

---

## AGREGATION DE DONNEES MULTIMODE : IMPACT SUR LA MODELISATION DES VARIABLES PRESENTANT UN EFFET DE MESURE

Mady Cissé, Christophe Barret (\*)

(\*) Céreq, Centre d'Etudes et de REcherches sur les Qualifications

mady.cisse@cereq.fr, christophe.barret@cereq.fr

**Mots-clés** (6 maximum) : Échantillonnage, calage, extension

---

### Résumé

Les enquêtes Génération du Céreq permettent d'étudier les parcours d'insertion professionnelle des jeunes à la sortie du système éducatif, et font l'objet d'un processus de rénovation dans lequel une collecte multimode Internet/Téléphone est envisagée. Des expérimentations de l'utilisation d'Internet comme mode de collecte ont donc été menées en parallèle des enquêtes par téléphone en 2013, 2015 et 2016.

La combinaison de plusieurs modes de collecte plutôt qu'un seul implique deux effets qui modifient l'estimation des paramètres d'intérêt. D'une part il y a des effets de mesure : les enquêtés ne répondent pas nécessairement de la même manière selon le mode de collecte. D'autre part il y a l'effet de sélection : selon le mode de collecte on va atteindre principalement les enquêtés qui y sont le plus réceptifs.

Une méthode de matching sur score de propension a été employée pour distinguer les effets de sélection des effets de mesure, et ainsi repérer les variables sujettes aux effets de mesure. Ces derniers sont problématiques lors de la réalisation d'estimations car, en l'absence de traitements spécifiques, ils sont sources de biais. Dans le cadre, d'enquêtes répétées, et toujours en l'absence de traitements spécifiques, ces biais sont sensibles à la variation de la répartition des modes de collecte entre les différentes enquêtes.

Cette présentation fait suite à celle présentée lors des dernières JMS<sup>[1]</sup> qui présentait la détection de ces effets au sein des enquêtes Génération ainsi que des éléments sur le protocole de collecte multimode. Il s'agira ici de comparer différentes approches d'agrégation des données, dont certaines tentent de corriger le biais lié à l'effet de mesure.

La première, l'agrégation simple, consiste à juxtaposer les données collectées dans chacun des deux modes sans appliquer de traitement spécifique au mode sur les réponses.

La seconde est une méthode proposée à titre expérimental par l'Insee<sup>[3]</sup> qui repose sur l'imputation des réponses des individus qui portent l'effet de mesure sur les variables d'intérêt. Les individus ayant des réponses, sur les variables présentant un effet de mode, trop éloignées de leur contrefactuel du matching évoqué précédemment sont identifiés puis ces réponses sont modifiées par hot-deck sur le contrefactuel afin d'annuler l'effet de mesure.

L'objectif de cette étude est alors de regarder l'impact de ces deux méthodes d'agrégation sur les estimations de paramètres de modélisations économétriques incluant des variables sujettes à des effets de mesure.

Il s'agira de discuter de la modélisation d'une variable présentant un effet de mesure à l'aide de variable sans effet de mesure, en fonction de la méthode d'agrégation. En particulier, l'introduction à

la modélisation de la variable indicatrice précisant le mode de collecte sera analysée de façon approfondie : permet-elle à elle seule de saisir l'effet de mesure dans le cas d'une agrégation simple ? Permet-elle de voir un effet de mesure résiduel dans le modèle après correction par imputation ? L'estimation des paramètres du modèle est-elle sensible à la méthode d'agrégation ?

Dans un second temps, l'analyse se portera sur l'ajout, dans la modélisation, des variables ayant servi à la modélisation du score de propension qui a permis le matching, c'est-à-dire celles captant l'effet de sélection. Réduisent-elles l'effet de mode présent dans le modèle ?

## **Abstract**

The Generation surveys main objective is to facilitate a regular evaluation of the education-to-work transition over the first three years of the young people's working lives. Experiments were conducted on the surveys of the 2010 and 2013 cohorts as part of the system's overhaul programme. The aim of these experiments was to investigate the impact of adding a self-administered, web-based data gathering tool on the quality of the responses. Two different matchings on propensity scores were tried in order to measure mode effects. A data gathering method based on response imputation of a part of the population is compared to a raw dataset. The effects of these imputations and matchings on econometric models are discussed: when using raw variables with mode effect, the mode variable must be included in the model. When using adjusted variables, the estimations slightly depend on the matching method and the reference mode.

## Introduction

Les enquêtes Génération du Centre d'études et de recherche sur les qualifications (Céreq) permettent d'étudier, l'insertion professionnelle des jeunes à l'issue du système éducatif. Il s'agit d'une enquête sur la France entière et sur tous les niveaux et spécialités de formations. Les jeunes d'une Génération sont systématiquement interrogés 3 ans après la sortie du système éducatif. Puis, pour une Génération sur deux, ces jeunes sont réinterrogés à 5 ans et 7 ans (voire 10 ans) après leur sortie.

Le mode de collecte historique est monomode téléphone. Après 20 ans d'existence, le Céreq a décidé de rénover en profondeur ce dispositif de la statistique publique tant d'un point de vue scientifique, sur le champ et le contenu du questionnaire, que méthodologique, sur le mode de collecte notamment. Ainsi, l'enquête Génération se déroulera, à partir de 2020, selon une collecte multimode téléphone-internet. L'introduction du mode internet répond à un double enjeu : un gage de modernité particulièrement adapté à la cible de cette enquête, les jeunes, qui ont même de plus en plus tendance à réclamer cette possibilité de répondre par internet. Par ailleurs, le coût d'une enquête réalisé sur internet est moindre

Pour réussir au mieux ce changement, le Céreq a organisé trois expérimentations en 2013, 2015 et 2016 en parallèle des enquêtes principales. La première en 2013 a consisté à réaliser une enquête monomode internet sur une première interrogation notamment pour se familiariser avec ce nouveau mode de collecte. La deuxième expérimentation en 2015 portait sur une réinterrogation et pour la première fois en mettant en place un protocole multimode internet-téléphone. Plus précisément, il y avait trois protocoles testés. Le premier était monomode internet, à savoir la réinterrogation des répondants à la première expérimentation web afin d'observer l'attrition dans le cadre d'un protocole 100 % web. Le second était un protocole multimode internet/téléphone « classique » incitant à répondre en priorité par internet avec l'introduction du téléphone comme mode de réponse alternatif pour ceux ne souhaitant pas répondre par internet. Le troisième protocole était nommé « échantillon embarqué » (ou imbriqué), qui consistait, sur un sous-échantillon, à affecter aléatoirement un des deux modes de collecte pour permettre de contrôler l'effet de sélection inhérent à une collecte multimode. La troisième expérimentation menée en 2016 est basée sur un protocole multimode « classique », légèrement révisée par rapport à 2015, mais cette fois-ci pour la première fois sur une première interrogation.

Ces expérimentations ont permis de délivrer un certain nombre d'enseignements <sup>[1], [2], [5]</sup> :

- La faiblesse des taux de réponse dans un protocole uniquement web avec notamment aucune relance téléphonique mais uniquement des mails de relance. Ainsi, dans une enquête de grande ampleur, il paraît indispensable de mixer internet avec un autre mode, en l'occurrence avec le téléphone pour Génération. L'animation téléphonique permet de mieux contacter les individus, les inciter à répondre sur internet tout en récupérant leur mail et s'ils sont réticents à répondre sur internet à passer le questionnaire par téléphone.
- L'importance de l'ergonomie des questionnaires (en particulier si elle est mauvaise) sur le risque d'abandon en cours d'enquête. En effet, la « survie » dans le questionnaire a été considérablement améliorée entre l'enquête 2013 et l'enquête 2015. Sans aucun doute, l'amélioration sensible de l'ergonomie entre ces deux expérimentations a contribué à limiter les abandons. L'autre explication est liée au protocole puisque les personnes qui interrompent leur questionnaire sont rappelées pour terminer l'enquête.
- La définition des protocoles multimode et leur efficacité
- La présence d'un effet de mode sur certaines variables a été mise en évidence et a concentré l'essentiel des travaux méthodologiques menés sur ces expérimentations. L'effet de mode se définit comme l'écart entre l'estimation issue des questionnaires internet et ceux issus du téléphone. Cet effet de mode se décompose en un effet de sélection et un effet de mesure.

L'effet de sélection résulte du fait que ce ne sont pas les mêmes structures de population qui répondent sur chaque mode (par exemple : la propension à répondre sur internet augmente avec le niveau du diplôme). L'effet de mesure résulte du fait qu'un même individu ne répond pas de la même manière selon le mode de collecte sur lequel il est interrogé (en autoadministré sur internet et interrogé par un enquêteur par téléphone). Pour mettre en évidence, l'effet de mesure deux méthodes ont pu être éprouvées. La correction de l'effet de sélection se faisant soit en utilisant la méthode économétrique du matching (i.e. en recherchant des contrefactuels sur l'autre mode selon le score de propension), soit en utilisant l'échantillon embarqué (ou imbriqué) construit pour gommer l'effet de sélection. L'échantillon embarqué et la méthode de matching semblent tout deux jouer leur rôle de contrôle de la sélection, avec des inconvénients pour chaque méthode (sensibilité au modèle de sélection pour le matching, taux de réponse dégradé pour l'échantillon embarqué lorsqu'on impose un changement de mode aux enquêtés). Si les effets de mesure détectés sont relativement cohérents entre les deux méthodes, l'évaluation de l'ampleur de l'effet de mesure est très sensible à la méthode choisie.

- La question de l'agrégation des données multimodes en présence d'effet de mesure est un sujet complexe. Plusieurs méthodes ont été proposées : agrégation simple, omission d'un mode, imputation totale ou partielle des données d'un mode... Sans surprise, le choix de la méthode influe beaucoup sur l'estimation.

Cet article propose, à partir de la dernière expérimentation multimode, de poursuivre les investigations sur les effets de mode. Une méthode de matching sur score de propension pour distinguer les effets de sélection des effets de mesure sera à nouveau mobilisée et discutée. Les données seront ensuite agrégées en utilisant différentes méthodes d'imputation pour corriger l'effet de mesure. Cette agrégation sur données imputées sera mise en regard d'une agrégation simple, qui consiste à juxtaposer les données collectées dans chacun des deux modes sans appliquer de traitement spécifique au mode sur les réponses.

Dans une dernière partie, en se mettant dans le rôle d'un chargé d'études qui souhaite réaliser des modélisations économétriques, l'objectif est de regarder l'impact de ces deux méthodes d'agrégation (simple et sur données imputées) sur les estimations de paramètres des modèles incluant des variables sujettes à des effets de mesure (que ce soit en explicatives ou en dépendantes).

## 1. Mise en évidence d'effet de mesure par méthode de matching

Usuellement utilisées dans le cadre de l'évaluation des politiques publiques, le but des méthodes de matching est d'évaluer l'effet d'un traitement T sur une variable d'intérêt Y. Dans le contexte de l'enquête Génération 2013, on considère que le traitement est le fait de répondre au questionnaire par Internet, par rapport à une réponse par téléphone ; les variables d'intérêt seront des variables du questionnaire susceptibles de présenter un effet de mesure, par exemple des variables portant sur l'opinion sur l'emploi des individus, les raisons de son arrêt des études, ou sur le sentiment de discrimination au sein de l'entreprise.

La population suivant le traitement a une variable de traitement T=1 tandis que celle ne le subissant pas a pour valeur T=0. Dans ce cadre, on considère deux variables latentes Y1 et Y0 qui correspondent aux valeurs de Y que l'individu aurait renseigné en répondant, respectivement, par Internet ou par téléphone. On n'observe qu'une seule de ces variables latentes, mais on a la relation :

$$Y = T * Y1 + (1 - T) * Y0.$$

L'hypothèse de cette approche par matching est de considérer que, conditionnellement aux variables qui expliquent la sélection, les variables latentes sont indépendantes du traitement. C'est-à-dire que conditionnellement aux covariables, il n'y a plus d'effet de sélection et l'assignation du traitement peut être considérée comme aléatoire.

L'objectif est d'alors d'estimer la valeur latente non observée. Pour cela on recherche, pour chaque individu qui a reçu le traitement, un contrefactuel qui n'a pas reçu le traitement et qui lui ressemble selon les variables qui expliquent la sélection. Dans le contexte de l'enquête multimode Génération 2013, la détermination d'un contrefactuel de chaque individu a été réalisée par la méthode d'appariement sur score de propension. Si on est dans le cas d'indépendance des variables latentes Y1 et Y0 par rapport aux covariables expliquant la sélection, alors il y a aussi indépendance par rapport à la probabilité de recevoir le traitement, le score de propension (Rosenbaum, Rubin, (1983)).

On estime alors la valeur de la variable latente inobservée Y0 par la variable Y du contrefactuel. Réciproquement, on peut estimer la valeur de la variable latente inobservée Y0 des individus qui n'ont pas suivi le traitement par la valeur Y de leur contrefactuel qui a suivi le traitement.

On peut alors estimer l'effet de mesure en prenant par l'effet moyen du traitement (Average Treatment effect on the Treated ou ATT) :

$$ATT = E(Y1 - Y0 | X, T = 1)$$

### 1.1. Score de Propension

Pour estimer la probabilité de répondre à l'enquête par Internet plutôt que par téléphone pour un individu répondant, un modèle de régression logistique faisant intervenir l'âge de l'individu à la sortie des études, son niveau et le type de son établissement de sortie, ainsi que son sexe a été employé. Les interactions de ces variables sont prises en compte deux à deux au sein du modèle.

Ce modèle fait ressortir que les femmes ont plus tendance à répondre sur Internet que les hommes. De plus, plus le niveau de sortie des études est élevé, plus la propension à répondre par Internet est élevée. L'âge de sortie a un effet différent selon le type d'établissement et le niveau de sortie des études : pour les sortants de formation post BAC, ce sont ceux qui sont sortis des études le plus jeune qui ont le plus tendance à répondre par Internet, alors que pour les sortants de formation de niveau inférieur au bac, la propension à répondre par Internet augmente avec l'âge de sortie des études.

La répartition des scores de propension individus des deux modes est la suivante :

Mode	Moyenne	Minimum	1er Quartile	Médiane	3e Quartile	Maximum
Téléphone	0.443	0.102	0.262	0.486	0.593	0.865
Internet	0.592	0.004	0.518	0.602	0.674	0.946

Tableau : Répartition des scores de propension selon le mode

En regardant la distribution des probabilités de réponse par internet en fonction du mode de réponse, on voit que le score de propension maximum pour les individus répondants par CATI est de 0,865. Les individus répondant par CAWI avec un score plus élevé ne peuvent donc pas trouver de contrefactuel ayant un score de propension proche du leur parmi les répondants par téléphone. On définit le support commun comme l'ensemble des individus qui ont leur score de propension entre le minimum des répondants par internet et le maximum des répondants par téléphone, c'est-à-dire pouvant trouver un contrefactuel ayant un score de propension similaire au leur dans le mode auquel ils n'ont pas répondu.

Ici, le score de propension minimum des répondants par internet est inférieur à celui des répondants par téléphone, cependant sur l'ensemble de la distribution, les scores des répondants Internet sont supérieurs à ceux des répondants par téléphone.

Le matching ne sera effectué que sur les individus du support commun, afin que les paires d'individus aient des scores de propension proches.

## 1.2. Différentes méthodes de Matching

Pour définir le contrefactuel d'un individu, plusieurs méthodes de matching peuvent être employées. Ici, deux méthodes différentes ont été expérimentées. Les deux approches reposent sur un appariement exact selon les variables de niveau de sortie des études de l'individu, et de situation professionnelle à la date d'enquête. Concrètement, le contrefactuel d'un individu sera sorti du système éducatif au même niveau que l'individu de référence, et sera dans la même situation professionnelle ; à savoir en emploi, en recherche d'emploi, en formation, en reprise d'études ou dans une autre situation.

Le matching exact permet de s'assurer que les individus appariés ont les mêmes caractéristiques selon ces variables qui sont déterminantes, à la fois pour la mesure des variables d'intérêt, mais aussi dans le cheminement du questionnaire. Les individus en emploi, par exemple, passent un module de questions sur leur emploi actuel. Au sein de ce module, certaines questions sont sujettes à des effets de mesure, il s'agit alors de pouvoir mesurer cet effet de mesure en appariant ensemble des individus qui ont passé ce module.

La première méthode consiste en un matching sans remise par la méthode du plus proche voisin. Séquentiellement, on apparie un individu répondant par Internet avec l'individu répondant par téléphone ayant le score de propension le plus proche qui n'a pas déjà été apparié auparavant.

Au fur et à mesure que la procédure s'effectue, il reste alors de moins en moins de choix pour appairer les individus car le nombre d'individus répondant par téléphone qui ne sont pas déjà appariés se restreint. Deux individus matchés à la fin de cette procédure peuvent alors avoir des scores de propensions trop éloignés, et donc ne pas être similaires au sens des variables expliquant la sélection. Pour éviter que cela n'arrive, une limite a été fixée dans l'écart entre les scores de propension des individus appariés. Ce caliper est ici de 0,02, c'est-à-dire que le score de propension du contrefactuel d'un individu diffère au plus de 0,02 du sien. Les individus étant en fin de procédure d'appariement risquent alors de ne pas trouver d'individu pouvant leur être apparié. Cette méthode n'apparie donc pas tous les individus. De plus, les individus appariés, ainsi que leur contrefactuel dépendent de l'ordre dans lequel les appariements sont réalisés, ce qui peut impliquer une différence

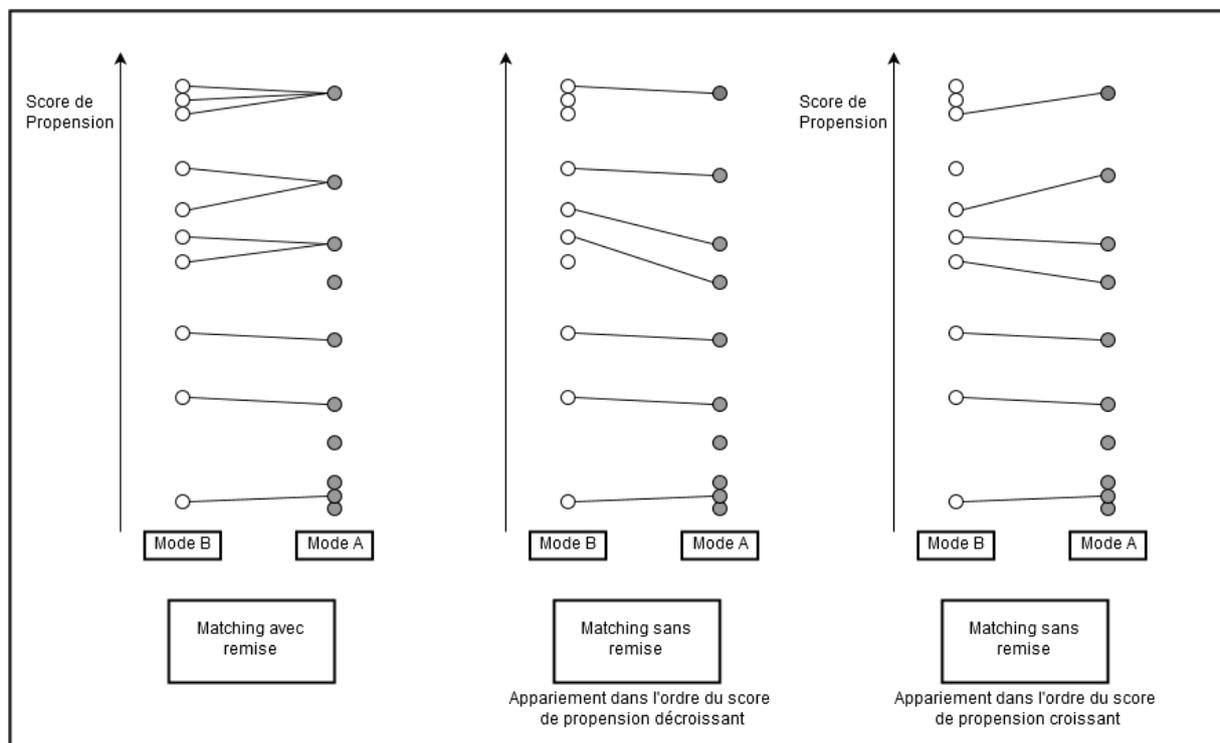
dans l'évaluation de l'effet de mesure. Cependant, le nombre total de paires appariées est peu dépendant de l'ordre et du mode de collecte qui est choisi comme référence pour le matching.

Une seconde méthode de matching consiste à appairer chaque individu répondant par Internet avec l'individu répondant par téléphone qui a le score de propension le plus proche, sans tenir compte des paires déjà existantes. Un même individu répondant par téléphone peut donc être le contrefactuel de plusieurs individus répondant par Internet. Un caliper a été fixé à 0,01.

Cette méthode n'est pas symétrique. Selon que l'on prenne comme variable de traitement « Réponse sur Internet » ou « Réponse par Téléphone », le nombre d'individus matchés, et leurs contrefactuels, sont différents. Pour la suite, deux matchings seront réalisés par cette méthode : un avec le mode internet comme référence, et l'autre avec le téléphone comme référence c'est-à-dire que pour le matching avec téléphone comme référence, il s'agit de trouver des contrefactuels à chaque individu répondant par internet.

### 1.3. Comparaison des différentes méthodes

#### 1.3.1. Exemple



Exemples : Appariement selon différentes méthodes

Dans l'exemple ci-dessus, les individus du mode B sont appariés à ceux du mode A par les différentes méthodes de matching. Le matching avec remise apparie chaque individu du mode B avec l'individu du mode A dont le score de propension est le plus proche. Certains individus du mode A sont appariés à plusieurs individus du mode A, et certains ne sont pas appariés.

Dans le matching sans remise, les individus du mode B sont appariés séquentiellement avec l'individu du mode A ayant le score de propension le plus proche s'il y en a un. Certains individus du mode B ne trouvent pas de contrefactuel car les individus de score de propension proche du mode A sont déjà appariés. Selon que l'appariement soit réalisé dans l'ordre croissant ou décroissant du score de propension, les paires ne sont pas les mêmes, et les individus non matchés non plus.

### 1.3.2. Nombre d'individus appariés

Le matching sans remise étant dépendant de l'ordre dans lequel les individus sont appariés, un tri aléatoire des individus est exercé avant la procédure de matching. Le tableau suivant présente le nombre d'individus appariés pour les différentes méthodes de matching présentées :

	Répondants	Appariés		
		Sans Remise	Avec Remise	
			Référence Téléphone	Référence Internet
Téléphone	1652	968	308	1360
Internet	1800	968	1567	317

Tableau : Nombre d'individus appariés par chaque méthode de matching.

Le matching sans remise ne permet d'apparier que 968 paires d'individus Internet-Téléphone, soit 59% des répondants par téléphone et 54% des répondants par Internet. Cependant il garantit que chaque individu n'est apparié qu'une unique fois.

Si le matching avec remise avec comme mode de référence le téléphone permet d'apparier 1567 individus répondant par internet, chaque individu répondant par téléphone apparié est en moyenne le contrefactuel de 5,4 individus. Si on prend internet pour référence, 1360 individus répondant par téléphone sont appariés, pour un nombre moyen de contrefactuels par répondant internet de 4,3.

Cependant, le nombre de paires auquel fait partie un individu répondant par le mode de référence n'est pas distribué équitablement. La répartition est la suivante :

Référence Téléphone		Référence Internet	
Nombre de contrefactuels Internet	Répondants par Téléphone	Nombre de contrefactuels Téléphone	Répondants par Internet
1	113	1	109
2	62	2	53
3	31	3	38
4	20	4	20
5 à 10	47	5 à 10	55
Plus de 10	34	Plus de 10	33

Tableau : Nombre de contrefactuels par individu du mode de référence

Plus de la moitié des individus du mode de référence sont appariés avec plus d'un individu. Avec le téléphone comme mode de référence, il y a un individu qui est le contrefactuel de 68 individus internet. Les individus étant le contrefactuel du plus grand nombre de répondants sont ceux qui sont les plus atypiques de leur mode. Par leurs nombreux appariements, ils ont une plus grande importance dans l'estimation de l'effet de mesure, en effet, dans le calcul de l'ATT, leur valeur compte d'autant plus que le nombre d'individus avec lesquels ils sont appariés est élevé.

### 1.3.3. Évaluation des effets de mesure

L'effet de mesure a été évalué sur de nombreuses variables d'intérêt dont les trois suivantes qui serviront d'exemples pour la suite. La première, a priori sans effet de mesure est la situation d'habitat de l'individu à la date d'enquête : chez les parents, seul ou en couple. Cette variable factuelle sert de témoin pour l'évaluation de l'effet de mesure. La seconde variable concerne la réponse à la question OP8 : « Nous allons parler de votre point de vue sur le travail en général. Les aspects suivants vous semblent-ils très importants, importants, peu important ou pas du tout important ? ». L'aspect étudié par la suite est le numéro 6 : « L'intérêt du poste ».

La dernière variable concerne la question Q25 : « Quelles sont les raisons pour lesquelles vous avez arrêté vos études en 2013 », qui est une question à réponses multiples présentée sur Internet sous la forme d'une liste d'items à sélectionner. L'item étudié ici est l'item 7 : « Pour entrer dans la vie active ».

L'évaluation des effets de mesure sur les variables d'intérêt dépend de la méthode de matching utilisée. Le matching sans remise étant aléatoire, l'effet de mesure a été calculé sur 100 appariements différents afin d'analyser la variabilité de cette méthode pour l'estimation de l'effet de mesure. Le tableau suivant montre l'effet de mesure mesuré pour chaque méthode pour chaque modalité de ces trois variables :

Variable	Modalité	Avec remise		Sans remise			
		Référence Téléphone	Référence Internet	Effet moyen	Ecart-type	Effet minimum	Effet maximum
Situation d'habitat à la date d'enquête	Vit chez ses parents	1.14	-5.86	0.02	0.75	-1.65	2.59
	Vit en couple	-4.78	-0.13	-1.18	0.82	-3.02	0.98
	Vit seul	3.64	6.00	1.15	0.74	-0.67	2.92
Intérêt du poste	Très important	-11.19	-17.24	-14.09	0.81	-16.05	-12.47
	Important	7.22	16.55	14.13	0.79	12.55	16.11
	Peu important	3.87	0.52	0.06	0.37	-0.86	0.93
	Pas du tout important	0.10	0.17	-0.10	0.11	-0.32	0.11
Arrêt des études pour entrer dans la vie active	Non	-15.25	-26.38	-23.26	1.28	-26.42	-20.40
	Oui	15.25	26.38	23.26	1.28	20.40	26.42

*Note de lecture : Les individus répondant par téléphone déclarent en moyenne 1,14 point de moins habiter seul que les individus répondant par Internet, à l'aide de l'estimation réalisée par matching avec remise avec pour référence le téléphone.*

Tableau : Effet de mesure

En moyenne, l'effet de mesure estimé à l'aide du matching sans remise est proche de zéro pour les trois modalités de la variable d'habitat à la date de l'enquête. Cependant, ce n'est pas le cas pour les matching avec remise. Une hypothèse pour expliquer ces différences serait qu'il reste une partie de l'effet de sélection que ces matching n'auraient pas réussi à capter en raison du faible nombre d'individus du mode de référence appariés par ces méthodes.

Pour la variable d'intérêt du poste, les estimations de l'effet de mesure vont dans le même sens pour chacune des modalités « Très important » et « Important » quelle que soit la méthode. Cependant, l'estimation de l'ampleur de cet effet n'est pas la même selon la méthode. Même au sein de la méthode sans remise, il y a 3,5 points de différence entre l'estimation de l'effet de mesure la plus élevée et la plus faible. Toutefois, chaque estimation montre un fort effet de mesure pour ces modalités : plus de 10 points de différence entre internet et téléphone. Les individus répondant par internet ont plus tendance à répondre « Très Important » que les individus répondant par téléphone, ce qui est conforme avec le fait que les individus aient plus tendance à répondre par les modalités extrêmes sur internet qu'au téléphone (Heerwegh, Loosveldt, 2008).

Les individus ayant peu choisi les modalités « Peu important » et « Pas du tout important » (6% au total que ce soit par internet ou téléphone), l'effet de mesure estimé est mécaniquement faible.

Les individus répondants par internet ont moins tendance à cocher « Pour entrer dans la vie active » comme raison d'arrêt des études, que les individus à répondre « Oui » par téléphone. Un fort effet de mesure est estimé par les trois approches de matching : plus de 15 points d'écart entre les deux modes. Deux effets cumulatifs peuvent expliquer cette différence. D'un côté, il peut y avoir un effet de désirabilité sociale : les individus peuvent voir le fait de vouloir entrer dans la vie active comme une raison positive d'arrêter ses études, et ainsi vouloir donner une bonne image d'eux à l'enquêteur au téléphone. D'un autre côté, il apparaît que les individus ont tendance à choisir moins de modalités sur Internet qu'au téléphone lors de questions à réponses multiples (Christian, Dillman, Smyth, 2008).

## 2. Agrégation en présence d'effet de mesure

### 2.1. Enjeux

Dans une enquête où plusieurs modes de collecte sont utilisés pour récolter les mêmes informations, des effets de mode peuvent apparaître et donner des estimations différentes en fonction du mode de collecte. Il faut alors trouver un moyen de réunir les jeux de données collectés selon les différents modes qui limite l'impact des effets de mesure sur les estimations.

Dans le cas des enquêtes Génération, des contraintes liées à l'exploitation de la base de données doivent être respectées. Le fichier d'utilisation doit pouvoir être facilement pris en main par un chargé d'études peu renseigné sur les problèmes liés à une collecte multimode. La base de données ne doit comporter qu'un nombre limité de poids pour chaque individu, ce qui empêche par exemple de pondérer spécifiquement chaque variable pour atténuer les erreurs de mesure.

Il n'est pas non plus envisageable de ne livrer aux chargés d'études uniquement les estimations des paramètres corrigés des effets de mesure, car les paramètres d'études peuvent être en nombre important ; le chargé d'études ne serait pas autonome dans l'exploitation des données s'il doit faire appel à une personne pour réaliser chaque estimation d'un paramètre complexe.

Par la suite, les deux méthodes qui sont détaillées présentent l'avantage de pouvoir livrer un jeu de données standard, semblable à celui d'une enquête monomode, à l'intention de personnes voulant analyser les données.

### 2.2. Différentes méthodes expérimentées

#### 2.2.1. Agrégation simple

Une première solution est de fusionner les deux jeux de données et de ne rien mettre en place pour corriger les effets de mesure. Cette solution se comprend si on considère que les réponses des individus ne sont pas des vérités absolues et que si un effet de mesure amène à des différences dans les estimations des paramètres en fonction du mode de réponse, la vraie valeur se trouve sûrement entre ces estimations. Autrement dit, chaque mode a un biais de mesure, et ceux-ci se compensent dans le cas de l'agrégation simple. Cette affirmation peut être remise en question, et est le point de départ du développement des autres méthodes d'agrégation des données.

L'agrégation simple est en pratique facile à mettre en œuvre pour la personne chargée de réaliser l'enquête et permet de conserver toutes les réponses données par les individus. Le fait qu'il n'y ait pas d'imputation permet de ne pas dégrader la variance des estimateurs. Un autre avantage de cette méthode est qu'elle ne nécessite pas de choisir un mode de référence, et donc de faire d'hypothèse sur un mode qui représenterait plus fidèlement la réalité que les autres.

Les liens entre les variables sont conservés, et un chargé d'études peut utiliser le fichier de données comme il l'aurait fait avec des données provenant d'une enquête monomode. De plus pour les variables ne présentant pas d'effet de mesure, cette méthode permet d'avoir la même précision que lors d'une enquête monomode, sans y ajouter de biais.

Cependant, pour les variables comportant un effet de mode, dans le cadre d'enquêtes répétées, cette méthode peut poser problème. En effet, il faut que la structure de la répartition entre les différents modes soit la même entre les enquêtes pour pouvoir les comparer en tenant compte de l'effet de mode, sinon l'estimation des variables sujettes aux erreurs de mesure sera sensible à l'évolution de la répartition des individus au sein des modes dans le temps. Une solution envisageable pour contrer ce problème serait de caler la structure des répondants au sein des modes. Cependant cela ne semble pas évident à réaliser en pratique, et dégraderait la précision des estimations et donc la qualité des comparaisons par rapport à d'autres méthodes si la structure évolue grandement.

## 2.2.2. Agrégation avec imputation pour corriger l'effet de mesure

### 2.2.2.1. Méthode

Cette méthode (Legleye, De Peretti, Razafindranovona (2018)) repose sur l'imputation des variables présentant un effet de mesure sur une partie de la population des répondants. Elle est plus compliquée à mettre en œuvre, et nécessite de choisir un mode de référence A dont on considère que les réponses sont plus fiables.

Ce choix n'est pas évident dans le contexte des enquêtes Génération. En effet, historiquement les enquêtes sont collectées par CATI uniquement. Dans le but de comparer les enquêtes futures avec celles ayant déjà eu lieu, considérer le téléphone comme référence semblerait judicieux. Cependant pour certaines questions d'opinion sujettes au biais de désirabilité sociale, il semble que les réponses par internet correspondent plus à la réalité que les réponses par téléphone. Choisir internet comme référence pour ces questions semble alors être une solution préférable. À des fins méthodologiques, dans la suite de l'article, les deux modes vont servir l'un après l'autre de référence.

La méthode consiste à repérer les individus dont les réponses sont spécifiques au mode alternatif B, puis de remplacer ces réponses par des réponses qui auraient pu être données dans le mode de référence.

Après avoir apparié les individus par une des méthodes de matching présentées auparavant, pour une variable d'intérêt Y, il y a deux types de paires d'individus matchés. D'un côté, les vrais jumeaux sont des individus matchés qui ont la variable Y identique, de l'autre côté, les faux jumeaux diffèrent en Y, bien qu'ils soient similaires selon les variables expliquant la sélection car appariés entre eux.

Pour annuler l'effet de mesure, deux propositions sont faites. La première consiste, pour chaque variable, à imputer tous les faux jumeaux du mode B par la valeur de leur contrefactuel. De ce fait, sur l'ensemble des individus matchés, toutes les paires deviennent des vrais jumeaux après imputation, et l'effet de mesure devient nul par construction sur la population des individus matchés. L'effet de mesure est alors mécaniquement réduit sur l'ensemble de la population.

Cependant, en cas d'effet de mesure important, cela peut amener une grande partie des individus du mode B à être imputés. Pour compenser cela, la deuxième proposition suggère de n'imputer que les individus nécessaires à annuler l'effet de mesure.

Le principe est d'imputer de manière stochastique Y une partie des faux-jumeaux du mode B selon un modèle réalisé sur l'ensemble des individus matchés du mode A. Il s'agit alors de choisir la bonne proportion d'individus à imputer pour, qu'avec ce modèle, on ait annulé l'effet de mesure sur l'ensemble des individus matchés ; c'est-à-dire que pour une variable catégorielle, les proportions de chaque modalité sont les mêmes sur le mode A et le mode B.

Cette approche est plus complexe à mettre en œuvre mais permet d'avoir moins d'individus imputés, et donc de conserver une plus grande partie des réponses données par les individus, tout en annulant l'effet de mesure sur les individus appariés.

### 2.2.2.2. Exemple

Avant Imputation			Après imputation		
Individus appariés du mode B			Première approche		
Individu	Y	Y contrefactuel	Individu	Y imputé	Y contrefactuel
1	1	1	1	1	1
2	0	0	2	0	0
3	1	1	3	1	1
4	1	0	4	0	0
5	1	0	5	0	0
6	1	0	6	0	0
7	0	1	7	1	1
8	0	1	8	1	1
<div style="background-color: #c8e6c9; padding: 2px; display: inline-block; margin-bottom: 5px;">Vrais jumeaux</div> <div style="background-color: #ffe0b2; padding: 2px; display: inline-block; margin-bottom: 5px;">Faux jumeaux</div>			Deuxième approche		
Individu	Y	Y contrefactuel	Individu	Y imputé	Y contrefactuel
1	1	1	1	1	1
2	0	0	2	0	0
3	1	1	3	1	1
4	1	0	4	1	0
5	1	0	5	0	0
6	1	0	6	1	0
7	0	1	7	0	1
8	0	1	8	0	1
			En rouge : données imputées		

Exemple : Imputation par les deux approches du même jeu de données.

Sur l'exemple, 5 répondants sur 8 du mode B ont Y=1, mais seuls 50% des individus du mode de référence ont Y=0. Il y a 3 paires de vrais jumeaux dont les valeurs de Y ne sont pas imputées.

Par la première approche, tous les faux jumeaux sont imputés à la valeur de leur contrefactuel, il y a donc 5 imputations pour passer de 62,5% à 50% de Y=0 dans le mode B.

Par la seconde approche, le modèle d'imputation va imputer en moyenne 50% des individus à imputer à la valeur Y=1. En effet, 50% des individus du mode de référence ont Y=1.

Parmi les faux jumeaux du mode B, il y a 3 individus dont Y=1, et 2 dont Y=0. En imputant ces 5 individus, il y aura alors en moyenne 2,5 individus dont Y=1 ; et donc en moyenne 56,25% de l'échantillon total qui aura Y=1, ce qui ne permet pas d'annuler l'effet de mesure entre les deux modes. Il s'agit alors de choisir quels individus imputer à l'aide de ce modèle pour atteindre 50% de valeur Y=1 dans l'échantillon du mode B après imputation.

En imputant 2 individus dont Y=1 par ce modèle, et conservant la valeur des 3 autres, il y a alors en moyenne 50% des imputations, soit un des deux individus, imputés à Y=1 ; 2 faux-jumeaux non-imputés avec Y=0, et un non-imputé avec Y=1, en plus des trois vrai-jumeaux. Après imputation, il y a alors en moyenne 50% des Y=1 sur l'ensemble de l'échantillon. Dans cet exemple, les individus imputés sont les individus 4 et 5.

Après imputation par la première approche, tous les individus ont la même valeur que leur contrefactuel, ce qui n'est pas le cas avec la seconde approche. Cependant la seconde approche permet de n'imputer que 2 individus au lieu des 5 de la première approche.

### 2.2.3. Estimations

L'objet de cette partie est d'étudier l'effet des différentes méthodes d'agrégation des données sur l'estimation des proportions des modalités d'une variable sujette à des effets de mesure.

La variable d'intérêt est ici OP8\_6, réponse à la question « Nous allons parler de votre point de vue sur le travail en général. Les aspects suivants vous semblent-ils très importants, importants, peu important ou pas du tout important ? », pour l'aspect « L'intérêt du poste ».

Les individus se ventilent de la manière suivante dans chacun des deux modes de collecte :

OP8_6	Téléphone	Internet
Très important	33.2	57.60
Important	60.32	37.15
Peu important	6.28	4.89
Pas du tout important	0.2	0.36

Tableau : Réponses de la population de chaque mode à la question OP8\_6

Comme vu précédemment, les individus répondant par internet sont plus nombreux que ceux répondant par téléphone à trouver l'intérêt du poste très important. Ces résultats serviront de base de comparaison pour les différentes méthodes d'agrégation.

En considérant l'agrégation simple, les résultats sont les suivants :

OP8_6	Pourcentage
Très important	43.17
Important	50.85
Peu important	5.71
Pas du tout important	0.26

Tableau : Proportions de réponse à la question OP8\_6 pour l'agrégation simple.

Les proportions de réponse de chaque modalité se trouvent entre celles données par téléphone et par Internet. En utilisant cette méthode, une hypothèse forte est implicitement formulée : aucun des deux modes utilisés n'est la vérité. La vérité se trouve entre les résultats donnés par chacun des modes. Autrement dit, chaque mode comporte un biais de mesure, et ceux-ci se compensent dans le cas de l'agrégation simple.

Le tableau suivant synthétise les résultats obtenus par la méthode d'agrégation avec imputation. Celle-ci peut être réalisée à partir d'un matching sans ou avec remise, de manière déterministe ou stochastique, et avec pour mode de référence Internet ou téléphone.

	Référence téléphone				Référence Internet			
Matching sans remise	Déterministe		Stochastique		Déterministe		Stochastique	
	OP8_6	Pourcent age						
	Très important	39.06	Très important	39.50	Très important	47.50	Très important	46.91
	Important	54.85	Important	54.47	Important	46.47	Important	46.48
	Peu important	5.86	Peu important	5.55	Peu important	5.72	Peu important	6.24
	Pas du tout important	0.23	Pas du tout important	0.48	Pas du tout important	0.31	Pas du tout important	0.37
	Nombre d'imputations :	506	Nombre d'imputations :	431	Nombre d'imputations :	506	Nombre d'imputations :	488
Matching avec remise	Déterministe		Stochastique		Déterministe		Stochastique	
	OP8_6	Pourcent age						
	Très important	39.13	Très important	37.45	Très important	51.35	Très important	51.78
	Important	53.46	Important	55.85	Important	43.00	Important	42.47
	Peu important	7.11	Peu important	6.29	Peu important	5.47	Peu important	5.08
	Pas du tout important	0.30	Pas du tout important	0.42	Pas du tout important	0.19	Pas du tout important	0.66
	Nombre d'imputations :	749	Nombre d'imputations :	406	Nombre d'imputations :	1359	Nombre d'imputations :	652

Tableau : Résultats de la variable OP8\_6 pour les différentes méthodes d'agrégation avec imputation

Un premier résultat qui apparaît en regardant ce tableau est que quels que soient le matching et la référence, les méthodes d'imputation déterministe et stochastique donnent des estimations très proches sur chaque modalité. Les deux approches permettent de diminuer l'effet de mesure de manière similaire sur l'ensemble de la population. Cependant, la méthode d'imputation stochastique réalise cette correction de l'effet de mesure à l'aide de moins d'imputations, et permet de conserver plus de valeurs réellement données par les individus enquêtés.

Une deuxième observation est que les méthodes d'agrégation réduisent le biais de mesure dans le sens de la référence, sans pour autant atteindre la valeur estimée sur la population de référence uniquement. En effet, pour les méthodes d'agrégation avec référence téléphone, le pourcentage de réponse « Très important » diminue par rapport à l'agrégation simple, mais n'atteint pas les 33,2 % estimés sur les répondants téléphone uniquement. C'est l'effet de sélection, qui n'est pas affecté par la correction de l'effet de mesure par cette méthode, qui crée cette différence. Ainsi, les 33,2 % de réponse « Très important » concernaient essentiellement des individus répondant par téléphone, avec un niveau de sortie inférieur à la moyenne de la population. Les plus diplômés considérant l'intérêt du poste plus souvent « Très important » que le reste de la population, l'estimation sur la population des répondants par téléphone sous-estime la réalité. L'estimation par la méthode d'imputation avec comme référence le téléphone, c'est-à-dire sur l'ensemble de la population, permet d'avoir une estimation sur la population globale, en ayant la même hypothèse : le téléphone permet d'avoir une qualité de réponse supérieure à internet, et est donc privilégié.

En observant les résultats observés pour les imputations avec internet comme référence, un constat est que les estimations réalisées avec le matching avec remise sont plus élevées que celles réalisées avec le matching sans remise. L'effet de mesure est plus corrigé en employant ces méthodes : l'estimation est plus proche de ce qui est obtenu sur les répondants par internet uniquement. En effet, la méthode d'agrégation des données par imputation ne corrige l'effet de mesure que sur les individus appariés. L'hypothèse implicite est alors que l'effet de mesure est porté par ces seuls individus. Or il peut paraître légitime de penser que tous les individus répondants participent de la même manière à l'enquête, et contribuent tout autant à l'effet de mesure. L'appariement avec remise créant plus de paires d'individus, le nombre d'individus sur lequel l'effet de mesure est corrigé est plus important, et ainsi la correction est plus importante. Cependant, le nombre d'individus du mode de référence apparié est plus faible que lors d'un matching avec remise. Il y a donc de plus fortes chances d'obtenir une proportion importante d'individus aux réponses atypiques qui viendraient polluer le modèle d'imputation.

### 3. Modélisation en présence d'effet de mesure

Dans la partie précédente, nous avons établi que la présence d'un effet de mesure sur une variable posait un vrai problème d'estimation et dépendait fortement du choix fait sur la méthode d'agrégation (agrégation simple VS agrégation après imputation).

L'objectif de cette partie est d'aller au-delà de la simple statistique descriptive et d'analyser si le problème subsiste lorsque l'on s'intéresse à un modèle économétrique. Est-ce que les paramètres estimés sont dépendants de la méthode d'agrégation choisie ? Peut-on obtenir des bonnes estimations avec l'agrégation simple en introduisant de différentes manières l'indicatrice du mode de collecte dans le modèle ? L'effet de cette indicatrice du mode a-t-il disparu lorsqu'on mobilise l'agrégation par imputation ?... Les problèmes sont-ils identiques selon la position de la variable présentant un effet de mesure, en dépendante ou en explicative ?... Pour aborder toutes ces questions, différents modèles de régression logistiques seront mis en œuvre et comparés.

#### 3.1. Variable dépendante avec un effet de mesure

Pour toute cette partie, nous mobilisons la variable indicatrice des individus qui déclarent que l'intérêt du poste est un critère très important dans le module « opinion sur l'emploi » (variable *int*). Cette variable présente un effet de mesure important (selon les méthodes, l'ATT se situe entre -17,2 et -11,2) et cela s'observe directement sur la statistique descriptive (qui cumule effet de sélection et effet de mesure) avec 24 points d'écart entre les deux modes.

L'intérêt du poste est très important	mode		Ensemble (agrégation simple)
	Téléphone	Internet	
non	66,8	42,4	56,8
oui	33,2	57,6	43,2

Nous allons modéliser cette variable selon plusieurs stratégies :

- Sans aucune précaution relative à l'effet de mesure, c'est-à-dire en mobilisant l'agrégation simple et sans introduire dans le modèle de variable représentant le mode de collecte (modèle M1).
- En introduisant la variable mode de collecte en indicatrice simple dans le modèle (modèle M2) puis dans un second temps en interaction avec les autres variables (modèle M3).
- Enfin nous réaliserons le même modèle mais sur données imputées en faisant varier la méthode de matching, la méthode d'imputation et la référence modale (modèles M4 à M11).

Les variables explicatives retenues ne présentent pas d'effet de mesure, seule la dépendante porte un effet de mesure.

##### 3.1.1. Agrégation simple avec l'introduction ou non de la variable mode de collecte

Le premier modèle logistique pourrait être un modèle construit classiquement par un chargé d'études qui disposerait du fichier d'enquêtes pondéré sur l'ensemble des individus (agrégation simple) sans être informé de la présence d'un effet de mesure sur cette variable et du mode de collecte utilisé par les individus.

Dans ce modèle M1, une sélection de variables explicatives significatives a été retenue. Le niveau de sortie influe fortement sur la variable *int* (intérêt du poste très important). Les autres variables qui entrent dans le modèle sont le sexe, la situation actuelle du jeune (en emploi, au chômage, en reprise d'études, en formation, autres situations), des variables synthétique sur le parcours sur le marché du travail (pourcentage de temps passé sur les trois premières années de vie active dans les différentes situations, nombre de périodes d'emploi dans des établissements différents, indicatrice sur l'accès au

premier emploi au-delà de la première année), le mode de cohabitation actuelle (seul, en couple, chez les parents), le fait de résider dans un quartier prioritaire de la politique de la ville (au moment de la sortie des études).

Le modèle retenu est très probablement améliorable mais l'objectif ici n'est pas de trouver le meilleur modèle explicatif mais simplement d'étudier comment les différents indicateurs liés à la modélisation (qualité du modèle, significativité des variables, odds ratios) sont perturbés ou non par la présence d'un effet de mesure et par les stratégies pour tenter de l'annuler ou du moins le limiter.

Dans tous les modèles ultérieurs, la même sélection de variables est retenue indépendamment de leur significativité dans le nouveau modèle.

		<b>Modèle M1</b>	<b>Modèle M2</b>	<b>Modèle M3</b>			
		Sans la variable mode de collecte	Avec l'indicatrice du mode en effet principal	Avec l'indicatrice du mode en effet principal et croisé			
Indicateur de qualité globale du modèle (aire sous la courbe de ROC)		67,2	68,5	68,8			
<b>Analyse des effets Type 3</b>							
<b>Variables explicatives</b>	<b>DDL</b>	<b>Khi-2 de Wald</b>	<b>Pr &gt; khi-2</b>	<b>Khi-2 de Wald</b>	<b>Pr &gt; khi-2</b>	<b>Khi-2 de Wald</b>	<b>Pr &gt; khi-2</b>
mode	1	.	.	170,3	<.0001	65,7	<.0001
niveau de sortie (nisor)	6	628,5	<.0001	400,7	<.0001	327,1	<.0001
sexe	1	49,9	<.0001	42,9	<.0001	41,1	<.0001
situation actuelle (sitde)	4	33,6	<.0001	33,8	<.0001	36,3	<.0001
Part de temps passé au chômage (pccho)	1	30,4	<.0001	25,1	<.0001	23,4	<.0001
Part de temps passé en inactivité (pcina)	1	28,8	<.0001	24,0	<.0001	22,5	<.0001
Part de temps passé en étude ou formation (pcetufor)	1	24,2	<.0001	20,3	<.0001	19,3	<.0001
part de temps passé en emploi (pcomp)	1	21,8	<.0001	17,8	<.0001	16,5	<.0001
nombre de périodes d'emploi (nsemp)	1	20,3	<.0001	13,5	0,0002	18,0	<.0001
mode de cohabitation actuel (habde)	2	14,0	0,0009	15,7	0,0004	17,7	0,0001
quartier prioritaire (QPV)	2	9,4	0,0093	3,3	0,1969	2,7	0,2609
accès à un premier emploi en plus de 12 mois (tapesup12)	1	6,5	0,0107	2,2	0,1361	1,7	0,1945
mode*nisor	6	.	.	.	.	59,1	<.0001
mode*nsemp	1	.	.	.	.	9,3	0,0023
mode*sexe	1	.	.	.	.	4,2	0,0396

Le second modèle (M2) reviendrait à informer le chargé d'études que la variable étudiée (*int*) présente un effet de mesure et à lui mettre à disposition l'indicatrice du mode de collecte<sup>1</sup>. Dans ce cas, pour tenir compte de l'effet de mesure, celui-ci décide d'introduire celle-ci en effet principal dans son modèle.

Le mode de collecte ressort sans surprise très nettement significative dans le modèle. Son introduction a modifié la lecture que l'on peut faire du modèle.

Le modèle M1 étant emboîté dans M2, la qualité du modèle M2 s'en retrouve mécaniquement amélioré, tout en restant dans le même ordre de grandeur, en regardant ici le critère de l'aire sous la courbe de Roc (mais le constat est le même avec d'autres indicateurs de qualité, tel que l'AIC par exemple).

Concernant la significativité des variables, l'introduction de l'indicatrice du mode vient perturber l'analyse. En effet, les variables *QPV* et *tapesup12*, jugées significatives dans le modèle M1, ne le sont désormais plus dans le modèle M2. L'influence de la variable *nisor* est toujours importante dans le modèle mais son effet a semble-t-il baissé. Dans une moindre mesure, d'autres variables sont concernées telles que *sexe* et *nsemp*. En réalité, l'omission de la variable significative *mode* dans le modèle M1 provoque le report de son « pouvoir explicatif » sur les variables avec lesquelles elle est liée. On retrouve donc sans surprise ici des variables qui expliquent l'effet de sélection lié au mode de collecte. Comme les deux modèles contiennent l'essentiel des variables expliquant l'effet de sélection, l'introduction de la variable *mode* dans M2 devrait essentiellement porter l'effet de mesure.

La baisse de significativité voire la non significativité de certaines explicatives est gênante en termes d'analyse et suggère que le modèle M1 n'est plus concevable en présence d'une dépendante avec effet de mesure. Ce constat alarmant doit cependant être nuancé avec l'examen des paramètres du modèle. Pour ce faire, le graphique suivant présente les oddsratios des deux modèles.

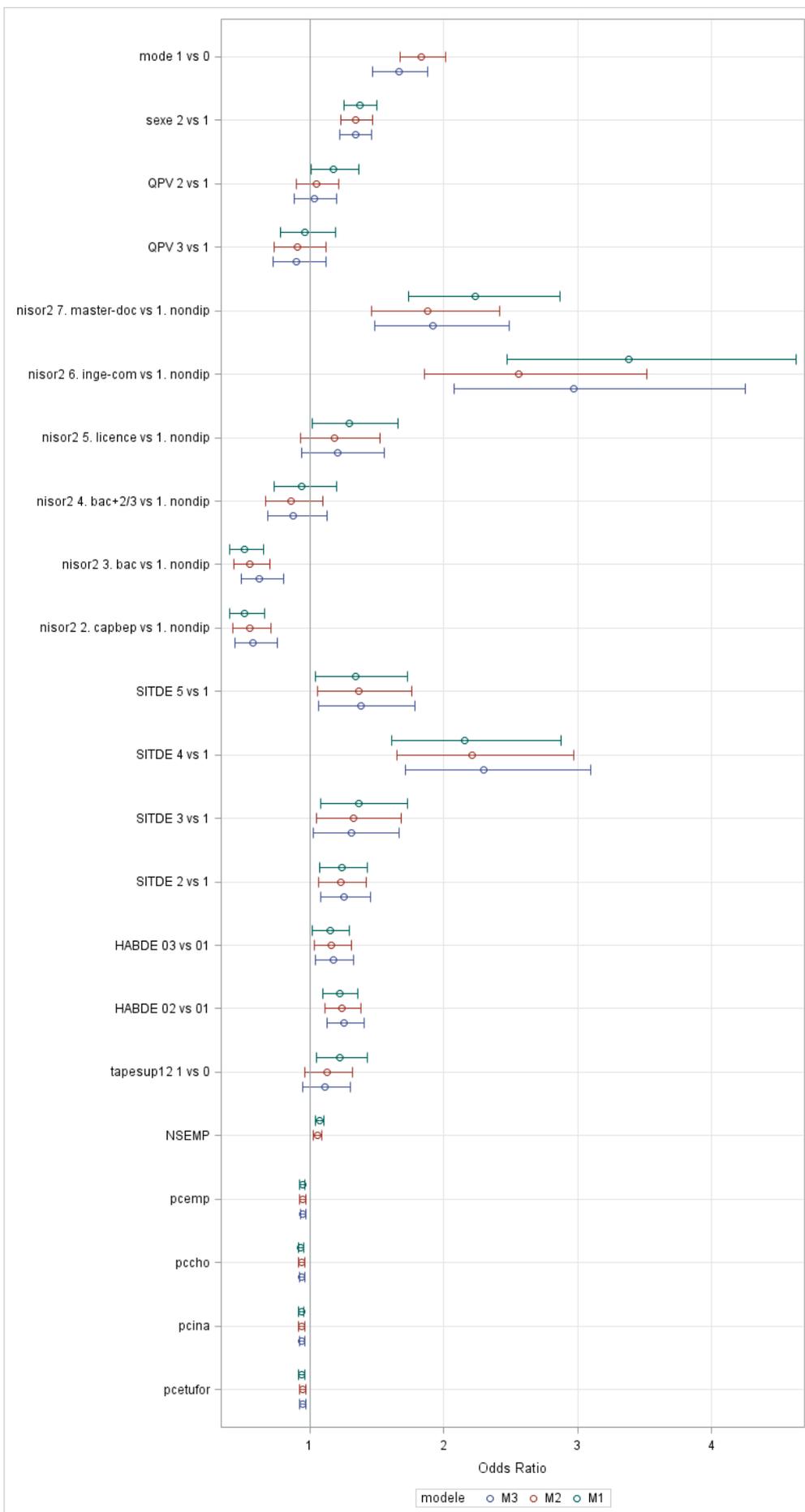
Ainsi, en comparant l'ensemble des oddsratios des deux modèles, on constate que ceux-ci sont très proches d'un modèle à l'autre (les intervalles de confiance de chaque oddsratios se recoupant très largement). Cela est rassurant en termes d'interprétation du modèle qui semblerait peu impacté par l'effet de mesure que l'on introduise la variable de mode ou non.

Au passage, cette modélisation permet de remarquer qu'un répondant sur internet multiplie par 1,8 le « risque » de déclarer que l'intérêt du poste est un critère très important par rapport à un répondant sur téléphone, toutes choses égales par ailleurs.

Ce modèle fait l'hypothèse que la variable indicatrice *mode* joue uniformément sur la probabilité de l'évènement *int=1* quelles que soient les autres caractéristiques présentes dans le modèle. Le modèle M3 ne fait plus cette hypothèse et autorise le fait que l'effet de la variable *mode* soit différencié selon les caractéristiques des individus. Pour ce faire, on introduit toutes les interactions entre les variables du modèle et l'indicatrice du *mode* (en laissant les effets principaux de chaque variable). Parmi, toutes les interactions testées, seules trois se révèlent significatives : *mode\*nisor*, *mode\*nsemp*, *mode\*sexe*.

---

<sup>1</sup> La grande majorité des répondants n'utilisent qu'un mode de collecte mais certains basculent d'un mode à l'autre en cours d'enquête. En toute rigueur, il faudrait donc utiliser une indicatrice pour chaque question qui indique quel mode a été utilisé pour y répondre. Pour éviter d'alourdir le raisonnement, une seule variable indicatrice du mode de collecte est créée (il s'agit du mode principal de réponse).



Graphique des oddsratio des 3 modèles avec agrégation simple

Le modèle M3 englobant le modèle M2 (qui lui-même englobe M1), sa « qualité » s'en trouve mécaniquement augmentée mais avec un gain tout de même limité. En termes de significativité de variables (autres que les trois interactions introduites), les conclusions sont les mêmes que sur M2, à savoir la non significativité des variables *QPV* et *tapesup12*. En termes d'interprétation, et en se référant à nouveau au graphe des oddsratios des trois premiers modèles (ci-dessus), le constat est en apparence toujours rassurant : l'introduction de la variable mode en interaction avec certaines explicatives ne vient pas (trop) perturber les oddsratios globaux (les intervalles de confiance se recoupant très largement encore).

Cela étant, dans le modèle M3, les oddsratios en présence d'interaction permettent de révéler des comportements différenciés selon le mode. Prenons l'exemple de la variable niveau de sortie dont l'interaction avec le mode ressort significatif dans le modèle. En s'intéressant à l'oddsratio des sortants bacheliers par rapport au sortants non diplômés, on constate dans le modèle M3 que sur le mode internet (mode=1), le comportement de réponse à la variable *int* est identique entre ces deux niveaux de sortie, toutes choses égales par ailleurs (OR=0,99). En revanche sur le mode téléphone, les bacheliers déclarent moins souvent que les non diplômés que l'intérêt du poste est un critère très important (*int*=1) (OR=0,63, avec un intervalle de confiance qui ne se recoupe pas avec celui de l'OR quand mode=1). Lorsqu'on regarde l'effet « moyen » de la variable sur l'ensemble des individus, l'OR passe à 0,63 et devient « comparable » avec les deux premiers modèles (0,52 et 0,55 avec les trois intervalles de confiance qui se recoupent). Sur cette base, on conclurait à un effet significatif de la modalité bachelier par rapport à la modalité non diplômé. Sans anticiper sur la suite, on peut introduire la notion de référence et considérer que les conclusions du chargé d'études ne seront pas les mêmes selon qu'il considère les réponses par internet plus justes, ou celles par téléphone.

MODELE	EFFET	OR	Orlow	Orup
M1	nisor2 3. bac vs 1. nondip	<b>0,52</b>	0,41	0,65
M2	nisor2 3. bac vs 1. nondip	<b>0,55</b>	0,43	0,7
M3	nisor2 3. bac vs 1. nondip dans <i>mode=1</i>	<b>0,99</b>	0,67	1,48
	nisor2 3. bac vs 1. nondip dans <i>mode=0</i>	<b>0,39</b>	0,29	0,53
	nisor2 3. bac vs 1. nondip (OR "moyen")	<b>0,63</b>	0,49	0,8

### 3.1.1. Modèle sur données imputées

L'idée est ici de faire le même modèle mais non plus sur la variable d'origine *int* mais sur la variable corrigée selon les différentes méthodes d'imputation présentées dans la partie précédente, en faisant varier les méthodes de matching et d'imputation ainsi que la référence modale.

Agrégation	Variable mode dans le modèle	Imputation	Matching	Référence	Modèle	Aire sous la courbe de ROC
simple	non				M1	67,2
	effet principal				M2	68,5
	effets croisés				M3	68,8
sur données imputées	effet principal	déterministe	sans remise	tel	M4	67,9
				web	M5	66,8
			avec remise	tel	M6	69,0
				web	M7	63,7
		stochastique	sans remise	tel	M8	69,2
				web	M9	67,7
			avec remise	tel	M10	67,8
				web	M11	64,6

Le tableau indique un indicateur de qualité de chaque modèle testé. Les modèles M4 jusqu'à M11 correspondent aux différentes variantes pour imputer la variable *int* présentant un effet de mesure. Ces modèles ne sont plus emboîtés car c'est à chaque fois la variable dépendante qui est modifiée, les explicatives restant inchangées. Sur la base de ce seul critère, la qualité des modèles logistiques sur données imputées varie sensiblement selon les différentes méthodes (entre 63,7 et 69,2). Par rapport au modèle de référence M2 ou M3, seuls deux modèles « font mieux » M8 (matching sans remise avec imputation stochastique avec référence téléphone) et M6 (matching avec remise, imputation déterministe, référence téléphone). Il est à noter que tous les modèles reposant sur la variable imputée en prenant comme référence le mode internet sont moins bons que ceux modèles utilisant la référence téléphone.

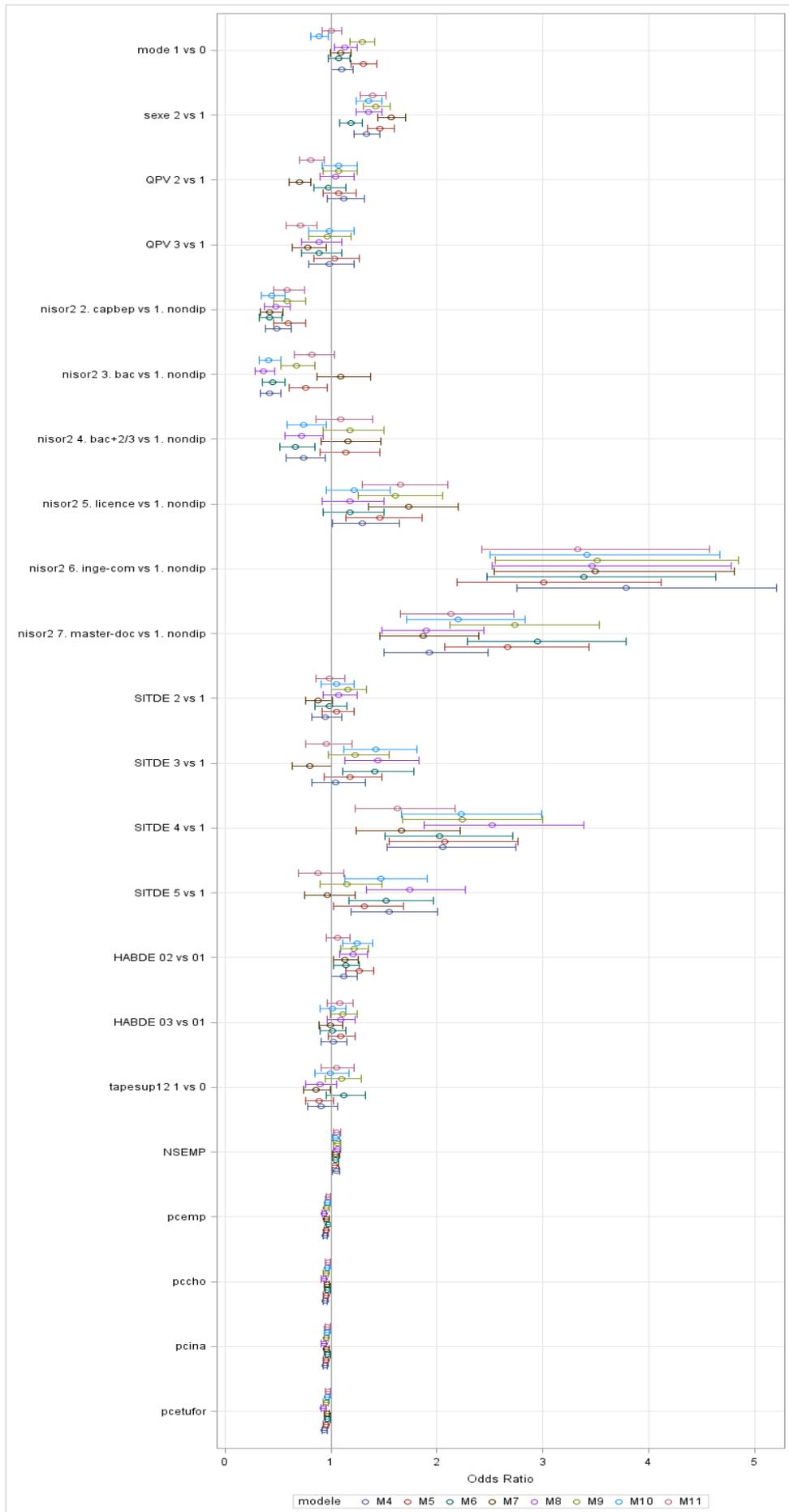
L'objectif de l'agrégation sur données imputées est de disposer de variables corrigées de l'effet de mesure. On devrait donc s'attendre à ce que l'indicatrice sur le mode de collecte (volontairement laissée dans les modèles) ne soit pas significative dans les modèles utilisant ces variables corrigées. La significativité des variables de chaque modèle est présente en annexe 2. Il ressort que sur les 8 modèles testés, l'influence de la variable mode de collecte a été drastiquement réduite. Pour autant, elle demeure significative sur 5 modèles. Seuls les modèles reposant sur un matching avec remise et imputation déterministe (M6 et M7) ainsi que le modèle M11 (matching avec remise, imputation stochastique et référence web) acceptent la non significativité de la variable *mode*. Cela semble confirmer l'hypothèse que l'effet de mesure n'est pas corrigé totalement par le matching sans remise, puisque le mode n'est plus significatif lors des matching avec remise.

Il faut tout de même noter que dans le modèle M7, les variables *QPV* et *tapesup12* deviennent significatives alors qu'elles sont jugées non significatives dans la quasi-totalité des modèles. C'est comme si, dans ce modèle, l'effet de mesure résiduel que devrait porter l'indicatrice du mode se reportait plutôt sur ces 2 variables. Dans le modèle M11, il se passe le même phénomène mais seule la variable *QPV* devient significative. Il n'y a que le modèle M6 qui rejette simultanément la significativité de la variable *mode* et des deux variables *QPV* et *tapesup12*. Toutes les autres variables demeurent significatives quelle que soit l'approche retenue.

La comparaison des oddsratios (ou, ce qui revient au même, des paramètres) de chaque modèle (voir graphe ci-après ou tableau en annexe 3) nous apporte plusieurs enseignements :

- Sur certaines modalités, il n'y a quasiment aucun impact du choix de la méthode et même du choix de la référence. Par exemple, il n'y a pas de différences significatives entre les personnes en emploi et ceux au chômage à la date d'enquête (Sitde 2 vs 1).
- Toujours, en restant sur cet exemple, en revanche, on constate un écart sensible avec les 3 premiers modèles M1, M2 et M3, basés sur l'agrégation simple (sans choix de référence) qui concluait que les personnes en recherche d'emploi avaient significativement plus de « chances » de déclarer  $int=1$  (OR compris entre 1,23 et 1,25).
- Sur certaines modalités, la méthode d'imputation n'a pas d'influence par contre le choix de la référence est déterminant. Ainsi, les sortants au niveau bac+2/3 déclarent significativement moins souvent que les non-diplômés que l'intérêt du poste est très important lorsque l'on choisit le mode de référence téléphone (OR compris entre 0,67 et 0,75). Avec le mode internet, il n'y aurait pas de différences significatives entre ces deux sous-populations (OR compris entre 1,09 et 1,16 avec des intervalles de confiance contenant la valeur 1).
- Globalement la méthode d'imputation (déterministe ou stochastique) ne semble pas trop influencer sur l'estimation des paramètres du modèle. En revanche, le choix du matching (avec ou sans remise) semble souvent influencer. Ainsi en prenant l'exemple des sortants au niveau bac contre les non diplômés, on constate que seuls deux modèles (M7 et M11, qui mobilisent la méthode avec remise et la référence web) conduisent à conclure à une non significativité de l'OR. Dans tous les autres modèles, les bacheliers ont une moindre propension que les non-diplômés à déclarer  $int=1$ , le phénomène étant encore plus marqué avec la référence téléphone (OR compris entre 0,37 et 0,45 dans ce cas). Le modèle M3, basé sur l'agrégation simple avec une interaction entre le mode et le niveau de sortie permettait de retrouver en partie ces résultats.

Ce travail fait sur cette variable a été reproduit sur une autre variable dépendante présentant un effet de mesure Q25\_7 (modalité « pour entrer dans la vie active » de la question à choix multiples sur le motif d'arrêt des études). Les conclusions sont relativement analogues à celles exposés ci-dessus.



### 3.2. Variable explicative avec effet de mesure

Dans cette partie, on va tenter de modéliser une variable qui ne présente pas de problème de mesure. On retient ici la variable dichotomique « être en emploi à la date de l'enquête » (*emp*).

Tableau : statistique descriptive sur la variable *emp* selon le mode de collecte

en emploi à 3 ans	mode		Ensemble
	téléphone	internet	
0	33,3	25,7	30,2
1	66,7	74,3	69,8

Le taux d'emploi à la date d'enquête, soit 3 ans après la sortie du système éducatif, est estimé à 70% (dans l'enquête expérimentale multimode). Les différences observées selon le mode (66,7% pour les répondants téléphone et 74,3% pour les répondants web) sont uniquement liées à l'effet de sélection (en particulier ce sont les jeunes les plus diplômés qui répondent le plus souvent sur internet et ce sont également eux qui ont les meilleurs taux d'emploi).

Dans un premier temps, nous allons faire un modèle avec uniquement des variables explicatives ne présentant pas d'effet de mesure. Puis nous allons introduire dans le modèle des variables présentant un effet de mesure en interaction ou non avec la variable mode et en testant les différentes versions de ces variables (variable d'origine pour l'agrégation simple et variables imputées selon les différentes méthodes et référence). Les différents modèles testés sont présentés dans le tableau suivant. L'introduction de la variable *q25\_7* (avoir déclaré comme motif d'arrêt des études vouloir entrer dans la vie active) améliore la qualité des modèles. Cette variable subjective présente un effet de mesure important. La méthode choisie semble en revanche très peu influencer sur la qualité générale du modèle (aire sous la courbe de ROC compris entre 73,6 et 74,5). Plus précisément, seul le choix de la référence a une légère influence sur la qualité du modèle et comme précédemment en défaveur de la référence web.

modèle	agrégation	variables explicatives	Imputation	Matching	Référence	variable mode	Aire sous la courbe de Roc	
_m2	simple	sans effet de mesure	déterministe	sans remise	tel	seul	72,8	
_m3						en interaction	73,0	
_m4		avec effet de mesure			tel	seul	74,1	
_m5						en interaction	74,2	
_m6	sur données imputées	avec effet de mesure		stochastique	sans remise	tel	seul	74,4
_m7							en interaction	74,5
_m8						web	seul	73,8
_m9					en interaction		73,9	
_m10			avec remise		tel		seul	74,4
_m11						en interaction	74,5	
_m12				web	seul	73,5		
_m13			en interaction		73,5			
_m14			stochastique		sans remise	tel	seul	74,4
_m15				en interaction			74,4	
_m16				web		seul	73,8	
_m17					en interaction	73,8		
_m18	avec remise	tel			seul	74,4		
_m19				en interaction	74,5			
_m20		web	avec remise	seul	73,6			
_m21	en interaction			73,6				

Tableau : liste des modèles testés et indicateur de qualité globale

Les variables ne présentant pas d'effet de mesure et qui interviennent dans l'ensemble des modèles sont : mode de cohabitation à la date d'enquête *HABDE*, nombre de séquences totales dans le calendrier d'activité *NSTOT*, résider en quartier prioritaire *QPV*, niveau de sortie *NISOR* et *sexe*. Ces variables sont significatives dans tous les modèles avec une relative stabilité de leur « pouvoir explicatif ». Ce résultat est rassurant car cela signifie que l'introduction dans le modèle d'une explicative présentant un effet de mesure ne semble pas venir le perturber, et ce quelle que soit la méthode d'agrégation utilisée.

Dans le modèle *\_m2*, qui ne comporte que des variables sans effet de mesure, la variable *mode* n'a strictement aucun effet sur le modèle. Cela s'explique très certainement par le fait que la plupart des variables expliquant l'effet de sélection sont présentes dans le modèle. La non significativité de la variable *mode* dans un tel modèle vient confirmer l'absence d'effet de mesure.

Dans les modèles avec une agrégation simple, dès lors que l'on introduit la variable avec effet de mesure *q25\_7*, la variable *mode* ressort significative que ce soit seule ou en interaction avec *q25\_7*. La variable *Q25\_7* et toutes ses variantes imputées ressortent systématiquement significatives dans les modèles. En revanche, dans tous les modèles sur données imputées (*\_m6* à *\_m21*), la variable *mode* en effet principal n'est jamais significative. En revanche le constat est moins net lorsque l'on regarde l'interaction entre le mode et *Q25\_7* (ou ses variantes imputées). En effet cette interaction ressort significative dans les modèles *\_m9*, *\_m11* et *\_m19* (mais le lien demeure mesuré). Pour tous les autres modèles, cette interaction n'est pas significative.

À ce stade de l'analyse, il semblerait que l'introduction d'une variable explicative présentant un effet de mesure ne vient pas perturber le modèle :

- à condition, dans un modèle reposant sur l'agrégation simple, de contrôler la variable *mode* (et plutôt en interaction)
- ou bien en remplaçant la variable incriminée par sa variante imputée.

Les paramètres (ou les oddsratio) des variables qui ne présentent pas d'effet de mesure sont relativement robustes d'une méthode à l'autre (voir tableau en annexe 4). L'effet de la variable *Q25\_7* qui présente un effet de mesure est relativement stable également : le fait d'avoir arrêté ses études pour entrer dans la vie active augmente les « chances » d'être en emploi 3 ans après par rapport à un individu qui ne retient pas ce motif (OR compris entre 1,6 et 1,9). Les variables imputées jouent bien leur rôle d'annulation de l'effet de mesure (puisque le mode ne ressort quasiment jamais significative).

Lorsque le *mode* ressort c'est uniquement en interaction (annexe 5) avec la variable *Q25\_7* (ou ses variantes imputées) (modèles *\_m5*, *\_m9*, *\_m11*, *\_m19*). L'interaction joue toujours de la même manière. C'est à dire que parmi ceux qui ne déclarent pas ce motif d'arrêt des études, les répondants sur internet ont plus de chances d'être en emploi. Ce n'est pas un phénomène majeur et l'interprétation générale du modèle n'est pas perturbée.

## Conclusion

Le Céreq, dans le cadre de la rénovation des enquêtes Génération, a décidé de basculer en 2020 vers une collecte multimode internet-téléphone. L'introduction d'internet comme mode de collecte supplémentaire présente plusieurs intérêts : un gage de modernité, un mode particulièrement adapté à la cible de l'enquête, les jeunes, une augmentation potentielle de la couverture en touchant des individus difficiles à joindre sur le mode historique (téléphone), une source d'économie puisqu'une enquête sur internet est moins coûteuse qu'une enquête passée avec un enquêteur par téléphone.

Cependant, le passage à une collecte multimode implique des effets de mesure entre les réponses des individus des différents modes. En effet, les individus ne répondent pas de la même manière aux mêmes questions selon qu'ils soient sur internet ou au téléphone.

Il est essentiel de pouvoir repérer quelles variables sont sujettes à des effets de mesure, afin de pouvoir le corriger s'il existe : pour des variables présentant un effet de mesure, les estimations seront différentes selon la part de répondants par téléphone et internet si aucune correction n'est apportée. Pour évaluer l'effet de mesure, la méthode d'appariement sur score de propension propose d'évaluer la différence de réponse entre les individus de chacun des modes à caractéristiques socio-démographiques homologues. Pour cela, les individus d'un mode sont appariés avec un individu de l'autre mode qui a des caractéristiques similaires. L'effet de mesure est alors estimé à l'aide de l'ATT.

Deux méthodes de matching sont ici comparées : avec et sans remise. La méthode avec remise a l'avantage d'apparier plus d'individus, mais certains individus du mode de référence peuvent avoir un gros poids dans l'estimation de l'ATT. La méthode sans remise apparie moins d'individus et est aléatoire ; elle a cependant l'avantage d'être symétrique et d'assurer qu'aucun individu n'aura de fort poids dans l'estimation de l'effet de mesure. Ces deux méthodes ne donnent pas la même valeur de l'estimation de l'effet de mesure mais sont concordantes sur le sens de celui-ci s'il existe.

Après avoir détecté un effet de mesure, la question de l'agrégation des données se pose, en vue de l'estimation de paramètres d'intérêt sur l'ensemble de la population. L'agrégation simple des données se comprend si on suppose que les effets de mesure entre les différents modes, par rapport à la « vérité », se compensent et que cette vérité se trouve entre les estimations de chacun des modes. Une méthode d'agrégation après imputation des données d'une partie de la population est proposée selon deux approches. Après avoir choisi un mode de référence, la répartition des individus appariés du mode de référence selon les variables d'intérêt sert de référence pour ajuster les individus du mode annexe.

La première approche est déterministe et consiste à imputer la valeur de leur contrefactuel à tous les individus du second mode ; mécaniquement, l'effet de mesure sur les individus appariés est annulé. La seconde approche, stochastique, consiste à imputer une partie des individus appariés dont la réponse est différente de leur contrefactuel afin d'obtenir une répartition similaire entre les deux modes sur les variables d'intérêt.

Les deux méthodes d'imputation donnent les mêmes estimations sur les fréquences des modalités de variables présentant un effet de mesure. Cependant, ces estimations diffèrent selon la référence choisie, et diffèrent de l'agrégation simple. De plus, il semble que le matching sans remise ne permet pas de corriger entièrement l'effet de mesure sur l'ensemble de la population, ce que le matching avec remise semble mieux réussir.

Les données des enquêtes Génération servant souvent à des études économétriques sur les sortants du système éducatif français, il est intéressant de se demander quels sont les effets de ces différences de mesure entre téléphone et internet, ainsi que des différentes méthodes d'agrégation des données, sur la modélisation de différentes variables.

La modélisation sans précaution d'une variable présentant un effet de mesure peut faire apparaître des liens fictifs entre cette variable et d'autres ne présentant pas d'effet de mesure, qui résultent uniquement d'un lien avec le mode de collecte. Pour pallier cet écueil, l'ajout de la variable de mode de collecte permet de capter l'effet de mesure. Ajouter son interaction avec les autres variables du modèle permet de tenir compte d'un éventuel effet de mesure différencié selon les modalités de ces variables, mais ne semble pas changer l'interprétation globale de ces variables.

La même modélisation à l'aide d'une variable imputée fait ressortir, à quelques exceptions près, les mêmes conclusions d'analyse : les oddsratios semblent indiquer le même effet. Une imputation stochastique ou déterministe ne semble pas avoir d'effet sur les paramètres du modèle, *a contrario* d'un changement de mode de référence qui peut influencer le modèle. Un effet de mesure résiduel peut être repéré par la significativité de la variable du mode sur les modèles basés sur le matching sans remise.

Lorsqu'on modélise une variable ne présentant pas d'effet de mesure en faisant intervenir une variable explicative en présentant un, le contrôle de l'effet de mesure par l'introduction de la variable de mode permet d'obtenir une interprétation similaire des effets de cette variable à celle obtenue avec les variables imputées. Celle-ci ne varie pas selon la méthode d'imputation, ni le matching utilisé.

Au regard de ces exemples, il semble important de sensibiliser les utilisateurs des données provenant d'enquêtes multimodes sur les effets de mesure, et de les avertir sur les conséquences éventuelles sur leur modélisation.

Deux approches pour diffuser les données semblent être possibles et équivalentes du point de vue de la modélisation (mais pas de la statistique descriptive) :

- Laisser les variables brutes uniquement en pointant celles présentant des effets de mesure, et sensibiliser les chargés d'études de l'intérêt de contrôler le mode de collecte (si possible en interaction) lorsqu'ils utilisent ces variables.
- Diffuser des variables imputées en plus des variables brutes en justifiant leur intérêt pour produire simplement des statistiques descriptives ou des modélisations. Cette approche nécessite cependant de faire le choix d'une référence modale.

## ANNEXE 1

Tableau des oddsratios selon les 3 modèles utilisant l'agrégation simple

EFFET	Modèle 1			Modèle 2			Modèle 3		
	OR	ORlow	ORup	OR	ORlow	ORup	OR	ORlow	ORup
mode 1 vs 0	.	.	.	<b>1,84</b>	1,68	2,01	<b>1,67</b>	<b>1,47</b>	<b>1,89</b>
sexe 2 vs 1	<b>1,37</b>	1,26	1,50	<b>1,34</b>	1,23	1,47	<b>1,34</b>	<b>1,23</b>	<b>1,47</b>
sexe 2 vs 1 dans mode=1	.	.	.	.	.	.	<b>1,47</b>	1,29	1,68
sexe 2 vs 1 dans mode=0	.	.	.	.	.	.	<b>1,22</b>	1,09	1,38
QPV 2 vs 1	<b>1,18</b>	1,01	1,37	<b>1,05</b>	0,90	1,22	<b>1,03</b>	0,89	1,21
QPV 3 vs 1	<b>0,97</b>	0,78	1,20	<b>0,91</b>	0,73	1,12	<b>0,90</b>	0,73	1,12
nisor2 7. master-doc vs 1. nondip	<b>2,23</b>	1,74	2,87	<b>1,88</b>	1,46	2,42	<b>1,92</b>	<b>1,49</b>	<b>2,49</b>
nisor2 7. master-doc vs 1. nondip dans mode=1	.	.	.	.	.	.	<b>2,22</b>	1,51	3,28
nisor2 7. master-doc vs 1. nondip dans mode=0	.	.	.	.	.	.	<b>1,66</b>	1,19	2,32
nisor2 6. inge-com vs 1. nondip	<b>3,38</b>	2,47	4,63	<b>2,56</b>	1,86	3,52	<b>2,97</b>	<b>2,08</b>	<b>4,25</b>
nisor2 6. inge-com vs 1. nondip dans mode=1	.	.	.	.	.	.	<b>2,78</b>	1,80	4,30
nisor2 6. inge-com vs 1. nondip dans mode=0	.	.	.	.	.	.	<b>3,17</b>	1,81	5,58
nisor2 5. licence vs 1. nondip	<b>1,30</b>	1,02	1,66	<b>1,19</b>	0,93	1,52	<b>1,21</b>	<b>0,94</b>	<b>1,55</b>
nisor2 5. licence vs 1. nondip dans mode=1	.	.	.	.	.	.	<b>1,15</b>	0,78	1,70
nisor2 5. licence vs 1. nondip dans mode=0	.	.	.	.	.	.	<b>1,27</b>	0,92	1,74
nisor2 4. bac+2/3 vs 1. nondip	<b>0,94</b>	0,74	1,20	<b>0,86</b>	0,67	1,10	<b>0,88</b>	<b>0,69</b>	<b>1,13</b>
nisor2 4. bac+2/3 vs 1. nondip dans mode=1	.	.	.	.	.	.	<b>1,03</b>	0,70	1,51
nisor2 4. bac+2/3 vs 1. nondip dans mode=0	.	.	.	.	.	.	<b>0,76</b>	0,55	1,04
nisor2 3. bac vs 1. nondip	<b>0,52</b>	0,41	0,65	<b>0,55</b>	0,43	0,70	<b>0,63</b>	<b>0,49</b>	<b>0,80</b>
nisor2 3. bac vs 1. nondip dans mode=1	.	.	.	.	.	.	<b>0,99</b>	0,67	1,48
nisor2 3. bac vs 1. nondip dans mode=0	.	.	.	.	.	.	<b>0,39</b>	0,29	0,53
nisor2 2. capbep vs 1. nondip	<b>0,52</b>	0,40	0,66	<b>0,55</b>	0,43	0,71	<b>0,58</b>	<b>0,45</b>	<b>0,76</b>
nisor2 2. capbep vs 1. nondip dans mode=1	.	.	.	.	.	.	<b>0,73</b>	0,47	1,12
nisor2 2. capbep vs 1. nondip dans mode=0	.	.	.	.	.	.	<b>0,47</b>	0,34	0,64
SITDE 5 vs 1	<b>1,34</b>	1,04	1,73	<b>1,36</b>	1,06	1,76	<b>1,38</b>	1,07	1,79
SITDE 4 vs 1	<b>2,16</b>	1,61	2,88	<b>2,21</b>	1,65	2,97	<b>2,30</b>	1,71	3,09
SITDE 3 vs 1	<b>1,37</b>	1,08	1,73	<b>1,33</b>	1,05	1,68	<b>1,31</b>	1,03	1,66
SITDE 2 vs 1	<b>1,24</b>	1,07	1,43	<b>1,23</b>	1,06	1,43	<b>1,25</b>	1,08	1,45
HABDE 03 vs 01	<b>1,15</b>	1,02	1,29	<b>1,16</b>	1,03	1,31	<b>1,18</b>	1,05	1,33
HABDE 02 vs 01	<b>1,22</b>	1,10	1,36	<b>1,24</b>	1,11	1,38	<b>1,26</b>	1,13	1,40
tapesup12 1 vs 0	<b>1,23</b>	1,05	1,43	<b>1,13</b>	0,96	1,32	<b>1,11</b>	0,95	1,30
NSEMP	<b>1,07</b>	1,04	1,11	<b>1,06</b>	1,03	1,09	.	.	.
NSEMP dans mode=1	.	.	.	.	.	.	<b>1,02</b>	0,98	1,06
NSEMP dans mode=0	.	.	.	.	.	.	<b>1,12</b>	1,07	1,17
pcomp	<b>0,95</b>	0,92	0,97	<b>0,95</b>	0,93	0,97	<b>0,95</b>	0,93	0,97
pccho	<b>0,94</b>	0,91	0,96	<b>0,94</b>	0,92	0,96	<b>0,94</b>	0,92	0,97
pcina	<b>0,94</b>	0,92	0,96	<b>0,94</b>	0,92	0,97	<b>0,94</b>	0,92	0,97
pcetufor	<b>0,94</b>	0,92	0,96	<b>0,95</b>	0,92	0,97	<b>0,95</b>	0,92	0,97

## ANNEXE 2

### Test de significativité des variables explicatives selon les différents modèles testés

Effect	DDL	M1 67,2		M2 68,5		M3 68,8	
		waldchisq1	probchisq1	waldchisq2	probchisq2	waldchisq3	probchisq3
mode	1	.	.	170,3	<.0001	65,7	<.0001
nisor2	6	628,5	<.0001	400,7	<.0001	327,1	<.0001
sexe	1	49,9	<.0001	42,9	<.0001	41,1	<.0001
SITDE	4	33,6	<.0001	33,8	<.0001	36,3	<.0001
pccho	1	30,4	<.0001	25,1	<.0001	23,4	<.0001
pcina	1	28,8	<.0001	24,0	<.0001	22,5	<.0001
pcetufor	1	24,2	<.0001	20,3	<.0001	19,3	<.0001
pcemp	1	21,8	<.0001	17,8	<.0001	16,5	<.0001
NSEMP	1	20,3	<.0001	13,5	0,0002	18,0	<.0001
HABDE	2	14,0	0,0009	15,7	0,0004	17,7	0,0001
QPV	2	9,4	0,0093	3,3	0,1969	2,7	0,2609
tapesup12	1	6,5	0,0107	2,2	0,1361	1,7	0,1945
mode*niso	6	.	.	.	.	59,1	<.0001
NSEMP*mod	1	.	.	.	.	9,3	0,0023
mode*sexe	1	.	.	.	.	4,2	0,0396

Effect	DDL	M4 67,9		M5 66,8		M6 69,0		M7 63,7		M8 69,2		M9 67,7		M10 67,8		M11 64,6	
		waldchisq4	probchisq4	waldchisq5	probchisq5	waldchisq6	probchisq6	waldchisq7	probchisq7	waldchisq8	probchisq8	waldchisq9	probchisq9	waldchisq10	probchisq10	waldchisq11	probchisq11
mode	1	4,3	0,0375	32,5	<.0001	2,3	0,1309	3,5	0,0623	6,7	0,0094	30,0	<.0001	5,9	0,0150	0,0	0,8425
nisor2	6	648,4	<.0001	436,8	<.0001	839,1	<.0001	442,1	<.0001	674,0	<.0001	531,9	<.0001	701,2	<.0001	379,6	<.0001
sexe	1	40,5	<.0001	74,7	<.0001	13,5	0,0002	105,6	<.0001	43,3	<.0001	62,8	<.0001	44,0	<.0001	58,2	<.0001
SITDE	4	31,5	<.0001	25,8	<.0001	32,7	<.0001	21,8	0,0002	49,0	<.0001	31,5	<.0001	36,0	<.0001	15,3	0,0041
pccho	1	22,5	<.0001	16,3	<.0001	6,1	0,0132	10,7	0,0011	33,2	<.0001	16,4	<.0001	9,0	0,0028	6,1	0,0132
pcina	1	20,8	<.0001	16,9	<.0001	6,1	0,0139	12,7	0,0004	33,2	<.0001	14,4	0,0001	9,2	0,0024	7,2	0,0074
pcetufor	1	24,9	<.0001	14,6	0,0001	6,4	0,0111	10,0	0,0015	33,5	<.0001	14,5	0,0001	7,8	0,0052	4,9	0,0276
pcemp	1	19,7	<.0001	14,2	0,0002	3,8	0,0523	11,2	0,0008	27,5	<.0001	11,8	0,0006	6,1	0,0136	4,6	0,0328
NSEMP	1	9,5	0,0021	5,4	0,0201	6,5	0,0107	8,3	0,0040	13,2	0,0003	13,8	0,0002	7,7	0,0057	12,6	0,0004
HABDE	2	4,6	0,1019	19,0	<.0001	6,5	0,0379	7,5	0,0238	11,3	0,0035	12,9	0,0016	19,2	<.0001	2,2	0,3392
QPV	2	4,4	0,1105	0,9	0,6354	1,4	0,5089	24,2	<.0001	3,8	0,1508	2,4	0,2976	1,6	0,4466	12,1	0,0024
tapesup12	1	1,3	0,2498	2,5	0,1119	2,1	0,1477	4,1	0,0441	1,6	0,2025	1,6	0,2069	0,0	0,9339	0,5	0,4846

## ANNEXE 3

Tableau des oddsratios des modèles sur les versions imputées de *int* (M4 à M11)

Effet	OR4	ORlow4	ORup4	OR5	ORlow5	ORup5	OR6	ORlow6	ORup6	OR7	ORlow7	ORup7	OR8	ORlow8	ORup8	OR9	ORlow9	ORup9	OR10	ORlow10	ORup10	OR11	ORlow11	ORup11
mode 1 vs 0	1,10	1,01	1,21	1,30	1,19	1,43	1,08	0,98	1,18	1,09	1,00	1,19	1,13	1,03	1,24	1,29	1,18	1,42	0,89	0,81	0,98	1,01	0,92	1,11
sexe 2 vs 1	1,34	1,22	1,46	1,47	1,34	1,60	1,19	1,08	1,30	1,57	1,44	1,71	1,35	1,24	1,48	1,43	1,31	1,56	1,36	1,24	1,49	1,39	1,28	1,52
QPV 3 vs 1	0,98	0,79	1,22	1,03	0,84	1,27	0,89	0,72	1,10	0,78	0,64	0,95	0,89	0,72	1,11	0,97	0,79	1,19	0,98	0,79	1,22	0,71	0,58	0,87
QPV 2 vs 1	1,12	0,96	1,31	1,07	0,92	1,24	0,98	0,84	1,14	0,70	0,60	0,81	1,05	0,90	1,22	1,08	0,93	1,25	1,07	0,92	1,25	0,81	0,70	0,94
nisor2 7. master-doc vs 1. nondip	1,93	1,50	2,48	2,67	2,07	3,44	2,95	2,29	3,79	1,87	1,46	2,40	1,90	1,48	2,44	2,74	2,13	3,53	2,20	1,71	2,83	2,13	1,66	2,73
nisor2 6. inge-com vs 1. nondip	3,79	2,75	5,21	3,01	2,19	4,12	3,39	2,47	4,63	3,50	2,54	4,81	3,47	2,52	4,78	3,52	2,55	4,85	3,42	2,50	4,68	3,33	2,42	4,58
nisor2 5. licence vs 1. nondip	1,29	1,01	1,65	1,46	1,14	1,86	1,18	0,93	1,50	1,73	1,36	2,21	1,18	0,92	1,50	1,61	1,26	2,05	1,22	0,95	1,56	1,65	1,30	2,10
nisor2 4. bac+2/3 vs 1. nondip	0,74	0,58	0,95	1,14	0,90	1,46	0,67	0,52	0,85	1,16	0,91	1,47	0,72	0,57	0,93	1,18	0,92	1,50	0,75	0,58	0,96	1,09	0,86	1,39
nisor2 3. bac vs 1. nondip	0,42	0,33	0,53	0,76	0,60	0,97	0,45	0,35	0,57	1,09	0,86	1,38	0,37	0,29	0,47	0,67	0,53	0,85	0,41	0,32	0,52	0,82	0,65	1,03
nisor2 2. capbep vs 1. nondip	0,49	0,38	0,63	0,60	0,46	0,76	0,42	0,33	0,54	0,43	0,33	0,55	0,48	0,37	0,61	0,59	0,46	0,76	0,44	0,34	0,57	0,59	0,46	0,75
SITDE 5 vs 1	1,55	1,19	2,01	1,32	1,03	1,69	1,52	1,17	1,97	0,96	0,75	1,23	1,74	1,34	2,27	1,15	0,89	1,48	1,47	1,13	1,91	0,88	0,69	1,13
SITDE 4 vs 1	2,05	1,53	2,75	2,07	1,55	2,77	2,02	1,51	2,72	1,66	1,24	2,23	2,52	1,88	3,39	2,24	1,68	3,00	2,23	1,66	2,99	1,63	1,23	2,17
SITDE 3 vs 1	1,05	0,82	1,33	1,18	0,94	1,48	1,41	1,11	1,79	0,80	0,64	1,01	1,44	1,13	1,83	1,23	0,97	1,55	1,43	1,12	1,81	0,95	0,76	1,19
SITDE 2 vs 1	0,95	0,82	1,10	1,05	0,91	1,21	0,99	0,85	1,15	0,88	0,76	1,01	1,07	0,92	1,25	1,16	1,00	1,34	1,05	0,90	1,22	0,99	0,86	1,13
HABDE 03 vs 01	1,02	0,91	1,15	1,10	0,98	1,23	1,01	0,90	1,14	0,99	0,89	1,12	1,09	0,97	1,23	1,11	0,99	1,25	1,01	0,90	1,14	1,08	0,97	1,21
HABDE 02 vs 01	1,12	1,00	1,25	1,26	1,14	1,40	1,14	1,02	1,27	1,14	1,02	1,26	1,21	1,08	1,35	1,22	1,09	1,35	1,25	1,12	1,39	1,06	0,96	1,18
tapesup12 1 vs 0	0,91	0,78	1,07	0,88	0,76	1,03	1,13	0,96	1,32	0,86	0,74	1,00	0,90	0,77	1,06	1,10	0,95	1,29	0,99	0,84	1,17	1,05	0,91	1,22
NSEMP	1,05	1,02	1,08	1,04	1,01	1,07	1,04	1,01	1,08	1,05	1,01	1,08	1,06	1,03	1,09	1,06	1,03	1,09	1,05	1,01	1,08	1,06	1,02	1,09
pcina	0,95	0,92	0,97	0,95	0,93	0,97	0,97	0,95	0,99	0,96	0,94	0,98	0,93	0,91	0,95	0,96	0,93	0,98	0,96	0,94	0,99	0,97	0,95	0,99
pctufor	0,94	0,92	0,96	0,95	0,93	0,98	0,97	0,95	0,99	0,96	0,94	0,99	0,93	0,91	0,95	0,95	0,93	0,98	0,97	0,94	0,99	0,97	0,95	1,00
pcemp	0,95	0,93	0,97	0,96	0,93	0,98	0,98	0,95	1,00	0,96	0,94	0,98	0,94	0,92	0,96	0,96	0,94	0,98	0,97	0,95	0,99	0,98	0,95	1,00
pccho	0,94	0,92	0,97	0,95	0,93	0,98	0,97	0,95	0,99	0,96	0,94	0,98	0,93	0,91	0,96	0,95	0,93	0,98	0,96	0,94	0,99	0,97	0,95	0,99

ANNEXE 4 Tableau des odds-ratio du modèle sur *emp* (\_m2 à \_m21) – hors variable avec effet de mesure

Effet	Modèle_m2			Modèle_m3			Modèle_m4			Modèle_m5		
	OR	ORLow	ORup									
mode 1 vs 0	1,01	0,91	1,11	.	.	.	1,15	1,03	1,28	.	.	.
QPV 3 vs 1	1,27	1,03	1,58	1,25	1,01	1,54	1,31	1,05	1,63	1,32	1,06	1,64
QPV 2 vs 1	1,43	1,23	1,66	1,38	1,19	1,59	1,48	1,27	1,72	1,48	1,28	1,72
NSTOT	0,93	0,91	0,94	0,93	0,91	0,94	0,94	0,92	0,95	0,94	0,92	0,96
HABDE 03 vs 01	1,98	1,75	2,24	1,98	1,75	2,24	2,06	1,81	2,34	2,08	1,83	2,36
HABDE 02 vs 01	2,51	2,25	2,82	2,54	2,27	2,85	2,50	2,22	2,80	2,51	2,23	2,82
sexe 2 vs 1	0,63	0,57	0,69	0,64	0,58	0,70	0,65	0,59	0,72	0,65	0,59	0,72
nisor2 7. master-doc vs 1. nondip	4,00	3,08	5,20	.	.	.	3,84	2,91	5,08	3,81	2,88	5,03
nisor2 6. inge-com vs 1. nondip	6,11	4,12	9,07	.	.	.	5,23	3,49	7,85	5,25	3,50	7,88
nisor2 5. licence vs 1. nondip	1,77	1,38	2,25	.	.	.	1,68	1,29	2,19	1,66	1,28	2,16
nisor2 4. bac+2/3 vs 1. nondip	4,27	3,32	5,48	.	.	.	3,84	2,94	5,01	3,82	2,92	4,99
nisor2 3. bac vs 1. nondip	1,74	1,38	2,20	.	.	.	1,58	1,23	2,03	1,57	1,22	2,02
nisor2 2. capbep vs 1. nondip	1,48	1,16	1,89	.	.	.	1,45	1,12	1,88	1,43	1,10	1,86

Effet	Modèle_m6			Modèle_m7			Modèle_m8			Modèle_m9			Modèle_m10			Modèle_m11			Modèle_m12			Modèle_m13		
	OR	ORLow	ORup	OR	ORLow	ORup	OR	ORLow	ORup	OR	ORLow	ORup	OR	ORLow	ORup									
mode 1 vs 0	1,04	0,94	1,16	.	.	.	1,06	0,96	1,18	.	.	.	1,05	0,95	1,17	.	.	.	1,08	0,97	1,19	.	.	.
sexe 2 vs 1	0,65	0,59	0,72	0,65	0,59	0,72	0,66	0,60	0,73	0,66	0,60	0,73	0,66	0,60	0,73	0,66	0,60	0,73	0,72	0,65	0,79	0,72	0,65	0,79
QPV 3 vs 1	1,31	1,05	1,64	1,32	1,06	1,64	1,33	1,06	1,65	1,34	1,07	1,67	1,31	1,06	1,64	1,32	1,06	1,65	1,22	0,97	1,52	1,22	0,98	1,52
QPV 2 vs 1	1,47	1,27	1,72	1,48	1,27	1,72	1,48	1,27	1,72	1,49	1,28	1,73	1,44	1,24	1,67	1,46	1,26	1,70	1,46	1,25	1,70	1,46	1,26	1,71
NSTOT	0,94	0,92	0,95	0,94	0,92	0,95	0,94	0,92	0,96	0,94	0,92	0,96	0,93	0,92	0,95	0,94	0,92	0,95	0,94	0,92	0,96	0,94	0,92	0,96
HABDE 03 vs 01	2,13	1,87	2,42	2,13	1,87	2,42	2,04	1,79	2,31	2,05	1,80	2,33	2,12	1,87	2,41	2,12	1,87	2,41	1,96	1,72	2,23	1,97	1,73	2,25
HABDE 02 vs 01	2,56	2,28	2,87	2,56	2,28	2,87	2,49	2,22	2,80	2,51	2,23	2,82	2,53	2,25	2,84	2,54	2,26	2,85	2,43	2,16	2,74	2,45	2,17	2,75
nisor2 7. master-doc vs 1. nondip	3,81	2,88	5,03	3,79	2,87	5,01	4,06	3,07	5,36	4,05	3,07	5,35	3,79	2,87	5,01	3,77	2,85	4,98	3,89	2,94	5,14	3,90	2,95	5,16
nisor2 6. inge-com vs 1. nondip	5,24	3,49	7,87	5,24	3,49	7,87	5,57	3,71	8,36	5,65	3,76	8,47	4,96	3,31	7,45	5,15	3,42	7,73	5,32	3,54	7,98	5,35	3,57	8,04
nisor2 5. licence vs 1. nondip	1,61	1,24	2,09	1,60	1,23	2,08	1,79	1,38	2,33	1,79	1,37	2,32	1,68	1,29	2,18	1,65	1,27	2,14	1,62	1,25	2,11	1,61	1,24	2,10
nisor2 4. bac+2/3 vs 1. nondip	3,82	2,93	5,00	3,81	2,92	4,98	3,94	3,01	5,14	3,93	3,01	5,14	3,82	2,92	4,99	3,81	2,91	4,98	3,68	2,81	4,81	3,67	2,80	4,79
nisor2 3. bac vs 1. nondip	1,59	1,23	2,05	1,58	1,23	2,04	1,56	1,21	2,01	1,55	1,20	1,99	1,63	1,26	2,09	1,60	1,24	2,05	1,34	1,04	1,73	1,33	1,03	1,72
nisor2 2. capbep vs 1. nondip	1,36	1,05	1,77	1,36	1,05	1,77	1,49	1,15	1,93	1,47	1,14	1,91	1,40	1,08	1,81	1,39	1,07	1,81	1,65	1,27	2,14	1,66	1,27	2,15

Effet	Modèle_m14			Modèle_m15			Modèle_m16			Modèle_m17			Modèle_m18			Modèle_m19			Modèle_m20			Modèle_m21		
	OR	ORLow	ORup																					
mode 1 vs 0	1,03	0,92	1,14	.	.	.	1,06	0,95	1,18	.	.	.	1,08	0,97	1,20	.	.	.	1,02	0,92	1,13	.	.	.
sexe 2 vs 1	0,65	0,59	0,72	0,65	0,59	0,72	0,65	0,59	0,72	0,65	0,59	0,72	0,66	0,60	0,73	0,66	0,60	0,73	0,64	0,58	0,70	0,64	0,58	0,70
QPV 3 vs 1	1,32	1,06	1,64	1,32	1,06	1,64	1,30	1,05	1,62	1,30	1,05	1,62	1,28	1,03	1,59	1,30	1,04	1,62	1,29	1,04	1,61	1,29	1,04	1,61
QPV 2 vs 1	1,49	1,28	1,73	1,49	1,28	1,73	1,48	1,28	1,72	1,49	1,28	1,73	1,46	1,26	1,70	1,47	1,27	1,71	1,48	1,27	1,72	1,48	1,27	1,72
NSTOT	0,94	0,92	0,95	0,94	0,92	0,95	0,94	0,92	0,95	0,94	0,92	0,95	0,94	0,92	0,95	0,94	0,92	0,95	0,94	0,92	0,95	0,94	0,92	0,95
HABDE 03 vs 01	2,09	1,84	2,38	2,09	1,84	2,38	2,08	1,83	2,36	2,08	1,83	2,36	2,08	1,83	2,37	2,10	1,85	2,38	2,06	1,81	2,33	2,05	1,81	2,33
HABDE 02 vs 01	2,53	2,25	2,84	2,53	2,25	2,85	2,53	2,25	2,84	2,53	2,26	2,85	2,50	2,23	2,81	2,52	2,24	2,83	2,49	2,22	2,80	2,49	2,22	2,80
nisor2 7. master-doc vs 1. nondip	3,85	2,91	5,09	3,84	2,90	5,07	3,97	3,01	5,24	3,97	3,01	5,23	3,70	2,80	4,89	3,71	2,81	4,90	4,08	3,10	5,39	4,08	3,10	5,39
nisor2 6. inge-com vs 1. nondip	5,43	3,62	8,16	5,42	3,61	8,13	5,46	3,64	8,18	5,47	3,65	8,20	5,05	3,36	7,58	5,16	3,43	7,74	5,54	3,70	8,31	5,54	3,69	8,30
nisor2 5. licence vs 1. nondip	1,63	1,25	2,12	1,63	1,25	2,11	1,75	1,35	2,27	1,74	1,34	2,26	1,67	1,29	2,17	1,65	1,27	2,14	1,75	1,35	2,27	1,75	1,35	2,27
nisor2 4. bac+2/3 vs 1. nondip	3,81	2,92	4,98	3,81	2,91	4,98	3,94	3,02	5,14	3,93	3,01	5,13	3,79	2,90	4,94	3,78	2,90	4,94	3,93	3,01	5,13	3,93	3,01	5,13
nisor2 3. bac vs 1. nondip	1,56	1,21	2,01	1,56	1,21	2,01	1,61	1,25	2,07	1,61	1,25	2,07	1,60	1,24	2,05	1,58	1,23	2,03	1,64	1,27	2,10	1,64	1,27	2,10
nisor2 2. capbep vs 1. nondip	1,39	1,07	1,81	1,39	1,07	1,81	1,45	1,12	1,88	1,45	1,12	1,88	1,44	1,11	1,87	1,42	1,09	1,84	1,50	1,16	1,95	1,50	1,16	1,95

## ANNEXE 5

Tableau des odds-ratio du modèle sur *emp* (\_m2 à \_m21) –variable avec effet de mesure, seul ou en interaction

modèle	Effet	ORlow	ORup	OR
_m4	Q25_7 1 vs 0	1,66	2,02	1,83
_m6	Q25_7* 1 vs 0	1,71	2,08	1,89
_m8	Q25_7* 1 vs 0	1,70	2,06	1,87
_m10	Q25_7* 1 vs 0	1,57	1,91	1,73
_m12	Q25_7* 1 vs 0	1,66	2,02	1,83
_m14	Q25_7* 1 vs 0	1,75	2,12	1,93
_m16	Q25_7* 1 vs 0	1,51	1,83	1,66
_m18	Q25_7* 1 vs 0	1,60	1,95	1,77
_m20	Q25_7* 1 vs 0	1,45	1,75	1,60

modèle	Effet	ORlow	ORup	OR
_m5	mode 1 vs 0 dans Q25_7=1	0,89	1,19	1,03
_m5	mode 1 vs 0 dans Q25_7=0	1,10	1,47	1,27
_m5	Q25_7 1 vs 0 dans mode=1	1,37	1,88	1,61
_m5	Q25_7 1 vs 0 dans mode=0	1,75	2,24	1,98
_m7	mode 1 vs 0 dans Q25_7*=1	0,87	1,13	0,99
_m7	mode 1 vs 0 dans Q25_7*=0	0,97	1,32	1,13
_m7	Q25_7* 1 vs 0 dans mode=1	1,49	2,03	1,74
_m7	Q25_7* 1 vs 0 dans mode=0	1,75	2,25	1,99
_m9	mode 1 vs 0 dans Q25_7*=1	0,80	1,09	0,93
_m9	mode 1 vs 0 dans Q25_7*=0	1,03	1,35	1,18
_m9	Q25_7* 1 vs 0 dans mode=1	1,38	1,89	1,61
_m9	Q25_7* 1 vs 0 dans mode=0	1,80	2,30	2,04
_m11	mode 1 vs 0 dans Q25_7*=1	0,79	1,04	0,91
_m11	mode 1 vs 0 dans Q25_7*=0	1,11	1,51	1,29
_m11	Q25_7* 1 vs 0 dans mode=1	1,19	1,63	1,39
_m11	Q25_7* 1 vs 0 dans mode=0	1,76	2,25	1,99
_m13	mode 1 vs 0 dans Q25_7*=1	0,83	1,13	0,97
_m13	mode 1 vs 0 dans Q25_7*=0	1,01	1,33	1,16
_m13	Q25_7* 1 vs 0 dans mode=1	1,40	1,92	1,64
_m13	Q25_7* 1 vs 0 dans mode=0	1,73	2,23	1,96
_m15	mode 1 vs 0 dans Q25_7*=1	0,87	1,14	1,00
_m15	mode 1 vs 0 dans Q25_7*=0	0,92	1,26	1,07
_m15	Q25_7* 1 vs 0 dans mode=1	1,57	2,15	1,84
_m15	Q25_7* 1 vs 0 dans mode=0	1,75	2,25	1,98
_m17	mode 1 vs 0 dans Q25_7*=1	0,88	1,19	1,03
_m17	mode 1 vs 0 dans Q25_7*=0	0,95	1,25	1,09
_m17	Q25_7* 1 vs 0 dans mode=1	1,37	1,88	1,61
_m17	Q25_7* 1 vs 0 dans mode=0	1,51	1,92	1,70
_m19	mode 1 vs 0 dans Q25_7*=1	0,82	1,08	0,94
_m19	mode 1 vs 0 dans Q25_7*=0	1,09	1,48	1,27
_m19	Q25_7* 1 vs 0 dans mode=1	1,26	1,72	1,47
_m19	Q25_7* 1 vs 0 dans mode=0	1,75	2,25	1,99
_m21	mode 1 vs 0 dans Q25_7*=1	0,88	1,20	1,03
_m21	mode 1 vs 0 dans Q25_7*=0	0,89	1,16	1,02
_m21	Q25_7* 1 vs 0 dans mode=1	1,38	1,88	1,61
_m21	Q25_7* 1 vs 0 dans mode=0	1,41	1,79	1,59

Q25\_7\* signifie qu'il s'agit d'une version imputée de la variable Q25\_7. Elle varie en fonction du modèle choisi

## Bibliographie

- [1] Barret C., Dzikowski C., « La collecte par Internet est-elle l'avenir des enquêtes Génération du Céreq ? », *12<sup>èmes</sup> Journées de Méthodologie Statistique*, 2015.
- [2] Barret C., Dzikowski C., « Évaluation d'un protocole multimode avec échantillon embarqué et agrégation des données en présence d'effet de mesure », *9<sup>e</sup> colloque francophone sur les sondages*, 2016
- [3] Legleye, S., « Effets de sélection, imputations et effets de mode : les dernières tendances en matière de multimode », *Séminaire de Méthodologie Statistique*, 2017.
- [4] Legleye S., de Peretti G., Razafindranovona T., «Agréger les échantillons d'une enquête multimode en limitant l'effet de mesure : une proposition d'imputation raisonnable et pragmatique», *13<sup>èmes</sup> Journées de Méthodologie Statistique*, 2018.
- [5] Barret C., Cissé M, Gaubert E., Mazari Z., Olaria M., «Efficacité d'un protocole multimode (téléphone et internet) », *13<sup>èmes</sup> Journées de Méthodologie Statistique*, 2018.
- [6] Rosenbaum, P. R., Rubin, D. B. «The central role of the propensity score in observational studies for causal effects». *Biometrika*, 1983, 70(1), 41-55, 1983.
- [7] Christian, L. M., Dillman, D. A., & Smyth, J. D. «The effects of mode and format on answers to scalar questions in telephone and web surveys». *Advances in telephone survey methodology*, 2008, 12, 250-275.
- [8] Heerwegh, D., Loosveldt, G. «Face-to-face versus web surveying in a high-internet-coverage population: Differences in response quality». *Public Opinion Quarterly*, 2008, 72(5), 836-846.
- [9] Barret, C., Cissé, M., Dzikowski, C. «Plan de sondage des enquêtes Génération : utilisation d'un calage pour suréchantillonner les extensions». *13<sup>èmes</sup> Journées de Méthodologie Statistique*, 2018.