



ined
INSTITUT
NATIONAL
D'ÉTUDES
DÉMOGRA
PHIQUES



Corriger ou pas les effets de mesure ? Présentation des travaux sur l'enquête ERFI 2

24 novembre 2025

Journée d'étude de l'observatoire des parcours de vie

Guillaume Carette

Institut national d'études démographiques (Ined)

Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du Programme d'Investissements d'Avenir portant la référence **ANR-21-ESRE-0037**.



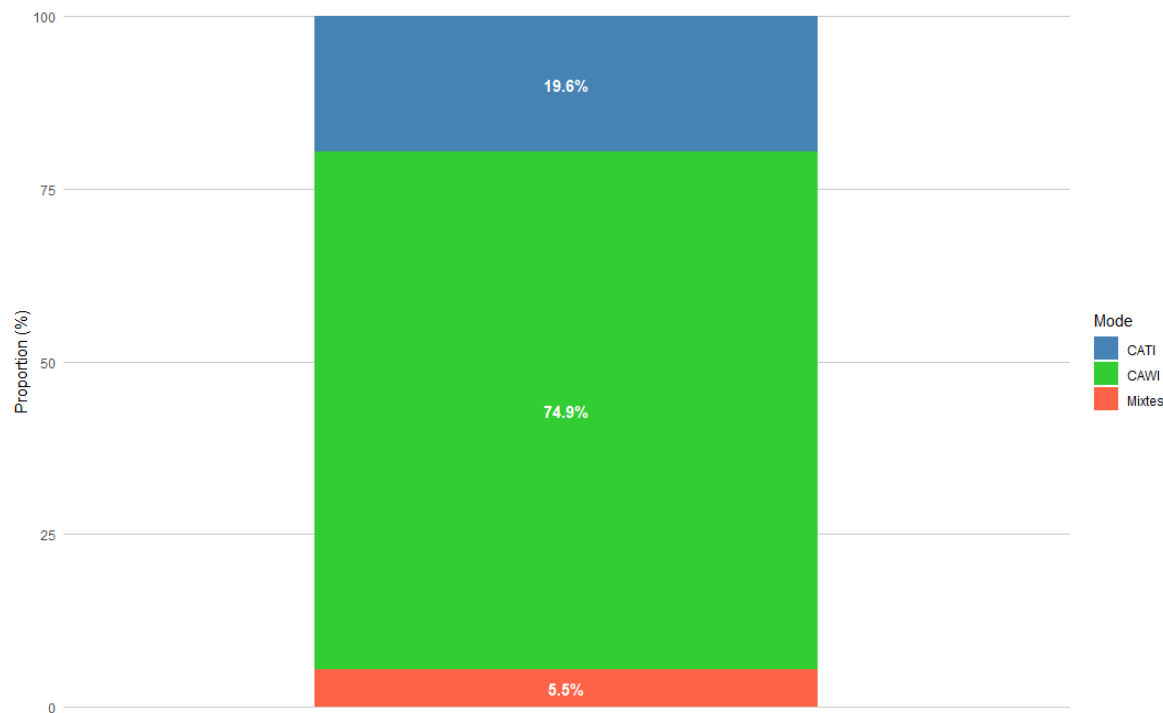
- **Présentation de l'enquête ERFI 2**
- **Etats des lieux des effets de mesure dans l'enquête**
- **Agrégation des données :**
 - Correction par imputation
 - Correction par calage
- **Conclusions et discussion**

L'enquête ERFI 2

Présentation de l'enquête

- **Etude des relations familiales et intergénérationnelles**
- **Enquête française** du programme international **Generations and Gender Programme (GGP)**
 - Fournir des données comparables sur les parcours de vie des individus et les dynamiques familiales et suivre leurs évolutions.
- **3 vagues**
 - 2024, 2027, 2030.
- **Champ**
 - **Hommes et Femmes** habitant en **France métropolitaine**, âgés de **18 à 79 ans**.
- **Protocole**
 - Une **collecte multimode concurrentielle** sur internet (CAWI) et par téléphone (CATI).

Répartition des répondants



	Effectifs	% total
Internet	9134	74,9%
Téléphone	2396	19,6%
Mixte	665	5,5%
TOTAL	12195	100%

Etats des lieux des effets de mesure dans l'enquête ERFI 2

- Une analyse par appariement sur score de propension sur l'ensemble des variables qualitatives du questionnaire (hors QCM et boucles sur les membres du foyer).
- 78 variables soumises à un effet de mode, soit 31% des variables qualitatives analysées :

	Base d'analyse		Variables avec effets de mode		
	Effectifs	%	Effectifs	% total variables avec effet	% catégorie
TOTAL VARIABLES QUALITATIVES	253	100%	78	100%	
Dont variables à échelle ordinale	96	37,94%	66	84,62%	68,75%
Dont variables catégorielles	95	37,55%	5	6,41%	5,26%
Dont variables binaires	62	24,51%	7	8,97%	11,29%

Variables présentant des effets significatifs d'au moins 4 points de pourcentage, sur au moins l'une des modalités de réponse.

- **Des variables à échelle ordinale fortement impactées par le biais de mesure, avec 3 types d'effet rencontrés :**
 - **Effet 1** : Positionnement plus important sur les modalités intermédiaires par Internet et sur les modalités extrêmes par téléphone (le plus souvent positives).
 - **Effet 2** : Recours plus important à la modalité médiane et neutre par Internet dans les questions avec un nombre impair de modalités de réponse.
 - **Effet 3** : Réponses plus négatives par Internet avec parfois un positionnement plus important sur les modalités extrêmes négatives.

- **Variables relatives à la santé mentale des répondants :**
 - Wel11a : Vous avez eu l'impression de ne pas parvenir à vous débarrasser de votre cafard, malgré l'aide de votre famille ou de vos amis.
 - Wel11b : Vous vous êtes senti(e) déprimé(e).
 - Wel11c : Vous avez pensé que votre vie était un échec.
 - Wel11d : Vous vous êtes senti(e) angoissé(e).
 - Wel11e : Vous vous êtes senti(e) triste.
- Un biais de **désirabilité sociale** chez les répondants CATI ?

Variables d'intérêt pour la correction (2/2)

- Estimation causale de l'effet du traitement sur les variables d'intérêt constitutives du score de santé mentale :

	wel11a	wel11b	wel11c	wel11d	wel11e
1. Jamais	0,150 ***	0,228 ***	0,157 ***	0,197 ***	0,276 ***
2. Occasionnellement	-0,114 ***	-0,209 ***	-0,142 ***	-0,159 ***	-0,235 ***
3. Assez souvent	-0,058 ***	-0,048 ***	-0,036 ***	-0,063 ***	-0,065 ***
4. Fréquemment ou tout le temps	0,022 ***	0,029 ***	0,022 ***	0,025 ***	0,023 ***

Lecture : Après appariement sur score de propension, les répondants téléphoniques recourent davantage à la modalité « Jamais » pour décrire les différents ressentis en termes de troubles mentaux, de manière significative. Cela se traduit par exemple par un effet de 15 points sur cette modalité entre les modes Internet et téléphone pour la variable wel11a.

*** : $pval < 0,001$

Agrégation des données : Correction par imputation

- **Pourquoi avoir recours à l'imputation ?**
 - L'effet de mesure est un biais qui repose sur un problème de réponses.
 - Hypothèse : il existe une « vraie » valeur pour la variable d'intérêt Y qui peut être observée à partir des réponses données sous un mode de collecte de référence.

- **Quel mode de collecte de référence choisir ?**
 - Celui dont les réponses sont jugées les plus fiables :
 - Mode CATI si suspicion de satisficing.
 - Mode CAWI si suspicion de désirabilité sociale.

Principes généraux (2/2)

- **Une approche parcimonieuse...**
 - Eviter l'imputation massive.
 - Corriger uniquement les observations les plus exposées à un effet de mesure.
- **... Basée sur l'estimation de valeurs contrefactuelles à partir d'un modèle causal.**
 - Appariement de répondants de modes de collecte différents à partir d'un score de propension.
 - Mêmes hypothèses que l'estimation de l'effet de mesure : Le traitement et les co-variables observées sont conditionnellement indépendants étant donné le score de propension (indépendance conditionnelle).

Un procédé en 4 étapes

1) Mécanisme d'allocation des modes de collecte et définition du support commun

- Estimation d'un score de propension à participer à l'un ou l'autre des modes de collecte.
- Définition du support commun du score estimé entre les deux modes.

2) Définition du support d'imputation

- Sous-ensemble d'observations du support commun qui feront l'objet de l'imputation (à partir d'un appariement).

3) Identification des observations affectées par le biais de mesure

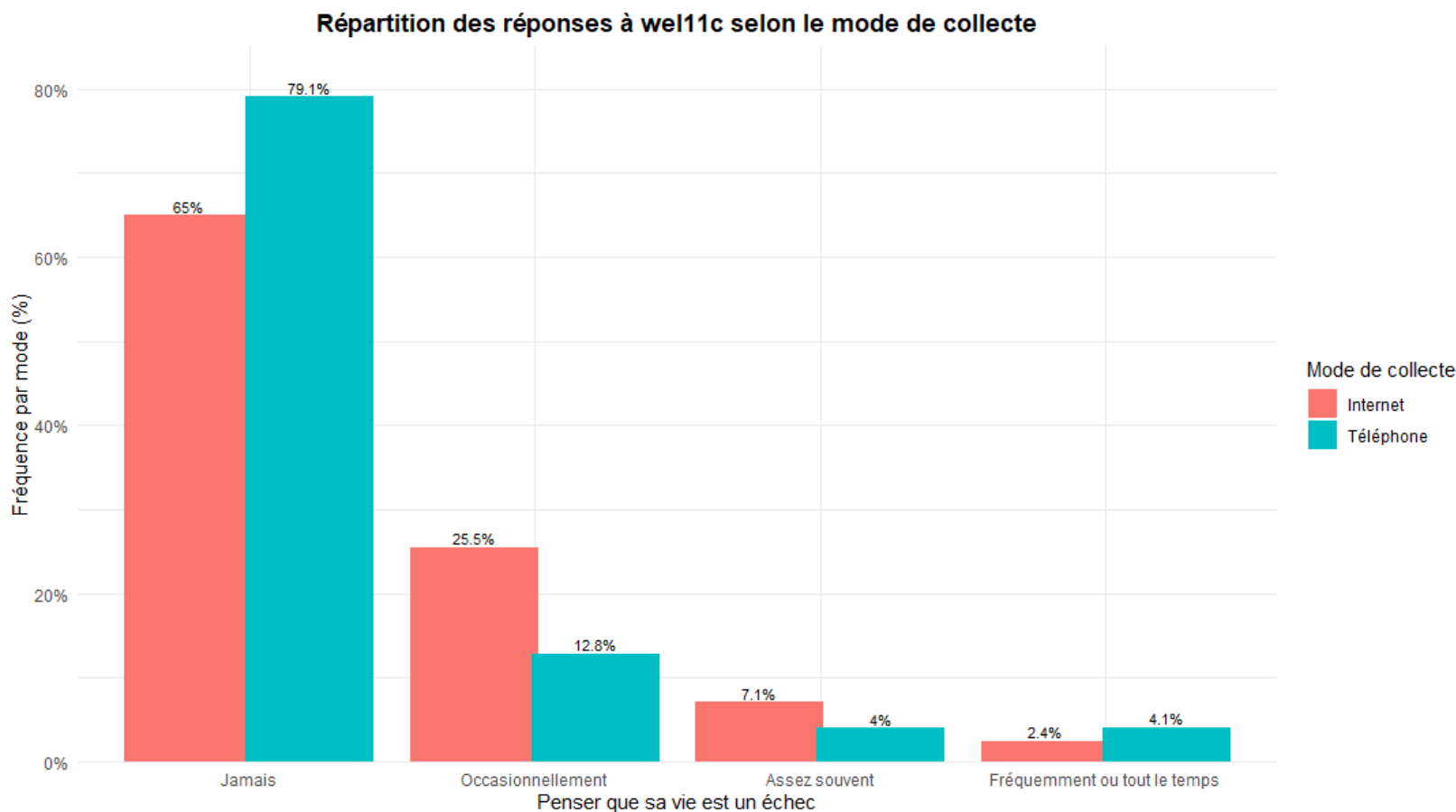
- Repérer les individus qui portent l'effet de mesure à partir du contrefactuel.

4) Imputation des valeurs

- Corriger les réponses du mode alternatif pour les rendre plus cohérentes avec celles observées dans le mode de référence.

Application pour une variable d'intérêt

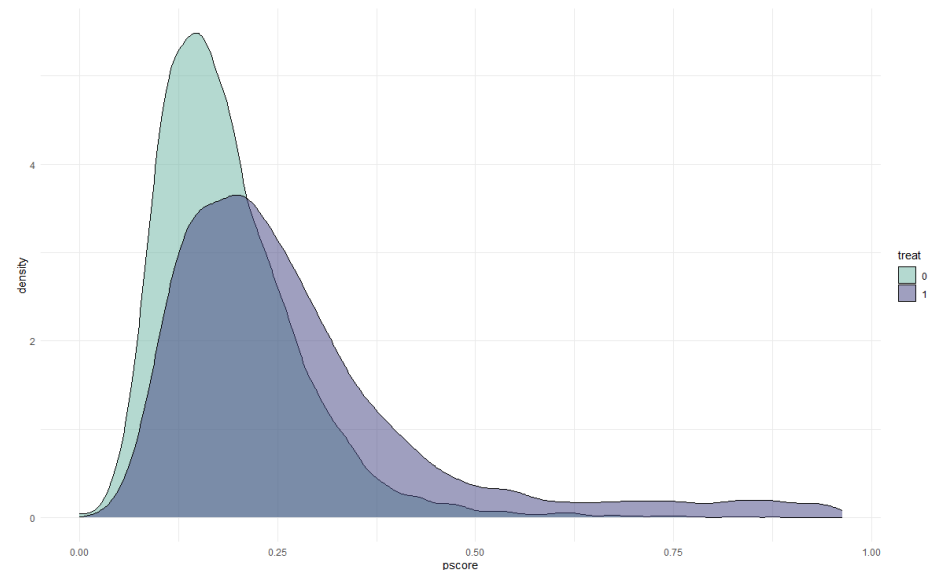
- Au cours de la semaine dernière, vous avez pensé que votre vie était un échec (wel11c).



Estimation d'un score de propension – Support commun

- **Estimer un score de propension à participer par téléphone :**
 - Des variables de la base de sondage : caractéristiques sociodémographiques, caractéristiques du ménage, type de revenus, coordonnées disponibles.
 - Des variables de questionnaire : lieu de naissance, niveau de diplôme, accès et usage d'Internet.

	CAWI	CATI
Moyenne	0,19	0,28
Min	0,001	0,04
Max	0,93	0,96



- **Support commun** : 11480 observations CAWI/CATI dans l'intervalle [0,04 ; 0,93] (50 unités exclues).

- **Appariement sur score de propension :**
 - Caliper 0,1 du score de propension.
 - Appariement 1:1 par paire sans contrainte d'égalité.
 - Vérification des propriétés équilibrantes.

	Unités de contrôle (CAWI)		Unités traitées (CATI)	
	Effectifs	% total	Effectifs	% total
Unités appariées	2283	25%	2283	96%
Unités exclues	6815	75%	99	4%
TOTAL	9098	100%	2382	100%

Les observations sur modes mixtes ne sont pas prises en compte.

- **Définir les candidats à l'imputation :**
 - Valeurs contrefactuelles estimées à partir d'un modèle multinomial sur les observations appariées du mode de collecte de référence (CAWI).
 - Cibler les modalités sur-représentées dans le mode de collecte CATI.
 - Classement des observations les moins bien prédites à partir du modèle dans le mode de référence.
- **Imputation des valeurs :**
 - Imputation séquentielle des observations les moins bien prédites vers la modalité la plus probable selon le modèle, jusqu'à équilibre.

Effet de mesure avant et après – Imputation déterministe

- Quel impact sur l'estimation de l'effet de mesure sur la variable wel11c ?

	Avant imputation			Après imputation		
	Effet téléphone	Std. Err	P-value	Effet téléphone	Std. Err	P-value
1. Jamais	0,16	0,008	< 0,001	0,02	0,008	0,084
2. Occasionnellement	-0,13	0,007	< 0,001	-0,01	0,006	0,248
3. Assez souvent	-0,05	0,004	< 0,001	-0,02	0,007	< 0,001
4. Fréquemment ou tout le temps	0,02	0,003	< 0,001	0,01	0,003	< 0,001

Lecture : Effet mesuré par régression pondérée par l'inverse du score de propension sur l'ensemble des observations du support commun ($n = 11\,480$)

Estimation finale après imputation déterministe

- **Estimations globales avant et après imputations déterministes sur wel11c**
 - **387 valeurs imputées**, soit 17% des unités CATI du support commun.

	Internet	Téléphone	Global
1. Jamais	65,20%	65,60%	65,80%
2. Occasionnellement	25,20%	24,60%	24,60%
3. Assez souvent	7,10%	6,10%	6,80%
4. Fréquemment ou tout le temps	2,50%	3,70%	2,80%

Lecture : Estimation sur les données agrégées à partir de l'ensemble des répondants à la question wel11c (n = 12195).

- **Objectif** : prendre en compte l'incertitude liée à l'imputation en générant m jeux de données qui reflètent les valeurs plausibles à imputer.
- Agrégation des données selon les règles de Rubin (Rubin, 1987) :

Estimation <i>poolée</i> (Q)	$\bar{Q} = \frac{1}{m} \sum_{k=1}^m \hat{Q}_k$
Variance intra-imputation (U)	$\bar{U} = \frac{1}{m} \sum_{k=1}^m \hat{U}_k$
Variance inter-imputation (B)	$B = \frac{1}{m-1} \sum_{k=1}^m (Q_k - \bar{Q})^2$
Augmentation relative de la variance (r)	$r = \frac{(1 + 1/m) B}{\bar{U}}$

Effet de mesure avant et après – Imputation multiple

- Quel impact sur l'estimation de l'effet de mesure sur la variable wel11c ?

	Avant imputation			Après imputation		
	Effet téléphone	Std. Err	P-value	Effet téléphone	Std. Err	P-value
1. Jamais	0,16	0,008	< 0,001	0,019	0,008	0,062
2. Occasionnellement	-0,13	0,007	< 0,001	-0,023	0,006	0,002
3. Assez souvent	-0,05	0,004	< 0,001	-0,010	0,007	0,003
4. Fréquemment ou tout le temps	0,02	0,003	< 0,001	0,014	0,003	< 0,001

Lecture : Effet mesuré par régression pondérée par l'inverse du score de propension sur l'ensemble des observations du support commun ($n = 11\,480$).

Estimation finale après imputation multiple

- **Estimations globales avant et après imputations multiples sur wel11c**
 - $m = 10$ jeux de données.
 - **395 valeurs imputées**, soit 17% des unités CATI du support commun.

	Avant imputation		Après imputation				
	Estimation globale	Estimation CAWI	Estimation poolée	Std. Err	U	B	r
1. Jamais	68,60%	65,20%	65,37%	0,008	< 0,05	< 0,05	0,001
2. Occasionnellement	22,30%	25,20%	24,63%	0,007	< 0,05	< 0,05	0,017
3. Assez souvent	6,30%	7,10%	7,19%	0,004	< 0,05	< 0,05	0,051
4. Fréquemment ou tout le temps	2,90%	2,50%	2,81%	0,003	< 0,05	< 0,05	0,066

Lecture : Estimation sur les données agrégées à partir de l'ensemble des répondants à la question wel11c ($n = 12195$).

Agrégation des données : Correction par calage

- **Le calage : une alternative à l'imputation**
 - Ne modifie pas les réponses des participants à l'enquête ;
 - Marge additionnelle correspondant à l'estimation pondérée d'une variable d'intérêt dans le mode de collecte de référence seul.
- **Objectif : une variable de pondération pour les 5 variables de santé mentale**
 - Marge additionnelle : estimation pondérée d'un score de santé mentale (0 à 15) sur le mode de collecte CAWI.

Score discrétisé	Estimation de référence
0 à 2	39,8%
3 à 5	37,7%
6 à 8	13,7%
9 à 15	8,8%

Lecture : Estimation pondérée du score de santé mentale sur le mode de référence CAWI seul.

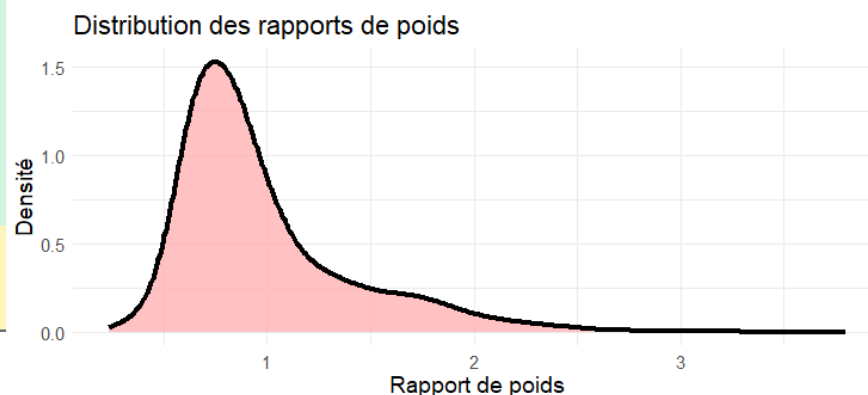
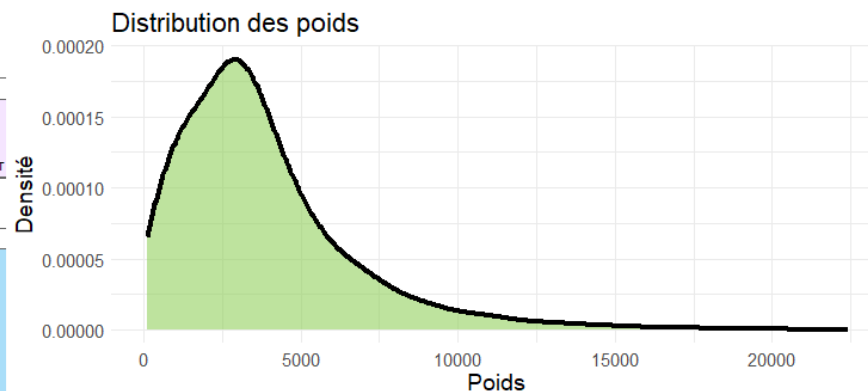
Calage sur score de santé mentale

- Caractéristiques des poids après calage avec marge additionnelle de score de santé mentale

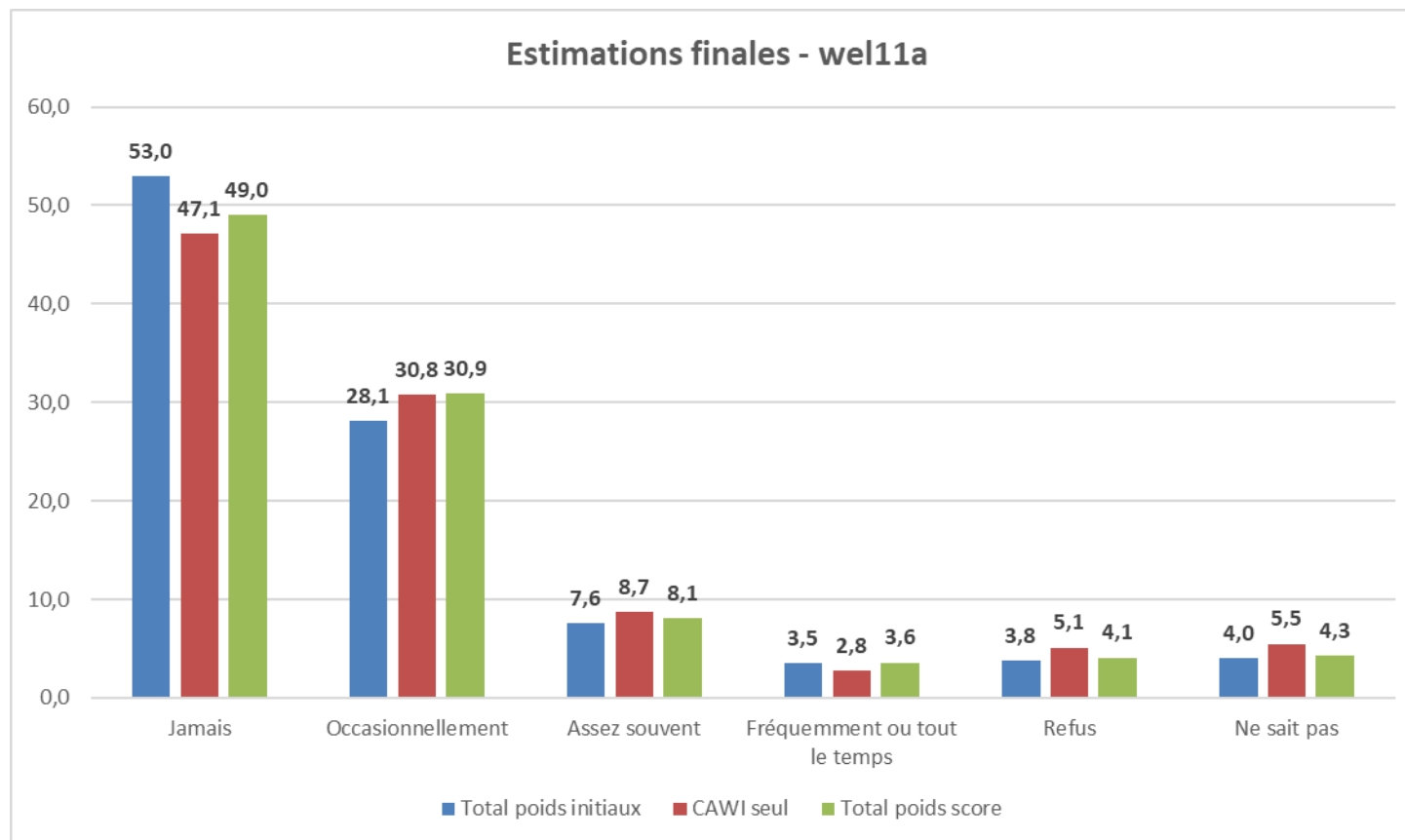
Calage sur échantillon complet

Critère	Détail
Méthode	raking
Bornes des rapports de poids	[0.24 ; 3.8]
Marges utilisées	.GE_D ; CMS ; DIPL4 ; ENFANT ; INAI ; MATR4 ; PCS4 ; TACT5 ; TUU6 ; ZEAT

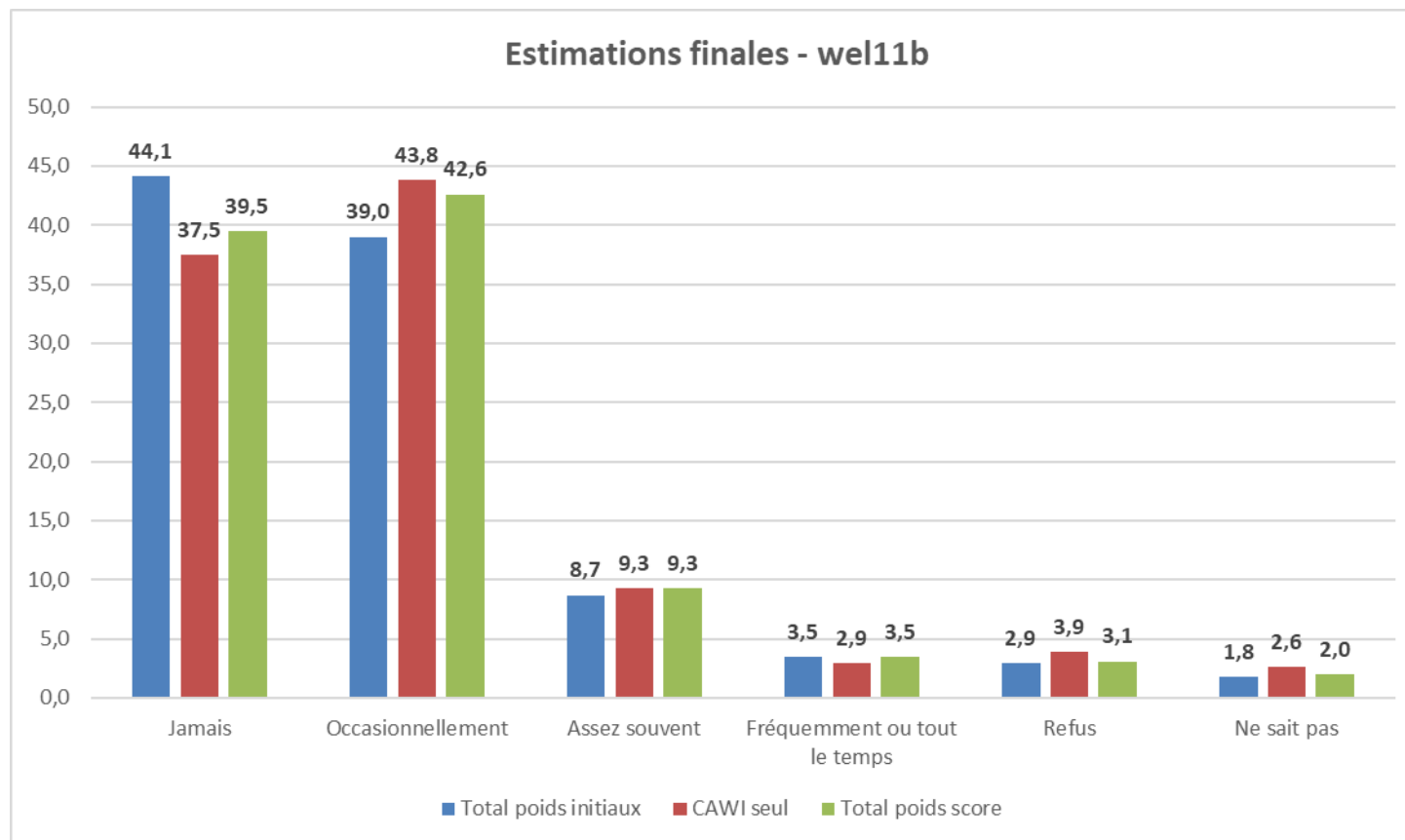
Indicateur	Statistiques
Somme	48 562 126,2
Ecart-type	2 827,6
Variance	7 995 199,1
Ecart interquartile	3 001,7
Moyenne	3 789,2
Coefficient de variation	74,6
Mode	3 162,3
Minimum	99,3
P1	214,5
P5	443,7
P10	895,0
Q1	1 878,4
Médiane	3 183,9
Q3	4 880,1
P90	7 333,2
P95	9 309,0
P99	13 993,7
Maximum	22 411,8
Q3/Q1	2,6
Q3/Min	49,1
P95/Min	93,7
P99/Min	140,9
Max/Min	225,6



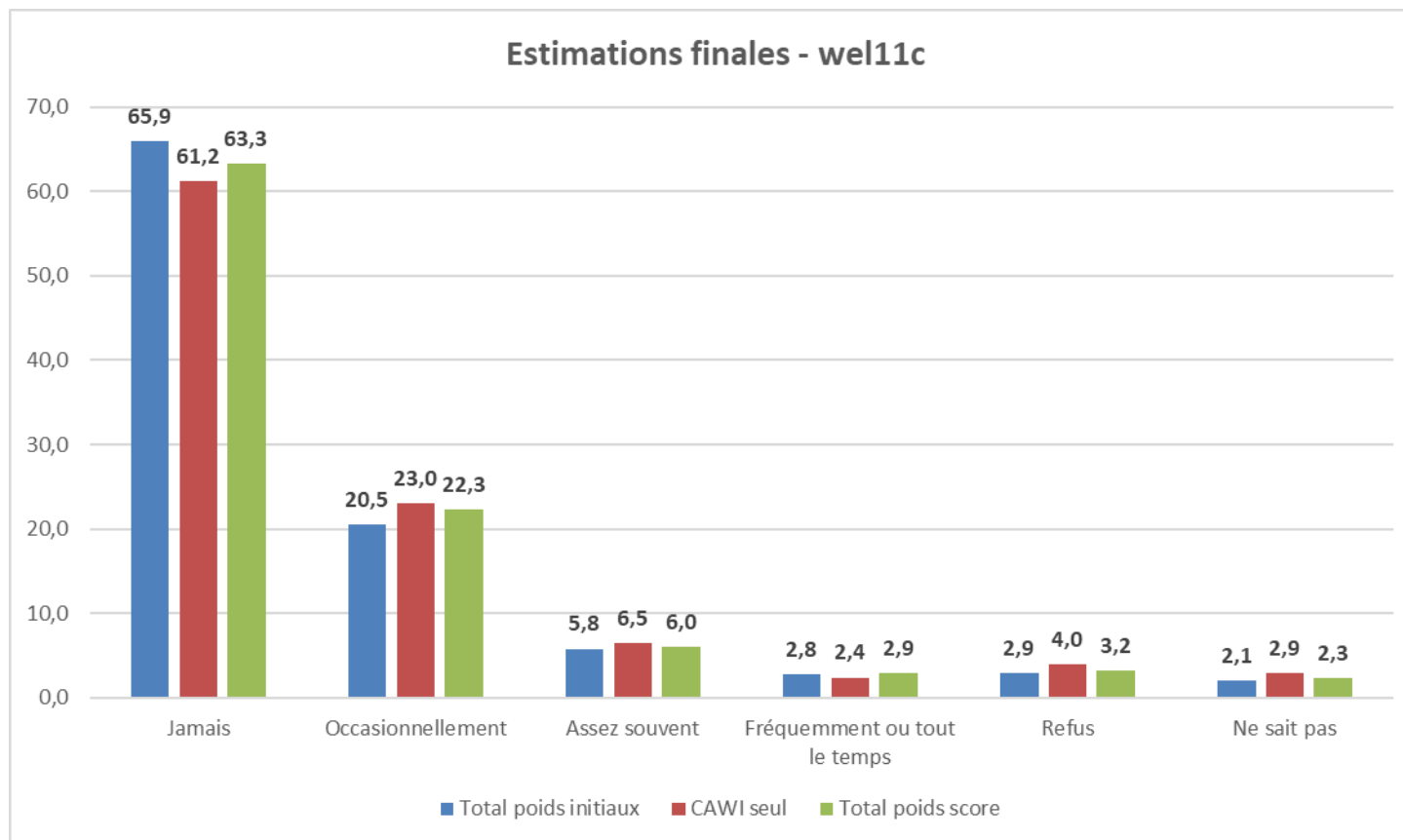
Impact sur les estimations totales



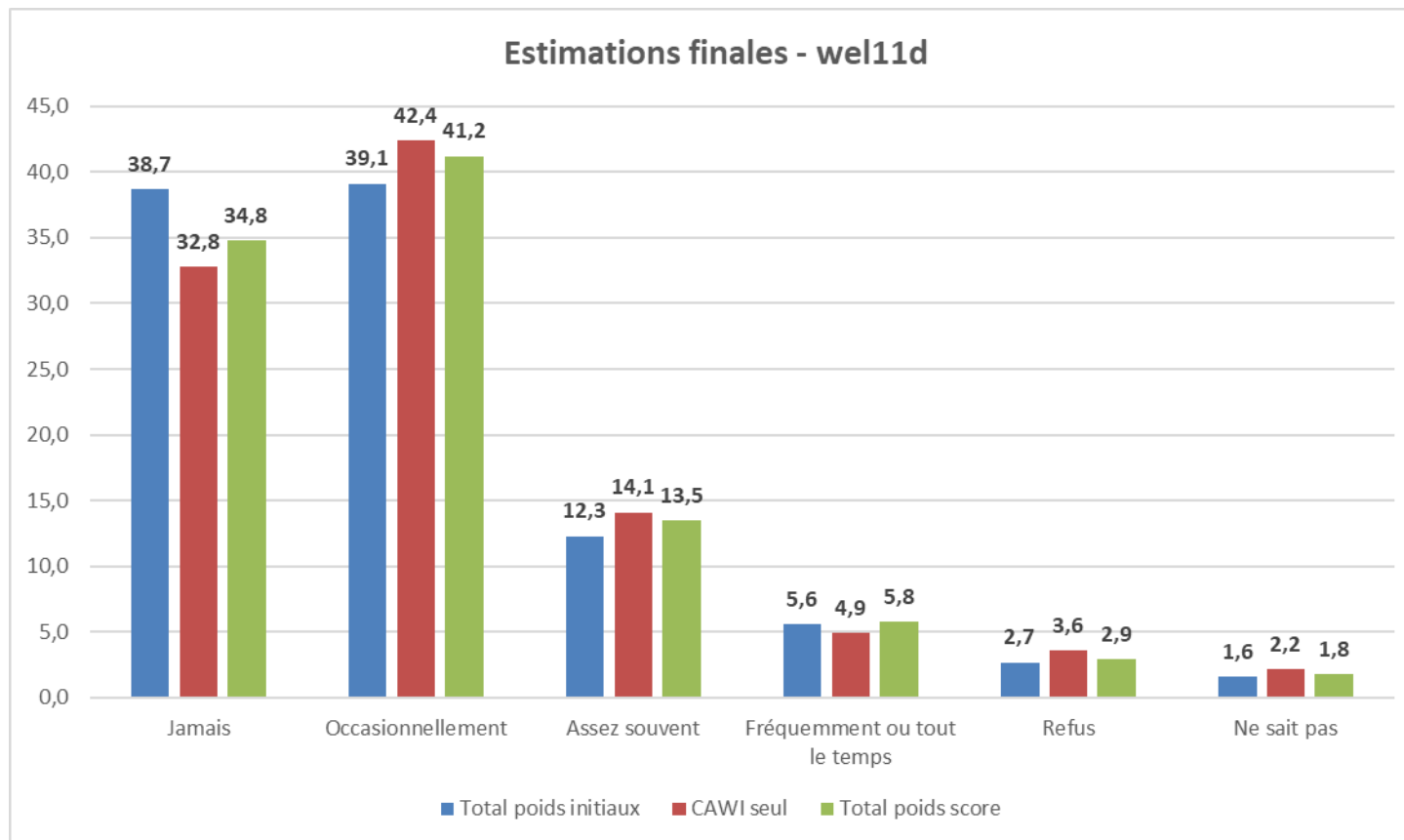
Impact sur les estimations totales



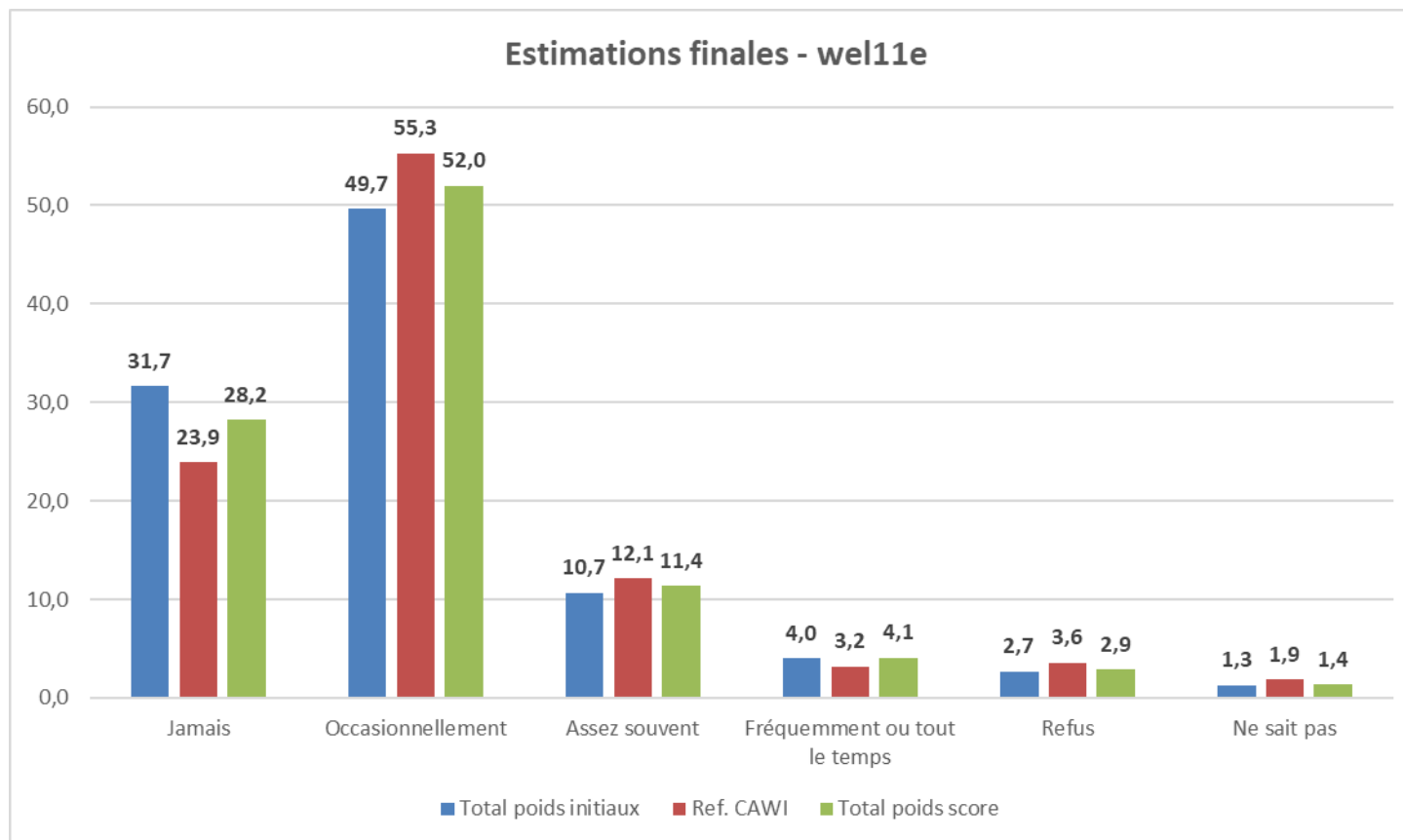
Impact sur les estimations totales



Impact sur les estimations totales



Impact sur les estimations totales



Conclusions et discussion

- **L'imputation : une méthode adaptée aux biais de mesure**
 - Une solution naturelle quand le biais porte sur la valeur des réponses.
 - Une méthode parcimonieuse basée sur l'effet causal et ciblée sur les observations les plus porteuses du biais ;
 - Une efficacité qui dépend aussi de la taille du support d'imputation : possibilité d'assouplir les contraintes d'appariement (remise, nombre de voisins).
 - Imputation multiple : Reflète l'incertitude liée à l'imputation (important lorsque le biais est élevé).

Conclusions et discussion (2/2)

- **... Mais qui présente certaines limites**
 - Une méthode déontologiquement contestable ;
 - Le choix du mode de collecte de référence : une hypothèse déterminante mais discutable ;
 - Ne prend pas en compte les potentiels biais de composition inobservés ;
 - Utilisation complexe dans le cadre de l'exploitation des données ;
 - Problématique pour des enquêtes longitudinales.

- **Le calage : une alternative à l'imputation**
 - Ne modifie pas les réponses des participants à l'enquête ;
 - Mais la correction par repondération n'est pas curative : elle permet d'obtenir une prévalence globale cohérente mais n'assure pas de neutraliser l'effet de mode.

Merci de votre attention

Bibliographie

- [1] Castell L., Clerc M., Croze D., Legleye S., Nougaret A., « Victimations déclarées et effets de mode : enseignements de l'expérimentation panel multimode de l'enquête Cadre de Vie et Sécurité », document de travail, Insee, 2023.
- [2] Cissé M., Barret C., « Agrégation de données multimode : Impact sur la modélisation des variables présentant un effet de mesure », 13èmes Journées de Méthodologie Statistique, 2018.
- [3] Cottrell G., Cot M., Mary J.-Y., « L'imputation multiple des données manquantes aléatoirement : concepts généraux et présentation d'une méthode Monte-Carlo », Revue d'épidémiologie et de Santé Publique, Volume 57, Issue 5, pages 361-372, 2009.
- [4] Legleye S., « Agrégation de données multimodes : un état des lieux et une proposition pragmatique pour contrôler l'effet de mesure », Insee, 2018.
- [5] Longford N. T., « Exemples d'imputations multiples », Recueil du symposium 2002 de Statistique Canada, Statistique Canada, 2002.
- [6] Quantin S., « Estimation avec le score de propension sous R », Méthodologie statistique, document de travail, Insee, 2018.