

Statbel

(Direction générale Statistique – Statistics Belgium)

Service Statistiques sociales (SSS)

Collecte de données Enquêtes Citoyens (CEC)

Service méthodologie (SM)

Enquête sur le budget des ménages (EBM) 2020

Note méthodologique

Personne de contact: Kelly Sabbe
Statbel (Direction générale Statistique – Statistics Belgium)
Chaque jour ouvrable entre 9h et 16h.



Boulevard du Roi Albert II
16
1000 Bruxelles

+32 (0) 800 120 33

statbel@economie.fgov.be
www.statbel.fgov.be



Table des matières

Liste des tableaux.....	5
Liste des figures.....	6
Liste des abréviations.....	7
1 Introduction générale.....	8
1.1 Qu'est-ce que l'Enquête sur le budget des ménages?.....	8
1.2 Base juridique.....	8
1.3 Quand l'EBM a-t-elle lieu?.....	8
1.4 Nombre de ménages participants.....	9
2 Les sous-échantillons pour EBM 2020.....	11
2.1 Introduction.....	11
2.2 Recrutement à partir des échantillons EFT – Source 1.....	13
2.2.1 Introduction.....	13
2.2.2 Degré 1 : tirage des unités géographiques (ou UPE)......	13
2.2.2.1 Cadre d'échantillonnage des UPE.....	13
2.2.2.2 Le plan de sondage pour le tirage des UPE.....	14
2.2.2.3 Répartition des UPE dans le temps.....	15
2.2.3 Degré 2 : tirage des ménages EFT (UPE).....	15
2.2.3.1 Le cadre d'échantillonnage des ménages EFT.....	15
2.2.3.2 Le plan de sondage pour le tirage des ménages EFT.....	16
2.2.3.3 Répartition des ménages-EFT dans le temps.....	17
2.2.4 Degré 3 : 'sélection' des ménages EBM (UTE).....	17
2.2.4.1 Le cadre d'échantillonnage et le plan de sondage pour la sélection des ménages EBM.....	17
2.2.4.2 Répartition des ménages EBM dans le temps.....	17
2.2.5 Probabilités de sélection et poids de sondage.....	18
2.3 Recrutement via l'échantillon EBM 2018 – Source 2.....	19
2.3.1 Introduction.....	19
2.3.2 Sous-échantillons trimestriels d'EBM 2020 de la source 2.....	19
2.3.2.1 Le sous-échantillon d'EBM 2020 T1 de la source 2.....	21
2.3.2.2 Le sous-échantillon d'EBM 2020 T2 de la source 2.....	21
2.3.2.3 Le sous-échantillon d'EBM 2020 T3 de la source 2.....	22
2.3.2.4 Le sous-échantillon d'EBM 2020 T4 de la source 2.....	22

2.3.3	Répartition des ménages d'EBM 2020 sur 24 demi-mois	22
2.4	Tirages complémentaires du RN – Source 3	23
2.4.1	Introduction	23
2.4.2	Degré 1 : tirage des unités géographiques (ou UPE)	23
2.4.3	Degré 2 : tirage des ménages EBM (USE)	23
2.4.4	Probabilités de sélection et poids de sondage	25
2.4.4.1	Pour les ménages de type EFT	25
2.4.4.2	Pour les ménages âgés	26
3	Taux de réponse	27
3.1	Taux de réponse initial de l'EFT - Source 1	27
3.2	Taux de réponse initial de l'EBM - Source 2	28
3.3	Taux de réponse initial tirage complémentaire du RN - Source 3	29
3.4	Echantillon initial total EBM 2020 par mois	30
3.5	Aperçu du travail de terrain et du taux de réponse	31
3.6	Raisons de la non-réponse	33
4	Collecte et traitement des données pour le calibrage	34
4.1	Collecte des données	34
4.2	Pandémie de coronavirus et impact sur le travail de terrain	34
4.3	Enregistrement des données	35
4.3.1	Questionnaire ménage	35
4.3.2	Carnet de dépenses	35
4.3.3	Fourchettes de prix	36
5	Analyse partie 1: détermination de la personne de référence du ménage sur la base du revenu	37
5.1	Introduction	37
5.2	Déterminer le revenu net	38
6	Calibrage et estimation	40
6.1	Introduction	40
6.2	Cadres de calibrage	40
6.2.1	Un échantillon EFT 2020 calibré	40
6.2.2	Population de ménages de type EFT issus du RN	41
6.2.3	Population de ménages âgés issus du RN	42
6.3	Calibrage d'EBM 2020	43
6.3.1	Calibrage de ménages de type EFT de l'EFT (source 1) et du RN (source 3.1)	43

6.3.2	Calibrage de ménages de type EFT d'EBM 2018 (source 2).....	43
6.3.3	Calibrage des ménages âgés (source 3.2).....	44
6.3.4	Pourcentage de propriétaires de logements dans la population.....	45
6.4	Estimation des paramètres.....	46
7	Analyse partie 2: traitement des données pour publication et diffusion.....	47
7.1	Représentativité du sondage.....	47
7.1.1	Répartitions au niveau du ménage.....	47
7.1.2	Répartitions au niveau individuel.....	48
7.2	Traitement des données.....	49
7.2.1	Identification des valeurs aberrantes.....	49
7.2.2	Imputation des dépenses énergétiques.....	50
7.2.3	Collecte des informations des carnets.....	50
7.2.4	Ventilation de factures groupées.....	51
7.2.5	Imputer les données manquantes.....	52
7.3	Loyer imputé.....	53
7.4	Extrapolation des données.....	54
8	Estimation de la variance.....	55
8.1	Introduction.....	55
8.2	Description du plan de sondage.....	56
8.2.1	Sources 1 et 3.1.....	56
8.2.2	Source 2.....	57
8.2.3	Source 3.2.....	57
8.3	Calibrage.....	58
8.3.1	Correction de la non-réponse.....	58
8.3.2	Combinaison des sources.....	58
8.4	Aperçu de l'estimation de la variance.....	59
9	Publication, diffusion et points d'attention.....	61
9.1	Publication sur le site web.....	61
9.2	Diffusion des microdonnées disponibles pour la recherche.....	61
9.3	Point d'attention: dépenses courantes / saisonnières.....	62
	Références.....	64
	Références aux pages web pertinentes.....	65

Liste des tableaux

Tableau 1: Liste des abréviations	7
Tableau 2: Répartition des échantillons par région	9
Tableau 3: Aperçu des sources pour les échantillons trimestriels (de T1 à T4) et bimensuels (de P01 à P24) de l'EBM 2020	12
Tableau 4: Degré 1 - tirage des UPE ou GRP, répartition des GRP entre les strates et trimestres ou GR - y compris la DEU	14
Tableau 5: Répartition uniforme des groupes EFT par trimestre ou par GR entre les semaines - y compris la DEU.....	15
Tableau 6: Répartition quasi uniforme des groupes EFT par trimestre ou par GR entre les semaines - hors DEU.....	15
Tableau 7: Degré 2 - tirage des MN: répartition des MN entre strates et trimestres ou GR - y compris la DEU.....	16
Tableau 8: Répartition quasi uniforme des ménages EFT par trimestre ou par GR entre les semaines - hors DEU.....	17
Tableau 9: Fractions approximatives de la répartition des ménages EBM par semaine entre les demi-mois	18
Tableau 10: Aperçu schématique de la création du sous-échantillon d'EBM 2020 via la source 2	20
Tableau 11: Les 4 échantillons trimestriels complémentaires d'EBM 2020: informations de base	23
Tableau 12: Les 4 échantillons trimestriels complémentaires d'EBM 2020: paramètres d'échantillon et distribution.....	24
Tableau 13: Les 4 échantillons trimestriels complémentaires d'EBM 2020: chiffres de la population (limités aux UPE tirées dans les GR8-15)	25
Tableau 14: Poids de sondage dans les échantillons complémentaires EBM 2020 de ménages de type EFT	25
Tableau 15: Poids de sondage dans les échantillons EBM 2020 complémentaires de ménages âgés	26
Tableau 16: Taux de réponse initial de l'EFT - Source 1	27
Tableau 17: Volonté de répondre selon le mode d'enquête	27
Tableau 18: Taux de réponse initial de l'EBM - Source 2	28
Tableau 19: Taux de réponse initial du tirage complémentaire du RN - Source 3.....	29
Tableau 20: Echantillon initial total EBM 2020 par mois	30
Tableau 21: Aperçu du travail de terrain et du taux de réponse pour la Belgique et les régions pour EBM 2020.....	31
Tableau 22: raison du refus de participer à l'EBM.....	33
Tableau 23: Validation du revenu pondéré au niveau du ménage collecté dans l'EBM par rapport aux données collectées dans EU-SILC	39

Tableau 24: Répartition du revenu selon les classes.....	39
Tableau 25: Sommes des poids calibrés pour l'échantillon de répondants calibré d'EFT 2020 vague 1	41
Tableau 26: Répartition en pourcentage et en chiffres absolus des classes de revenus du ménage par région - variable INK_5_G	42
Tableau 27: Nombre de ménages âgés selon la région et la taille du ménage (RN du 1/07/2020)	42
Tableau 28: Echantillon de répondants EBM 2020 de ménages âgés: nombre de répondants et sommes des poids calibrés selon la province, la taille du ménage et le trimestre	45
Tableau 29: Estimation du nombre de propriétaires de logements (EBM 2020)	46
Tableau 30: Répartition selon la taille des ménages participants.....	47
Tableau 31: Répartition selon le nombre d'actifs par ménage des ménages participants.....	47
Tableau 32: Répartition selon le sexe des membres des ménages participants.....	48
Tableau 33: Répartition selon l'âge des membres des ménages participants	48
Tableau 34: Différence de montant après l'application de la méthode RFM pour ECOICOP 1 et 2 (non pondéré)	49
Tableau 35: Nombre de données manquantes par catégorie pour les dépenses énergétiques.....	50
Tableau 36: Nombre de données imputées des carnets pour les dépenses énergétiques	50
Tableau 37: Aperçu des factures groupées pour les dépenses énergétiques.....	51
Tableau 38: Nombre de données imputées sur la base des modèles de régression pour les dépenses énergétiques.....	52
Tableau 39: Validation du loyer réel et imputé par rapport à l'enquête EU-SILC (pondéré)	53
Tableau 40: Aperçu de l'estimation de la variance pour les catégories ECOICOP jusqu'au 2e niveau pour la Belgique et les régions	59
Tableau 41: Données au niveau du ménage et au niveau individuel.....	61

Liste des figures

Figure 1: Flowchart des étapes du traitement des données et du service impliqué	10
---	----

Liste des abréviations

Tableau 1: Liste des abréviations

ANT	Anvers
BRU	Région de Bruxelles-Capitale
BWA	Brabant wallon
CAPI	Computer-Assisted Personal Interviewing
CATI	Computer-Assisted Telephone Interviewing
CAWI	Computer-Assisted Web Interviewing
ECOICOP	European Classification of Individual Consumption according to Purpose
CEC	Collecte de données Enquêtes Citoyens
DEU	Communauté germanophone
SM	Service méthodologie
SSS	Service Statistiques sociales
EU-SILC	European Union Statistics on Income and Living Conditions
UCM	Unité de consommation modifiée
HAI	Hainaut
EBM	Enquête sur le budget des ménages
HBS	Household Budget Survey
MN	Ménage(s)
EFT	Enquête sur les forces de travail
LIE	Liège
LIM	Limbourg
LUX	Luxembourg
NAM	Namur
OVL	Flandre orientale
PS	Post-stratification
PPSI	Paper-and-pencil Self-administered Interviewing
UPE	Unité(s) primaire(s) d'échantillonnage
GR	Groupe(s) de rotation
RN	Registre national
USE	Unité(s) secondaire(s) d'échantillonnage
UTE	Unité(s) tertiaire(s) d'échantillonnage
VLA	Région flamande
VLB	Brabant flamand
WAL	Région wallonne
WVL	Flandre occidentale

1 Introduction générale

1.1 Qu'est-ce que l'Enquête sur le budget des ménages?

Statbel réalise l'Enquête sur le budget des ménages (EBM, en anglais HBS: Household Budget Survey) depuis 1853. L'objectif de cette enquête est d'analyser les dépenses des ménages privés et de fournir à ce sujet des données descriptives en lien avec certaines caractéristiques des ménages. Comme cette enquête collecte une grande quantité d'informations uniques, ses résultats sont souvent utilisés afin de définir, suivre et évaluer le comportement d'achat des ménages.

L'Enquête sur le budget des ménages est une enquête nationale axée sur les dépenses supportées par les ménages en vue de l'acquisition de biens et de services. Elle dresse un tableau des conditions de vie dans l'Union européenne (UE). Elle est menée par chaque État membre et est utilisée pour définir la pondération d'importants indicateurs macroéconomiques, comme les indices des prix à la consommation (qui mesurent l'inflation) et les comptes nationaux.

Les données des enquêtes sont ventilées en fonction des caractéristiques des ménages: revenus, caractéristiques socioéconomiques, taille et composition, degré d'urbanisation et région.

La collecte des données combine la réalisation d'un ou plusieurs entretiens et la tenue d'un journal complété quotidiennement par les ménages.

1.2 Base juridique

Il n'y a pas de règlement-cadre européen pour UE-EBM. L'enquête n'est donc pas obligatoire pour les États membres de l'UE. UE-EBM est réalisée sur la base d'un 'gentlemen's agreement'. Cet accord a été obtenu lors de la réunion de 1989 du Comité du programme statistique.

Le *gentlemen's agreement* contient des directives concernant la production d'EBM qui sont conseillées à tous les États membres de l'UE afin que les résultats soient le plus possible comparables au niveau international. Il s'agit par exemple de l'échantillonnage, des définitions des concepts et des variables, des classifications, etc.

L'objectif de l'enquête sur le budget des ménages est de dégager un cadre global pour la production de données statistiques 'communautaires' relatives à la consommation des ménages (UE-EBM) sur la base de données transversales (montant, composition,...) sur les postes de leur budget.

Un [arrêté royal de janvier 1996](#) fixe le montant des indemnités aux ménages et un autre [arrêté royal du 17 mai 2006](#) le montant des indemnités accordées aux enquêteurs.

1.3 Quand l'EBM a-t-elle lieu?

L'EBM a été réformée en 2012 pour devenir une enquête bisannuelle afin de produire des chiffres de meilleure qualité, plus adaptés aux demandes de nos principaux utilisateurs et ce, dans des délais plus courts après la période de collecte de données.

L'EBM 2014 a été organisée de la même manière que l'EBM 2012, avec quelques améliorations au niveau du questionnaire ménage (interviews réalisées en CAPI). L'EBM 2016 a été organisée de la même manière qu'en 2014.

Afin d'alléger la charge d'enregistrement des ménages et d'être plus en phase avec les autres pays européens, la période d'enregistrement a été modifiée à partir d'EBM 2018. Les ménages ne doivent plus enregistrer leurs dépenses pendant un mois mais pendant un demi-mois, du 1^{er} au 15^e jour ou du 16^e jour à la fin du mois.

1.4 Nombre de ménages participants

Depuis 2012, la taille visée de l'échantillon est de 5.000 ménages. La personne de référence du ménage est la personne ayant le revenu net le plus élevé. L'échantillon ne comprend pas de ménages collectifs (couvents, maisons de repos, prisons, etc.). Le Tableau 2 donne un aperçu du nombre de ménages par région pour la période à partir de 2012.

Tableau 2: Répartition des échantillons par région

Nombre de ménages	2012	2014	2016	2018	2020
Région de Bruxelles-Capitale (BRU)	723	633	531	713	680
Région flamande (VLA)	3.254	2.893	2.175	3.103	3.155
Région wallonne (WAL)	2.604	2.609	1.784	2.320	2.270
Au niveau national	6.581	6.135	4.490	6.136	6.105

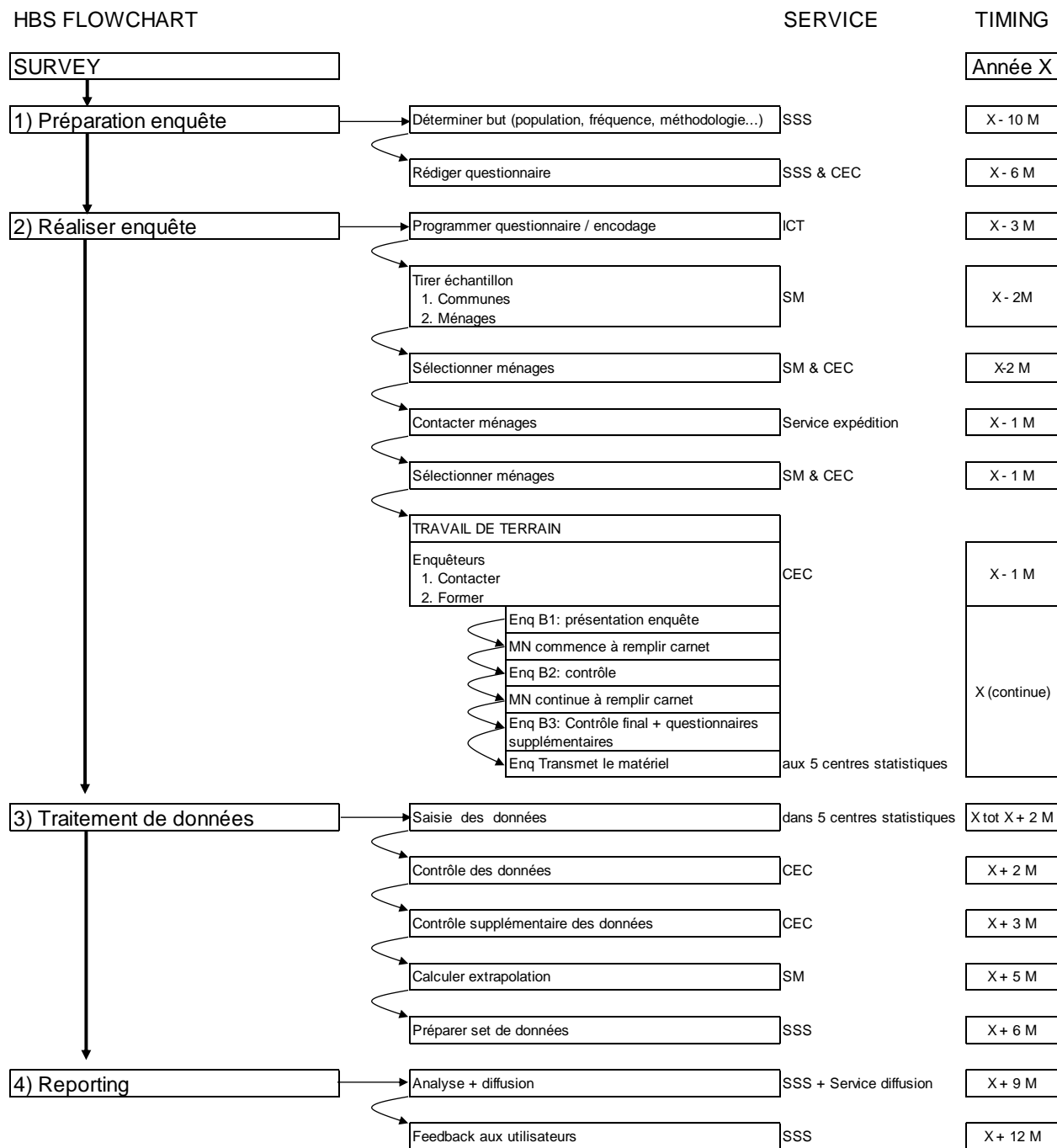
Les FAQ fournissent des informations générales complémentaires sur EBM et son histoireⁱ.

La suite du document donne une description aussi chronologique que possible du tirage de l'échantillon, de la réponse initiale et définitive, de la collecte et du traitement des données pour le calibrage, du calibrage, du traitement des données pour publication et diffusion et de la manière dont se déroulent la publication et la diffusion. La Figure 1 montre quel est le service impliqué lors de chaque étape du traitement.

Remarque

L'intégralité de la méthodologie d'EBM 2020 est identique à celle de 2018. Par conséquent, pour certaines parties qui sont une répétition littérale, il est fait référence à l'*Enquête sur le budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b)) pour plus d'informations.

Figure 1: Flowchart des étapes du traitement des données et du service impliqué



2 Les sous-échantillons pour EBM 2020

2.1 Introduction

L'échantillon de ménages pour EBM 2020 est composé à partir de 3 sources¹:

- Source 1: Échantillons trimestriels EFT, ou *groupes de rotation* (GR);
- Source 2: l'échantillon de ménages répondants d'EBM 2018, limité aux ménages de type EFT, c'est-à-dire les ménages des sources 1 et 3.1 d'EBM 2018;
- Source 3: un échantillon complémentaire, tiré directement du Registre national (RN), avec
 - Source 3.1: ménages de type EFT dans le RN,
 - Source 3.2: ménages 'âgés' dans le RN.

Le recrutement de la source 1 se fait parmi les répondants EFT de la vague 4; pour les GR 8 à 15 (voir Tableau 3), l'interrogation de la vague 4 a eu lieu du trimestre 2018 T3 au trimestre 2020 T2 inclus, respectivement. L'EFT est également toujours basée sur un échantillon de ménages du RN. La source 1 est discutée au paragraphe 2.2.

On entend par **ménage de type EFT** (dans la base de sondage), un ménage dont au moins un membre appartient à la classe d'âge des 15-76² ans, à savoir un ménage qui serait aussi éligible pour le tirage d'EFT. Le critère d'âge doit être valable au moment du tirage. On entend par **ménage âgé** (dans la base de sondage) un ménage dont tous les membres ont au moins 77 ans au moment du tirage. Dans un premier temps, nous parlons de ménages de type EFT et de ménages âgés dans le contexte du tirage supplémentaire du RN (source 3). Mais les ménages qui ont déjà été tirés pour EFT ou recrutés pour EBM 2016 via des échantillons EFT peuvent également être considérés comme des ménages de type EFT.

La source 2 est une partie de l'échantillon de répondants d'EBM 2018. Ce sous-échantillon provient aussi indirectement du RN. Cette source est discutée au paragraphe 2.3.

La source 3 est un échantillon complémentaire du RN et est abordée au paragraphe 2.4.

De cette manière, tous les ménages d'EBM 2020 proviennent du RN, via des échantillons aléatoires, mais le parcours que ces ménages ont suivi pour arriver dans l'échantillon d'EBM 2020 diffère fortement d'une source à l'autre. Tout cela signifie que la composition de l'échantillon EBM 2020 est complexe et applique un plan de sondage compliqué, ce qui complexifie le traitement des données (calibrage, estimation, estimation de la variance).

Le passage à la tenue d'un carnet pendant un demi-mois³, introduit depuis l'EBM 2018, a été maintenu dans l'EBM 2020 (Sabbe, Delclite, & Geenens, 2021). Les ménages recrutés pour l'enquête EBM 2020 doivent donc être affectés à l'un des 24 *échantillons bimensuels*. Nous désignerons ces quinzaines par P01, P02, ... P24.

¹ Il convient de prêter attention à la notion de 'source(s)': si nous voulons dire une source d'EBM 2020, nous indiquons en bref 'source 1', etc. Cependant, si nous faisons référence à une source d'EBM 2018, nous écrivons en entier 'source 1 d'EBM 2018', etc. Attention: EBM 2018 est elle-même 'source 2', c'est-à-dire l'une des trois sources d'EBM 2020.

² Si nous parlons d'une personne du groupe d'âge A-B, cela signifie que cette personne a au moins A ans, mais n'a pas encore B+1 ans.

³ La notion de 'demi-mois' ne doit pas être prise au sens strict: l'année est scindée en 24 périodes sur la base des mois calendrier. Chaque mois, une première période commence le 1^{er} jour et se termine le 15^e jour du mois, et une deuxième période commence le 16^e jour et se termine le dernier jour du mois.

Le Tableau 3 présente un bref aperçu des sources d'EBM 2020 et de la réalisation des 24 sous-échantillons EBM bimensuels.

Tableau 3: Aperçu des sources pour les échantillons trimestriels (de T1 à T4) et bimensuels (de P01 à P24) de l'EBM 2020

Sources			EBM 2020
EFT panel (*)	EBM 2018	RN	
Source 1: Répondants EFT			
GR8 & GR9			T1 ~ P01-P06
GR10 & GR11			T2 ~ P07-P12
GR12 & GR13			T3 ~ P13-P18
GR14 & GR15			T4 ~ P19-P24
Source 2: Répondants d'EBM 2018 (°)			
	T1		T1 ~ P01-P06
	T2		T2 ~ P07-P12
	T3		T3 ~ P13-P18
	T4		T4 ~ P19-P24
Source 3: Registre national (RN)			
		UPE pour les GR8 & GR9	T1 ~ P01-P06
		UPE pour les GR10 & GR11	T2 ~ P07-P12
		UPE pour les GR12 & GR13	T3 ~ P13-P18
		UPE pour les GR14 & GR15	T4 ~ P19-P24
(*) Le recrutement se fait parmi les répondants EFT qui ont participé à la 4 ^e (et dernière) vague pour les GR mentionnés. Il convient de noter que le scénario du panel pour tous les GR mentionnés ici (de 8 à 15) est de 2(2)2.			
(°) Pour la source 2, ce tableau 3 diffère du tableau 3 du document <i>Enquête sur le budget des ménages 2018 - Note méthodologique</i> (Statbel (2019a) ou Statbel (2019b)): ici, nous ne mentionnons pas de quels échantillons EFT les échantillons d'EBM 2018 sont issus. Il convient toutefois de noter que certains répondants d'EBM 2018 ne sont pas issus d'EFT, mais directement du RN pour l'EBM 2018. Nous expliquons cela plus en détails au paragraphe 2.3.			

Le recrutement à partir des échantillons EFT (source 1) est comparable à l'approche de EBM 2018 (voir *Enquête sur le budget des ménages 2018 - Note méthodologique* Statbel (2019a) ou Statbel (2019b)) mais l'introduction d'un panel pour l'EFT - depuis 2016 T3 en ce qui concerne l'échantillonnage - a eu un impact sur la méthode de travail pour obtenir les échantillons EBM bimensuels. Nous expliquons cela plus en détails au paragraphe 2.2. Comme pour l'EBM 2018, la source 2 doit être considérée comme un complément nécessaire au recrutement à partir de la source 1. Un autre complément aux sources 1 et 2 est le recrutement à partir du RN (source 3). Le recrutement à partir d'EBM 2018 (source 2) est expliqué au paragraphe 2.3. Le tirage des échantillons complémentaires du RN (source 3) est expliqué au paragraphe 2.4.

L'introduction des paragraphes 2.2, 2.3 et 2.4 suivants est identique à celle des paragraphes correspondants dans *Enquête sur le budget des ménages 2018 - Note méthodologique* (voir Statbel (2019a) ou Statbel (2019b)). Nous voulons ainsi souligner le parallèle entre le recrutement pour EBM 2018 et celui pour EBM 2020. Certains sous-paragraphes n'ont pas été développés ou ont été (fortement) raccourcis dans le document, car ils correspondent aux sous-paragraphes correspondants de la note méthodologique sur EBM 2018; dans ce cas, nous faisons explicitement référence au document *Enquête sur le budget des ménages 2018 - Note méthodologique* (voir Statbel (2019a) ou Statbel (2019b)).

2.2 Recrutement à partir des échantillons EFT – Source 1

2.2.1 Introduction

Les échantillons ou groupes de rotation (GR) EFT trimestriels sont des *échantillons à deux degrés*. Pour EBM, un recrutement complémentaire de ménages est effectué dans plusieurs GR. Ce recrutement est décrit ci-dessous comme le *troisième degré*. Une première partie (soit la partie de la source 1) des échantillons EBM pour 2020 est donc présentée ici comme des *échantillons à trois degrés*:

- Degré 1: tirage des *unités primaires d'échantillonnage* (UPE) à partir d'un cadre d'échantillonnage dérivé du RN, et composé d'UPE qui sont des unités géographiques construites (voir paragraphe 2.2.2.1);
- Degré 2: tirage des *unités secondaires d'échantillonnage* (USE) d'un cadre d'échantillonnage dérivé du RN composé d'USE, qui sont les ménages *éligibles*, appartenant aux UPE sélectionnées au degré 1;
- Degré 3: 'tirage' des *unités tertiaires d'échantillonnage* (UTE) de l'échantillon de ménages du degré 2; les unités tertiaires sont donc aussi des ménages.

Les échantillons d'EFT résultent donc des tirages au degré 1 et au degré 2. Les échantillons d'EBM sont obtenus au degré 3, parmi les répondants EFT qui ont participé à la vague 4 (v4) du panel EFT. Cette méthode de travail a permis d'obtenir 4 échantillons EBM trimestriels de la manière suivante (voir aussi Tableau 3):

- EBM 2020 T1 des GR8 &GR9 d'EFT (v4 en 2018 T3 & 2018 T4, respectivement),
- EBM 2020 T2 des GR10 &GR11 d'EFT (v4 en 2019 T1 & 2019 T2, respectivement),
- EBM 2020 T3 des GR12 &GR13 d'EFT (v4 en 2019 T3 & 2019 T4, respectivement),
- EBM 2020 T4 des GR14 &GR15 d'EFT (v4 en 2020 T1 & 2020 T2, respectivement).

Remarquez que l'interrogation d'EBM 2020 a eu lieu au moins quatre trimestres après la dernière interrogation (v4) d'EFT.

2.2.2 *Degré 1 : tirage des unités géographiques (ou UPE).*

2.2.2.1 Cadre d'échantillonnage des UPE

Les cadres d'échantillonnage d'UPE pour les GR 8 à 14, c-à-d pour les échantillons trimestriels d'EFT 2017 T2 à 2018 T4, ont été construits de la même manière que le cadre d'échantillonnage d'UPE pour le GR7: ce point est abordé dans *Enquête sur le budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b), paragraphe 2.2.2.1). Les cadres d'échantillonnage d'UPE pour les GR 8 à 10 sont dérivés d'une extraction du RN datant du 10/09/2016, et ceux des GR 11 à 14 d'une extraction du RN datant du 09/09/2017. Il convient de noter que la méthode de travail pour la construction de ces cadres d'échantillonnage d'UPE peut être brièvement résumée comme suit:

- (1) les sections, caractérisées par le nombre de ménages privés - c'est-à-dire la 'taille' de la section - sont éventuellement fusionnées en UPE suffisamment grandes au sein des parties de communes. La fusion éventuelle des sections en UPE se fait de manière automatisée;
- (2) pour les UPE, outre la 'taille' ou le nombre de ménages privés, les caractéristiques taux de chômage et revenu moyen des ménages sont ajoutées;

(3) les trois caractéristiques d'UPE sont converties en quintiles respectifs pour chaque région.

À partir du GR 15, c'est-à-dire à partir d'EFT 2019 T1, la méthode de travail a quelque peu changé:

- (1) les sections - caractérisées par leur taille (c'est-à-dire le nombre de ménages privés), le taux de chômage et le revenu moyen des ménages - sont fusionnées manuellement en UPE suffisamment grandes (toujours au sein des parties de communes), qui sont également aussi homogènes que possible en termes de taux de chômage et de revenu moyen des ménages;
- (2) comme ci-dessus;
- (3) comme ci-dessus.

2.2.2.2 Le plan de sondage pour le tirage des UPE

Sur ce point, nous renvoyons le lecteur au paragraphe 2.2.2.2 correspondant de *Enquête sur le budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b)): la description dans ce paragraphe du GR 7, tiré pour la première interrogation au trimestre 2017 T1, est valable pour tous les GR 8 à 15. Le nouveau Tableau 4 devient donc:

Tableau 4: Degré 1 - tirage des UPE ou GRP, répartition des GRP entre les strates et trimestres ou GR - y compris la DEU

strate UPE		Trimestres EFT et GR							
		2017			2018				2019
		T2	T3	T4	T1	T2	T3	T4	T1
		GR8	GR9	GR10	GR11	GR12	GR13	GR14	GR15
1	ANT	31	31	31	31	31	31	31	31
2	BRU	39	39	39	39	39	39	39	39
3	WVL	27	27	27	27	27	27	27	27
4	OVL	29	29	29	29	29	29	29	29
5	HAI	28	28	28	28	28	28	28	28
6	LIE	25	25	25	25	25	25	25	25
7	LIM	22	22	22	22	22	22	22	22
8	LUX	16	16	16	16	16	16	16	16
9	NAM	17	17	17	17	17	17	17	17
10	VLB	25	25	25	25	25	25	25	25
11	BWA	16	16	16	16	16	16	16	16
12	DEU	11	11	11	11	11	11	11	11
Total		286	286	286	286	286	286	286	286

2.2.2.3 Répartition des UPE dans le temps

Par trimestre (ou période de référence EFT), les sélections d'UPE ou les groupes EFT sont toujours répartis uniformément entre les 13 semaines de référence du trimestre. Conformément à la situation simplifiée telle que proposée dans le Tableau 4, les tableaux 5 et 6 d'Enquête sur le budget des ménages 2018 - Note méthodologique (Statbel (2019a) ou Statbel (2019b), paragraphe 2.2.2.3) peuvent également être simplifiés:

Tableau 5: Répartition uniforme des groupes EFT par trimestre ou par GR entre les semaines - y compris la DEU

Année	Trim	GR	Semaine du trimestre													Tot.
			1	2	3	4	5	6	7	8	9	10	11	12	13	
2017	2	8	22	22	22	22	22	22	22	22	22	22	22	22	22	286
	3	9	22	22	22	22	22	22	22	22	22	22	22	22	22	286
	4	10	22	22	22	22	22	22	22	22	22	22	22	22	22	286
2018	1	11	22	22	22	22	22	22	22	22	22	22	22	22	22	286
	2	12	22	22	22	22	22	22	22	22	22	22	22	22	22	286
	3	13	22	22	22	22	22	22	22	22	22	22	22	22	22	286
	4	14	22	22	22	22	22	22	22	22	22	22	22	22	22	286
2019	1	15	22	22	22	22	22	22	22	22	22	22	22	22	22	286
Total			176	176	176	176	176	176	176	176	176	176	176	176	176	2288

Tableau 6: Répartition quasi uniforme des groupes EFT par trimestre ou par GR entre les semaines - hors DEU

Année	Trim	GR	Semaine du trimestre													Tot.
			1	2	3	4	5	6	7	8	9	10	11	12	13	
2017	2	8	21	22	21	21	21	21	21	21	21	22	21	21	21	275
	3	9	21	22	21	21	21	21	21	21	21	22	21	21	21	275
	4	10	21	22	21	21	21	21	21	21	21	22	21	21	21	275
2018	1	11	21	22	21	21	21	21	21	21	21	22	21	21	21	275
	2	12	21	22	21	21	21	21	21	21	21	22	21	21	21	275
	3	13	21	22	21	21	21	21	21	21	21	22	21	21	21	275
	4	14	21	22	21	21	21	21	21	21	21	22	21	21	21	275
2019	1	15	21	22	21	21	21	21	21	21	21	22	21	21	21	275
Total			168	176	168	168	168	168	168	168	168	176	168	168	168	2200

2.2.3 Degré 2 : tirage des ménages EFT (UPE)

2.2.3.1 Le cadre d'échantillonnage des ménages EFT

Les échantillons de ménages sont tirés trimestre après trimestre et un cadre d'échantillonnage de ménages est à nouveau construit chaque trimestre, chaque fois sur la base d'une extraction aussi récente que possible du RN:

- pour l'EFT 2017 T2 (soit GR8 et EBM 2020 T1): extraction datant du 11/02/2017
- pour l'EFT 2017 T3 (soit GR9 et EBM 2020 T1): extraction datant du 20/05/2017
- pour l'EFT 2017 T4 (soit GR10 et EBM 2020 T2): extraction datant du 12/08/2017
- pour l'EFT 2018 T1 (soit GR11 et EBM 2020 T2): extraction datant du 18/11/2017
- pour l'EFT 2018 T2 (soit GR12 et EBM 2020 T3): extraction datant du 10/02/2018

- pour l'EFT 2018 T3 (soit GR13 et EBM 2020 T3): extraction datant du 19/05/2018
- pour l'EFT 2018 T4 (soit GR14 et EBM 2020 T4): extraction datant du 11/08/2018
- pour l'EFT 2019 T1 (soit GR15 et EBM 2020 T4): extraction datant du 17/11/2018

Le cadre d'échantillonnage de ménages (USE) pour un trimestre donné, dérivé de l'extraction correspondante du RN dans laquelle on peut trouver tous les ménages belges, se compose toujours de ménages qui répondent aux critères suivants:

- [1] Seuls les ménages qui appartiennent aux UPE sélectionnées au degré 1 sont retenus.
- [2] Nous nous limitons aux ménages privés; les ménages collectifs comme les maisons de repos, les prisons,... sont donc exclus, conformément aux instructions d'Eurostat pour EFT.
- [3] Les ménages, qui ont été sélectionnés dans un échantillon EFT pour l'un des 10 trimestres précédents, n'entrent plus en ligne de compte pour une sélection.
- [4] Au moins un membre du ménage a atteint au moins l'âge de 15 ans et au maximum l'âge de 76 ans. L'âge des membres du ménage est calculé le dimanche de la semaine de référence (c'est-à-dire le dernier jour de la semaine de référence), déterminée selon le paragraphe 2.2.2.3 pour le groupe EFT auquel le ménage appartient.

2.2.3.2 Le plan de sondage pour le tirage des ménages EFT

Sur ce point, nous renvoyons le lecteur au paragraphe 2.2.3.2 correspondant d'*Enquête sur le budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b)): la description dans ce paragraphe du GR 7, tiré pour la première interrogation au trimestre 2017 T1, est valable pour tous les GR 8 à 15. Le nouveau Tableau 4 devient donc:

Tableau 7: Degré 2 - tirage des MN: répartition des MN entre strates et trimestres ou GR - y compris la DEU

Strate UPE		Taille du groupe	Trimestres EFT et GR							
			2017			2018				2019
			T2	T3	T4	T1	T2	T3	T4	T1
			GR8	GR9	GR10	GR11	GR12	GR13	GR14	GR15
1	ANT	23	713	713	713	713	713	713	713	713
2	BRU	26	1014	1014	1014	1014	1014	1014	1014	1014
3	WVL	23	621	621	621	621	621	621	621	621
4	OVL	23	667	667	667	667	667	667	667	667
5	HAI	23	644	644	644	644	644	644	644	644
6	LIE	23	575	575	575	575	575	575	575	575
7	LIM	23	506	506	506	506	506	506	506	506
8	LUX	23	368	368	368	368	368	368	368	368
9	NAM	23	391	391	391	391	391	391	391	391
10	VLB	23	575	575	575	575	575	575	575	575
11	BWA	23	368	368	368	368	368	368	368	368
12	DEU	23	253	253	253	253	253	253	253	253
Total		-	6695	6695	6695	6695	6695	6695	6695	6695

2.2.3.3 Répartition des ménages-EFT dans le temps

Au paragraphe 2.2.2.3, nous avons abordé la répartition des groupes EFT entre les semaines de référence. On en déduit facilement la répartition des ménages EFT entre les semaines de référence. Le Tableau 8 suivant montre cette répartition, sans tenir compte des ménages de la DEU, étant donné qu'ils n'entrent pas en ligne de compte pour l'EFT.

Tableau 8: Répartition quasi uniforme des ménages EFT par trimestre ou par GR entre les semaines - hors DEU

Année	Trim	GR	Semaine du trimestre													Tot.	
			1	2	3	4	5	6	7	8	9	10	11	12	13		
2017	2	8	436	459	436	436	436	436	436	436	436	436	459	436	436	436	5714
	3	9	436	459	436	436	436	436	436	436	436	436	459	436	436	436	5714
	4	10	436	459	436	436	436	436	436	436	436	436	459	436	436	436	5714
2018	1	11	436	459	436	436	436	436	436	436	436	436	459	436	436	436	5714
	2	12	436	459	436	436	436	436	436	436	436	436	459	436	436	436	5714
	3	13	436	459	436	436	436	436	436	436	436	436	459	436	436	436	5714
2019	4	14	436	459	436	436	436	436	436	436	436	436	459	436	436	436	5714
	1	15	436	459	436	436	436	436	436	436	436	436	459	436	436	436	5714
Total			3488	3672	3488	3488	3488	3488	3488	3488	3488	3488	3672	3488	3488	3488	45712

2.2.4 Degré 3 : 'sélection' des ménages EBM (UTE)

2.2.4.1 Le cadre d'échantillonnage et le plan de sondage pour la sélection des ménages EBM

Au paragraphe 2.2.1, nous avons déjà indiqué que les 4 échantillons trimestriels EBM 2020 de ménages sont dérivés des 8 échantillons EFT de GR8 à GR15. La correspondance est simple: quatre paires de GR donnent chacune un échantillon EBM 2020 trimestriel.

Les ménages qui appartiennent aux échantillons EFT forment à proprement parler le cadre d'échantillonnage de la 'sélection' de ménages pour EBM 2020 au degré 3. Il s'agit ici d'une *autosélection* de ménages, dans le sens où seuls les ménages qui ont entièrement terminé la dernière interrogation EFT prévue (c-à-d la 4^e) ont été invités à participer à l'EBM. La probabilité qu'un ménage qui a été sélectionné pour EFT (GR8 à GR15) appartienne à l'échantillon EBM 2020 correspond donc à sa probabilité de réponse à l'EFT.

2.2.4.2 Répartition des ménages EBM dans le temps

Il nous reste encore seulement à expliquer comment les ménages sont répartis entre chacun des échantillons trimestriels entre chaque période de 3 demi-mois.

Les ménages de chaque GR sont répartis de manière pratiquement équitable entre les 13 semaines du trimestre durant lequel le GR est lancé. Le Tableau 9 montre comment les ménages sont attribués – de manière systématique – aux 3 demi-mois (ou 3 'périodes') durant les mois qui correspondent au GR. L'objectif est de parvenir à une répartition aussi uniforme que possible des ménages sur l'ensemble de l'année 2020.

Tableau 9: Fractions approximatives de la répartition des ménages EBM par semaine entre les demi-mois

Demi-Mois	Semaine du trimestre													
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	1	1	1	1	0,33									4,33
2					0,66	1	1	1	0,66					4,33
3									0,33	1	1	1	1	4,33
	1	1	1	1	1	1	1	1	1	1	1	1	1	13

Répartir équitablement les ménages en 13 semaines sur 3 demi-mois peut être réalisé de manière approximative en attribuant à chaque demi-mois les ménages en 4,33 semaines.⁴ Cela peut se faire de manière systématique en attribuant 100% des ménages des 4 premières semaines et 33% des ménages de la semaine 5 au premier demi-mois. Les 66% restants des ménages de la semaine 5, nous les attribuons au deuxième demi-mois.⁵ Au deuxième demi-mois, nous attribuons ensuite aussi tous les ménages (100%) des semaines 6, 7 et 8 et 66% des ménages de la semaine 9. Nous poursuivons de cette manière – systématique – l’attribution des ménages de la semaine 9 (33%) et des semaines 10 à 13 au troisième demi-mois.

Remarquez que tous les ménages d’un même groupe EFT (de 23 ou 26 ménages) sont attribués à la même semaine de référence EFT. Attribuer tous les ménages d’une semaine au même demi-mois équivaut donc à attribuer tous les groupes EFT d’une semaine donnée au même demi-mois. Quand les ménages d’une semaine donnée sont ventilés sur 2 demi-mois, cette opération s’effectue en fait au niveau des groupes si bien que tous les ménages d’un même groupe EFT sont attribués au même demi-mois. Une répartition équitable approximative des ménages EBM sur les 24 demi-mois ou périodes est importante afin de garantir que chaque période d’observation est suffisamment représentée dans les données collectées, afin de limiter ainsi au maximum un éventuel biais en raison de fluctuations saisonnières du comportement d’achat.

Enfin, il convient de souligner que la répartition des ménages au fil du temps discutée ici ne joue aucun rôle dans le calcul des probabilités d’inclusion et des poids de sondage, qui jouent pourtant un rôle essentiel dans le calibrage. La période d’observation ne sera pas non plus reprise comme variable dans le calibrage (afin d’atteindre éventuellement un équilibre entre les périodes dans l’échantillon calibré, via des variables de contraste).

2.2.5 Probabilités de sélection et poids de sondage

Les principes de détermination des probabilités de sélection et des poids de sondage (corrigés) pour les répondants d’EBM 2020 - éventuellement conditionnels aux étapes successives allant de la sélection du RN à l’inclusion dans l’échantillon EMB 2020 - sont les mêmes que pour les répondants d’EBM 2018. C’est pourquoi nous vous renvoyons au paragraphe 2.2.5 d’*Enquête sur le budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b)) pour plus de détails à ce sujet.

⁴ Nous entendons par ‘ménages d’une semaine (déterminée)’ les ménages dont la semaine de référence EFT est cette semaine (déterminée).

⁵ Les pourcentages et fractions de ce paragraphe et au Tableau 9 sont approximatifs car les *nombres* par semaine (p.ex. 19) doivent être répartis en *nombres* sur deux demi-mois (p.ex. 5 et 14, ce qui se rapproche de la répartition 25-75%).

2.3 Recrutement via l'échantillon EBM 2018 – Source 2

2.3.1 Introduction

L'échantillon de répondants d'EBM 2018 est une deuxième source pour obtenir les répondants d'EBM 2020: les répondants à EBM 2018, à l'exclusion des ménages âgés et des ménages qui avaient déjà participé à EBM 2016, ont été invités à participer à nouveau à EBM 2020.

L'échantillon EBM 2018 a vu le jour de manière très complexe, à partir de différentes sources, comme expliqué en détail dans *Enquête sur le budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b)) au paragraphe 2. Au Tableau 10 ci-après, nous présentons de manière schématique la création du sous-échantillon d'EBM 2020 via EBM 2018, qui est donc la deuxième source pour EBM 2020. La complexité de la composition de l'échantillon d'EBM 2018 se reflète dans le Tableau 10, qui montre clairement comment la source 2 d'EBM 2020 est répartie sur quatre sous-échantillons EBM 2020 trimestriels, et aide à rédiger les formules des poids de sondage (approximatifs).

2.3.2 Sous-échantillons trimestriels d'EBM 2020 de la source 2

Étant donné que les formules des poids de sondage approximatifs, qui sont présentées ci-après aux sous-paragraphe 2.3.2.1 à 2.3.2.4, sont analogues aux formules indiquées dans *Enquête sur le budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b)), nous ne présenterons pas ici de dérivation détaillée de ces formules. Toutefois, à des fins de clarification, les définitions suivantes s'appliquent à certaines probabilités de réponse:

p_{nr} = probabilité de réponse à EFT, c'est-à-dire la probabilité qu'un ménage sélectionné pour participer à EFT y participe

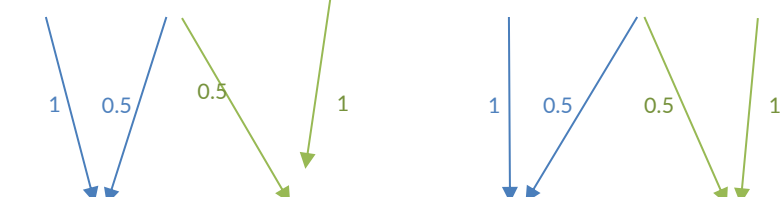
$p_{t,LFS}$ = probabilité qu'un ménage qui a participé à EBM 2018 Tt, après avoir précédemment participé à EFT, participe à EBM 2020 Tt (t= 1,...,4)

$p_{t,RR}$ = probabilité qu'un ménage qui a participé à EBM 2018 Tt, après sélection dans un échantillon supplémentaire pour EBM 2018, participe à EBM 2020 Tt (t= 1,...,4)

Les probabilités de réponse p_{nr} ne sont pas mentionnées dans le Tableau 10. Cependant, elles sont utilisées dans les formules ci-dessous pour tenir compte du fait que seuls les ménages *répondants* des échantillons EFT (les GR) sont invités à participer à EBM. Ceci peut être considéré comme une étape supplémentaire juste avant l'étape 3, pour les ménages des échantillons EFT GR1 à GR7. Il convient de noter que ces probabilités p_{nr} , tout comme les probabilités $p_{t,LFS}$ et $p_{t,RR}$, peuvent être 'modélisées'. Les probabilités p_{nr} ont effectivement été estimées au moyen d'un *mixed effects logistic regression model* (voir paragraphe 2.2.5 de *Enquête sur le budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b)) pour un problème similaire d'estimation des probabilités de réponse à EFT). Comme pour EBM 2018, les probabilités $p_{t,LFS}$ et $p_{t,RR}$ sont les probabilités de réponse observées par province.

Enfin, nous remarquons que les 'probabilités de sélection conditionnelles' 1 ou 0,5, comme mentionnées au degré 3 du Tableau 10, indiquent que 100% ou 50% respectivement d'un GR (éventuellement la fusion des GR3 et GR4) ou d'un échantillon bimensuel correspondant (Jan-Fév, ...) ont été attribuées à un échantillon trimestriel EBM 2018.

Tableau 10: Aperçu schématique de la création du sous-échantillon d'EBM 2020 via la source 2

		Commentaires
	RN	Plusieurs extractions du RN utilisées (*)
Degrés 1 & 2	$\frac{n_{rel}}{N}$ $\frac{\tilde{n}_{rel}}{\tilde{N}}$	Probabilités de sélection: <ul style="list-style-type: none"> • n_{rel}/N Pour les ménages d'EBM 2018 source 1 • $\tilde{n}_{rel}/\tilde{N}$ Pour les ménages d'EBM 2018 source 3.1 (varie également selon le GR, en raison des différentes extractions du RN)
	GR1 GR2 GR3 & GR4 GR5 GR6 GR7 Jan-Fév Mar-Avr Mai-Juin Jui-Août Sep-Oct Nov-Déc	GR 1-7 ~ 3 échantillons trimestriels EFT (EBM 2018 source 1) (**) 6 échantillons complémentaires (EBM 2018 source 3.1) (**) Probabilités de sélection conditionnelles
Degré 3		
	EBM 2018 T1 EBM 2018 T2 EBM 2018 T3 EBM 2018 T4	
Degré 4	$p_{1,LFS}$ $p_{1,RR}$ $p_{2,LFS}$ $p_{2,RR}$ $p_{3,LFS}$ $p_{3,RR}$ $p_{4,LFS}$ $p_{4,RR}$	Probabilités d'(auto)sélection conditionnelles ~ Probabilités de réponse par source <ul style="list-style-type: none"> • varie éventuellement selon le trimestre • peuvent être 'modélisées'
	EBM 2020 T1 EBM 2020 T2 EBM 2020 T3 EBM 2020 T4	Echantillons trimestriels de répondants EBM 2020
(*) 9 extractions du RN ont été effectuées: 3 pour les échantillons trimestriels EFT: EFT 2016 T3 (qui se compose des GR1, GR2 et GR3), EFT 2016 T4 (qui se compose des GR4, GR5 et GR6) et EFT 2017 T1 (qui se compose uniquement du GR7). Et 6 pour les échantillons bimensuels complémentaires EBM 2018 Jan-Fév à Nov-Déc. On obtient ainsi les chiffres de population N_1, \dots, N_3 en $\tilde{N}_1, \dots, \tilde{N}_6$, respectivement, utilisés dans les formules des paragraphes 2.3.2.1 à 2.3.2.4. (**) La numérotation des GR et des échantillons bimensuels supplémentaires impliquent les indices de n_{rel} et \tilde{n}_{rel} , respectivement, dans les formules ci-dessous.		

2.3.2.1 Le sous-échantillon d'EBM 2020 T1 de la source 2

L'échantillon EBM 2020 T1 est constitué via 4 canaux, soit du RN via EFT, soit du RN via EBM 2018. Chaque canal peut être vu comme un échantillonnage à plusieurs degrés, ce qui conduit à une formule (approximative) pour un poids de sondage des ménages d'EBM 2020 T1 recrutés via EBM 2018. En bref:

- RN → GR1 (partie d'EFT 2016 T3) → EBM 2018 T1 → EBM 2020 T1
 - poids initial à calibrer $(N_1/n_{rel,1})/p_{nr}/1/p_{1,LFS}$
- RN → EBM 2018 Jan-Fév → EBM 2018 T1 → EBM 2020 T1
 - poids initial à calibrer $(\tilde{N}_1/\tilde{n}_{rel,1})/1/p_{1,RR}$
- RN → GR2 (partie d'EFT 2016 T3) → EBM 2018 T1 → EBM 2020 T1
 - poids initial à calibrer $(N_1/n_{rel,2})/p_{nr}/0,5/p_{1,LFS}$
- RN → EBM 2018 Mar-Avr → EBM 2018 T1 → EBM 2020 T1
 - poids initial à calibrer $(\tilde{N}_2/\tilde{n}_{rel,2})/0,5/p_{1,RR}$

2.3.2.2 Le sous-échantillon d'EBM 2020 T2 de la source 2

Analogue pour les ménages d'EBM 2020 T2 recrutés via EBM 2018, mais en tenant compte maintenant de 6 canaux (parce que nous devons traiter les GR3 et GR4 séparément ici):

- RN → GR2 (partie d'EFT 2016 T3) → EBM 2018 T2 → EBM 2020 T2
 - poids initial à calibrer $(N_1/n_{rel,2})/p_{nr}/0,5/p_{2,LFS}$
- RN → EBM 2018 Mar-Avr → EBM 2018 T2 → EBM 2020 T2
 - poids initial à calibrer $(\tilde{N}_2/\tilde{n}_{rel,2})/0,5/p_{2,RR}$
- RN → GR3 (partie d'EFT 2016 T3) → EBM 2018 T2 → EBM 2020 T2
 - poids initial à calibrer $(N_1/n_{rel,3})/p_{nr}/1/p_{2,LFS}$
- RN → EBM 2018 Mai-Juin → EBM 2018 T2 → EBM 2020 T2
 - poids initial à calibrer $(\tilde{N}_3/\tilde{n}_{rel,3})/1/p_{2,RR}$
- RN → GR4 (partie d'EFT 2016 T4) → EBM 2018 T2 → EBM 2020 T2
 - poids initial à calibrer $(N_2/n_{rel,4})/p_{nr}/1/p_{2,LFS}$
- RN → EBM 2018 Mai-Juin → EBM 2018 T2 → EBM 2020 T2
 - poids initial à calibrer $(\tilde{N}_3/\tilde{n}_{rel,4})/1/p_{2,RR}$

2.3.2.3 Le sous-échantillon d'EBM 2020 T3 de la source 2

Analogue pour les ménages de EBM 2020 T3 recrutés via EBM 2018:

- RN → GR5 (partie d'EFT 2016 T4) → EBM 2018 T3 → EBM 2020 T3
 - poids initial à calibrer $(N_2/n_{rel,5})/p_{nr}/1/p_{3,LFS}$
- RN → EBM 2018 Juil-Août → EBM 2018 T3 → EBM 2020 T3
 - poids initial à calibrer $(\tilde{N}_4/\tilde{n}_{rel,5})/1/p_{3,RR}$
- RN → GR6 (partie d'EFT 2016 T4) → EBM 2018 T3 → EBM 2020 T3
 - poids initial à calibrer $(N_2/n_{rel,6})/p_{nr}/0,5/p_{3,LFS}$
- RN → EBM 2018 Sep-Oct → EBM 2018 T3 → EBM 2020 T3
 - poids initial à calibrer $(\tilde{N}_5/\tilde{n}_{rel,6})/0,5/p_{3,RR}$

2.3.2.4 Le sous-échantillon d'EBM 2020 T4 de la source 2

Analogue pour les ménages de EBM 2020 T4 recrutés via EBM 2018:

- RN → GR6 (partie d'EFT 2016 T4) → EBM 2018 T4 → EBM 2020 T4
 - poids initial à calibrer $(N_2/n_{rel,6})/p_{nr}/0,5/p_{4,LFS}$
- RN → EBM 2018 Sep-Oct → EBM 2018 T4 → EBM 2020 T4
 - poids initial à calibrer $(\tilde{N}_5/\tilde{n}_{rel,6})/0,5/p_{4,RR}$
- RN → GR7 (= EFT 2017 T1) → EBM 2018 T4 → EBM 2020 T4
 - poids initial à calibrer $(N_3/n_{rel,7})/p_{nr}/1/p_{4,LFS}$
- RN → EBM 2018 Nov-Déc → EBM 2018 T4 → EBM 2020 T4
 - poids initial à calibrer $(\tilde{N}_6/\tilde{n}_{rel,7})/1/p_{4,RR}$

2.3.3 Répartition des ménages d'EBM 2020 sur 24 demi-mois

Dans cette partie de l'échantillon, nous écrivons à nouveau aux ménages qui ont participé à l'enquête EBM en 2018 pour leur demander s'ils sont intéressés à participer également à l'enquête EBM 2020. Nous nous limitons toutefois aux ménages qui ont participé à EBM pour la première fois en 2018. Ainsi, 4478 ménages ont reçu une lettre. La lettre d'invitation aux ménages mentionne pendant quelle période ou 'demi-mois' il doivent remplir les carnets d'EBM 2020. Nous avons choisi de conserver la période à travers les vagues. Les ménages qui ont participé à une certaine période en 2018 sont sollicités pour la même période en 2020.

2.4 Tirages complémentaires du RN – Source 3

2.4.1 Introduction

Tout comme le recrutement des groupes de rotation de l'EFT (source 1 – voir Tableau 3, colonnes 'panel EFT' et 'EBM 2020', ainsi que le paragraphe 2.2.1) a donné en premier lieu 4 échantillons trimestriels d'EBM 2020, qui ont ensuite été répartis entre 6 échantillons bimensuels d'EBM 2020, un échantillon complémentaire a été tiré directement du RN tous les trois mois. Il était ensuite réparti entre 6 échantillons bimensuels complémentaires d'EBM 2020 (voir Tableau 3, colonnes 'RN' et 'EBM 2020').

Les 4 échantillons trimestriels complémentaires (voir colonne [1] du Tableau 11 ont été ajustés au recrutement issu d'EFT. Ceci est expliqué au Tableau 11.

Tableau 11: Les 4 échantillons trimestriels complémentaires d'EBM 2020: informations de base

Échantillon complémentaire EBM	GR EFT sous-jacents	Nombre d'UPE (hors DEU)	Nombre d'UPE vides	Nombre de MN éligibles
[1]	[2]	[3]	[4]	[5]
T1	GR8 & GR9	493	16	884.814
T2	GR10 & GR11	510	0	1.124.983
T3	GR12 & GR13	507	0	1.141.902
T4	GR14 & GR15	550	49	1.047.465

- La colonne [2] indique quel GR EFT a servi de point de départ pour chaque échantillon trimestriel EBM.
- La colonne [3] indique combien d'UPE différentes, à l'exclusion des UPE des communes germanophones, sont tirées pour former les GR pour l'EFT.
- La colonne [4] indique combien d'UPE, mentionnées à la colonne [3], ne contiennent pas les ménages privés éligibles ou sont 'vides'. Pour la définition d'un 'ménage éligible', nous vous renvoyons au paragraphe 2.4.3 de *Enquête sur le budget des ménages 2018 - Note méthodologique* (voir Statbel (2019a) ou Statbel (2019b)).
- La colonne [5] donne le nombre de ménages éligibles dans les UPE non vides (et non germanophones).

Les tirages trimestriels complémentaires de ménages du RN, qui ont été abordés dans ce paragraphe, peuvent tous être considérés comme un échantillonnage à deux degrés avec la sélection des UPE au degré 1, suivie par la sélection au degré 2 des ménages parmi les UPE sélectionnées au degré 1.

2.4.2 Degré 1 : tirage des unités géographiques (ou UPE).

Le tirage des UPE n'est pas un nouveau tirage: les UPE avaient déjà été tirées pour l'EFT (GR8 à GR15) comme expliqué au paragraphe 2.2.2.

2.4.3 Degré 2 : tirage des ménages EBM (USE)

L'objectif des tirages complémentaires directs du RN est quadruple:

- Comme pour EBM 2018, nous voulions encore atteindre le nombre de 400 'ménages âgés'. Étant donné que nous nous attendions à ce que, pour l'EBM 2020, un plus grand nombre de ménages

de type EFT que pour EBM 2018 puisse être atteint grâce aux sources 1 et 2, nous avons initialement visé (seulement) 500 ménages de type EFT dans EBM 2020.

- Nous voulions répartir ces deux nombres équitablement entre les 4 échantillons trimestriels.

De plus, nous voulions répartir équitablement les 4 échantillons supplémentaires entre les UPE concernées. Pour atteindre ces objectifs, un tirage aléatoire stratifié a été réalisé. Les strates d'échantillonnage sont les cellules au croisement de deux variables:

- Une variable qui détermine pour chaque ménage à quelle UPE il appartient.
- Une variable qui détermine pour chaque ménage s'il est de type EFT.

Compte tenu du nombre de ménages souhaité (500 de type EFT et 400 ménages âgés, sur base annuelle) et de la stratification décrite ci-dessus, la taille initiale de l'échantillon et l'allocation, c'est-à-dire la répartition de l'échantillon entre les strates, ont ensuite été calculées par échantillon trimestriel. À cette fin, les probabilités de réponses attendues (au niveau des ménages) devaient encore être déterminées. Le Tableau 12 montre, aux colonnes [2], [3a] et [3b], les paramètres qui varient selon les échantillons EBM trimestriels et le type de ménage:

- La colonne [2] montre le nombre de groupes EFT dans les UPE non vides par trimestre (à l'exclusion des UPE ou groupes EFT des communes germanophones). Il convient de noter que si une UPE est sélectionnée dans les deux GR associés à un trimestre, cette UPE est comptée deux fois dans la détermination du nombre de groupes EFT. La raison en est que nous ne voulons pas sélectionner un nombre fixe de ménages de type EFT et un nombre fixe de ménages âgés en plus par UPE, mais par groupe EFT.
- La ventilation des colonnes [3a] et [3b], ainsi que des colonnes [4a] et [4b] ainsi que [5a] et [5b] correspond aux deux catégories pour la deuxième variable de stratification (ménages de type EFT versus ménages âgés). Les probabilités de réponse attendues dans les colonnes [3a] et [3b] pour les T1 et T2 sont basées sur les probabilités de réponse obtenues dans EBM 2018. Les probabilités de réponse attendues ont été revues à la baisse pour les T3 et T4 en raison de la crainte que la pandémie de COVID-19 ne fasse fortement chuter les probabilités de réponse.

Tableau 12: Les 4 échantillons trimestriels complémentaires d'EBM 2020: paramètres d'échantillon et distribution

Échantillon complément aire EBM	Nombre GRP	Probabilité de réponse (Niveau du MN)		Distribution par GRP		Distribution totale (réalisée)	
		Type EFT	Âgé	Type EFT	Âgé	Type EFT	Âgé
[1]	[2]	[3a]	[3b]	[4a]	[4b]	[5a]	[5b]
T1	534	5%	3,5%	5	5	2.666	2.650
T2	550	5%	3,5%	5	5	2.750	2.749
T3	550	4,5%	1,5%	12	12	6.600	6.427
T4	483	4,5%	1,5%	14	14	6.718	6.246
Total	2.117	-	-	-	-	18.734	18.072

L'allocation par GRP des colonnes [4a] et [4b] a été calculée selon les mêmes principes que pour l'enquête EBM 2018. De plus amples informations à ce sujet sont disponibles au paragraphe 2.4.3 de *Enquête sur le budget des ménages 2018 - Note méthodologique* (voir Statbel (2019a) ou Statbel (2019b)).

La définition des strates d'échantillonnage et le calcul du nombre de ménages à sélectionner par strate complètent le plan de sondage pour les tirages supplémentaires du RN, et permettent de tirer l'échantillon. Le tirage au sein de chaque strate est aléatoire.

En principe, nous pouvons calculer les nombres des colonnes [5a] et [5b] à partir des colonnes [2], [4a] et [4b], par exemple pour T2: $550 \times 5 = 2.750$ dans la colonne [5a]. Certains des nombres des colonnes

[5a] et [5b] ne sont pas exactement des produits de ce type. Pour le T1, $534 \times 5 = 2.670 \neq 2.666$ s'applique, car dans ce cas, une des UPE ne semblait contenir que 1 et non au moins 5 ménages de type EFT éligibles. C'est la raison pour laquelle nous avons repris aux colonnes [5a] et [5b] non pas les tailles d'échantillon calculées mais les tailles d'échantillon réalisées. Le Tableau 13 suivant clôture ce paragraphe.

Tableau 13: Les 4 échantillons trimestriels complémentaires d'EBM 2020: chiffres de la population (limités aux UPE tirées dans les GR8-15)

Échantillon complémentaire EBM	Nombre de MN éligibles		
	Type EFT	Âgé	Total
[1]	[2a]	[2b]	[3]
T1	810.120	74.694	884.814
T2	1.027.840	97.143	1.124.983
T3	1.040.795	101.107	1.141.902
T4	954.628	92.837	1.047.465

Les chiffres de la population de la colonne [3] sont repris de la colonne [5] dans le Tableau 11. Les colonnes [2a] et [2b] montrent la ventilation des chiffres de la colonne [3] selon le type de ménage.

2.4.4 Probabilités de sélection et poids de sondage

2.4.4.1 Pour les ménages de type EFT

Les principes du calcul des probabilités de sélection et poids de sondage ont été expliqués au paragraphe 2.4.4.1 de *Enquête sur le budget des ménages 2018 - Note méthodologique* (voir Statbel (2019a) ou Statbel (2019b)).

Le Tableau 14 montre les poids de sondage $\tilde{N}/\tilde{n}_{rel}$ (approximatifs) pour les ménages de type EFT.

Tableau 14: Poids de sondage dans les échantillons complémentaires EBM 2020 de ménages de type EFT

Strate UPE		Échantillons trimestriels			
		Trim 1	Trim 2	Trim 3	Trim 4
		GR8-9	GR10-11	GR12-13	GR14-15
1	ANT	2.326,24	2.336,20	975,89	941,49
2	BRU	1.322,88	1.332,31	557,27	520,46
3	WVL	1.740,25	1.745,94	729,32	688,05
4	OVL	2.220,24	2.045,63	854,46	864,49
5	HAI	2.461,97	1.939,11	809,74	807,25
6	LIE	1.806,53	1.811,74	756,16	728,30
7	LIM	1.512,45	1.518,31	634,70	611,73
8	LUX	696,83	699,01	292,04	275,19
9	NAM	1.162,89	1.166,95	487,72	523,26
10	VLB	1.715,01	1.726,80	721,33	787,20
11	BWA	951,13	953,88	398,42	376,09

2.4.4.2 Pour les ménages âgés

Les principes du calcul des probabilités de sélection et poids de sondage ont été expliqués au paragraphe 2.4.4.2 de *Enquête sur le budget des ménages 2018 - Note méthodologique* (voir Statbel (2019a) ou Statbel (2019b)).

Le Tableau 15 montre les poids de sondage (approximatifs) $\tilde{N}/\tilde{n}_{rel}$ pour les ménages âgés.

Tableau 15: Poids de sondage dans les échantillons EBM 2020 complémentaires de ménages âgés

Strate UPE		Échantillons trimestriels			
		Trim 1	Trim 2	Trim 3	Trim 4
		GR8-9	GR10-11	GR12-13	GR14-15
1	ANT	253,06	252,42	105,35	105,34
2	BRU	92,19	93,04	39,72	36,95
3	WVL	238,29	239,43	100,35	105,30
4	OVL	246,51	225,71	94,72	102,60
5	HAI	219,88	172,51	72,67	76,09
6	LIE	171,82	171,11	73,82	75,85
7	LIM	158,88	160,02	68,05	67,15
8	LUX	67,63	64,43	33,40	33,44
9	NAM	106,81	106,69	47,38	56,14
10	VLB	191,81	192,68	81,37	90,50
11	BWA	91,22	91,75	40,28	39,60

3 Taux de réponse

3.1 Taux de réponse initial de l'EFT - Source 1

Depuis 2016, l'EFT est organisée sous la forme d'un panel. La première interview avec le ménage est réalisée à domicile par un enquêteur (CAPI). Une deuxième interrogation est réalisée le trimestre suivant, mais à distance cette fois-ci. Une autre enquête à distance suit un an après le premier entretien. L'interview finale suit un trimestre plus tard et a également lieu à distance. Cette enquête à distance est d'abord un questionnaire web en ligne (CAWI). En cas d'absence de réponse, l'enquêteur prend contact par téléphone pour réaliser l'enquête par téléphone (CATI). C'est lors de la quatrième (et donc dernière) interview EFT que la question est posée de savoir si la personne est intéressée par une participation à l'enquête EBM.

Tableau 16: Taux de réponse initial de l'EFT - Source 1

Enquête EFT Vague 4	Enquête EBM	Nombre de MN EFT positifs W4	Nombre de MN EFT positifs	% (Cf. 2018: 27%/ 2016: 33%)
EFT2018 T3	EBM 2020 JAN	2184	494	22,62
EFT2018 T3+T4	EBM 2020 FÉV	2334	541	23,18
EFT2018 T4	EBM 2020 MAR	2660	630	23,68
EBM2019 T1	EBM 2020 AVR	2588	621	24,00
EFT2019 T1+T2	EBM 2020 MAI	2467	572	23,19
EFT2019 T2	EBM 2020 JUIN	2044	489	23,92
EFT2019 T3	EBM 2020 JUIL	2344	548	23,38
EFT2019 T3+T4	EBM 2020 AOÛT	2302	528	22,94
EFT2019 T4	EBM 2020 SEP	2285	548	23,98
EFT2020 T1	EBM 2020 OCT	2584	689	26,66
EFT2020 T1+T2	EBM 2020 NOV	2409	640	26,57
EFT2020 T2	EBM 2020 DÉC	2249	528	23,48
	Total annuel	28450	6828	24,00

La volonté générale de participer à l'enquête EBM diminue de 27% pour EBM 2018 à 24% pour EBM 2020. Le rapport méthodologique de 2018 mentionnait des différences de taux de réponse selon le mode d'interview. Les répondants qui achevaient l'enquête de manière indépendante via CAWI étaient en effet moins disposés à poursuivre leur participation que les répondants qui répondaient à l'enquête par téléphone auprès d'un enquêteur. Cet effet était aussi perceptible en 2020. C'est ce qui ressort du Tableau 17 suivant.

Tableau 17: Volonté de répondre selon le mode d'enquête

% volonté de répondre	EFT 2018		EFT 2018				EFT 2020	
	T3	T4	T1	T2	T3	T4	T1	T2
4e vague téléphonique	28	32	32	29	28	28	30	26
4e vague web	18	20	20	19	19	23	26	27

Comme la volonté de répondre des répondants EFT à l'EBM 2020 a diminué par rapport au recrutement de 2018, il a été décidé d'adapter la formulation des questions dans l'enquête EFT. Selon les enquêteurs, la formulation des questions n'était en effet pas très claire. Ce changement a été introduit à partir d'EFT 2020 T1. Les premiers chiffres montrent en tout cas une amélioration, surtout pour les répondants qui ont répondu en ligne.

Ancienne formulation des questions

Nous organisons en 2020, une autre enquête pour laquelle le ménage est payé pour participer, l'Enquête sur le budget des ménages.

L'objectif de l'enquête est de noter toutes vos dépenses dans un carnet de ménage pendant un demi-mois durant le mois ^pMonth_N. Cela peut se faire sur papier ou sur internet. Ils reçoivent 45 euros pour leur participation avec un bonus éventuel de 5 euros s'ils remplissent le carnet entièrement sur internet.

Si vous êtes intéressé(e), un enquêteur vous rendra visite dans le courant du mois ^pMonth_visit_N afin de vous expliquer l'enquête en détails et vous remettre les documents nécessaires Etes-vous intéressé(e)?

- Oui. 1

- Non. 2

Nouvelle formulation des questions

Ceci était votre dernière participation à l'Enquête sur les forces de travail.

Nous vous proposons, pour un montant de 50 euros (45 euros si vous participez sur papier), de participer à une enquête sur le budget des ménages. Cette enquête sur le budget des ménages est importante, entre autres, pour le calcul de l'indice des prix à la consommation qui est utilisé pour adapter les revenus et les allocations au coût de la vie.

Pour ce faire, vous devez noter toutes vos dépenses pendant deux semaines et compléter un questionnaire sur vos principales dépenses. Pour vous, ces deux semaines tombent pendant le mois de ^pMonth_N ^Year_N. Avant le début de cette période, un enquêteur vous contactera afin de vous donner toutes les explications.

Souhaitez-vous nous aider en participant à cette enquête sur le budget des ménages?

- Oui. 1

- Non. 2

3.2 [Taux de réponse initial de l'EBM - Source 2](#)

Les ménages d'EBM 2020 ont été répartis sur les 24 périodes de référence d'EBM 2020.

Tableau 18: Taux de réponse initial de l'EBM - Source 2

Enquête EBM	Envoi				Résultats		
	Date 1 ^{er} envoi	Nombre de lettres	Date rappel	Nombre de rappels	Nombre négatifs	Nombre de MN positifs	% oui / total
Trimestre 1	03/10/2019	1242	22/10/2019	670	220	590	47,5
Trimestre 2	13/01/2020	1127	05/02/2020	648	225	521	46,2
Trimestre 3	01/04/2020	936	23/4/2020	514	149	470	50,2
Trimestre 4	01/07/2020	1173	22/7/2020	511	186	512	43,7
TOTAL		4478		2343	780	2093	46,7

Chaque trimestre, un courrier a été adressé aux ménages des six périodes de référence afin de les inviter à participer à l'EBM. Il leur était demandé de renvoyer un formulaire sur lequel ils pouvaient indiquer s'ils étaient intéressés ou non par une participation. S'ils étaient intéressés, ils pouvaient indiquer leurs coordonnées.

3.3 Taux de réponse initial tirage complémentaire du RN - Source 3

Le SM fournissait un échantillon tous les deux mois. Le CEC était ensuite chargé d'attribuer la période de référence aux ménages et de calculer un numéro de ménage. Chaque mois, un courrier était adressé aux ménages des deux périodes de référence afin de les inviter à participer à l'EBM. Il leur était demandé de renvoyer un formulaire sur lequel ils pouvaient indiquer s'ils étaient intéressés ou non par une participation. S'ils étaient intéressés, ils pouvaient indiquer leurs coordonnées.

- Nombre de ménages par groupe EFT et par UPE: 5 de la population EFT + 5 de la population âgée
- A partir de l'échantillon de juillet: nombre par groupe: 12 de la population EFT + 12 de la population âgée
- A partir de l'échantillon d'octobre: nombre par groupe: 14 de la population EFT + 14 de la population âgée

Remarque

Ces chiffres sont des objectifs: chaque tirage s'est avéré avoir des UPE dans lesquelles il n'y avait pas assez de ménages pour atteindre ces chiffres.

Tableau 19: Taux de réponse initial du tirage complémentaire du RN - Source 3

Mois EBM	Nombre total de lettres	Nombre de rappels	Nombre de réactions négatives	Nombre de réactions positives	% oui / total
Janvier	1643	1469	314	84	5,11
Février	1839	1616	288	92	5,00
Mars	1833	1657	287	84	4,58
Avril	1899	1679	341	86	4,53
Mai	1910	1707	328	101	5,28
Juin	1690	1504	267	88	5,21
Juillet	4502	4007	404	239	5,30
Août	4512	4038	749	279	6,18
Septembre	4006	3619	445	204	5,09
Octobre	5270	4731	460	291	5,52
Novembre	4372	3939	380	231	5,28
Décembre	3312	2980	390	171	5,16
TOTAL	36788	32946	4653	1950	5,19

La décision d'augmenter l'échantillon à partir de juillet 2020 s'explique par la pandémie de coronavirus. L'impact de la pandémie de coronavirus sur le taux de réponse à EBM a été difficile à évaluer. C'est pourquoi on a choisi de créer une réserve afin de disposer de suffisamment de ménages dans l'échantillon final de 2020.

3.4 Echantillon initial total EBM 2020 par mois

Chaque mois, le CEC établit un échantillon, composé des ménages positifs de l'EFT, des nouveaux ménages du RN et des ménages du panel de l'enquête EBM 2018 qui ont accepté de participer. Cet échantillon couvre alors les deux périodes de référence de ce mois.

Tableau 20: Echantillon initial total EBM 2020 par mois

Enquête EBM	Nombre de MN EFT positifs	Nombre de MN EFT positifs 2018	Nombre de nouveaux MN positifs (RN)	Total
EBM 2020 JAN	494	168	84	746
EBM 2020 FÉV	541	200	92	833
EBM 2020 MAR	630	213	84	927
EBM 2020 AVR	621	165	86	872
EBM 2020 MAI	572	172	101	845
EBM 2020 JUIN	489	186	88	763
EBM 2020 JUIL	548	140	239	927
EBM 2020 AOÛT	528	148	279	955
EBM 2020 SEP	548	175	204	927
EBM 2020 OCT	689	175	291	1155
EBM 2020 NOV	640	174	231	1045
EBM 2020 DÉC	528	183	171	882
Total	6828	2099	1950	10877

Lorsque l'échantillon mensuel est prêt, le CEC crée un dossier pour chaque enquêteur reprenant ses listes d'adresses et le nombre de carnets nécessaire.

3.5 Aperçu du travail de terrain et du taux de réponse

Tableau 21: Aperçu du travail de terrain et du taux de réponse pour la Belgique et les régions pour EBM 2020

Partie travail de terrain / taux de réponse	D'un point de vue géographique	EFT	EBM 2018	Nouvelles MN de type EFT	Nouvelles MN âgés	Total
Recrutement						
Potentiel	Belgique	28.450	4.490	18.725	18.063	69.728
	Flandre	13.937	2.247	9.103	8.913	34.200
	Wallonie	10.855	1.692	6.913	6.463	25.923
	Bruxelles	3.658	551	2.709	2.687	9.605
Recrutés	Belgique	6.828	2.098	1.416	534	10.876
	Flandre	2.795	1084	763	266	4.908
	Wallonie	2.796	806	486	198	4.286
	Bruxelles	1237	208	167	70	1.682
Taux de réponse recrutement = recrutés/potentiels	Belgique	24,00%	46,73%	7,56%	2,96%	15,60%
	Flandre	20,05%	48,24%	8,38%	2,98%	14,35%
	Wallonie	25,76%	47,64%	7,03%	3,06%	16,53%
	Bruxelles	33,82%	37,75%	6,16%	2,61%	17,51%
Travail de terrain (1)						
Non réalisé	Belgique	14	5	0	0	19
	Flandre	0	0	0	0	0
	Wallonie	14	5	0	0	19
	Bruxelles	0	0	0	0	0
<p>Potentiels: nombre de ménages qui ont été invités à participer à l'EBM 2020.</p> <ul style="list-style-type: none"> • Pour la partie de l'échantillon qui provient de l'EFT, il s'agit de tous les ménages qui ont effectivement participé à l'EFT. • Pour la partie de l'EBM 2018, il s'agit de tous les ménages qui ont participé à l'EBM en 2018, à l'exclusion de ceux issus de l'EBM 2016. • Pour la partie du nouvel échantillon (tant du type EFT que les ménages âgés), il s'agit de tous les ménages qui ont été tirés du RN. <p>Recrutés: nombre de ménages qui ont confirmé leur intérêt à participer à l'EBM 2020. Ils ont réagi 'positivement' à notre invitation.</p> <p>Non effectués: nombre de ménages qui n'ont pas été contactés, par manque d'enquêteur par exemple.</p>						

Partie travail de terrain / taux de réponse	D'un point de vue géographique	EFT	EBM 2018	Nouvelles MN de type EFT	Nouvelles MN âgés	Total
Travail de terrain (2)						
Refus de participer	Belgique	3713	186	327	144	4370
	Flandre	1270	90	154	71	1585
	Wallonie	1600	72	121	53	1846
	Bruxelles	843	24	52	20	939
Aucune information	Belgique	27	6	4	4	41
	Flandre	9	2	1	1	13
	Wallonie	13	2	2	2	19
	Bruxelles	5	2	1	1	9
Refusés	Belgique	135	162	26	18	341
	Flandre	54	80	11	10	155
	Wallonie	57	59	12	4	132
	Bruxelles	24	23	3	4	54
Entièrement finalisés	Belgique	2.939	1.739	1.059	368	6.105
	Flandre	1.462	912	597	184	3155
	Wallonie	1112	668	351	139	2270
	Bruxelles	365	159	111	45	680
Taux de réponse travail de terrain =Entièrement finalisés/Recrutés	Belgique	43,04%	82,89%	74,79%	68,91%	56,13%
	Flandre	52,31%	84,13%	78,24%	69,17%	64,28%
	Wallonie	39,77%	82,88%	72,22%	70,20%	52,96%
	Bruxelles	29,51%	76,44%	66,47%	64,29%	40,43%
Recrutement + travail de terrain						
Réponse globale =Entièrement finalisés/potentiels	Belgique	10,33%	38,73%	5,66%	2,04%	8,76%
	Flandre	10,49%	40,59%	6,56%	2,06%	9,23%
	Wallonie	10,24%	39,48%	5,08%	2,15%	8,76%
	Bruxelles	9,98%	28,86%	4,10%	1,67%	7,08%
<p>Refus de participer = nombre de ménages qui n'ont finalement pas participé.</p> <p>Aucune info: nombre de ménages pour lesquels l'enquêteur n'a pas transmis toutes les données.</p> <p>Refusés: nombre de ménages qui ont rempli un carnet et un questionnaire ménage, mais qui n'ont pas été acceptés comme répondant, parce que les contrôles démontrent que la qualité des données n'est pas satisfaisante.</p> <p>Entièrement finalisés: nombre de ménages pour lesquels la qualité des données est suffisante pour être acceptés comme répondant.</p>						

3.6 Raisons de la non-réponse

Dans la lettre d'invitation qui est envoyée aux ménages, ceux-ci ont également la possibilité de donner la raison pour laquelle ils ne veulent pas participer. Le Tableau 22 donne un aperçu de ces raisons. Si l'on analyse ces raisons, les principales raisons invoquées par les ménages âgés sont qu'ils se considèrent trop âgés pour participer ou qu'ils ont un problème de santé qui les empêche de participer ou qu'ils considèrent que leur propre situation n'est pas pertinente (par exemple parce qu'ils ne font pas les courses eux-mêmes et n'ont donc pas une vue d'ensemble de leurs dépenses). Les ménages issus de l'échantillon de type EFT et du panel EBM indiquent comme raison principale de ne pas participer le fait qu'ils n'ont pas le temps ou qu'ils ne sont pas disponibles pendant la période proposée.

Tableau 22: raison du refus de participer à l'EBM

	Panel EBM (%)	Nouvelles MN de type EFT (%)	Nouvelles MN âgés (%)
Raison du refus (du recrutement et du travail de terrain)	Pourcentage		
Pas le temps, trop occupé, pas disponible pendant la période	38,94	34,92	4,76
Déjà (trop) participé	8,61	0,87	0,31
Je suis trop âgé / malade	17,61	10,30	59,74
Raison familiale (un proche est malade / décédé, je viens de déménager, de me séparer, etc.)	11,15	2,71	5,20
Autre	2,54	2,82	2,49
Pas intéressé	11,55	19,41	8,30
Atteinte à la vie privé / trop privé	3,72	9,65	3,45
Doute quant à l'utilité de l'enquête	0,59	1,52	0,48
Ma situation n'est pas intéressante (isolé, indépendant, faibles revenus, cultive mes propres légumes, ne travaille pas, etc.)	1,57	11,17	6,55
En maison de repos	0,20	0,33	7,12
Problèmes linguistiques	0,00	3,25	0,44
Négatif envers les autorités (impôts, ils savent déjà tout, etc.)	2,35	1,52	0,52
Corona	1,17	1,52	0,66
Total	100,00	100,00	100,00
Nombre total de refus reçus	511	922	2290

Remarque

Il est frappant de constater que très peu de personnes font des remarques liées au coronavirus.

4 Collecte et traitement des données pour le calibrage

4.1 Collecte des données

Le travail de terrain est mené par des enquêteurs, qui ont trois contacts avec les ménages après avoir envoyé une lettre d'introduction à laquelle le ménage a répondu positivement.

1. Lors de la première visite, l'enquêteur se rend au domicile du ménage pour remettre tous les documents et fournir des explications.
2. Le deuxième contact se déroule soit de manière physique, si les ménages remplissent leur carnet sur papier (PPSI) ou s'ils en manifestent le besoin, soit par téléphone si les ménages remplissent leur carnet en ligne. Il existe en effet depuis 2012 une application numérique (CAWI) pour le carnet des ménagesⁱⁱ.
3. La troisième visite a lieu après la période de référence pendant laquelle les ménages ont consigné leurs dépenses. L'enquêteur réalise alors une interview (CAPI depuis 2014). Dans des circonstances exceptionnelles, ce troisième contact peut également se dérouler par téléphone (CATI).

Après la période de référence, les enquêteurs envoient les carnets des ménages collectés par la poste aux centres de traitement. Les collaborateurs des centres de traitement enregistrent les carnets reçus et procèdent à un contrôle de qualité visuel. Si les carnets sont de qualité suffisante, ils vérifient s'ils ont reçu le questionnaire ménage. Si ces deux éléments sont en ordre, il est procédé au paiement du ménage. Le paiement des enquêteurs est également effectué à ce moment. Les enquêteurs sont payés par visite et peuvent donc également être (partiellement) payés pour des enquêtes qui ont finalement été interrompues.

Un [arrêté royal du 17 janvier 1996](#) fixe le montant de l'indemnité aux ménages. Les ménages reçoivent pour leur participation un montant de 45 euros, ainsi qu'éventuellement un bonus de 5 euros s'ils ont complété le carnet entièrement en ligne. Un [arrêté royal du 17 mai 2006](#) fixe le montant de l'indemnité aux enquêteurs. Pour une enquête entièrement terminée, 4 visites sont rémunérées. Comme la dernière visite est assez longue parce qu'il faut remplir le questionnaire ménage et vérifier le carnet, cette visite sera facturée deux fois. L'enquêteur reçoit également une prime par questionnaire ménage complété pour avoir posé les questions supplémentaires sur l'énergie ajoutées depuis 2016.

4.2 Pandémie de coronavirus et impact sur le travail de terrain

L'épidémie de coronavirus et les mesures prises par le gouvernement ont eu un impact sur le déroulement du travail de terrain, tant pour les enquêteurs sur le terrain que pour le fonctionnement interne du CEC.

Le plus grand changement pour les enquêteurs est le travail à distance, sans contact personnel avec les personnes interrogées. Le 19 mars 2020, Statbel a envoyé l'instruction aux enquêteurs EBM que le contact en face à face avec les ménages était désormais interdit. Les enquêteurs devaient:

- Prendre contact par téléphone ou par e-mail pour se présenter.
- Déposer les carnets dans la boîte aux lettres (ou les envoyer par la poste) et donner les explications par téléphone.
- Encourager les ménages à utiliser l'application web pour remplir les carnets au lieu des carnets papier.
- Interroger les ménages par téléphone uniquement. Pour ce faire, Statbel a mis à disposition une liste à remplir qui peut être partagée avec les ménages. Cela permet aux ménages de se préparer à l'entretien afin qu'il se déroule le mieux possible par téléphone.
- Le contact avec Statbel peut se faire autant que possible par e-mail (par exemple, en renvoyant les bordereaux par e-mail plutôt que par la poste).

4.3 Enregistrement des données

Tous les membres du ménage doivent consigner leurs dépenses dans un carnet des ménagesⁱⁱⁱ pendant une période de référence donnée. Cette période de référence est définie à l'avance. En 2012, 2014 et 2016, cette période de référence était d'un mois calendrier. Depuis 2018, la période de référence a été ramenée à 15 jours, soit du 1^{er} au 15 du mois, soit du 16 à la fin du mois. A la fin de la période de référence, la personne de référence du ménage est interrogée au moyen d'un questionnaire ménage.^{iv}

4.3.1 Questionnaire ménage

Le questionnaire ménage aborde les sujets suivants^v:

1. Composition du ménage
2. Questions sur l'habitation (résidence principale et seconde résidence) (depuis 2016: y compris le module énergie)
3. Dépenses périodiques (énergie et logement)
4. Équipement du logement
5. Achat de biens durables
6. Assurances
7. Revenus
8. Avantages en nature
9. Relations entre les membres du ménage
10. Absences

Les dépenses qui sont demandées dans le questionnaire ménage sont les dépenses périodiques et couvrent les dépenses effectuées au cours des trois mois qui précèdent la période de référence et le mois de la période de référence, sauf indication contraire. Selon les feedbacks des enquêteurs, le questionnaire ménage prend en moyenne 40 à 50 minutes à remplir.

4.3.2 Carnet de dépenses

Pendant la période de référence, les ménages enregistrent toutes leurs dépenses dans le carnet. Cela peut se faire dans l'application en ligne ou sur un carnet papier. Les encodeurs des centres de traitement régionaux introduisent les carnets papier dans l'application en ligne.

Les dépenses supportées par les ménages aux fins de l'acquisition de biens et de services sont consignées au prix réellement payé, y compris les impôts indirects (TVA et accises) à charge de l'acheteur.

Le carnet consiste en un enregistrement quotidien de toutes les dépenses ordinaires et un enregistrement (limité) des produits issus de la propre culture ou de la propre entreprise du ménage. Une description détaillée de chaque produit est demandée dans les différentes colonnes:

1. Description du produit, la plus détaillée possible;
2. Les montants dépensés au moment de l'achat;
3. Le solde qui doit encore être payé ultérieurement;
4. La quantité;
5. L'unité (litre, kilogramme, ou pièces);
6. Le lieu de l'achat;
7. Le pays d'achat;
8. Les dépenses mixtes pour indépendants.

L'application en ligne présente l'avantage que le produit acheté peut être placé immédiatement dans la catégorie correspondante de la nomenclature ECOICOP. Cela se fait au moyen d'une liste de termes établie au préalable dans laquelle chaque article est lié à un code de la nomenclature. Un autre avantage

de l'application en ligne est qu'un certain nombre de contrôles de qualité peuvent être effectués immédiatement. La première étape consiste à vérifier si une quantité a été encodée. Ensuite, il y a aussi un contrôle des prix des produits. Il est vérifié si le prix indiqué est plausible par rapport à la quantité indiquée. Ce contrôle est effectué sur la base de fourchettes de prix déterminées à l'avance.

4.3.3 Fourchettes de prix

Des fourchettes de prix sont utilisées dans le programme d'enregistrement en ligne des dépenses. Chaque catégorie de la nomenclature ECOICOP se voit attribuer un prix minimum et un prix maximum par unité achetée. Si les ménages enregistrent une dépense dont le prix d'achat se situe en dehors des fourchettes de prix, ils reçoivent un avertissement. L'avertissement consiste en un triangle rouge avec un point d'exclamation qui apparaît à côté de la ligne de la dépense.

Avant le début de chaque édition de l'EBM, les fourchettes de prix sont déterminées en utilisant les enregistrements de l'édition précédente. Pour chaque catégorie de la ECOICOP, les centiles des dépenses par unité sont calculés. La limite inférieure est C5, la limite supérieure C95. Les centiles sont ensuite arrondis à 1 eurocent. Un contrôle manuel des fourchettes de prix obtenues est ensuite effectué. Afin d'effectuer ce contrôle, la qualité des anciennes fourchettes de prix de l'édition précédente est également vérifiée. À cette fin, on examine combien de fois un produit a été enregistré lors de l'édition précédente en dehors des fourchettes de prix (et n'a pas été corrigé), à la fois en nombres absolus et en termes relatifs (pourcentage).

Le contrôle manuel s'effectue dans un tableau Excel. Toutes les fourchettes de prix sont vérifiées et peuvent être ajustées, mais certaines catégories sont marquées de manière à être certainement contrôlées. Le marquage indique que:

- Plus de 15% des enregistrements dans cette catégorie de la ECOICOP se situaient en dehors des anciennes fourchettes de prix.
- Plus de 10% des enregistrements dans cette catégorie de l'ECOICOP se situaient en dehors des anciennes fourchettes de prix, ce qui représentait plus de 250 enregistrements en chiffres absolus.
- Très peu de dépenses d'un certain produit ont été enregistrées et le calcul n'est donc pas très fiable (nombre de dépenses inférieur à 70).
- Il y a une grande différence entre les anciennes fourchettes de prix et les nouvelles (la différence est supérieure à 50%).
- Le prix minimum est inférieur à 5 cents par unité.

La personne responsable a la possibilité d'ajuster les fourchettes de prix dans la liste Excel. Pour les nouvelles catégories de l'ECOICOP, les prix seront trouvés sur Internet. Ensuite, la personne responsable détermine une unité standard ainsi qu'un prix minimum et un prix maximum.

5 Analyse partie 1: détermination de la personne de référence du ménage sur la base du revenu

5.1 Introduction

L'unité de base pour la collecte et l'analyse des données dans l'EBM est le ménage. Il importe d'identifier la personne de référence (le chef du ménage, généralement) dont les caractéristiques personnelles pourront être utilisées à des fins de classement et d'analyse des informations relatives au ménage. Les résultats sont classés et présentés en fonction de la catégorie socioéconomique, de l'emploi exercé et de la situation professionnelle, du revenu, du sexe et de l'âge de la personne de référence. La personne de référence selon EBM est la personne qui participe pour la plus grande part aux revenus du ménage. Cette définition de la personne de référence (chef de ménage) ne correspond pas nécessairement à celle du RN.

Dans le cadre des enquêtes sur les conditions sociales et le revenu, comme l'enquête EBM, le ménage^{vi} est défini comme une unité sociale domestique ou opérationnelle, en raison de l'existence d'accords mutuels, du partage des dépenses et des besoins quotidiens du ménage ou de la vie dans une résidence commune.

Il se compose d'une personne vivant seule ou d'un groupe de personnes non obligatoirement liées par un lien de parenté résidant à la même adresse et consommant collectivement certains biens ou services, c'est-à-dire partageant au moins un repas par jour ou partageant un salon ou une pièce de séjour.

Il existe également, par opposition aux ménages privés, des ménages collectifs ou ménages institutionnels, par exemple: les hôpitaux, les maisons de retraite, les établissements d'hébergement collectifs, les prisons, les casernes militaires, les institutions religieuses, les internats et les résidences mises à disposition des travailleurs par les employeurs. Ceux-ci ne sont pas interrogés dans le cadre de l'EBM.

Le revenu disponible équivalent^{vii} correspond au revenu total d'un ménage, après impôt et autres déductions, disponible en vue d'être dépensé ou épargné, divisé par le nombre de membres du ménage converti en équivalents adultes. L'équivalence entre les membres du ménage est obtenue par pondération en fonction de l'âge, à partir de l'échelle d'équivalence modifiée de l'OCDE.^{viii}

Le revenu disponible équivalent est calculé en trois étapes:

1. D'abord, tous les revenus monétaires perçus par chaque membre du ménage, peu importe leur origine, sont additionnés. Il s'agit des revenus du travail et des placements et des prestations sociales, auxquels s'ajoute tout autre revenu perçu par le ménage, tous nets d'impôts et de cotisations sociales.
2. Ensuite, afin de prendre en compte les différences de taille et de composition des ménages, le revenu total (net) du ménage est divisé par le nombre d'équivalents adultes au moyen d'une échelle standard: l'échelle modifiée de l'OCDE. Cette échelle assigne une pondération à tous les membres du ménage (dont l'addition constitue la taille équivalente du ménage):
 - a. 1,0 au premier adulte;
 - b. 0,5 à chaque membre âgé de 14 ans et plus;
 - c. 0,3 aux enfants de moins de 14 ans.
3. Enfin, le chiffre obtenu, appelé «revenu disponible équivalent», est distribué à parts égales entre chaque membre du ménage.

5.2 Déterminer le revenu net

Dans l'EBM, la personne de référence est la personne ayant le revenu net le plus élevé. Cette personne doit d'abord être déterminée avant de pouvoir effectuer le calibrage.

Jusqu'en 2010, les revenus étaient inscrits dans le carnet de manière exhaustive, tout comme les dépenses. Depuis 2012, les revenus sont collectés de manière plus globale dans un questionnaire afin de limiter la charge de travail des répondants. Cette méthode de travail offre la possibilité de comparer les dépenses des ménages en fonction de leur niveau de vie (p.ex. quartiles de revenu) mais les tableaux détaillés sur la composition des revenus ne seront plus basés sur l'EBM.

Pour toutes les études sur les revenus, nous recommandons d'utiliser l'enquête EU-SILC (enquête communautaire sur les revenus et conditions de vie), qui est également organisée par Statbel^{ix}.

Pour le calcul du revenu disponible, tous les revenus du ménage sont additionnés, aussi bien les revenus communs que les revenus individuels des membres du ménage. Il s'agit plus spécifiquement des revenus mensuels nets suivants au niveau individuel:

- R1: salaire net
- R2: revenu net indépendant
- R3: revenu net des investissements (y compris les intérêts, les dividendes...)
- R4 : revenu net d'immeubles loués (loyers nets reçus)
- R5: pension de vieillesse
- R6: pré-pension (employeur + chômage)
- R7: allocations de chômage
- R8 : pension de veuf et d'orphelin
- R9: allocations familiales
- R10: allocations maladie/invalidité
- R11: revenu de l'activité secondaire
- R12: autre revenu
- R13: avantages en nature
- R14: remboursement mutualité
- R15: remboursement assurance
- R16: remboursement impôts

Les revenus sont vérifiés en détails par catégorie (identification des valeurs aberrantes, vérification des ménages qui n'ont pas rempli correctement les informations : questionnaire ménage vide, revenus trop faibles ou non conformes aux informations sur le statut socioprofessionnel, secteur d'activité ...).

Ensuite, un revenu est calculé par personne en prenant la somme de toutes les catégories. Le revenu du ménage est calculé en prenant la somme de tous les revenus des individus du ménage.

La somme de chaque type de revenu R1 à R16 est comparée au revenu déclaré par le ménage, en partant du principe que le revenu auto-déclaré par le ménage est plus précis parce que les ménages le connaissent mieux.

Un premier contrôle de qualité est effectué en validant le revenu EBM moyen par rapport au revenu moyen déclaré par les ménages dans l'enquête EU-SILC. Le Tableau 23 donne un aperçu de la comparaison des valeurs statistiques de base pour le revenu au niveau du ménage dans l'EBM avec le revenu au niveau du ménage trouvé dans l'enquête EU-SILC pour la même année de référence.

Tableau 23: Validation du revenu pondéré au niveau du ménage collecté dans l'EBM par rapport aux données collectées dans EU-SILC

Weighted Basic Statistical Measures (Location & Variability)	HBS 2020	EU-SILC 2020
Mean	38159.68	42798.78
Std Deviation	640083	846258
Median	32324.76	35268.36
Variance	4.09707 ^{E11}	7.16153 ^{E11}
Mode	19200.00	15744.27
Range	345510	887360
Interquartile Range	29640	35424

Un autre contrôle de qualité est effectué en répartissant les revenus en catégories ou en classes. Si aucun revenu ne peut être trouvé pour un ménage à partir des revenus déclarés entre R1 et R16 et qu'aucune donnée n'est disponible dans l'EFT, les classes sont utilisées pour imputer les valeurs médianes des valeurs manquantes. Les valeurs des classes obtenues sont à nouveau comparées avec les valeurs trouvées dans l'enquête EU-SILC. Le Tableau 24 donne un aperçu des classes de revenus pour EBM et EU-SILC pour la même année de référence.

Tableau 24: Répartition du revenu selon les classes

Income class	Income HBS 2020		Income EU-SILC 2020	
	Percent	Cumulative Percent	Percent	Cumulative Percent
[0-250[0.19	0.19	0.66	0.66
[250-500[0.00	0.19	0.30	0.96
[500-1000[2.29	2.48	2.66	3.62
[1000-1500[13.93	16.41	13.19	16.81
[1500-2000[15.69	32.10	13.08	29.89
[2000-2500[13.62	45.71	11.08	41.69
[2500-3000[9.97	55.68	9.50	51.20
[3000-5000[29.33	85.01	26.50	77.70
[5000+	14.99	100.00	22.30	100.00

Remarque

Pour le tableau ci-dessus, le revenu total disponible des ménages a été divisé par 12 pour obtenir la distribution des fréquences du revenu du SILC.

6 Calibrage et estimation

6.1 Introduction

Pour le calibrage de l'échantillon de répondants EBM 2020, nous nous basons sur le modèle de calibrage développé pour EBM 2018. La raison est la recherche d'une continuité dans la méthodologie, afin d'éviter les ruptures dans les statistiques publiées dues aux modifications méthodologiques.

La complexité du calibrage d'EBM 2020 est comparable à celle d'EBM 2018. Comme indiqué au paragraphe 2, les répondants sont issus de trois sources et les sous-échantillons de répondants qui en résultent seront calibrés indépendamment selon leurs modèles de calibrage respectifs, tout comme pour EBM 2018.

Le calcul de ces pondérations initiales (c'est-à-dire les poids de sondage, voir paragraphe 2) comprend l'estimation de toutes sortes de probabilités de réponse et d'auto-sélection. Ces probabilités représentent les étapes successives du processus de recrutement – variant selon la source – des ménages invités à participer à l'EBM 2020. L'ensemble des ménages invités à participer à l'EBM 2020 peut être considéré comme l'échantillon EBM 2020 initial. Le poids initial correspondant à chacun de ces ménages peut être considéré comme un poids de sondage (bien que *poids de recrutement* serait un terme plus approprié, pour insister sur la complexité de la procédure de recrutement).

Chaque ménage recruté et invité à participer à l'EBM 2020 a une probabilité de réponse intrinsèque, qui sera abordée au paragraphe 2. Les poids initiaux (ou poids de sondage ou poids de recrutement) mentionnés au paragraphe précédent ne sont pas corrigés sur la base de ces probabilités de réponse intrinsèques. Les poids initiaux sont donc directement corrigés par le calibrage, ce qui suppose de manière implicite que le calibrage corrige suffisamment la non-réponse à partir du recrutement et de l'invitation à participer à l'EBM 2020.

Au par. 6.2 ci-après, nous aborderons les cadres de calibrage à partir desquels sont calculés les totaux de calibrage ou benchmarks. Ensuite, au paragraphe 6.3, nous aborderons les modèles de calibrage qui seront finalement utilisés pour les sous-échantillons d'EBM 2020. Pour les détails techniques et l'interprétation, nous nous référons à *Enquête sur le budget des ménages 2016 - Note méthodologique* ([Statbel \(2017a\)](#) ou [Statbel \(2017b\)](#)), où les modèles renouvelés sont discutés en détail.

6.2 Cadres de calibrage

6.2.1 Un échantillon EFT 2020 calibré

Afin d'estimer les totaux de calibrage appropriés pour le calibrage de l'échantillon de répondants EBM 2020, l'échantillon de répondants d'EFT 2020 est calibré sur la population, tant au niveau des individus que des ménages. Le modèle de calibrage de l'EFT 2020 est le même que celui d'EFT 2018 et EFT 2016 (voir paragraphe 5.1.3 de *Enquête sur le budget des ménages 2016 - Note méthodologique* ([Statbel \(2017a\)](#) ou [Statbel \(2017b\)](#)):

$$\{HH; STRAT12; SEX \times AGE // HHSIZE; Logit\}$$

Ce modèle (avec notamment la légende entre parenthèses) a été abordé en détail au par. 5.1.3 dans [Statbel \(2017a\)](#) ou [Statbel \(2017b\)](#).

Remarques

- EFT 2020 a été calibrée sur une population de 10.776.236 individus dans 4.542.861 ménages privés de type EFT. Étant donné que les sous-échantillons d'EBM 2020 issus de différentes sources seront calibrés séparément, l'échantillon calibré de l'EFT 2020 ne fournira que les totaux de calibrage pour

les calibrages des sous-échantillons d'EBM 2020 contenant des ménages de type EFT, c'est-à-dire les sources 1, 2 et 3.1.

- Les poids de sondage de 17.122 ménages répondants d'EFT 2020 (vague 1) ont été corrigés à l'aide d'une probabilité de réponse estimée. Les poids de sondage corrigés de la non-réponse sont donc les poids qui ont été corrigés à nouveau lors du calibrage. Voir [Vanderhoeft et al \(2021\)](#) pour une discussion du *mixed effects logistic regression model* pour l'estimation des probabilités de réponse ; les prédicteurs utilisés dans le modèle de régression sont *HHtype*, *Origin*, *Province* et *Urbanisation*.

Les limites inférieure et supérieure des *g-weights* - c'est-à-dire les facteurs de correction déterminés par le calibrage - lors de l'application de la méthode logit sont, tout comme en 2016 et 2018, respectivement 0 et 10.

Les totaux de calibrage du modèle de calibrage de l'EFT 2020 (vague 1) - comme par exemple ceux de la ligne "Total" dans le Tableau 25 ci-après - sont calculés sur la base d'une extraction du RN du 1/07/2020.

Le Tableau 25 donne les sommes des poids calibrés par région et selon la variable de 'poststratification' PS, qui joueront un rôle dans le calibrage de l'échantillon d'EBM 2020 (voir paragraphes 6.3.1 et 6.3.2).

Tableau 25: Sommes des poids calibrés pour l'échantillon de répondants calibré d'EFT 2020 vague 1

PS : Post-strate (Au sein de la Région)	Région			Total
	BRU	VLA	WAL	
Salarié, 1P	101.248,48	319.959,99	158.944,60	580.153,07
Salarié, 2P, <=1act	36.639,81	189.141,45	112.254,67	338.035,93
Salarié, 2P, 2act	32.085,09	233.970,94	87.648,37	353.704,40
Salarié, 3+P, <=1act	55.115,36	195.162,72	116.062,17	366.340,26
Salarié, 3P, 2+act	19.941,87	202.466,85	109.402,76	331.811,48
Salarié, 4+P, 2+act	39.230,78	337.831,20	167.251,52	544.313,50
Indépendant	63.893,44	293.381,38	136.400,48	493.675,30
Non actif, <=59, 1P	49.307,62	95.197,99	115.645,58	260.151,19
Non actif, <=59, 2+P	39.294,14	76.969,32	101.140,61	217.404,07
Non actif, 60-69, 1P	32.778,92	146.200,89	102.643,63	281.623,44
Non actif, 60-69, 2+P	12.687,98	163.372,82	87.531,37	263.592,17
Non actif, 70+, 1P	25.072,23	110.177,47	77.486,90	212.736,61
Non actif, 70+, 2+P	13.755,27	195.230,97	90.333,34	299.319,58
Total	521.051,00	2.559.064,00	1.462.746,00	4.542.861,00

Les ventilations estimées de la population de l'échantillon EFT 2020 calibré seront calculées et seront utilisées comme benchmarks (totaux de calibrage) dans le calibrage d'EBM 2020: voir paragraphe 6.3.1. Il s'agit ici des variables qui sont utilisées dans la construction de la variable de 'poststratification' PS, à savoir le statut socio-professionnel et l'âge de la personne de référence du ménage, et le nombre de membres du ménage qui appartiennent ou non à la classe d'âge active. Le modèle de calibrage d'EFT 2020 est supposé réduire suffisamment le biais dans l'échantillon de répondants EFT, pour que le calibrage de l'EFT n'introduise pas de biais dans le calibrage de l'EBM.

6.2.2 Population de ménages de type EFT issus du RN

L'extraction du RN datant du 1/07/2020, qui a été complétée par le couplage au registre fiscal reprenant le revenu net imposable de 2018 (pour les individus et les ménages) a permis la construction d'un deuxième cadre de calibrage, dans le but de calculer des totaux de calibrage afin que certaines variables de revenu puissent être incluses dans les modèles de calibrage pour EBM 2020.

Ce deuxième cadre de calibrage est limité aux ménages de type EFT (et donc privés) dans lesquels les variables géographiques (région et province) et le revenu (net imposable) du ménage sont transformés en variables de classe de revenu INK_5, INK_5_G et INK_5_P (respectivement au niveau global, régional et provincial). La construction de ces variables est similaire au calibrage de l'EBM 2016 et 2018 et est expliquée en détail au paragraphe 5.2.3 de *Enquête sur le budget des ménages 2016 - Note méthodologique* ([Statbel \(2017a\)](#) ou [Statbel \(2017b\)](#)). Comme pour EBM 2018, nous nous limitons également aux ménages de type EFT pour EBM 2020.

Le '_5' dans le nom des variables des classes de revenu fait référence aux quintiles de population sous-jacents K_1, K_2, K_3, K_4 , qui sont calculés pour l'ensemble de la population ainsi que par région et par province, et qui sont également nécessaires pour calculer les variables INK_5, INK_5_G et INK_5_P pour les MN répondants d'EBM 2020 (voir paragraphe 5.2.3 de *Enquête sur le budget des ménages 2016 - Note méthodologique* ([Statbel \(2017a\)](#) ou [Statbel \(2017b\)](#))).

Le Tableau 26 présente la distribution relative (en pourcentage) de INK_5_G par région, telle qu'elle est dérivée de la population des ménages de type LFS dont il est question ici. La distribution absolue (par région) est obtenue en multipliant les totaux du Tableau 25 par les pourcentages. De cette façon, il y a une cohérence immédiate entre les totaux de calibrage dans le Tableau 25 et le Tableau 26.

Tableau 26: Répartition en pourcentage et en chiffres absolus des classes de revenus du ménage par région - variable INK_5_G

INK_5_G	BRU		VLA		WAL	
	Abs.	%	Abs.	%	Abs.	%
1	98.975,37	19,00	90.257,88	3,53	80.594,82	5,51
2	84.414,73	16,20	493.760,82	19,29	276.430,24	18,90
3	84.415,73	16,20	493.761,82	19,29	276.430,24	18,90
4	84.414,73	16,20	493.759,82	19,29	276.430,24	18,90
5	84.415,73	16,20	493.762,82	19,29	276.430,24	18,90
6	84.414,73	16,20	493.760,83	19,29	276.430,26	18,90
Total	521.051,00	100,00	2.559.064,01	100,00	1.462.746,02	100,00

Les ventilations de INK_5 et INK_5_P ne sont pas présentées parce que ces variables n'ont finalement pas été retenues dans les modèles de calibrage d'EBM 2020.

6.2.3 Population de ménages âgés issus du RN

Le (petit) sous-échantillon d'EBM 2020 de ménages âgés (voir paragraphe 2.4) sera calibré séparément selon une sous-population correspondante de ménages âgés selon le RN. A partir d'une extraction du RN du 1/07/2020 - la même que celle utilisée pour le calibrage d'EFT 2020 (paragraphe 6.2.1) - nous obtenons une population de 462.262 ménages privés âgés, répartis de la manière suivante selon la région et la taille du ménage (personne seule vs. ménage comptant plusieurs personnes).

Tableau 27: Nombre de ménages âgés selon la région et la taille du ménage (RN du 1/07/2020)

	Seul	Pas seul	Total	%
BRU	29.260	6.861	36.121	7,81%
VLA	199.865	93.442	293.307	63,45%
WAL	99.983	32.851	132.834	28,74%
Total	329.108	133.154	462.262	100,00%
%	71,20%	28,80%	100,00%	

Les nombres du Tableau 27 sont les totaux de calibrage du sous-échantillon EBM 2020 de ménages âgés (source 3.2); voir paragraphe 6.3.3.

6.3 Calibrage d'EBM 2020

6.3.1 *Calibrage de ménages de type EFT de l'EFT (source 1) et du RN (source 3.1)*

Le sous-échantillon EBM 2020 de ménages répondants qui sont issus des sources 1 et 3.1 contient au total 3.998 ménages répondants dont 2.939 de la source 1 et 1.059 de la source 3.1. Le modèle de calibrage d'EBM 2018 a également été appliqué pour ce sous-échantillon d'EBM 2020:

$$\{HH; REGION \times PS + PROV \times (C12 + C13 + C14) + REGION \times INK_5_G; TrLin\}.$$

Pour des explications détaillées sur ce modèle (et les variantes), nous vous renvoyons au paragraphe 5.2 de *Enquête sur le budget des ménages 2016 - Note méthodologique* ([Statbel \(2017a\)](#) ou [Statbel \(2017b\)](#)) et au paragraphe 5.2 de *Enquête sur le budget des ménages 2018 - Note méthodologique* ([Statbel \(2019a\)](#) ou [Statbel \(2019b\)](#)). Pour rappel:

- Le premier terme $REGION \times PS$, constitue une post-stratification classique par région et type de ménage (selon le statut d'occupation de la personne de référence et des autres membres du ménage, la taille du ménage et l'âge de la personne de référence : la variable PS est exactement telle que définie pour EBM 2018, avec 13 catégories (cf. Tableau 25));
- Le deuxième terme, $PROV \times (C12 + C13 + C14)$, apporte un équilibre entre les trimestres (via les variables de contraste C12, C13 en C14) au niveau provincial.
- Le troisième terme, $REGION \times INK_5_G$, permet d'aligner le ratio propriétaires/locataires sur celui estimé à partir du SILC : les modèles avec revenus dans le calibrage corrigent bien la surestimation des propriétaires par les modèles sans revenus. L'alternative plus détaillée $PROV \times INK_5_P$ n'a pas pu être appliquée pour la raison purement technique que, pour certaines provinces, la catégorie missing pour INK_5_P n'est pas représentée dans l'échantillon des répondants.

Les totaux de calibrage (benchmarks) ont été obtenus comme suit :

- Pour le premier terme, $REGION \times PS$, ils sont calculés à partir d'un échantillon EFT 2020 calibré; voir paragraphe 6.2.1 et Tableau 25;
- Pour le deuxième terme, $PROV \times (C12 + C13 + C14)$, ils sont tous égaux à 0, car ce terme concerne les variables de contraste: voir par. 5.2.2 dans [Statbel \(2017a\)](#) ou [Statbel \(2017b\)](#) pour une description détaillée de la construction des variables de contraste C12, C13 et C14, et pourquoi le 0 des totaux de calibrage correspondants;
- Pour le troisième terme, $REGION \times INK_5_G$, ils sont calculés à partir d'une population de ménages de type EFT du RN (complétée par les revenus nets imposables) : voir par. 6.2.2 et Tableau 26.

Le modèle ci-dessus a également été appliqué aux échantillons des sources 1 et 3.1 séparément et a livré des résultats analogues.

6.3.2 *Calibrage de ménages de type EFT d'EBM 2018 (source 2)*

Le calibrage du sous-échantillon EBM 2020 de la source 2 (EBM 2018) a donc été effectué séparément, tout comme pour EBM 2018. Les raisons en sont expliquées au paragraphe 5.3.2 de *Enquête sur budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b)).

Formellement, le modèle de calibrage appliqué aux ménages répondants de la source 2 est le même que pour le calibrage du sous-échantillon des sources 1 et 3.1 prises ensemble (voir paragraphe précédent 6.3.1); c'est-à-dire

$$\{HH; REGION \times PS + PROV \times (C12 + C13 + C14) + REGION \times INK_5_G; TrLin\}$$

Les totaux de calibrage ont été calculés exactement de la même manière qu'au par. 6.3.1. Le calcul des poids initiaux pour le calibrage tient compte du cheminement complexe: sélection aléatoire dans les échantillons de l'EFT et dans les échantillons supplémentaires contenant des ménages de type EFT issus du RN, à partir de là, recrutement pour l'EBM 2018, cfr. source 1 et source 3.1 de EBM 2018, puis recrutement pour l'EBM 2020 à partir des ménages répondants à l'EBM 2018; voir Tableau 10 au paragraphe 2.3. À cette fin, les probabilités d'(auto)-sélection $p_{i,RR}$ et $p_{i,RR}$ du **Fout! Verwijzingsbron niet gevonden.** ont simplement été 'modélisées' en fonction de la province dans laquelle vit le ménage, le trimestre au cours duquel le ménage a participé à l'EBM en 2018 et le trimestre au cours duquel il est invité à participer à l'EBM en 2020. Le 'modèle' est simplement un croisement des trois variables mentionnées, et le calcul de la valeur $p_{t,LFS}$ ou de la valeur $p_{i,RR}$ observée par cellule dans ce croisement

6.3.3 Calibrage des ménages âgés (source 3.2)

L'échantillon de répondants d'EBM 2020 de ménages privés âgés ne contient que 368 ménages. C'est la raison pour laquelle le modèle de calibrage est également plus limité que les modèles discutés ci-dessus pour le calibrage des sous-échantillons des ménages de type EFT :

$$\{HH; REGION \times PS^* + PROV \times (C12 + C13 + C14); TrLin\}$$

La variable PS^* a peu à voir avec la variable (de post-stratification) PS (qui compte 13 classes) utilisée dans les paragraphes précédents :

- $PS^* = 1$: Ménage d'une personne (single)
- $PS^* = 2$: Ménage de 2 ou plusieurs personnes (pas single)

En forçant un peu, PS^* peut être considéré comme une forte simplification de PS ; d'où le maintien de la notation ' PS '. Cette simplification est évidemment liée au fait que:

- tous les membres des ménages 'âgés' ont au moins 77 ans;
- la majorité des membres sont inactifs;
- la majorité de ces ménages ne comptent que 1 ou 2 membres.

Le revenu du ménage a bien été pris en considération pour l'intégrer dans le calcul via le calibrage. Un problème technique empêche toutefois l'utilisation de cette variable. Dans la population, il y a des ménages âgés qui ne sont pas repris dans le registre fiscal (donc sans revenu). Cependant, dans le petit échantillon de 368 ménages âgés, tous les ménages ont un revenu (net imposable) figurant dans le registre fiscal.

Le Tableau 28 ci-dessous montre la répartition de l'échantillon de répondants non pondéré et calibré selon les variables de calibrage. Notez que les sommes de poids en italique et en gras sont des totaux de calibrage correspondant au terme $REGION \times PS^*$ dans le modèle de calibrage (voir aussi Tableau 27; les sommes uniquement en italiques sont des totaux de calibrage implicites en raison du deuxième terme supplémentaire $PROV \times (C12 + C13 + C14)$).

Tableau 28: Echantillon de répondants EBM 2020 de ménages âgés: nombre de répondants et sommes des poids calibrés selon la province, la taille du ménage et le trimestre

Région	Province	Nombres de répondants						Sommes des poids calibrés				
		Trimestre EBM				Taille du ménage (PS*)		Total	par EBM trimestre	Taille du ménage (PS*)		Total
		T1	T2	T3	T4	Seul	En aucun e mesure Seul			Seul	Pas seul	
BRU	BRU	5	8	14	18	32	13	45	9.030,25	29.260,00	6.861,00	36.121,00
VLA	ANT	6	11	13	20	33	17	50	22.809,66	61.659,89	29.578,76	91.238,65
	WVL	3	7	12	8	17	13	30	12.870,21	37.208,80	14.272,05	51.480,86
	OVL	6	5	14	15	24	16	40	17.307,01	48.462,75	20.765,30	69.228,05
	LIM	5	4	5	8	14	8	22	6.692,62	18.754,09	8.016,41	26.770,49
	VLB	4	6	17	15	21	21	42	13.647,24	33.779,47	20.809,47	54.588,94
	Total	24	33	61	66	109	75	184	73.326,75	199.865,00	93.442,00	293.307,00
WAL	HAI	2	3	15	7	21	6	27	8.135,58	28.889,96	3.652,35	32.542,31
	LIE	10	9	12	9	25	15	40	13.522,30	42.286,79	11.802,41	54.089,20
	LUX	6	2	6	3	9	8	17	1.856,93	4.677,02	2.750,69	7.427,71
	NAM	3	4	7	8	10	12	22	4.330,47	10.463,26	6.858,63	17.321,89
	BWA	7	4	10	12	18	15	33	5.363,22	13.665,97	7.786,92	21.452,89
	Total	28	22	50	39	83	56	139	33.208,50	99.983,00	32.851,00	132.834,00
Total		57	63	12	12	135	144	368	115.565,50	329.108,00	133.154,00	462.262,00

6.3.4 Pourcentage de propriétaires de logements dans la population

L'ajout du terme REGION x INK_5_G était déjà, dans le cadre du calibrage de l'EBM 2016, une manière d'aligner le pourcentage de propriétaires de logements estimé sur la base d'EBM avec l'estimation de SILC, qui s'élève à 67,8% pour SILC 2020. Le Tableau 29 montre que nous y parvenons également dans une large mesure pour EBM 2020: le pourcentage global estimé des propriétaires de logements s'élève à 71,23%. Cette estimation s'obtient de la façon suivante (voir également Exemple 3 au par. **Fout! Verwijzingsbron niet gevonden.**).

Les sources 1+3.1 et 2 sont toutes deux considérées comme représentant la population de 4.542.861 ménages de type EFT. Par conséquent, les pourcentages de propriétaires de logements de 69,55% et 72,90% respectivement sont des estimations du même paramètre. Pour obtenir une moyenne, nous utilisons les nombres respectifs de répondants comme poids:

$$\frac{69,55\% \times 3998 + 72,90\% \times 1739}{3998 + 1739} = 69,55\% \times 0,6969 + 72,90\% \times 0,3031 = 70,57\%$$

Pour obtenir une estimation pour toute la population de 5.005.123 ménages privés, nous travaillons avec les nombres absolus estimés de propriétaires de logements:

$$\frac{70,57\% \times 4.542.861 + 77,73\% \times 462.262}{5.005.123} = \frac{3.205.897,00 + 359.316,25}{5.005.123} = \frac{3.565.213,25}{5.005.123} = 71,23\%$$

Cette méthode est appliquée pour l'estimation de tous les paramètres similaires (nombres et proportions/pourcentages).

Tableau 29: Estimation du nombre de propriétaires de logements (EBM 2020)

Source	Nombre MN répondants	Population de MN (total de calibrage)	Estimation des Propriétaires de logements	
			Nombre	%
MN de type EFT des sources 1 (EFT) + 3.1 (RN)	3998 (=2939 + 1059)	4.542.861	3.159.559,8	69,55%
MN de type EFT de la source 2 (EBM 2018)	1739	4.542.861	3.311.745,7	72,90%
MN âgés de la source 3.2 (RN)	368	462.262	359.316,3	77,73%
Global	6.105	5.005.123	3.565.213,3	71,23%
Global sans la source 3.2	5.737	4.542.861	3.205.897,0	70,57%

Le Tableau 29 montre que l'effet des répondants de la source 3.2 (ménages privés âgés) est relativement faible. À cet égard, il convient de noter que l'échantillon de 368 ménages âgés répondants se compose principalement de ménages ayant un revenu supérieur (qui sont plus susceptibles d'être propriétaires).

Nous constatons que l'estimation du pourcentage de propriétaires de logements s'élève à 72,90% si nous nous limitons aux ménages répondants de la source 2 (EBM 2018). Nous supposons qu'il s'agit d'un effet cumulé de l'autosélection: dans EBM 2018, le nombre de revenus plus élevés (et, par conséquent, le nombre de propriétaires de logements) était déjà surreprésenté, ce qui a entraîné une correction via le terme REGION x INK_5_G dans le modèle de calibrage. Et cette réponse sélective se prolonge d'EBM 2020 à EBM 2020.

À cet égard, il convient de noter que les sources 1+3.1 et 2 doivent également contenir des ménages 'âgés': initialement, c-à-d lors du tirage des échantillons EFT, tous ces ménages étaient de type EFT, mais depuis ce ou ces tirages initiaux, beaucoup de temps a passé et certains de ces ménages ont 'vieilli' si bien qu'ils pourraient être considérés comme des ménages 'âgés' lors de la collecte de données pour EBM 2020.

6.4 Estimation des paramètres

Les méthodes de calcul des estimations des paramètres EBM n'ont pas changé par rapport à EBM 2018. Pour des explications et des exemples, nous vous renvoyons donc au paragraphe 5.4 d'*Enquête sur le budget des ménages 2018 - Note méthodologique* (Statbel (2019a) ou Statbel (2019b)).

7 Analyse partie 2: traitement des données pour publication et diffusion

7.1 Représentativité du sondage

7.1.1 Répartitions au niveau du ménage

La composition de l'échantillon est contrôlée afin de s'assurer que les données sont représentatives. À cette fin, la distribution au niveau des ménages pour certaines variables est comparée à la distribution effective de cette variable dans la population sur la base des sources administratives disponibles, telles que le Registre national.

Tableau 30: Répartition selon la taille des ménages participants

Taille du ménage	EBM 2020				Rapport
	Non pondérée		Pondérée		
	Périodicité	%	Périodicité	%	
1	1.939	31,76	1.784.475	35,65	1,12
2	2.217	36,31	1.577.510	31,52	0,87
3	809	13,25	713.847	14,26	1,08
4	788	12,91	646.046	12,91	1,00
5	271	4,44	209.217	4,18	0,94
6	71	1,46	64.093	1,28	0,88
7	8	0,13	6.676	0,13	1,00
8	1	0,02	2.921	0,06	3,00
9+	1	0,02	338	0,01	0,50
Total	6.105	100,00	5.005.123	100,00	1,00

Remarques

- Dans l'échantillon des ménages répondants, la taille maximale du ménage n'est pas du même ordre de grandeur que la taille maximale du ménage dans la population.
- La distribution relative pondérée et non pondérée de la taille du ménage diffèrent, uniquement en raison d'un calibrage limité selon la taille de ménage dans l'EBM.

Tableau 31: Répartition selon le nombre d'actifs par ménage des ménages participants

Nombre d'actifs par ménage	EBM 2020				Rapport
	Non pondérée		Pondérée		
	Périodicité	%	Périodicité	%	
0	2.464	40,36	1.899.223	37,95	0,94
1	1.800	29,48	1.589.389	31,76	1,08
2	1.740	28,50	1.438.068	28,73	1,01
3	90	1,47	72.574	1,45	0,99
4	11	0,18	5.869	0,12	0,67
Totaal	6.105	100,00	5.005.123	100,00	1,00

7.1.2 Répartitions au niveau individuel

Certaines répartitions au niveau individuel sont comparées aux distributions de la population des sources de données administratives afin de vérifier la représentativité au niveau individuel.

Tableau 32: Répartition selon le sexe des membres des ménages participants

Sexe	EBM 2020				Rapport
	Non pondérée		Pondérée		
	Périodicité	%	Périodicité	%	
Homme	6.667	48,28	5.263.210	47,12	0.98
Femme	7.142	51,72	5.906.813	52,88	1.02
Total	13.809	100,00	11.170.023	100,00	1.00

Remarques

- Parce que l'EBM n'est pas calibrée au niveau individuel, le total pondéré n'est pas exactement égal au chiffre de la population. Les distributions relatives diffèrent également pour cette même raison.
- Le poids individuel calibré est égal au poids calibré du ménage auquel appartient l'individu.

Tableau 33: Répartition selon l'âge des membres des ménages participants

Âge	EBM 2020				Rapport
	Non pondérée		Pondérée		
	Périodicité	%	Périodicité	%	
- 25	3.907	28,29	3.406.829	30,50	1.08
26-45	3.465	25,09	3.181.039	28,48	1.14
46-64	3.558	25,77	2.643.063	23,66	0.92
65-74	1.825	13,22	978.414	8,76	0.66
75 +	1.054	7,63	960.678	8,60	1.13
Total	13.809	100,00	11.170.023	100,00	1.00

Remarques

- Parce que l'EBM n'est pas calibrée au niveau individuel, le total pondéré n'est pas exactement égal au chiffre de la population. Les distributions relatives diffèrent également pour cette même raison.
- Le poids individuel calibré est égal au poids calibré du ménage auquel appartient l'individu.

7.2 Traitement des données

Après la collecte des données, celles-ci sont à nouveau contrôlées.

7.2.1 *Identification des valeurs aberrantes*

Pour les dépenses mentionnées dans le carnet ménage, des programmes SAS ont été développés pour le contrôle et l'identification des valeurs aberrantes. Pour l'alimentation et les boissons non alcoolisées, nous avons opté pour une méthode intitulée 'resistant fences method' (RFM). Les données ont au préalable subi une transformation de Box-Cox (avec un paramètre λ qui varie en fonction de la catégorie de produit) puisque la symétrie de la distribution améliore la qualité de la détection avec la méthode RFM. Les valeurs aberrantes sont identifiées et remplacées par la médiane de la catégorie de dépenses.

Dans la méthode RFM univariée, la marge de tolérance est déterminée à l'aide de quartiles:

$$[Q1 - \text{coeff.low} * \max(Q3 - Q1, |a \times Q2|), Q3 + \text{coeff.up} * \max(Q3 - Q1, |a \times Q2|)].$$

Le terme $a \times Q2$ représente une valeur minimale de la distance. Les coefficients (low et up) ont été fixés à 3, comme habituellement dans la pratique et à 0,05 pour le paramètre a.

Pour l'alimentation et les boissons non alcoolisées, les valeurs aberrantes sont détectées automatiquement. Ce n'est pas le cas pour les autres catégories. À l'aide de la méthode RFM, nous corrigeons les dépenses alimentaires ECOICOP 01 par un facteur de 0,9%, ce qui est raisonnable.

Tableau 34: Différence de montant après l'application de la méthode RFM pour ECOICOP 1 et 2 (non pondéré)

Année	ECOICOP	Montant avant	Montant après application RFM	Différence
2020	01	1.449.527	1.445.448	4.079
	02	219.801	214.497	5.304

Si nous utilisons la même méthode pour ECOICOP 12 par exemple, 'Soins corporels et services', la baisse des dépenses totales s'établirait à environ 25 %, ce qui est irréaliste.

C'est pourquoi nous préférons contrôler les valeurs données par catégorie de dépenses pour les autres catégories de la COICOP et remplacer les montants réellement divergents par la médiane, sans effectuer de correction automatique.

Lors de l'enregistrement des dépenses dans un carnet de dépenses pendant un demi-mois, il est difficile d'avoir une idée des dépenses périodiques et importantes (rares ou non) d'un ménage. C'est pourquoi des questions sur ces deux types de dépenses sont reprises dans un questionnaire ménage distinct^x.

Différents contrôles de cohérence sont effectués en comparant les dépenses enregistrées dans le carnet du ménage et celles du questionnaire ménage.

Toutes les dépenses enregistrées sont ramenées sur un demi-mois: les dépenses du carnet restent telles quelles. Les dépenses du questionnaire ménage (questions sur les 4 derniers mois et questions sur la possession des biens durables) sont divisées par 8 et les dépenses périodiques sont divisées par 2. Pour le tableau final, toutes ces dépenses sont recalculées pour refléter les dépenses sur base annuelle, c'est-à-dire multipliées par 24.

7.2.2 Imputation des dépenses énergétiques

Le questionnaire ménage a été entièrement rempli sous la forme d'une interview CAPI. Cela signifie que lors de la dernière visite au ménage, l'enquêteur lit les questions sur son écran d'ordinateur et encode immédiatement les réponses données. Il suit l'ordre des questions (routing) tel que programmé à l'avance afin de s'assurer que toutes les questions pertinentes soient posées. Avant 2014, un questionnaire ménage sur papier était utilisé. Dans la pratique, c'était souvent le ménage lui-même qui le remplissait, après quoi l'enquêteur effectuait seulement une vérification. Beaucoup de données étaient dès lors manquantes, notamment sur les dépenses énergétiques. Les incohérences dans les réponses étaient également nombreuses (p. ex. des ménages qui déclarent que leur principale source de chauffage fonctionne au mazout de chauffage, mais ne mentionnent une dépense que pour du gaz naturel). Ces informations manquantes et ces incohérences engendraient beaucoup de travail de correction par la suite. L'introduction des questionnaires ménage CAPI a grandement amélioré la qualité des données, réduisant ainsi le travail de vérification ex post.

Tableau 35: Nombre de données manquantes par catégorie pour les dépenses énergétiques

Dépense	Nombre de données manquantes en 2020
Eau	1131
Électricité	167
Gaz (conduite de gaz)	11

Il est également possible d'encoder des factures combinées. Dans une seule facture, par exemple, un ménage paie la facture d'électricité et de gaz, ou la consommation d'électricité et d'eau. Cependant, les ménages et les enquêteurs devaient quand même ventiler les montants autant que possible. Dans la suite de ce document, on appelle également imputation la ventilation d'une telle combinaison de types d'énergie. Même pour les ménages qui n'ont pas fourni d'informations dans leur questionnaire ménage, des sources d'information sur leurs dépenses énergétiques sont disponibles. Les carnets, par exemple, fournissent des informations sur les dépenses pendant la période de référence.

Il a été décidé d'utiliser d'abord les informations des carnets pour estimer les dépenses manquantes, puis d'essayer de ventiler les dépenses mixtes et d'imputer ensuite les dépenses énergétiques encore manquantes en utilisant ensuite un modèle de régression linéaire.

7.2.3 Collecte des informations des carnets

La première étape consiste à utiliser les informations enregistrées dans les carnets pendant la période de référence comme base pour imputer une consommation annuelle manquante dans le questionnaire ménage. Nous ne faisons cette imputation que pour l'eau, le gaz et l'électricité parce que ces dépenses sont effectuées selon une fréquence fixe. En estimant cette périodicité, les dépenses qui ont eu lieu au cours d'un mois donné peuvent être converties en montants annuels. Cette estimation a été faite sur la base de montants moyens basés sur les données des questionnaires par type de ménage. Les données imputées se sont vu attribuer un 'drapeau' :

Tableau 36: Nombre de données imputées des carnets pour les dépenses énergétiques

Catégorie	Drapeau	2020
Facture d'eau	FL_W	122
Facture d'électricité	FL_E	9
Facture de gaz	FL_G	11

7.2.4 Ventilation de factures groupées

Il y a différentes combinaisons possibles des dépenses énergétiques. La combinaison la plus fréquente est la facturation conjointe de gaz et d'électricité.

Tableau 37: Aperçu des factures groupées pour les dépenses énergétiques

Description	2020
Eau, gaz et électricité	32
Eau et électricité	19
Eau et gaz	31
Electricité et gaz	1523

Les combinaisons 'eau et électricité' et 'eau et gaz' ont été ventilées en établissant un modèle de régression qui estime la consommation d'eau. Le modèle de régression a été basé sur les variables indépendantes taille du ménage (exprimée en unités équivalentes), région, présence d'un système de récupération des eaux de pluie et d'une machine à laver. La présence d'un puits n'a pas eu d'incidence significative sur les estimations et n'a pas été incluse dans la régression. La consommation de l'autre source d'énergie (l'électricité et le gaz respectivement) a ensuite été calculée en faisant la différence entre la facture commune et l'estimation de la consommation d'eau.

La combinaison 'électricité et gaz' a été ventilée en établissant un modèle de régression qui estime la consommation d'électricité. Le modèle était basé sur la taille du ménage (en unités équivalentes), la région et la présence d'une machine à laver et de panneaux solaires photovoltaïques. La présence d'un réfrigérateur ou congélateur ou l'utilisation de l'électricité pour un chauffage d'appoint n'a pas eu d'incidence significative sur les estimations et n'ont donc pas été retenues pour le modèle de régression. La consommation de gaz a été calculée en faisant la différence entre la facture commune et la consommation d'électricité estimée.

La combinaison 'eau, gaz et électricité' a été ventilée en effectuant les deux estimations, d'abord pour l'eau et ensuite pour l'électricité. La consommation de gaz a été ensuite calculée en faisant la différence entre la facture commune et les deux consommations estimées d'eau et d'électricité.

Remarque

Toutes les factures communes n'ont pas pu être ventilées de cette façon. Le montant de la consommation restante après déduction de l'estimation de la facture commune était parfois trop faible. Dans ce cas, aucune ventilation n'a été effectuée.

7.2.5 Imputer les données manquantes

Les données encore manquantes ont été complétées par une estimation fondée sur un modèle de régression par catégorie de données. Les dépenses manquantes dans la catégorie BOIS (ECOICOP=04549A) n'ont pas été imputées. Le bois de chauffage peut être obtenu gratuitement ou un stock peut avoir été constitué pour plusieurs années.

Tableau 38: Nombre de données imputées sur la base des modèles de régression pour les dépenses énergétiques

Catégorie	Drapeau	2020	2018	2016
Eau	FL_WR	1131	1219	320
Électricité	FL_ER	167	141	148
Gaz	FL_Gas	674	203	155
Fioul domestique (mazout)	FL_oil	341	105	91
Butane	FL_butane	.	.	11
Propane	FL_propane	.	.	78
Butane et propane	FL_butprop	40	9	.
Charbon	FL_coal	.	1	4
Pellets	FL_pellets	28	5	9
Pétrole	FL_petroleum	.	2	.

- On suppose que chaque ménage doit payer une facture d'eau. Les montants manquants sont remplis à cette étape à l'aide du même modèle de régression que celui utilisé pour la ventilation de la facture commune.
- Comme pour la consommation d'eau, nous supposons que chaque ménage participant consomme de l'électricité. La méthode de régression est la même que celle utilisée pour ventiler les factures groupées.
- Un montant est imputé pour le gaz pour tous les ménages qui déclarent se chauffer au gaz et qui n'ont enregistré aucune dépense annuelle de gaz dans le questionnaire ménage ou le carnet. Les variables indépendantes pertinentes sont la taille du ménage (exprimée en unités équivalentes), la région et le fonctionnement du système principal de chauffage au gaz.
- Les montants pour le gasoil de chauffage ont également été estimées sur la base d'un modèle de régression linéaire. Ceci pour tous les ménages qui ont indiqué qu'ils utilisaient du gasoil pour se chauffer, mais qui n'ont pas enregistré de dépenses annuelles en gasoil. Les variables indépendantes pertinentes sont la taille du ménage (exprimée en unités équivalentes), habitant de la Région de Bruxelles-Capitale et le fonctionnement du système principal de chauffage au gasoil.
- Aucun modèle de régression n'a pu être élaboré pour l'imputation du butane en raison de l'absence de données complètes. L'imputation a ensuite été effectuée sur la base des dépenses annuelles moyennes en butane de ces ménages qui avaient déclaré un montant. L'imputation du propane a donc été effectuée sur la base de la moyenne.
- L'imputation du charbon aux ménages qui utilisent le charbon pour se chauffer n'a pas pu être effectuée sur la base d'un modèle de régression. Le nombre de cas complets dans le set de données était en effet trop faible pour cela. L'imputation a dès lors été effectuée sur la base des dépenses annuelles moyennes en charbon de ces ménages qui avaient déclaré une dépense de charbon.
- La consommation de pellets n'a pas pu être prévue de manière significative par un modèle de régression. L'imputation a donc été effectuée sur la base de la moyenne.
- En ce qui concerne la consommation de pétrole, aucune donnée n'était manquante, et donc aucune imputation n'a été nécessaire.

7.3 Loyer imputé

Le 'loyer imputé' est un loyer fictif attribué aux propriétaires et aux ménages qui habitent un logement à titre gratuit.

D'après le SEC (système européen des comptes), l'achat d'une habitation est principalement considéré comme un investissement/une formation de capital, et pas comme une consommation^{xi}. La propriété de l'habitation fournit un service de logement/hébergement consommé par les ménages au fil du temps. Par conséquent, un prix doit être estimé. Étant donné qu'il n'y a pas de transaction monétaire, cette estimation doit se faire à l'aide d'une imputation. Ce loyer imputé fait partie de la consommation. Dès lors, afin de rendre l'EBM cohérente avec le SEC, les mêmes principes sont appliqués.

Pour EU-SILC, le loyer imputé est obligatoirement transmis à Eurostat depuis 2007. Pour EBM, cette variable est déjà demandée depuis 2005.

Afin d'estimer le prix auquel leur logement pourrait être loué, et donc avoir une idée de l'avantage financier lié au statut de propriétaire, nous utilisons la méthode recommandée par Eurostat: un modèle de Heckman à deux phases. Nous utilisons d'abord une régression probit pour estimer la probabilité d'être locataire en fonction des caractéristiques du ménage et de ses membres (notamment l'âge, le niveau d'instruction, le nombre de membres actifs, etc.) Nous estimons ensuite à l'aide d'une régression le loyer avec les caractéristiques du logement présentes dans l'enquête, telles que le type et l'âge du logement, le nombre de pièces, la présence d'un jardin, d'un garage, d'un balcon, du chauffage central Et naturellement, en tenant compte du résultat de la première régression.

Cette méthode est utilisée par Statbel tant pour l'enquête SILC que pour EBM. Par ailleurs, les résultats d'EBM et de SILC sont comparés pour chaque année d'enquête.

Tableau 39: Validation du loyer réel et imputé par rapport à l'enquête EU-SILC (pondéré)

	EBM 2020 (€)	SILC 2020 (€)
Loyer mensuel par ménage locataire	603	587
Loyer imputé mensuel par ménage propriétaire	751	747

Remarques

- Pour le tableau ci-dessus, la moyenne de la variable 'loyer actuel lié au logement occupé' (Current rent related to occupied dwelling - The variable refers to the total monthly rent that is currently paid on the main residence of the household) a été utilisée pour le loyer mensuel par ménage locataire de SILC 2020.
- Pour le 'loyer imputé mensuel par ménage propriétaire', la variable 'loyer imputé' de SILC 2020 (niveau annuel divisé par 12) a été utilisée, la moyenne étant prise uniquement pour les propriétaires.

7.4 Extrapolation des données

Les dépenses sont agrégées au niveau des ménages et extrapolées aux dépenses annuelles selon la nomenclature ECOICOP^{xii}. La Classification of Individual Consumption according to Purpose (COICOP) est une nomenclature des dépenses de consommation gérée par les Nations unies (COICOP (ONU)). La European Classification of Individual Consumption according to Purpose (ECOICOP)^{xiii} est l'adaptation européenne de cette nomenclature et compte 5 niveaux. Cette nomenclature est à la base des nomenclatures utilisées pour les dépenses de consommation notamment dans les comptes nationaux^{xiv}, l'indice des prix à la consommation^{xv} et EBM.

L'ECOICOP se compose de douze divisions décrivant la consommation individuelle des ménages, divisées en 47 groupes et 117 classes. La nomenclature est ensuite sous-divisée en sous-classes. Pour la Belgique, un sixième niveau est ajouté à l'ECOICOP: elle devient donc la COICOP-BE. Les données pour la Belgique sont publiées jusqu'au sixième niveau, et jusqu'au quatrième niveau à l'échelle régionale, en raison de la faible consommation dans certaines catégories. La nomenclature COICOP des Nations Unies fait actuellement l'objet d'une mise à jour.^{xvi}

Cette nomenclature contient 12 catégories ECOICOP de dépenses:

01. Produits alimentaires et boissons non alcoolisées
02. Boissons alcoolisées, tabac et stupéfiants
03. Vêtements et chaussures
04. Logement, eau, gaz, électricité et autres combustibles
05. Meubles, appareils et ustensiles ménagers, entretien courant
06. Santé
07. Transport
08. Communication
09. Culture et temps libre
10. Éducation
11. Horeca
12. Soins et services personnels

Chacune de ces catégories est subdivisée en sous-catégories. Le niveau européen va jusqu'à 5 positions. Il y a un 6^e niveau qui est une subdivision nationale supplémentaire qui fournit davantage de détails sur les catégories. Les résultats nationaux sont diffusés jusqu'au sixième niveau. Les résultats régionaux sont diffusés jusqu'au quatrième niveau. Il n'y a pas de diffusion d'un niveau plus détaillé dans les régions car les résultats obtenus ne sont pas suffisamment précis en raison de la taille limitée des échantillons par région.

Toutes les dépenses ECOICOP sont extrapolées de la même manière, en estimant d'abord les dépenses totales pour l'ensemble de la population par catégorie ECOICOP, et en divisant ensuite ce total par le nombre total de ménages dans la population, c'est-à-dire pour 2018: 4.921.414 ménages.

Étant donné que nous donnons une dépense moyenne pour tous les ménages (pas seulement pour les achats des ménages), le nombre de ménages dans le dénominateur est toujours le même. Il est donc facile de recalculer, sur la base des tableaux que nous fournissons comportant les dépenses moyennes, le montant total des dépenses par catégorie ECOICOP.

8 Estimation de la variance

8.1 Introduction

L'estimation de la variance doit prendre en compte au mieux toutes les étapes de l'enquête:

- Le plan de sondage, et le tirage de l'échantillon. La théorie des sondages définit pour chaque étape du plan de sondage des formules pour estimer la variance de l'estimateur d'un total. Quand un plan de sondage est à plusieurs degrés, les formules se combinent pour tenir compte de chaque degré du plan de sondage.
- La non-réponse et la manière dont elle est corrigée. La non-réponse est une étape de l'enquête non contrôlée par le statisticien. Dans le cadre de l'estimation de la variance, la non-réponse est modélisée par une seconde phase du plan de sondage, la première phase étant le plan de sondage déterminé pour l'enquête. La probabilité de réponse est par définition inconnue. On doit donc l'estimer. Dans l'enquête EBM, la non-réponse n'est pas corrigée explicitement. On va donc estimer la probabilité de répondre par le rapport entre le poids final (poids après calage) et le poids de sondage. Le processus de non-réponse est alors modélisé comme un tirage de Poisson, la probabilité de tirage étant estimée par la probabilité de réponse définie ci-dessus.
- Le calage. Le calage est l'étape qui consiste à modifier les poids de sondage (ou les poids corrigés pour la non-réponse) de sorte que les totaux de certaines variables connues sur l'univers entier soient estimés exactement sur base de l'échantillon et des poids adaptés, et ce tout en conservant les poids les plus proches possible des poids initiaux. Pour l'estimation de la variance de l'estimateur d'un total par calage, l'estimateur de la variance s'obtient en substituant la variable d'intérêt par le résidu de la régression linéaire de cette variable sur les variables de calage.
- L'estimateur utilisé. Les formules habituelles d'estimation de la variance sont définies pour des estimations de totaux sur l'univers. Quand le paramètre estimé est un ratio de totaux, on substitue à la variable d'intérêt une variable que l'on dit 'linéarisée'. L'estimation du total de la variable linéarisée a asymptotiquement la même variance que le ratio auquel on s'intéresse.

Le calcul de la variance est réalisé par le logiciel POULPE. POULPE est un ensemble de macro programmées sous SAS par l'INSEE. Il permet de calculer la variance pour différents plans de sondage, et de prendre en compte la non-réponse et le calage. POULPE permet d'estimer la variance pour des totaux, mais offre aussi une solution pour les variables calculées à partir de totaux (par exemple les moyennes, mais aussi la différence de deux totaux...). POULPE permet donc d'estimer la variance en prenant en compte toutes les étapes du traitement de l'enquête.

Le plan de sondage de l'enquête HBS est complexe. La combinaison de diverses sources pour l'échantillon a conduit à un estimateur complexe (calage des différentes sources, suivi de la combinaison des différentes sources pour obtenir un estimateur final). Cette complexité a conduit à devoir simplifier la modélisation du processus pour rendre l'estimation possible.

8.2 Description du plan de sondage

L'échantillon de EBM est constitué de quatre sources (voir les détails dans la partie 2) de sélection des ménages. Le plan de sondage, pour chacune de ces sources, est le suivant:

- Premier degré: sélection (éventuellement avec répétition) selon un tirage systématique proportionnel à la taille (SYSPPSWR, *systematic probability proportional to size with replacement*) d'unités primaires (UP) (les unités primaires sont des unités géographiques). La taille d'une unité primaire est le nombre de ménages éligibles pour l'enquête LFS dans l'unité primaire, selon le Registre national.
- Second degré: sélection par un sondage aléatoire simple (SRS) d'un nombre fixe de ménages dans les unités sélectionnées au premier degré.

En raison du tirage systématique au premier degré, il n'existe pas d'estimateur non biaisé de la variance pour ce type de plan de sondage. Ceci découle du fait que certaines paires d'unités primaires voisines (les unités primaires de taille inférieure au pas de tirage) n'ont aucune chance d'être tirées en même temps. Les probabilités d'inclusion doubles sont donc nulles pour certaines unités primaires, et donc on ne peut pas définir un estimateur non biaisé de la variance.

POULPE propose une solution approximative, pour les tirages systématiques à probabilité égale, sans répétition. Cette solution n'est pas applicable ici, en raison du tirage avec répétition, et des probabilités de tirage inégales.

Dans le cadre de l'estimation de la variance de l'enquête LFS, nous avons pu montrer par simulation que dans le cadre de ce plan de sondage (SYSPPSWR suivi d'un SRS de taille fixe par UP sélectionnée) la variance pouvait être estimée de manière satisfaisante en considérant que chaque groupe est le résultat d'un sondage aléatoire simple de taille égale à la taille du groupe, dans une strate de taille égale à la taille du pas de tirage du tirage systématique des unités primaires. Ceci fournit une solution pour estimer la variance due au plan de sondage. Il suffit pour chaque groupe de déterminer la taille de l'échantillon (connu) et la taille du pas, utilisé comme taille de la strate. Ces deux éléments étant suffisants pour estimer la variance d'un total dans le cas d'un SRS stratifié.

8.2.1 Sources 1 et 3.1

Les ménages sélectionnés pour les sources 1 et 3.1 sont issus des mêmes unités primaires sélectionnées pour l'enquête EFT. Les ménages de la source 1 sont les ménages sélectionnés pour EFT et qui ont accepté de participer à EBM. Les ménages de la source 3.1 sont les ménages sélectionnés dans les mêmes UP. Ces ménages ont été sollicités par courrier pour participer à EFT. Comme ces ménages sont issus des mêmes UP, ils ont été traités ensemble au niveau du calage (après détermination d'un poids avant calage, intégrant en partie le processus de non-réponse lié à l'enquête EFT; voir 6.3.1).

Pour les groupes de la source combinée, on a donc comme taille de la strate la taille du pas de sélection des UP, et comme taille de l'échantillon le nombre de ménage tirés à l'origine pour l'enquête LFS plus le nombre de ménages tirés pour l'échantillon supplémentaire.

8.2.2 Source 2

Les ménages sélectionnés pour la source 2 sont d'anciens ménages ayant répondu en 2018 à l'enquête EBM. Ces ménages ont donc été tirés à l'origine pour l'enquête EFT, ont ensuite participé à l'enquête EBM 2018, et ont accepté de répondre à nouveau à EBM 2020.

Comme pour la source 1, la modélisation du plan de sondage devrait être une strate par groupe tiré à l'origine pour EFT, la taille de la strate égale au pas de tirage utilisé à l'origine, et la taille de l'échantillon égale à la taille du groupe tiré à l'origine pour EFT.

Cependant, cette stratégie ne peut être appliquée. En effet, l'estimateur de la variance dans le cadre d'un sondage aléatoire simple stratifié est donnée par

$$\hat{V}_{ST}(\hat{t}) = \sum_{h=1}^H \hat{V}(\hat{t}_h) = \sum_{h=1}^H N_h^2 \left(1 - \frac{n_h}{N_h}\right) s_h^2$$

où

$$s_h^2 = \sum_{i=1}^{n_h} \frac{(y_i - \bar{y}_h)^2}{n_h - 1}$$

En cas de non-réponse, s_h^2 est calculé sur les répondants. Il faut donc au minimum deux répondants par groupe pour pouvoir réaliser l'estimation.

Or les 1739 répondants de la source 2 se répartissent entre 967 groupes différents ayant au moins un répondant, parmi lesquels 487 groupes ayant seulement 1 répondant. Pour permettre que tous les répondants soient pris en compte dans l'estimation de la variance, on a procédé au regroupement des groupes. Les strates sont donc redéfinies en regroupant tous les groupes d'un trimestre par province. Les 44 strates ainsi définies (Prov12*TRIM) comprennent de 18 à 62 répondants.

8.2.3 Source 3.2

La source 3.2 couvre les ménages composés exclusivement de personnes âgées de plus de 75 ans. Les unités primaires utilisées sont les mêmes que pour les sources 1 et 3.1. Le nombre de ménages sélectionnés par groupe varie de 5 à 8 selon la période. On pourrait a priori utiliser la même modélisation du plan de sondage que pour la source 1. Cependant, le taux de réponse dans cet échantillon est extrêmement faible. Comme pour la source 2, on a été contraint de fortement simplifier la modélisation du plan de sondage. Dans le cadre de l'estimation de la variance, on considère que l'échantillon est le résultat d'un sondage aléatoire simple dans l'univers des ménages 75+.

8.3 Calibrage

Pour l'estimation de la variance de l'estimateur d'un total par calage, l'estimateur de la variance s'obtient en substituant la variable d'intérêt par le résidu de la régression linéaire de cette variable sur les variables de calage. Cette fonctionnalité est intégrée dans POULPE.

Le calage est réalisé indépendamment sur les trois regroupements mentionnés ci-dessus. Les variables de calage sont éventuellement différentes pour les trois groupements utilisés pour le calage. Or POULPE utilise l'entièreté de l'échantillon répondant comme un seul échantillon. L'on doit donc étendre les variables de calage en leur donnant une valeur 0 en dehors de la source sur laquelle elles ont en premier lieu été définies.

8.3.1 Correction de la non-réponse

Dans le processus d'estimation de la variance, la non-réponse est modélisée comme une seconde phase de tirage, selon un processus de Poisson (on suppose que chaque ménage répond indépendamment des autres ménages, selon une probabilité p_i . Ici aussi, l'hypothèse n'est certainement pas complètement vérifiée, la non-réponse pouvant par exemple dépendre de l'enquêteur, et donc ne pas être totalement indépendante pour des ménages d'un même groupe). Les probabilités de réponse des ménages sont bien sûr inconnues, mais sont estimées, dans le cadre de l'estimation de la variance, pour les ménages répondants par $\tilde{p}_i = d_i/w_i$ où d_i est le poids de sondage et où w_i est le poids de calage.

Cette seconde phase de sondage, nécessaire pour modéliser la non-réponse, est intégrée dans les fonctionnalités de POULPE.

8.3.2 Combinaison des sources

Les trois sources ci-dessus sont chacune calées sur le total de la population correspondante. Pour obtenir une estimation sur la population totale, on combine les résultats de la manière suivante (voir 6.4):

$$\begin{aligned}\hat{T} &= \frac{\hat{T}_1 \times 3998 + \hat{T}_2 \times 1739}{3998 + 1739} + \hat{T}_3 \\ &= \alpha \hat{T}_1 + (1 - \alpha) \hat{T}_2 + \hat{T}_3\end{aligned}$$

La 'mise à l'échelle' des totaux 1 et 2 se justifie parce que chaque source a été traitée indépendamment, et calée sur le total de l'univers des ménages type EFT. Le total 3 est traité indépendamment et calé sur le total de l'univers 75+. Donc

$$\hat{V}(\hat{T}) = \hat{V}(\alpha \hat{T}_1 + (1 - \alpha) \hat{T}_2 + \hat{T}_3) = \hat{V}(\hat{T}'_1 + \hat{T}'_2 + \hat{T}_3)$$

où, p.ex.

$$\hat{T}'_1 = \alpha \hat{T}_1 = \alpha \sum_{i \in r_1} w_i y_i = \sum_{i \in r_1} \alpha d_i \tilde{p}_i y_i$$

Et où r_1 est l'échantillon des ménages répondants venant des sources 1 et 3.1.

Pour estimer la variance d'un total sur des trois sources combinées, on a considéré chaque source comme une strate séparée, avec les poids corrigés des facteurs α et $(1 - \alpha)$.

8.4 Aperçu de l'estimation de la variance

Tableau 40: Aperçu de l'estimation de la variance pour les catégories ECOICOP jusqu'au 2e niveau pour la Belgique et les régions

COICOP	Géographique	Variance	Consommation	Ecart type	Coefficients de variation	Intervalle de confiance	
						Minimum	Maximum
00	Belgique	115167,60	35208,78	339,36	0,01	34543,63	35873,93
	Flandre	249613,60	36446,91	499,61	0,01	35467,67	37426,16
	Wallonie	249314,50	34095,98	499,31	0,01	33117,33	35074,64
	Bruxelles	709348,10	32057,03	842,23	0,03	30406,26	33707,80
01	Belgique	5871,20	5596,67	76,62	0,01	5446,49	5746,85
	Flandre	11359,70	5835,72	106,58	0,02	5626,82	6044,62
	Wallonie	9325,78	5305,61	96,57	0,02	5116,34	5494,89
	Bruxelles	99550,77	5206,36	315,52	0,06	4587,94	5824,77
02	Belgique	572,97	760,50	23,94	0,03	713,59	807,42
	Flandre	1101,43	755,83	33,19	0,04	690,78	820,88
	Wallonie	1668,08	826,01	40,84	0,05	745,96	906,06
	Bruxelles	3695,78	596,84	60,79	0,10	477,68	715,99
03	Belgique	1602,78	1201,36	40,03	0,03	1122,89	1279,83
	Flandre	3787,03	1364,11	61,54	0,05	1243,49	1484,73
	Wallonie	2001,68	916,45	44,74	0,05	828,76	1004,14
	Bruxelles	13708,69	1184,06	117,08	0,10	954,57	1413,54
04	Belgique	3214,58	11205,45	56,70	0,01	11094,33	11316,58
	Flandre	6297,09	11351,90	79,35	0,01	11196,36	11507,43
	Wallonie	8491,98	10983,96	92,15	0,01	10803,34	11164,58
	Bruxelles	13708,69	1184,06	117,08	0,10	954,57	1413,54
05	Belgique	33954,24	2348,11	184,27	0,08	1986,95	2709,28
	Flandre	97502,63	2777,04	312,25	0,11	2165,02	3389,06
	Wallonie	19271,58	1840,32	138,82	0,08	1568,23	2112,41
	Bruxelles	28407,70	1606,45	168,55	0,10	1276,10	1936,80
06	Belgique	5358,05	1803,36	73,20	0,04	1659,90	1946,83
	Flandre	11795,74	1820,92	108,61	0,06	1608,05	2033,80
	Wallonie	10730,30	1799,40	103,59	0,06	1596,37	2002,43
	Bruxelles	35340,47	1724,83	187,99	0,11	1356,37	2093,29
07	Belgique	30173,02	3420,08	173,70	0,05	3079,62	3760,54
	Flandre	51724,38	3331,39	227,43	0,07	2885,63	3777,15
	Wallonie	115474,30	3923,72	339,82	0,09	3257,68	4589,76
	Bruxelles	132436,00	2431,81	363,92	0,15	1718,53	3145,09
08	Belgique	91,05	1157,13	9,54	0,01	1138,43	1175,83
	Flandre	134,13	1164,87	11,58	0,01	1142,17	1187,57
	Wallonie	318,29	1161,55	17,84	0,02	1126,58	1196,52
	Bruxelles	1220,60	1104,84	34,94	0,03	1036,36	1173,32
09	Belgique	4755,49	2209,57	68,96	0,03	2074,40	2344,73
	Flandre	11126,55	2337,42	105,48	0,05	2130,68	2544,17
	Wallonie	8077,48	2022,37	89,87	0,04	1846,22	2198,53
	Bruxelles	26030,52	2091,07	161,34	0,08	1774,85	2407,30
10	Belgique	934,99	194,64	30,58	0,16	134,71	254,58
	Flandre	2472,33	250,15	49,72	0,20	152,70	347,61
	Wallonie	454,97	79,26	21,33	0,27	37,45	121,07
	Bruxelles	6957,01	240,90	83,41	0,35	77,42	404,38

COICOP	Géographique	Variance	Consommation	Ecart type	Coefficients de variation	Intervalle de confiance	
						Minimum	Maximum
11	Belgique	1955,13	1560,46	44,22	0,03	1473,80	1647,13
	Flandre	4170,55	1746,94	64,58	0,04	1620,36	1873,51
	Wallonie	4080,07	1245,46	63,88	0,05	1120,26	1370,65
	Bruxelles	15040,25	1507,92	122,63	0,08	1267,55	1748,29
12	Belgique	7288,90	3751,44	85,38	0,02	3584,10	3918,77
	Flandre	9507,12	3710,62	97,50	0,03	3519,51	3901,73
	Wallonie	35708,64	3991,86	188,97	0,05	3621,49	4362,24
	Bruxelles	46091,14	3271,89	214,69	0,07	2851,11	3692,68

Dans le Tableau 40 sont présentés la variance, le coefficient de variation et l'intervalle de confiance pour la consommation générale (ECOICOP 0) et les 12 catégories principales ECOICOP pour la Belgique et les régions.

Pour la Belgique, ces variables par catégorie ECOICOP sont également calculées par groupe d'âge (<30 ans, 30-44 ans, 45-59 ans ou 60+), statut professionnel (ouvrier dans l'industrie et le secteur des services, employé dans l'industrie et le secteur des services, indépendant, chômeur ou 'autre' inactif) et type de ménage (isolé, deux adultes, trois adultes ou plus, parents de familles monoparentales avec enfants à charge, deux adultes avec enfants dépendants ou trois adultes ou plus avec enfants dépendants). Ces données peuvent être obtenues sur demande.

9 Publication, diffusion et points d'attention

9.1 Publication sur le site web

A l'issue de chaque période de collecte de données, les données collectées sont encodées, nettoyées, calibrées, agrégées et contrôlées, de sorte que les résultats puissent être publiés 9 mois après la période de référence.

Les résultats d'EBM sont diffusés sur le site web via des tableaux Excel^{xvii}. Cela concerne:

- Les résultats sur les dépenses par catégorie ECOICOP pour la Belgique et les régions.
- Les résultats sur la possession de biens durables pour la Belgique et les régions.

Les données sont collectées 1 fois tous les 5 ans pour Eurostat (2005 - 2010 - 2015 - 2020 - ...). Eurostat ne publie que quelques années plus tard car la collecte de données n'a pas lieu la même année dans chaque pays.

Les chercheurs peuvent demander des données agrégées (et donc anonymisées). Des microdonnées anonymisées peuvent également être obtenues sous certaines conditions.

9.2 Diffusion des microdonnées disponibles pour la recherche

Il existe deux niveaux de variables dans l'EU-HBS: les variables au niveau du ménage et les variables au niveau individuel. Vous trouverez ci-dessous la liste des variables par niveau dans l'enquête.

Tableau 41: Données au niveau du ménage et au niveau individuel

Données au niveau du ménage	
DONNEES DE BASE	Données de base sur le ménage
LOGEMENT	Type de logement, statut de propriété et caractéristiques du logement
	Équipements du logement
	Coût de l'hébergement
VOITURES	Voiture principale, nombre de voitures et nombre de voitures à disposition de l'employeur
BIENS DURABLES	Possession de biens durables
DÉPENSES	Dépenses des ménages selon la nomenclature ECOICOP d'Eurostat.
	Les quantités sont demandées mais ne sont exploitées que comme variables de contrôle. La qualité est encore insuffisante pour les publier.
Données au niveau individuel	
DONNEES DE BASE	Données individuelles de base
	Données démographiques
REVENUS	Données de base sur le revenu déclaré des ménages

Un manuel des codes fournit des informations plus détaillées sur les fichiers et les variables existants. Il se trouve dans le catalogue des données^{xviii}.

Les microdonnées belges peuvent être demandées sur le site de Statbel via les étapes décrites dans la rubrique 'Microdonnées pour la recherche'^{xix} et les fichiers européens auprès d'Eurostat^{xx}

9.3 Point d'attention: dépenses courantes / saisonnières

Les tableaux Excel montrent par produit ou catégorie ECOICOP les dépenses annuelles moyennes pour un ménage de la population belge de ménages privés, y compris les ménages qui n'ont pas effectué d'achat dans cette catégorie ECOICOP. L'estimation de ces moyennes a été discutée ci-dessus. En raison de l'inclusion des ménages qui n'ont pas enregistré d'achats, certains résultats semblent étranges ou pas directement utilisables. Un problème similaire se présente pour l'estimation de la proportion de ménages qui effectuent un achat dans une certaine catégorie ECOICOP.

Les ménages remplissent quotidiennement un carnet de dépenses pendant un demi-mois. Ce carnet est accompagné d'un questionnaire ménage. Le carnet et le questionnaire ménage reprennent tous les deux tant les dépenses courantes que saisonnières.

L'estimation des dépenses annuelles totales et moyennes sur la base de l'EBM repose sur la supposition que le comportement du ménage pendant une certaine période de référence (d'un demi-mois) reste le même toute l'année, en d'autres termes, que ces dépenses x_i^0 effectuées pendant la période de référence multipliées par 24 fournissent le total x_i de ses dépenses pendant l'année. Pour les dépenses courantes, c'est une hypothèse plausible. En effet, pour certains ménages répondants, x_i sera une bonne approximation des dépenses annuelles totales, alors que pour d'autres ménages, on obtiendra soit une surestimation (si par hasard une dépense supérieure à la moyenne est effectuée pendant la période de référence), soit une sous-estimation (si par hasard une dépense inférieure à la moyenne (ou aucune dépense!) a été effectuée pendant la période de référence) des dépenses annuelles totales x_i . Comme les ménages sont répartis assez uniformément sur les 24 périodes de référence, un nombre suffisamment important de ménages peut être considéré ensemble pour donner une bonne image des dépenses courantes totales et moyennes annuelles. Les dépenses saisonnières sont également traitées officiellement de cette manière.

Les dépenses moyennes de tous les ménages (y compris de ceux qui n'en ont pas effectuées) telles que publiées sur le site web sont la meilleure façon d'estimer les dépenses pour une catégorie de produits particulière. Nous démontrons ci-dessous que cela s'applique également à (certains) produits saisonniers ; d'autre part, nous montrons également que l'estimation des proportions et du nombre de ménages qui effectuent des dépenses pour un produit saisonnier est problématique.

Nous illustrons cela à l'aide des dépenses pour un sapin de Noël (naturel) (catégorie de produits fictive, car un 'sapin de Noël naturel' fait partie de la catégorie ECOICOP 09322F). Supposons:

- H1: que les ménages répondants qui ont acheté un sapin de Noël, ne l'ont fait que pendant la première moitié de décembre;
- H2: que précisément 50 % des ménages répondants ont acheté un sapin de Noël, avec la première moitié de décembre comme période de référence;
- H3: que chacun de ces ménages ont payé $x_i^0 = €50,00$ ($x_i^0 = €0,00$ pour tous les autres ménages répondants).

Des hypothèses de travail supplémentaires, uniquement pour simplifier les calculs ci-dessous, sont

- H4: que tous les ménages répondants ont le même poids calibré ajusté $\tilde{w}_i = \tilde{w}$;
- H5: qu'exactly le même nombre de ménages répondants a été obtenu pour l'ensemble des 24 périodes de référence.

Il ressort immédiatement de l'hypothèse H4 que $N = n\tilde{w}$ (n est le nombre total de ménages répondants; N est le nombre total de ménages dans la population étudiée). Il ressort de l'hypothèse 5 que n est un multiple de 24, que $n/24$ ménages répondants ont la première moitié de décembre comme période de référence, et il ressort des hypothèses H1 et H2 que $n/48$ ménages répondants ont effectivement

acheté un sapin de Noël. La dépense moyenne pour un sapin de Noël pour l'ensemble de la population de ménages privés belges est alors estimée comme suit:

$$\begin{aligned}
 \hat{X} &= \frac{1}{N} \sum_{i=1}^n \tilde{w}_i x_i \\
 &= \frac{1}{n\bar{w}} \sum_{i=1}^n \tilde{w}_i \times (24 \times x_i^0) = \frac{1}{n} \times 24 \times \sum_{i=1}^n x_i^0 \\
 &= \frac{1}{n} \times 24 \times \left(50 \times \frac{n}{48} + 0 \times \frac{47n}{48} \right) \\
 &= 24 \times 50 \times \frac{1}{48} = 25
 \end{aligned}$$

Et la dépense annuelle totale estimée pour les sapins de Noël dans la population belge de ménages privés est de $25N$. Cette moyenne et le total qui en découle sont des estimations parfaitement acceptables, sous les hypothèses H1-H3 (complétées par les hypothèses de travail H4-H5).

L'estimation de la proportion et du nombre de ménages privés belges qui achètent un sapin de Noël est en revanche incertaine si l'on utilise les formules générales. En effet, de nouveau sous les hypothèses H1-H5, la variable dummy x , qui indique si un ménage répondant i achète ou non un sapin de Noël, a la valeur

$$\begin{aligned}
 x_i &= 1 \quad \text{pour } n/48 \text{ ménages qui ont bien enregistré un achat de sapin de Noël} \\
 &= 0 \quad \text{pour } 47n/48 \text{ ménages qui n'ont pas enregistré d'achat de sapin de Noël}
 \end{aligned}$$

Par conséquent, la proportion est estimée comme suit:

$$\begin{aligned}
 \hat{X} &= \frac{1}{N} \sum_{i=1}^n \tilde{w}_i x_i = \frac{1}{n\bar{w}} \sum_{i=1}^n \tilde{w}_i x_i = \frac{1}{n} \sum_{i=1}^n x_i \\
 &= \frac{1}{n} \times \frac{n}{48} = \frac{1}{48} = 2,08\%
 \end{aligned}$$

Et le nombre estimé de ménages privés belges qui achètent un sapin de Noël est de $\frac{1}{48}N$. Cette proportion et le nombre qui en découle sont absolument inacceptables: nous attendons que 50 % de tous les ménages de la population achètent un sapin de Noël, car le 'comportement' à l'égard de l'achat d'un sapin de Noël qui est démontré par les ménages répondants, avec la première moitié de décembre comme période de référence, est supposé pouvoir être projeté aux autres ménages répondants et à l'ensemble de la population.

Enfin, les dépenses annuelles moyennes estimées pour un sapin de Noël pour les ménages qui font effectivement cette dépense, sont égales à $25N / \left(\frac{1}{48}N\right) = 25 \times 48 = \text{€}1.200$. Il s'agit clairement d'une surestimation (nous attendons en effet une moyenne de 50,00€!), ce qui est naturellement la conséquence du fait que le dénominateur $\frac{1}{48}N$ est une nette sous-estimation du nombre de ménages qui achètent effectivement un sapin de Noël.

Conclusion

Nous ne publions pas les dépenses moyennes par ménage avec dépense effective ni la part des ménages avec dépense effective, car bien que cela fournisse des informations intéressantes sur certains biens de consommation quotidiens dont la consommation est répartie équitablement, ce n'est pas le cas pour tous les biens (même l'alimentation peut avoir un caractère saisonnier).

Références

Arrêté royal du 17 janvier 1996.

http://www.ejustice.just.fgov.be/cgi_loi/change_lg.pl?language=fr&la=F&cn=1996011731&table_name=loi

Arrêté royal du 17 mai 2006.

http://www.ejustice.just.fgov.be/cgi/article_body.pl?language=fr&caller=summary&pub_date=06-06-14&numac=2006011243

Statbel (2017a) *Huishoudbudgetonderzoek 2016 – Methodologische Nota*, Statbel, Brussel.

https://statbel.fgov.be/sites/default/files/files/documents/Huishoudens/10.1%20Huishoudbudget/Methodologische%20nota%20HBS%202016%20NL_finaal.pdf

Statbel (2017b) *Enquête sur le budget des ménages 2016 – Note Méthodologique*, Statbel, Brussel.

https://statbel.fgov.be/sites/default/files/files/documents/Huishoudens/10.1%20Huishoudbudget/Note%20methodologique%20HBS%202016%20FR_final.pdf

Statbel (2019a) *Huishoudbudgetonderzoek 2018 – Methodologische Nota*, Statbel, Brussel.

<https://statbel.fgov.be/sites/default/files/files/documents/Huishoudens/10.1%20Huishoudbudget/Methodologische%20nota%202018.pdf>

Statbel (2019b) *Enquête sur le budget des ménages 2018 – Note Méthodologique*, Statbel, Brussel.

<https://statbel.fgov.be/sites/default/files/files/documents/Huishoudens/10.1%20Huishoudbudget/Note%20methodologique%202018.pdf>

Sabbe, K, Delcité, T, Geenens, G. (2021) Wijziging van de duur van het huishoudbudgetonderzoek: analyse van de overgang naar 15 dagen, *Statbel analyse*, no. 15.

https://statbel.fgov.be/sites/default/files/files/documents/Analyse/NL/Analyse_15_Huishoudbudgetonderzoek_STATBEL_v2.pdf

Sabbe, K, Delcité, T, Geenens, G. (2021) Modification de la durée de l'enquête sur le budget des ménages : analyse d'une transition à 15 jours, *Statbel analyse*, no. 15.

https://statbel.fgov.be/sites/default/files/files/documents/Analyse/FR/Analyse_15_Budget%20des%20m%C3%A9nages_STATBEL_v2.pdf

Sabbe, K, Delcité, T, Geenens, G. (2021) Change in the duration of the household budget survey: analysis of a transition to 15 days, *Statbel analyse*, no. 15.

https://statbel.fgov.be/sites/default/files/files/documents/Analyse/EN/Analyse_15_Household%20budget%20survey_STATBEL_v2.pdf

Termote, A. en Depickere, A. (2018a) De hervormde enquête naar de arbeidskrachten in 2017, *Statbel Analyse*, nr. 4.

https://statbel.fgov.be/sites/default/files/Over_Statbel_FR/Analyse_eak_2017_nl_20181220.pdf

Termote, A. et Depickere, A. (2018b) Réforme de l'enquête sur les forces de travail en 2017, *Statbel Analyse*, no. 4.

https://statbel.fgov.be/sites/default/files/Over_Statbel_FR/Analyse_eak_2017_fr_20181220.pdf

Särndal, C.-E., Swensson, B. and Wretman, J. (1992) *Model Assisted Survey Sampling*, New-York: Springer.

Références aux pages web pertinentes

i Page FAQ d'EBM:

<https://statbel.fgov.be/fr/ebm-faq>

ii Carnet des dépenses en ligne

<https://hbs.statdata.be/hbs/login.xhtml;jsessionid=59EB4B588AE1EA4F5B93FE1BC87F9728.worker12>

iii Carnet des dépenses papier:

https://statbel.fgov.be/sites/default/files/files/documents/Huishoudens/10.1%20Huishoudbudget/carnet2016_FR.pdf

iv Questionnaire ménage:

https://statbel.fgov.be/sites/default/files/files/documents/Huishoudens/10.1%20Huishoudbudget/Questionnaire_m%C3%A9nage_2018.pdf

v Questionnaire ménage:

https://statbel.fgov.be/sites/default/files/files/documents/Huishoudens/10.1%20Huishoudbudget/Questionnaire_m%C3%A9nage_2018.pdf

vi Eurostat Statistics Explained: Glossaire: Ménage – statistiques sociales

https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Household_social_statistics/fr

vii Eurostat Statistics Explained: Glossaire: Revenu disponible équivalent

https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Equivalised_disposable_income/fr

viii Adjusting household incomes: equivalence scales:

<http://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf>

ix EU-SILC (European Union Statistics on Income and Living Conditions):

<https://statbel.fgov.be/fr/themes/menages/pauvrete-et-conditions-de-vie>

x Questionnaire ménage:

https://statbel.fgov.be/sites/default/files/files/documents/Huishoudens/10.1%20Huishoudbudget/Questionnaire_m%C3%A9nage_2018.pdf

xi Household Budget Surveys in the EU: Methodology and recommendations for harmonisation – 2003:

<https://ec.europa.eu/eurostat/documents/3859598/5875361/KS-BF-03-003-EN.PDF/42a95cc0-cb48-48c7-8d3a-dfc5fa265eff?version=1.0>

xii EUROSTAT RAMON - Reference And Management Of Nomenclatures – European Classification of Individual Consumption according to Purpose (ECOICOP):

https://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_CLS_DLD&StrNom=COICOP_5&StrLanguageCode=FR&StrLayoutCode=HIERARCHIC

xiii EUROSTAT RAMON - Reference And Management Of Nomenclatures – Metadata European Classification of Individual Consumption according to Purpose (ECOICOP):

https://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=COICOP_5

xiv Banque Nationale de Belgique: Comptes nationaux/régionaux:

<https://www.nbb.be/fr/statistiques/comptes-nationauxregionaux>

xv Statbel – Indice des prix à la consommation:

<https://statbel.fgov.be/fr/themes/prix-la-consommation/indice-des-prix-la-consommation>

^{xvi} United Nations Statistical Division – COICOP revision:

https://unstats.un.org/unsd/class/revisions/coicop_revision.asp

^{xvii} Chiffres de l'enquête sur le budget des ménages:

<https://statbel.fgov.be/fr/themes/menages/budget-des-menages/plus>

^{xviii} Catalogue des données EBM pour demander les microdonnées pour la recherche:

<https://statbel.fgov.be/fr/propos-de-statbel/vie-privee/microdonnees-pour-la-recherche/catalogue-des-donnees>

^{xix} Demander des microdonnées pour la recherche

<https://statbel.fgov.be/fr/propos-de-statbel/vie-privee/microdonnees-pour-la-recherche>

^{xx} Microdonnées européennes pour EBM 2005, 2010 et 2015:

<https://ec.europa.eu/eurostat/fr/web/microdata/household-budget-survey>