réduite. Et de fait, les taux de collecte sont très bas dans ce sous-échantillon (39,3 % en 2021, dont 46 % pour les logements avec une adresse courriel et 34 % pour ceux qui en sont dépourvus, contre respectivement 53,5 % et 73,5 % pour les sous-échantillons méthodologique et téléphonique). Certains individus sélectionnés pourraient avoir tendance à ignorer l'enquête lorsqu'ils sont sollicités pour répondre par Internet et sont peu à l'aise avec cet outil ou lorsque le sujet de l'enquête ne les mobilise pas (cela est obvie pour l'équipement : les personnes non équipées ne peuvent répondre par Internet). Dans ce cas, les différences que l'on pourrait retrouver toutes choses égales par ailleurs entre des répondants Internet et face-à-face devraient être imputables à une sélection non-ignorable, c'est-à-dire à une autosélection

des répondants sur Internet en fonction de leurs équipements, pratiques et compétences en Internet et en informatique (Castell and Sillard 2021).

Enfin, en plus du biais de non-réponse totale évoqué ci-dessus (qui est un biais de sélection classique), la passation via le papier ou Internet ne permet pas de s'assurer de l'identité du répondant. On suspecte ainsi qu'au sein des ménages participant en auto-administré, la personne la plus intéressée par le sujet des TIC ou la plus à l'aise sur Internet répondrait en lieu et place de la personne Kish, ce qui constitue un second biais de sélection (qui n'est pas de non-réponse).

Enfin, il n'est pas totalement possible d'ignorer le risque de biais de mesure, ici essentiellement un biais de désirabilité sociale (Kreuter, Presser et al. 2008). La présence d'un enquêteur pourrait en effet, malgré le caractère apparemment anodin des questions de l'enquête, inciter les gens à se présenter comme plus souvent équipés qu'ils ne le perçoivent, plus souvent utilisateurs quotidiens d'Internet ou plus compétents qu'ils ne sont en réalité. Deux mécanismes peuvent jouer. Comme le numérique et Internet ont conquis toutes les sphères de l'existence et que l'aisance et les compétences dans le domaine sont valorisées socialement, cela peut passer par une volonté directe de valorisation de soi de la part des répondants, qui choisiraient d'eux-mêmes les modalités de réponse les plus avantageuses. On peut toutefois imaginer que le mécanisme puisse en fait provenir activement des enquêteurs eux-mêmes, au moins pour certaines catégories d'enquêtés : en face de certaines catégories de personnes en apparence peu connectées et versées dans les nouvelles technologies, ils pourraient fournir des incitations à compter une connexion Internet qui ne sert jamais comme existante et active, à considérer que les usages d'Internet sont plus réguliers que ce qu'en perçoivent ou déclarent spontanément les répondants et enfin à aider les personnes à répondre au questionnaire traitant des compétences informatiques et communicationnelles, ce qui les classera comme non illettrés. Il s'agit d'un biais de désirabilité sociale de la part de l'enquêteur.

Il est finalement possible que les répondants des modes auto-administrés connaissent des difficultés plus grandes pour répondre. Si le questionnement n'est pas sensible, il n'en est pas moins technique, soit sur des appareils et dispositifs physiques, soit sur des logiciels et applications. Le questionnaire est donc assorti de nombreuses consignes que l'enquêté est censé lire, comprendre et appliquer. La lecture et la compréhension du questionnaire est donc une charge cognitive qui peut générer lassitude et sentiment d'incompétence. Ce transfert classique de compétences et de charge de l'enquêteur vers l'enquêté, classique, pourrait ainsi générer sur certaines questions des imprécisions et des erreurs qui n'ont pas cours en situation d'intermédiation. Les réponses choisies par l'enquêté en mode auto-administré auront ainsi plus de chances d'être imprécises, systématiques ou approximatives. L'enquêté se contente alors de réponses non optimales : il se satisfait de réponses qu'il juge suffisantes eu égard à la difficulté des questions et de son intérêt pour la problématique (Krosnick, Narayan et al. 1996). Ce biais de *satisficing* a été clairement démontré dans l'expérimentation CVS panel (Castell, Clerc et al. 2021).

Peut-on distinguer les biais de sélection et de mesure ? On peut penser que le biais de désirabilité sociale est relativement constant en intensité en cas d'intervention d'un enquêteur, donc qu'il ne diffère pas dans les échantillons Téléphone et Face-à-face (relativement à un même mode de collecte, Internet, par exemple) : la comparaison de ces échantillons toutes choses égales par ailleurs devrait ainsi permettre d'exhiber essentiellement un biais de sélection. De la même manière, les répondants Internet de l'échantillon Internet / papier peuvent être comparés aux répondants Internet de l'échantillon méthodologique : interrogés sur le même mode, leur comparaison toutes choses égales par ailleurs devrait refléter un biais de sélection. En revanche, dès lors que l'on compare un échantillon de répondants issu d'un mode auto-administré à un autre issu d'un mode intermédié, le risque de confusion des deux est important.

1.4. Apports de l'enquête face-à-face

Le biais de sélection dans TIC est donc susceptible d'être de nature non-ignorable dans la mesure où ce sont bien les caractéristiques que l'on cherche à estimer qui pourraient causer une large partie de la non-réponse. Au-delà des considérations statistiques propres aux techniques classiques de redressement (Little and Vartivarian 2005, Haziza and Lesage 2016), disposer d'un échantillon de contrôle permet d'apporter une réponse directe à la question de la fiabilité des estimations. C'est ce que permet la collecte en face-à-face. Elle est d'une part censée obtenir un meilleur taux de collecte ainsi qu'une meilleure représentativité, en particulier des ménages pauvres et sans coordonnées de contact dans Fideli. D'autre part, le fait qu'elle soit monomode et en présentiel permet d'assurer la sélection du répondant suivant les critères Kish ordinaires, ce qui n'est pas assuré pour les collectes auto-administrées. La comparaison avec l'enquête en face-à-face permettra de vérifier si globalement et pour des sous-populations particulières, la procédure de redressement actuelle de l'enquête multimode classique assure bien ce travail.

1.5. Objectifs

Nous décrivons les principaux éléments des bilans de collecte : taux de réponse, représentativité des échantillons et sous-échantillons de répondants de l'enquête classique et les comparons à ceux calculés pour les répondants de l'enquête face-à-face. Dans un second temps, nous procédons à une comparaison des niveaux de trois variables d'intérêt de l'enquête (ou « outcomes ») : équipement Internet du ménage, usage quotidien au cours des trois derniers mois et illettrisme. Les deux premiers font l'objet d'une correction des effets de mode dans l'enquête classique, mais pas le troisième.

Les différents échantillons de répondants à l'enquête classique sont comparés à chaque étape du redressement avec l'échantillon de l'enquête face-à-face, afin de mettre en évidence le mécanisme global de réponse à l'enquête. La comparaison des indicateurs d'intérêt obtenus dans l'enquête classique avec ceux obtenus en face-à-face (après les redressements des deux enquêtes) renseigne sur la qualité de l'enquête classique et de son redressement. Dans un dernier temps, nous identifions quelques variables de l'enquête qui influent sur l'estimation des outcomes.

2. Méthodes

2.1. Représentativité des échantillons

La représentativité des échantillons est exprimée en distances standardisées à chacune des bases de sondage et au moyen d'indicateurs R (Bethlehem, Cobben et al. 2009). Pour une indicatrice mesurée parmi les répondants (R) et dans la base de sondage (BS) la formule est la suivante :

$$d_i = \frac{100 * (p_R - p_{BS})}{\sqrt{\frac{p_R(1 - p_R) + p_{BS}(1 - p_{BS})}{2}}}$$

Pour une variable catégorielle, nous avons choisi de prendre la moyenne des valeurs absolues des différences calculées sur l'ensemble des modalités, afin de simplifier les tableaux.

Pour une variable continue ou un taux, la formule est :

$$d = \frac{100 * (m_R - m_{BS})}{\sqrt{\frac{s_R^2 + s_{BS}^2}{2}}}$$

Les distances standardisées sont communément utilisées pour juger de la proximité d'échantillons indépendants du point de vue de la distribution de variables clefs. Elles permettent de se concentrer sur l'importance des écarts de distribution en tenant compte de l'imprécision des mesures observées mais en s'affranchissant d'un test statistique dont la puissance est dépendante de la taille de l'échantillon (qui conduit à ne jamais considérer comme significatifs des écarts forts sur des petits échantillons et à l'inverse à toujours considérer significatifs des écarts faibles sur de grands échantillons). Il n'y a pas de seuil universellement reconnu pour juger de la signification d'une distance standardisée, mais la littérature suggère de retenir 20 comme indication pertinente d'un déséquilibre important (Rubin 2007, Austin and Stuart 2015).

Les distances standardisées permettent d'apprécier les écarts de distribution modalité par modalité ou variable par variable (dichotomique ou continue) mais n'offrent pas de perspective multivariée. C'est ce que fait l'indicateur R qui propose d'évaluer non pas directement la proximité de distribution des marges de l'échantillon de répondants et de la base de sondage, mais la variabilité de la probabilité de répondre conditionnellement à des variables clefs (souvent les marges des bases et échantillons). Plus la variabilité est élevée, moins la réponse est uniforme, plus des problèmes de « représentativité » (sur et sous-représentations relatives) sont à suspecter. Le calcul repose sur une modélisation de la probabilité de répondre conditionnellement aux variables clefs que l'on suppose influencer sur celle-ci (les variables du plan de sondage, les principales marges sociodémographiques des bases et échantillons et plus généralement toutes variables pertinentes connues pour les répondants et non-répondants). L'indicateur se calcule ainsi : $R = 1 - 2Std(ps)$, où ps est le score de propension à répondre à l'enquête, soit la probabilité prédite estimée par le modèle et Std son écart-type (Bethlehem, Cobben et al. 2009).

2.2. Variables sociodémographiques

Les variables de la base de sondage communes aux deux enquêtes sont en nombre relativement réduit, du fait du changement d'algorithme de tirage des échantillons intervenu entre les deux enquêtes et du délai très court pour le tirage de l'enquête Face-à-face. On en compte 11 principales : sexe et âge (en continu) de la personne de référence du ménage, taille du ménage (1, 2 ou 3, 4 ou 5, 6 et plus), taille d'unité urbaine (en 5 catégories, de rural à l'agglomération de Paris), revenu fiscal du ménage (en quatre tranches), type de logement (maison, autre), indicatrice de ménage pauvre financièrement, indicatrice de ménage résidant en quartier prioritaire de la ville. À ces variables sociodémographiques s'ajoutent deux variables de coordonnées de contact présentes dans la base

de sondage : la présence d'un numéro de téléphone pour le ménage (fixe ou mobile, présent dans Fideli ou dans l'annuaire téléphonique) ; la présence d'un courriel dans Fideli.

Pour ce qui est des caractéristiques des répondants, nous avons retenu l'âge, le sexe, le niveau de diplôme (4 catégories), les équipements en téléphone fixe et smartphone (deux variables binaires). Nous avons laissé de côté les pratiques de filtrage ou de refus systématique de prise d'appel téléphonique interrogées dans l'enquête, car ces variables nous semblent susceptibles d'être affectées d'un biais de mesure suivant que les questions afférentes sont posées par un enquêteur ou bien auto-administrées et suivant que l'interrogation a notamment lieu au téléphone ou non.

2.3. Variables d'intérêt (outcomes)

Les variables d'intérêt étudiées sont au nombre de trois (toutes binaires) : équipement Internet du ménage, usage d'Internet au cours des trois derniers mois et illettrisme.

2.4. Analyses multivariées

Les analyses multivariées choisies sont des modèles de Cox. Ces modèles sont originellement destinés à des analyses longitudinales et produisent des hazard ratio, ou ratios de risque instantanés. Toutefois, appliqués à des enquêtes transversales, en fixant arbitrairement la survenue de l'événement d'intérêt à un moment ou « âge » identique pour tous les individus, les hazard ratios deviennent des ratios de risque. Ces derniers sont préférables aux odds ratios pour leur simplicité d'interprétation (Diaz-Quijano 2012). Nous avons choisi une modélisation directe de l'outcome pour estimer les effets de l'interrogation sur les différents modes de collecte plutôt qu'une méthode de repondération par l'inverse du score de propension (éventuellement suivie par une analyse multivariée). En effet, nous avons considéré que les principaux facteurs de confusion (variables affectant la non-réponse à chacun des modes de collecte ainsi que les outcomes) étaient disponibles dans les données des bases de sondage ou les questionnaires (notamment les équipements téléphoniques). Autrement dit, nous privilégions l'approche de Pearl (Pearl 2009) à celle de Rubin (Imbens and Rubin 2015) (les deux approches devant donner des résultats proches lorsque les conditions de validité sont réunies). Le lecteur intéressé pourra trouver des compléments et une approche assez complète mais non technique dans (Hernan and Robins 2020).

3. Résultats

3.1. Taux de réponse des enquêtes

Le face-à-face a permis d'obtenir un meilleur taux de collecte (62 %) que l'enquête classique (54 %). Il faut noter que le taux du face-à-face est sous-estimé relativement à celui du téléphone. En effet, il est possible d'interroger un ménage ayant déménagé par téléphone ou Internet ou courrier, mais pas en face-à-face. Il est plus vraisemblable que le taux réel du face-à-face approche ou dépasse les 64 % (si on retient qu'environ 12 % des ménages ayant changé d'adresse n'ont pu pour cette raison être interrogés en face-à-face, mais qu'on peut estimer que le taux de collecte en face-à-face pour eux pourrait atteindre celui du téléphone, soit 75 %)².

Cette supériorité du face-à-face est marquée pour les ménages sans téléphone (47 % vs 41 %). Pour les ménages avec numéro, le taux atteint 57 % dans l'enquête classique contre 70 % dans l'enquête face-à-face. Dans l'enquête classique, les ménages avec numéros pouvaient être interrogés par Internet (échantillon méthodologique) ou bien directement au téléphone. Le taux de collecte des premiers atteint 54 % mais celui des deuxièmes atteint 75 %, soulignant clairement la supériorité du mode intermédiaire pour cette sous-population (par ailleurs largement majoritaire).

D'emblée, le face-à-face fait donc mieux que le multimode pour atteindre et faire participer des personnes n'ayant pas laissé de numéro dans Fideli. En revanche, il fait moins bien que le téléphone pour les ménages dont un numéro est connu.

3.2. Représentativité sociodémographique des échantillons de répondants

Calculé à partir des variables de la base de sondage, l'indicateur R pour l'enquête classique vaut 0,696 ; celui pour l'enquête Face-à-face 0,706, suggérant un léger avantage de cette dernière. Le détail (Tableau 1) montre que l'équilibrage est meilleur pour 6 variables sur 11 (signalé par une distance standardisée plus réduite en face-à-face). C'est le cas pour les ménages pauvres ($d=-7,9$), les revenus ($d=-2,0$), pour l'âge de la personne de référence du ménage ($d=-5,2$), la résidence en QPV ($d=-2,7$), la fourniture d'un courriel à l'administration fiscale ($d=-4,0$) et enfin le sexe de la personne de référence ($d=-0,9$).

² Environ 12 % des ménages ont changé d'adresse. Il est possible de réestimer le taux de réponse à plus de 72 % si on applique le taux de réponse de l'échantillon téléphonique à la portion des ménages disposant d'un numéro (ce qui est sans aucun doute encore une sous-estimation car le face-à-face fait pratiquement toujours mieux que le téléphone).

Le face-à-face a également plus de succès dans la représentation des répondants sans numéro ni courriel connu (6,6 % dans la base de TIC classique et 10,7 % dans celle de TIC Face-à-face qui les surreprésentait quelque peu) : $d=-2,2$.

Tableau 1 : valeurs absolues des distances standardisées aux bases de sondage des échantillons de répondants des enquêtes Face-à-face et Classique

	Classique	FàF	Écart
Âge personne de référence	17.7	12.5	-5.2
Taille d'unité urbaine	1.6	6.1	4.5
Sexe pers. de référence	6.8	5.8	-0.9
Type de logement	14.2	17.6	3.3
Revenus (tranches)	8.2	6.2	-2.0
I_Pauvre	11.1	3.2	-7.9
Région	1.0	2.1	1.1
I_QPV	8.0	5.4	-2.7
Nb pers. ménage	4.4	6.9	2.5
Num téléphone (Fideli/annuaire)	13.7	17.7	4.1
Courriel (Fideli)	13.4	12.7	-0.7
Ni Tel ni Courriel	14.9	12.7	-2.2

Les données sont pondérées par les poids de sondage.

Plus la valeur est élevée, plus la distribution de la variable dans l'échantillon de répondants est éloignée de celle de son échantillon d'origine. Un écart négatif signe une meilleure représentation de l'enquête face-à-face.

En opérant ces comparaisons par coordonnées de contact (Tableau 2), on confirme le constat global : le face-à-face fait légèrement mieux que TIC classique pour interroger ces populations, en particulier les ménages sans coordonnées, *a priori* les plus difficiles à interroger. Néanmoins, les indicateurs R montrent des résultats multivariés mitigés : pour les ménages avec numéro, $R=0,771$ dans TIC classique contre $R=0,781$ dans le FaF ; pour les ménages sans numéro connu, $R=0,773$ dans TIC classique contre $R=0,704$ dans le FaF. Enfin pour les ménages sans numéro ni courriel connu, $R=0,788$ dans le classique et $R=0,687$ en Face-à-face.

Le face-à-face fait donc globalement un peu mieux que l'enquête multimode classique pour interroger les personnes *a priori* difficiles à joindre (parce que leurs coordonnées de contact nous sont inconnues) ou celles disposant de faibles revenus. Toutefois, il ne fait pas nettement mieux en termes de représentativité globale de celles-ci. Les écarts sont par exemple très importants pour l'âge de la personne de référence du ménage, où le face-à-face se révèle « moins bon » que l'enquête classique.

Tableau 2 : Différences des valeurs absolues des distances standardisées aux bases de sondage entre les échantillons de répondants des enquêtes Face-à-face et Classique pour trois sous-populations

	Avec numéro	Sans numéro	Sans numéro ni courriel
Âge personne de référence	-7.6	18.7	7.4
Taille d'unité urbaine	3.4	5.1	1.7
Sexe pers. de référence	-2.3	-2.5	-2.4
Type de logement	1.3	4.3	-9.5
Revenus (tranches)	-2.8	-4.9	-10.0
I_Pauvre	-6.6	-5.9	3.4
Région	0.7	2.4	1.2
I_QPV	-2.3	-5.8	-4.1
Nb pers. ménage	-0.9	0.2	-5.7
Courriel (Fideli)	-1.0	-13.9	
Moyenne D=	-1.8	-0.2	-2.0

Les données sont pondérées par les poids de sondage.

Note : Une valeur négative (resp. positive) signe une meilleure (resp. plus mauvaise) représentation de la catégorie dans l'enquête face-à-face.

Exemple : pour les revenus, la représentativité de l'échantillon de répondants du Face-à-face (relativement à son échantillon tiré) est meilleure que celle du téléphone pour les ménages sans numéro ni courriel ($d=-10,0$).

3.3. Comparaison des enquêtes par sous-échantillon

Les répondants de l'échantillon téléphonique sont très proches de ceux de l'échantillon méthodologique (Tableau 3). Ils sont plus jeunes (ce qui est surprenant : $d>10$) mais sont moins diplômés ($d<10$) et déclarent plus souvent posséder un smartphone ($d>10$).

Les répondants du Face-à-face appartenant à des ménages dont un numéro de téléphone est connu sont plus vieux que ceux de l'échantillon téléphonique ($d>10$), plus diplômés ($d>10$) mais possèdent des revenus moins élevés ($d<10$). La comparabilité de ces deux échantillons est très bonne, et en particulier, on ne note pas d'écart important concernant les déclarations d'équipement en téléphone fixe ou en smartphone.

Enfin, si l'on compare ces répondants du face-à-face (pour lesquels on dispose d'un numéro de téléphone) à leurs homologues de l'échantillon méthodologique (dont on connaît un numéro mais qui ont répondu par Internet), ils apparaissent disposer de revenus plus faibles ($d<10$), mais déclarent cependant plus souvent être équipés d'un smartphone ($d>10$).

Ces écarts perdurent quelle que soit la pondération : la correction des effets de mode opérée entre la pondération intermédiaire et finale ne modifie que peu ces constats relativement à la pondération corrigée de la non-réponse totale. Cela se voit sur la plupart des variables ainsi que dans le calcul des indicateurs R.

La même analyse peut être faite à propos de la comparaison des répondants de l'échantillon Internet/papier et du téléphone ou du face-à-face (Tableau 4). Les premiers sont redressés pour ressembler aux seconds du point de vue des taux d'équipement Internet du ménage et du taux d'usage quotidien d'Internet durant les trois derniers mois. Toutefois les répondants téléphoniques apparaissent toujours plus jeunes, posséder des revenus plus élevés, provenir d'un ménage plus nombreux et plus souvent posséder un smartphone ou un téléphone fixe. Ils sont aussi plus nombreux à avoir fourni une adresse courriel à l'administration fiscale.

Tableau 3 : Équilibrage des échantillons de répondants (méthodologique, téléphone et face-à-face **avec** numéro de téléphone)

	Méthodologique			Téléphonique			Face-à-face			Meth_tel			Meth_Faf			Tel_Faf		
	S	I	F	S	I	F	S	I	F	S	I	F	S	I	F	S	I	F
Âge (moy.)	56.1	50.2	50.2	53.4	45.2	45.0	53.1	49.1	49.1	-10.5	-24.9	-25.7	-17.9	-6.0	-5.8	-1.3	19.0	20.0
Hommes	50.9%	48.3%	48.2%	47.0%	47.5%	47.5%	47.1%	48.5%	48.5%	-7.9	-1.6	-1.3	-7.6	0.5	0.6	0.3	2.0	2.0
I_pauvre	8.8%	11.4%	11.5%	10.9%	14.0%	14.0%	10.7%	11.9%	11.9%	7.0	7.9	7.5	6.6	1.8	1.5	-0.4	-6.1	-6.0
I_QPV	4.3%	5.9%	6.0%	5.2%	6.6%	6.6%	5.2%	6.4%	6.4%	4.1	2.8	2.2	4.2	2.0	1.4	0.1	-0.8	-0.8
Revenus (moy. 1-4)	2.5	2.4	2.4	2.4	2.5	2.5	2.3	2.3	2.3	-3.7	2.3	3.5	-26.7	-12.7	-12.2	-12.2	-14.9	-15.6
Nb pers. ménage	2.1	2.4	2.4	2.1	2.5	2.5	2.2	2.5	2.5	-2.8	6.4	6.4	3.0	6.5	6.4	4.4	0.2	0.0
Courriel	84.8%	85.7%	85.0%	83.5%	84.5%	84.9%	85.8%	87.3%	87.3%	-3.6	-3.4	-0.1	2.9	4.7	6.7	6.5	8.1	6.8
Tel. fixe	80.8%	79.0%	78.8%	80.8%	81.0%	80.9%	77.2%	77.6%	77.6%	-0.1	5.0	5.2	-9.0	-3.6	-3.1	-8.9	-8.6	-8.3
Smartphone	73.6%	77.8%	77.1%	82.8%	84.6%	84.8%	82.3%	83.7%	83.7%	22.4	17.5	19.9	21.1	14.9	16.7	-1.4	-2.6	-3.2
Diplôme (moy. 1-4)	2.8	2.7	2.7	2.7	2.5	2.5	2.8	2.7	2.7	-5.6	-15.4	-13.8	-4.8	1.3	1.3	2.2	16.7	15.1
<i>Equip. Internet</i>	<i>91.4%</i>	<i>93.6%</i>	<i>92.6%</i>	<i>89.8%</i>	<i>92.5%</i>	<i>92.9%</i>	<i>93.3%</i>	<i>94.6%</i>	<i>94.6%</i>	<i>-5.5</i>	<i>-4.5</i>	<i>1.0</i>	<i>7.0</i>	<i>4.2</i>	<i>8.2</i>	12.4	8.7	7.3
<i>Illectronisme</i>	<i>12.4%</i>	<i>10.4%</i>	<i>12.3%</i>	<i>13.5%</i>	<i>10.8%</i>	<i>10.3%</i>	<i>11.6%</i>	<i>10.4%</i>	<i>10.4%</i>	<i>3.4</i>	<i>1.2</i>	<i>-6.1</i>	<i>-2.5</i>	<i>-0.2</i>	<i>-6.0</i>	<i>-5.9</i>	<i>-1.5</i>	<i>0.1</i>
R sans fixe/smart										0.89	0.91	0.91	0.83	0.90	0.90	0.85	0.88	0.88
R avec										0.84	0.90	0.90	0.81	0.89	0.89	0.84	0.87	0.87

Meth et Tel : échantillons méthodologique et téléphonique de TIC classique (tous deux avec un numéro de téléphone) ; FaF : échantillon face-à-face, avec ou sans numéros.
 S : poids de sondage ; I : poids corrigé de la non-réponse sans correction des effets de mode ; F : poids final, corrigé des effets de mode et calé sur la structure nationale.
 En gras, les distances de valeur absolue supérieure ou égale à 10 ; en rouge celles de valeur absolue dépassant 20.

Les données en italique ne participent pas au calcul des indicateurs R

Exemple : avec le poids de sondage, la moyenne des répondants est de 56,1 ans dans l'échantillon méthodologique et 53,4 ans dans l'échantillon téléphonique, pour une différence standardisée $d=-10,5$; ces valeurs sont respectivement 50,2, 45,0 et $d=-25,7$ avec la pondération finale.

Tableau 4 : Équilibrage et distances standardisées des échantillons de répondants Internet/papier relativement aux répondants Téléphone et face-à-face (**sans** numéro de téléphone)

	Int (%)			FaF		d : Int_Tel			d : Int_Faf		
	S	I	F	S	I (ou F)	S	I	F	S	I	F
Âge (moy.)	56.6	49.9	50.2	59.1	49.3	-12.8	-22.6	-24.9	18.4	-3.0	-4.4
Hommes	48.5%	46.4%	46.6%	40.5%	44.5%	-3.1	2.2	1.8	-16.3	-3.9	-4.3
I_pauvre	15.9%	18.8%	19.1%	16.6%	19.9%	-14.7	-13.1	-13.9	2.0	2.7	2.0
I_QPV	5.9%	6.8%	6.8%	7.1%	8.6%	-3.0	-1.0	-0.9	5.0	6.6	6.7
Revenus (moy. 1-4)	2.2	2.2	2.2	1.8	1.8	17.0	24.6	25.8	-62.6	-34.0	-33.5
Nb pers. ménage	1.9	2.0	2.0	1.8	2.1	14.2	35.7	35.9	-13.8	3.0	3.1
Courriel	56.4%	57.9%	56.9%	43.2%	46.8%	61.9	61.5	64.8	-26.6	-22.2	-20.3
Tel. fixe	69.6%	67.3%	67.0%	73.7%	70.8%	26.1	31.7	32.1	9.2	7.4	8.1
Smartphone	63.8%	71.5%	70.0%	62.7%	72.6%	44.0	32.0	36.0	-2.3	2.5	5.9
Diplôme (moy. 1-4)	2.5	2.5	2.5	2.2	2.3	13.0	-3.9	-1.1	-45.4	-23.4	-22.1
Equip. Internet	80.4%	85.4%	83.8%	74.1%	82.9%	26.7	22.7	28.6	-15.1	-7.0	-2.3
Illlectronisme	24.0%	18.1%	20.5%	31.9%	22.8%	-27.1	-20.7	-28.5	17.6	11.9	5.6
R sans fixe/smart						0.0	0.0	0.0	0.80	0.84	0.84
R avec						0.0	0.0	0.0	0.79	0.83	0.82

Int : échantillon Internet/papier ; Tel : échantillon téléphone ; Faf : échantillon face-à-face.

S : poids de sondage ; I : poids corrigé de la non-réponse sans correction des effets de mode ; F : poids final, corrigé des effets de mode et calé sur la structure nationale.

En gras, les distances de valeur absolue supérieure ou égale à 10 ; en rouge celles de valeur absolue dépassant 20.

Note : les distances et indicateurs R sont calculés pour les comparaisons des répondants des échantillons Internet / papier d'une part et téléphone et face-à-face d'autre part. Pour la comparaison des échantillons Internet / papier et téléphone, elles sont telles que les indicateurs R sont inférieurs à 0,1.

Un zoom sur les plus de 70 ans, population au sein de laquelle sont soupçonnés les plus importants biais, est instructif :

- Avant la correction des effets de mode, les répondants du face-à-face dont un numéro de téléphone est connu diffèrent peu des répondants de l'échantillon méthodologique. Ils sont plus jeunes, vivent dans des ménages plus nombreux aux revenus inférieurs, mais ils sont plus souvent équipés en smartphone ($d > 20$). Avec la pondération finale, l'écart pour le smartphone persiste et l'écart pour la communication d'une adresse courriel à l'administration fiscale augmente.
- Les écarts entre les répondants du face-à-face et les répondants de l'échantillon téléphonique sont négligeables ($d < 0,1$) sauf en ce qui concerne le sexe, les revenus (plus faibles en face-à-face), les diplômes (plus élevés en face-à-face) et l'équipement téléphonique dont le smartphone (les taux sont plus faibles en face-à-face).

Il existe ainsi des différences entre l'échantillon de référence de l'enquête classique (répondants de l'échantillon téléphone âgés de 70 ans et plus) et ceux de l'échantillon méthodologique d'un côté ou du face-à-face de l'autre. C'est essentiellement du point de vue des revenus (variable de la base de sondage) et de l'équipement en smartphone (variable collectée via le questionnaire) que les écarts sont importants. Toutefois, les écarts de revenus sont largement réduits par la correction des effets de mode, ce qui n'est pas le cas de l'équipement en smartphone.

La présence d'un courriel est associée à une participation plus élevée à chacune des enquêtes. Les distances standardisées sont modestes ($10 < d < 20$), mais l'effet se manifeste en analyse multivariée $RR = 1.26 [1.21 : 1.32]$ dans l'enquête classique et $RR = 1.09 [1.01 : 1.17]$ en face-à-face³. Or on peut montrer de même que la présence d'un courriel est associée à chacun des outcomes. Cette variable de la base de sondage mérite donc d'être prise en compte dans les pondérations, ce qui n'est pas fait en standard dans TIC classique (ni TIC Face-à-face).

³ En contrôlant les variables de la base de sondage plus le fait de connaître un numéro de téléphone pour le ménage.

Tableau 5 : Distances standardisées des principales variables sociodémographiques des répondants âgés de 70 ans et plus (échantillon méthodologique, téléphonique et face-à-face avec numéro de téléphone)

	Faf (% ou moy.)		d : Meth_Tel			d : Meth_Faf			d : Tel_Faf		
	S	I (ou F)	S	I	F	S	I	F	S	I	F
Âge (moy.)	77.4	77.6	-4.8	-3.8	-11.3	-15.4	-7.9	-11.1	-5.4	-4.3	0.0
Hommes	41.2%	46.2%	-15.5	-22.2	-21.4	-19.7	-2.3	-1.7	-4.1	19.9	19.7
I_pauvre	6.1%	5.5%	0.1	-3.6	-6.2	-0.1	-6.3	-8.6	-0.2	-2.7	-2.4
I_QPV	2.7%	3.1%	-1.7	-5.9	-8.6	-4.9	-5.5	-7.8	-3.3	0.4	0.9
Revenus (moy.)	2.0	2.0	-7.2	-0.9	6.4	-23.0	-9.7	-5.4	-8.5	-9.0	-11.7
Nb pers. ménage	1.6	1.7	-2.4	5.8	8.4	-1.7	11.4	13.1	1.3	4.7	3.7
Courriel	64.1%	65.3%	0.0	-2.7	8.3	-1.3	4.6	13.0	-1.4	7.4	4.7
<i>Filtre (fixe)</i>	62.1%	60.7%	-5.8	-4.6	-1.7	0.7	-1.4	0.8	6.5	3.1	2.5
<i>Filtre (mobile)</i>	59.2%	58.1%	-17.5	-18.3	-17.4	-16.9	-19.5	-18.9	0.6	-1.2	-1.5
Tel. fixe	90.5%	92.0%	25.6	28.8	32.3	6.9	14.3	17.1	-18.9	-14.9	-15.7
Smartphone	55.3%	54.9%	41.2	41.3	50.3	22.0	20.6	28.0	-18.8	-20.2	-21.6
Diplôme (1-4)	2.2	2.1	-9.9	-15.0	-6.9	-2.3	3.4	5.9	7.7	18.5	13.0
<i>Equip. Internet</i>	77.0%	78.7%	-12.9	-15.8	1.9	5.5	8.1	22.2	18.4	23.9	20.2
<i>Illectronisme</i>	38.9%	40.6%	14.4	21.6	0.5	8.0	11.5	-6.1	-6.4	-10.1	-6.5
R sans fixe/smart			0.88	0.91	0.91	0.82	0.88	0.88	0.85	0.86	0.87
R avec			0.73	0.84	0.83	0.78	0.86	0.85	0.82	0.84	0.84

S : poids de sondage ; I : poids corrigé de la non-réponse sans correction des effets de mode ; F : poids final, corrigé des effets de mode et calé sur la structure nationale.

En gras, les distances de valeur absolue supérieures ou égales à 10 ; en rouge, celles de valeur absolue dépassant 20.

3.3.1. Coordonnées de contact et réponse à l'enquête

Les coordonnées de contact (téléphone et courriel déclarés à l'administration fiscale) sont assez fortement liées à la non-réponse à l'enquête TIC. Cela est vrai pour l'enquête classique : la proportion de courriels dans Fideli atteint 81,7 % parmi les répondants contre 70,0 % parmi les non-répondants (d=27,7) ; la proportion de ménages avec un numéro atteint 92,3 % parmi les répondants et 81,4 % parmi les non-répondants (d=31,6). Dans l'enquête face-à-face : 80,3 % contre 70,1 % (avec d=23,8) pour le courriel et 87,2 % contre 72,3 % pour le numéro (d=37,5)⁴. Cela reste vrai lorsqu'on contrôle les variables de la base de sondage. Dans l'enquête classique, l'effet du courriel est important (RR=1.28 [1.22 ; 1.34]), tout comme celui du numéro (RR=1.49 [1.43 ; 1.55]). Dans l'enquête face-à-face, RR=1.10 [1.03 ; 1.18] pour le courriel et RR=1.44 [1.32 ; 1.51] pour le numéro.

3.3.2. Coordonnées de contact et estimation des outcomes

Dans l'enquête classique, globalement, le courriel est fortement associé à la déclaration d'un équipement Internet dans le ménage (RR=1,26 [1,23 ; 1,29]), à son usage au cours des trois derniers mois (RR=1,36 [1,31 ; 1,41]) et à l'illectronisme (RR=0,45 [0,40 ; 0,51]). La connaissance d'un numéro de téléphone joue faiblement sur l'équipement déclaré (RR=1,02 [1,00 ; 1,04]), mais pas sur l'usage d'Internet (RR=1,01 [0,98 ; 1,03]) et tend en revanche à réduire le risque d'illectronisme (RR=0,94 [0,86 ; 1,03]).

Dans l'enquête face-à-face, les résultats sont similaires. La connaissance d'un courriel joue fortement sur l'équipement RR=1.20 [1.15 ; 1.26], l'usage (RR=1.26 [1.20 ; 1.36]) et l'illectronisme (RR=0.54 [0.43 ; 0.67]). Le numéro joue faiblement sur l'équipement 1.03 [1.00 ; 1.07]), mais pas sur l'usage (RR=1.02 [0.97 ; 1.06]) ni sur l'illectronisme (RR=0.94 [0.78 ; 1.13]).

3.3.3. Coordonnées de contact, équipement téléphonique et estimation des outcomes

Dans l'enquête classique, la connaissance d'un numéro de téléphone pour joindre le ménage n'est, toutes choses égales par ailleurs, que significativement associée à la qualification d'illectronisme (Tableau 6). En revanche, la connaissance d'une adresse courriel se révèle fortement associée à la déclaration d'un équipement Internet au sein du ménage, à l'usage d'Internet au cours des trois derniers mois et à la qualification d'illectronisme. L'équipement téléphonique déclaré par le répondant et en particulier la possession d'un smartphone ont un effet

⁴ Ces chiffres et tous les suivants sont obtenus avec les poids de sondage, mais les résultats ne sont pas nettement changés avec la pondération intermédiaire ou finale.

similaire. L'analyse de l'enquête en face-à-face révèle des résultats très proches. Il semble donc important de considérer les déclarations d'équipement téléphonique, autant que les coordonnées de contact, dans la comparaison des échantillons pour l'estimation des effets de mode.

Tableau 6 : Effets des coordonnées de contact et des équipements téléphoniques sur les mesures des outcomes (ratios de risques IC95%)

		Équipement			Internet 3 mois			Illectronisme		
		RR	IC95%		RR	IC95%		RR	IC95%	
TIC classique	Courriel	1.22	1.19	1.25	1.30	1.26	1.35	0.50	0.45	0.56
	Numéro	<u>1.01</u>	0.99	1.03	1.00	0.98	1.02	0.97	0.88	1.06
	Tel fixe	1.08	1.06	1.10	1.05	1.03	1.07	0.77	0.70	0.85
	Smartphone	1.19	1.16	1.22	1.29	1.24	1.33	0.45	0.40	0.51
TIC Face-à-face	Courriel	1.15	1.10	1.20	1.19	1.12	1.26	0.65	0.53	0.79
	Numéro	<u>1.03</u>	0.99	1.06	1.01	0.97	1.05	1.00	0.84	1.19
	Tel fixe	1.07	1.04	1.10	1.03	1.00	1.06	0.70	0.56	0.87
	Smartphone	1.35	1.27	1.44	1.75	1.57	1.94	0.22	0.16	0.30

Les variables d'ajustement sont celles de la base de sondage (sauf l'âge et le sexe, qui sont ceux du répondant et non de la personne de référence). Pour l'enquête classique, est en plus contrôlé le mode de réponse (Internet, papier ou téléphone).

En gras les RR significativement différents de 1 au seuil de 5%

Exemple : toutes choses égales, les répondants de l'enquête classique ayant déclaré posséder un smartphone ont 0,68 fois plus de chances de se révéler illectronistes que les autres : la proportion d'illectronistes en leur sein est 32 % inférieure à celle estimée parmi les personnes qui n'ont pas déclaré posséder de smartphone.

3.4. Estimations des outcomes et mode de collecte dans TIC classique

Pour l'enquête face-à-face, afin d'identifier une situation comparable à celle de l'enquête classique, les enquêtés ont été interrogés sur l'équipement de leur domicile en Internet, à la fois au moment de la collecte (c'est-à-dire au quatrième trimestre 2021) et en avril 2021 (période de collecte de l'enquête TIC classique). Dans les faits, il y a très peu de différences entre les gens qui ont déclaré être équipés en avril et au quatrième trimestre (1 point maximum quelle que soit la variable socio-démographique étudiée). Dans la suite, c'est donc l'équipement au mois d'avril qui est étudié.

Dans l'ensemble, les ménages enquêtés en face-à-face sont légèrement plus connectés que ceux enquêtés dans l'enquête classique : ils déclarent un peu plus souvent être équipés (92,6 % contre 90,9 %), sont un peu plus nombreux à avoir utilisé Internet dans les trois derniers mois (87,5 % contre 86,2 %) et moins souvent illectronistes (12,5 % contre 13,5 %). Bien que cet écart soit modeste, il n'en reste pas moins surprenant dans la mesure où le face-à-face représente mieux les ménages modestes et pauvres.

L'écart se creuse pour certaines populations spécifiques, normalement plus fragiles vis-à-vis des nouvelles technologies :

- Ainsi, pour l'équipement, 62,5 % des plus de 75 ans déclarent posséder une connexion à domicile dans l'enquête Face-à-face contre seulement 54,4 % dans l'enquête classique. La différence s'accroît lorsqu'on prend en compte la présence de numéro de téléphone dans les bases fiscales ou dans l'annuaire : les individus ayant donné leur numéro de téléphone sont plus souvent équipés en Internet dans Face-à-face que dans TIC classique (+11,6 points). En revanche, la différence est beaucoup moins marquée pour les individus sans numéro ni mail déclaré. La différence est moins présente pour l'usage et l'illectronisme.

- Pour les ménages aux revenus les plus modestes (première tranche), le taux d'équipement est à nouveau meilleur sur TIC Face-à-face. Cette différence s'accroît pour les individus ne disposant ni de mail ni de numéro de téléphone (+9,8 points), l'enquête en Face-à-face captant davantage d'individus connectés dans cette population que ne le fait l'enquête classique. Il en va de même pour l'illectronisme ou pour l'utilisation d'Internet dans les 3 mois, où la présence de mail dans la base de sondage introduit une grande différence dans l'utilisation selon que la collecte soit en face-à-face ou non (63,3 % de personnes dans la première tranche de revenu sont utilisatrices d'Internet contre 53,0 % dans TIC classique).

- Les personnes vivant seules, avec de faibles diplômes ou dans une zone rurale déclarent elles aussi être davantage connectées lorsqu'elles sont enquêtées en face-à-face que dans l'enquête classique.

Lorsqu'on regarde par mode de collecte, les différences sont aussi importantes : si le taux d'équipement est quasiment identique entre les enquêtés en face-à-face et par téléphone (respectivement 92,6 % et 92,9 %), les

ménages équipés sont logiquement plus nombreux dans les répondants Internet (97,2 %) et bien moins nombreux chez les répondants papier (81,9 %), pour lesquels ce mode est la seule possibilité de réponse à l'enquête s'ils n'ont pas Internet ou savent peu s'en servir. Il en va de même pour l'utilisation d'Internet dans les trois derniers mois : la proportion d'utilisateurs fréquents est plus importante sur Internet (93,7 %), puis par téléphone (89,3 %) puis en face-à-face (87,5 %) puis par papier (75,2 %). Le constat est le même pour l'indicateur d'illectronisme.

Tous ces résultats sont obtenus avec la pondération finale des enquêtes censée corriger les biais de sélection et de mesure suspectés dans TIC classique. Ces résultats donnent ainsi à penser que les personnes devant répondre ou ayant choisi de répondre via le papier dans l'enquête classique ont tendance à utiliser l'enquête pour signaler qu'elles ne sont pas équipées, qu'elles ne sont pas utilisatrices d'Internet et sont illectronistes. Les sections suivantes approfondissent la comparaison des deux enquêtes en distinguant chacun des sous-échantillons et en procédant à des analyses multivariées.

3.5. Équipement, usage d'Internet et illectronisme

3.5.1. L'échantillon Téléphone comme référence dans TIC classique

Nous commençons par vérifier que l'échantillon téléphonique, qui est la référence pour les estimations des outcomes, est bien comparable au face-à-face. Les écarts en termes de représentativité étaient relativement faibles (Tableau 3), ce qui est une condition nécessaire. Pour cette analyse, l'hypothèse la plus plausible concernant les biais est celle du biais de sélection : en effet, comme rappelé plus haut, dans les deux collectes, la présence de l'enquêteur (sonore et physique) doit conduire à éliminer tout biais de mesure relatif majeur. Par conséquent, si les principaux facteurs de confusion sont pris en compte dans une analyse toutes choses égales par ailleurs, l'effet du mode de collecte devrait refléter l'effet global des biais différentiels de sélection propres aux deux protocoles de collecte.

En données brutes et toutes choses égales par ailleurs, les différences d'estimation des outcomes sont de fait très faibles, ce qui est rassurant quant au statut de référence de l'échantillon téléphonique dans TIC classique (Tableau 7). Le rôle de la correction des effets de mode est logement tout à fait négligeable (car l'échantillon téléphonique est la référence à cet égard au sein de l'enquête classique) et on observe que les résultats sont pratiquement inchangés, que l'on raisonne en bivarié (colonnes « Brut ») ou en multivarié (colonnes « Socio-démographique », « + Coord » et « + Équipement »).

Cela reste notamment vrai alors que les taux d'équipement en téléphone fixe déclaré en face-à-face est inférieur⁵.

Tableau 7 : Comparaisons des échantillons Téléphone et Face-à-face (avec numéro) : pourcentages et ratios de risques associés

Pondération		Brut				Socio démographique			+ Coord		+ Équipement				
		Tel (%)	Faf (%)	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%				
Inter.	Équipement	92,5	94,6	1.02	1.01	1.04	1.03	1.02	1.05	1.02	1.01	1.04	1.03	1.01	1.04
	Usage_3mois	88,9	89,6	1.01	0.99	1.03	1.03	1.01	1.05	1.02	1.00	1.04	1.01	0.99	1.03
	Illectronisme	10,8	10,4	0.96	0.79	1.16	0.92	0.78	1.08	0.99	0.85	1.16	0.92	0.79	1.06
Finale	Équipement	92,9	94,6	1.02	1.00	1.04	1.03	1.02	1.05	1.02	1.01	1.04	1.02	1.01	1.04
	Usage_3mois	89,3	89,6	1.00	0.98	1.03	1.02	1.00	1.04	1.01	0.99	1.03	1.01	0.99	1.03
	Illectronisme	10,4	10,4	1.00	0.83	1.21	0.91	0.77	1.06	0.98	0.84	1.15	0.90	0.78	1.05

Note : 2239 répondants Téléphone et 2290 répondants face-à-face.

La pondération intermédiaire corrige de la non-réponse dans chaque sous-échantillon (téléphonique, méthodologique, Internet/papier), mais pas des effets de mode. La pondération finale corrige les effets de mode parmi les 70 ans et plus et opère un dernier calage de l'échantillon de tous les répondants sur les marges nationales.

En gras les RR significativement différents de 1 au seuil de 5% ; en souligné ceux pour lesquels p-valeur < 0.1.

- Sociodémographique : L'ajustement comprend sexe, âge, diplôme, ménage pauvre, résidence en QPV, taille d'unité urbaine agrégée, type de logement, taille du ménage ;
- + coordonnées : ajoute les coordonnées de contact (courriel et téléphone du ménage) ;
- + équipement téléphone : ajoute l'équipement en téléphone fixe et smartphone)

Exemple : avec la pondération intermédiaire, on obtient respectivement 92,5 % et 94,6 % d'individus équipés d'Internet à domicile dans l'échantillon Téléphone et dans l'échantillon Face-à-face, conduisant à des ratios de risque sans ajustement (brut), avec ajustement sur les variables socio-démographiques et avec les variables d'équipement téléphonique, de 1,02, 1,03 et 1,03 (respectivement).

⁵ Avec la pondération finale : RR=0.96 [0.92 ; 1.00] en brut, RR=0.93 [0.90 ; 0.97] en contrôlant les caractéristiques socio-démographiques et les coordonnées de contact. Les taux d'équipement déclarés en smartphone sont en revanche extrêmement proches.

La même comparaison restreinte aux 70 ans produit des résultats similaires : la correction des effets de modes est encore sans effet et les résultats sont à peine changés suivant que l'on raisonne en bivarié ou en multivarié. On observe toutefois que parmi les 70 et plus, l'estimation du taux d'équipement en Internet au domicile est plus élevée en face-à-face qu'au téléphone ce quel que soit le type d'analyse. Malgré les écarts de taux d'équipement téléphonique déclaré par les répondants entre les deux sous-populations, on constate que l'ajustement sur celui-ci ne modifie substantiellement rien l'estimation des écarts entre les deux enquêtes. Ainsi, les écarts de distribution bruts observés au tableau 5 persistent pour les l'équipement, tandis que la tendance à une moindre prévalence de l'illectronisme en face-à-face est conservée ($d=-6.5$ dans le tableau 5, $RR_{brut}=0.93$ et $RR+équipement=0.88$).

Tableau 8 : Comparaisons des échantillons Téléphone et Face-à-face (avec numéro) parmi les 70 ans et plus : pourcentages et ratios de risques associés

Pondération		Brut				Socio démographique			+ Coord			+ Équipement téléphonique			
		Tel (%)	Faf (%)	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%		
Inter.	Équipement	68,2	78,7	1.15	1.06	1.26	1.13	1.05	1.22	1.12	1.05	1.20	1.13	1.06	1.21
	Usage_3mois	51,2	59,3	1.10	0.97	1.24	1.05	0.94	1.18	1.05	0.94	1.17	1.06	0.95	1.17
	Illectronisme	45,6	40,6	0.89	0.76	1.05	0.93	0.80	1.08	0.97	0.84	1.13	<u>0.88</u>	0.76	1.01
Finale	Équipement	69,9	78,7	1.13	1.04	1.22	1.12	1.04	1.21	1.12	1.04	1.19	1.12	1.05	1.20
	Usage_3mois	55,9	59,3	1.06	0.94	1.20	1.05	0.94	1.17	1.05	0.94	1.16	1.05	0.95	1.17
	Illectronisme	43,8	40,6	0.93	0.78	1.09	0.93	0.80	1.08	0.97	0.84	1.13	<u>0.88</u>	0.76	1.01

Note : 523 répondants Téléphone et 519 répondants face-à-face.

La pondération intermédiaire corrige de la non-réponse dans chaque sous-échantillon (téléphonique, méthodologique, Internet/papier), mais pas des effets de mode. La pondération finale corrige les effets de mode parmi les 70 ans et plus et opère un dernier calage de l'échantillon de tous les répondants sur les marges nationales.

En gras les RR significativement différents de 1 au seuil de 5% ; en souligné ceux pour lesquels p-valeur <0.1.

- Sociodémographique : L'ajustement comprend sexe, âge, diplôme, ménage pauvre, résidence en QPV, taille d'unité urbaine agrégée, type de logement, taille du ménage ;
- + coordonnées : ajoute les coordonnées de contact (courriel et téléphone du ménage) ;
- + équipement téléphone : ajoute l'équipement en téléphone fixe et smartphone)

Si l'on considère le face-à-face comme le mode le plus indépendant de la thématique de TIC (équipements en outils de communications et en usage des technologies de l'information) alors notre analyse souligne que l'échantillon téléphonique de l'enquête classique est d'une qualité comparable à celui de l'enquête en face-à-face. Toutefois, il subsiste quelques écarts en termes d'équipement pour les personnes âgées. Relativement à l'illectronisme, les écarts sont négligeables. On note cependant que pour les 70 ans et plus, la prise en compte de l'équipement téléphonique tend à augmenter l'écart au profit de l'enquête classique ($RR=0.88$ [0.76 ; 1.01]), ce qui reflète le fait que les taux d'équipement en téléphone fixe et en smartphone sont nettement inférieurs dans le Face-à-face au sein de cette population âgée⁶. Le Face-à-face parvient ainsi à interroger des personnes se déclarant moins équipées en téléphone que l'échantillon téléphonique qui utilise précisément le téléphone pour interroger les personnes. On retrouve là sans doute un biais de sélection du protocole téléphonique.

3.5.2. Échantillons Méthodologique et Téléphone

La correction des effets de mode a un impact mineur sur le taux d'équipement et le niveau d'usage quotidien mais elle augmente l'écart entre l'échantillon téléphone et l'échantillon méthodologique pour l'illectronisme : avant, l'écart est négligeable ($RR=1.04$) ; après, la proportion est nettement plus faible parmi les répondants Téléphone ($RR=0,84$). Pour l'équipement et l'usage, ces résultats sont peu affectés par les différents ajustements multivariés. En revanche, l'illectronisme s'avère davantage associé à l'équipement téléphonique, dont on a vu (Tableau 3) qu'il était nettement plus élevé dans l'échantillon téléphonique. Comme pressenti au vu du lien souligné précédemment entre ces variables d'équipement et les outcomes, il n'est pas surprenant que leur prise en compte affecte l'écart entre les deux échantillons.

⁶ Cela reste vrai dans un modèle multivarié pour ce qui est de la déclaration de possession d'un téléphone fixe : $RR=0.96$ [0.92 ; 0.99], d'un téléphone mobile ($RR=0.87$ [0.83 ; 0.91]), et d'un smartphone : $RR=0.85$ [0.76 ; 0.85].

Tableau 9 : Comparaisons des échantillons Méthodologique et Téléphone (tous âges) : pourcentages et ratios de risque associés

Pondération		Brut				Socio démographique		+ Coordonnées		+ Équipement téléphonique					
		Meth (%)	Tel (%)	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%				
Intermédiaire	Équipement	93,6	92,4	0.99	0.97	1.00	0.97	0.96	0.99	0.98	0.96	0.99	0.97	0.96	0.98
	Usage_3mois	93,4	88,9	0.99	0.98	1.01	0.98	0.96	0.99	0.98	0.96	1.00	0.97	0.96	0.99
	Illectronisme	10,4	10,8	1.04	0.89	1.21	<u>1.13</u>	0.99	1.28	1.09	0.96	1.23	1.29	1.14	1.47
Finale	Équipement	92,6	92,9	1.00	0.99	1.02	0.98	0.97	1.00	0.99	0.97	1.00	0.98	0.97	0.99
	Usage_3mois	87,5	89,3	1.02	1.00	1.04	1.00	0.98	1.01	1.00	0.98	1.02	<u>0.99</u>	0.97	1.01
	Illectronisme	12,3	10,3	0.84	0.72	0.98	0.97	0.86	1.11	0.96	0.85	1.09	1.14	1.00	1.30

Note : 2239 répondants Téléphone et 9390 répondants Méthodologique.

En gras les RR significativement différents de 1 au seuil de 5% ; en souligné ceux pour lesquels p-valeur <0.1.

- Sociodémographique : L'ajustement comprend sexe, âge, diplôme, ménage pauvre, résidence en QPV, taille d'unité urbaine agrégée, type de logement, taille du ménage ;
- + coordonnées : ajoute les coordonnées de contact (courriel et téléphone du ménage) ;
- + équipement téléphone : ajoute l'équipement en téléphone fixe et smartphone)

En restreignant l'analyse aux 70 ans et plus, on note que la pondération finale gomme bien tous les écarts d'estimation des outcomes entre les deux échantillons : pour l'illectronisme par exemple, l'écart est de RR=1,30 en brut avec la pondération intermédiaire mais de RR=1,01 avec la pondération finale. En revanche, conformément aux précédentes analyses de représentativité (

Tableau 5) qui soulignaient les différences de mesure de coordonnées de contact d'équipement entre les deux échantillons, leur prise en compte affecte fortement la différence de mesure d'illectronisme entre les deux échantillons : 1,09 avec l'ajustement sur les variables socio-démographiques et les coordonnées et RR=1,32 avec l'ajustement supplémentaire sur l'équipement téléphonique, avec la pondération finale.

Tableau 10 : Comparaisons des échantillons Méthodologique et Téléphone pour les répondants de 70 ans et plus : pourcentages et ratios de risque associés

Pondération		Brut				Socio démographique			+ Coordonnées			+ Équipement téléphonique			
		Meth (%)	Tel (%)	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%		
Inter.	Équipement	75,3	68,2	0.91	0.84	0.97	0.91	0.86	0.97	0.91	0.86	0.97	0.88	0.83	0.93
	Usage_3mois	65,0	54,2	0.83	0.76	0.92	0.84	0.77	0.92	0.84	0.78	0.92	0.81	0.74	0.88
	Illectronisme	35,0	45,6	1.30	1.15	1.47	1.31	1.17	1.47	1.30	1.16	1.45	1.57	1.40	1.76
Finale	Équipement	69,0	69,9	1.01	0.94	1.09	0.98	0.92	1.05	0.97	0.91	1.02	0.92	0.87	0.98
	Usage_3mois	56,3	55,9	0.99	0.90	1.10	0.96	0.88	1.05	0.94	0.87	1.02	0.88	0.81	0.96
	Illectronisme	43,6	43,8	1.01	0.89	1.14	1.07	0.96	1.20	<u>1.09</u>	0.98	1.22	1.32	1.17	1.48

Note : 523 répondants Téléphone et 2663 répondants Méthodologique.

La pondération intermédiaire corrige de la non-réponse dans chaque sous-échantillon (téléphonique, méthodologique, Internet/papier), mais pas des effets de mode. La pondération finale corrige les effets de mode parmi les 70 ans et plus et opère un dernier calage de l'échantillon de tous les répondants sur les marges nationales.

En gras les RR significativement différents de 1 au seuil de 5% ; en souligné ceux pour lesquels p-valeur <0.1.

- Sociodémographique : L'ajustement comprend sexe, âge, diplôme, ménage pauvre, résidence en QPV, taille d'unité urbaine agrégée, type de logement, taille du ménage ;
- + coordonnées : ajoute les coordonnées de contact (courriel et téléphone du ménage) ;
- + équipement téléphone : ajoute l'équipement en téléphone fixe et smartphone)

Exemple : sur la population des 70 ans et plus, avec la pondération intermédiaire, on obtient respectivement 75,3 % et 68,2 % d'individus équipés d'Internet à domicile dans l'échantillon méthodologique et dans l'échantillon téléphonique. Avec la pondération finale, on obtient respectivement 69,0 % et 69,9 %.

Il est difficile de tirer une conclusion ferme quant à la qualité de la correction des effets de mode opérée dans TIC classique entre l'échantillon méthodologique et le téléphone. La correction est complète si la déclaration de l'équipement téléphonique est affectée d'un biais de mesure relatif important entre les échantillons : dans ce cas en effet, sa considération ne pourrait que fausser la correction et l'analyse toutes choses égales par ailleurs. En revanche, si la déclaration d'équipement possède l'invariance de mesure désirée, alors la prise en compte des variables afférentes est indispensable au vu de l'effet qu'elles ont sur les estimations des outcomes. Un doute plus préoccupant subsiste en ce qui concerne les coordonnées de contact (courriel et téléphone du ménage). Si la prise en compte des coordonnées n'affecte pas les estimations des différences de niveaux d'équipement et d'usage d'Internet, elle modifie plus substantiellement celle de l'illectronisme. Or les coordonnées de contact ne sont pas recueillies au moment de l'enquête mais connues à l'avance dans la base de sondage : elles sont donc a priori exemptes de biais de sélection ou de biais de mesure lié à l'enquête elle-même.

3.5.3. Échantillons Internet/papier et Face-à-face

De façon surprenante, les écarts entre les échantillons Internet/papier et Face-à-face (sans numéro) sont très réduits, notamment après prise en compte de la pondération définitive, même parmi les 70 ans et plus (

Tableau 12). Cela est vrai également pour la sous-population de ménages sans courriel (En gras les RR significativement différents de 1 au seuil de 5% ; en souligné ceux pour lesquels p-valeur <0.1.

- Sociodémographique : L'ajustement comprend sexe, âge, diplôme, ménage pauvre, résidence en QPV, taille d'unité urbaine agrégée, type de logement, taille du ménage ;
- + coordonnées : ajoute les coordonnées de contact (courriel et téléphone du ménage) ;
- + équipement téléphone : ajoute l'équipement en téléphone fixe et smartphone)

Tableau 13), c'est-à-dire la population dépourvue de coordonnées de contact, donc *a priori* la plus difficile à joindre et qui doit effectuer les efforts les plus importants pour répondre.

Tableau 12 : Comparaison des échantillons Internet/papier et face-à-face (sans numéro de téléphone) : pourcentages et ratios de risque associés parmi les 70 ans et plus

Pondération		Brut				Socio démographique			+ Coordonnées			+ Équipement téléphonique			
		Int/Pap	Tel	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%		
		(%)	(%)												
Inter.	Équipement	57,0	46,9	0.82	0.69	0.98	0.94	0.80	1.10	0.96	0.82	1.13	<u>0.89</u>	0.78	1.02
	Usage_3mois	44,8	30,5	0.70	0.55	0.91	0.85	0.70	1.05	0.88	0.72	1.07	0.79	0.66	0.94
	Illectronisme	55,1	68,5	1.24	1.09	1.42	1.12	1.01	1.25	1.11	1.00	1.23	1.15	1.04	1.27
Finale	Équipement	50,6	46,9	0.93	0.77	1.11	1.07	0.90	1.26	1.09	0.93	1.29	0.97	0.84	1.13
	Usage_3mois	36,7	31,5	0.86	0.66	1.11	1.02	0.84	1.25	1.05	0.86	1.28	0.89	0.75	1.06
	Illectronisme	68,8	68,5	1.09	0.96	1.24	1.01	0.92	1.11	1.00	0.91	1.10	1.05	0.96	1.15

Note : 284 répondants FaF et 787 répondants Internet.

En gras les RR significativement différents de 1 au seuil de 5% ; en souligné ceux pour lesquels p-valeur <0.1.

- Sociodémographique : L'ajustement comprend sexe, âge, diplôme, ménage pauvre, résidence en QPV, taille d'unité urbaine agrégée, type de logement, taille du ménage ;
- + coordonnées : ajoute les coordonnées de contact (courriel et téléphone du ménage) ;
- + équipement téléphone : ajoute l'équipement en téléphone fixe et smartphone)

Tableau 13 : Comparaison des échantillons Internet/papier sans courriel et face-à-face (sans numéro ni courriel) : pourcentages et ratios de risque

Pondération	Outcome	Brut				Socio démographique			+ Équipement téléphonique			
		Int/Pap	FaF	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%	
		(%)	(%)									
Intermédiaire	Équipement	71,6	73,3	1.02	0.95	1.11	1.00	0.93	1.07	0.95	0.89	1.02
	Usage_3mois	66,2	66,4	1.00	0.91	1.10	0.99	0.91	1.07	0.94	0.87	1.02
	Illectronisme	33,6	33,5	1.00	0.83	1.20	0.99	0.87	1.12	1.06	0.94	1.19
Finale	Équipement	68,8	73,3	<u>1.07</u>	0.98	1.15	1.03	0.96	1.11	0.97	0.91	1.04
	Usage_3mois	62,3	66,4	1.07	0.97	1.18	1.03	0.95	1.13	0.98	0.90	1.06
	Illectronisme	37,5	33,5	0.89	0.74	1.08	0.92	0.81	1.03	1.00	0.89	1.12

Note : 468 répondants FaF et 1164 répondants Internet/papier.

En gras les RR significativement différents de 1 au seuil de 5%

- Sociodémographique : L'ajustement comprend sexe, âge, diplôme, ménage pauvre, résidence en QPV, taille d'unité urbaine agrégée, type de logement, taille du ménage ;
- + équipement téléphone : ajoute l'équipement en téléphone fixe et smartphone)

3.6. Comparaison globale des estimations des outcomes dans les deux enquêtes

Les analyses précédentes soulignent que les répondants du face-à-face diffèrent quelque peu de leurs homologues de TIC classique (notamment sur le plan des revenus), des coordonnées de contact et des déclarations d'équipement téléphonique. Cela a-t-il un effet sur les estimations des outcomes au niveau global ? Quelle que soit la pondération utilisée dans TIC classique, les écarts avec le face-à-face sont très réduits. La prise en compte de la pondération n'a qu'un effet négligeable dans les estimations des outcomes dans l'enquête classique : par exemple, le taux d'équipement des ménages passe de 91,8 % à 90,9 % après correction des principaux effets de mode (Tableau 14).

On note des écarts significatifs mais de faible ampleur pour les taux d'équipement et d'usage ; mais les écarts les plus préoccupants sont relevés pour l'illectronisme : après ajustement sur les variables socio-démographiques, on note que la proportion estimée dans le face-à-face serait de 9% inférieure (RR=0.91 [0.83 ; 1.00]) ; l'écart se comble un peu après prise en compte des coordonnées socio-démographiques (RR=0.94 [0.86 ; 1.04]) et n'est plus du tout visible après prise en compte supplémentaire de l'équipement téléphonique (RR=1.00). Ces différences ne sont pas toujours significatives, leurs intervalles de confiance sont compatibles avec l'infériorité avérée des mesures dans le face-à-face.

Tableau 14 : Comparaison des enquêtes : pourcentages et ratios de risque

Pondération		Brut				Socio démographique			+ Coordonnées			+ Équipement téléphonique			
		Classique	Face-à-face	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%	RR	IC95%		
		(%)	(%)												
Inter.	Équipement	91.8	92.6	1.01	1.00	1.02	1.01	1.00	1.02	1.01	1.00	1.02	1.00	0.99	1.01
	Usage_3mois	87.8	87.5	1.00	0.98	1.01	1.00	0.99	1.02	1.00	0.98	1.01	0.99	0.98	1.00
	Illectronisme	12.0	12.5	1.04	0.93	1.17	1.02	0.93	1.12	1.04	0.95	1.15	1.10	1.00	1.20

	Équipement	90.9	92.6	1.02	1.01	1.03	1.02	1.01	1.03	1.02	1.01	1.03	1.01	1.00	1.02
Finale	Usage_3mois	86.2	87.5	1.01	1.00	1.03	1.02	1.00	1.03	1.02	1.00	1.03	1.00	0.99	1.02
	Illectronisme	13.5	12.5	0.93	0.82	1.04	0.91	0.83	1.00	0.94	0.86	1.04	1.00	0.92	1.09

En gras les RR significativement différents de 1 au seuil de 5% ; en souligné ceux pour lesquels p-valeur <0.1.

- Sociodémographique : L'ajustement comprend sexe, âge, diplôme, ménage pauvre, résidence en QPV, taille d'unité urbaine agrégée, type de logement, taille du ménage ;
- + coordonnées : ajoute les coordonnées de contact (courriel et téléphone du ménage) ;
- + équipement téléphone : ajoute l'équipement en téléphone fixe et smartphone)

4. Discussion et remarques conclusives

Les mesures de représentativité des échantillons de répondants des deux enquêtes sont relativement proches. Toutefois, on peut noter que le face-à-face parvient à interroger proportionnellement davantage de personnes issues de ménages pauvres ou à faibles revenus. C'est le cas notamment pour les ménages dont aucune coordonnée de contact n'est connue avant la collecte (sauf l'adresse postale). Parmi les personnes âgées de 70 ans et plus, l'enquête face-à-face parvient également à interroger des personnes se déclarant moins souvent équipées d'un téléphone fixe et surtout d'un smartphone que l'échantillon téléphonique de l'enquête classique, qui justement repose sur un protocole mobilisant le téléphone pour interroger les personnes. Il est peu plausible que cet écart résulte d'un biais de mesure entre ces deux protocoles mobilisant un enquêteur. Par conséquent, la conclusion la plus probable est que l'échantillon téléphonique sélectionne, ce qui est naturel, des répondants mieux équipés téléphoniquement que le face-à-face au sein de la population âgée (notamment en smartphone). Toutefois, globalement, l'échantillon Téléphone de TIC classique est très proche de l'échantillon correspondant du Face-à-face : d'une part, les écarts de représentativité sont très réduits entre les deux pour les variables socio-démographiques et les équipements téléphoniques et d'autre part, les écarts d'estimations des outcomes sont très réduites. On note toutefois quelques écarts chez les plus de 70 ans, le face-à-face interrogeant des personnes plus souvent équipées d'Internet et tendanciellement moins souvent en situation d'illectronisme. L'approche qui consiste à prendre comme référence l'échantillon de répondants téléphoniques de l'échantillon Téléphone de TIC classique est donc fondée globalement mais pourrait être examinée avec davantage d'attention pour les 70 ans et plus. Le gain en termes de mesure serait sans doute modeste mais perceptible eu égard à la taille de la population concernée.

Les deux enquêtes fournissent des estimations pratiquement identiques des niveaux d'équipement des ménages, d'usagers d'Internet au cours des trois derniers mois et de personnes en situation d'illectronisme. Pour cette dernière variable, l'écart est plus marqué bien que significatif uniquement lorsque les caractéristiques socio-démographiques sont contrôlées. La prise en compte des coordonnées de contact en plus tend à réduire la différence modélisée entre les deux enquêtes.

De manière générale, les coordonnées de contact connues dans les bases de sondage et les déclarations d'équipement téléphonique et notamment en smartphone dans les enquêtes sont liées d'une part à la réponse à l'enquête (pour les coordonnées) et d'autre part aux mesures des outcomes, en analyses bivariées et multivariées. Cela souligne l'intérêt qu'il y aurait à prendre en compte les coordonnées pour les corrections de la non-réponse totale dans les enquêtes et d'autre part de considérer la prise en compte des déclarations d'équipement téléphonique pour l'estimation des effets de mode dans les enquêtes.

Il semble donc y avoir très peu de biais de sélection entre le face-à-face et l'échantillon téléphone ; mais il subsiste des doutes (relativement mineurs) pour l'échantillon méthodologique relativement au téléphone parmi les 70 ans et plus, qui se manifestent par des écarts significatifs de mesure d'illectronisme une fois les coordonnées de contact prises en compte.

De façon surprenante, les estimations des outcomes dans l'échantillon Internet/papier de l'enquête classique, et notamment du sous-échantillon sans courriel connu, sont très proches de celles observées en face-à-face. Les craintes relatives à l'existence d'un fort biais de sélection sont donc à écarter pour cette population difficile à contacter (le taux de collecte y est inférieur) parce qu'elle se rend moins joignable (pas de coordonnées de contact courriel ou téléphone connues).

Des investigations plus poussées devraient cependant être menées pour comprendre pourquoi l'estimation de l'illectronisme est (un peu) inférieure dans l'enquête en face-à-face. L'écart est sensible et significatif pour l'échantillon téléphone et sensible mais non significatif pour l'échantillon Internet/papier.

Bibliographie

- Austin, P. C. and E. A. Stuart (2015). "Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies." Stat Med **34**(28): 3661-3679.
- Bethlehem, J., F. Cobben and B. schouten (2009). "Des indicateurs de la représentativité aux enquêtes." Techniques d'Enquêtes(Recueil du symposium 2008 de Statistique Canada): 1-10.
- Castell, L., M. Clerc and S. Legleye (2021). Victimations déclarées et effets de mode: l'expérimentation CVS panel. 11ème colloque francophone sur les sondages. Bruxelles.
- Diaz-Quijano, F. A. (2012). "A simple method for estimating relative risk using logistic regression." BMC Medical Research Methodology **12**(14).
- Haziza, D. and E. Lesage (2016). "A discussion of weighting procedures for unit nonresponse." Journal of Official Statistics **32**(1): 129-145.
- Hernan, M. A. and J. M. Robins (2020). Causal Inference: What If, Chapman & Hall/CRC.
- Imbens, G. W. and D. B. Rubin (2015). Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge, MA, Cambridge University Press
- Kreuter, F., S. Presser and R. Tourangeau (2008). "Social Desirability Bias In CATI, IVR, And Web Surveys. The Effects Of Mode And Question Sensitivity." Public Opinion Quarterly **72**(5): 847-865.
- Krosnick, J. A., S. Narayan and W. R. Smith (1996). Satisficing in surveys: initial evidence. Advances in survey research. M. Braveman and J. Slater. San Francisco, Jossey-Bass: 29-44.
- Legleye, S. and A. Rolland (2019). "Une personne sur six n'utilise pas Internet, plus d'un usager sur trois manque de compétences numériques de base [One in six people do not use the Internet, more than one in three users lack basic digital skills]." Insee Première **1780**.
- Little, R. J. A. and S. Vartivarian (2005). "Does weighting for nonresponse increase the variance of survey means?" Survey methodology **31**(2): 161–168.
- Pearl, J. (2009). Causality, Cambridge University Press.
- Rubin, D. B. (2007). "The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trials." Statistics in Medicine **26**: 20-36.