

**Rapport intermédiaire
sur la qualité des données
françaises
EU-SILC 2008**

TABLE DES MATIÈRES

Table des matiÈres	2
1 Indicateurs transversaux communs de l'Union européenne	4
1.1 Indicateurs communautaires transversaux communs fondés sur la dimension transversale des statistiques EU-SILC	4
1.1.1 Taux de pauvreté après transferts sociaux	4
1.1.2 Inégalité de la distribution des revenus Q80/Q20, ratio interquintile	9
1.1.3 Intensité de la pauvreté	9
1.1.4 Dispersion autour du seuil de pauvreté	9
1.1.5 Taux de pauvreté ancré à une date fixée (2005).....	10
1.1.6 Taux de pauvreté avant transferts	11
1.1.7 Inégalité de la distribution des revenus : coefficient de Gini	11
1.2 Autres indicateurs	11
1.2.1 Niveau de vie moyen	11
1.2.2 Écart de rémunération entre les sexes non ajusté	11
2 PrÉcision	12
2.1 Plan d'échantillonnage.....	12
2.1.1 Type de plan d'échantillonnage.....	12
2.1.2 Unités d'échantillonnage.....	12
2.1.3 Critères de stratification et de sous-stratification.....	12
2.1.4 Taille de l'échantillon et critères d'attribution.....	13
2.1.5 Systèmes de sélection de l'échantillon.....	13
2.1.6 Répartition de l'échantillon au cours du temps et renouvellement de l'échantillon : groupes de rotation	13
2.1.7 Pondération.....	13
2.1.8 Substitutions	19
2.2 Erreurs d'échantillonnage	20
2.3 Erreurs non dues à l'échantillonnage	23
2.3.1 Erreurs dans la base de sondage et la couverture.....	23
2.3.2 Erreurs de mesure et de traitements	24
2.3.3 Erreurs de non-réponse.....	27
2.4 Mode de collecte des données	35
2.5 Durée de l'entretien.....	36
2.6 Les imputations.....	36
3 Comparabilité	41
3.1 Concepts et définitions de base.....	41
3.2 Composantes du revenu.....	42
3.2.1 Les différences entre les définitions nationales et les définitions type des statistiques EU-SILC	42

3.2.2	La source ou la procédure utilisée pour la collecte des variables du revenu.....	45
3.2.3	La forme sous laquelle ont été obtenues les variables du revenu au niveau des composantes.....	45
3.2.4	La méthode utilisée pour obtenir les variables cibles du revenu sous la forme requise (c'est-à dire en tant que valeur brute).....	45
4	CohÉrence	47
4.1	Comparaison des variables cibles avec des sources extérieures	47
4.1.1	Salaire et revenu assimilé (PY010N).....	47
4.1.2	Les allocations chômage (PY090N)	48
4.1.3	Retraites de régime privé (PY080N), allocations vieillesse (PY100N) et pension au survivant (PY110N).....	49
4.1.4	Bénéfices en espèces ou perte de trésorerie en rapport avec une activité indépendante 49	
4.1.5	Revenus sociaux (HY050G, HY060G et HY070G),	50
4.1.6	Prestations familiales (HY050N),.....	50
4.1.7	Impôts (impôt sur le revenu des personnes physiques et taxe d'habitation).	51
4.1.8	Loyers	51
4.1.9	Intérêts, dividendes, bénéfices tirés d'investissements en capital dans les opérations incorporelles (HY090N).....	51
4.1.10	Revenu disponible du ménage et niveau de vie des individus.....	52

1 INDICATEURS TRANSVERSAUX COMMUNS DE L'UNION EUROPÉENNE

Les indicateurs ont été estimés grâce aux programmes fournis par Eurostat.

1.1 Indicateurs communautaires transversaux communs fondés sur la dimension transversale des statistiques EU-SILC

1.1.1 Taux de pauvreté après transferts sociaux

Le taux de pauvreté après transferts sociaux est défini comme la proportion de personnes ayant un niveau de vie inférieur au seuil de pauvreté. Tous les individus d'un ménage ont le même niveau de vie, qui est égal au revenu disponible divisé par le nombre d'unités de consommation du ménage. Le seuil de pauvreté est fixé à 60 % de la médiane du niveau de vie des individus.

A partir des données de revenus 2007, soit pour la collecte ayant eu lieu en 2008, les données d'enquête ont été appariées avec les fichiers administratifs et sociaux. Auparavant, les données de revenu étaient collectées par voie d'enquête. Le changement de collecte des revenus permet une amélioration de la qualité des données, un allègement de la charge des enquêtes et une meilleure cohérence avec l'enquête de référence à l'INSEE sur les revenus, soit l'enquête sur les revenus fiscaux et sociaux ERFS (enquête emploi appariée avec les fichiers fiscaux et sociaux).

Ce changement de collecte permet de disposer en particulier d'une meilleure information sur les revenus du patrimoine et foncier qui étaient mal recensés par voie de collecte et une meilleure qualité globale pour les autres revenus. L'impact de l'appariement avec la source fiscale avait été évalué à 0,2 point sur le taux global de pauvreté (cf. Comparability Study, Laurence Dauphin) mais celui de l'appariement avec les données des organismes gestionnaires de prestations sociales (Caisse nationale d'allocations familiales, Caisse nationale d'allocation vieillesse, Caisse Centrale de Mutualité Sociale Agricole) n'a pas pu être évalué.

1.1.1.1 Taux de pauvreté selon l'âge et le sexe

	Total			0-17	0 - 64			18+			18 - 64		
	Total	F	H	Total	Total	F	H	Total	F	H	Total	F	H
Pauvres	13.4	14.0	12.7	17.4	13.8	14.5	13.2	12.3	12.8	11.7	12.6	13.1	12.1
Non pauvres	86.6	86.0	87.3	83.6	86.2	85.5	86.8	87.7	87.2	88.3	87.3	86.9	87.9

	18 - 24			25 - 49			50 - 64			65+		
	Total	F	H	Total	F	H	Total	F	H	Total	F	H
Pauvres	20.5	20.2	20.7	11.6	12.7	10.4	10.7	10.5	11.0	11.1	11.8	10.2
Non pauvres	79.5	79.8	79.3	88.4	87.3	89.6	89.3	89.5	89.0	88.9	88.2	89.8

Effectifs non pondérés

	Total			0 - 17	0 - 64			18+		
	Total	F	H	Total	Total	F	H	Total	F	H
Total	25510	13187	12323	6122	21566	10989	10577	19388	10170	9218
Pauvres	3314	1793	1521	1018	2894	1541	1353	2296	1261	1035
Non pauvres	22196	11394	10802	5104	18672	9448	9224	17092	8909	8183

	18 - 64			18 - 24			25 - 49			50 - 64		
	Total	F	H									
Total	15444	7972	7472	2306	1136	1170	8229	4290	3545	4908	2546	2362
Pauvres	1876	1009	867	438	216	222	929	535	394	509	258	251
Non pauvres	13568	6963	6605	1868	920	948	7300	3755	3939	4399	2288	2111

	65+		
	Total	F	H
Total	3944	2198	1746
Pauvres	420	252	168
Non pauvres	3524	1946	1578

Répartition de la population totale selon le sexe¹

Total	F	H
100,0	51.5	48.5

Répartition de la population totale selon l'âge et le sexe

	0 - 17	0 - 64	18+	18 - 64	18 - 24	25 - 49	50 - 64	65+	Total
Total	21.5	83.5	78.5	62,0	=8.8	34.4	18.8	16.5	100,0
F	20,4	81,5	79,6	61,0	8,4	33,8	18,9	18,6	100,0
H	22,6	85,6	77,4	63,0	9,3	35,1	18,6	14,4	100,0

Répartition de la population pauvre selon le sexe

Total	F	H
100,0	53,8	46,2

Répartition de la population pauvre selon l'âge et le sexe

	0 - 17	0 - 64	18+	18 - 64	18 - 24	25 - 49	50 - 64	65+	Total
Total	27.9	86.2	72.0	58.3	13.5	29.8	15.0	13.7	100,0
F	27,4	84,3	72,6	56,9	12,1	30,7	14,1	15,7	100,0
H	28,6	88,5	71,4	59,9	15,1	28,7	16,1	11,5	100,0

1.1.1.2 Taux de pauvreté selon la situation professionnelle vis à vis de l'emploi la plus fréquente et le sexe

	Total			En emploi			Sans emploi : total			Sans emploi : Chômeurs		
	Total	F	H	Total	F	H	Total	F	H	Total	F	H
Pauvres	12.0	12.6	11.3	6.8	6.3	7.3	17.9	18.7	16.9	39.2	36.3	42.1
Non pauvres	88.0	87.4	88.7	93.2	93.7	92.7	82.1	81.3	83.1	60.8	63.7	57.9

	Sans emploi : retraités			Sans emploi : autres inactifs		
	Total	F	H	Total	F	H
Pauvres	9.5	9.5	9.5	27.4	27.3	27.7
Non pauvres	90.5	90.5	90.5	73.6	73.7	72.3

Sont exclus du champ les personnes de moins de 18 ans au 1^{er} janvier de l'année de l'enquête et les individus qui n'ont pas gardé le même statut d'occupation pendant plus de la moitié de l'année.

¹ Les tableaux en % sont pondérés avec la variable rb050

Effectifs non pondérés

	Total			En emploi			Sans emploi : total			Sans emploi : chômeurs		
	Total	F	H	Total	F	H	Total	F	H	Total	F	H
Total	18 998	9 979	9 019	10 233	4 927	5 306	8 765	5 252	3 713	897	463	433
Pauvres	2197	1225	972	646	282	364	1 551	943	608	337	154	183
Non pauvres	16 801	8 754	8 047	9 587	4 645	4 942	7 214	4 309	3 105	560	309	300

	Sans emploi : retraités			Sans emploi : autres inactifs		
	Total	F	H	Total	F	H
Total	4 909	2 469	2 440	2 959	2 119	840
Pauvres	440	215	225	774	574	200
Non pauvres	4 469	2254	2 215	2185	1 545	640

Répartition de la population totale selon la situation professionnelle la plus fréquente et le sexe

	Total	En emploi	Sans emploi : total	Sans emploi : chômeurs	Sans emploi : retraités	Sans emploi : autres inactifs
F	100,0	47,5	52,5	4,6	26,3	21,3
H	100,0	56,7	43,3	5,0	27,7	10,6
Total	100,0	51.9	48.1	4.8	27.0	16.2

Répartition de la population pauvre selon le statut professionnel le plus fréquent et le sexe

	Total	En emploi	Sans emploi : total	Sans emploi : chômeurs	Sans emploi : retraités	Sans emploi : autres inactifs
F	100,0	23,4	76,6	13,2	19,8	43,8
H	100,0	36,1	63,9	18,3	22,9	22,7
Total	100,0	29.2	70.8	15.5	21.2	34.2

1.1.1.3 Taux de pauvreté selon le type de ménage

	Sans enfant : total	Personnes seules (total)	2 adultes, moins de 65 ans	2 adultes, dont au moins un de 65 ans ou +	Sans enfant : autres
Pauvres	10.9	15.9	7.6	8.5	11.4
Non pauvres	89,1	84,1	92,4	91,5	88,6

	Avec enfant : total	Familles monoparentales	2 adultes, 1 enfant	2 adultes, 2 enfants	2 adultes, 3 enfants et +	Avec enfants : autres
Pauvres	15.7	26.6	10.2	8.7	23.0	24.0
Non pauvres	84,3	73,4	89,8	91,3	77	76

Effectifs non pondérés

	Sans enfant : total	Personnes seules (total)	2 adultes, moins de 65 ans	2 adultes, dont au moins un de 65 ans ou +	Sans enfant : autres
Total	10950	2865	3716	2814	1555
Pauvres	1146	468	294	224	160
Non pauvres	9804	2397	3422	2590	1395

Total	Avec enfant : total	Familles monoparentales	2 adultes, 1 enfant	2 adultes, 2 enfants	2 adultes, 3 enfants et +	Avec enfants : autres
Total	14499	1536	2949	5100	3400	1514
Pauvres	2150	402	264	432	709	343
Non pauvres	12349	1134	2685	4668	2691	1171

Taux de pauvreté des personnes seules selon l'âge ou le sexe

	F	H	< 65	65+
Pauvres	16.0	15.8	16.8	14.5
Non pauvres	84,0	84,2	83,2	85,5

Effectifs non pondérés

	F	H	< 65	65+
Total	1750	1115	1615	1250
Pauvres	283	185	284	184
Non pauvres	1467	930	1331	1066

Répartition de la population totale selon le type de ménage

Sans enfant : total	Personnes seules (total)	2 adultes, moins de 65 ans	2 adultes, dont au moins un de 65 ans ou +	Sans enfant : autres
48.6	15.4	15.8	11.0	6.4

Avec enfants : total	Familles monoparentales	2 adultes, 1 enfant	2 adultes, 2 enfants	2 adultes, 3 enfants et +	Avec enfants : autres
51.4	5.3	11.5	17.7	11.1	5.6

Répartition des personnes seules selon l'âge ou le sexe

F	H	< 65	65+
57.0	43.0	54.2	45.8

Répartition de la population pauvre selon le type de ménage

Sans enfant : total	Personnes seules (total)	2 adultes, moins de 65 ans	2 adultes, dont au moins un de 65 ans ou +	Sans enfant : autres
39.7	18.3	8.9	7.0	5.5

Avec enfants : total	Familles monoparentales	2 adultes, 1 enfant	2 adultes, 2 enfants	2 adultes, 3 enfants et +	Avec enfants : autres
60.3	10.6	8.7	11.5	19.1	10.4

Répartition des personnes seules pauvres selon l'âge ou le sexe

F	H	< 65	65+
57.2	42.8	53.0	47.0

1.1.1.4 Taux de pauvreté selon le statut d'occupation du logement

Les propriétaires ont un risque de pauvreté plus faible que les locataires même quand le taux de pauvreté est estimé comme ici sans prendre en compte les loyers imputés.

	Total	Locataires	Propriétaires ou logés à titre gratuit
Pauvres	13.4	22.4	8.6
Non pauvres	86.6	77.6	91.4

Effectifs non pondérés

	Total	Locataires	Propriétaires ou logés à titre gratuit
Total	25510	7962	17548
Pauvres	3314	1887	1427
Non pauvres	22196	6075	16121

Répartition de la population totale selon le statut d'occupation du logement

Total	Locataires	Propriétaires ou logés à titre gratuit
100,0	34.5	65.5

Répartition de la population pauvre selon le statut d'occupation du logement

Total	Locataires	Propriétaires ou logés à titre gratuit
100,0	42.2	57.8

1.1.1.5 Taux de pauvreté selon le type de ménage et l'intensité d'emploi

L'intensité d'emploi dans le ménage (W) mesure la part du temps passé en emploi pour les personnes du ménage en âge de travailler (entre 18 et 64 ans) sur une période donnée. Il s'agit précisément du rapport entre le nombre de mois travaillés et le nombre de mois travaillables (mois travaillés, au chômage, en retraite ou en inactivité) pour l'ensemble des personnes du ménage. L'intensité d'emploi est donc comprise entre 0 et 1. Une valeur de 0 correspond à l'absence d'emploi au sein du ménage sur la période, une valeur de 1 à un nombre de mois travaillés égal au nombre de mois travaillables (emploi saturé).

	Sans enfant W=0	Sans enfant 0<W<1	Sans enfant W=1	Avec enfants W=0	Avec enfants 0<W<0,5	Avec enfants 0,5<W<1	Avec enfants W=1
Pauvres	20.4	10.0	4.9	71.7	53.3	19.9	5.1
Non pauvres	79.6	90.0	95.1	28.3	46.7	80.1	94.9

Sont exclus du champ les ménages composés seulement de personnes de moins de 18 ans, de plus de 65 ans ou n'ayant pas travaillé au moins un mois au cours de l'année de référence ou d'étudiants.

Effectifs non pondérés

	Sans enfant W=0	Sans enfant 0<W<1	Sans enfant W=1	Avec enfants W=0	Avec enfants 0<W<0,5	Avec enfants 0,5<W<1	Avec enfants W=1
Total	1757	2681	3190	772	692	4652	8413
Pauvres	351	253	144	540	391	847	374
Non pauvres	1406	2428	3046	232	301	3805	8039

Répartition de la population totale selon le type de ménage et l'intensité d'emploi

Total	Sans enfant W=0	Sans enfant 0<W<1	Sans enfant W=1	Avec enfants W=0	Avec enfants 0<W<0,5	Avec enfants 0,5<W<1	Avec enfants W=1
100,0	9.3	13.4	17.5	3.0	2.9	19.6	34.3

Répartition de la population pauvre selon le type de ménage et l'intensité d'emploi

Total	Sans enfant W=0	Sans enfant 0<W<1	Sans enfant W=1	Avec enfants W=0	Avec enfants 0<W<0,5	Avec enfants 0,5<W<1	Avec enfants W=1
100,0	14.2	10.0	6.4	16.0	11.5	28.8	13.1

1.1.1.6 Seuil de pauvreté (valeurs indicatives)

En 2007, le seuil de pauvreté s'élève à 10 540 euros par an. Une personne seule est pauvre si elle a un revenu disponible inférieur à 878 euros par mois et un couple avec deux enfants de moins de 14 ans, s'il a un revenu disponible inférieur à 22 134 euros (soit 1 845 euros par mois).

1.1.2 Inégalité de la distribution des revenus Q80/Q20, ratio interquintile

Le rapport interquintile est le rapport entre la somme des niveaux de vie des 20 % de personnes ayant les niveaux de vie les plus élevés et celle des 20 % de personnes ayant les niveaux de vie les plus faibles.

Le cinquième de la population le plus riche possède 4,2 fois plus de niveaux de vie, en masse, que le cinquième de la population le plus pauvre.

1.1.3 Intensité de la pauvreté

L'intensité de la pauvreté est estimée comme l'écart relatif de la médiane des niveaux de vie de la population pauvre au seuil de pauvreté. Si les pauvres n'avaient aucun revenu, cet écart serait de 100 %. S'ils avaient tous un niveau de vie très proche du seuil de pauvreté, l'écart serait presque nul.

Total			0 - 17	18+			18 - 64			65+		
Total	F	H	Total	Total	F	H	Total	F	H	Total	F	H
18.1	17.2	18.7	15.0	19.4	18.6	20.6	21.2	20.2	22.9	12.1	12.9	11.4

Effectifs non pondérés

Total			0 - 17	18+			18 - 64			65+		
Total	F	H	Total	Total	F	H	Total	F	H	Total	F	H
3314	1793	1521	1018	2296	1261	1035	1876	1009	867	420	252	168

1.1.4 Dispersion autour du seuil de pauvreté

Le seuil de pauvreté pourrait aussi être fixé à 40, 50 et 70 % de la médiane du niveau de vie, L'étude des taux de pauvreté ainsi obtenus permet d'apprécier la concentration des niveaux de vie autour du seuil de pauvreté.

	Total	F	H
<i>40 % de la médiane</i>			
Pauvres	3.4	3.5	3.3
Non pauvres	96.6	96.5	96.7
<i>50 % de la médiane</i>			
Pauvres	7.1	7.1	7.1
Non pauvres	92.9	92.9	92.9
<i>70 % de la médiane</i>			
Pauvres	21.2	21.8	20.5
Non pauvres	78.8	78.2	79.5

Effectifs non pondérés

	total	F	H
Total	25510	13187	12323
<i>40 % de la médiane</i>			
Pauvres	833	446	387
Non pauvres	24677	12741	11936
<i>50 % de la médiane</i>			
Pauvres	1793	952	841
Non pauvres	23717	12235	11482
<i>70 % de la médiane</i>			
Pauvres	5214	2795	2419
Non pauvres	20296	10392	9904

1.1.5 Taux de pauvreté ancré à une date fixée (2005)

Il s'agit du taux de pauvreté calculé en 2007, en prenant comme seuil celui de 2005, corrigé de l'inflation entre ces deux dates. Compte tenu de la rupture de série sur les revenus 2007, cet indicateur n'est pas pertinent.

	Total			0-17	0 - 64			18+			18 - 64		
	Total	F	H	Total	Total	F	H	Total	F	H	Total	F	H
Pauvres	12.0	12.6	11.5	15.6	12.6	13.2	12.0	11.1	11.5	10.6	11.6	11.9	11.2
Non pauvres	88.0	87.4	88.5	84.4	87.4	86.8	88.0	88.9	88.5	89.4	88.4	88.1	88.8

	18 - 24			25 - 49			50 - 64			65+		
	Total	F	H	Total	F	H	Total	F	H	Total	F	H
Pauvres	19.4	19.1	19.7	10.5	11.4	9.5	9.9	9.6	10.4	9.2	10.1	8.0
Non pauvres	80.6	80.9	80.3	89.5	88.6	90.5	90.1	90.4	89.6	90.8	89.9	92.0

Effectifs non pondérés

	Total			0-17	0 - 64			18+			18 - 64		
	Total	F	H	Total	Total	F	H	Total	F	H	Total	F	H
Pauvres	3002	1622	1380	918	2650	1407	1243	2084	1138	946	1732	923	809
Non pauvres	22508	11565	10943	5204	21566	10989	9334	17304	9032	8272	13712	7049	6663

	18 - 24			25 - 49			50 - 64			65+		
	Total	F	H	Total	F	H	Total	F	H	Total	F	H
Pauvres	414	204	210	846	484	362	472	235	237	352	215	137
Non pauvres	1892	932	960	7384	3806	3578	4436	2311	2125	3592	1983	1609

1.1.6 Taux de pauvreté avant transferts

Les transferts considérés sont les prestations sociales, y compris les revenus de remplacement (chômage, retraites). Le seuil de pauvreté est celui estimé à partir du revenu disponible après transferts.

1.1.6.1 Taux de pauvreté avant tout transfert social selon l'âge et le sexe

	Total			0 - 17	18+			18 - 64			65+		
	Total	F	H	Total	Total	F	H	Total	F	H	Total	F	H
Pauvres	44.7	46.9	42.4	35.7	47.2	49.6	44.6	34.8	36.1	33.5	93.6	94.1	92.8
Non pauvres	55,3	53,1	57,6	64,3	52,8	50,4	55,4	65,2	63,9	66,5	6,4	5,9	7,2

Effectifs non pondérés

	Total			0 - 17	18+			18 - 64			65+		
	Total	F	H	Total	Total	F	H	Total	F	H	Total	F	H
Total	25510	13187	12323	6122	19388	10170	9218	15444	7972	7472	3944	2198	1746
Pauvres	10867	5890	4977	2112	8755	4837	3918	5090	2780	2310	3665	2057	1608
Non pauvres	14643	7297	7346	4010	10633	5333	5300	10354	5192	5162	279	141	138

1.1.6.2 Taux de pauvreté avant transferts sociaux autres que retraites selon l'âge et le sexe

	Total			0 - 17	18+			18 - 64			65+		
	Total	F	H	Total	Total	F	H	Total	F	H	Total	F	H
Pauvres	23.1	23.5	22.8	34.2	20.1	20.6	19.6	21.9	22.6	21.2	13.4	14.2	12.4
Non pauvres	76,9	76,5	77,2	65,8	79,9	79,4	80,4	78,1	77,4	78,8	86,6	85,8	87,6

Effectifs non pondérés

	Total			0 - 17	18+			18 - 64			65+		
	Total	F	H	Total	Total	F	H	Total	F	H	Total	F	H
Total	25510	13187	12323	6122	19388	10170	9218	15444	7972	7472	3944	2198	1746
Pauvres	5838	3099	2739	2031	3807	2087	1720	3320	1795	1525	487	292	195
Non pauvres	19672	10088	9584	4091	15581	8083	7498	12124	6177	5947	3457	1906	1551

1.1.7 Inégalité de la distribution des revenus : coefficient de Gini

Le coefficient de Gini permet de mesurer l'écart de la distribution des revenus à une distribution purement égalitaire. Il s'élève à 28.1 % en 2007.

1.2 Autres indicateurs

1.2.1 Niveau de vie moyen

En 2007, le niveau de vie des individus s'élève en moyenne à 19 995 euros par unité de consommation, soit 1 666 euros par mois.

1.2.2 Écart de rémunération entre les sexes non ajusté

Ils sont fournis par l'enquête Emploi qui dispose d'un plus grand échantillon et de données plus précises sur le temps de travail.

2 PRÉCISION

2.1 Plan d'échantillonnage

2.1.1 Type de plan d'échantillonnage

Le système d'échantillonnage de l'enquête SILC en France repose sur un échantillon-maître. Ce dernier constitue par définition une réserve de logements regroupant, d'une part des logements recensés en mars 1999, d'autre part des logements achevés après cette date, appelés « logements neufs ».

L'unité d'échantillonnage gérée dans ces bases de sondage est le logement. Ce qui se passe au sein du logement relève du terrain : en l'occurrence, dans cette enquête, lorsque l'échantillon est entrant, on interroge l'ensemble des individus de chaque ménage présent dans le logement. L'échantillonnage de logements est stratifié, à plusieurs degrés : selon les strates, il y a deux ou trois degrés de tirage. Dans tout échantillon entrant, puisqu'on est exhaustif au dernier degré de sondage, on peut dire qu'il s'agit d'un sondage d'individus en grappes.

2.1.2 Unités d'échantillonnage

Au premier degré, l'unité d'échantillonnage est un groupe de communes contiguës. En zone rurale, ce groupe correspond à un découpage ad-hoc. En revanche, en zone urbaine l'unité primaire est, soit un ensemble d'unités urbaines, soit une unité urbaine (l'unité urbaine est une entité regroupant des communes, définie à partir des données du recensement). Au second degré, l'unité secondaire est, soit un groupe ad-hoc de communes lorsqu'on se situe en zone rurale, soit un logement si on est en zone urbaine. Le troisième degré n'existe qu'en zone rurale, et les unités de tirage correspondantes sont des logements.

2.1.3 Critères de stratification et de sous-stratification

Le partage rural/urbain constitue une première stratification préalable au premier degré de tirage. Elle est enrichie de deux manières. D'une part, la strate urbaine distingue en fait trois niveaux :

- unités urbaines de moins de 20 000 habitants (« petit urbain »),
- unités urbaines comprises entre 20 000 et 100 000 habitants (« moyen urbain »),
- unités urbaines de plus de 100 000 habitants (« grand urbain »),

D'autre part, chaque strate ainsi définie -y compris la strate rurale- est elle-même ventilée en sous-strates selon la région administrative. Comme il y a 22 régions administratives, la stratification préalable comporte 88 strates (dont 2 sont vides).

En zone rurale, il n'y a pas de stratification préalable au second degré. En revanche, dans les strates urbaines, il y a une sous-stratification constituée par des groupes de communes. Il n'y a pas de critères sociodémographiques utilisés pour définir ces groupes de communes, on peut seulement dire qu'il s'agit de communes contiguës regroupées afin d'atteindre une certaine taille - mais cela permet au moins de distinguer ville-centre et banlieue.

Enfin, l'ultime stratification est préalable au tirage des logements dans les groupes de communes tirés. Elle distingue quatre types de logements :

- les logements recensés de type « résidence principale » ;
- les logements recensés de type « résidence secondaire » ;
- les logements recensés de type « résidence vacante » ;
- les logements neufs,

2.1.4 Taille de l'échantillon et critères d'attribution

L'échantillon 2004 comprenait 16 000 logements (ventilés en 9 sous-échantillons), et les sous-échantillons entrant des années 2005 à 2008 en comprenait chacun 3 000. Les sous-échantillons entrants des années à venir devraient être également calibrés à 3 000 logements.

2.1.5 Systèmes de sélection de l'échantillon

L'échantillon de logements de SILC est tiré à probabilités égales (taux f) à l'exception des logements recensés secondaires qui sont tirés avec la probabilité $f/4$ et des logements recensés vacants en strate rurale qui sont tirés au taux $f/2$.

Cas de la strate rurale

Le tirage des 128 unités primaires de l'échantillon-maître parmi les 227 recensées a été effectué proportionnellement à la taille définie en nombre de résidences principales, selon un algorithme de tirage équilibré (l'équilibrage a porté sur des structures par âge et sur le revenu déclaré aux services fiscaux). Le tirage de 719 unités secondaires de l'échantillon de EU-SILC sur les 3 217 unités de l'échantillon-maître a également été proportionnel à la taille en nombre de résidences principales. Le tirage des logements est effectué au sein du groupe de communes par un algorithme systématique sur fichier trié.

Le choix des tailles des échantillons d'unités primaires et secondaires est dicté par des considérations de coût de déplacement et de charge moyenne des enquêteurs, sachant que la règle générale consiste à affecter un enquêteur à une unité primaire.

Cas des strates urbaines jusqu'à 100 000 habitants

Dans le petit urbain, l'échantillon-maître comprend 75 unités primaires parmi les 975 recensées et dans le moyen urbain, 93 unités primaires parmi 180.

L'échantillonnage des unités primaires s'effectue selon les mêmes modalités qu'en strate rurale. Même chose pour le tirage des logements au sein des groupes de communes. Les tailles d'échantillons sont déterminées par les mêmes règles qu'en strate rurale.

Cas du « grand urbain »

Les 53 unités urbaines de plus de 100 000 habitants sont retenues. Dans chaque groupe de communes de l'unité urbaine, le tirage des logements est effectué par un algorithme systématique sur fichier trié. La taille de l'échantillon découle directement de la probabilité de sélection des logements.

2.1.6 Répartition de l'échantillon au cours du temps

L'échantillon est de type rotatif : chaque année, on interroge neuf sous-échantillons, tirés chacun selon la méthode exposée ci-dessus. Chaque sous-échantillon est donc un panel interrogé neuf années de suite. Les tirages des sous-échantillons sont indépendants d'une année sur l'autre mais, par construction, ils sont toujours effectués dans les mêmes unités primaires puisque celles-ci constituent un échantillon-maître.

2.1.7 Renouvellement de l'échantillon : groupes de rotation

Pour la collecte de 2008, l'échantillon entrant était constitué de 3 000 logements tirés de la façon décrite plus haut. L'échantillon des réinterrogés était constitué des logements où habitaient en 2007 au moins une personne répondante et n'appartenant pas à l'échantillon sortant. Il comprenait 10 455 logements.

2.1.8 Pondération

La pondération se fait au niveau logement.

En 2008, nous avons estimé pour la quatrième année des poids pour un échantillon comprenant à la fois des entrants et des réinterrogés. Nous avons suivi la méthodologie présentée par Pascal Ardilly et Pierre Lavallée (Symposium 2003, Stat Canada).

2.1.8.1 Effet du plan d'échantillonnage (entrants)

Les poids sont définis au niveau logement et sont égaux à l'inverse de la probabilité d'inclusion qui découle du plan de sondage. Les ménages et les individus ont la même pondération que le logement auquel ils appartiennent. Chaque année, les pondérations initiales, notées WI, varient dans un rapport de 1 à 4 puisque les logements recensés secondaires ont quatre fois moins de chance d'être tirés que les logements recensés principaux ou que les logements neufs (cf. 2.1.5).

Si l'on regarde l'ensemble des sous-échantillon, le poids WI varie dans un rapport de 1 à 7 en raison de la différence de taille entre l'échantillon de 2004 et les échantillons suivants.

2.1.8.2 Ajustements pour non-réponse

Pour les réinterrogés, il y a deux types de non-réponse totale : la non-réponse des individus au moment du tirage « initial », c'est-à-dire lorsqu'on échantillonne le panel entrant, et la non-réponse à la date courante. Les modèles de non-réponse ont été estimés pour les deux types de non-réponse.

La probabilité de non-réponse initiale est notée PI.

La correction de la non-réponse initiale

La correction de la non-réponse initiale pour les entrants en 2008

Les modèles de non-réponse ont été estimés sur les logements de l'échantillon qui faisaient partie du champ en 2008 : les enquêteurs les ont reconnus comme résidence habituelle d'au moins une personne. Nous avons différencié trois types de modèles de non-réponse totale selon les informations disponibles dans les bases de sondage ou recueillies par l'enquêteur en 2008 :

- les résidences principales au dernier recensement de la population (RP) et en 2008,
- les résidences recensées secondaires, occasionnelles ou vacantes au RP,
- les logements neufs (construits après mars 1999).

Toutes catégories de logements confondues, les motifs principaux de non-réponse sont liés à :

- la région,
- le fait d'habiter une maison ou un appartement en 2008,
- le nombre de pièces du logement en 1999,
- l'âge de la personne de référence (plus ou moins de 60 ans) en 1999,

À partir de ces résultats, nous avons formé différentes sous-populations de répondants en croisant les motifs significatifs de la non-réponse (certains groupes à faible effectif ont été regroupés pour assurer la robustesse des résultats). Au final, nous avons réparti les ménages répondants dans 11 groupes et nous avons supposé le mécanisme de réponse homogène à l'intérieur de sous-populations. La probabilité de réponse initiale d'un ménage dans une sous-population donnée vaut donc :

$$PI = \frac{\text{Nombre de ménages répondants dans la sous-population}}{\text{Nombre de ménages échantillonnés dans la sous population}}$$

La correction de la non-réponse initiale pour les réinterrogés

La probabilité de non-réponse initiale a été estimée de la même façon sur les échantillons entrants en 2004, 2005, 2006 et 2007 : le nombre de groupe et les motifs principaux de non réponse variant à la marge. Par ailleurs en 2004, compte tenu de la taille d'échantillon plus grande (tous les échantillons étaient entrants, ce sont cinq modèles de non-réponses qui ont été estimés et non trois (cf. Rapport final sur la qualité des données françaises EU-SILC 2006).

Pour l'ensemble des ménages, PI est comprise entre 0,3 et 0,9. Le poids corrigé de la non-réponse initiale (WI/PI) varie dans un rapport de 1 à 10.

La correction de la non-réponse en réinterrogation

Les modèles de non-réponse sont estimés au niveau individuel. Une probabilité de réponse est attribuée à chaque individu panel qui est encore dans le champ en 2008. Est considéré comme répondant tout individu qui appartient à un ménage répondant (DB135='1').

On estime quatre (2008-2004) modèles sur tous les individus qui répondent en même interrogation depuis 2004, soit :

- un modèle pour les individus interrogés en 2ème vague
 - o sur les entrants 2004 encore dans le champ 2005
 - o sur les entrants 2005 encore dans le champ 2006
 - o sur les entrants 2006 encore dans le champ 2007
 - o sur les entrants 2007 encore dans le champ 2008
- un modèle pour les individus interrogés en 3ème vague :
 - o sur les entrants 2004 encore dans le champ 2006
 - o sur les entrants 2005 encore dans le champ 2007
 - o sur les entrants 2006 encore dans le champ 2008
- un modèle pour les individus interrogés en 4ème vague :
 - o sur les entrants 2004 réinterrogés en 2007
 - o sur les entrants 2005 réinterrogés en 2008
- un modèle pour les individus interrogés en 5ème vague :
 - o sur les entrants 2004 réinterrogés en 2008

Cela suppose une stabilité des comportements de non réponse dans le temps.

Probabilité de répondre en 2ème vague sachant qu'on a répondu à la première

Un modèle de non-réponse a été estimé sur les adultes réinterrogés en 2^{ème} vague. L'information utilisée est celle de la première vague. Les facteurs explicatifs de la non-réponse sont :

- le fait d'avoir déménagé avec l'ensemble de son ménage depuis la dernière vague,
- le fait d'habiter dans une maison ou non
- la situation familiale (seul/non),
- avoir un contrat à durée indéterminée ou non
- la localisation géographique (agglomération parisienne/autre),
- le quartile de niveau de vie (1er, 2d ou au-dessus de la médiane) du ménage.

À partir de ces résultats, nous avons formé différentes sous-populations de répondants en croisant les motifs significatifs de la non-réponse (certains groupes à faible effectif ont été regroupés pour assurer la robustesse des résultats). Au final, les adultes réinterrogés en 2ème vague ont été réparti en 7 groupes et nous avons supposé le mécanisme de réponse homogène à l'intérieur de sous-populations. La probabilité PR(2ème vague) vaut donc :

$$\frac{\text{Nombre d'individus répondants dans la sous-population}}{\text{Nombre d'individus panel dans la sous population}}$$

La probabilité PR(2^{ème} vague/1^{ère} vague) est comprise entre 0,5 et 0,9.

Probabilité de répondre à la 3ème vague sachant qu'on a répondu à la 1ère vague

Un modèle de non-réponse a été estimé sur les adultes réinterrogés en 3^{ème} vague.. L'information utilisée est celle de la première vague. Les facteurs explicatifs de la non-réponse sont :

- le fait qu'il y ait eu un ou plusieurs départs dans le ménage depuis la dernière vague,
- le fait d'habiter dans une maison ou non
- le quartile de niveau de vie (1^{er} ou plus) du ménage ,
- la localisation géographique (agglomération parisienne/autre),
- la situation familiale (seule/non),
- avoir un contrat à durée indéterminée ou non

À partir de ces résultats, nous avons formé différentes sous-populations de répondants en croisant les motifs significatifs de la non-réponse (certains groupes à faible effectif ont été regroupés pour assurer la robustesse des résultats). Au final, nous avons réparti les individus dans 7 groupes et nous avons supposé le mécanisme de réponse homogène à l'intérieur de sous-populations. La probabilité PR(3ème vague/1ère vague) vaut donc :

$$\frac{\text{Nombre d'individus répondants dans la sous-population}}{\text{Nombre d'individus panel dans la sous population}}$$

La probabilité PR(3^{ème} vague/1^{ère} vague) est comprise entre 0,5 et 0,9.

Probabilité de répondre en 4ème vague sachant qu'on a répondu à la 1ère vague

Un modèle de non-réponse a été estimé sur les adultes panel interrogés en 4ème vague et répondants en 1ère vague . L'information utilisée est celle de la première vague. Les facteurs explicatifs de la non-réponse sont :

- le fait qu'il y ait eu un ou plusieurs départs dans le ménage depuis la dernière vague,
- la nationalité (magrébine/autre),
-
- le quartile de niveau de vie (1^{er} ou plus) du ménage,

À partir de ces résultats, nous avons formé différentes sous-populations de répondants en croisant les motifs significatifs de la non-réponse (certains groupes à faible effectif ont été regroupés pour assurer la robustesse des résultats). Au final, nous avons réparti les individus dans 5 groupes et nous avons supposé le mécanisme de réponse homogène à l'intérieur de sous-populations. La probabilité PR(4ème vague/1ère vague) vaut donc :

$$\frac{\text{Nombre d'individus répondants dans la sous-population}}{\text{Nombre d'individus panel dans la sous population}}$$

La probabilité PR(4^{ème} vague/1^{ère} vague) est comprise entre 0,5 et 0,9.

Probabilité de répondre en 5ème vague sachant qu'on a répondu à la 1ère vague

Un modèle de non-réponse a été estimé sur les adultes panel interrogés en 5ème vague et répondants en 1ère vague . L'information utilisée est celle de la première vague. Les facteurs explicatifs de la non-réponse sont :

- le fait d'avoir déménagé depuis 2004,
- le fait qu'il y ait eu un ou plusieurs départs dans le ménage depuis la dernière vague,
- avoir en contrat à durée indéterminée ou non
- le niveau de diplôme

À partir de ces résultats, nous avons formé différentes sous-populations de répondants en croisant les motifs significatifs de la non-réponse (certains groupes à faible effectif ont été regroupés pour assurer la robustesse des résultats). Au final, nous avons réparti les individus dans 5 groupes et nous avons supposé le mécanisme de réponse homogène à l'intérieur de sous-populations. La probabilité PR(5ème vague/1ère vague) vaut donc :

$$\frac{\text{Nombre d'individus répondants dans la sous-population}}{\text{Nombre d'individus panel dans la sous population}}$$

La probabilité PR(5^{ème} vague/1^{ère} vague) est comprise entre 0,3 et 0,9.

Partage des poids

Soit :

- $a_{2007,k}$ le sous-échantillon panel à enquêter l'année t en $k^{\text{ième}}$ interrogation
- $W_i(2008,k)$ le poids de sondage « brut » de l'individu i : il s'agit du poids de sondage du logement dans lequel se trouve i à la date de son tirage en tant qu'individu-panel, soit en 2008- $k+1$.
- $\Phi_i(2008,k)$ la probabilité de répondre en première interrogation des ménages du logement l , c'est-à-dire en $t - k + 1$.
- $\theta_i(2008,k)$ la probabilité de répondre en 2008 sachant qu'on appartient à l'échantillon $a_{2007,k}$ et qu'on a répondu en 2008 - $k + 1$.
- $N_{pan_l}(2008)$ est le nombre d'individus panel du logement l en 2008.
- $N_{champ_l^{t,t'}}(2008)$ est le nombre d'individus habitant le logement l en 2008 qui étaient dans le champ en t' .
- $N_{bchamp_l^{t,t'}}(2008)$ est la somme des années passées dans le champ de t à t' par les individus appartenant au logement l en 2008.
- N_{pers08} est le nombre d'individus dans le logement en 2008.

En 2008, on attribue à chaque individu panel répondant le poids suivant :

$$W_i^{*2008} = \frac{W_l(2008, k)}{9} \times \frac{1}{\Phi_l(2008, k)} \times \frac{1}{Npan_l(2008)} \sum_{\substack{j \in l \\ j \in r, k}} \frac{1}{\theta_j(2008, k)}$$

$$\times \frac{9 \times Npan_l(2008)}{5 \times Nchamp_l^{2004}(2008) + Nchamp_l^{2005, 2007}(2008) + Npers_l(2008)}$$

Tous les individus répondant d'un même logement ont donc le même poids qu'on peut décomposer en le poids de tirage corrigé de la non-réponse initiale puis de la non-réponse en réinterrogation des individus panel présents en 2008 et du partage des poids.

Où la correction relative au partage des poids correspond à :

$$\frac{9 \times Npan_l(2008)}{5 \times Nchamp_l^{2008}(2008) + Nchamp_l^{2005, 2007}(2008) + Npers_l(2008)}$$

2.1.8.3 Ajustements aux données extérieures (niveau, variables utilisées et sources) et « troncature »

L'enquête, réalisée en mai-juin 2008, est, comme toutes les enquêtes après des ménages de l'Insee, calée sur les marges issues de l'enquête Emploi de 2007. En effet, toutes les enquêtes ménages d'une année civile N sont calées sur la dernière enquête Emploi disponible au début de l'année N. Toutes les enquêtes dont la collecte a lieu pendant l'année civile N sont donc calées sur les marges constituées à partir des 4 enquêtes Emploi trimestrielles de l'année N-1.

Contrairement aux autres enquêtes ménages de l'Insee, l'enquête Emploi ignore la notion de budget séparé, et un logement n'est occupé que par un seul ménage. Il a donc fallu, dans un premier temps, reconstituer cette notion de ménage-logement et déterminer la personne de référence afin de pouvoir réaliser le calage.

Le calage et la correction de la non-réponse de seconde interrogation ont été effectués avec le logiciel Calmar 2. Les poids en entrée sont ceux issus du partage des poids. Nous avons utilisé la fonction de distance « logit ».

Les variables et les modalités introduites dans le calage sont les suivantes. Il s'agit du :

- nombre de ménages par tranche d'âge de la personne de référence (5 modalités, des moins de 31 ans aux 76 ans et plus). L'âge est mesuré au 31/12/2007.
- nombre de ménages par tranche de densité d'habitat : rural, unité urbaine de moins de 20 000 habitants, unité urbaine de 20 000 à 100 000 habitants, unité urbaine de plus de 100 000 habitants, Région parisienne.
- nombre de ménages par type :
 - personne seule,
 - couple sans enfant,
 - couple avec 1 enfant,
 - couple avec 2 enfants ou plus,
 - famille monoparentale,
 - autre configuration ;

- nombre d'hommes par tranche d'âge (6 modalités, des moins de 15 ans aux 76 ans et plus) ;
- nombre de femmes par tranche d'âge (6 modalités) ;
- nombre de ménages selon le diplôme de la personne de référence :
 - sans diplôme, non déclaré,
 - diplôme inférieur baccalauréat (CAP, BEPC),
 - Baccalauréat, bac+2,
 - diplôme supérieur ;
- nombre de ménages selon la catégorie socioprofessionnelle de la personne de référence (activité actuelle ou ancienne activité) :
 - agriculteurs (retraités ou non),
 - indépendants et professions libérales (retraités ou non),
 - professeurs et instituteurs actifs,
 - professeurs et instituteurs retraités,
 - employés et ouvriers actifs,
 - employés et ouvriers retraités,
 - autres ;
 -

2.1.8.4 Pondération transversale finale

Les pondérations (au niveau logement) obtenues à l'issue du calage sont affectées à tous les ménages répondants qui occupent ce logement et à tous les individus qui appartiennent à un ménage répondant.

Pour les répondants, le poids est modifié pour corriger de la non-réponse individuelle. Les pondérations des individus répondants sont corrigées de façon à garder la même structure par âge et par sexe (16 modalités : croisement du sexe par l'âge en 8 modalités, des moins de 16 ans aux 76 ans et plus) sur l'ensemble de la population. Elle ne respecte donc pas la structure par âge et par sexe au sein d'un ménage. Comme le taux de non-réponse individuel n'est que de 0,3 %, l'impact de cette correction est très limité.

Pour les enfants âgés de 2 à 12 ans, le poids est modifié de façon à avoir la même structure par année de naissance que dans l'enquête Emploi.

Il existe donc trois systèmes de poids pour les individus : le premier (RB050) est parfaitement cohérent avec le système de poids des ménages, les autres permettent de limiter les biais des estimateurs calculés seulement sur certains individus (PB040 pour les questions du questionnaire individuel et RL070 pour les questions relatives aux modes de garde).

2.1.9 Substitutions

Sans objet dans le cas de la France.

2.2 Erreurs d'échantillonnage

Les écarts-type calculés estiment l'écart-type causé par l'échantillonnage corrigé de la non-réponse et du calage. Ils sont linéarisés puis estimés en utilisant le logiciel Poulpe développé à l'INSEE.

La variance du taux de pauvreté obtenue est supérieure de 5 % à celle obtenue par un tirage aléatoire simple de même taille. Cet effet de sondage s'explique par l'utilisation d'un tirage à plusieurs degrés.

En raison du trop grand nombre de valeurs nulles, nous ne sommes pas parvenus à linéariser le taux de pauvreté avant tout transfert et nous n'avons estimé ni écart-type ni effet de sondage pour cet indicateur.

Indicateurs	Valeur	Écart-type	Coefficient de variation	Effet de sondage	Effectif atteint	Taille effective de l'échantillon
Taux de pauvreté après prestations sociales (en pourcentage)						
Total	13.4	0.4	3.0	1.04	25510	24529
Hommes total	12.7	0.5	3.9	1.08	12323	11410
Femmes total	14.0	0.5	3.6	1.03	13187	12803
0-17 ans	17.4	0.8	4.6	0.88	6122	6957
18-24 ans	20.5	1.1	5.4	1.19	2306	1938
25-49 ans	11.6	0.5	4.3	1.04	8230	7913
50-64 ans	10.7	0.8	7.5	1.41	4908	3481
65+ ans	11.1	1.0	9.0	1.71	3944	2306
18+ ans	12.3	0.4	3.3	1.27	19388	15266
18-64 ans	12.6	0.4	3.2	1.15	15444	13430
0-64 ans	13.8	0.4	2.9	0.97	21566	22233
Hommes 18-24 ans	20.7	1.4	6.8	1.19	1170	983
Hommes 25-49 ans	10.4	0.5	4.8	1.14	3940	3456
Hommes 50-64 ans	11.0	1.0	9.1	1.54	2362	1534
Hommes 65+ ans	10.2	1.3	12.7	1.74	1746	1003
Hommes 18+ ans	11.7	0.5	4.3	1.34	9218	6879
Hommes 18-64 ans	12.1	0.5	4.1	1.26	7472	5930
Hommes 0-64 ans	13.2	0.5	3.8	1.02	10577	10370
Femmes 18-24 ans	20.2	1.4	6.9	1.19	1136	955
Femmes 25-49 ans	12.7	0.5	3.9	1.02	4290	4206
Femmes 50-64 ans	10.5	0.9	8.6	1.33	2546	1914
Femmes 65+ ans	11.8	1.1	9.3	1.58	2198	1391
Femmes 18+ ans	12.8	0.5	3.9	1.26	10170	8071
Femmes 18-64 ans	13.1	0.5	3.8	1.13	7972	7055
Femmes 0-64 ans	14.5	0.5	3.4	0.96	10989	11447

Indicateurs	Valeur	Écart-type	Coefficient de variation	Effet de sondage	Effectif atteint	Taille effective de l'échantillon
Taux de pauvreté après prestations sociales (en pourcentage)						
Actifs occupés	6.8	0.4	5,9	0.94	10233	10886
Chômeurs	39.2	2.2	5,6	1.01	897	888
Retraités	9.5	0.7	7,4	1.74	4909	2821
Autres inactifs	27.4	1.2	4,4	0.94	2959	3148
Hommes, actifs occupés	7.3	0.4	5,5	1.0	5306	5306
Hommes, chômeurs	42.1	2.7	6,4	1.06	433	408
Hommes, retraités	9.5	0.9	9,5	1.72	2440	1419
Hommes, autres inactifs	27.7	1.6	5,8	1.0	840	840
Femmes, actives occupées	6.3	0.4	6,3	0.91	4927	5414
Femmes, chômeuses	36.3	2.9	8,0	1.07	464	434
Femmes, retraitées	9.5	0.8	8,4	1.53	2469	1614
Femmes, autres inactives	27.3	1.3	4,8	0.99	2119	2140
Personnes seules, < 65 ans	16.7	1.0	6,0	1.38	1615	1170
Personnes seules, 65+ ans	14.5	1.5	10,3	1.50	1250	833
Personnes seules, hommes	15.8	1.3	8,2	1.61	1115	693
Personnes seules, femmes	16.0	1.1	6,9	1.37	1750	1277
Personnes seules, total	15.9	0.9	5,7	1.41	2865	2032
2 adultes, sans enfant, tous les deux < 65	7.6	0.7	9,2	1.41	3716	2635
2 adultes, sans enfant, au moins 1 65+	8.5	1.5	17,6	1.96	2814	1436
Autres ménages sans enfant	11.4	2.2	19,3	1.70	1555	915
Familles monoparentales	26.6	2.1	7,9	0.84	1536	1829
2 adultes, 1 enfant	10.2	1.1	10,8	0.84	2949	3511
2 adultes, 2 enfants	8.7	1.0	11,5	0.81	5100	6296
2 adultes, 3+ enfants	23.0	2.1	9,1	0.79	3400	4304
Autres ménages avec enfants	24.0	3.1	12,9	0.87	1514	1740
Ménages sans enfant	10.9	0.5	4,6	1.55	10950	7065
Ménages avec enfants	15.7	0.7	4,5	0.91	14499	15933
Propriétaires	8.6	0.5	5,8	1.13	17458	15450
Locataires	22.5	0.9	4,0	0.98	7962	8124
Ménages sans enfant, w = 0	20.4	1.7	8,3	1.67	1757	1052
Ménages sans enfant, 0 < w < 1	10.0	1.0	10,0	1.28	2681	2095
Ménages sans enfant, w = 1	4.9	0.6	12,2	1.45	3190	2200
Ménages avec enfants, w = 0	71.7	3.4	4,7	0.76	772	1016
Ménages avec enfants, 0 < w < 0,5	53.2	3.5	6,6	0.79	692	876
Ménages avec enfants, 0,5 < w < 1	19.9	1.4	7,0	0.81	4652	5743
Ménages avec enfants, w = 1	5.1	0.6	11,8	0.85	8413	9898
Médiane des niveaux de vie (en euros par unité de consommation)	17564	69.5	0,4	1.22	25510	20910
Seuil de pauvreté - personnes seules (en euros)	10538	30.9	0,3	1.41	25510	18092
Seuil de pauvreté - 2 adultes, 2 enfants (en euros)	22130	51.0	0,2	0.81	25510	31494

Indicateurs	Valeur	Écart-type	Coefficient de variation	Effet de sondage	Effectif Atteint	Taille effective de l'échantillon
Inégalité de la distribution des revenus : rapport Q80/Q20	4.2	0.2	4.8	0.88	25510	28989
Écart de risque de pauvreté (en pourcentage)						
Total	18.1	0.7	3.9	0.96	3314	3452
Hommes total	18.7	0.9	4.8	0.99	1521	1536
Femmes total	17.2	0.8	4.7	0.97	1793	1848
0-17 ans	15.0	0.9	6.0	0.81	1018	1257
18-64 ans	21.1	0.9	4.3	1.10	1876	1705
65+ ans	12.1	1.8	14.9	1.73	420	243
18+ ans	19.4	0.9	4.6	1.24	2296	1852
Hommes, 18-64 ans	22.8	1.1	4.8	1.21	867	717
Hommes, 65+ ans	11.4	2.2	19.3	1.61	168	104
Hommes, 18+ ans	20.6	1.1	5.3	1.33	1035	778
Femmes, 18-64 ans	20.2	0.9	4.5	1.08	1009	934
Femmes, 65+ ans	12.9	2.1	16.3	1.83	252	138
Femmes, 18+ ans	18.6	0.9	4.8	1.22	1261	1034
Taux de pauvreté selon le seuil de pauvreté (en pourcentage)						
40 % de la médiane	3.4	0.2	5.9	1.02	25510	25010
50 % de la médiane	7.1	0.3	4.2	1.03	25510	24767
70 % de la médiane	21.2	0.5	2.4	1.11	25510	22982
Taux de pauvreté avant prestations sociales autres que retraites (en pourcentage)						
Total	23.1	0.4	1.7	1.05	25510	24295
Homes total	22.8	0.5	2.2	1.10	12323	11203
Femmes total	23.5	0.5	2.1	0.99	13187	13320
0-17 ans	34.2	0.9	2.6	0.89	6122	6879
18-64 ans	21.9	0.6	2.7	1.20	15444	12870
65+ ans	13.4	0.9	6.7	1.66	3944	2376
18+ ans	20.1	0.4	2.0	1.23	19388	15763
Hommes, 18-64 ans	21.2	0.6	2.8	1.32	7472	5661
Hommes, 65+ ans	12.4	1.2	9.7	1.80	1746	970
Hommes, 18+ ans	19.6	0.5	2.6	1.37	9218	6728
Femmes, 18-64 ans	22.6	0.5	2.2	1.14	7972	6993
Femmes, 65+ ans	14.2	1.0	7.0	1.48	2198	1485
Femmes, 18+ ans	20.6	0.5	2.4	1.15	10170	8843
Coefficient de Gini (en pourcentage)	28.1	0.9	3.2	0.82	25510	31110
Niveau de vie moyen (en euros par unité de consommation)	19995				25510	

Composantes du revenu disponible	Moyenne par ménage (en euros)	Écart-type	Coefficient de variation	Effet de sondage	Effectif atteint	Taille effective de l'échantillon
Revenu disponible total (HY020)	31138	398	1.3	0.96	10418	10852
Revenus tirés de la location d'une propriété ou de terres (HY040N)	803	76	9.5	1.44	10418	7235
Allocation famille/enfant (HY050N)	742	12	1.6	0.89	10418	11706
Exclusion sociale non classée ailleurs (HY060N)	256	19	7.4	1.37	10418	7604
Aides au logement (HY070N)	433	11	2.5	1.35	10418	7717
Transferts interménages perçus régulièrement en espèces (HY080N)	158	13	8.2	0.91	10418	11448
Intérêts, dividendes, bénéfices tirés d'investissements en capital dans les opérations incorporelles (HY090N)	929	362	39.0	0.84	10418	12402
Revenus perçus par les personnes de moins de 16 ans (HY110N)	4	1	25.0	0.52	10418	20035
Impôts réguliers sur la fortune (HY120N)	82	43	52.4	1.41	10418	7389
Transferts interménages versés régulièrement en espèces (HY130N)	226	27	11.9	1.13	10418	9219
Remboursements consécutifs à des ajustements fiscaux (HY145N)	1893	89	4.7	0.67	10418	15549
Salaire ou revenu assimilé (PY010N)	18191	225	1.2	1.34	10418	7775
Bénéfices en espèces ou perte de trésorerie en rapport avec une activité indépendante (PY050N)	1988	189	9.5	0.97	10418	10740
Allocations chômage(PY090N)	757	54	7.1	0.90	10418	11576
Retraites (PY100N)	8538	135	1.6	1.60	10418	6511
Allocations au survivant (PY110N)	61	30	49.2	1.65	10418	6314
Indemnités de maladie (PY120N)	250	20	8.0	1.04	10418	10017
Pensions d'invalidité (PY130N)	186	25	13.4	1.61	10418	6471
Allocations d'éducation (PY140N)	43	4	9.3	1.10	10418	9471

2.3 Erreurs non dues à l'échantillonnage

2.3.1 Erreurs dans la base de sondage et la couverture

Deux bases de sondages ont été utilisées conjointement :

- Le recensement général de la population de mars 1999, qui est reconnu comme étant de très bonne qualité, même s'il y a quelques omissions de logements (la qualité du recensement est certes moins bonne pour les personnes physiques - mais la base utilisée ici est une base de logements). L'ancienneté du recensement explique qu'un certain nombre de logements de la base sont hors champ et que nous sommes obligés de prendre dans l'échantillon des résidences vacantes ou secondaires au moment du recensement alors que notre champ se limite aux résidences principales.
- Une base qui a mis à jour la précédente fin 2005, dite « base de sondage de logements neufs » (BSLN). Pour la constituer, on part de l'ensemble des logements appartenant aux permis de construire délivrés par les mairies. Ces permis sont gérés dans un fichier administratif appelé SITADEL, qui comprend environ 300 000 logements chaque année. Un extrait de SITADEL est obtenu par sondage, ce qui donne lieu à un échantillon de logements initialement « fictifs » dont la construction est suivie sur le terrain par des enquêteurs jusqu'à ce qu'il y ait achèvement du logement. À la constatation de l'achèvement, le logement entre dans la BSLN. La qualité de SITADEL et du processus de suivi sur le terrain n'est pas finement quantifiable, mais on considère que la BSLN est tout à fait satisfaisante en matière de couverture de la construction neuve.

2.3.2 Erreurs de mesure et de traitements

Description des différentes causes d'erreurs de mesure susceptibles d'être détectées dans l'enquête

Jusqu'en 2007, l'ensemble des données étaient récupérées par voie d'enquête. Pour limiter les erreurs de mesures (cf. rapports intermédiaires des années 2004 à 2007) et gagner en qualité sur la mesure des revenus individuels dans SRCV, les données de revenus (imposables) et les prestations sociales sont désormais récupérées par appariement avec les données fiscales et sociales (DGFIP, CNAF, CNAV et CCMSA), sur le modèle de l'Enquête Revenus Fiscaux et sociaux en France.

L'appariement n'est cependant pas exhaustif : du fait que les appariements se font à partir des adresses, les jeunes adultes âgés de 18 à 25 ans faisant leur déclaration de revenus sur celles de leurs parents à une adresse autre que celle de la collecte ne peuvent être retrouvés (pour ces jeunes, le questionnaire restera sous la même forme que les années précédentes afin de récupérer les données de revenu). De la même façon, les personnes ayant récemment déménagé peuvent être difficilement retrouvées. Pour ces cas de non appariements (environ 4% des individus), les erreurs de mesures constatées les années précédentes restent (confusion euros/francs, confusion entre les différentes périodes de référence dans l'enquête...).

Enfin, en ce qui concerne les revenus, seuls les revenus imposables sont récupérés par appariement. Par conséquent, le questionnaire comporte encore quelques questions sur les revenus exonérés d'impôt sur le revenu.

Description de la façon dont le questionnaire a été élaboré, utilisation éventuelle d'un laboratoire cognitif, test sur le terrain du questionnaire, incidence de sa conception, de son contenu et de sa formulation.

Structure du questionnaire

L'enquête est composée d'un questionnaire Ménage (destiné à l'ensemble du ménage) et d'un questionnaire individuel posé à toutes les personnes du ménage âgées de 16 ans ou plus (au 1^{er} janvier de l'année d'enquête). Le questionnaire ménage est précédé du Tronc commun des ménages, qui constitue le socle de l'ensemble des enquêtes auprès des ménages conduites par l'Insee.

Grâce à l'appariement, on a pu alléger sensiblement les questionnaires (ménage et individu), en ce qui concerne les "prestations familiales", les "aides au logement", le bloc "impôt sur le revenu", "prime pour l'emploi", le bloc "impôts locaux", salaires, "Allocations de chômage, allocations de préretraite, pensions et retraites" ou "Aide sociale".

Le temps dégagé du fait de l'appariement a permis d'introduire des questions sur les conditions de vie des ménages (insertion sociale, loisirs, culture, santé, travail...) permettant ainsi de développer les mesures de l'exclusion/inclusion sociale dans SRCV, conformément aux objectifs de SILC.

L'architecture du questionnaire est la suivante :

- Tronc commun des enquêtes ménages (TCM) :
 - Identification du logement, contact ;
 - Tableau des habitants du logement et contour des unités de vie (THL),
- Questionnaire ménage (renseigné par un adulte quelconque du ménage « unité de vie ») :
 - Ressources et charges en période courante ;
 - Changements récents et jeunes enfants ;
 - Conditions de logement - résidence principale ;
 - Revenus non individualisables sur la période de référence ;
 - Endettement et confort financier,
- Questionnaire individuel (renseigné par chaque adulte du ménage, le recours à un proxy étant toléré) :
 - Biographie et ressources culturelles ;

- Activité, emploi, profession ;
- Revenus individualisables sur la période de référence ;
- Santé ;
- Indicateurs sociaux : Santé, participation sociale et relations professionnelles, vie associative.

Pour limiter les erreurs de collecte, l'équipe de conception a d'une part introduit dans CAPI des filtres et des contrôles, d'autre part autorisé une collecte en francs pour de nombreux montants, en particulier lorsque les documents sont anciens (tableau d'amortissement pour la collecte des intérêts d'emprunt, prix de vente du logement).

Exhaustivité et absence de doubles comptes dans la collecte des revenus

L'appariement avec les sources fiscales et sociales permet de garantir une bonne exhaustivité et de limiter les doubles comptes. Pour les revenus non imposables et certains individus non appariés (jeunes adultes), la stratégie de questionnement sur les revenus reste la même que les années précédentes. Comme l'enquêteur n'est pas en mesure de connaître à priori toutes les sources de revenus du ménage et que ce dernier peut oublier d'en mentionner une, on met en œuvre une stratégie générale de « balayage » des revenus. Cette stratégie consiste à :

- recenser les différents types de revenus perçus par le ménage au cours de l'année de référence (année civile 2007), avant de collecter les montants correspondants,
- vérifier que le montant n'a pas été inclus antérieurement dans un autre revenu pour éviter les double-comptes,
- essayer systématiquement, lorsque l'enquêté n'a pas pu ou voulu renseigner un montant, d'obtenir une information en tranches.

Utilisation des documents administratifs et fiscaux

Le questionnement est fonction du type de revenus et des documents utilisés. Lorsque le questionnement complet est passé : parmi ces documents, la ou les déclarations des revenus de l'année de référence 2007 (pré-remplie par l'administration fiscale, reçue en mai 2007) sont privilégiées pour les revenus d'activité et de remplacement des adultes, qui sont en quasi-totalité imposables. La déclaration de revenus présente l'avantage d'un mode de recueil des revenus annuels à la fois plus facile et plus fiable. D'une part, les récapitulatifs annuels sont déjà faits. D'autre part, le concept mesuré (revenu déclaré) est homogène avec celui mesuré dans l'enquête Revenus fiscaux et sociaux (enquête de référence à l'Insee pour la distribution des revenus).

Dans le questionnaire individuel, l'enquêté indique combien de déclarations de revenus 2006 il a effectué, et s'il consent à s'y reporter. Pour les revenus d'activité et de remplacement, le questionnement prend deux formes différentes selon la réponse :

- une boucle (dans CAPI) indexée sur le nombre de déclarations de revenus en cas de consentement,
- sinon, une boucle indexée sur le nombre de sources de revenus (nombre d'employeurs pour les salaires, d'entreprises pour les revenus d'indépendants, de caisses ou organismes de versement pour les retraites ou pensions...).

Dans le premier cas, on obtient un montant annuel de revenu déclaré, qui doit être complété par une interrogation synthétique sur l'existence et le montant de revenus complémentaires non déclarés.

Dans le second cas, on reconstitue un montant annuel de revenu perçu cumulé par source de manière plus analytique : durée de perception, régularité du montant mensuel, et montant mensuel en cas de versements réguliers ou montant annuel en cas de versements irréguliers.

Codification de la profession

L'intégration du logiciel de codification SICORE dans le questionnaire CAPI permet de reconnaître les libellés de profession : SICORE (Système Informatique de COdage des Réponses aux Enquêtes) est un système de chiffrage automatique développé par l'Insee. Les enquêteurs sont avertis lorsque le libellé renseigné est inconnu de la base des professions. Ils peuvent alors saisir d'autres libellés qui sont tous enregistrés. Ce ci permet de corriger les erreurs de saisie et de préciser les libellés.

Questionnement longitudinal

Le questionnement longitudinal a commencé en 2005. Les personnes réinterrogées ont répondu à un questionnaire un peu différent de celui des entrants si elles avaient accepté lors de la dernière vague que les informations qu'elles avaient fournies leur soient restituées l'année suivante.

Dans ce cas, le questionnement tient compte des réponses données l'année précédente dans la formulation des questions, voire en ne reposant pas certaines questions, ce qui allège l'interview.

Des contrôles de cohérence peuvent être effectués d'une vague à l'autre, ce qui réduit le bruit des données. Ils portent :

- sur les dates de certains événements (calendrier d'activité par exemple),
- sur la non-perception d'un type de revenu perçu l'année précédente,
- sur l'évolution de montants détaillés de composantes du revenu ou du coût du logement (fourchette entre -20 % et + 30 %),
- sur l'évolution de montants agrégés de revenus (impôts, salaires, retraites).

Tests du questionnaire

Le questionnaire transversal de 2008 a bénéficié des bilans des collectes 2004 à 2007. Le bilan de la collecte 2008 a été effectué grâce aux questionnaires papiers remplis par l'ensemble des enquêteurs, des synthèses rédigées par les directions régionales et des comptes-rendus des réunions de bilans organisées en Basse-Normandie, Bourgogne et Lorraine.

Le questionnaire de 2008 a donné lieu à un test sur le terrain qui s'est déroulé en novembre 2007 dans les directions régionales de Franche-Comté et de PACA. Il a été mené auprès de 116 ménages (répondants). Il avait pour objectif de tester le questionnaire transversal amélioré suite à la collecte de 2007 et aménagé suite à la refonte ainsi que le module secondaire.

Enfin, un test en bureau a été mené en février 2008. Ce type de test permet de se placer dans des configurations particulières qui n'ont pas été rencontrées lors des tests sur le terrain.

À l'issue de chaque test et suite aux bilans de collecte et aux différentes remarques des enquêteurs, des modifications du questionnaire (structure, formulation) sont proposées au maître d'ouvrage par l'équipe de conception. Ces modifications du questionnaire sont validées à l'occasion d'un comité de pilotage du projet.

Les tests servent aussi à mesurer les durées de collecte qui permettent de calculer la rémunération des enquêteurs pour l'enquête. L'équipe de conception fournit aux enquêteurs une feuille permettant le relevé manuel des temps d'interview, ces relevés venant compléter les temps mesurés dans CAPI.

Informations sur l'intensité et l'efficacité de la formation des enquêteurs: nombre de jours de formation, test des compétences avant d'entreprendre le travail sur le terrain (taux de succès, etc).

L'Insee dispose d'un réseau d'enquêteurs stable. L'Institut a recours dans la mesure du possible aux mêmes enquêteurs d'une vague sur l'autre, ce qui présente au moins deux avantages : les enquêteurs connaissent bien l'enquête, et les enquêtés sont plus facilement fidélisés par un enquêteur qu'ils connaissent déjà.

Formation des enquêteurs

Pour les tests, l'équipe de conception assure directement la formation des enquêteurs.

En revanche, pour la collecte en grandeur réelle (mai-juin 2008), l'équipe de conception ne forme pas directement les enquêteurs mais les gestionnaires responsables de l'enquête dans les directions régionales. Les gestionnaires des directions régionales forment ensuite les enquêteurs SILC. Cependant, la formation des enquêteurs reproduit à l'identique la formation des gestionnaires. L'équipe de conception SILC fournit aux gestionnaires des directions régionales (en plus des documents de collecte) les transparents et les exercices CAPI. La formation dispensée aux gestionnaires et aux enquêteurs a duré trois jours pour ceux qui ne connaissaient pas encore l'enquête et 1,5 jours pour ceux qui avaient déjà participé à la collecte précédente.

Lors de la collecte, chaque enquêteur est accompagné au moins une fois par un agent de l'Insee (gestionnaire de l'enquête, concepteur...).

Documents de collecte

En vue de la collecte, l'équipe de conception prépare les documents de collecte fournis aux enquêteurs. Ceux-ci sont constitués :

- d'un argumentaire (utilisé pour convaincre les ménages récalcitrants),
- d'une lettre avis envoyée aux enquêtés pour les prévenir de la venue d'un enquêteur,
- d'une plaquette de présentation du dispositif envoyée avec la lettre avis
- d'une instruction aux enquêteurs,
- d'un cahier des cartes (les cartes sont utilisées dans le cas où l'énumération des modalités d'une question par l'enquêteur serait trop longue ou fastidieuse),
- d'un fascicule « revenus », compléments d'information pour les enquêteurs sur certaines aides et prestations existantes et sur les conditions à remplir pour en bénéficier, réalisé afin d'aider les enquêteurs à retrouver un type de revenus lorsque le ménage ne sait pas exactement ce qu'il perçoit,
- d'une architecture du questionnaire,
- de fiches de prise de contact et de suivi. Les fiches de suivi comportent quelques informations (numéros de téléphone fixe ou portable, adresse internet, personne-relais) pour anticiper les déménagements ;
- d'un document présentant les résultats de l'enquête de l'année précédente, donné à tous les ménages répondant.

2.3.3 Erreurs de non-réponse

2.3.3.1 Taille de l'échantillon obtenu

Les deux tableaux ci-dessous fournissent le nombre de ménages, puis le nombre d'individus correspondants, pour lesquels l'entretien a été accepté suite à la collecte 2008. Ces nombres sont répartis selon les 9 sous-échantillons, la France ayant opté pour un échantillon rotatif par neuvième (9 interrogations pour un individu panel). La variable « Nombre d'années dans le panel » correspond au nombre d'interrogations annuelles, pour le sous-échantillon concerné, restant à mener juste avant la collecte 2008. Ainsi le sous-échantillon pour lequel le « Nombre d'années dans le panel » vaut 1 correspond aux ménages qui ont été interrogés pour la dernière fois en 2008. L'échantillon pour lequel cette variable vaut 9 correspond à l'échantillon entrant en 2008.

Nombre de ménages pour lesquels un entretien est accepté pour la base de données	
Ventilation par sous-échantillon	
Nombre d'années dans le panel	Nombre de ménages
1	831
2	716
3	761
4	723
5	751
6	1 471
7	1 610
8	1 661
9	1 894
Total	10 418

Nombre de personnes âgées de 16 ans et plus, membres des ménages dont l'entretien est accepté pour la base de données, pour lesquels un entretien personnel est achevé,	
Ventilation par sous-échantillon	
Nombre d'années dans le panel	Nombre de personnes
1	1 617
2	1 396
3	1 442
4	1 402
5	1 442
6	2 879
7	3 098
8	3 208
9	3 531
Total	20 014

2.3.3.2 Non-réponse des unités

Le taux de non-réponse des ménages est donné de manière détaillée dans le tableau ci-dessous. On obtient un taux de non-réponse des ménages de 17,5 %. Il est de 26,0 % pour les entrants et de 15,4 % pour les réinterrogés (10,8 % pour les réinterrogations obligatoires en 2^{ème}, 3^{ème} et 4^{ème} année et 20,4 % pour les réinterrogations non obligatoire de la 5^{ème} année). Il a augmenté par rapport à 2007, du fait pour la première année d'interrogations de ménages en 5^{ème} vague non obligatoire (seules les quatres premières interrogations sont obligatoires).

Adresses contactées avec succès DB120=11	12 594
Total DB120	13 112
Adresse inexistante DB120=23	478
Adresses validées sélectionnées	12 634
Taux de contact à l'adresse Ra	99,7 %
Entretiens de ménages achevés et acceptés pour la base de données DB135=1	10 418
Ménages éligibles aux adresses contactées	12 594
Total DB130	
Proportion d'entretiens de ménages achevés et acceptés pour la base de données Rh	82,7 %
Taux de non-réponse des ménages NRh	17,5 %

Le taux de non-réponse individuelle, ainsi que le taux global, et le détail des calculs sont donnés ci-dessous.

Nombre entretiens individuels achevés RB250 = 11	20 014
Nombre d'individus éligibles dans les ménages dont les entretiens ont été achevés et acceptés pour la base de données RB245 = 1	20 125
Rp	99,4 %
Taux de non-réponse individuelle NRp	0,6 %
Taux global de non-réponse individuelle *NRp	18,0 %

On obtient ainsi un taux global de non-réponse individuelle de 18 % très proche du taux de non-réponse des ménages à 17,5 %.

2.3.3.3 Ventilation des ménages selon l'« enregistrement du contact à l'adresse », le « résultat du questionnaire ménage », et l'« acceptation de l'entretien » pour chaque groupe de rotation et le total

Ventilation des ménages selon l' « enregistrement du contact à l'adresse » (DB120)

Nombre d'années dans le panel	Enregistrement du contact à l'adresse			Total :
	Adresse contactée	Impossible de localiser l'adresse	Adresse inexistante	
1	989 99,8 %	2 0,2 %	0 0,0%	991
2	926 99,9 %	1 0,1%	0 0,0%	927
3	971 100,0 %	0 0,0%	0 0,0%	971
4	918 99,8 %	2 0,2 %	0 0,0%	920
5	948 99,9 %	1 0,1 %	0 0,0 %	949
6	1 669 99,8 %	3 0,2 %	0 0,0 %	1 672
7	1 794 99,8 %	3 0,2 %	0 0,0 %	1 797
8	1 845 99,9 %	1 0,1 %	0 0,0 %	1 846
9	2 534 83,4 %	27 0,9 %	478 15,7 %	3 039
Total	12 594 96,0 %	40 0,3 %	478 3,7 %	13 112

Ventilation des ménages selon le « résultat du questionnaire ménage » (DB130)

DB130		Résultat du questionnaire "ménage"				Total
Nombre d'années dans le panel	Questionnaire rempli	Refus de coopérer	Ensemble du ménage absent	Ménage incapable de répondre	Autres raisons	
1	831 84,0 %	94 9,5 %	23 2,3 %	13 1,3 %	28 2,8 %	989
2	716 77,3 %	131 14,1 %	34 3,7 %	22 2,4 %	23 2,5 %	926
3	761 78,4 %	149 15,3 %	24 2,5 %	16 1,6 %	21 2,2 %	971
4	725 78,9 %	137 14,9 %	17 1,9 %	22 2,4 %	17 1,9 %	918
5	753 79,4 %	133 14,0 %	22 2,3 %	22 2,3 %	18 1,9 %	948
6	1475 88,4 %	71 4,2 %	46 2,8 %	37 2,2 %	40 2,4 %	1669
7	1612 89,9 %	71 3,9 %	46 2,6 %	34 1,9 %	31 1,73 %	1794
8	1665 90,2 %	74 4,0 %	31 1,7 %	29 1,6 %	46 2,5 %	1845
9	1895 74,8 %	299 11,8 %	201 7,9 %	90 3,6 %	49 1,9 %	2534
Total	10 443 82,9 %	1 159 9,2 %	444 3,5 %	285 2,3 %	273 2,2 %	12 594

Ventilation des ménages selon l'« acceptation de l'entretien » (DB135) :

DB135	Acceptation de l'entretien		Total
Nombre d'années dans le panel	1 : Entretien accepté pour la base de données	2 : Entretien rejeté pour la base de données	
1	831 100,0 %	0 0,0 %	831
2	716 100,0 %	0 0,0 %	716
3	761 100,0 %	0 0,0 %	761
4	723 99,7 %	2 0,3 %	725
5	751 99,7 %	2 0,3 %	753
6	1471 99,7 %	4 0,3 %	1475
7	1610 99,9 %	2 0,1 %	1612
8	1661 99,8 %	4 0,2 %	1665
9	1894 99,9 %	1 0,1 %	1895
Total	10418 99,9 %	15 0,1 %	10433

2.3.3.4 Ventilation des unités remplacées

Sans objet dans le cas de la France.

2.3.3.5 Non-réponse à certaines questions

Les tableaux suivants fournissent par composante du revenu net disponible la part d'unités percevant cette composante, le pourcentage de valeur manquante et le taux d'information partielle.

Ventilation des non-réponses à certaines questions

	% de ménages...		
	... ayant perçu un montant (après imputation)	... avec valeur manquante (avant imputation)	... avec information partielle (avant imputation)
<i>Ménages</i>			
Revenu disponible total (HY020)	100.0	0.4	97.9
Revenu disponible des ménages avant transferts sociaux à l'exception des allocations de vieillesse et des pensions au survivant (HY022)	99.2	1.2	92.4
Revenu disponible des ménages avant transferts sociaux (HY023)	97.1	9.8	81.3
Revenus tirés de la location d'une propriété ou de terres (HY040N)	12.4	-	-
Allocation famille/enfant (HY050N)	27.4	27.4	-
Exclusion sociale non classée ailleurs (HY060N)	6.8	3.1	0.6
Aides au logement (HY070N)	23.7	2.3	12.3
Transferts interménages perçus régulièrement en espèces (HY080N)	5.2	-	-
Intérêts, dividendes, bénéfices tirés d'investissements en capital dans les opérations incorporelles (HY090N)	79.1	0.1	69.1
Revenus perçus par les personnes de moins de 16 ans (HY110N)	1.8	-	-
Impôts réguliers sur la fortune (HY120N)	1.5	-	-
Transferts interménages versés régulièrement en espèces (HY130N)	6.4	5.0	0.8
Remboursements consécutifs à des ajustements fiscaux (HY145N)	87.1	12.3	58.6

	% de personnes de plus de 16 ans...		
	... ayant perçu un montant (après imputation)	... avec valeur manquante (avant imputation)	... avec information partielle (avant imputation)
<i>Individus</i>			
Salaire ou revenu assimilé (PY010N)	54.7	5.0	0.6
Avantages en nature (PY020N)	11.8	-	-
Bénéfices en espèces ou perte de trésorerie en rapport avec une activité indépendante (PY050N)	4.6	-	-
Allocations chômage (PY090N)	6.8	-	5.7
Retraites (PY100N)	28.3	-	14.0
Allocations au survivant (PY110N)	0.5	-	0.2
Indemnités de maladie (PY120N)	4.6	-	4.6
Pensions d'invalidité (PY130N)	1.9	-	-
Allocations d'éducation (PY140N)	1.9	-	1.9

2.3.3.6 Nombre total des non-réponses

La non-réponse se décompose en non-réponse partielle (non-réponse à une question) et en non-réponse totale. La non-réponse au niveau ménage (cf. partie 2.3.3.2) concerne 2 694 ménages, soit 17,5 % des ménages.

Taille d'échantillon et non-réponse totale par type d'indicateur

	Taille d'échantillon obtenu (nb de personnes prises en compte dans le calcul)	Non réponse partielle	Hors champ	Non-réponse totale des individus	Taux de non réponse individuelle
Taux de pauvreté après transferts selon l'âge et le sexe	25 510	0	0	0	0 %
Taux de pauvreté après transferts selon le statut professionnel	18 998	0	6 401 ¹	111 ²	
Taux de pauvreté après transferts selon le type de ménage	25 510	0	0	0	0 %
Taux de pauvreté après transferts selon le statut d'occupation du logement	25 510	0	0	0	0 %
Taux de pauvreté après transferts selon l'intensité d'emploi	22 157	0	3 353 ³	0	0 %
Médiane du revenu disponible par unité de consommation - après transferts sociaux	25 510	0	0	0	0 %
Seuil de pauvreté - après transferts sociaux - personnes seules	25 510	0	0	0	0 %
Seuil de pauvreté - après transferts sociaux - 2 adultes, 2 enfants	25 510	0	0	0	0 %
Inégalité de la distribution de revenu - Rapport du partage de revenus S80/S20	25 510	0	0	0	0 %
Écart moyen relatif du taux de pauvreté	3 214	0	22 196 ⁴	0	0 %
Dispersion autour du seuil de pauvreté	25 510	0	0	0	0 %
Taux de pauvreté avant transferts sociaux autres que retraites	25 510	0	0	0	0 %
Taux de pauvreté avant tout transfert social	25 510	0	0	0	0 %
Coefficient de Gini	25 510	0	0	0	0 %
Revenu disponible par uc moyen	25 510	0	0	0	0 %

¹ dont 5 385 enfants de moins de 16 ans, 725 entre 16 et 18 ans, et 291 individus qui n'ont pas gardé le même statut professionnel plus de la moitié de l'année.

² 111 individus n'ont pas rempli de questionnaire individuel, leur statut d'occupation est donc inconnu.

³ personnes dans un ménage d'étudiants ou dans un ménage avec seulement des enfants dépendants ou des personnes âgées de plus de 65 ans.

⁴ 22196 non-pauvres.

2.4 Mode de collecte des données

Le taux d'acceptation des individus appartenant aux ménages répondants est très élevé : 99,4 %. En effet, les ménages répondants ont tous rempli au moins un questionnaire individuel et les proxys sont autorisés.

Ventilation des membres de ménages répondants âgés du plus de 16 ans selon le statut des données

RB250		Statut des données		
Nombre d'années dans le panel		Information collectée par enquête	Pas de contact	Total
1	Nombre Pourcentage	1 617 99,5 %	8 0,5 %	1 625
2	Nombre Pourcentage	1 396 99,4 %	8 0,6 %	1 404
3	Nombre Pourcentage	1 442 99,6 %	6 0,4 %	1 448
4	Nombre Pourcentage	1 401 99,3 %	10 0,7 %	1 411
5	Nombre Pourcentage	1 442 99,5 %	7 0,5 %	1 449
6	Nombre Pourcentage	2 879 99,5 %	16 0,5 %	2 895
7	Nombre Pourcentage	3 098 99,1 %	27 0,9 %	3 125
8	Nombre Pourcentage	3 208 99,5 %	17 0,5 %	3 225
9	Nombre Pourcentage	3 531 99,7 %	12 0,3 %	3 543
Total	Nombre Pourcentage	20 014 99,4 %	111 0,6%	20 125

Ventilation des membres de ménages répondants âgés de plus de 16 ans selon le type d'entretien

Seuls les membres du ménage ont le droit de répondre à la place d'un individu. Le tableau suivant détaille le taux de recours au proxy, qui s'élève globalement à 28 %.

RB260		Type d'entretien	
Nombre d'années dans		2	5
		Face à face CAPI	Proxy
1	Nombre Pourcentage	1 143 70,7 %	474 29,3 %
2	Nombre Pourcentage	1 026 73,5 %	370 26,5 %
3	Nombre Pourcentage	1 044 72,4 %	398 27,6 %
4	Nombre Pourcentage	989 70,6 %	412 29,4 %
5	Nombre Pourcentage	991 68,7 %	451 31,3 %
6	Nombre Pourcentage	2 026 70,4 %	853 29,6 %
7	Nombre Pourcentage	2 246 72,5 %	852 27,5 %
8	Nombre Pourcentage	2 315 72,2 %	893 27,8 %
9	Nombre Pourcentage	2 608 73,9 %	923 26,1 %
Total	Nombre Pourcentage	14 388 71,9 %	5 626 28,1 %

2.5 Durée de l'entretien

La durée moyenne de l'entretien pour la collecte 2008 est de 46 minutes (58 minutes pour les entrants et 52 pour les réinterrogés). La durée moyenne (moins 8 minutes) est plus faible que l'année précédente du fait du module secondaire plus long et de l'intégration de questions spécifiques relatives à la sécurité incendie cette année là dans le questionnaire français.

2.6 Les imputations

L'enquête réalisée en 2008 est différente des précédentes dans le mode de collecte des revenus. En effet, à de très rares exceptions les ménages ne sont plus questionnés sur le montant des revenus. Ces derniers sont obtenus à partir des fichiers fiscaux de la Direction Générale des Finances Publiques ou des organismes gestionnaires de prestations pour les revenus sociaux. Le rapprochement du fichier d'enquête et du fichier fiscal est opéré à l'aide de l'adresse du ménage, dans le cas où les adresses ne correspondent pas, les revenus ne seront pas retrouvés. C'est le cas des jeunes en particulier qui peuvent déclarer leur impôt avec leurs parents s'ils sont âgés de moins de 21 ans, ou bien s'ils sont étudiants de moins de 25 ans. Si ces jeunes ont décohabité, le risque de ne pas retrouver leurs revenus est élevé, c'est pourquoi nous les interrogeons toujours suivant l'ancienne méthode.

Les prestations familiales et les minima sociaux sont obtenus auprès des organismes les détenant, c'est toujours l'adresse du ménage qui permet la recherche et les caractéristiques démographiques de l'allocataire (sexe, date de naissance) : pour diverses raisons certaines adresses ne correspondent pas, ces revenus sont alors imputés.

L'imputation est d'abord nécessaire parce qu'il existe des données manquantes ou en tranches. Les données manquantes proviennent des individus dont la déclaration fiscale n'a pu être retrouvée. Les

données en tranches concernent les enfants rattachés fiscalement au foyer fiscal de leurs parents et ne résidant pas avec eux-ci. Avant toute chose, il convient de présenter certaines difficultés rencontrées. Les revenus dans l'EU-SILC 2008 sont relatifs à plusieurs dates. L'impôt payé a pour assiette les revenus imposables perçus au cours de l'année 2006. Les revenus collectés sont relatifs à l'année 2007. Enfin certains revenus concernent la période courante de l'enquête, soit mai 2008, ou bien une moyenne estimée sur une période relativement récente. Nous devons faire l'hypothèse d'une certaine continuité, pour pouvoir comparer ces différents revenus et les contrôler.

Nous décrivons maintenant les opérations concernant les revenus principaux.

Les deux méthodes retenues pour les imputations

L'imputation des revenus individuels est menée de deux façons différentes, selon que le ménage est enquêté pour la première fois ou non. Dans le premier cas, l'imputation est transversale : une équation du revenu est estimée sur les répondants et permet d'imputer le revenu des non-répondants (pour les jeunes) ou non appariés. Dans le second cas, nous utilisons le revenu donné par l'individu à une date précédente pour estimer le revenu manquant perçu à l'autre date. Pour ce faire, nous estimons une équation du ratio entre les revenus des deux années sur les répondants aux deux vagues. Ce ratio est ensuite estimé pour les individus n'ayant répondu qu'à une enquête afin d'attribuer le revenu manquant. Cette méthode est appliquée pour les imputations des salaires et des retraites.

Salaires ou revenu assimilé (PY010N)

Il est nécessaire dans un premier temps de définir sur quelles données l'imputation va porter puis nous comparons le salaire imputé avec les maxima observés dans l'ERFS en tenant compte du sexe du salarié et de sa catégorie socioprofessionnelle (sur une position). Nous attribuons aussi un salaire aux individus déclarant en percevoir et qui ne figurent pas dans le fichier des impôts, ainsi qu'aux individus ayant répondu par un montant en tranches (cas des jeunes).

L'imputation est menée par strates. Huit strates sont créées à partir du sexe, de l'emploi qualifié ou non, et du secteur d'emploi, privé ou public. Nous sélectionnons différentes variables pouvant expliquer le salaire dans chaque strate. Un tronc commun de variables explicatives est formé par l'ancienneté dans la profession et son carré, l'emploi atypique ou non, et le diplôme du salarié. Pour les salariés du privé nous y ajoutons le type de contrat, le fait d'avoir un emploi en Île-de-France ou pas, la proportion de femmes dans le secteur et le fait d'être cadre ou pas. Enfin pour les salariés du public, en plus des variables du tronc commun, nous complétons avec le fait d'être enseignant ou pas, fonctionnaire d'État ou pas et le grade.

Le salaire mensuel ou le ratio entre les salaires des deux années consécutives est imputé. Nous tenons ensuite compte du nombre de mois d'activité déclaré à l'enquête pour estimer le salaire annuel. Un travail particulier est nécessaire pour les salariés à temps partiel.

Les salaires des non-salariés

Nous traitons également les salaires d'individus dont l'activité principale n'est pas salariée. Le petit nombre de cas et le manque d'information, nous ont amenés à imputer pour ces individus des salaires moyens de personnes ayant les mêmes caractéristiques.

Préretraites

Un petit nombre de préretraites sont à imputer : nous nous sommes contentés de leur attribuer le montant moyen de préretraites d'individus ayant des caractéristiques similaires.

Allocations de vieillesse (PY100N) ou pension au survivant (PY110N)

Les retraites sont traitées différemment des autres revenus. L'hypothèse majeure justifiant ce choix est que les retraites sont quasi stables. Comme nous ne disposons pas d'une retraite courante (relative au mois de l'enquête), nous utilisons le revenu courant comme élément de contrôle et nous le comparons à la somme des revenus de l'année de référence du ménage. Pour les ménages concernés, la retraite est un élément prépondérant du revenu total, ce qui justifie la comparaison. Nous ajoutons un autre contrôle, cette fois entre les années de revenu 2006, 2005, 2004 et 2003, toujours sous la même hypothèse nous comparons les impôts sur le revenu déclarés en 2006, 2005, 2004 et 2003 que nous calculons. Deux contrôles sont donc possibles avant de prendre une décision.

À l'aide d'un logiciel approprié nous pouvons observer, pour le ménage, l'ensemble des revenus cumulés, le nombre de mois de retraite déclaré, le nombre d'individus et leur description : âge, catégorie socioprofessionnelle, etc. L'expertise a fait ressortir en 2004 des cas types d'erreurs, ce qui permet de corriger automatiquement une grande partie des retraites erronées. Les erreurs les plus fréquentes portent sur les confusions francs/euros, ainsi que sur les périodes de perception.

Lorsqu'il n'a pas été possible de déterminer l'origine de l'erreur, le montant de la retraite est imputé. Deux strates sont utilisées, suivant que le conjoint de la personne retraitée est vivant ou pas. Pour les personnes dont le conjoint n'est pas décédé les variables explicatives du montant de la retraite perçue sont le sexe, le secteur d'activité (privé ou public), la qualification, le diplôme, l'âge et son carré, et l'ancienneté dans la profession. Pour les retraités dont le conjoint est décédé, ces variables sont complétées par le secteur d'activité de l'ex-conjoint ainsi que sa qualification.

Selon le rang d'interrogation de l'individu, la retraite ou le ratio des retraites des deux années consécutives est estimé afin d'imputer un montant.

Bénéfices en espèces ou perte de trésorerie en rapport avec une activité indépendante

Ces revenus sont collectés sous deux formes : d'une part la forme fiscale, comprenant les amortissements et autres abattements, et d'autre part la forme privée, correspondant au revenu net déterminé par le ménage. Le revenu fiscal est jugé peu réaliste, c'est pourquoi le revenu privé est privilégié. Ainsi si les deux revenus sont renseignés et si le revenu privé est vraisemblable, le revenu privé est le seul pris en compte. À défaut de revenu privé, le revenu fiscal est retenu. Si les deux types de revenus sont manquants ou peu crédibles, l'imputation se fait par hot-deck.

Prestations familiales

Les prestations familiales sont, depuis la collecte 2008 sur les prestations 2007, récupérées par appariement auprès des organismes de prestations sociales. Dans le cas où l'allocataire n'a pu être apparié, les prestations sont calculées sur barème. La principale difficulté est la période de référence des revenus pour les prestations sous condition de ressources. Pour les Caisses d'allocations familiales (Caf), jusque juillet d'une année N, les revenus retenus pour le calcul des aides est celui de l'année N-2 ; à partir de juillet N, les revenus retenus sont ceux de N-1, Nous utilisons uniquement les revenus de l'année N pour imputer les prestations de l'année N.

Les imputations des prestations familiales ne posent pas de problèmes majeurs, les résultats sont conformes aux données des CAF. Les valeurs calculées sont imputées si l'individu n'a pas été retrouvé dans les fichiers sociaux.

Aides au logement

Pour les allocataires non appariés, les aides au logement sont calculées sur barème pour tous les locataires et les accédants à la propriété. Nous utilisons le loyer déclaré à l'enquête, et à défaut, un loyer imputé. La masse des allocations collectées est inférieure de 10 % aux données de la Cnaf corrigées de la différence de champ (ménages en institutions) . Des aides sont donc attribuées à certains ménages, de façon aléatoire, afin de disposer du bon nombre de bénéficiaires. Le calcul des aides pour les locataires ne pose pas de problème majeur même si, comme pour les prestations familiales, la période des revenus n'est pas exactement celle retenue par les CAF. Le calcul des aides aux accédants à la propriété diffère néanmoins du calcul des mêmes aides dans l'ERFS. Nous avons en effet choisi d'appliquer le barème locatif aux accédants, en nous aidant du loyer fictif imputé.

Minima sociaux

Pour les allocataires non appariés, trois minima sociaux sont imputés dans SILC : le revenu minimum d'insertion (RMI), le minimum vieillesse et l'allocation de parent isolé (API). Selon les données brutes, SILC comprend 50 % des bénéficiaires du RMI, 15 % des bénéficiaires du minimum vieillesse et 50 % des bénéficiaires de l'aide au parent isolé (API).

Les méthodes utilisées sont proches des méthodes d'imputation utilisées dans l'ERFS. Elles présentent une limite. Le revenu retenu pour le calcul du RMI est un revenu trimestriel que nous ne connaissons pas. Le RMI imputé est égal à la différence entre le plafond du RMI et les revenus de l'année 2004. Ce plafond dépend du type de famille et du nombre de personnes à charge. Ainsi calculé, aucun minimum ne peut être imputé à une famille dont les revenus annuels sont supérieurs

au plafond. Cette famille a pourtant pu être éligible, si les revenus d'un trimestre se sont avérés insuffisants.

Nous n'avons pas traité les dispositifs d'intéressement à la prise ou à la reprise d'un emploi.

Loyer imputé

Un loyer fictif est imputé aux propriétaires, aux accédants, aux usufruitiers, aux logés gratuitement et aux locataires payant un loyer inférieur au prix du marché.

La méthode retenue comprend quatre étapes, dont deux régressions :

- 1) Estimation d'une équation de loyer à partir des données de l'enquête logement 2002. L'estimation porte sur les logements locatifs du parc privé hors loi de 19482. Les variables explicatives sont les caractéristiques du logement (surface, confort, sanitaires, équipement, état du logement ...) et de localisation (tranche de taille d'agglomération, zone climatique, typologie socio-économique de Nicole Tabard ...). Deux équations distinctes sont estimées, l'une pour les appartements (variable expliquée : le logarithme du loyer au m²) et l'autre pour les maisons (variable expliquée : le logarithme du loyer, la surface figurant parmi les variables explicatives).
- 2) Les équations précédemment estimées sont utilisées pour imputer un loyer fictif aux propriétaires occupants ainsi qu'aux ménages logés gratuitement et un loyer de marché aux locataires du parc social ou en Loi de 1948 de l'enquête logement. On a rajouté à la valeur issue de l'équation un résidu tiré selon une procédure de hot-deck stratifié.
- 3) Ce loyer imputé est régressé sur deux types de variables : des variables du tronc commun des enquêtes ménages de l'Insee d'une part, et des variables géographiques d'autre part. À caractéristiques sociodémographiques et de localisation identiques, les logements occupés par les accédants à la propriété sont d'une qualité moyenne supérieure à ceux des propriétaires sans charge de remboursement, qui sont eux-mêmes de meilleure qualité que ceux du parc social. Estimer une seule équation aurait pu biaiser les estimations. Aussi huit régressions distinctes ont-elles été estimées sur des segments relativement homogènes du parc :
 - appartements, propriétaires sans charge de remboursement et ménages logés gratuitement
 - appartements, accédants à la propriété ;
 - appartements, locataires du parc social ou loi de 1948 ;
 - appartements, locataires du parc libre louant vide ;
 - maisons, propriétaires sans charge de remboursement et ménages logés gratuitement ;
 - maisons, accédants à la propriété ;
 - maisons, locataires du parc social ou loi de 1948 ;
 - maisons, locataires du parc libre louant vide.
- 4) Les huit équations estimées sont exportées vers l'enquête SILC pour y imputer soit un loyer fictif soit un loyer manquant. Lors de l'imputation on rajoute à la valeur prédite un résidu tiré selon une procédure de hot-deck stratifié.
- 5) Les résultats avant prise en compte de l'augmentation des loyers sont présentés ici.

² La crise du logement consécutive à la Seconde Guerre mondiale a obligé les autorités à prendre des mesures contre la flambée des loyers en les fixant par décret. C'est l'objet de la loi du 1^{er} septembre 1948. Elle est applicable aux immeubles construits antérieurement à cette date principalement dans des communes de plus de 4 000 habitants.

Loyers mensuels réels ou fictifs en appartement

	Propriétaires		Accédants		Locataires parc social		Locataires parc libre	
	Enquête Logement 2006	EU-SILC	Enquête Logement 2006	EU-SILC	Enquête Logement 2006	EU-SILC	Enquête Logement 2006	EU-SILC
Effectif	2 441 483	2 141 765	1 167 597	1 233 461	4 238 130	3 913 061	3 316 355	3 359 684
Moyenne	726	697	696	673	323	330	495	609
Minimum	59	79	119	120	14	9	33	1
Q5	267	267	290	284	169	161	238	230
Q10	324	318	343	333	197	195	278	280
Q25	439	414	446	424	244	250	356	360
Q50	598	583	607	565	300	305	444	470
Q75	883	829	835	843	373	385	570	600
Q90	1222	1228	1116	1160	463	500	714	723
Q95	1484	1505	1421	1366	538	585	854	845
Maximum	10780	5691	4098	3028	3756	1568	4081	2290

Loyers mensuels réels ou fictifs en maison individuelle

	Propriétaires		Accédants		Locataires parc social		Locataires parc libre	
	Enquête Logement 2006	EU-SILC	Enquête Logement 2006	EU-SILC	Enquête Logement 2006	EU-SILC	Enquête Logement 2006	EU-SILC
Effectif	8 428 662	8 470 295	3 972 366	4 514 311	820 883	1 179 988	1 505 235	1 638 388
Moyenne	498	472	664	613	360	358	516	498
Minimum	31	38	51	78	38	17	37	1
Q5	162	149	267	274	164	105	195	195
Q10	208	197	340	337	203	151	269	270
Q25	303	282	467	440	276	265	364	350
Q50	428	410	609	564	355	350	477	480
Q75	599	569	792	733	432	452	625	610
Q90	825	776	1030	940	503	556	785	790
Q95	1018	948	1790	1119	558	632	904	900
Maximum	498	4 712	664	3 094	360	1 100	516	1 700

3 COMPARABILITÉ

3.1 Concepts et définitions de base

Population de référence

La population de référence de l'enquête est constituée par l'ensemble des ménages ordinaires (hors institutions) dont la résidence principale se situe en France métropolitaine. En 1999, 2,2 % de la population de la France métropolitaine vivait en collectivité, en institution ou était sans domicile fixe³. Les habitants des DOM-TOM représentaient alors 2,8 % des personnes vivant en ménage ordinaire.

Définition du ménage privé

Un ménage est dans l'enquête française SILC « une personne vivant seule ou un groupe de personnes vivant ensemble qui partagent les dépenses et participent à une économie ménagère commune ». On considère que c'est en contribuant aux dépenses de l'unité de vie que les membres du ménage mettent en commun leurs ressources.

Moins de 1% des logements dans l'enquête SILC contiennent plusieurs ménages qui constituent des unités de vie indépendantes.

Appartenance au ménage

Font partie d'un même ménage les individus qui résident dans une même résidence habituelle en faisant budget commun. Un ménage n'est enquêté que dans sa résidence principale. En première vague, nous n'interrogeons que les unités de vie ayant pour résidence principale un logement de l'échantillon.

Période(s) de référence du revenu utilisée(s), périodes utilisées pour les impôts sur le revenu et les cotisations sociales

L'enquête est conduite en mai-juin 2008. L'année de référence est 2007. On collecte au niveau individuel et ménage les revenus perçus au cours de l'année 2007. On collecte au niveau du ménage les impôts payés en 2007 au titre des revenus perçus au cours de l'année 2006. Les cotisations sociales sont relatives aux revenus perçus en 2007.

Période de référence pour les impôts sur la fortune

On considère le montant de l'impôt de solidarité sur la fortune payé en 2007. Il est relatif au patrimoine détenu au 1^{er} janvier 2007.

Écart entre la période de référence du revenu et les variables actuelles

Les variables actuelles (par exemple les variables relatives au coût du logement) portent sur la période de collecte, c'est-à-dire les mois de mai et juin 2008, ce qui rajoute une troisième période de référence après l'année 2007 (revenus déclarés) et l'année 2006 (assiette des impôts sur le revenu).

Durée totale de la collecte de données pour l'échantillon.

La collecte SILC s'est déroulée du 13 mai au 28 juin 2008.

Informations générales sur l'activité professionnelle durant la période de référence.

La catégorie socioprofessionnelle des actifs ou des anciens actifs est codée automatiquement après la collecte par le système SICORE. 95 % des professions sont codées automatiquement, les autres professions étant « reprises » à la main par une équipe spécialisée à l'Insee.

Le codage en ISCO se fait à partir du code des professions et catégories socioprofessionnelles (PCS) à quatre chiffres, de la nomenclature d'activité française en 712 postes et du nombre de salariés de l'établissement. Dans 95 % des cas, la matrice de passage élaborée pour l'enquête Emploi permet de coder la profession de l'enquêté en un code ISCO à deux positions. Sinon, la codification en ISCO a dû se faire par imputation statistique. Le code ISCO est alors affecté aléatoirement en fonction de la

³ Une enquête a été menée en 2001 auprès des sans domiciles fixes. Les résultats de cette enquête ont été publiés dans le n° 391-392 de la revue *Économie et Statistique*.

répartition du code ISCO par PCS à deux chiffres dans l'enquête Emploi du deuxième trimestre de 2007.

Dans 1,6 % des cas, il y a imputation en raison d'un manque d'information dans l'enquête (PCS sur 2 positions au lieu de 4, activité codée seulement en 62 postes). Dans 3,8 % des cas, il s'agit d'un défaut de la matrice de passage entre les PCS et ISCO. Dans ces cas, l'enquête Emploi fournit un codage en une position, ce que le règlement SILC n'autorise pas alors que l'information collectée sur la profession est plus précise dans l'enquête Emploi que dans SILC.

3.2 Composantes du revenu

3.2.1 Les différences entre les définitions nationales et les définitions type des statistiques EU-SILC

Les fichiers français respectent les définitions des statistiques EU-SILC.

Loyers imputés (HY030N), avantages en nature (PY020N), et valeur des biens produits pour la consommation propre (PY070N)

Ces trois composantes du revenu sont obligatoires à partir de 2008. Il n'est pas possible d'isoler des salaires l'avantage en nature que représente la voiture de fonction. La variable PY021 n'est donc pas renseignée.

Revenus de la propriété (HY040N et HY090N)

Pas de classification particulière.

Allocations famille/enfants (HY050)

Elles incluent l'ensemble des allocations versées par les Caisses d'allocations familiales hormis les allocations logement et le RMI. La plupart sont sans conteste des prestations en espèces versées pour faire face aux dépenses liées à l'éducation des enfants ou à la perte de revenu consécutive à l'arrêt d'activité d'un des parents pour élever ses enfants :

- l'allocation d'adoption (ADA) versée pendant 21 mois après l'adoption sous condition de ressources ;
- La prime de naissance de la prestation d'accueil du jeune enfant (PAJE). Elle est versée sous condition de ressources au 7^{ème} mois de grossesse ;
- L'allocation de base de la PAJE. Sous condition de ressources, elle est versée jusqu'aux 3 ans du dernier né ;
- Le complément de libre-choix d'activité de la PAJE. Il vise à compenser l'arrêt total ou partiel d'activité d'un des deux parents pour élever son enfant âgé de moins de 3 ans. La durée maximale de l'arrêt est de 6 mois pour un premier enfant et 3 ans pour les suivants. Son montant est modulé en fonction de la perception ou non de l'allocation de base ;
- l'allocation parentale d'éducation (APE) versée aux parents qui arrêtent ou ont arrêté totalement ou partiellement de travailler pour élever leurs enfants. L'APE de rang 2 (2 enfants à charge) est versée à condition d'avoir travaillé deux ans (consécutifs ou non) dans les 5 ans précédant l'arrivée du dernier enfant. L'APE de rang 3 (3 enfants ou plus à charge) est versée à condition d'avoir travaillé deux ans au cours des dix ans précédant l'arrivée du dernier-né ;
- les allocations familiales qui sont versées à toutes les familles d'au moins deux enfants à charge ;
- le complément familial versé aux familles de trois enfants ou plus dont le benjamin est âgé d'au moins trois ans. Cette allocation est versée sous condition de ressources ;
- l'allocation parent isolé (API) est un minimum social versé aux personnes élevant seules un enfant. Cette allocation est versée durant douze mois ou bien jusqu'au mois précédant le troisième anniversaire du dernier-né ;
- l'allocation de soutien familial (ASF) est versée quand un parent ne concourt pas aux besoins de l'enfant ;
- l'allocation de rentrée scolaire (ARS) est versée en début d'année scolaire sous condition de ressources ;

- l'allocation d'éducation spéciale (AES) pour enfants handicapés ;
- l'allocation journalière de présence parentale (AJPP) pour enfants handicapés ou malades en cas d'arrêt total ou partiel de travail des parents.

L'inclusion du complément de mode de garde de la PAJE ou de l'aide à la famille pour l'emploi d'une assistante agréée (AFEAMA) ou de l'allocation de garde d'enfant à domicile (AGED) dans les allocations familiales est plus discutable. Ces prestations pourraient en effet être classées en prestations en nature puisqu'elles couvrent une partie des dépenses liées à la garde des enfants. Nous avons cependant choisi de les inclure dans les prestations familiales. La mise en place de la PAJE se passe de la façon suivante : les ménages qui ont un nouveau-né après le 1^{er} janvier 2004 passent entièrement sous le régime de la PAJE, les autres continuent de bénéficier des anciennes allocations.

En outre, les indemnités pour congé maternité ne sont pas incluses dans les allocations familiales, mais dans les indemnités maladie.

Dépenses liées à l'exclusion sociale non classées ailleurs (HY060)

Elles incluent le RMI et les aides financières sociales versées par les collectivités locales.

Aides au logement (HY070)

Pas de classification particulière.

Transferts interménages perçus/versés régulièrement en espèces (HY080 et HY130)

Ces transferts excluent les versements exceptionnels mais incluent le paiement de loyer par un tiers. Ce paiement a en effet été considéré comme un paiement en espèces.

Revenus perçus par les personnes de moins de 16 ans (HY110N)

Les enfants ne travaillant pas avant 16 ans, ils ne reçoivent donc pas de revenus d'activité. En outre, cet item doit exclure les transferts entre ménages et donc les pensions alimentaires. Cette composante est ainsi égale aux bourses d'école reçues par les élèves de famille à faibles revenus.

Impôts réguliers sur la fortune (HY120)

Cet item est égal à l'impôt de solidarité sur la fortune.

Impôts sur le revenu et cotisations sociales (HY140G) et remboursements/encaissements liés à des ajustements d'impôt (HY145N)

L'agrégat HY140G comprend le HY145N, les contributions et cotisations sociales. Ces contributions sont assises sur les salaires, les pensions, les revenus d'indépendants, les allocations chômage, les retraites, les prestations familiales et les allocations logement.

L'agrégat HY145N comprend l'impôt sur le revenu des personnes physiques (IRPP), la taxe d'habitation et les impôts payés à l'étranger. La prime pour l'emploi est comptée négativement. L'IRPP assis sur les revenus de l'année N est payé l'année N+1 sous forme d'acomptes et d'un solde. Il n'est donc pas prélevé à la source et porte sur plusieurs types de revenus tous perçus l'année précédente, c'est pourquoi nous l'avons intégralement enregistré comme ajustement d'impôt. La taxe d'habitation est payée par les personnes occupant un logement au premier janvier, elle est assise sur la valeur locative du logement mais son montant dépend aussi des revenus du ménage. La prime pour l'emploi est versée aux actifs aux revenus faibles, elle est déduite de l'IRPP.

Salaires ou revenu assimilé (PY010N)

Les salaires sont nets des cotisations sociales salariales. Depuis 2008, les salaires étant récupérées par appariement, ils comprennent les avantages en nature déclarés aux impôts et les allocations de préretraite imposables (ceux-ci ne peuvent pas être distingués des salaires).

Cotisations sociales des employeurs (PY030G)

Elles comprennent les cotisations sociales versées au régime général, mais aussi certaines taxes assises sur les salaires (versement transport, Fond national d'aide au logement, taxe d'apprentissage, formation professionnelle).

Bénéfices en espèces ou pertes de trésorerie en rapport avec une activité indépendante (y compris honoraires) (PY050N)

Nous avons privilégié les prélèvements pour usage propre. Le revenu fiscal n'est utilisé que comme approximation en cas de non-disponibilité de la première information.

Valeur des biens produits pour la consommation propre (PY070N)

Il s'agit de la production agricole du ménage au prix du marché : le montant est estimé par le ménage.

Allocations de chômage (PY090N)

Elles contiennent les allocations chômage.

Allocations de vieillesse (PY100N)

Ce sont les retraites de droit propre comme de droit dérivé, les préretraites (à l'exclusion de celles incluses dans les allocations chômage et les pensions d'invalidité) et le minimum vieillesse servies à 65 ans ou plus. Le minimum vieillesse est un minimum social versé sous condition de ressources aux personnes ayant atteint 65 ans. Les retraites facultatives sont incluses également car elles ne sont pas dissociables des allocations de vieillesse dans les déclarations des impôts.

Pensions au survivant (PY110N)

Ce sont les pensions de réversion, c'est-à-dire les pensions reçues par les survivants n'ayant pas atteint 65 ans. Y sont inclus également les retraites facultatives de ces personnes.

Retraites non obligatoires (PY080N)

Ce sont les retraites facultatives perçues par les personnes ne percevant ni de droit direct ni de droit dérivé (soit avec PY100N et PY110N =0) et quel sur soit leur âge.

Indemnités de maladie (PY120N)

Elles contiennent les indemnités journalières pour congés maladie, accidents du travail, congés maternité et les pensions d'accidents du travail.

Pensions d'invalidité (PY130N)

Les allocations suivantes sont comptées dans cette composante du revenu :

- l'allocation adulte handicapé (AAH) qui est versée aux handicapés de plus de 20 ans sous condition de ressources,
- les pensions militaires d'invalidité ou de victime de guerre,
- les pensions d'invalidité,
- les préretraites pour cause de diminution de la capacité de travail,

Sont aussi comptabilisées dans cet agrégat des aides permettant de financer la garde de personnes invalides ou dépendantes :

- l'aide personnalisée à l'autonomie (APA) qui est versée aux personnes âgées dépendantes qui ont recours à une aide à domicile,
- la prestation spécifique dépendance (PSD) qui remplit le même rôle que l'APA,
- l'allocation compensatrice pour tierce personne (ACTP), qui est versée aux handicapés de moins de 60 ans qui ont besoin d'une aide à domicile.

Les pensions versées aux personnes ayant atteint l'âge de la retraite sont incluses dans cet agrégat.

Allocations d'études (PY140N)

Il s'agit de l'ensemble des bourses versées aux plus de 16 ans et déclarées.

3.2.2 *La source ou la procédure utilisée pour la collecte des variables du revenu*

Les revenus sont récupérés par appariement avec les fichiers fiscaux et sociaux principalement et collectés par enquête pour ceux non imposables ou pour les personnes pour lesquels nous savons que l'appariement ne sera pas possible (cf. la description de la procédure utilisée pour recueillir les revenus, partie 2.3.2)

3.2.3 *La forme sous laquelle ont été obtenues les variables du revenu au niveau des composantes*

Les variables de revenu sont déclarées nettes des cotisations sociales mais y compris des contributions sociales non déductibles. En l'absence d'abattement, il a été facile de réestimer les montants nets.

3.2.4 *La méthode utilisée pour obtenir les variables cibles du revenu sous la forme requise (c'est-à-dire en tant que valeur brute)*

Les impôts sont récupérés par appariement. Les cotisations sociales ont été imputées sur barème.

Des valeurs brutes et nettes ont été estimées pour les agrégats suivants.

Salaire ou revenu assimilé (PY010G et PY030G)

Le salaire est récupéré par appariement. Dans le cas où il est collecté, il est dans la plupart des cas celui déclaré au fisc.

À partir de cette valeur nous calculons le salaire brut correspondant. Pour cela trois groupes de salariés sont constitués : les agents de l'État, les salariés du privé non cadres et les cadres salariés du privé. En fonction du montant du salaire nous calculons sur barème le salaire brut correspondant au salaire déclaré. Nous cumulons les cotisations dues pour la partie du salaire inférieure à un plafond de la sécurité sociale, puis celles pour la fraction comprise entre un et trois plafonds de la sécurité sociale, puis entre trois et quatre plafonds, entre quatre et huit plafonds et enfin au-delà de huit plafonds. Les taux de certaines cotisations sont en effet différents selon le montant du salaire. Nous avons ainsi reconstitué le salaire brut à partir duquel les diverses cotisations sociales salariales et employeurs sont calculées à l'aide du barème.

Sont ainsi estimés la contribution sociale de solidarité (CSG) et la contribution de remboursement de la dette sociale (CRDS) non déductibles, la CSG déductible, les cotisations du régime général (maladie, vieillesse, veuvage, chômage et famille), les cotisations aux retraites complémentaires salariées (reçues par l'Agirc) et cadres (reçues par l'Arrco), le versement transport, la cotisation au Fond national d'aide au logement, la taxe d'apprentissage, la participation des employeurs à la formation professionnelle. Afin de contrôler les cotisations obtenues, des cas types ont été testés sur le logiciel de l'Union de recouvrement des cotisations de sécurité sociale et d'allocations familiales (URSSAF) donnant des résultats satisfaisants.

Bénéfices en espèces ou perte de trésorerie en rapport avec une activité indépendante (PY050)

Les cotisations relatives aux revenus des indépendants sont estimées en répartissant les revenus en trois groupes : bénéfices agricoles, bénéfices industriels et commerciaux, bénéfices non commerciaux. Les régimes sociaux des indépendants sont très divers ; des paramètres législatifs moyens ont donc été utilisés.

Retraites de régimes privés, allocations chômage (PY090), de vieillesse (PY100) et pension au survivant (PY110)

La valeur brute est égale à la valeur nette à laquelle on ajoute la CSG et la CRDS. Ces contributions sont simples à estimer parce que leur taux ne dépend que du type de revenus et non du montant de revenu.

Loyers imputés (HY030)

La valeur nette est égale à la valeur brute, qui est imputée, diminuée de la taxe foncière payée sur la résidence principale, qui est collectée.

Prestations familiales (HY050) et allocations logement (HY070)

La valeur brute est égale à la valeur nette, collectée, à laquelle s'ajoute la CRDS, imputée.

Cotisations à un régime de retraite privé (PY035), valeur des biens produits pour la consommation propre (PY070), indemnités maladie (PY130), transferts entre ménage (HY080 et HY130), remboursement de prêts hypothécaires (HY100)

La valeur brute est égale à la valeur nette.

4 COHÉRENCE

4.1 Comparaison des variables cibles avec des sources extérieures

Les données externes ayant servi à contrôler les éléments constitutifs du revenu sont diverses. Nous avons utilisé l'enquête Revenus fiscaux et sociaux et l'enquête Logement. L'enquête Revenus fiscaux et sociaux des ménages est considérée comme la source de référence sur la distribution des revenus et des niveaux de vie en France. Elle consiste en un appariement entre l'enquête Emploi et les déclarations fiscales. Les prestations sociales y étaient imputées sur barèmes jusqu'en 2005 puis collectées auprès des organismes sociaux depuis 2006.

Nous avons aussi mobilisé les statistiques de la Caisse nationale des Allocations familiales. Enfin certains dossiers sont venus compléter cette aide extérieure, entre autres le dossier sur les revenus sociaux 2007 publié par la Direction de la recherche, des études, de l'évaluation et des statistiques.

4.1.1 Salaires et revenus assimilés (PY010N)

Année de revenus	EU-SILC		ERFS	
	2006	2007	2006	2007
Salaires (milliards d'€)	448	486	509	527
Bénéficiaires (millions)	25,4	26,4	26,4	26,8
Moyenne	17 622	18 434	19 258	19 680
Minimum	10	10	1	1
D10	3 869	3 774	3 487	3 660
D20	7937	8 361	8 041	8 260
D30	11 607	12 129	12 505	12 676
D40	13 929	14 576	14 927	15 372
D50	15 925	16 487	17 034	17 532
D60	18 000	18 676	19 276	19 849
D70	20 649	21 445	22 058	22 653
D80	24 021	25 084	26 103	26 630
D90	30 840	32 082	33 761	34 161
Maximum	1 590 389	594 622	765 766	2 282 103

Les écarts entre les deux enquêtes sont de l'ordre de 8 % en masses (- 5 points par rapport à 2007) et de 1,5 % en nombre de bénéficiaires. Ces variations trouvent en partie leur origine du fait que la csg non déductible et la crds sont comprises dans le salaire ERFS, tandis qu'elles ont été enlevées du salaire SRCV.

4.1.2 Les allocations chômage (PY090N)

	EU-SILC		ERFS	
	2006	2007	2006	2007
Revenus				
Chômage (milliards d'€)	24,8	20,2	25,11	24,28
Bénéficiaires (millions)	3,75	3,34	4,46	4,22
Moyenne	6 612	6 055	5 629	5 751
Minimum	28	50	8	1
D10	800	867	621	634
D20	1 576	1 445	1 325	1 337
D30	2 544	2 254	2 144	2 243
D40	3 575	3 230	3 059	3 112
D50	4 788	4 478	4 078	4 145
D60	5 600	5 242	5 346	5 328
D70	7 460	6 587	6 510	6 455
D80	9 600	8 531	8 833	8 736
D90	12 653	11 144	11 739	12 178
Maximum	157 154	433 555	74 702	113 039

Les écarts sur l'estimation du nombre de bénéficiaires sont plus importants sur ce poste. La différence observée s'explique par des différences de définition. SILC ne prend en compte que les allocations chômage, tandis que ERFS cumule dans ce poste les allocations chômage, toutes les allocations de préretraites du privé ou les allocations de congés de fin d'activité du secteur public, les indemnités de départ à la retraite, les primes de départ en retraite, les indemnités de membre de divers conseils, les indemnités des élus et des parlementaires.

4.1.3 Retraites de régime privé (PY080N), allocations vieillesse (PY100N) et pension au survivant (PY110N)

Année de revenus	EU-SILC		ERFS	
	2006	2007	2006	2007
Retraites (milliards d'€)	202	230	201	212
Bénéficiaires (millions)	13,6	14,9	14,6	15,0
Moyenne	14 850	15 483	13 732	14 115
Minimum	31	22	1	4
D10	4 340	3 960	2 696	2 738
D20	7 313	7 499	5 090	5 362
D30	9 245	9 736	7 778	8 149
D40	11 282	11 852	9 847	10 243
D50	13 327	13 878	12 139	12 547
D60	15 189	15 992	14 484	14 817
D70	17 405	18 319	16 922	17 273
D80	20 707	21 902	20 307	20 768
D90	26 073	27 281	25 839	26 399
Maximum	241 363	247 768	230 839	234 108

L'appariement avec les données fiscales a permis d'améliorer la collecte des revenus de retraites plus sensibles aux effets mémoires des enquêtés. Par contre, en 2008, les montants de retraites sont ceux déclarés aux impôts et récupérés par appariement ; ils ne contiennent pas la majoration de retraite pour trois enfants ou plus qui n'est pas imposable. Dans SILC, 10 %⁴ de majoration a donc été rajoutée aux retraités ayant eu au cours de leur vie 3 enfants ou plus : le nombre d'individus percevant cette majoration et la masse des majorations est conforme à ce qui est connu par ailleurs. Cela amène une différence avec l'enquête ERFS qui ne tient pas compte de cette majoration (la descendance n'étant pas connu dans l'enquête).

⁴ Si le montant de la majoration versée par la caisse nationale d'assurance vieillesse, qui verse les retraites des salariés du privé, est fixée à 10 %, ce n'est pas le cas de toutes les caisses de retraite : certaines appliquent une majoration variable selon le nombre d'enfants. Nous ne disposons cependant pas dans l'enquête du nombre et du nom des caisses qui versent une pension aux retraités : seuls le dernier emploi est connu, alors qu'il faudrait idéalement disposer de l'ensemble du parcours professionnel pour estimer une majoration. Seule une estimation grossière pourrait être faite.

Bénéfices en espèces ou perte de trésorerie en rapport avec une activité indépendante

Année de revenus	EU-SILC		ERFS	
	2006	2007	2006	2007
Revenus d'indépendants (milliards d'€)	51,7	53,1	54,2	56,1
Bénéficiaires (millions)	2,16	2,33	2,02	2,00
Moyenne	23 947	22 780	26 758	28 109
Minimum	100	2	1	1
D10	3 000	1 298	1 486	1 432
D20	7 000	3 429	3 662	3 824
D30	10 000	6 000	6 675	7 007
D40	13 000	9 200	9 847	10 185
D50	17 000	12 000	13 200	14 521
D60	20 603	15 994	17 816	20 715
D70	25 200	20 615	23 915	29 228
D80	33 060	29 500	36 355	41 210
D90	48 837	49 772	59 449	66 610
Maximum	548 816	1 500 000	2 410 096	1 414 689

Les revenus des indépendants sont abordés de façon différente dans l'ERFS et le panel, aussi les comparaisons ne sont pas possibles. Le panel privilégie le revenu privé, à savoir les sommes que l'indépendant prélève sur les ressources de l'entreprise à des fins privées, pour consommer ou épargner. L'ERFS enregistre un résultat fiscal ayant pour base un forfait, un bénéfice réel ou un régime micro-entreprise, résultant d'abattements, d'amortissements et de déductions. Ce résultat peut être négatif, le cas échéant il n'a pas été retenu pour la comparaison.

4.1.4 Revenus sociaux (HY050N, HY060N et HY070N)

Année de revenus	EU-SILC		ERFS	
	2006	2007	2006	2007
Sociaux (milliards d'€)	41,25	38,3	42,4	44,9
Bénéficiaires (millions)	10,34	11,0	11,0	12,1

Les bénéficiaires se comptent en ménages.

4.1.5 Prestations familiales (HY050N),

Année de revenus	EU-SILC		Cnaf	
	2006	2007	2006	2007
Prestations familiales (milliards d'€)	22,00	19,8	25,3	25,9
Bénéficiaires (millions)	6,3	6,1		

Les différentes allocations versées pour les familles sont regroupées ici. La comparaison est faite avec les statistiques de la Caisse nationale des Allocations familiales. Le champ couvert par la Cnaf est plus large que celui de SILC : les bénéficiaires ne vivant pas en ménage ordinaire y sont inclus. Cette différence de champ ne peut cependant expliquer l'ensemble de l'écart.

La Cnaf ne publie que le nombre de bénéficiaires au 31 décembre 2007 qui ne peut être comparé avec le nombre de personnes déclarant dans SILC avoir perçu une prestation au cours de l'année 2007.

4.1.5.1 Aides au logement (HY070N)

Année de revenus	EU-SILC		Cnaf	
	2006	2007	2006	2007
Allocations logements (milliards d'€)	14,4	11.6	13,2	13.2
Bénéficiaires (millions)	6,5	6.9		

La Cnaf ne publie que le nombre de bénéficiaires au 31 décembre 2007 qui ne peut être comparé avec le nombre de personnes déclarant dans SILC avoir perçu une prestation au cours de l'année 2007.

4.1.6 Impôts (impôt sur le revenu des personnes physiques et taxe d'habitation).

Année de revenus	EU-SILC		ERFS	
	2006	2007	2006	2007
Impôts (milliards d'€)	53,7	50.6	53,9	57.2

Les deux enquêtes diffèrent sur la masse de l'impôt sur le revenu. La période de référence n'est pas la même : l'ERFS retient l'impôt de l'année des revenus tandis que le panel collecte l'impôt de l'année précédant les revenus. Mais ceci n'explique qu'une partie de la différence observée.

4.1.7 Loyers

Les comparaisons montrent que les loyers collectés dans EU-SILC sont de bonne qualité.

Un loyer est attribué à chaque locataire dans EU-SILC. Le loyer est collecté par enquête. En cas de non-réponse ou de loyer en dessous des prix du marché, un loyer fictif lui est attribué (cf. partie 2.6).

Nous présentons les comparaisons avec l'enquête logement 2006.

	Enquête Logement 2006	EU-SILC 2008
Effectif	10 353 930	10 039 062
Moyenne	400	417
Q0	14	1
Q5	175	165
Q10	211	216
Q25	269	280
Q50	356	382
Q75	475	510
Q90	617	652
Q95	728	770
Q100	3 951	2 930

4.1.8 Intérêts, dividendes, bénéfices tirés d'investissements en capital dans les opérations incorporelles (HY090N)

Dans le panel EU-SILC, les revenus du patrimoine financier sont calculés à partir des stocks collectés et appariés. Six postes sont distingués lors de la collecte, les livrets d'épargne exonérés, les livrets bancaires soumis à l'impôt sur le revenu, l'épargne logement, les revenus de valeurs mobilières, les produits d'assurance vie, les autres placements financiers. Un taux de rendement issu de l'enquête Patrimoine de 2004 est appliqué aux stocks collectés afin d'estimer le revenu.

La différence sur les masses financières entre EU-SILC et ERFS résulte du fait que les revenus financiers non imposables sont imputés dans ERFS tandis qu'ils ne sont que collectés par enquête dans SRCV : or les montants déclarés dans les enquêtes sur le patrimoine sont très généralement sous-estimés.

Année des revenus	EU-SILC		ERFS	
	2006	2007	2006	2007
Revenus du patrimoine financier (milliards d'€)	20,21	24,8	53,9	55,8
Bénéficiaires (millions)	19,84	20,8	23,7	23,9
Moyenne	1 019	1 197	2 275	2 328
Minimum	2	1	1	1
D10	25	12	31	37
D20	67	23	91	103
D30	92	85	182	206
D40	253	85	319	357
D50	325	242	508	565
D60	513	242	816	882
D70	788	463	1 318	1423
D80	1 371	793	2 305	2386
D90	2 680	1 847	5 105	5134
Maximum	20 040	417 761	324 131	429 550

4.1.9 Revenu disponible du ménage et niveau de vie des individus.

Année des revenus	EU-SILC		ERFS	
	2006	2007	2006	2007
Revenus disponible des ménages (milliards d'€)	759,62	832,44	842,52	870,01
Bénéficiaires (millions)	26,27	26,7	25,88	26,8
Moyenne	28 911	31 138	32 557	32 462
Minimum	52	- 576 713	9	- 67 626
D10	11 238	11 947	12 207	11 725
D20	14 386	15 493	15 707	15 495
D30	17 450	18 783	19 127	18 994
D40	21 006	22 336	22 950	22 894
D50	24 905	26 299	27 152	27 142
D60	29 107	30 524	31 701	31 791
D70	33 861	35 579	36 617	37 269
D80	40 442	42 314	43 782	44 655
D90	50 469	53 976	56 290	57 224
Maximum	1 041 917	812 767	2 146 463	1 438 606

Revenus	EU-SILC		ERFS	
	2006	2007	2006	2007
Niveau de vie des individus (milliards d'€)	487,44	533,38	537,07	557,38
Bénéficiaires (millions)	59,57	60,67	59,46	60,55
Moyenne	18 588	19 886	20 605	20 859
Minimum	52	- 274 625	9	- 67 626
D10	9 321	9 425	9 724	9 709
D20	11 488	11 902	12 091	12 208
D30	13 192	13 906	13 962	14 198
D40	14 803	15 744	15 756	16 077
D50	16 623	17 493	17 598	18 031
D60	18 440	19 460	19 685	20 216
D70	20 861	21 844	21 241	22 864
D80	24 073	25 302	25 803	26 533
D90	29 627	31 822	33 194	33 716
Maximum	520 959	541 844	1 430 975	939 683

Ces comparaisons montrent que les revenus estimés dans EU-SILC sont davantage proches des revenus estimés par l'ERFS depuis que les revenus sont collectés par appariement dans SILC.

Annexe méthodologique

En présence de non-réponse, les estimateurs de sondage ne sont pas calculables tels quels. Une possibilité est de procéder à des imputations. On distingue souvent les méthodes dites « déterministe » ou « aléatoire ». Dans le cadre d'un modèle linéaire, une imputation aléatoire simple, par opposition à l'imputation déterministe, consiste à ajouter un nombre aléatoire simulé à la moyenne linéaire en les covariables, ce nombre correspond à une réalisation du résidu. Une imputation déterministe revient à remplacer la donnée manquante par la moyenne de la loi (dans le cas du modèle linéaire la moyenne est une moyenne conditionnelle à l'observation de covariables).

On peut justifier une telle approche lorsque l'on dit que l'on souhaite effectuer une prédiction de l'estimateur de sondage non calculable. Fournir cette prédiction est optimale au sens que nous obtenons in fine l'estimateur le moins risqué⁵. De ce point de vue la prédiction de données individuelles ne donne pas le résultat que l'on souhaite lorsque l'estimateur de sondage est non linéaire en les variables d'intérêt⁶. L'estimateur de sondage d'une médiane, d'un quantile ou d'un indicateur d'inégalité sont des quantités non linéaires en les variables d'intérêt. Le choix d'une imputation aléatoire, autrement dit par simulation, a donc été retenu. En effet, produire l'estimateur qui correspondrait à la moyenne d'estimateurs de sondage calculés sur plusieurs jeux indépendants de données simulées approcherait la prédiction de l'estimateur de sondage. Afin de satisfaire la demande d'Eurostat nous avons uniquement procédé à une imputation aléatoire simple. Lorsque le taux de valeurs simulées est très inférieur à celui des valeurs renseignées et conservées, la différence est négligeable. La différence est également d'autant plus faible que nous utilisons des informations auxiliaires par conditionnement (informations en fourchettes et jusqu'à un certain point l'information fournie par des covariables).

Nous allons détailler désormais les différentes étapes des simulations des montants manquants. La démarche est de procéder en deux étapes. La première étape est une étape d'estimation d'un modèle. La seconde étape est une étape de simulation des variables d'intérêt manquantes conformément au modèle préalablement estimé.

Nous avons modélisé les montants comme issus de tirages dans une loi log-normale dont la moyenne du logarithme est linéaire en des covariables et avec un facteur d'échelle constant. Les montants sont en général observés soit en clair soit en tranches. Bien qu'il soit possible de faire une estimation par maximisation de la vraisemblance il nous a semblé préférable de ne mener l'estimation que sur le groupe des répondants en clair avec une procédure d'estimation robuste (dans l'état actuel de nos connaissances nous ne savions pas mener une estimation robuste utilisant ces deux informations). Nous faisons implicitement l'hypothèse, qu'au sein des deux groupes, la loi (conditionnelle à l'observation des covariables) du montant est la même. Il est possible d'imaginer la présence de biais de sélection, mais nous avons pris le parti de les négliger. Les biais de sélection sont d'autant plus faibles que nous avons utilisé le conditionnement par des observations de covariables.

Le phénomène le plus préoccupant dans l'enquête EU-SILC est la présence d'erreurs. Une seule erreur pourrait en théorie induire un biais « infini » avec des estimations de type maximum de vraisemblance ou MCO. Nous avons donc utilisé des méthodes de régressions robustes à une contamination inférieure à 50% par des erreurs. La comparaison de différentes méthodes disponibles sous SAS IML nous a fait opter pour le CALL LMS (Least Median Squares de P.J. Rousseeuw). En quelque sorte, les paramètres correspondent à ceux d'un modèle qui rassemblerait les 50 % des valeurs les plus crédibles de l'échantillon. L'estimation a été faite sur des sous-groupes lorsque cela était nécessaire (défaut d'additivité) et car, pour des problèmes de complexité algorithmique, le call LMS nécessite de restreindre drastiquement le nombre de covariables (à peu près 15 covariables dichotomisées). Cette étape d'estimation a été menée sur le sous-échantillon des réponses en clair qui n'étaient pas déclarées suspectes au vu de règles comptables prenant en compte les autres déclarations faites par le ménage.

Dans un second temps nous avons fait l'approximation que les estimateurs des paramètres sont les vrais paramètres et nous avons produit des valeurs simulées.

⁵ Pour le risque quadratique calculé avec l'aléa de modèle.

⁶ Dans certains cas seulement la différence peut être négligeable.

Si aucune information n'est disponible sur la valeur du montant manquant, la simulation est faite dans la loi log-normale. Lorsqu'une plage de valeurs est disponible, ou reconstituée à partir de la connaissance d'autres variables, la simulation se fait dans la loi précédente conditionnelle au fait que la variable est dans la plage de valeurs prescrite. Dans ce second cas la simulation revient à la simulation de lois normales tronquées. La simulation dans la loi normale tronquée est effectuée par inversion de la fonction de répartition dès que cela est possible. Il est arrivé parfois que des problèmes numériques (plage loin de la valeur prédite par le modèle...) nous fassent utiliser une méthode d'acceptation/rejet avec optimisation de la probabilité d'acceptation. Cette méthode est une variante de la méthode de la macro « résidus simulés » mais est beaucoup plus efficace⁷.

⁷ Elle revient à tirer conjointement dans une loi uniforme et dans une famille à un paramètre de lois exponentielles translatées en un point de la troncature ou dans une loi uniforme dans certains cas où il y a troncature à gauche et à droite. Le critère d'arrêt découle du résultat général du principe d'acceptation-rejet. L'optimisation sur le paramètre permet de maximiser la probabilité d'acceptation en 1 coup. Dans ce cas à peu près 5 itérations suffisent pour obtenir une simulation pour toutes les données manquantes sans avoir à envisager de traitement adapté en cas de « non convergence ». Du point de vu du temps de calcul la simulation est quasiment instantanée mais légèrement plus longue que celle basée sur l'inversion de la fonction de répartition.