
CORRECTION DE LA SOUS-ESTIMATION DE LA FÉCONDITÉ PAR NIVEAU DE VIE MESURÉE À PARTIR DE L'EDP

Didier REYNAUD

Insee, Direction des statistiques démographiques et sociales

didier.reynaud@insee.fr

Mots-clés : Indices de fécondité, niveau de vie, appariements, Échantillon Démographique Permanent (EDP), rang de naissance

Domaine concerné : 14. Concepts et mesures / 14.1 Indicateurs, échelles, indices

Résumé

Suite à la publication d'une première étude de la fécondité selon le niveau de vie (Insee Première n°1826 « Les femmes les plus modestes et les plus aisées ont le plus d'enfants », novembre 2020), ce document de travail explicite la méthodologie utilisée et propose des améliorations méthodologiques. Il actualise également les résultats de l'étude grâce à ces améliorations.

La première partie présente une méthode destinée à corriger la sous-estimation de la fécondité telle que constatée dans l'étude publiée en 2020, réalisée à partir de l'EDP - l'Échantillon Démographique Permanent. La sous-estimation de la fécondité provient d'un appariement incomplet des naissances de l'état Civil aux femmes présentes dans les données socio-fiscales intégrées à l'EDP, appariement nécessaire afin d'étudier la fécondité selon le niveau de vie. C'est la réinjection de naissances non appariées initialement, qui, après ajustement, permet de retrouver un niveau de fécondité comparable au niveau de référence du Bilan démographique. Cette méthode permet, à partir de l'EDP, de disposer d'une base plus appropriée pour mener des études sur la fécondité selon différentes caractéristiques. Plus généralement, le travail réalisé montre qu'un appariement incomplet n'est pas forcément un obstacle, et que des variables non identifiantes peuvent être mobilisées pour combler l'information manquante.

La deuxième partie explicite l'indicateur choisi pour mesurer la fécondité selon le niveau de vie. Il s'agit de l'ICFRA – indicateur conjoncturel de fécondité tenant compte du rang et de l'âge, construit à partir des probabilités d'agrandissement observées une année donnée pour chaque âge et chaque rang de naissance. Compte tenu de certaines contraintes liées à l'EDP, une optimisation particulière du champ de l'étude est nécessaire afin de pouvoir disposer d'un échantillon de femmes suffisamment important tout en pouvant mobiliser la variable portant sur le nombre d'enfants déjà nés.

Les troisième et quatrième parties réviseront les résultats. Ce document les complètera également, avec deux années plus récentes et des croisements de variables plus détaillés.

Abstract en anglais

Following the publication of a first study of fertility by standard of living (Insee Première n°1826 "The most modest and well-off women have the most children", November 2020), this working document explains the methodology used and proposes methodological improvements. It also updates the results of the study with these improvements.

The first part presents a method intended to correct the underestimation of fertility as observed in the previous study, carried out from the EDP – the Permanent Demographic Sample, a large-scale socio-demographic panel established in France. The underestimation of fertility stems from an incomplete matching of births from the Civil Registry with the women present in the socio-fiscal data of the EDP, a matching necessary in order to study fertility according to standard of living. It is the reinjection of initially unmatched births which, after adjustment, makes it possible to find a level of fertility comparable to the reference level of the demographic report. This method makes it possible, from the EDP, to have a more appropriate basis for carrying out studies on fertility according to different characteristics. More generally, the work carried out shows that an incomplete match is not necessarily an obstacle, and that non-identifying variables can be used to fill in the missing information.

The second part explains the indicator chosen to measure fertility according to standard of living. This is the ICFRA – total fertility rate taking into account the birth order and age, constructed from the fertility rates by age observed in a given year for each birth order. Given certain constraints linked to the EDP, special optimization of the field of the study is necessary in order to have a sufficiently large sample of women while being able to use the variable relating to the number of children already born.

The forthcoming third and fourth parts will revise the results, without calling into question the conclusions of Insee Première on fertility according to standard of living. In particular, the results at the bottom of the living standards scale and for foreign-born women are significantly revised upwards. This document will also complete them, with two more recent years or more detailed cross-tabulations of variables.

Introduction

Une première étude sur la fécondité selon le niveau de vie a été menée en 2020 à partir de l'échantillon Démographique Permanent (EDP). Elle a donné lieu le 25 novembre 2020 à la publication de *l'Insee Première* n°1826 « Les femmes les plus modestes et les plus aisées ont le plus d'enfants ». Il était précisé que « Pour la période 2012-2017, la fécondité mesurée à partir des seuls taux de fécondité par âge s'établit en France métropolitaine à 1,94 enfant par femme, contre 1,78 pour les estimations avec l'EDP ». En conséquence, *l'Insee Première* annonçait des travaux complémentaires.

Les approfondissements méthodologiques annoncés ont été entrepris de façon à corriger la sous-fécondité constatée à partir de l'EDP et sont présentés dans ce document de travail.

Le présent document comporte quatre parties :

La **partie 1** présente les sources utilisées et le champ, ainsi que la méthodologie destinée à corriger la sous-estimation de la fécondité constatée.

La **partie 2** détaille les choix retenus et la méthode utilisée pour estimer la fécondité selon le niveau de vie, ainsi qu'une optimisation du champ de l'étude.

La **partie 3** actualise les principaux résultats de *l'Insee Première* sur la fécondité selon le niveau de vie, compte tenu des avancées méthodologiques présentées dans les parties 1 et 2 (disponible ultérieurement).

La **partie 4** présente des résultats supplémentaires plus détaillés (disponible ultérieurement).

Les résultats présentés dans ce document de travail (parties 3 et 4) ont été établis en moyenne sur la période de 2012 à 2017. Mais la méthode a également été testée sur les années 2011, 2018 et 2019, ces deux dernières années ayant été disponibles ultérieurement. L'analyse de l'évolution annuelle de la fécondité selon le niveau de vie a ainsi pu être complétée avec l'année 2011 (figures 47 à 53) et ainsi que les années 2018 et 2019 (figure 50).

1. Méthodologie de la correction de la sous-estimation de la fécondité des femmes

Après avoir rappelé le constat, nous détaillerons comment la fécondité a été mesurée pour l'étude publiée, les limites rencontrées, et les solutions mises en œuvre ici pour corriger la base des données.

1.1 Constat de départ : une fécondité sous-estimée à partir de l'EDP

L'étude intitulée « Les femmes les plus modestes et les plus aisées ont le plus d'enfants » (*Insee Première* n°1826), parue en novembre 2020, a été réalisée sur le champ des femmes âgées de 15 à 50 ans domiciliées en métropole et nées un 1^{er} ou 4 octobre, afin de disposer dans l'Échantillon démographique permanent (EDP) de l'historique des enfants qu'elles ont déjà eus (la profondeur historique sur les enfants déjà nés n'est en effet pas la même selon le jour de naissance de l'EDP) et d'estimer la fécondité par rang de naissance. Par ailleurs les femmes pour lesquelles le niveau de vie du ménage n'est pas disponible étaient exclues du champ de l'étude.

Si la source utilisée dans ce document de travail est la même (EDP), le champ d'analyse est toutefois différent de celui de l'étude déjà parue. Dans cette première partie, le champ est étendu à l'ensemble des femmes de 15 à 50 ans domiciliées en métropole, et donc nées l'un des 16 jours qui caractérisent l'EDP et pas seulement nées un 1^{er} ou un 4 octobre (soient 2 jours parmi les 16 jours EDP)¹. Par ailleurs, les femmes pour lesquelles le niveau de vie du ménage n'est pas disponible sont réintégrées dans le champ, grâce à une imputation des niveaux de vie.

Pour l'actualisation des résultats de l'Insee première (parties 3 et 4), le champ est par ailleurs optimisé (*cf.* le paragraphe 4 de la partie 2), afin de pouvoir mobiliser le rang de naissance sur un champ qui ne soit pas trop réduit.

Seul l'EDP permet de comparer la fécondité entre différents groupes de femmes selon leur niveau de vie, leur lieu de naissance et de diplôme.

La base étude 2017 de l'EDP a donc été utilisée. Pour les femmes présentes dans les données socio-fiscales de l'EDP, on apparie les naissances d'enfants provenant de l'État Civil, source également présente dans l'EDP. La base ainsi construite est nommée ci-après « base initiale ». Les données socio-fiscales de l'EDP sont mobilisées pour les années 2012 à 2017, pour lesquelles on dispose des revenus perçus et des niveaux de vie des années 2011 à 2016.

Les taux de fécondité par âge ont été calculés en rapportant les naissances d'enfants nés d'une mère EDP d'un âge donné à l'effectif des femmes de cet âge d'après les données fiscales. La somme sur toutes les femmes des taux par âges ainsi calculés pour une année donnée donne l'indicateur conjoncturel de fécondité (ICF) pour cette année : c'est le nombre moyen d'enfants qu'auraient les femmes si elles avaient durant toute leur vie féconde la fécondité par âge observée cette année-là.

Les taux de fécondité ainsi construits sous-estiment la fécondité.

L'Indicateur Conjoncturel de Fécondité calculé à partir de l'EDP s'établit en effet en moyenne à 1,77 enfant par femme pour la période de 2012 à 2017 et sur le champ de la France métropolitaine. Ce niveau est relativement faible, puisque par comparaison, l'ICF de référence publié dans le Bilan démographique est de 1,94 sur la période et un champ identique (Figure 1). L'ICF moyen calculé à partir de l'EDP s'avère ainsi en moyenne inférieur de 8,3 % à celui du Bilan démographique.

¹ Car le revenu disponible, donc le niveau de vie, n'est disponible pour les DOM qu'à partir de l'année 2016
14^e édition des Journées de méthodologie statistique de l'Insee (JMS 2022)

Figure 1 - Indicateur conjoncturel de fécondité, selon la source et l'année

	2011	2012	2013	2014	2015	2016	2017
EDP base initiale (*)	1,87	1,82	1,83	1,81	1,77	1,74	1,68
Bilan démographique	2,00	1,99	1,97	1,97	1,93	1,89	1,86
Écart	-0,13	-0,17	-0,14	-0,16	-0,16	-0,15	-0,18
Écart en %	-6,4%	-8,5%	-7,4%	-8,3%	-8,1%	-8,1%	-9,7%

	2011-2017	2012-2017
EDP base initiale (*)	1,79	1,77
Bilan démographique	1,94	1,94
Écart	-0,15	-0,17
Écart en %	-8,0%	-8,3%

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent & Insee, Bilan démographique 2019

Champ du bilan démographique : France métropolitaine

* EDP base initiale = base constituée avant la réinjection des naissances qui permet de corriger la sous-estimation de la fécondité

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

1.2 Présentation générale de l'EDP et des données socio-fiscales

Remédier au problème constaté de sous-estimation de la fécondité dans l'EDP nécessite de revenir en détail sur le mode de construction de l'EDP et la sélection des observations mobilisées pour étudier la fécondité selon les niveaux de vie.

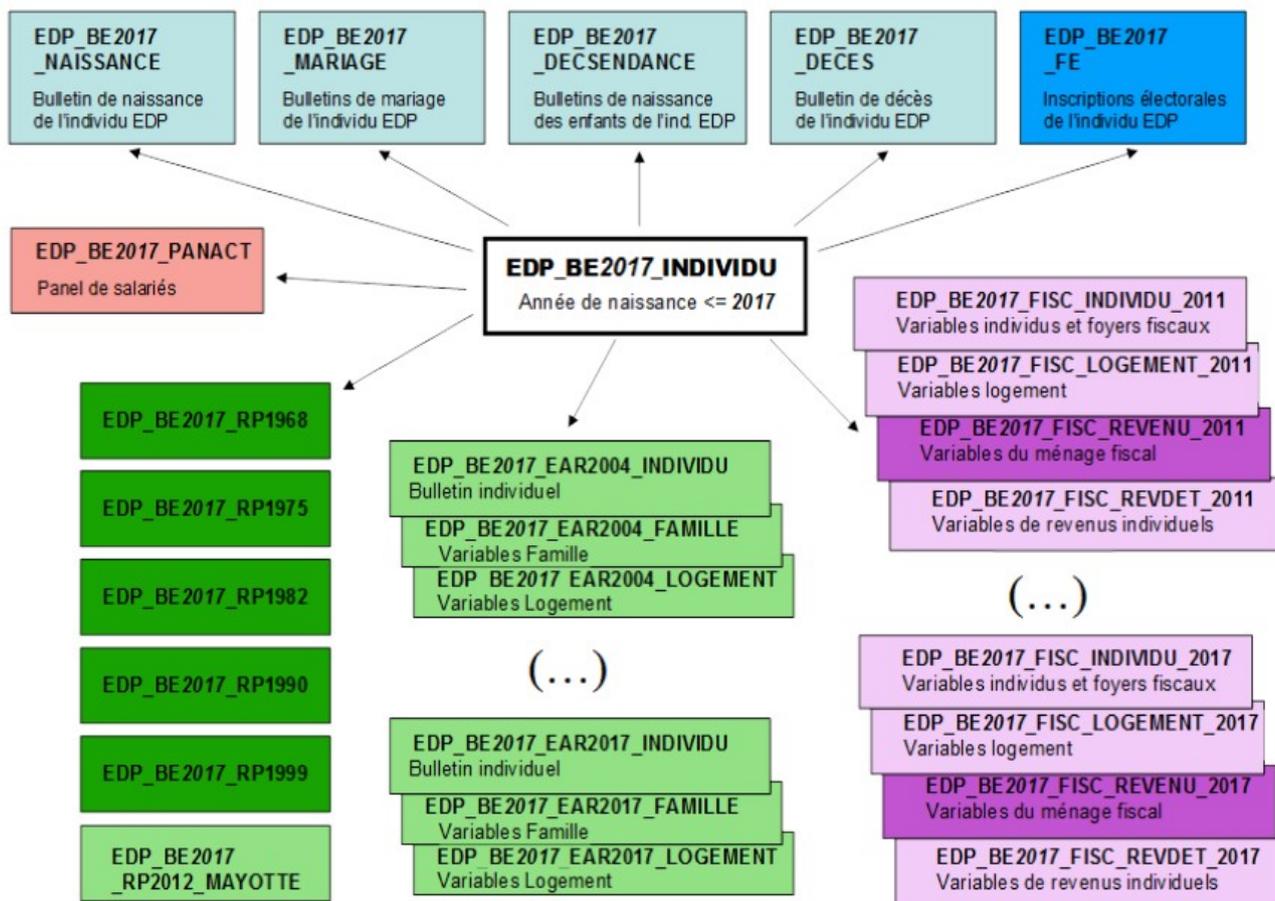
Comme l'*Insee Première* précédemment publié, ce document de travail porte sur la base étude 2017 de l'Échantillon démographique permanent. L'EDP compile, pour un échantillon d'individus nés certains jours dans l'année, dit « jours EDP », des données issues de différentes sources, notamment :

- l'État Civil, qui permet de repérer si les personnes ont eu un enfant une année donnée ;
- les sources socio-fiscales, qui permettent d'estimer le niveau de vie de l'année précédente ;
- les enquêtes annuelles de recensement, qui donnent le niveau de diplôme.

La base étude de l'EDP est constituée de différentes tables (Figure 2). Les tables mobilisées dans ce document de travail sont les suivantes :

- la table INDIVIDU qui est la table centrale de l'EDP, avec les principales caractéristiques des individus EDP ;
- la table DESCENDANCE, constituée, à partir de l'État Civil, des bulletins de naissances des enfants nés en France et ayant un parent né un des jours EDP ;
- les tables EARXXXX_INDIVIDU (XXXX = 2004 à 2017), qui regroupent les bulletins individuels des enquêtes annuelles de recensement de l'année XXXX, et apportent des informations socio-démographiques, tel que le diplôme des individus EDP (mais aussi des personnes qui vivent dans le même logement lors de l'interrogation du recensement) ;
- les tables FISC_INDIVIDU_XXXX (XXXX = 2011 à 2017), qui, pour chaque année fiscale compilent les informations sociales et fiscales des individus EDP et des autres individus du logement (Source : *fichier démographique des logements et des individus - fidéli*),
- les tables FISC_REVENU_XXXX (XXXX = 2011 à 2017), qui, pour chaque année fiscale, compilent les informations sur les revenus et les niveaux de vie des ménages constitués d'au moins un individu EDP (Source : *fichier localisé social et fiscal - filosofi*).

Figure 2 - Description de la base étude de l'EDP



Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

La table FISC_INDIVIDU_XXXX, construite à partir des données du dispositif Fidéli, est centrale pour l'étude de la fécondité selon le niveau de vie.

Elle regroupe les seuls individus des logements connus de Fidéli dans lesquels réside un individu EDP. C'est dans cette table que sont sélectionnées les femmes de 15 à 50 ans, nées un jour EDP ayant une adresse fiscale de taxation en métropole. Des coefficients de pondérations (variable poids_fideli) permettent de reconstituer l'effectif total de femmes présentes dans Fidéli, en se calant sur les marges de l'année fiscale considérée.

Le fichier individuel du répertoire statistique des logements (Fidéli), qui permet de construire la table FISC_INDIVIDU_XXXX, est alimenté à partir :

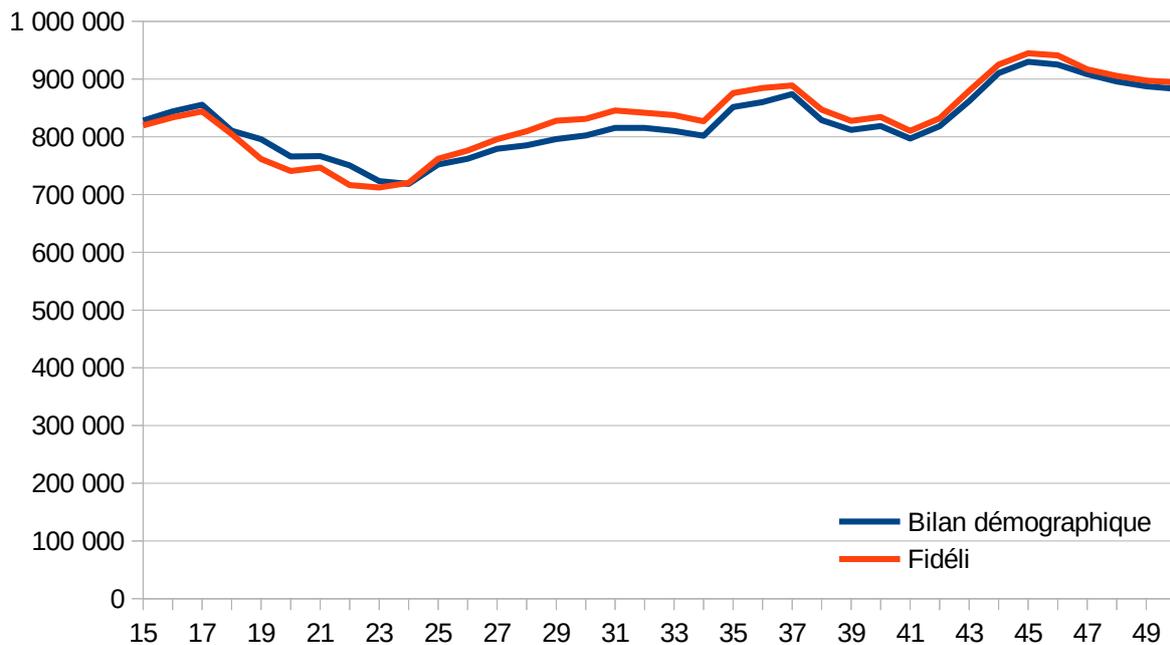
- du fichier d'imposition des personnes (FIP), rapproché du fichier des déclarations de revenus (POTE) afin de récupérer les informations des personnes à charge,
- et du fichier de la taxe d'habitation (PLFC).

Il est ainsi constitué des individus connus soit au titre de l'impôt sur le revenu, soit au titre de la taxe d'habitation principale. Les individus fiscaux « EDP » sont identifiés grâce aux informations nominatives dont dispose l'administration fiscale (nom, prénom, sexe, date de naissance, lieu de naissance).

De manière générale, le champ de Fidéli, qui s'appuie sur les données fiscales, est relativement exhaustif. Les effectifs en résultant sont proches de ceux publiés dans le Bilan démographique (Figure 3 et Encadré 1).

La table FISC_INDIVIDU_2017, restreinte aux femmes de 15 à 50 ans, nées un jour EDP et ayant une adresse fiscale de taxation en métropole, représente 14,555 millions de femmes de ces âges avec une adresse fiscale de taxation en métropole (Figure 4).

Figure 3 - Comparaison des effectifs de Fidéli par âge, de 15 à 50 ans, aux effectifs du Bilan démographique en 2018



Champ : France, individus âgés de 15 à 50 ans

Source : Insee

Encadré 1 : Détail des écarts d'effectifs entre Fidéli et le Bilan démographique

Le bilan démographique établit chaque année la population par sexe et âge. C'est la référence.

« Globalement, dans Fidéli, la population comprend davantage de personnes d'âge intermédiaire (25 à 49 ans) mais moins de jeunes (moins de 25 ans) et de personnes âgées (65 ans ou plus) que dans le bilan démographique.

Une raison à l'excès de population d'âge intermédiaire pourrait être que, malgré les filtres mis en place [note : pour établir des statistiques], les données fiscales contiennent des travailleurs non-résidents pour ces tranches d'âge. Par ailleurs, le regroupement de toutes les occurrences fiscales d'une même personne pour ne pas avoir de doublons comporte des limites.

Il manque dans la source des jeunes adultes de 18 à 22 ans. Il s'agit probablement de personnes vivant dans des communautés (et qui ne paient pas de taxe d'habitation) et ne déclarant pas de revenus, y compris dans le foyer parental. Les imputations mises en place pour corriger ces insuffisances ne permettent pas de compenser totalement les déficits.

Il manque aussi un peu de jeunes de moins de 18 ans, car la source identifie sans doute mieux les déclarants et conjoints des foyers que les personnes à charge. »

Extrait de la note sur les « Caractéristiques démographiques du millésime 2018 de Fidéli »

Figure 4 - Effectifs EDP mobilisés, après pondération à partir de la variable intitulée « poids_fideli », pour l'année 2017

Femmes de 15 à 50 ans, 2017	effectif	%
ensemble	14 554 903	100,0 %
nées en France	12 532 514	86,1 %
nées à l'étranger	2 022 389	13,9 %
diplôme inférieur au bac	4 339 189	29,8 %
diplôme de niveau bac	2 348 816	16,1 %
diplôme de niveau supérieur au bac	3 713 318	25,5 %
variable diplôme non disponible	4 153 580	28,5 %

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

Un premier appariement avec les tables FISC_REVENU_XXXX, permet de récupérer le revenu, qui provient du dispositif Filosofi.

Un second appariement avec les enquêtes annuelles de recensement de 2004 à 2017 permet de récupérer le niveau de diplôme d'un maximum de femmes. Pour certaines d'entre elles (29 %, figure 4), notamment celles qui n'ont pas été interrogées sur la période dans le cadre du recensement, le niveau de diplôme reste inconnu.

Un troisième appariement avec la table DESCENDANCE, alimentée par les bulletins de naissance de l'État Civil, permet de récupérer l'ensemble des naissances des enfants de femmes nées un « jour EDP », dans l'optique d'étudier leur fécondité.

1.3 Imputation des niveaux de vie manquants

Il s'avère que le niveau de vie du ménage est manquant dans près d'un cas sur vingt (Figure 5).

Figure 5 - Part des femmes dont le niveau de vie du ménage est manquant

2011	2012	2013	2014	2015	2016	2017
4,7%	4,7%	4,4%	4,5%	4,6%	4,4%	4,7%

Note : il s'agit des effectifs pondérés (avec la variable poids_fideli)

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

Pour ces femmes, le niveau de vie va être imputé.

Pour une femme dont le niveau de vie est manquant est attribué le niveau de vie d'un ménage comprenant une femme, tirée aléatoirement, parmi celles partageant des caractéristiques équivalentes au regard des variables suivantes² :

- âge
- lieu de naissance (France, Europe hors France, Maghreb, Afrique subsaharienne, Asie, Amériques, Océanie)
- jours de naissance (née les 1 & 4 octobre : oui / non). La profondeur historique sur les enfants déjà nés n'est en effet pas la même selon le jour de naissance de l'EDP,
- nombre d'enfants déjà nés
- niveau de diplôme (inférieur, égal ou supérieur au bac ou bien inconnu)

Hormis la variable de jour de naissance, qui servira à préciser ultérieurement le champ d'étude pour la fécondité par rang de naissance, les autres variables sont corrélées au niveau de vie. Des clés de tri aléatoire sont utilisées.

Figure 6 - Résultats de l'imputation

année	données avant imputation		données imputées		données après imputation	
	niveau de vie médian	niveau de vie moyen	niveau de vie médian	niveau de vie moyen	niveau de vie médian	niveau de vie moyen
2011	18 698	21 311	17 956	20 480	18 665	21 286
2012	18 981	21 711	18 144	20 713	18 941	21 677
2013	19 201	21 872	18 238	20 950	19 157	21 844
2014	19 406	21 822	18 599	20 989	19 371	21 799
2015	19 583	22 062	18 488	21 103	19 539	22 037
2016	19 692	22 210	18 662	21 311	19 650	22 185
2017	19 818	22 431	18 872	21 442	19 771	22 401

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

Les femmes dont le niveau de vie est manquant se voient imputer des niveaux de vie inférieurs, en médiane comme en moyenne, aux niveaux de vie des femmes pour lesquelles il est renseigné (Figure 6).

Cette opération permet au final de travailler sur l'ensemble du champ, en intégrant les femmes pour lesquelles le niveau de vie du ménage n'est pas disponible au départ.

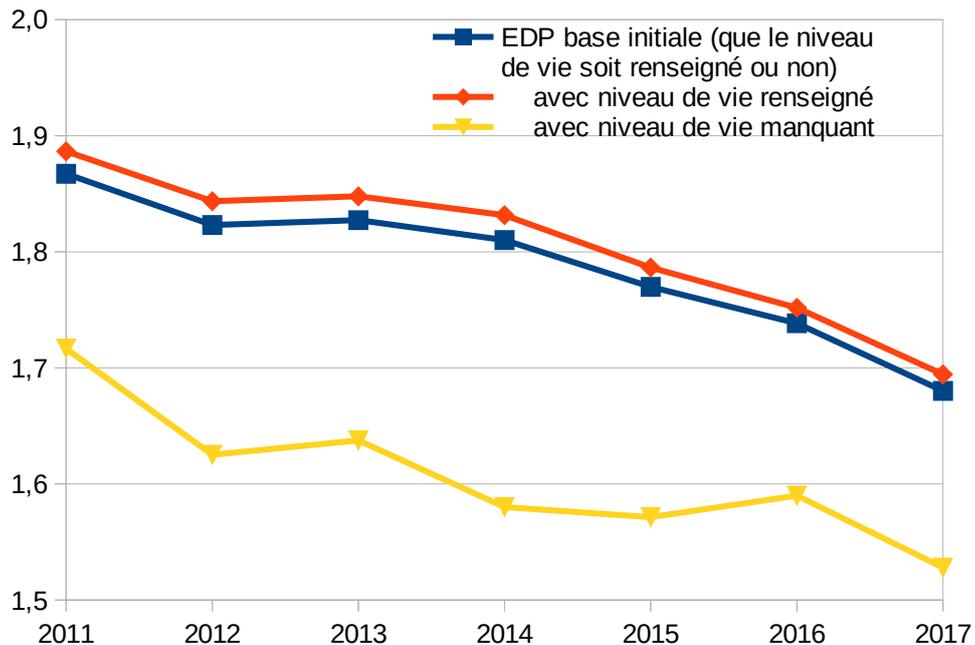
² Le niveau de vie est fortement corrélé à l'âge et au niveau de diplôme, tandis que le nombre d'enfants déjà nés entre directement dans le calcul du niveau de vie via le nombre des unités de consommation. Le lieu de naissance a été introduit dans la perspective de mieux suivre la fécondité des femmes nées à l'étranger. Le jour de naissance est pris en compte, car l'étude a été menée, selon les variables dont on a besoin, tantôt sur le champ des nées les 1 & 4 octobre, tantôt sur le champ complet.

Précédemment, ces femmes avaient été écartées du champ de l'Insee Première, ce qui a introduit un biais en termes de niveau de fécondité. Dans ce document de travail, ces femmes sont intégrées à l'étude : un niveau de vie leur est imputé en cohérence avec leurs caractéristiques socio-démographiques, ce qui supprime ce biais.

Ces femmes, pour lesquels le niveau de vie du ménage a donc été imputé, ont en effet un ICF nettement inférieur à celui des autres femmes (figure 7).

Cela est semble-t-il le signe que la qualité de l'appariement avec la table DESCENDANCE est de moins bonne qualité pour celles-ci. Nous verrons par la suite que les femmes dont le niveau de vie n'est pas disponible sont surreprésentées parmi celles pour lesquelles une naissance va être « réinjectée » (cf. figure 14).

Figure 7 - Comparaison de l'ICF selon que le niveau de vie est manquant ou non



Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

EDP Base initiale = base constituée avant la réinjection des naissances

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

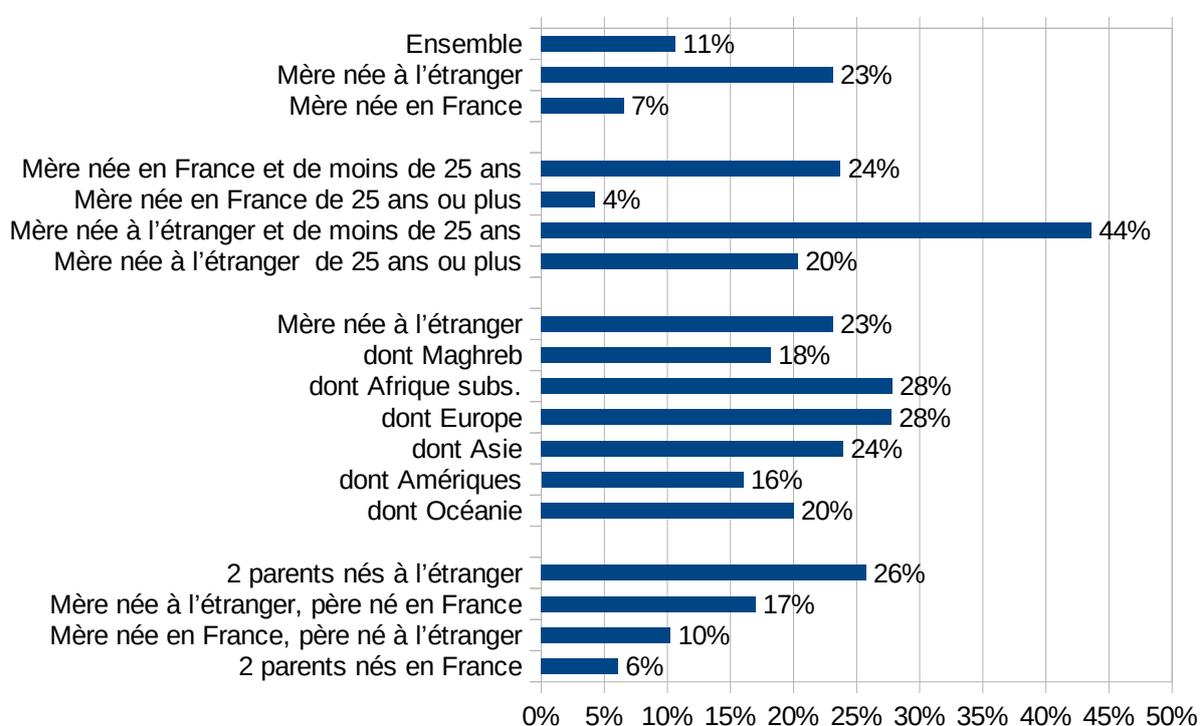
1.4 Analyse des naissances non appariées de la table DESCENDANCE

La table DESCENDANCE de l'EDP comprend l'ensemble des naissances des individus ayant un parent EDP. Ces informations sont issues des bulletins de naissance de l'État Civil. Un appariement est effectué entre les tables FISC_INDIVIDU et DESCENDANCE, afin de récupérer l'ensemble des naissances des femmes du champ de l'étude.

Après appariement, un certain nombre de naissances de la table DESCENDANCE ne correspondent à aucune des femmes présentes dans la table FISC_INDIVIDU.

En 2017, le taux de naissances non appariées est de 11 % au niveau global (Figure 8). Il varie en revanche fortement selon le lieu de naissance et l'âge de la mère. Il est de 23 % pour les mères nées à l'étranger et de 24 % pour les mères nées en France et de moins de 25 ans. Par opposition, il est de seulement 4 % pour les mères nées en France de 25 ans ou plus. Parmi les femmes nées à l'étranger, ce sont celles nées en Afrique et en Europe (hors France) pour lesquelles le taux de naissances non appariées est le plus important (28 %).

Figure 8 - Part des naissances 2017 de la table DESCENDANCE non appariées à une femme de la table FISC_INDIVIDU



Note : les caractéristiques des mères sont celles issues de l'État Civil, et proviennent donc de la table DESCENDANCE

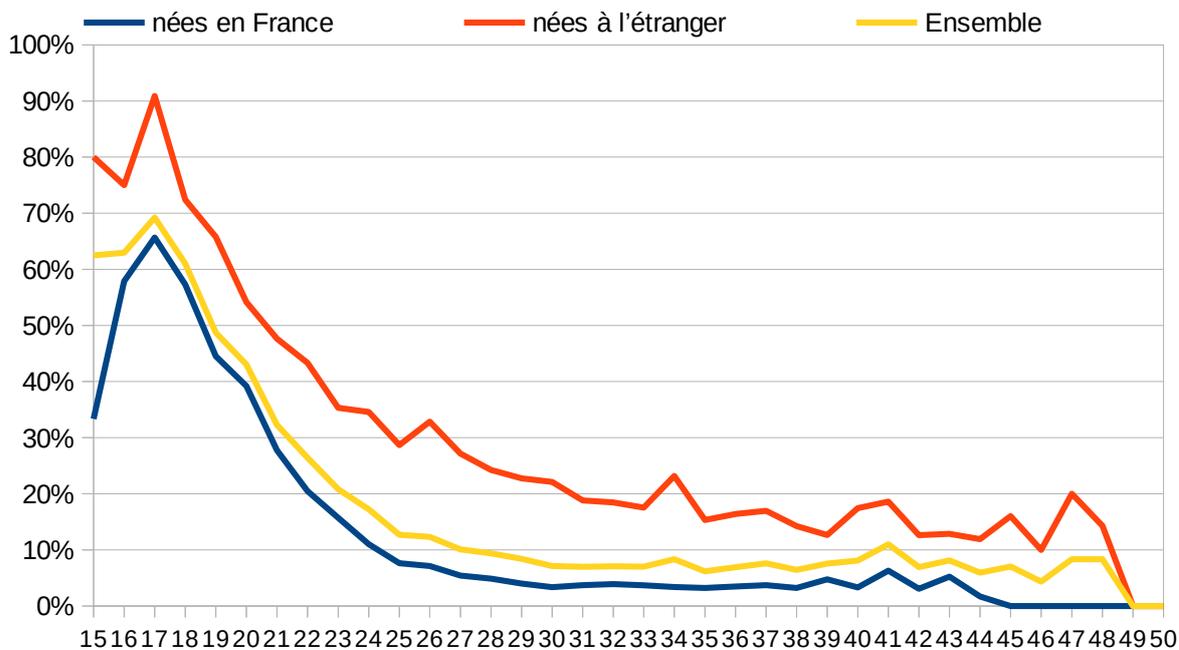
Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : naissances en France métropolitaine d'une mère née un jour EDP

La part des naissances de la table DESCENDANCE non appariées est très élevée pour les mères les plus jeunes (notamment en lien avec le déficit de jeunes adultes dans Fidéli, cf. figure 3). Cette part tend à décroître selon l'âge de la mère, et ce quel que soit le lieu de naissance (figure 9). À partir de 28 ans, les naissances non appariées représentent moins de 10 % du total.

Par ailleurs, pour les femmes de 26 ans ou plus nées en France, près de 4 % des naissances ne sont pas appariées.

Figure 9 - Part des naissances 2017 de la table DESCENDANCE non appariées, selon l'âge de la mère



Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

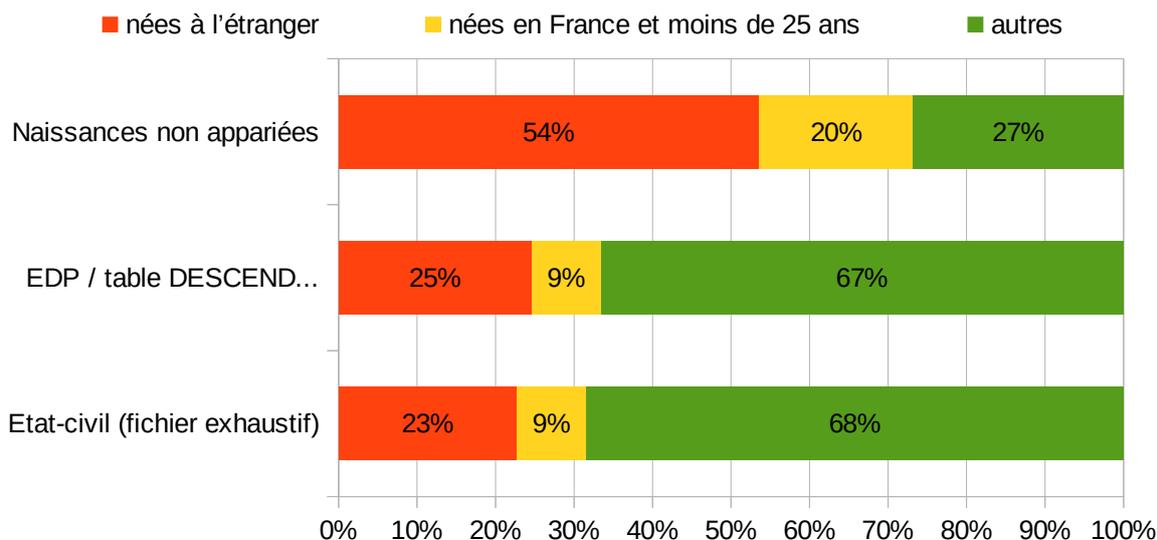
Champ : naissances en France métropolitaine d'une mère née un jour EDP

Parmi les naissances non appariées, deux catégories de mères sont ainsi surreprésentées :

- les mères nées à l'étranger,
- les mères nées en France de moins de 25 ans.

En 2017, plus de la moitié des naissances non appariées (54 %) sont relatives à des mères nées à l'étranger (Figure 10). Et 20 % des naissances non appariées sont relatives à des femmes nées en France de moins de 25 ans.

Figure 10 - Naissances 2017 non appariées : comparaison des caractéristiques des mères



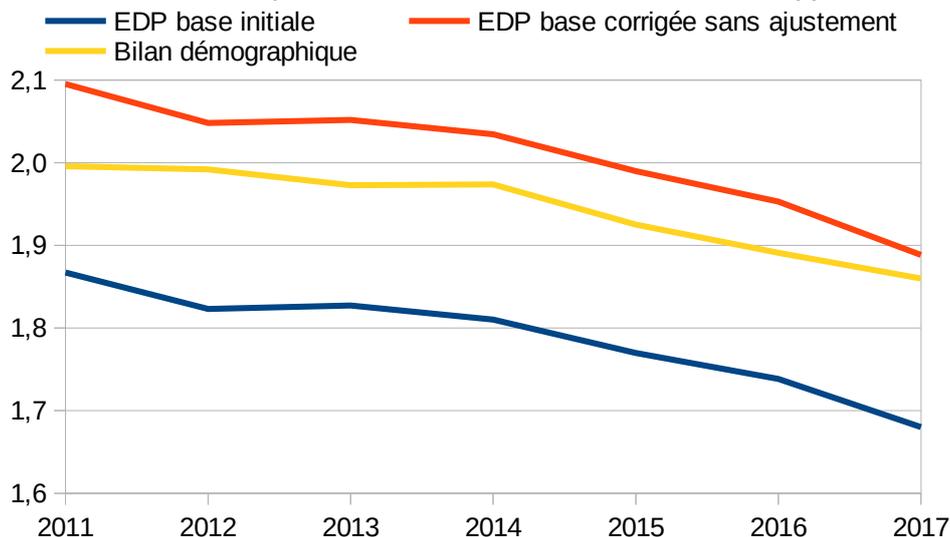
Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : naissances en France métropolitaine d'une mère née un jour EDP

NB : Les naissances de mères nées à l'étranger sont légèrement surreprésentées dans l'EDP, par rapport aux données exhaustives de l'État Civil. Cela est à relier à l'excédent de naissances sur certains jours EDP pour les femmes nées à l'étranger (cf. annexe 1).

Une solution consiste à réinjecter ces naissances non appariées en les attribuant aléatoirement à des femmes de la table FISC_INDIVIDU pour lesquelles aucune naissance n'a été appariée l'année considérée. Mais réinjecter l'ensemble des naissances non appariées conduit en définitive à une fécondité supérieure à la fécondité de référence (figure 11). Il faut donc procéder différemment et introduire un ajustement (cf. points 5 et 6 ci-après).

Figure 11 - ICF résultant de la réinjection de l'ensemble des naissances non appariées



Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

EDP Base initiale = base constituée avant la réinjection des naissances

EDP Base corrigée = base constituée après la réinjection des naissances

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

1.5 Évaluation de l'exhaustivité de la table FISC_INDIVIDU de l'EDP

Les échecs d'appariement entre la base DESCENDANCE et la base FISC_INDIVIDU proviennent, du moins en partie, du fait que la table FISC_INDIVIDU de l'EDP est incomplète : seules entrent dans l'EDP des personnes que l'on a pu « identifier ». Dans le processus de production de l'EDP, l'identification des individus, soit l'attribution d'un identifiant, est réalisée à partir du NIR (numéro d'identification au Répertoire national d'identification des personnes physiques, RNIPP³). Or le NIR n'est pas disponible dans les données socio-fiscales et doit être reconstitué. Cela est généralement possible à partir d'un certain nombre de variables caractérisant l'individu, mais pour certains individus dans les données fiscales il n'a pas été possible de leur affecter un NIR. Parmi ceux-ci, les mineurs ou jeunes adultes, ainsi que les personnes nées à l'étranger sont surreprésentés. Ne pouvant disposer de leur NIR, ces individus ne vont pas pouvoir être intégrés à la table FISC_INDIVIDU de l'EDP. Des coefficients de pondération (variable poids_fideli de la table FISC_INDIVIDU) ont été établis de manière à corriger cette sous-représentativité au niveau des effectifs. Ces poids sont issus d'un calage sur les marges des données fiscales exhaustives (Fidéli), en tenant compte des variables suivantes : le sexe, l'âge, la situation conjugale, le département de résidence, le lieu de naissance (France métropolitaine, DOM, étranger), le nombre de foyers par logement et le nombre de personnes à charge dans le foyer. Cependant, la prise en compte de ces coefficients de pondération ne redresse pas la sous-fécondité constatée à partir de l'EDP (cf. de fait l'Insee Première de novembre 2020), puisqu'on ne peut pas affecter une naissance de l'État Civil à une femme absente de la table FISC_INDIVIDU. Ces coefficients permettent en revanche d'obtenir à partir de la base FISC_INDIVIDU une image fidèle des données socio-fiscales (et notamment des effectifs par sexe et âge dans ces données, dont on a vu qu'elles différaient cependant un peu des effectifs de population de référence, i.e. ceux des Estimations de population présentées dans les Bilans démographiques, notamment pour les moins de 25 ans, cf. Figure 3).

Afin de pouvoir ajuster le nombre de naissances à réinjecter, nous allons évaluer, à partir des coefficients de pondération disponibles (variable poids_fideli), l'exhaustivité de la table FISC_INDIVIDU pour certaines catégories de femmes, en particulier les femmes les plus jeunes et celles nées à l'étranger.

Un coefficient de pondération de 22,83 (= 365,25 / 16) correspond à un coefficient de pondération « normal » puisque seules sont intégrées à l'EDP les femmes nées les 4 premiers jours de chaque trimestre, soit 16 jours de l'année sur 365,25 (en tenant compte des années bissextiles).

Un coefficient de pondération supérieur à 22,83 est le signe d'une sous-représentation, qu'il vise à compenser. Si par exemple, pour une catégorie de femmes donnée, le coefficient de pondération moyen constaté est de 24, on en déduit que seules 95,12 % (22,83 / 24) des femmes de cette catégorie figurent en moyenne dans la table FISC_INDIVIDU. Si pour une autre catégorie, le poids moyen constaté est de 30, alors on en déduit que seules 76,09 % des femmes de cette catégorie sont globalement prises en compte.

La formule générale, pour une catégorie de femmes donnée, est la suivante :

$$\text{exhaustivité apparente du champ (en \%)} = (365,25 / 16) / \text{poids moyen constaté} * 100$$

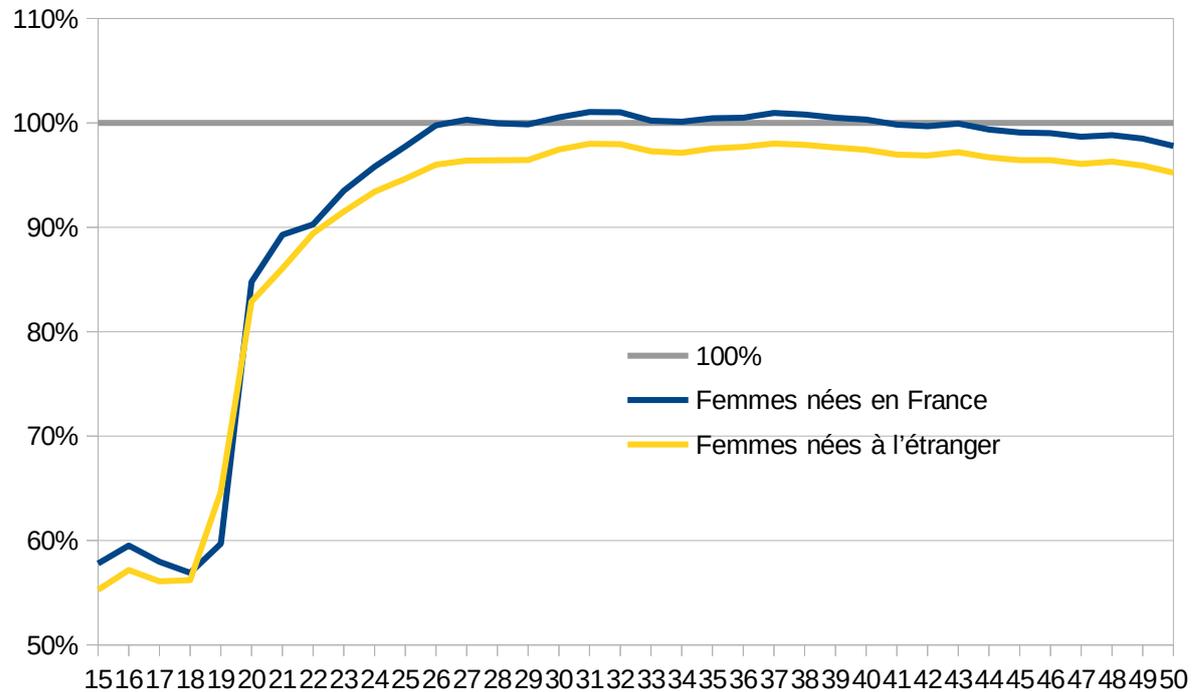
Les femmes sont réparties en sous-groupes, selon leur âge, leur date de naissance (nées les 1 et 4 octobre ou non) et leur lieu de naissance (France / Europe hors France / Maghreb / Afrique subsaharienne / Asie / Amériques / Océanie). Pour chaque sous-groupe est calculé le coefficient de pondération moyen et l'exhaustivité de la table FISC_INDIVIDU va ainsi être évaluée.

D'après les coefficients de pondération calculés à partir des deux seules variables suivantes, l'âge et le lieu de naissance (France / étranger), il s'avère que les femmes les plus jeunes sont nettement sous-représentées (Figure 12). Il manquerait ainsi près de 40 % des femmes de 15 à 18 ans. La part du

³ Cet identifiant est celui utilisé lors de la production de l'EDP, pour relier les sources entre elles pour un même individu. Il n'est pas conservé dans les bases d'études accessibles pour les chargés d'études et chercheurs. Un identifiant anonymisé (ID_DIFF) est introduit : non signifiant, il permet d'effectuer les appariements pour les études, sans recourir au NIR.

champ représenté progresse ensuite au fur et à mesure que l'âge augmente. Au-delà de 26 ans, le champ serait exhaustif pour les femmes nées en France, tandis qu'il manquerait environ 3 % des femmes nées à l'étranger dans la table FISC_INDIVIDU.

Figure 12 - Part des femmes présentes dans la table FISC_INDIVIDU de l'EDP, d'après les coefficients de pondération



Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

1.6 Détermination du nombre de naissances à réinjecter

Le constat d'un déficit d'appariement des naissances (base DESCENDANCE) aux individus de la base fiscale de l'EDP (base FISC_INDIVIDU) étant fait, que les pondérations poids_fideli ne permettent pas de combler totalement, une dernière opération consiste à ajuster le nombre de naissances manquantes à réinjecter, et à les intégrer dans la nouvelle table d'études destinée à analyser, dans ce document de travail, la fécondité selon le niveau de vie.

Le nombre de naissances, appariées ou non, de la table DESCENDANCE est connu. L'exhaustivité de la table FISC_INDIVIDU de l'EDP, évaluée à partir des coefficients de pondération est également connue, selon les différentes variables vues précédemment. Le nombre de naissances à réinjecter va désormais pouvoir être déterminé.

Si, par exemple pour un âge et lieu de naissance donné, X% des femmes figurent dans la table FISC_INDIVIDU de l'EDP, alors on doit réinjecter suffisamment de naissances (non appariées) pour avoir au total X % des naissances de la table DESCENDANCE effectivement appariées à une femme.

Les naissances réinjectées vont être attribuées aléatoirement à des femmes de la table FISC_INDIVIDU, pour lesquelles aucune naissance n'a été appariée l'année considérée, et qui partagent des caractéristiques identiques (des caractéristiques de la mère issues du bulletin de naissance). Les variables prises en considération sont les suivantes :

- âge
- lieu de naissance (France, Europe hors France, Maghreb, Afrique subsaharienne, Asie, Amériques, Océanie)
- jours de naissance (née les 1 & 4 octobre : oui / non)
- nombre d'enfants déjà nés
- niveau de diplôme (inférieur, égal ou supérieur au bac ou bien inconnu)

Hormis la variable de jour de naissance, qui servira à préciser ultérieurement le champ d'étude, les autres variables sont des variables qui influent sensiblement sur la fécondité.

Figure 13 - Détail de l'opération de réinjection de naissances

	2011	2012	2013	2014	2015	2016	2017
naissances table DESCENDANCE	35 229	34 994	35 040	34 804	34 369	33 695	32 719
dont naissances appariées	31 826	31 628	31 572	31 321	30 880	30 242	29 235
dont naissances non appariées	3 398	3 358	3 463	3 480	3 489	3 452	3 478
dont écarts divers liés aux données*	5	8	5	3	0	1	6
cible de naissances ajustée	34 228	34 164	34 385	34 161	33 639	33 129	32 431
ajout nécessaire	2 444	2 572	2 835	2 864	2 787	2 932	3 217
ajout possible	2 273	2 418	2 580	2 653	2 742	2 849	3 042
naissances ajoutées	2 258	2 401	2 574	2 630	2 715	2 833	3 029
naissances totales	34 084	34 029	34 146	33 951	33 595	33 075	32 264
% d'atteinte de la cible	99,6%	99,6%	99,3%	99,4%	99,9%	99,8%	99,5%

* : par exemple, la variable SEXE du parent peut dans de très rares cas diverger entre la table DESCENDANCE et la table INDIVIDU.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Note : il s'agit de la somme des calculs opérés dans chaque sous-groupe, d'où un résultat pouvant différer par rapport au même calcul réalisé globalement sur l'ensemble de la population.

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

Pour chaque sous-groupe de femmes défini selon l'âge, le lieu de naissance et le jour de naissance sont calculés :

- la cible de naissances ajustée, qui est déterminée à partir du total de naissances de la table DESCENDANCE, ajusté compte tenu de l'exhaustivité de la table FISC_INDIVIDU.
- l'ajout nécessaire, égal à la différence entre la cible ajustée et le nombre de naissances appariées, lorsque celle-ci est positive.
- l'ajout possible, qui est égal au minimum entre l'ajout nécessaire et le nombre de naissances non appariées, puisqu'on ne réinjecte pas plus de naissances que le nombre de naissances non appariées.
- le nombre de naissances effectivement ajoutées, qui est inférieur étant donné qu'il arrive de ne pas toujours trouver de femmes partageant des caractéristiques exactement identiques, auxquelles attribuer les naissances restant à appairer.

Dans la pratique, toutes les naissances non appariées sont réinjectées lorsqu'une femme sans naissance l'année considérée et aux caractéristiques similaires est trouvée. Puis dans un second temps, certaines naissances qui ont été réinjectées sont annulées, afin d'atteindre la cible lorsqu'elle est dépassée.

Au final, pour chacune des années et pour chacun des âges, suffisamment de naissances ont pu être réinjectées pour atteindre en général plus de 99 % de la cible.

Après contrôle des résultats, il apparaît que le niveau de vie était plus fréquemment manquant parmi les femmes pour lesquelles une naissance est réinjectée (figure 14). Cela justifie d'autant plus l'opération d'imputation du niveau de vie.

L'âge moyen des mères pour lesquelles une naissance est réinjectée est de 28,5 ans, contre un âge moyen de 33,2 ans pour l'ensemble des femmes du champ.

Plus de la moitié des naissances réinjectées concernent des femmes nées à l'étranger. Par ailleurs, le niveau de diplôme est plus souvent inconnu pour les femmes auxquelles est attribuée une naissance « fictive ».

Figure 14 - Analyse des caractéristiques des mères après imputation du niveau de vie et réinjection des naissances

moyenne 2011-2017	% niveau de vie manquant	niveau de vie médian*	niveau de vie moyen*	âge moyen
ensemble des femmes	4,6%	19 299	21 890	33,2
avec naissance (réelle) dans l'année	4,5%	20 293	22 341	30,6
avec naissance réinjectée (avant ajustement)	10,4%	15 878	18 380	27,1
avec naissance réinjectée (après ajustement)	9,7%	16 083	18 527	28,5
avec naissance dans l'année (réelle ou fictive)	4,9%	19 929	22 027	30,4

moyenne 2011-2017	% nées à l'étranger	% diplôme inférieur au bac	% diplôme égal au bac	% diplôme supérieur au bac
ensemble des femmes	13,1%	31,4%	17,5%	27,7%
avec naissance (réelle) dans l'année	20,0%	23,8%	16,7%	33,9%
avec naissance réinjectée (avant ajustement)	48,6%	32,0%	8,0%	8,8%
avec naissance réinjectée (après ajustement)	52,7%	29,6%	8,7%	10,9%
avec naissance dans l'année (réelle ou fictive)	22,7%	24,3%	16,1%	32,0%

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

1.7 Résultats en matière d'ICF

Suite à ces différentes opérations, nous disposons désormais d'une table de femmes, construite à partir de la table FISC_INDIVIDU de l'EDP, et comprenant leurs naissances correspondantes, qu'elles soient réelles ou imputées (de vraies naissances, mais appariées à des femmes qui n'en avaient *a priori* pas). Dans cette nouvelle base, notre base finale, les niveaux de vie manquants sont imputés et le déficit de naissances est résorbé.

Dans cette nouvelle base d'études, dite base corrigée par la suite, l'ICF est désormais très proche de celui publié dans le Bilan démographique : alors que sur la période 2012-2017, il était inférieur de 8,3 % à celui du Bilan démographique dans la base d'études initiale, il en est désormais très proche dans la base corrigée : + 0,2 % avec des écarts allant de - 0,8 % à + 1,0 %.

Figure 15 - Indicateur conjoncturel de fécondité, selon la source

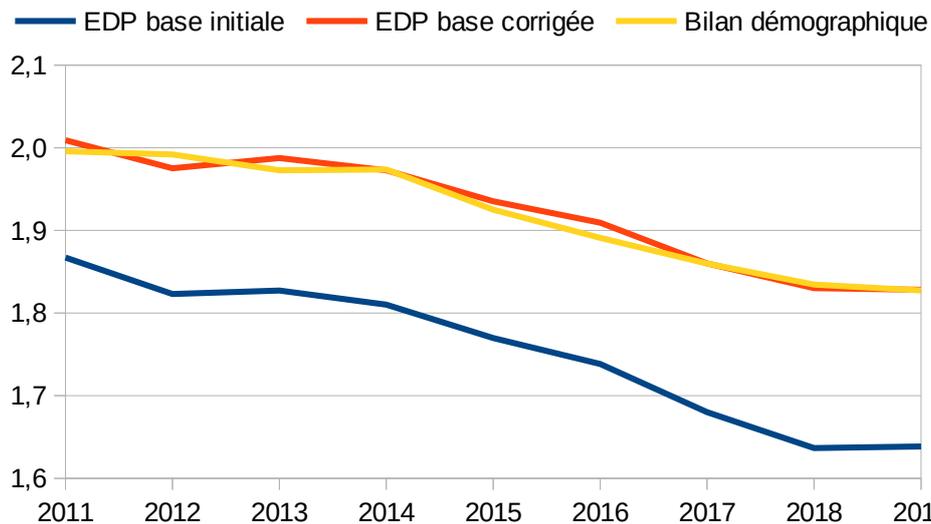
	2011	2012	2013	2014	2015	2016	2017
EDP base initiale	1,87	1,82	1,83	1,81	1,77	1,74	1,68
EDP base corrigée	2,01	1,98	1,99	1,97	1,94	1,91	1,86
Bilan démographique	2,00	1,99	1,97	1,97	1,93	1,89	1,86
Écart base initiale - bilan démographique	-0,13	-0,17	-0,14	-0,16	-0,16	-0,15	-0,18
en %	-6,4%	-8,5%	-7,4%	-8,3%	-8,1%	-8,1%	-9,7%
Écart base corrigée - bilan démographique	0,01	-0,01	0,02	0	0,01	0,02	0
en %	0,7%	-0,8%	0,8%	-0,1%	0,5%	1,0%	0,0%

	2011-2017	2012-2017
EDP base initiale	1,79	1,77
EDP base corrigée	1,95	1,94
Bilan démographique	1,94	1,94
Écart base initiale - bilan démographique	-0,15	-0,17
en %	-8,0%	-8,3%
Écart base corrigée - bilan démographique	0,01	0
en %	0,3%	0,2%

Source : Insee Bilan démographique & Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

Figure 16 - Comparaison de l'ICF avant et après correction de la base



Note : contrairement aux autres graphiques, les années 2018 et 2019, calculées à partir de la base étude 2019 ont été ajoutées à celui-ci.

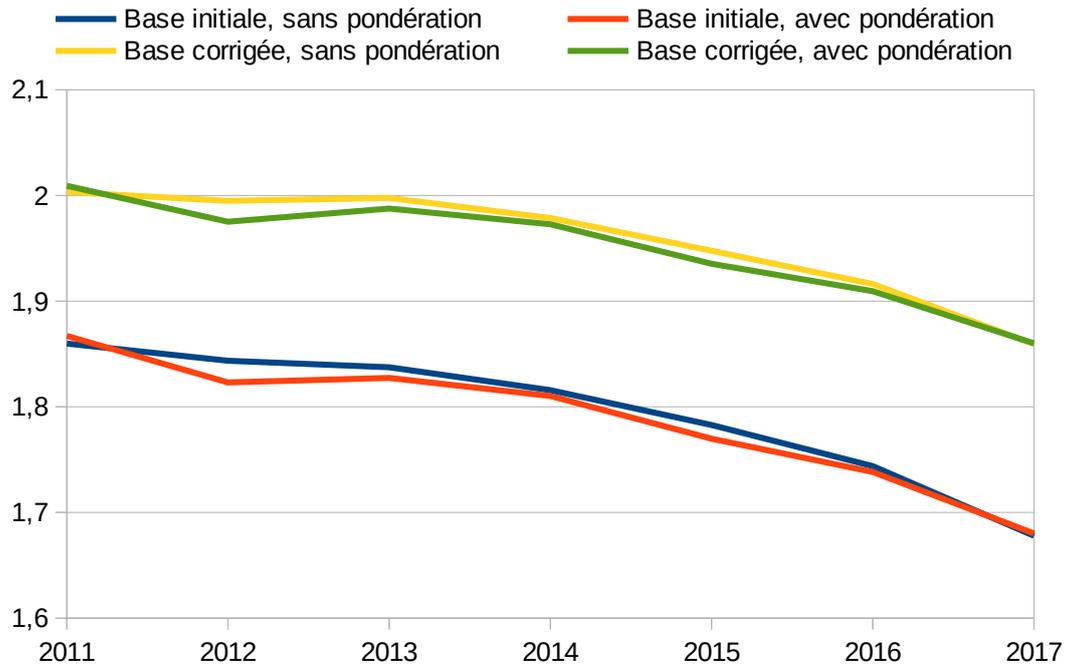
Source : Insee Bilan démographique & Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

C'est à partir de cette base corrigée que nous mettons à jour les résultats de l'étude sur la fécondité selon le niveau de vie publiée en novembre 2020. Dans la suite de ce document de travail, le champ est quelque peu réduit par rapport à ce qui vient d'être présenté dans le document, afin de pouvoir mobiliser la variable relative au rang de naissance, même s'il a été étendu par rapport à celui de l'IP.

Après vérification, la pondération à travers la variable « poids_fideli » ne corrige aucunement le déficit de fécondité (figure 17) : elle corrige les effectifs des personnes, mais sans pouvoir tenir compte du fait qu'elles ont ou non des enfants. L'opération de réinjection (avec ajustement) des naissances manquantes pour l'ensemble des femmes ayant entre 15 et 50 ans (y compris celles ayant un niveau de vie inconnu) était donc nécessaire et justifiée.

Figure 17 - ICF selon la prise en compte ou non des coefficients de pondérations



Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

Enfin, près de 4 % des naissances ne sont pas appariées (Figure 9) pour les femmes nées en France de 26 ans et plus. Or d'après la figure 12, la table FISC_INDIVIDU est exhaustive pour celles-ci.

Le problème d'identification lié à la construction de la table FISC_INDIVIDU (Fidéli) n'explique ainsi que partiellement les échecs d'appariement avec la table DESCENDANCE (État Civil).

2. Méthodologie de l'étude sur la fécondité selon le niveau de vie

2.1 Détermination des quantiles et construction des sous-groupes

Le niveau de vie est égal au revenu disponible du ménage divisé par le nombre d'unités de consommation (UC). Le revenu disponible d'un ménage comprend les revenus d'activité (nets des cotisations sociales), les revenus du patrimoine, les transferts en provenance d'autres ménages et les prestations sociales (y compris les pensions de retraite et les indemnités de chômage), nets des impôts directs. Le niveau de vie est donc le même pour tous les individus d'un même ménage. Les unités de consommation sont calculées selon l'échelle d'équivalence dite de l'OCDE modifiée qui attribue 1 UC au premier adulte du ménage, 0,5 UC aux autres personnes de 14 ans ou plus et 0,3 UC aux enfants de moins de 14 ans.

Le niveau de vie dépend généralement de l'âge des personnes qui composent le ménage. Les quantiles de niveau de vie ont ainsi été établis âge par âge. Une femme sera donc considérée comme « aisée » si elle fait partie des plus aisées de son âge.

À chaque âge entre 15 et 50 ans, les femmes sont donc classées du niveau de vie le plus faible au niveau de vie le plus élevé, puis réparties en q sous-groupes (n°1, n°2, ...n°q), avec q = 5, 10 ou 20, selon qu'on opte pour les quintiles, déciles ou vingtiles. L'ensemble des sous-groupes n°1 de chaque âge sont ensuite regroupés. Puis de même pour les sous-groupes n°2 de chaque âge, et ainsi de suite. On obtient au final q sous-groupes de femmes, dont la structure par âge est identique. La moyenne des niveaux de vie de chacun des q sous-groupes est ensuite calculée.

Exemple :

		Répartition par âge dans les sous-groupes				
		1 ^{er} cinquième	2 ^{ème} cinquième	3 ^{ème} cinquième	4 ^{ème} cinquième	5 ^{ème} cinquième
Âge	15	20%*	20%	20%	20%	20%
	16	20%	20%	20%	20%	20%
	17	20%	20%	20%	20%	20%
	18	20%	20%	20%	20%	20%
	19	20%	20%	20%	20%	20%
	...	20%	20%	20%	20%	20%

* : Il s'agit des 20 % des femmes âgées de 15 ans, les moins aisées en matière de niveau de vie.

Une hypothèse de stabilité de leur position dans l'échelle des niveaux de vie tout au long de la vie féconde

On calcule ensuite l'indicateur conjoncturel de fécondité des femmes dans chaque sous-groupe de niveau de vie. Cela revient à faire l'hypothèse qu'une femme reste dans le même sous-groupe tout au long de sa vie, c'est-à-dire dans la même position de niveau de vie : par exemple, l'ICF des plus aisées correspond à la fécondité des femmes qui seraient parmi les plus aisées à 15 ans, les plus aisées à 16 ans, ..., les plus aisées à 50 ans. En pratique, ces femmes peuvent ne pas être toute leur vie féconde dans la même position sur l'échelle des niveaux de vie. L'analyse de la fécondité par quantile de niveaux de vie repose ainsi sur des cas fictifs (stabilité de la position dans la distribution des niveaux de vie, fécondité par âge égale à celle constatée sur la période 2012-2017 tout au long de la vie féconde).

2.2 Choix de l'indicateur en matière de fécondité

L'indicateur conjoncturel de fécondité (ICF) a été calculé de deux manières différentes, l'approche classique et une approche tenant compte du rang [Rallu, Toulemon, 1993]. La fécondité (ICFRA – indicateur conjoncturel de fécondité tenant compte du rang et de l'âge) est mesurée ici à partir des probabilités d'agrandissement par âge observées une année donnée pour chaque rang de naissance : c'est-à-dire la probabilité d'avoir un premier enfant à un âge donné parmi les femmes sans enfant à cet âge, probabilité d'avoir un deuxième enfant parmi celles en ayant déjà un, etc. [Levy, 1986]. En combinant ces probabilités, on obtient le nombre moyen d'enfants d'une génération fictive de femmes qui auraient, tout au long de leur vie féconde, la fécondité par âge et rang de naissance observée cette année-là. La descendance ainsi estimée est répartie par nombre d'enfants : combien de femmes n'auraient pas d'enfant au cours de leur vie, en auraient un seul, deux, etc.

Exemple :

Faisons l'hypothèse que la probabilité pour une femme de 15 ans d'avoir un premier enfant dans l'année est de 0,002.

Alors pour 100 femmes de 15 ans, on aura 0,2 naissances.

Faisons-les vieillir d'un an. Et faisons l'hypothèse que la probabilité pour une femme de 16 ans d'avoir un premier enfant dans l'année est de 0,003. Et que, pour une femme de 16 ans avec un enfant, la probabilité d'en avoir un second dans l'année est de 0,05.

Alors on aura 99,8 femmes de 16 ans sans enfant qui auront $99,8 \times 0,003 = 0,29$ naissances et 0,2 femme de 16 ans avec un enfant qui aura $0,2 \times 0,05 = 0,01$ naissance, soit un total de 0,3 naissances.

Et ainsi de suite jusqu'à 50 ans.

Pour 100 femmes de 15 ans, et en appliquant à chaque âge les probabilités d'agrandissement, on obtiendra le nombre de femmes sans enfant à 50 ans, le nombre de femmes avec un seul enfant à 50 ans... Cela permet ainsi d'estimer la descendance finale de ces 100 femmes.

Pourquoi utilise-t-on l'ICFRA plutôt que l'ICF ?

La fécondité est plus tardive pour les femmes les plus aisées (Reynaud, 2020). A un âge donné, notamment aux âges jeunes, les femmes ayant un niveau de vie élevé ont en moyenne moins d'enfants déjà nés. Et ayant moins d'enfants déjà nés, elles mettent au monde plus d'enfants (la fécondité diminue avec le nombre d'enfants déjà nés, à âge égal).

Les différents sous-groupes de niveau de vie constitués vont ainsi avoir une structure différente en matière de nombre d'enfants déjà nés, ce qui a un impact sur la fécondité.

L'ICFRA découle des probabilités d'agrandissement déterminées par âge et par rang de naissance (équivalent au nombre d'enfants déjà nés). Contrairement à l'ICF, il permet de contourner cet effet de structure dans l'analyse des différences de fécondité selon le niveau de vie. Avec l'ICFRA, il en résulte une courbe de fécondité moins croissante selon le niveau de vie. L'ICFRA est ainsi plus adapté pour l'étude de la fécondité par niveau de vie.

Dans ce document de travail, comme précédemment avec l'Insee Première sur la fécondité selon le niveau de vie, nous avons vu que les quantiles étaient établis par âge. Un test a été réalisé, en déterminant les déciles non plus par âge mais globalement, quel que soit l'âge des femmes.

En définitive, les résultats obtenus s'avèrent assez proches. Les niveaux de vie moyens de chaque sous-groupe différent peu et les deux courbes relatives aux indices de fécondité présentent une allure comparable.

Nous poursuivons en privilégiant désormais l'ICFRA.

Figures 18 et 19 (disponibles ultérieurement)

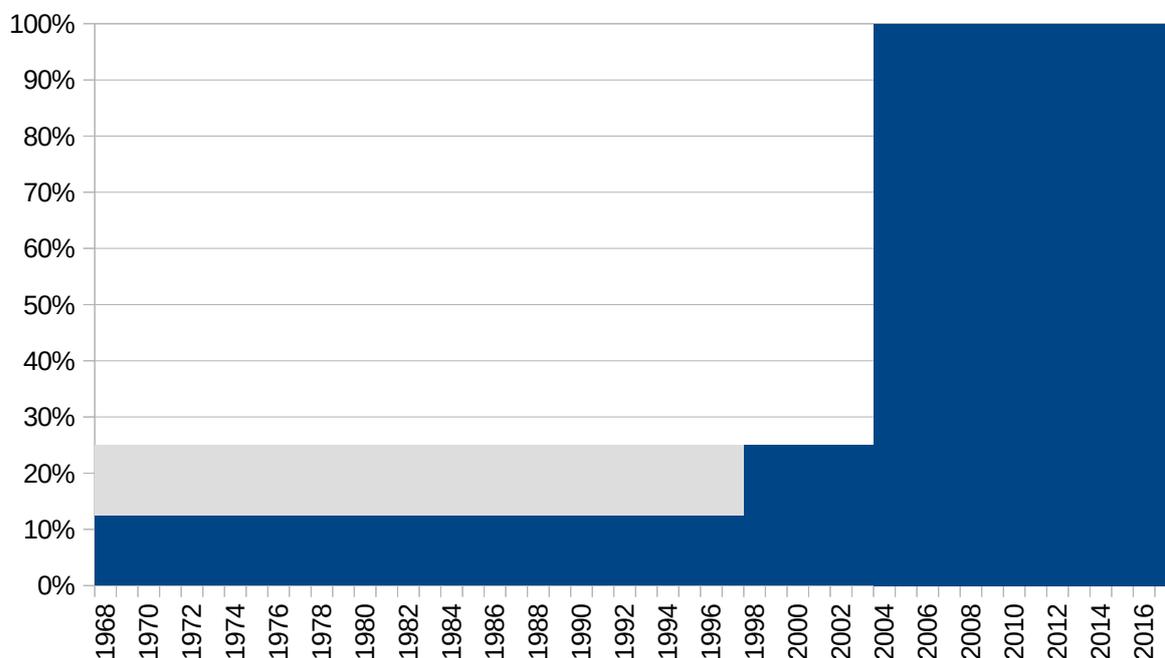
2.3 Contraintes de champ

Le choix de l'ICFRA comme indicateur requiert de disposer du nombre d'enfants déjà nés de chaque femme, en comptabilisant pour chaque femme, le nombre de naissances passées par rang. Or dans les données historiques de la table DESCENDANCE, un certain nombre de naissances manquent.

Pour les naissances antérieures à 1982, et bien que ces naissances figurent dans la table DESCENDANCE, le jour de naissance des parents est rarement renseigné.

Jusqu'en 1997, un nombre non négligeable de naissances de l'État Civil de parents EDP ne sont pas disponibles dans la table DESCENDANCE du fait de trous de collecte concernant l'EDP certaines années⁴. Cependant les naissances sur cette période sont exhaustives pour les femmes nées les 1er ou 4 octobre, soit la moitié de l'échantillon d'alors. Jusqu'en 2003, l'EDP comprenait les seuls individus nés les 4 premiers jours d'octobre. À partir de l'année 2004, la taille de l'EDP a été multiplié par 4 et le champ a été élargi aux femmes nées les 4 premiers jours de chaque trimestre (du 2 au 5 janvier pour le premier trimestre). Par ailleurs, les naissances dans les DOM (hors Mayotte) sont également disponibles à partir de l'année 2004.

Figure 20 - Part des naissances présentes dans l'EDP par construction, selon l'année de naissance



Note : table DESCENDANCE DE L'EDP. 100 % correspond au recueil pour les EDP nés les 4 premiers jours de chaque trimestre ; 25 % correspond à un recueil uniquement pour les EDP nés les 4 premiers jours d'octobre.

Champ : France métropolitaine

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Le nombre d'enfants déjà nés et le rang de naissance sont des variables fiables, seulement sur le champ restreint des femmes nées les 1 et 4 octobre. Par rapport au champ complet de l'EDP, cela revient à diviser la taille de l'échantillon par huit, limitant ainsi la précision et augmentant la variabilité des résultats.

Le choix de l'ICFRA, l'indicateur conjoncturel de fécondité tenant compte du rang, implique donc de travailler sur un champ restreint. Mais nous allons voir que nous pouvons optimiser ce champ.

⁴ pour des questions de moyens de gestion

2.4 Optimisation du champ de l'étude

L'objectif est de travailler sur le champ complet de l'EDP, à savoir l'ensemble des femmes nées les 4 premiers jours de chaque trimestre, lorsque c'est possible. Or nous allons voir que c'est possible, la variable rang de naissance restant fiable sur le champ complet jusqu'à une certaine génération de femmes. Pour les femmes au-delà d'une certaine génération, il est nécessaire de rester sur le champ restreint des femmes nées les 1 & 4 octobre, afin de pouvoir continuer à mobiliser la variable rang de naissance.

Nous allons comparer pour les femmes nées les 1 & 4 octobre d'un côté, et celles nées un autre jour EDP de l'autre, les répartitions selon le nombre d'enfants déjà nés, par âge.

Par exemple en 2017, pour les femmes de 50 ans, donc nées en 1967, les répartitions selon le nombre d'enfants déjà nés sont les suivantes :

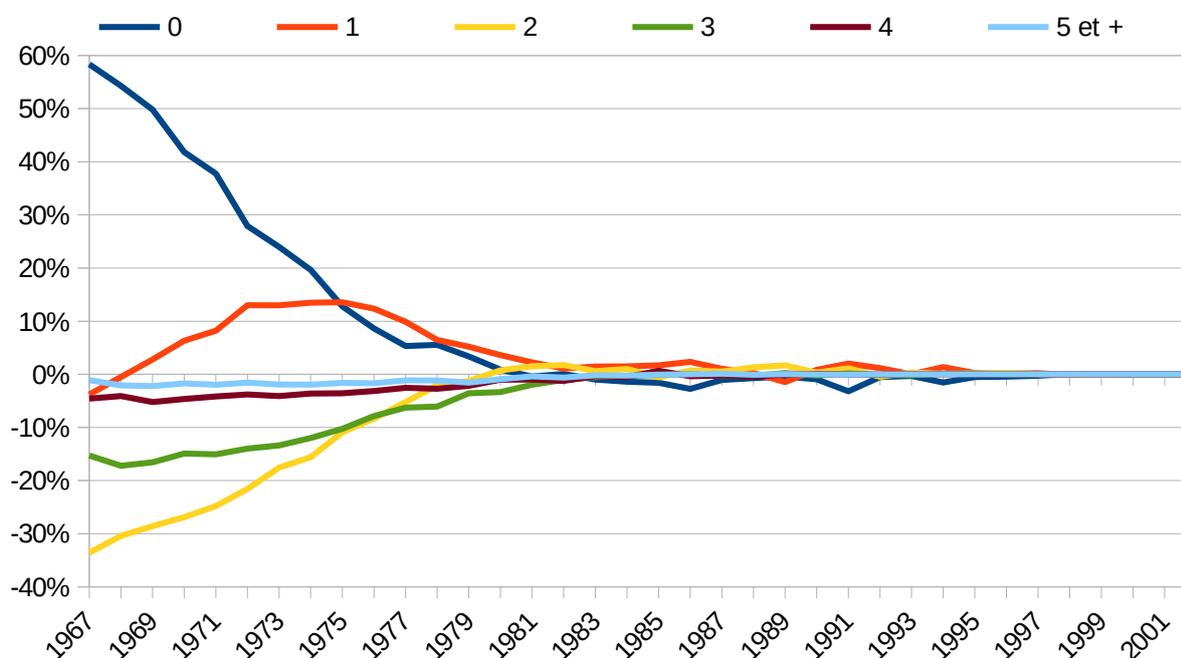
	Nombre d'enfants déjà nés					
	0	1	2	3	4	5 et +
Mère née les 1 & 4 oct	21 %	20 %	38 %	16 %	5 %	1 %
Mère née un autre jour EDP	79 %	16 %	4 %	1 %	0 %	0 %
<i>écart en points entre les deux</i>	58	-4	-34	-15	-5	-1

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

En reportant l'ensemble des écarts par nombre d'enfants déjà nés et par année de naissance, on obtient le graphique suivant :

Figure 21 - Écart en points de la répartition du nombre d'enfants déjà nés, entre les femmes nées un jour autre que le 1 ou 4 octobre et celles nées un 1 ou 4 octobre, selon l'année de naissance de la mère (données 2017)



Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : ensemble des femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

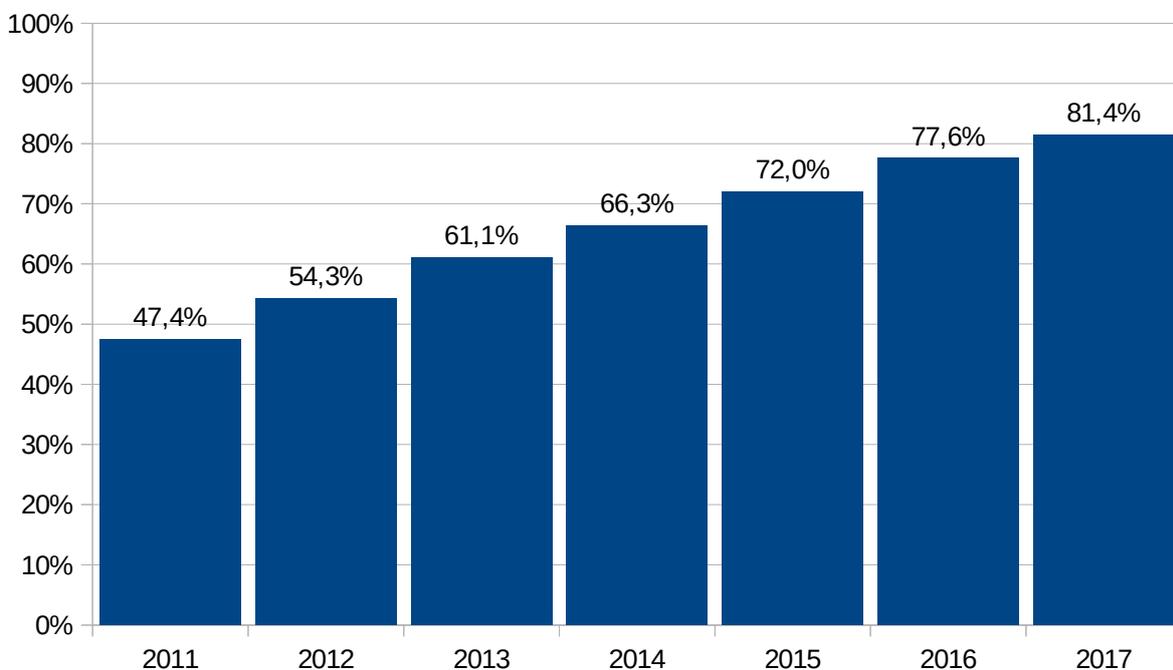
La répartition selon le nombre d'enfants déjà nés est par construction plus complète pour les femmes nées les 1 & 4 octobre (on recueillait bien dans l'EDP toutes les naissances de leurs enfants). Suite à la comparaison des répartitions pour chaque génération de femmes, selon qu'elles soient nées les 1 & 4 octobre ou non, un écart important est visible pour celles les plus âgées. Cet écart décroît et reste perceptible jusqu'à l'année de naissance 1981 environ. Au-delà, on peut en déduire que la répartition selon le nombre d'enfants déjà nés est correcte sur l'ensemble du champ et qu'elle n'est pas affectée par l'historique incomplet des naissances.

Dès lors, il est possible d'optimiser le champ :

- pour les femmes nées avant 1982, le champ est restreint aux seules femmes nées les 1 et 4 octobre.
- pour celles nées à partir de 1982, le champ sera complet (16 jours EDP).

Par rapport à l'Insee Première sur la fécondité selon le niveau de vie, dans lequel le champ avait été restreint aux seules femmes nées un 1^{er} ou un 4 octobre, cette opération permet d'augmenter sensiblement la taille de l'échantillon sur lequel établir les statistiques. Sur la période de 2012 à 2017, les taux de fécondité calculés sur le champ complet, donc sur un échantillon huit fois plus grand, contribuent à plus de la moitié de l'ICF, et ce de manière croissante au fil des années. En 2017, 81,4 % de l'ICF est calculé sur le champ complet, améliorant ainsi la qualité des estimations.

Figure 22 - Part de l'ICF calculé sur le champ complet (16 jours), correspondant aux années de naissance



Lecture : En 2017 81,4 % de l'ICF a été calculé sur le champ des 16 jours et 18,6 % sur le champ des 2 jours.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

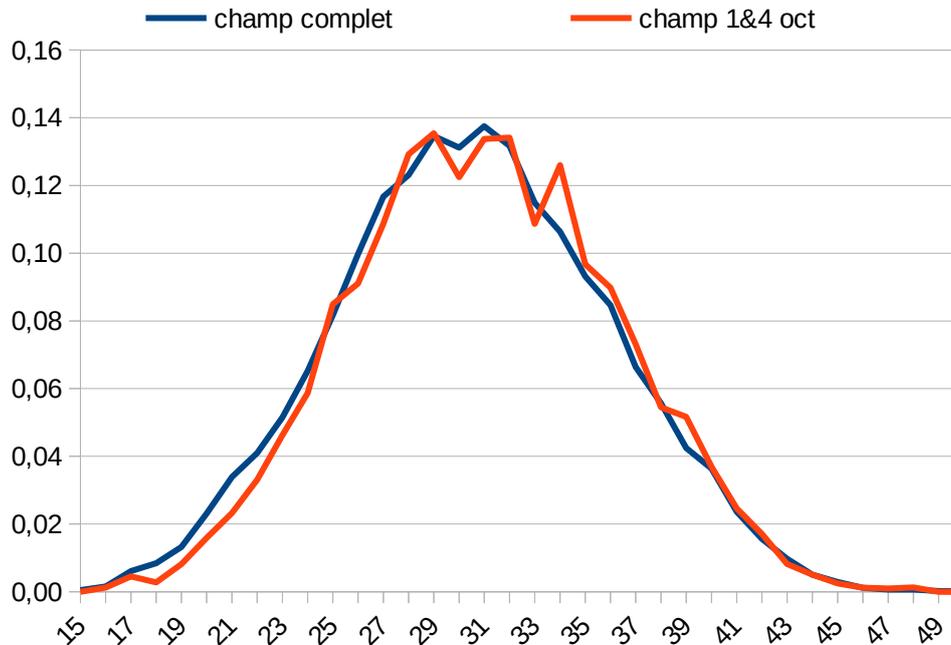
Champ : femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine, le champ étant restreint aux seules femmes nées les 1 et 4 octobre pour les générations nées avant 1982.

Compte tenu du fait que moins de 50 % de l'ICF est construit sur le champ complet en 2011, nous écarterons cette année dans la suite, notamment pour l'actualisation des résultats de l'Insee Première.

2.5 Analyse et recalage des taux de fécondité

La comparaison des taux de fécondité met en évidence un léger décalage temporel entre les courbes de l'ICF sur le champ complet de l'EDP et sur deux jours seulement. D'un côté, les femmes du champ complet sont nées le 19 mai en moyenne, soit le 139^e jour de l'année. De l'autre, les femmes nées les 1 et 4 octobre sont nées entre les 274 et 277^e jours de l'année. Les femmes nées les 1 et 4 octobre sont donc plus jeunes de 136,5 jours en moyenne, soit 0,375 année, par rapport à l'ensemble des femmes de l'EDP. Un décalage de 0,375 année existe donc entre les deux courbes.

Figure 23 - Comparaison des taux de fécondité par âge selon le champ (données 2017)



Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

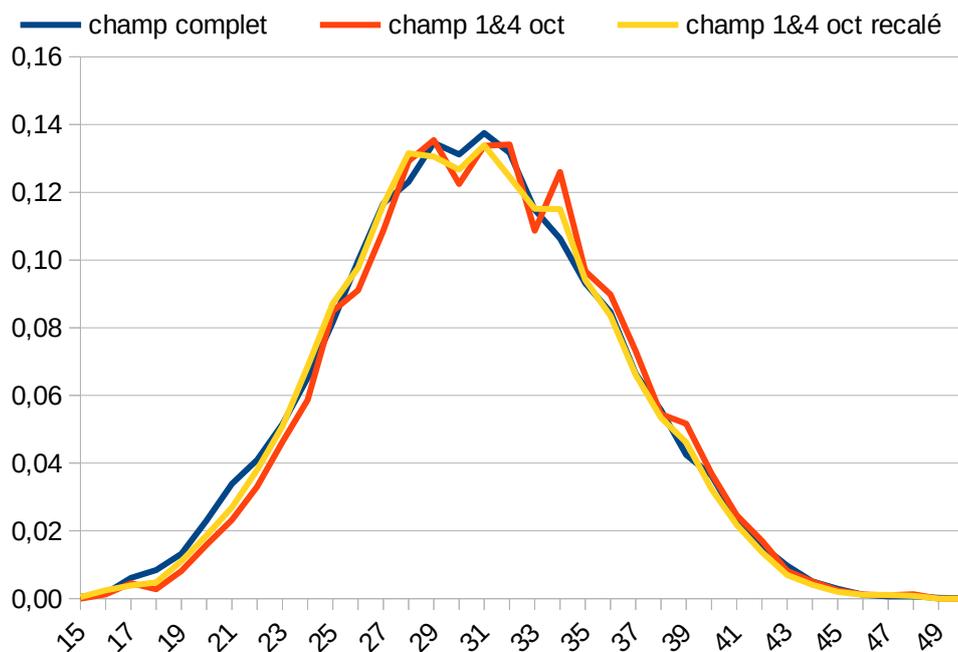
Champ : femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine, avec une distinction selon le jour de naissance

En rajeunissant de 0,375 en moyenne les femmes nées les 1&4 octobre, soit en décalant de 0,375 année vers la gauche la courbe des taux de fécondité des femmes nées les 1 et 4 octobre, les courbes vont coïncider. Pour ce faire, on va rajeunir 37,5 % des femmes de chaque génération d'une année. Sur le champ des femmes nées les 1 et 4 octobre, le taux de fécondité Tf d'un âge a donné, va ainsi être recalculé à partir de la formule suivante :

$$Tf'_a = 62,5 \% \times Tf_a + 37,5 \% \times Tf_{a-1}$$

À noter par ailleurs qu'il existe, sur le champ complet de l'EDP, un léger biais concernant l'âge, puisque les femmes sont nées en moyenne le 139^e jour de l'année, soit 43,5 jours en moins par rapport au milieu de l'année (365/2).

Figure 24 - Comparaison des taux de fécondité par âge selon le champ après recalage (2017)

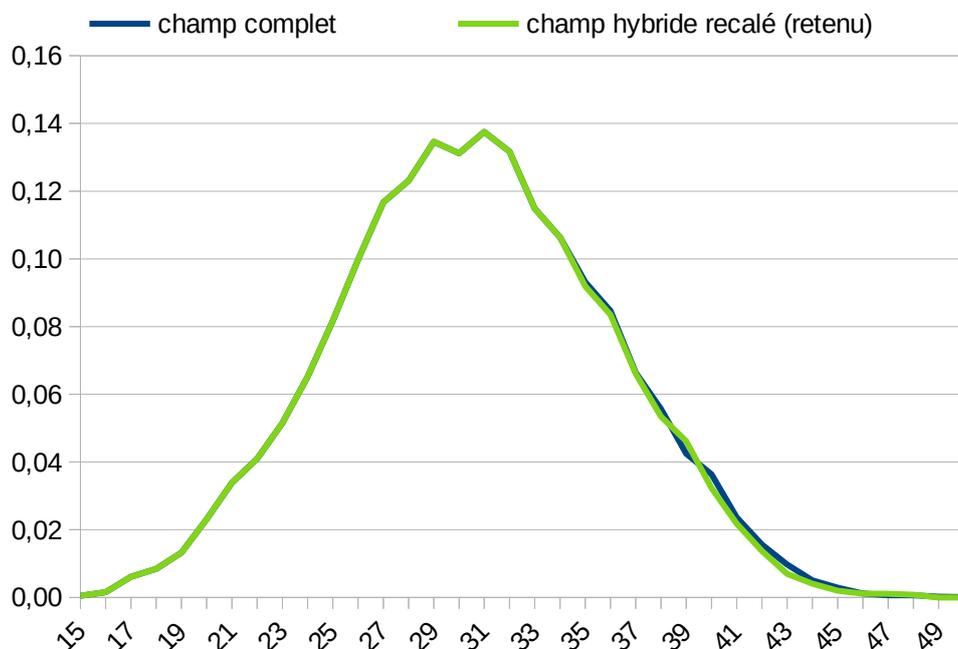


Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine, avec une distinction selon le jour de naissance

Le champ retenu est ainsi le champ hybride suivant : pour les femmes nées à partir de 1982 (femmes de 35 ans ou moins en 2017), il s'agit du champ complet. Pour les femmes nées avant 1982 (au-delà de 35 ans en 2017), le champ est restreint aux seules femmes nées les 1 et 4 octobre, avec un recalage temporel des taux de fécondité.

Figure 25 - Taux de fécondité par âge retenus (données 2017)



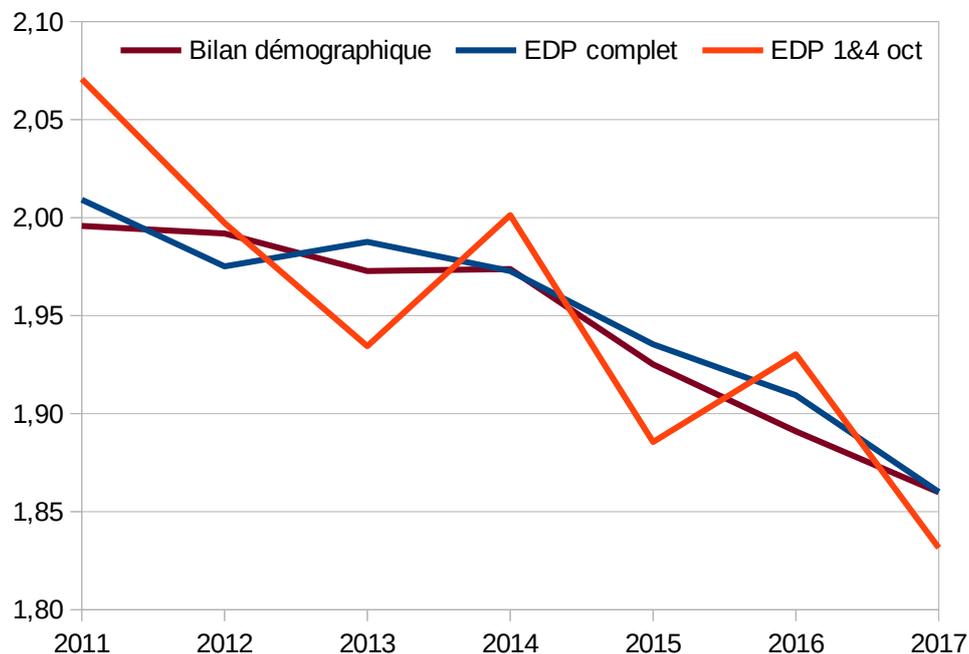
Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine, avec une distinction selon le jour de naissance

2.6 Résultats en matière d'ICF

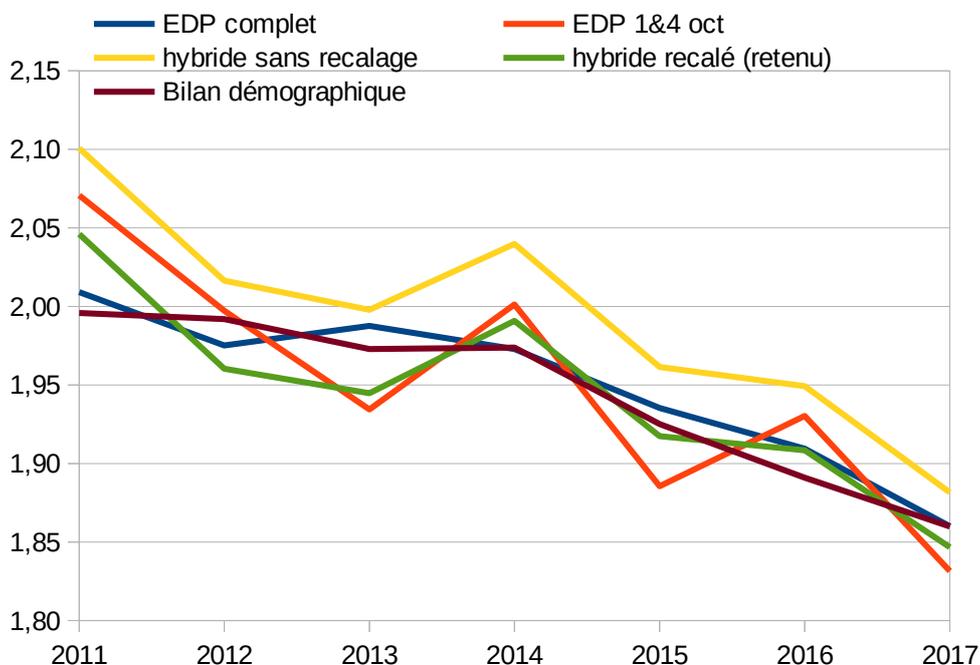
La taille du champ complet est égale à huit fois celui des femmes nées les 1 et 4 octobre. Il est ainsi normal que la variabilité de l'ICF calculé sur le champ réduit soit plus élevée.

Figure 26 - Comparaison des ICF de 2011 à 2017 selon le champ



Source : Insee Bilan démographique & Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent
 Champ : femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine, avec une distinction selon le jour de naissance

Figure 27 - Comparaison des ICF de 2011 à 2017 selon le champ (suite)



Source : Insee Bilan démographique & Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent
 Champ : femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine, avec une distinction selon le jour de naissance

Le champ retenu (champ complet jusqu'à un certain âge et champ réduit au-delà) permet de se placer pour un certain nombre de générations de femmes sur le champ complet de l'EDP. Pour ces générations, qui contribuent à plus de la moitié de l'ICFRA à partir de 2012, cela revient à multiplier l'échantillon par huit par rapport à celui de l'Insee Première sur la fécondité selon le niveau de vie. Celui-ci avait en effet été intégralement réalisé sur le champ des seules femmes nées les 1 & 4 octobre.

La figure 27 montre également que l'écart en termes de niveau de fécondité induit par le fait de se placer sur un champ hybride (champ complet jusqu'à un certain âge et champ réduit au-delà), est résorbé grâce au recalage temporel des taux de fécondité. Cela permet de rester à un niveau d'ICF proche de celui du bilan démographique.

Le champ retenu constitue un compromis, permettant à la fois :

- de disposer du nombre d'enfants déjà nés de chaque femme,
- de maintenir un niveau de fécondité proche de celui du bilan démographique
- tout en limitant la variabilité du fait de l'optimisation de la taille de l'échantillon.

Nous allons travailler sur ce champ pour l'actualisation les résultats de l'Insee Première et détailler un peu plus les résultats.

Figure 28 - Bilan en matière d'ICF

	ICF 2012-2017	écart au bilan démographique
bilan démographique	1,94	
champ complet (16 jours)	1,94	0,2%
champ des femmes nées les 1 & 4 oct (2 jours)	1,93	-0,3%
champ hybride (16 jours jusqu'à un certain âge puis 2 jours) sans recalage temporel	1,97	2,0%
champ hybride recalé (retenu)	1,93	-0,4%

Source : Insee Bilan démographique & Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine, avec une distinction selon le jour de naissance

Limites de l'étude

Pour finir, il subsiste une imperfection résiduelle : le nombre d'enfants déjà nés des femmes est issu de l'appariement imparfait avec la table DESCENDANCE et est, dans certains cas, égal à tort à zéro dans la base lorsque l'appariement ne se fait pas.

Or la réinjection des naissances non appariées s'appuie sur le nombre d'enfants déjà nés puisqu'elle porte sur les femmes sans enfant de l'échantillon. Et l'indicateur conjoncturel de fécondité privilégié dans cette étude, l'ICFRA, indicateur conjoncturel de fécondité tenant compte du rang et de l'âge, repose lui aussi sur le nombre d'enfants déjà nés, à travers le rang de naissance.

Aussi un test de robustesse a été effectué (cf. annexe 2). Il montre que cette imperfection est acceptable : sa correction, qui reposerait sur des hypothèses fortes, ne modifie que marginalement les résultats du document de travail.

Conclusion

La méthode proposée corrige au bout du compte la sous-estimation de la fécondité telle que constatée dans l'étude réalisée en novembre 2020 sur la fécondité par niveau de vie à partir de l'EDP (Insee Première n°1826). C'est la réinjection de naissances non appariées initialement, qui, après ajustement, permet de retrouver un niveau de fécondité comparable au niveau de référence du Bilan démographique. Cette méthode permet de disposer d'une base plus appropriée pour mener des études sur la fécondité selon différentes caractéristiques.

La sous-estimation de la fécondité provient d'un appariement incomplet des naissances de l'État Civil aux femmes présentes dans les données socio-fiscales de l'EDP, appariement nécessaire afin d'étudier la fécondité selon le niveau de vie.

Il s'explique en partie par un problème d'identification des individus lors de la construction des données socio-fiscales de l'EDP à partir de la source Fidéli, qui concerne plus spécifiquement les jeunes adultes et les personnes nées à l'étranger. Les coefficients de pondération (variable poids_fideli) élaborés de sorte à pallier ce problème d'identification, permettent de redresser les effectifs mais pas le déficit de fécondité constaté.

Plus généralement, le travail réalisé montre qu'un appariement incomplet n'est pas forcément un obstacle, et que des variables non identifiantes peuvent être mobilisées pour combler l'information manquante.

RÉSULTATS ANNEXES

A.1 Focus sur l'excès de naissances dans la table Descendance de l'EDP

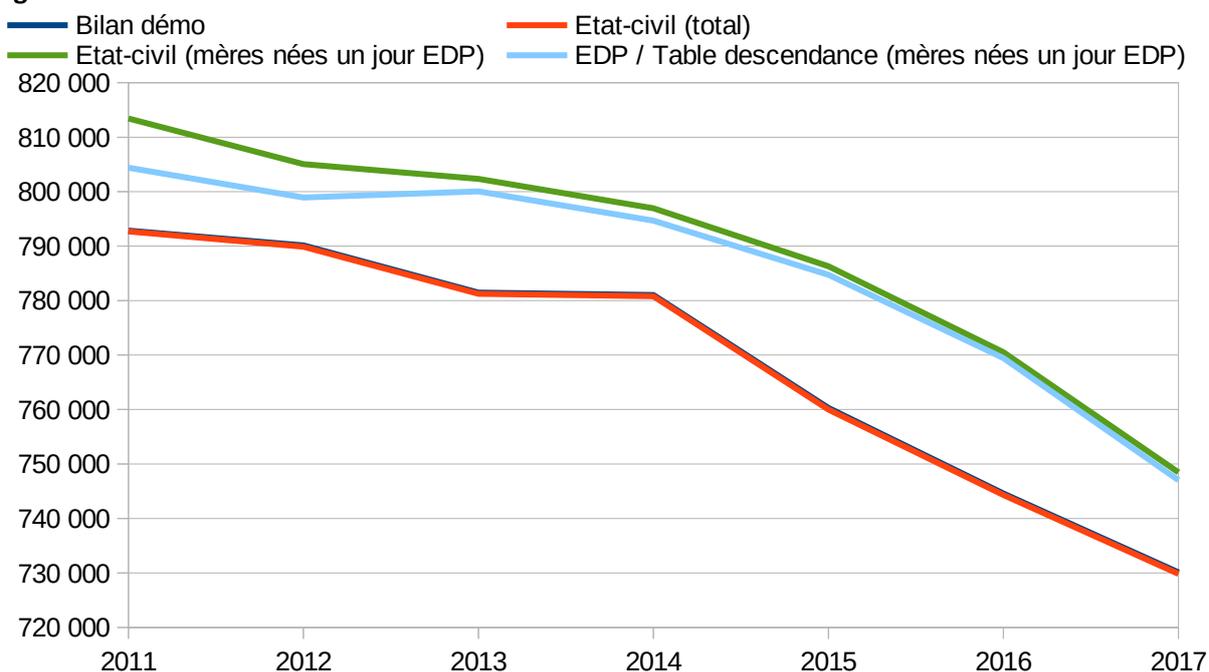
Nous comparons ici le nombre de naissances annuelles selon les différentes sources, sur le champ de la métropole.

Le nombre total de naissances coïncide par construction entre le bilan démographique et les données exhaustives de l'État Civil.

En revanche dans la table DESCENDANCE de l'EDP, le nombre de naissances relatives à des mères nées un jour EDP, soit 16 jours sur 365, semble surestimé. Il en est de même lorsqu'on sélectionne dans l'État Civil les naissances des seules mères nées l'un des 16 jours EDP.

Cet écart est relativement stable dans le temps.

Figure A1 - Nombre total de naissances : écarts selon les différentes sources

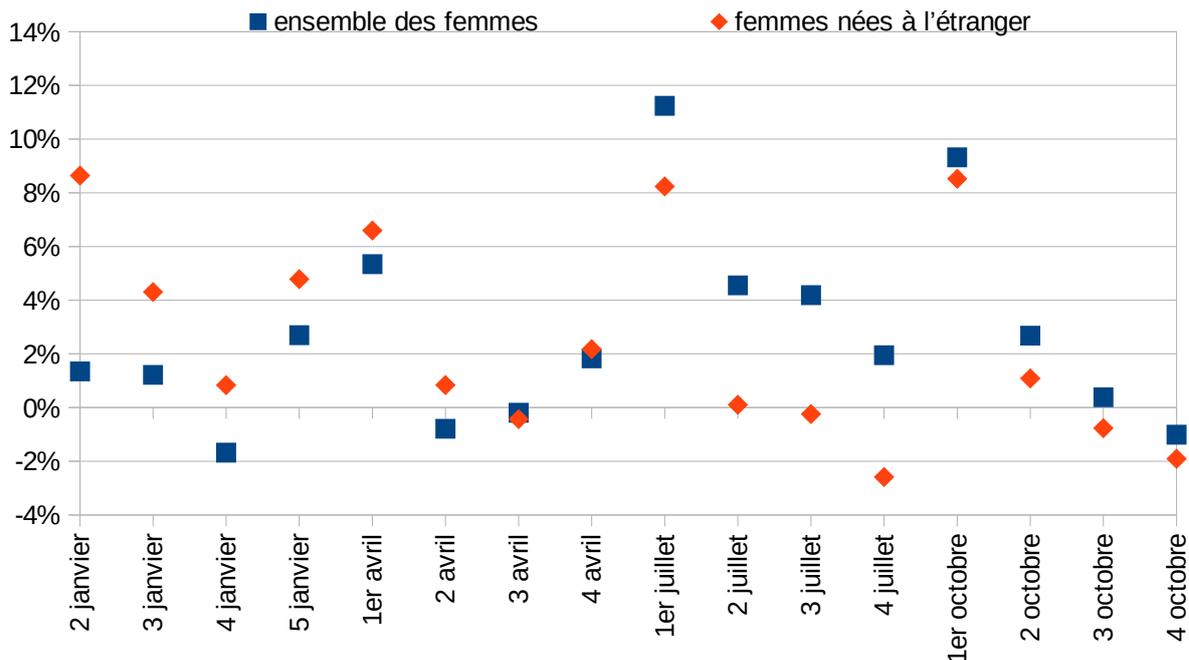


Source : Insee Bilan démographique et État Civil, Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent
Champ : France métropolitaine

Il s'avère que les naissances sont plus nombreuses pour les mères nées certains jours de l'année, et notamment le 1^{er} jour de chaque trimestre. Cela provient en particulier des mères nées à l'étranger pour lesquelles le jour de naissance figurant dans l'État Civil n'est pas toujours fiable⁵. Cela revient à majorer le nombre de femmes nées à l'étranger dans l'EDP, et le nombre de naissances correspondantes. L'incidence sur l'ICF est nulle si le numérateur (nombre de naissances) et le dénominateur (nombre de femmes) sont majorés dans la même proportion.

⁵ Lorsque la date de naissance exacte d'une personne n'est pas connue, mais uniquement le mois, l'administration de l'État-civil lui assigne une date de naissance correspondant au premier jour de ce mois. Si seule l'année de naissance est connue, il lui est souvent attribué la date anniversaire du 1^{er} janvier. C'est notamment pour cette raison que la date du 1^{er} janvier est exclue des « jours EDP ».

Figure A2 - Surplus de naissances selon le jour de naissance de la mère (jours EDP), par rapport au nombre de naissances journalier moyen, et contribution des femmes nées à l'étranger (en %, période de 2011 à 2017)



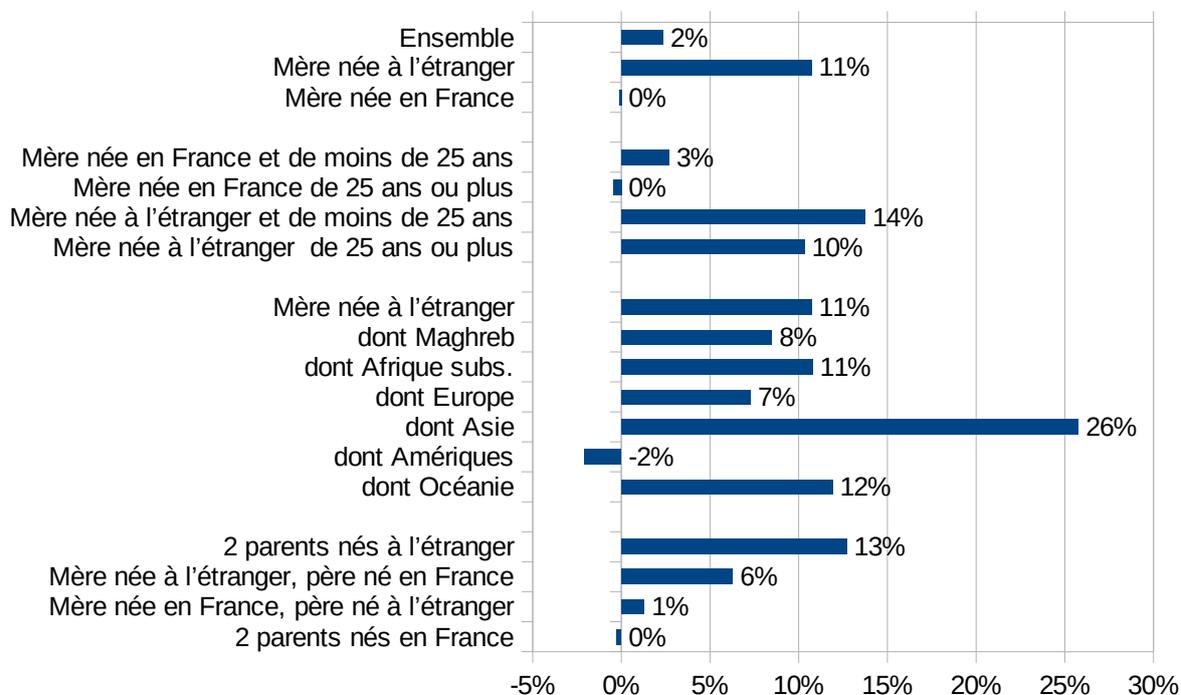
Lecture : pour les mères nées le 1^{er} juillet, on observe dans l'EDP 11 % de naissances en plus par rapport au nombre de naissances journalier moyen. Sur ces 11 %, la contribution des femmes nées à l'étranger est de 8 %.

Note : il s'agit des naissances figurant dans la table DESCENDANCE.

Source : Insee État Civil

Champ : France métropolitaine

Figure A3 - Surestimation des naissances dans l'EDP par rapport à l'État Civil (2017)



Note : il s'agit des naissances figurant dans la table DESCENDANCE.

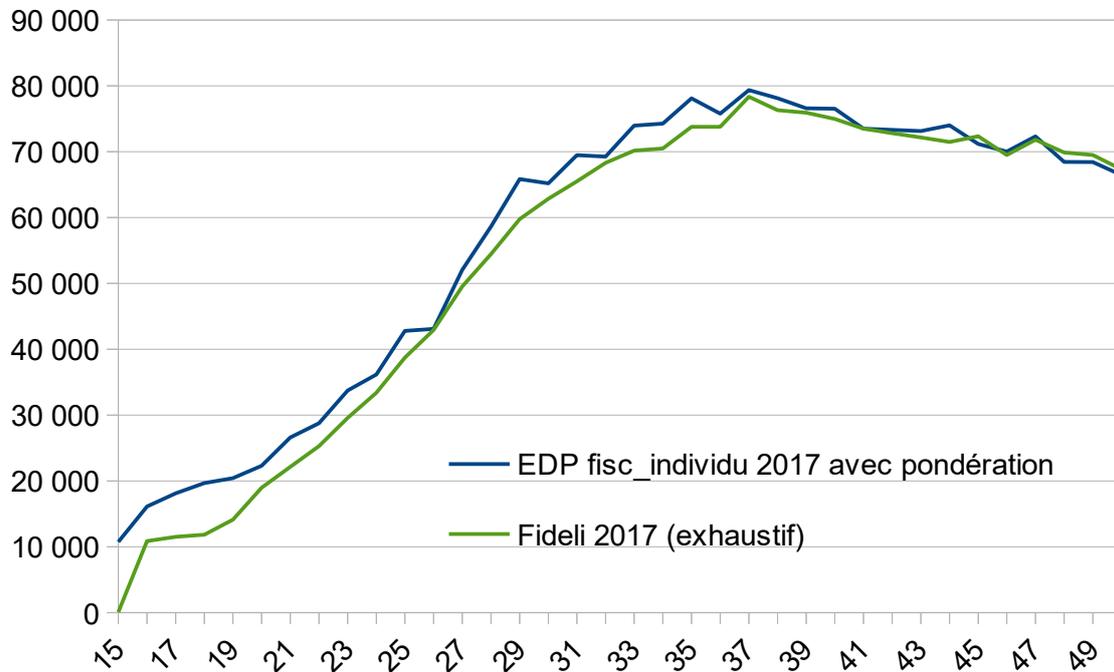
Source : Insee État Civil, Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : France métropolitaine

Au final, le nombre de femmes nées à l'étranger est majoré dans l'EDP, compte tenu du manque de fiabilité de leur jour de naissance (cf. figure A4). Il en est de même pour la proportion de femmes nées à l'étranger (cf. figure A5). Par ailleurs, la proportion de femmes nées à l'étranger n'est pas modifiée par la prise en compte des coefficients de pondération.

Or cette proportion a une incidence sur la fécondité globale, compte tenu de leur fécondité plus élevée que celle des femmes nées en France.

Figure A4 – Nombre de femmes nées à l'étranger selon l'âge

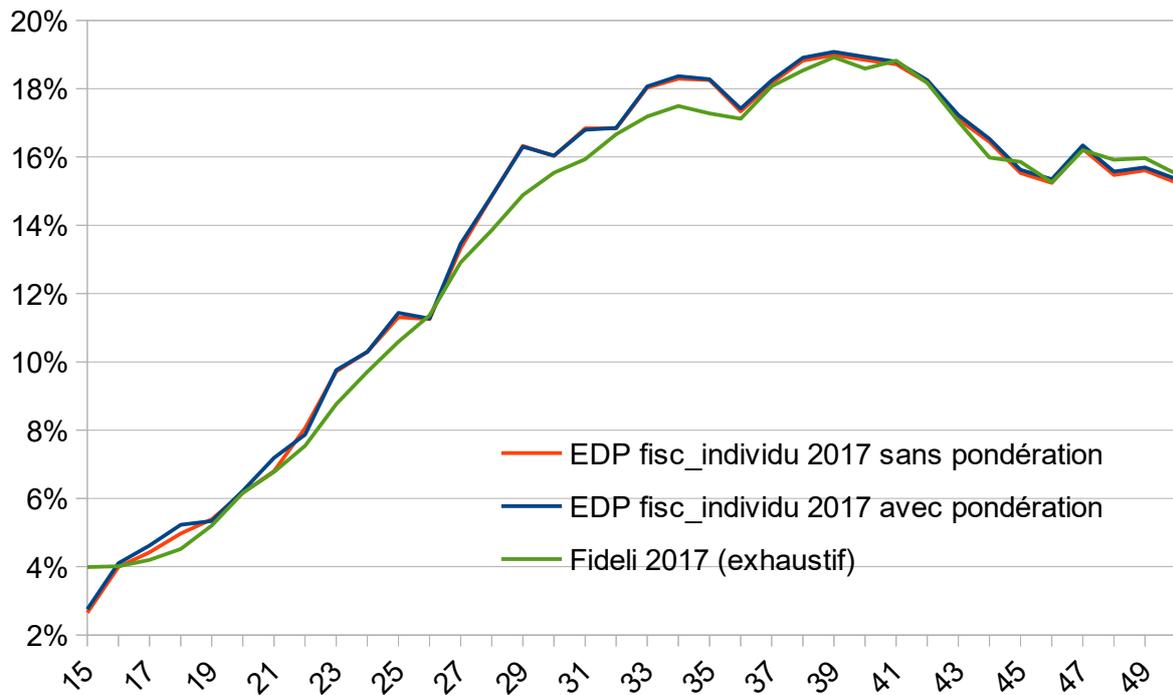


Source : Insee Fidéli, Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : France métropolitaine

Note : il s'agit de la pondération poids_fideli de la table fisc_individu

Figure A5 – Part des femmes nées à l'étranger selon l'âge



Source : Insee Fidéli, Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : France métropolitaine

Note : il s'agit de la pondération poids_fideli de la table fisc_individu

Figure A6 et A7 (disponibles ultérieurement)

A.2 Focus sur la détermination du « nombre d'enfants déjà nés »

Le nombre d'enfants déjà nés est une variable calculée d'une grande importance dans cette étude, puisqu'elle est corrélée au nombre d'enfants à venir.

Le nombre d'enfants déjà nés intervient ainsi dans les opérations d'imputation du niveau de vie et de réinjection des naissances. Équivalent au rang de naissance, il intervient également pour le calcul de l'ICFRA.

Or le nombre d'enfants (déjà nés) est calculé à partir de l'ensemble des naissances figurant dans la table DESCENDANCE et relatives à une même mère

Dans un certain nombre de cas, le nombre d'enfants déjà nés va ainsi être égal à 0 (ou minoré), non pas du fait qu'une mère n'a pas eu d'enfants mais parce que cette mère n'a pas été retrouvée dans la table DESCENDANCE bien qu'elle ait déjà eu des enfants. La valeur 0 comprend donc bien les personnes n'ayant pas encore d'enfant, mais aussi quelques femmes avec enfant pour lesquelles l'identification dans l'EDP et donc l'appariement avec la table DESCENDANCE n'a pas fonctionné.

Une analyse plus fine de cette variable a donc été effectuée, en comparant notamment le nombre d'enfants avec celui déterminé avec la variable « lien_familial » de la table FISC_INDIVIDU (issue cette fois-ci non plus de l'État-civil mais de Fidéli).

D'une part, le nombre total d'enfants issu de Fidéli n'est pas suffisant seul pour corriger complètement la sous-fécondité constatée précédemment.

D'autre part, le nombre d'enfants issu de chacune des sources coïncide globalement dans plus de 4 cas sur 5, et dans près de 70 % des cas pour les femmes nées à l'étranger (figure A6).

Figure A6 - Comparaison du nombre d'enfants issu de l'appariement avec l'État-civil (EC) et celui issu de Fidéli (nombre d'enfants au 1^{er} janvier 2017)

	ensemble	nées en France	nées à l'étranger
coïncide	81,4%	83,3%	69,2%
non nul dans Fidéli et nul d'après l'appariement EC	1,7%	0,5%	9,0%
non nul d'après l'appariement EC et nul dans Fidéli	8,2%	8,3%	7,3%
non nul dans les deux sources et écart de 1	6,3%	5,7%	9,6%
non nul dans les deux sources et écart supérieur à 2	2,5%	2,1%	4,9%

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, Échantillon démographique permanent

Champ : femmes de l'EDP âgées de 15 à 50 ans et domiciliées fiscalement en France métropolitaine

Lorsque le nombre d'enfants est non nul dans Fidéli et nul d'après l'appariement EC, on peut supposer que l'appariement avec l'État-civil ne fonctionne pas. Dans certains cas, il est également possible que cela provienne d'une divergence de champ, ou d'un décalage temporel entre les sources.

En tous cas, ce cas de figure est peu fréquent dans l'ensemble (1,7 %) mais représente 9 % des cas pour les femmes nées à l'étranger (cf. tableau ci-dessus).

À l'inverse, on constate que quel que soit le lieu de naissance, il arrive que le nombre d'enfants soit nul dans Fidéli et non nul d'après l'appariement avec l'état-civil (8,2 % en moyenne). Le nombre d'enfants dans Fidéli s'avère être alors non exhaustif.

Dans le cadre de l'étude, un test de robustesse des résultats a été effectué en corrigeant le nombre d'enfants issu de l'appariement avec l'État-civil, lorsque celui-ci était nul, à partir de celui issu de Fidéli, lorsque celui-ci était non nul.

Globalement, le niveau de fécondité varie peu selon qu'on corrige ou non la variable portant sur le nombre d'enfants, y compris pour les femmes nées à l'étranger. De même, les résultats en termes de descendance finale restent assez proches.

Cela montre au final une certaine robustesse des résultats : la variable calculée relative au nombre d'enfants déjà nés n'est certes pas parfaite du fait de défauts d'appariement, mais corriger cette variable avec des informations (elles aussi imparfaites) sur les naissances telles que apparaissant dans les données fiscales ne changent pas les grands résultats de l'étude.

A.3. Programmes SAS (disponibles ultérieurement)

Bibliographie

- [1] Reynaud D., « Les femmes les plus modestes et les plus aisées ont le plus d'enfants », Insee Première n°1826, 2020.
<https://www.insee.fr/fr/statistiques/4982628>
- [2] Volant S., Pison G., Héran F., « La France a la plus forte fécondité d'Europe. Est-ce dû aux immigrées ? », Population et Sociétés n° 568, Ined, 2019.
- [3] Blanpain N., « L'espérance de vie par niveau de vie », Insee Première n° 1687, 2018.
Daguet F., « En 2016, les femmes cadres ont un peu moins d'enfants que les employées », Insee Première n° 1769, 2016.
- [4] Blanpain N., Lincot L., « Avoir trois enfants ou plus à la maison », Insee Première n° 1531, 2015.
- [5] Costemalle V., « Parcours conjugaux et familiaux des hommes et des femmes selon les milieux sociaux et les générations », in Couples et Familles, coll. « Insee Références », édition 2015.
- [6] Masson L., « Avez-vous eu des enfants ? Si oui, combien ? », in France, portrait social, coll. « Insee Références », édition 2013.
- [7] Pison G., « Les maternités précoces en recul dans le monde », Population et Sociétés n° 490, Ined, 2012.
- [8] Eudeline J.-F., Garbinti B., Lamarche P., Roucher D., Tomasini M., « L'effet d'une naissance sur le niveau de vie du ménage », in Les revenus et le patrimoine des ménages, coll. « Insee Références », édition 2011.
- [9] Davie E., Mazuy M., « Fécondité et niveau d'études des femmes en France à partir des enquêtes annuelles de recensement », Population vol. 65 n° 3, 2010.
- [10] Héran F., Pison G., « Deux enfants par femme dans la France de 2006 : la faute aux immigrées ? », Population et Sociétés n° 432, Ined, 2007.
- [11] Toulemon L., « La fécondité des immigrées : nouvelles données, nouvelle approche », Population et Sociétés n° 400, Ined, 2004.
- [12] Rallu J.-L., Toulemon L., 1993, « Les mesures de la fécondité transversale. I — Constructions des différents indices », Population, 48, 1, 7-26 & « Les mesures de la fécondité transversale. II — Application à la France de 1946 à 1989 », Population, 48, 2, 369-404.
- [13] Levy M.-L., « Rang des enfants, taille des familles et probabilité d'agrandissement », Population et Sociétés n° 206, Ined, 1986.