

114

2002

DOCUMENTS DE TRAVAIL

Situations familiales et fécondité selon le milieu social

Résultats à partir de l'enquête EHF de 1999

Magali MAZUY





SITUATIONS FAMILIALES ET FECONDITE SELON LE MILIEU SOCIAL
RESULTATS A PARTIR DE L'ENQUETE EHF DE 1999

Rapport final rédigé pour le compte de la
Direction de la Population et des Migrations dans le
cadre de la convention DPM-INSEE-INED N° 20000074

Magali MAZUY

Juillet 2002

Institut national d'études démographiques

133, Boulevard DAVOUT

75 980 PARIS Cedex 20

Mél : mazuy@ined.fr

Tel : 01 56 06 22 26

Sommaire

Présentation générale.....	5
Première partie.....	7
Résultats de l'enquête en ménages.....	7
1 - Les situations familiales : parents, grands-parents, arrière-grands-parents	7
1.1 – Situations familiales et niveau d'études.....	8
1.2 – Situations familiales, niveau d'études et célibat définitif	11
1.3 – Situations familiales, niveau d'études et nombre d'enfants.....	15
1.4 – Situations familiales, niveau d'études et âge au premier enfant	21
1.5 – Régression logistique sur le risque d'être grand-parent.....	28
2 – Analyse de la fécondité : milieu social et sexe du locuteur	33
2.1 – Fécondité des hommes et des femmes selon leur propre P.C.S.	34
2.2 – Fécondité des hommes et des femmes selon la P.C.S. de leur conjoint.....	38
Seconde partie.....	41
Disponibilité et qualité des données.....	41
1 – Déclarations relatives aux enfants et protocole d'enquête.....	42
2 – Disponibilité et pondérations du fichier.....	50
2.1 – À quoi servent les pondérations ?	50
2.2 – Le plan de sondage	51
2.3 – Les non-réponses	52
2.4 – La post-stratification.....	52
2.5 – Les pondérations finales	53
2.6 – Quelle pondération utiliser ?.....	54
2.7 – Les pondérations des enfants	55
Bibliographie	57
Annexe : documentation du fichier.....	59

Présentation générale

Ce rapport présente des résultats issus de l'enquête Etude de l'histoire familiale (EHF) réalisée en 1999, appelée encore « Enquête Famille ». Il s'inscrit dans le cadre de la convention DPM-INSEE-INED 20000074, relative aux premières exploitations faites à partir du fichier provisoire. Il fait suite à deux précédents rapports : un rapport méthodologique présentant l'enquête et énumérant les différents aspects de production et de validation des données dans le cadre de la convention. Le deuxième rapport présentait différents travaux de validation des données, ainsi que les premiers résultats de l'enquête, portant sur des variables démographiques : situation familiale au moment de l'enquête, taille des familles et niveau de la fécondité générale. Il a donné lieu à un numéro des *Dossiers et recherches de l'Ined* (n° 93, février 2001) intitulé « Etude de l'histoire familiale, premiers résultats de l'enquête en ménages ».

Les résultats présentés ici portent également sur la population vivant en ménages ordinaires. Ils complètent certains résultats mis en évidence lors du précédent travail par un enrichissement avec des variables de position sociale. Les thèmes abordés sont à nouveau les situations familiales (parentalité, grand-parentalité et arrière-grand-parentalité), l'âge auquel on perd ses parents et la fécondité. Il s'agit cette fois de la fécondité biologique, sans intégrer les beaux-enfants de la personne. On s'intéresse surtout aux différentiels sociaux de fécondité et à la fécondité différentielle selon le sexe.

Enfin, on présentera des travaux parallèles portant sur la qualité des données recueillies dans l'enquête. Cette source étant récente, de multiples travaux d'enrichissement ont été effectués. Le fichier a maintenant une pondération définitive. D'autre part, nous avons effectué un travail de comparaison des déclarations relatives aux enfants, que nous présenterons en seconde partie.

Première partie

Résultats de l'enquête en ménages

1 - Les situations familiales : parents, grands-parents, arrière-grands-parents

Les situations familiales sont largement dépendantes de l'âge. Nous avons vu dans le précédent travail que, parmi les personnes âgées de 50 ans, on compte environ 20% de grands-parents et que, parmi celles âgées de 65 ans, la proportion atteint 75%. Elle plafonne à 80% parmi les personnes les plus âgées (certaines personnes restent sans aucune descendance ou bien ont des enfants mais aucun petit-enfant). Parmi les personnes survivant jusqu'à leur 65^{ème} anniversaire, la majorité sont grands-parents. Les arrière-grands-parents sont quant à eux beaucoup moins nombreux en raison du caractère tardif de cet événement ; il s'agit de ce fait d'une population largement féminine.

Outre les âges charnières qui dessinent les différents contours du cycle de vie d'un individu, de multiples facteurs vont apporter des nuances dans la probabilité d'accéder plus ou moins vite, ou pas du tout, à la grand-parentalité et arrière-grand-parentalité. Le sexe, la catégorie sociale, l'âge de fin d'études, le nombre d'enfants, l'âge au premier enfant, sont autant de variables interdépendantes qui modifient les caractéristiques de la parentalité et de la grand-parentalité. Ceci revient à analyser la grand-parentalité comme un événement personnel, même s'il s'agit d'un événement qui ne dépend pas des enquêtés eux-mêmes, mais de la parentalité de leurs propres enfants.

1.1 – Situations familiales et niveau d'études

Le niveau d'instruction a été choisi comme indicateur de position sociale. Il a l'avantage d'être également un indicateur du moment à l'entrée dans la vie active. Nous avons choisi un découpage en quatre catégories. Il oppose une situation générale à des situations plus particulières (tableau 1.1) : sont placées dans un premier groupe les personnes qui n'ont eu aucun diplôme (correspondant à celles qui ont arrêté l'école très jeunes) ; dans un deuxième groupe les personnes qui ont un diplôme inférieur au baccalauréat (certificat d'études ou diplômes professionnels), qui regroupe la moitié de l'échantillon. Nous avons ensuite séparé les diplômés et les très diplômés en deux groupes distincts, prenant comme critère le bac (professionnel ou général) pour le troisième groupe ; le dernier groupe est formé des diplômés du supérieur. Le fait d'avoir le bac ou d'avoir fait des études supérieures implique des âges à l'entrée sur le marché du travail ainsi qu'une autonomie financière et des perspectives de carrière très différents du reste de la population. De même, les personnes qui ont fait des études supérieures étaient dans ces générations assez peu nombreuses (très minoritaires pour les femmes) et appartenaient aux classes les plus aisées de la population. Nous comparons donc le groupe majoritaire aux non diplômés et aux diplômés.

Tableau 1.1 :

Répartition par diplôme des personnes âgées de 50 ans et plus

Diplôme		50-54 ans	55-59 ans	60-64 ans	65-69 ans	70-74 ans	75 ans et +
HOMMES	NR	6,1	6,5	8,0	8,0	7,8	9,0
	sans diplôme	11,7	14,5	18,7	25,1	24,3	24,4
	inf. bac	53,0	51,2	52,7	50,3	51,9	51,4
	bac	10,3	10,7	8,9	7,3	7,1	6,7
	sup. bac	18,9	17,1	11,7	9,3	8,8	8,5
	ensemble	100,0	100,0	100,0	100,0	100,0	100,0
Diplôme		50-54 ans	55-59 ans	60-64 ans	65-69 ans	70-74 ans	75 ans et +
FEMMES	NR	6,4	8,1	9,4	9,8	9,4	9,9
	sans diplôme	13,0	16,1	20,6	25,6	23,7	31,3
	inf. bac	53,6	54,1	55,5	54,4	57,6	51,8
	bac	11,0	9,4	7,3	5,5	5,2	4,1
	sup. bac	16,0	12,3	7,2	4,7	4,1	3,0
	ensemble	100,0	100,0	100,0	100,0	100,0	100,0

Tous les résultats concernent la population répondante, postulant qu'elle est représentative de la population non-répondante en terme de situation familiale.

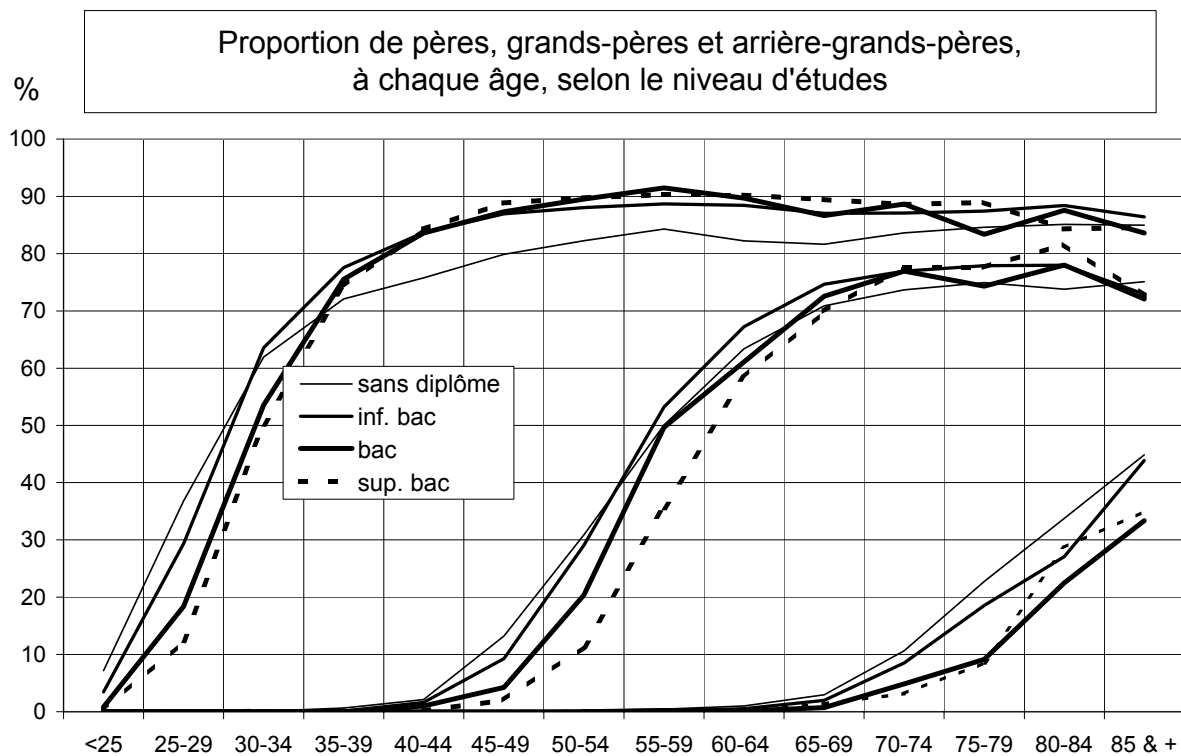
L'indicateur de niveau social fait apparaître des contrastes différents et parfois même opposés selon le sexe, en termes de situation parentale. Ainsi les hommes non diplômés sont moins souvent pères, un peu moins souvent grands-pères et plus souvent arrière-grands-pères. Les femmes non diplômées sont quant à elles plus souvent mères et grands-mères et beaucoup plus souvent arrière-grands-mères. Plus les femmes sont diplômées plus la grand-parentalité est faible, contrairement aux hommes (graphiques 1.1 et 1.2).

La prise en compte du diplôme fait apparaître beaucoup plus d'hétérogénéité chez les femmes que chez les hommes. Ainsi, parmi les générations anciennes, la fréquence de la grand-parentalité diffère de quelques points lorsque l'on est un homme peu ou très diplômé, alors que les femmes âgées de 70 ans sont grands-mères pour deux tiers d'entre elles lorsqu'elles sont très diplômées et huit fois sur dix quand elles n'ont eu aucun diplôme.

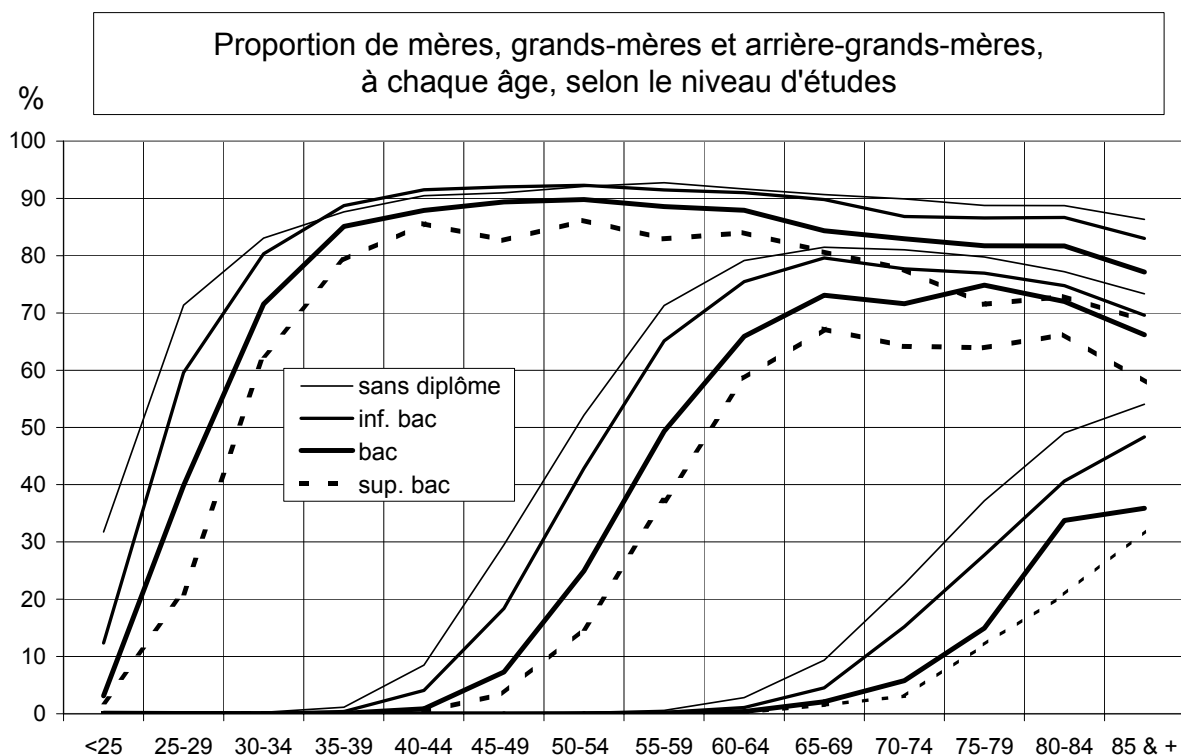
Enfin, le calendrier des personnes moins diplômées est plus précoce que celui des plus diplômées, aussi bien pour les hommes que pour les femmes avec une fois encore une hétérogénéité plus importante pour les femmes. Les différences sont particulièrement significatives pour les arrière-grands-mères puisque parmi les femmes les plus âgées, un tiers des femmes les plus diplômées ont au moins un arrière-petit-enfant alors que c'est le cas pour plus de la moitié des femmes qui n'ont aucun diplôme.

La fréquence du célibat définitif et des personnes restées sans enfant, le nombre d'enfants et l'âge au premier enfant expliquent ces différences.

Graphique 1.1



Graphique 1.2



1.2 – Situations familiales, niveau d'études et célibat définitif

La fréquence du célibat et des individus restés sans enfant a une influence directe sur la grand-parentalité. Nous avons mis en évidence dans le précédent rapport que les taux de célibat masculin et féminin par génération étaient sensiblement différents (graphique 2.6 p.22, *Dossiers et Recherches N°93*). Nous voyons ici aussi que la proportion de personnes n'ayant jamais vécu en couple, qui ne dépasse pas 6%, baisse avec l'âge pour les hommes alors qu'elle augmente légèrement pour les femmes (graphiques 1.3 et 1.4).

Nous employons le terme de célibat par extension, mais il s'agit des personnes n'ayant jamais vécu en union, ce qui est moins restrictif que le célibat définitif. Nous traitons des personnes âgées de plus de 65 ans, âge auquel la majorité des personnes sont grands-parents, car nous nous intéressons dans un premier temps à l'intensité de la grand-parentalité.

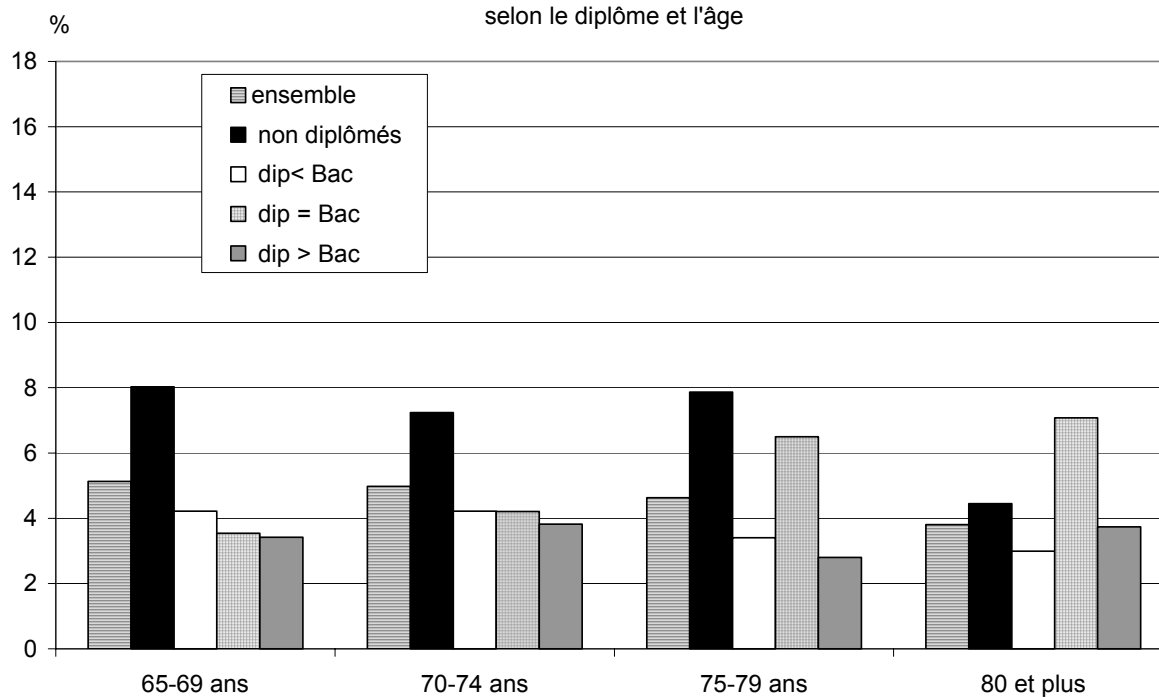
Le fait d'inclure le diplôme introduit une variabilité dans la fréquence observée du célibat. Variabilité différente selon le sexe et plus faible pour les hommes. Les hommes qui n'ont jamais vécu en couple sont plus nombreux parmi les moins diplômés pour les générations âgées de 65 à 80 ans au moment de l'enquête, alors que les femmes qui n'ont jamais vécu en union sont beaucoup plus nombreuses parmi les plus diplômées. Ceci va dans le même sens que des résultats issus d'une autre source présentant des taux de célibat définitif (Toulemon, Lapierre-Adamcyk, 2000). Néanmoins nous trouvons une proportion moins élevée d'hommes non diplômés n'ayant jamais vécu en couple, la différence s'expliquant peut-être par le fait que ces hommes ont eu une histoire conjugale ne s'étant jamais concrétisée par un mariage. La fréquence des hommes n'ayant jamais vécu en couple est très faible parmi les hommes diplômés. Si les proportions les plus élevées ne dépassent pas les 8% pour les hommes non diplômés, elles atteignent 16% pour les femmes les plus diplômées.

Une fois neutralisé cet effet des personnes n'ayant jamais vécu en couple, on remarque pour les plus de 65 ans que l'effet du diplôme sur la grand-parentalité ne joue plus pour les hommes ; pour les femmes, il garde une influence mais beaucoup plus limitée qu'auparavant (graphiques 1.5a et 1.5b).

Il reste l'effet des personnes qui ont vécu en couple mais restées sans enfant, stables pour les hommes, mais variables selon la génération et le diplôme pour les femmes (cf. graphiques 1.6a et 1.6b).

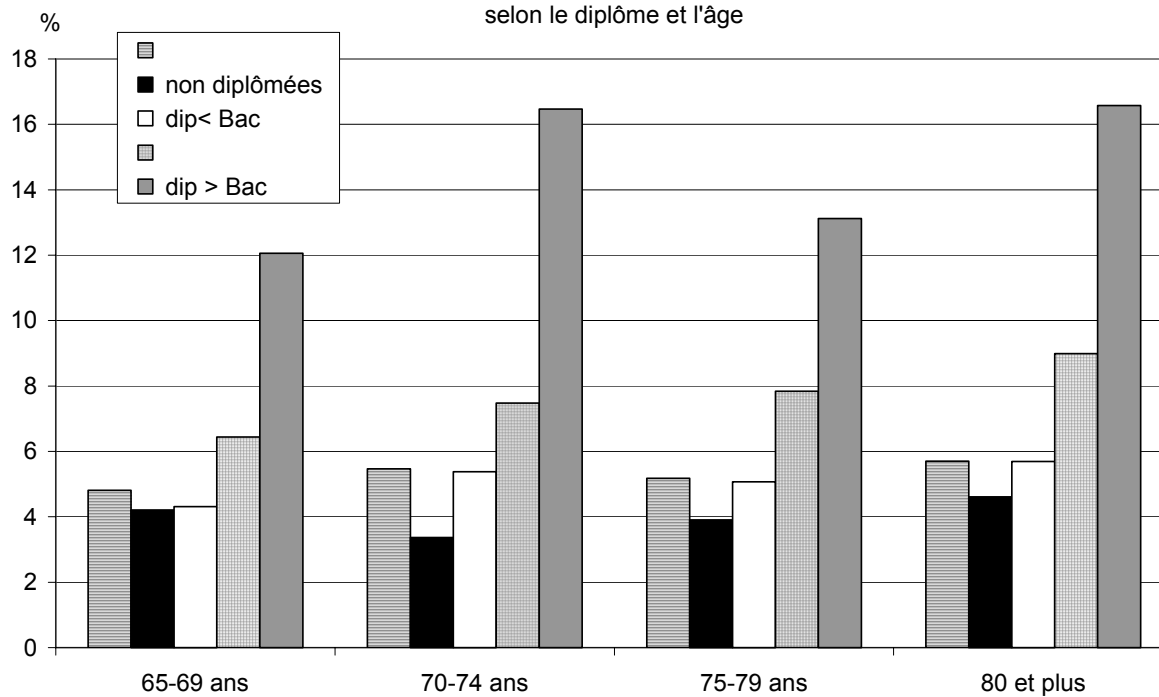
Graphique 1.3

Proportion d'hommes n'ayant jamais vécu en couple
selon le diplôme et l'âge

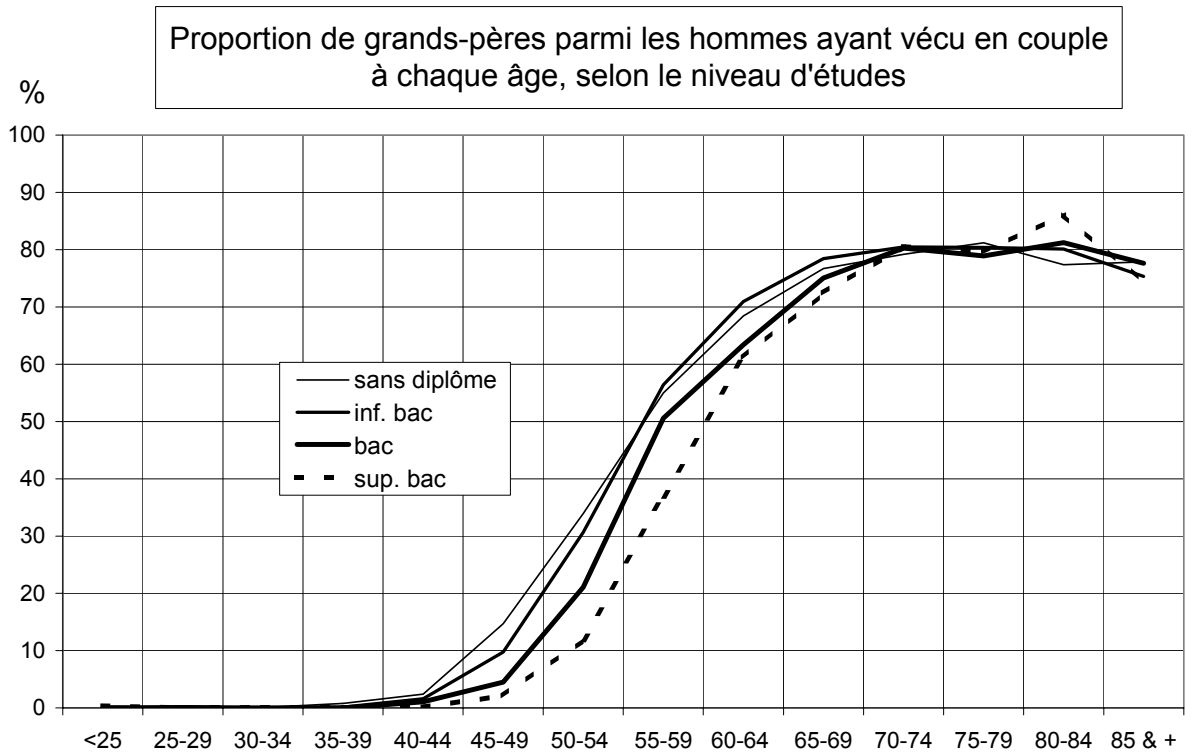


Graphique 1.4

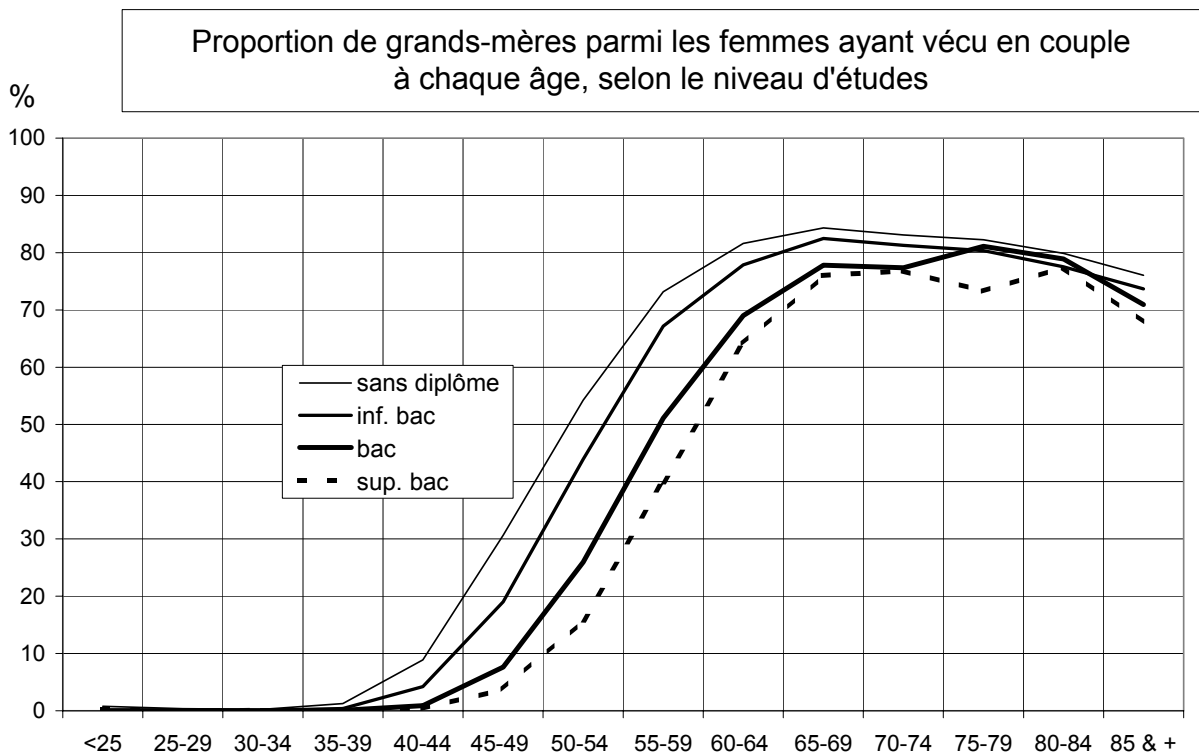
Proportion de femmes n'ayant jamais vécu en couple
selon le diplôme et l'âge



Graphique 1.5a

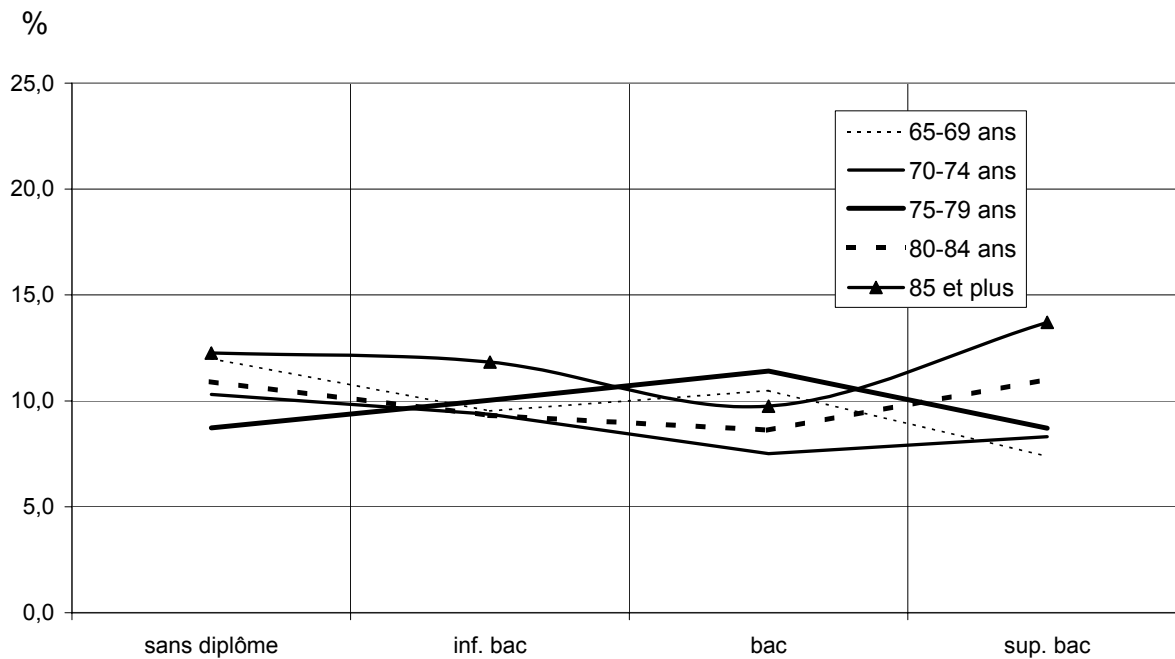


Graphique 1.5b



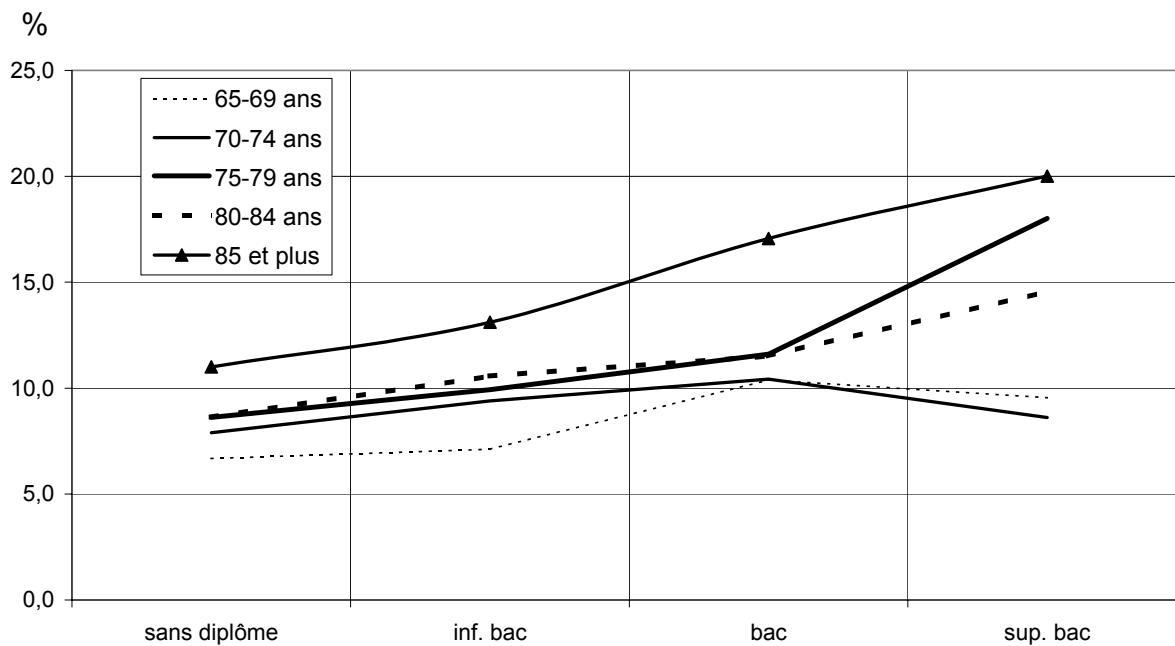
Graphique 1.6a

Proportion d'hommes ayant vécu en couple restés sans enfants



Graphique 1.6b

Proportion de femmes ayant vécu en couple restées sans enfants



Nous avons vu lors de nos premières analyses que les personnes qui sont le moins souvent grands-parents sont bien entendu des personnes restées sans enfant ou encore des personnes qui ont eu un seul enfant. Plus le nombre d'enfants est élevé, plus les chances d'être grand-parent et arrière-grand-parent augmentent. Le fait de n'avoir qu'un enfant réduit considérablement la probabilité d'être grand-parent, quel que soit le niveau de diplôme. Nous allons donc poursuivre les analyses sur la population des parents.

1.3 – Situations familiales, niveau d'études et nombre d'enfants

Les graphiques 1.7a à 1.7d et 1.8a à 1.8d présentent respectivement les proportions de grands-pères et de grands-mères selon le nombre d'enfants, pour chaque niveau d'études.

Pour les hommes comme pour les femmes, le fait de n'avoir eu qu'un seul enfant est le critère le moins favorable, surtout lorsqu'ils sont diplômés.

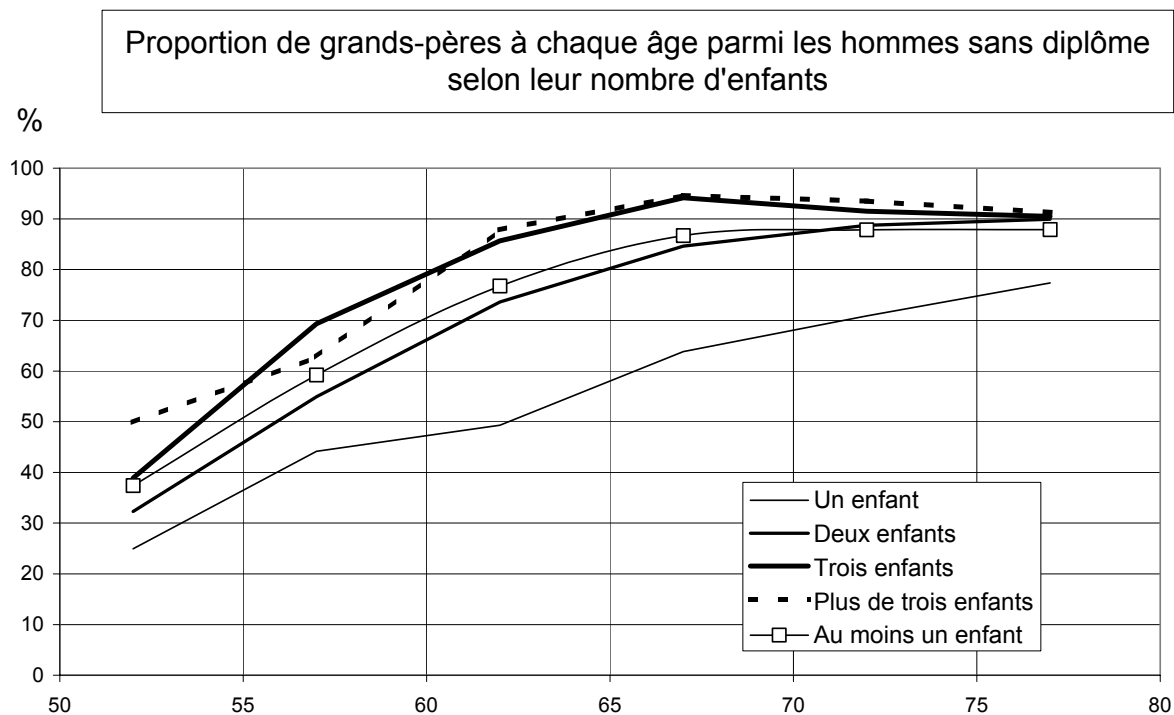
Parmi les générations plus âgées, les différences dues au diplôme sont peu importantes. C'est surtout aux âges jeunes que les différences sont importantes, avec des calendriers beaucoup plus précoces pour les femmes peu ou non diplômées. L'influence du nombre d'enfants agit également en terme de calendrier.

Les femmes les moins diplômées cumulent forte fécondité et précocité du calendrier (graphique 1.8d). Ainsi on trouve beaucoup moins de grands-mères précoces (à 50-54 ans) parmi les plus diplômées, même lorsqu'elles ont eu une famille nombreuse.

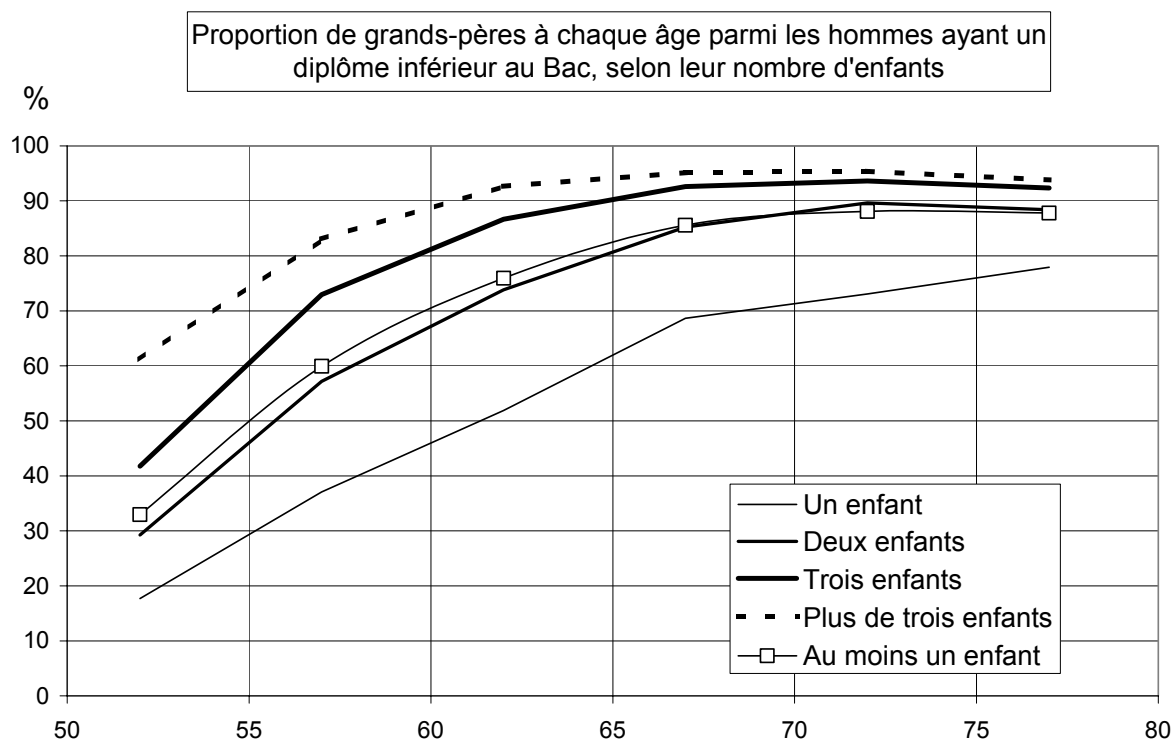
En ce qui concerne les arrière-grands-parents, les différences sont beaucoup plus marquées pour les femmes que pour les hommes. Le diplôme a une influence en raison du caractère tardif de l'événement. Les femmes sans diplôme ayant une famille nombreuse sont arrière-grands-mères plus d'une fois sur deux (et les hommes une fois sur trois), alors que les personnes les plus diplômées ayant eu un seul enfant sont arrière-grands-parents une fois sur dix (graphiques 1.9 et 1.10).

Ces différences sont sans doute dues également au fait que les personnes qui ont le plus d'enfants ont eu leur premier enfant à des âges jeunes, c'est pourquoi nous allons tester l'effet de l'âge à la primo-paternité et maternité.

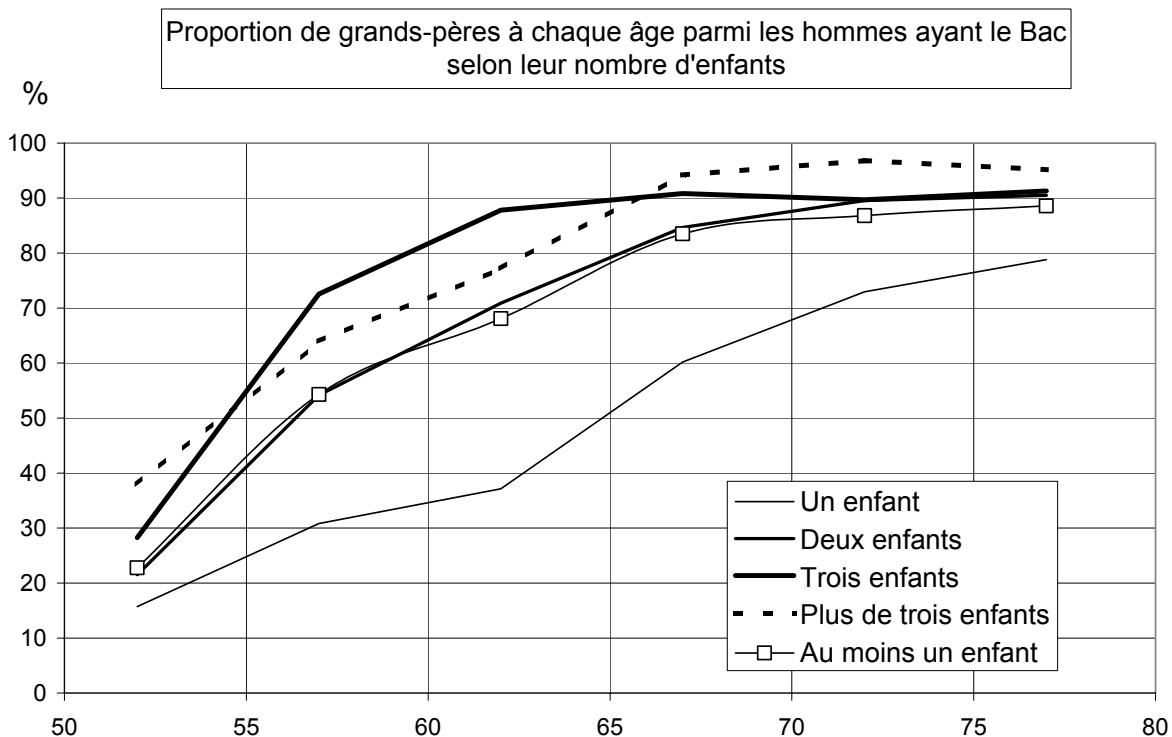
Graphique 1.7a



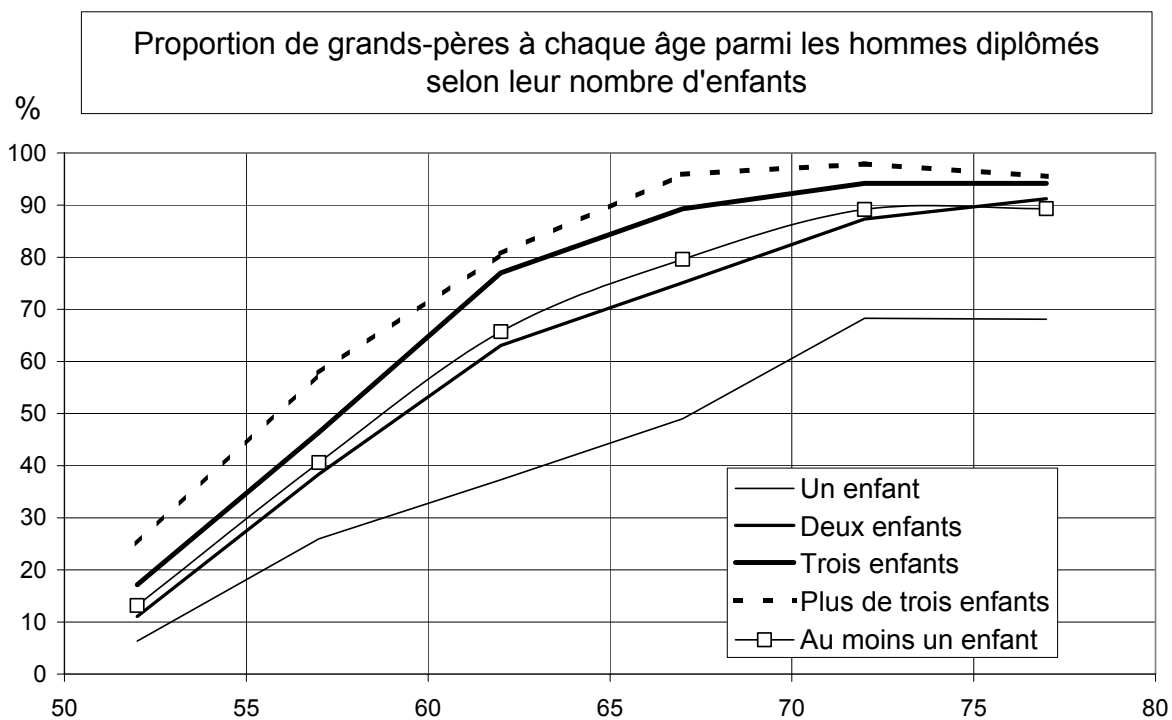
Graphique 1.7b



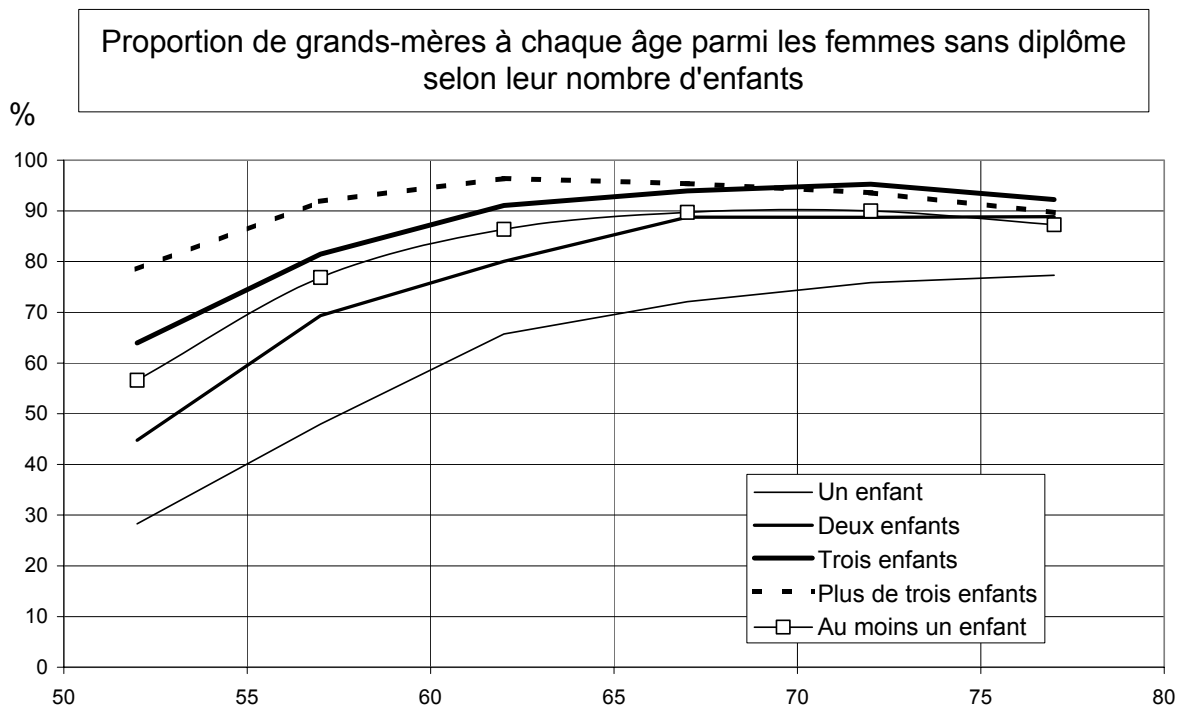
Graphique 1.7c



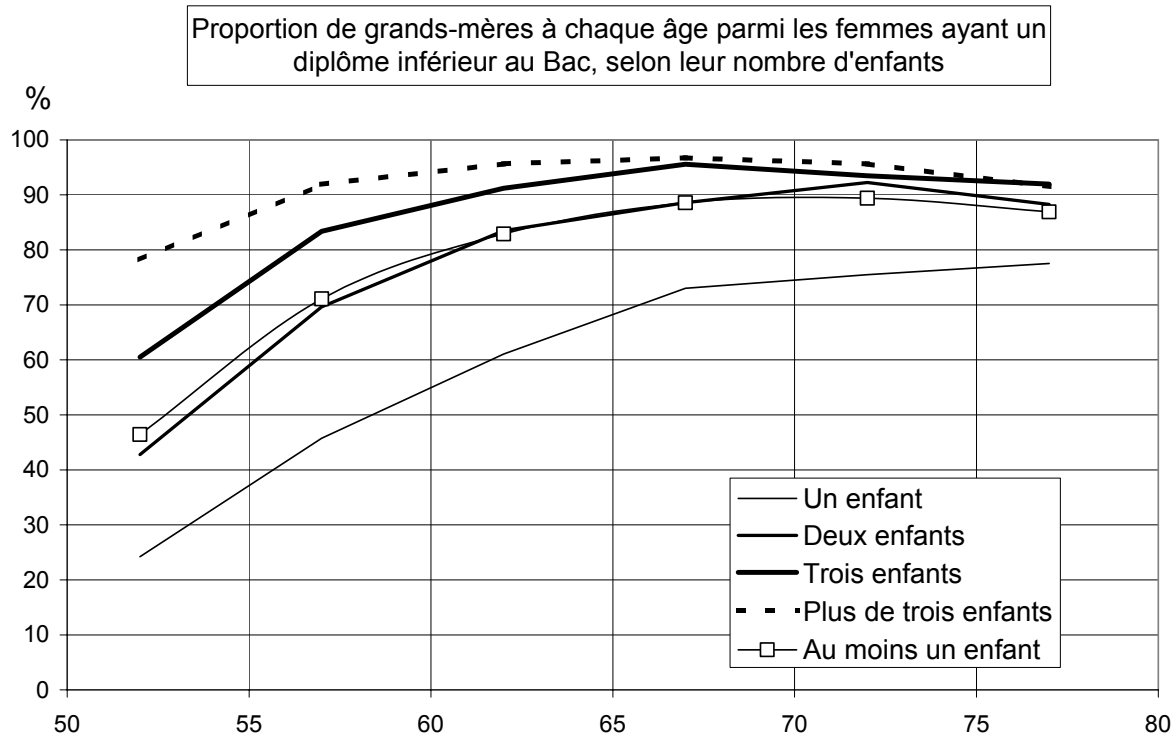
Graphique 1.7d



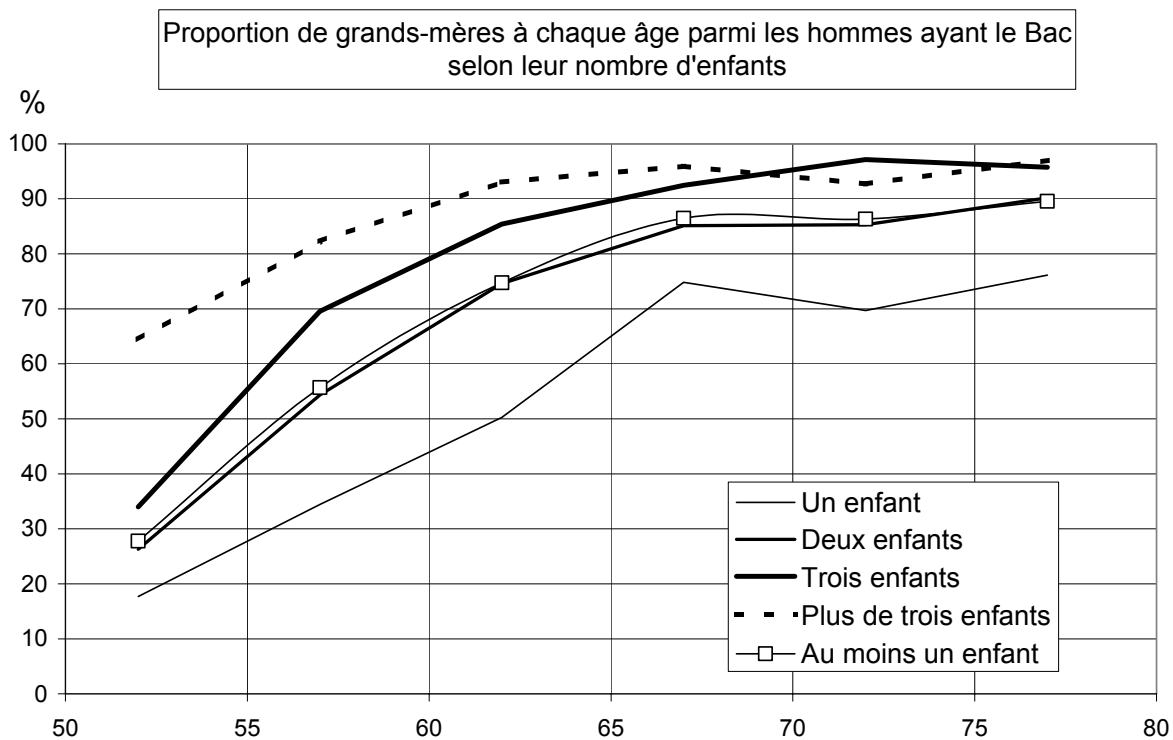
Graphique 1.8a



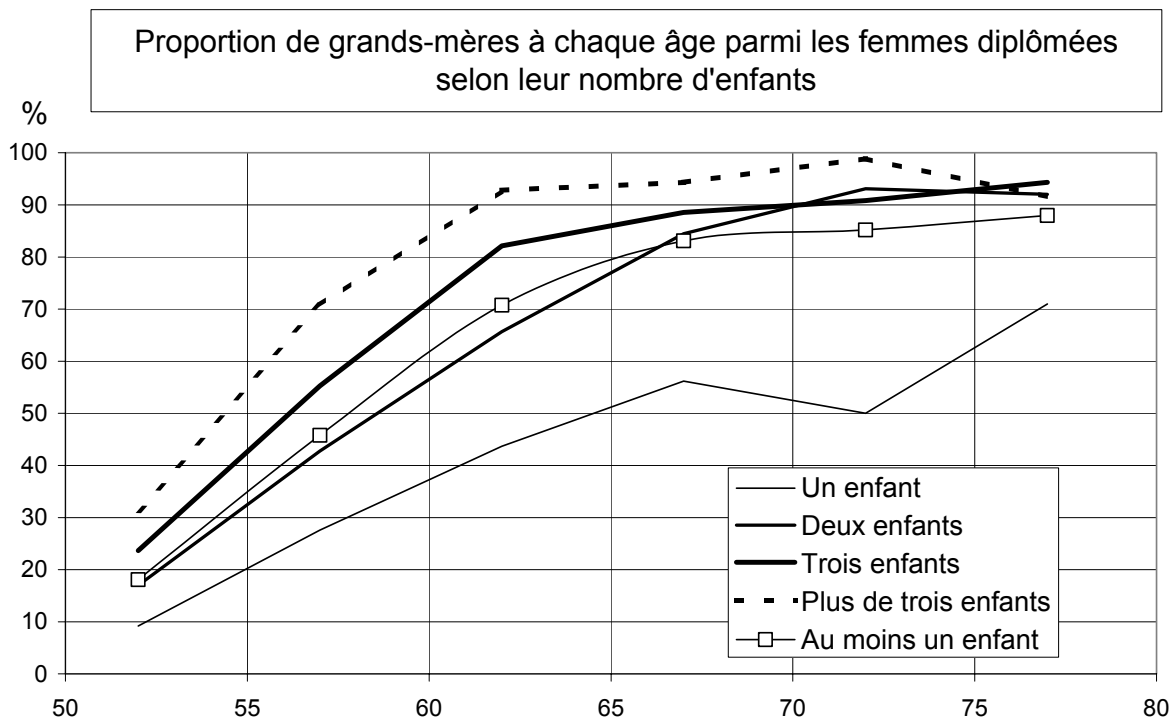
Graphique 1.8b



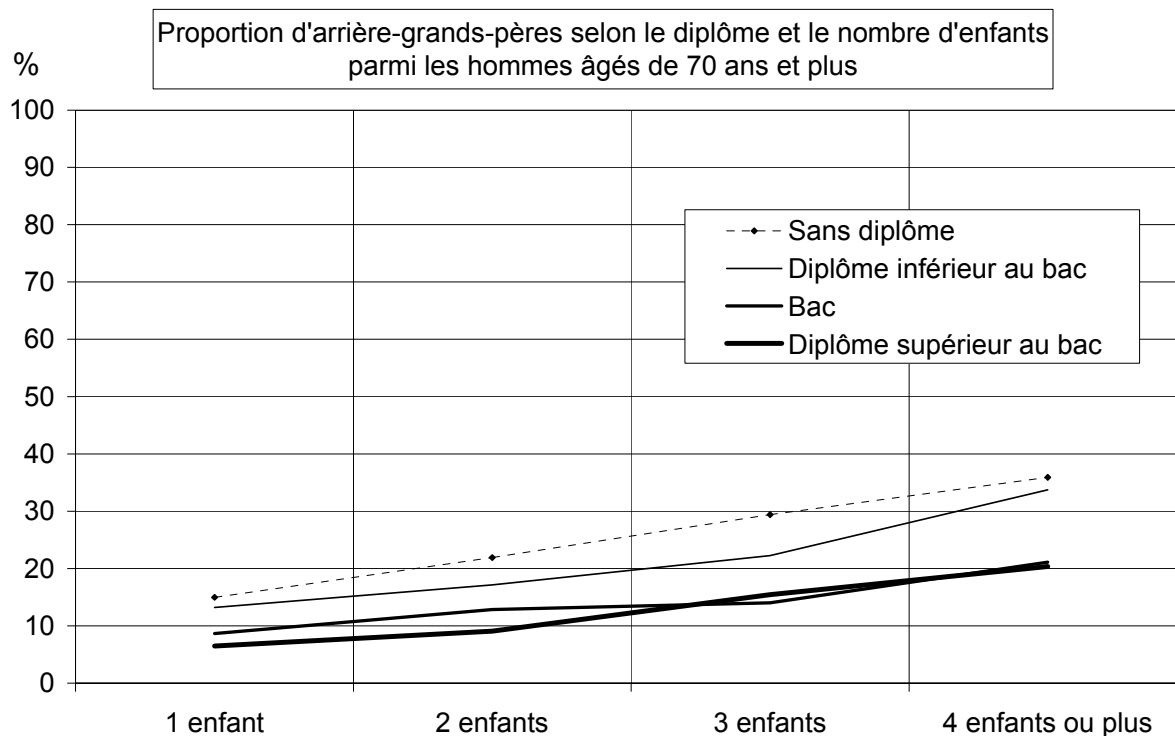
Graphique 1.8c



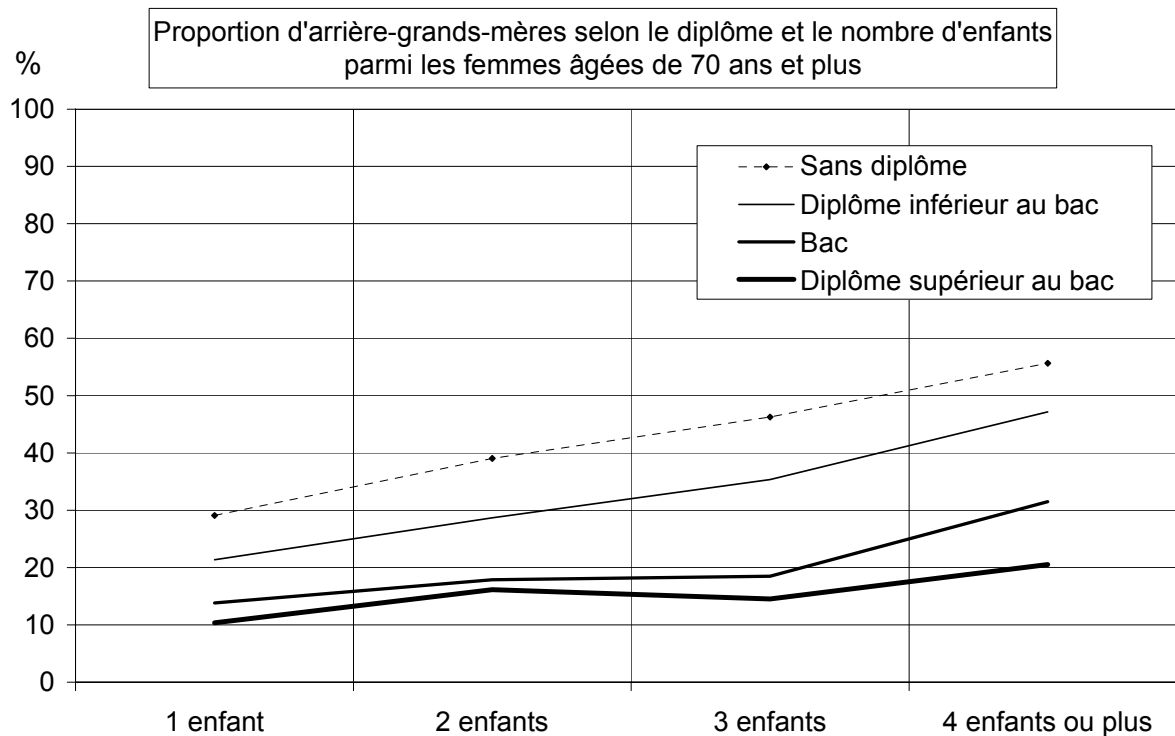
Graphique 1.8d



Graphique 1.9



Graphique 1.10



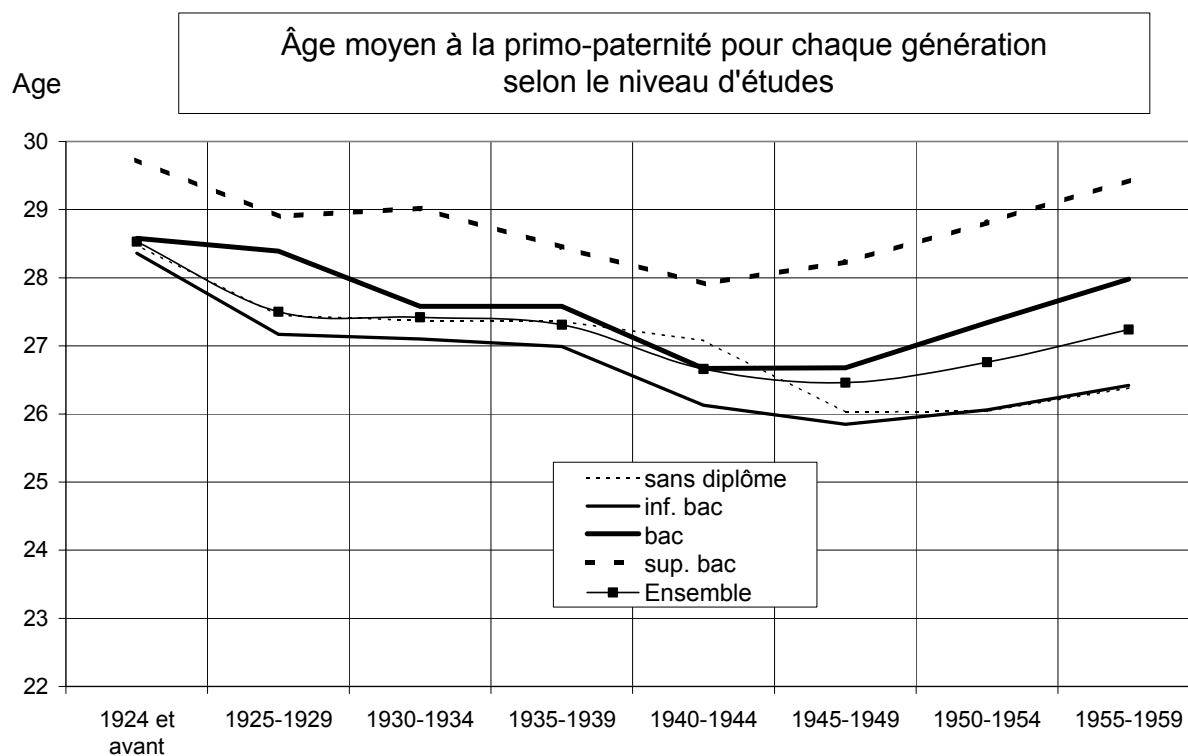
1.4 – Situations familiales, niveau d'études et âge au premier enfant

L'âge moyen au premier enfant est très différent selon le niveau de diplôme et les écarts croissent avec les générations (graphiques 1.11 et 1.12), en gardant la même hiérarchie. Les différentiels sont surtout marqués chez les femmes. Les femmes les moins diplômées ont leurs enfants beaucoup plus jeunes, alors que, pour les hommes, les différences dues au diplôme sont beaucoup plus atténuées. La baisse de l'âge à la venue du premier enfant s'interrompt avec les générations nées après la Seconde Guerre mondiale, pour ensuite reculer. La précocité du recul parmi les femmes les plus diplômées, analysée antérieurement pour les femmes (Daguet, 2000), se confirme comme on pouvait s'y attendre pour les hommes (graphique 1.11).

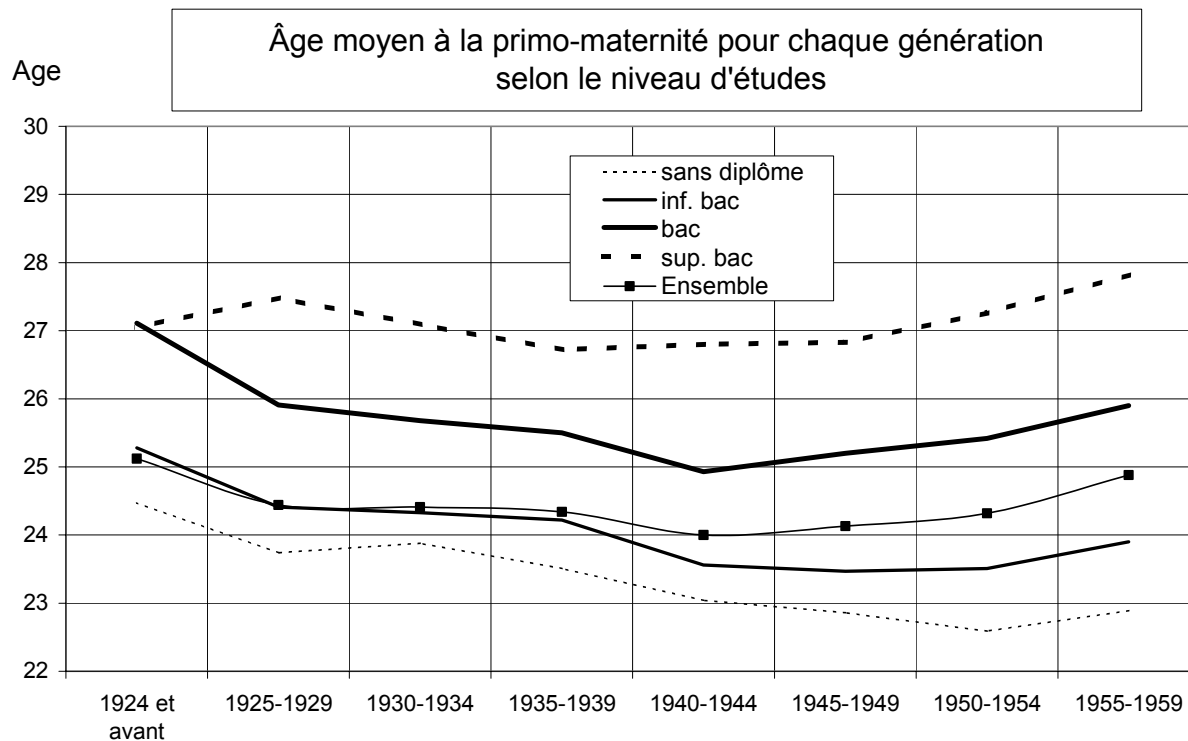
Si les différences relatives à l'âge au premier enfant sont importantes selon le diplôme, les différences observées en terme de grand-parentalité sont principalement dues au retard d'entrée dans la parentalité (graphiques 1.13 et 1.14). Retard que les individus ne rattrapent pas complètement, d'autant plus s'ils ont eu leur premier enfant après trente ans. Les deux graphiques montrent qu'il y a très peu de différences entre hommes et femmes. Les courbes se distinguent uniquement aux âges élevés : les femmes nées au début du siècle, diplômées du supérieur et ayant eu leur premier enfant après l'âge de trente ans, sont moins souvent grands-mères. Elles sont sans doute mères d'enfants uniques.

Les proportions décroissantes selon les générations âgées de plus de 65 ans ne s'expliquent plus par une fréquence des personnes restées sans enfant plus forte car nous raisonnons maintenant au niveau des parents, mais par la moindre proportion de personnes restées sans enfant parmi la population de leurs enfants.

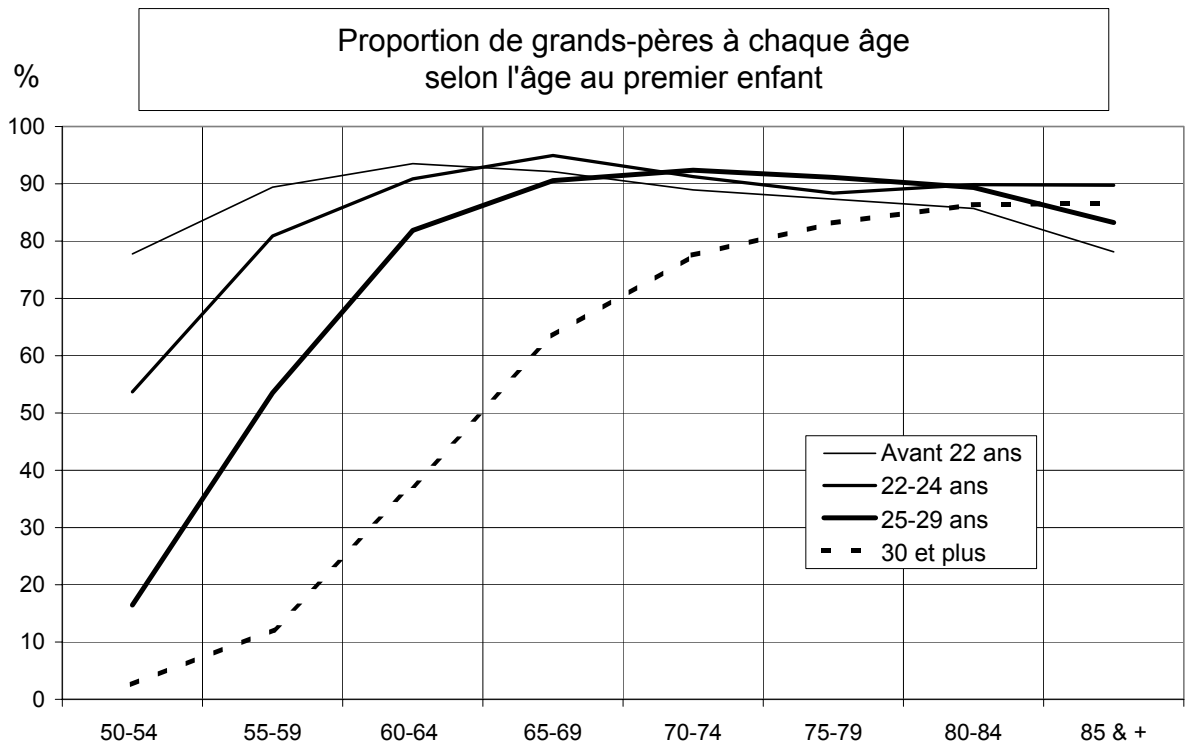
Graphique 1.11



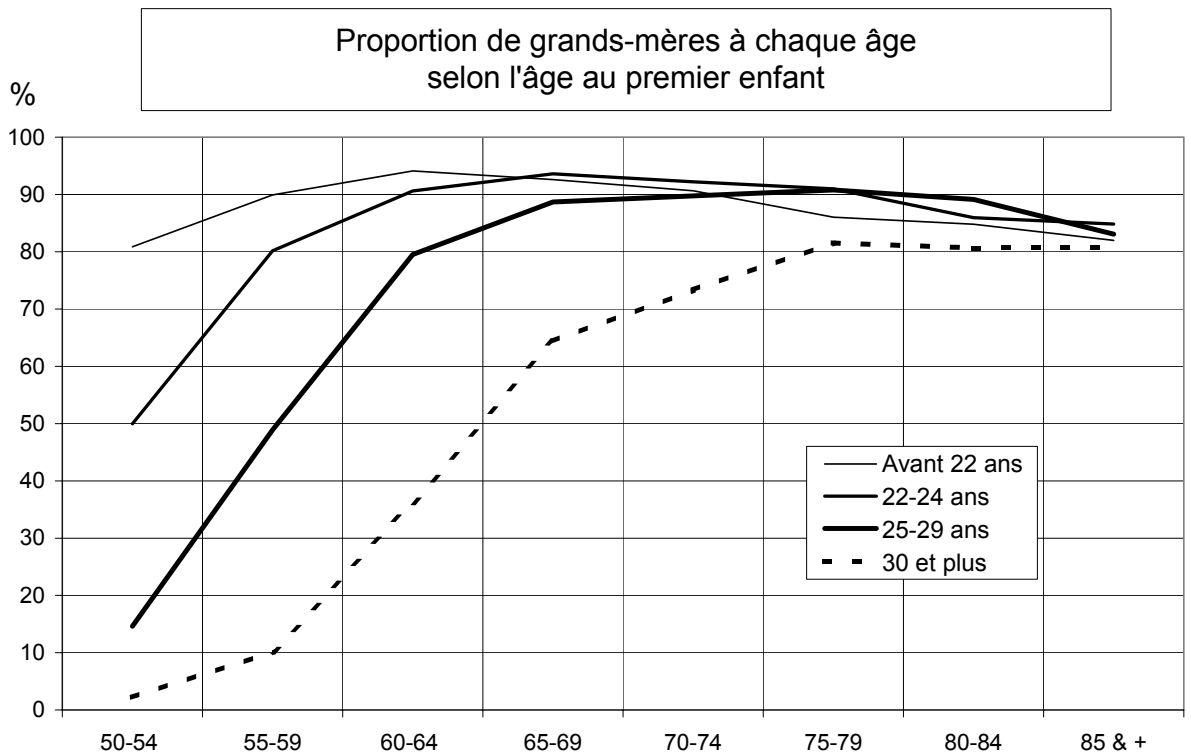
Graphique 1.12



Graphique 1.13



Graphique 1.14



Les quatre graphiques ci-dessous montrent les différences selon le diplôme, en termes d'âge au premier enfant et de nombre d'enfants.

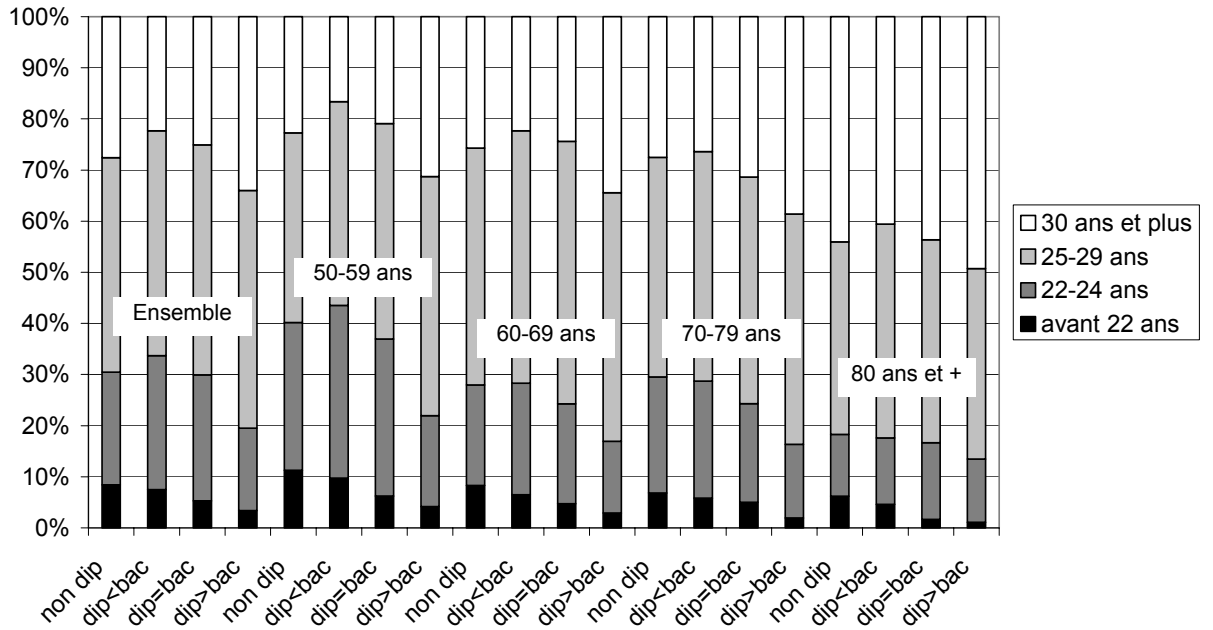
Comme nous venons de le voir, les hommes ont leurs enfants plus tard que les femmes et les différentiels sont plus marqués pour les femmes. Parmi les générations les plus âgées, presque la moitié des hommes avaient leur premier enfant après l'âge de trente ans, alors que, quelle que soit la génération, les femmes ont eu leur premier enfant le plus fréquemment avant 25 ans (graphiques 1.15 et 1.16). La hiérarchie selon le diplôme est toujours la même, quels que soient le sexe et la génération d'appartenance : les plus diplômés ont leur premier enfant de manière plus tardive. Ainsi, les hommes diplômés ont moins souvent que les autres leur premier enfant avant 25 ans et les femmes les plus diplômées sont rarement mères avant 22 ans.

Cette précocité se remarque également par les nombres moyens d'enfants : parmi les moins de trente ans, les hommes et femmes diplômés ont peu d'enfants alors que les femmes non diplômées ont en moyenne plus d'un enfant (graphiques 1.17 et 1.18). Les différences s'estompent pour les hommes mais pas pour les femmes.

Si l'on croise maintenant diplôme et âge au premier enfant, on s'aperçoit que l'influence du diplôme sur la grand-parentalité est beaucoup moins forte (graphiques 1.19 et 1.20). C'est bien le fait d'avoir son premier enfant tôt qui va impliquer que l'on est moins souvent grand-parent. Cependant, parmi les personnes âgées de 50 à 59 ans, l'effet du diplôme est toujours visible, et particulièrement parmi les personnes qui ont eu leur premier enfant après 30 ans. On peut donc penser que leurs enfants ont également eu une fécondité plus tardive.

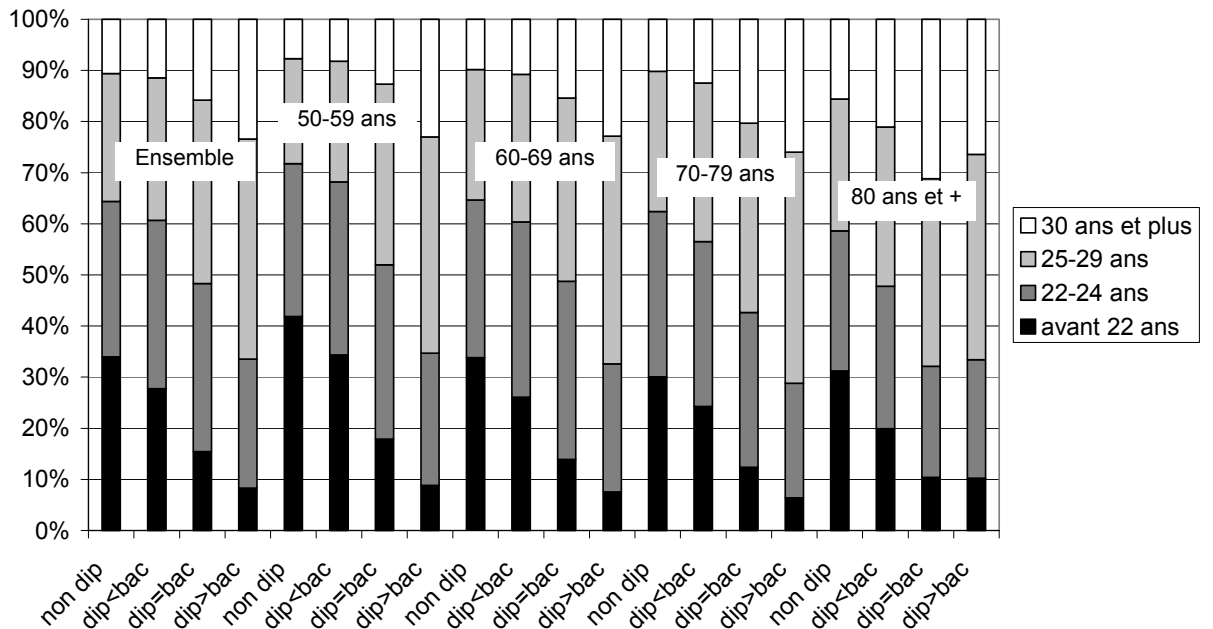
Graphique 1.15

Répartition des pères (âgés de 50 ans et plus) par âge au premier enfant, selon le niveau d'études



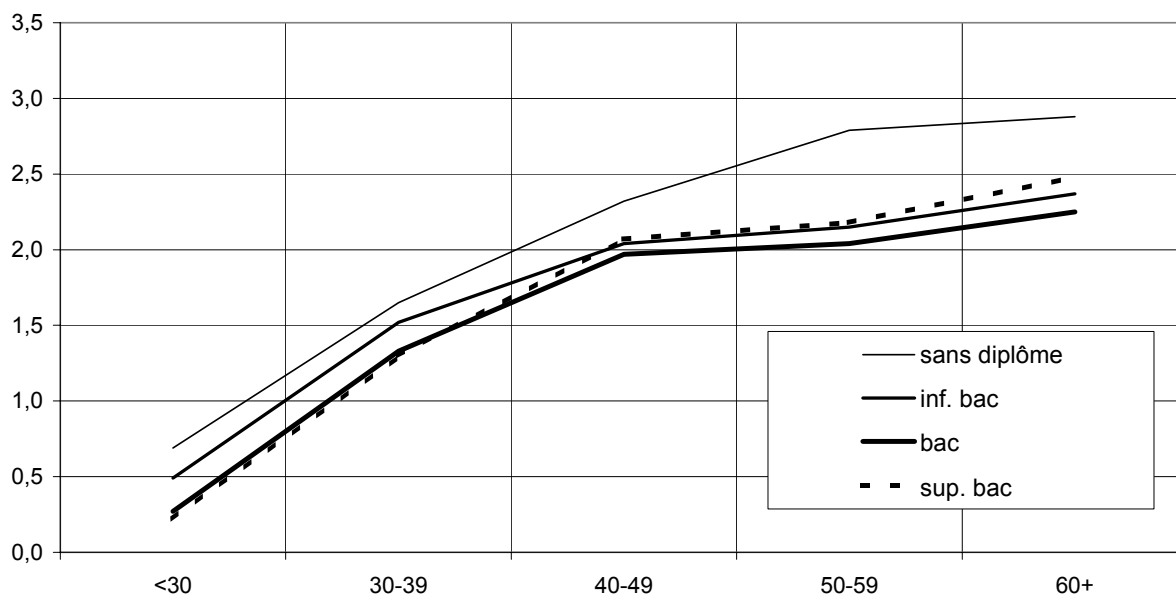
Graphique 1.16

Répartition des mères (âgées de 50 ans et plus) par âge au premier enfant, selon le niveau d'études



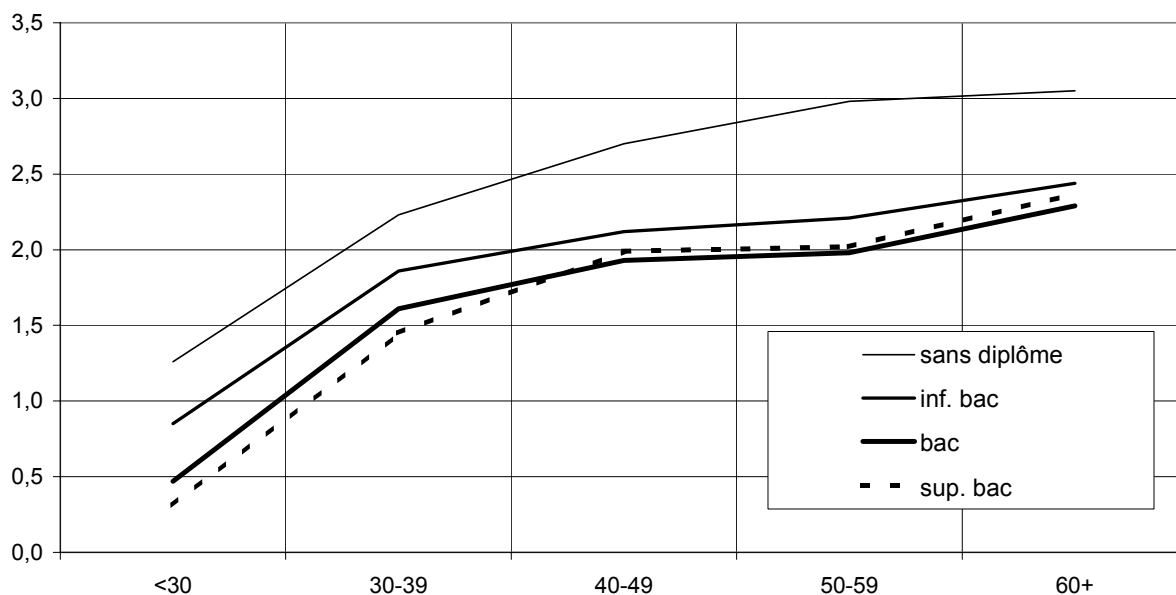
Graphique 1.17

Nombre moyen d'enfants des hommes en union (rompue ou en cours)
selon le diplôme et l'âge

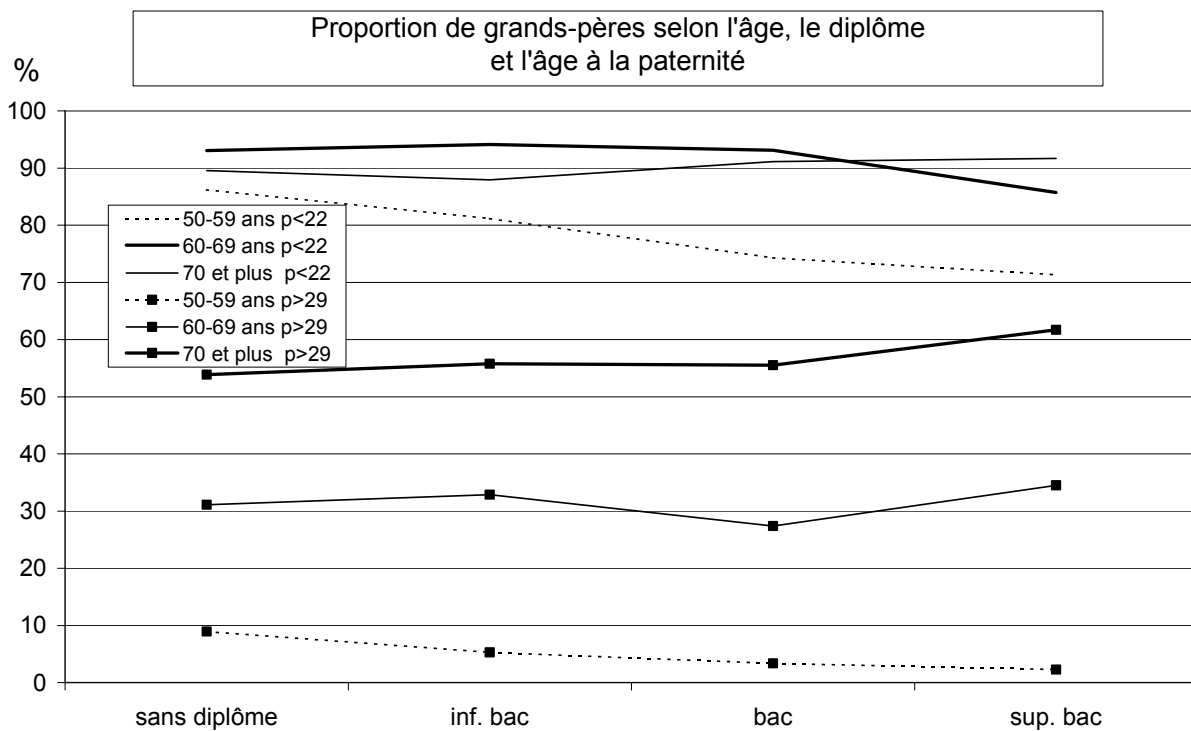


Graphique 1.18

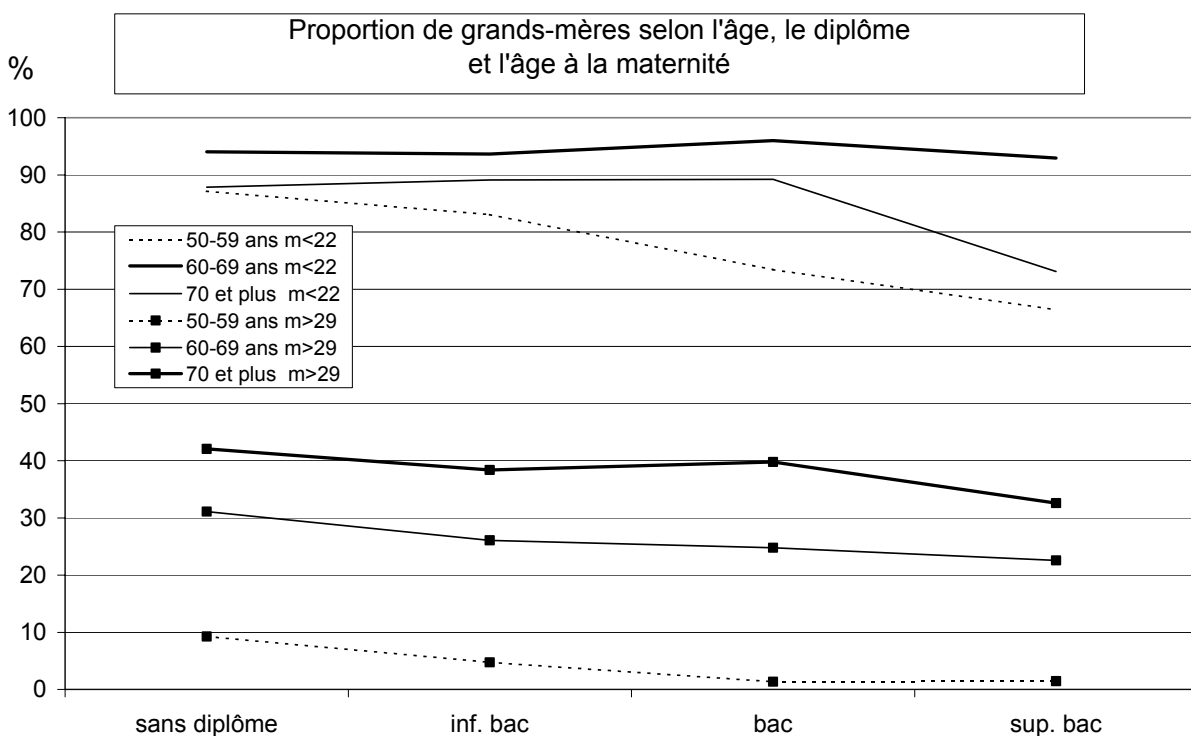
Nombre moyen d'enfants des femmes en union (rompue ou en cours)
selon le diplôme et l'âge



Graphique 1.19



Graphique 1.20



1.5 – Régression logistique sur le risque d'être grand-parent

Nous présentons ici plusieurs modèles de régression qui synthétisent les résultats vus ci-dessus.

Dans un premier temps nous avons calculé le risque sur l'ensemble de la population (tableau 1.2), ensuite nous avons effectué une régression uniquement sur les personnes qui ont eu des enfants, afin de mesurer l'effet du nombre d'enfants et de l'âge au premier enfant (tableau 1.3). Nous avons choisi de faire une régression pour chaque sexe, étant donné que certains facteurs ne jouent pas dans le même sens pour les hommes et pour les femmes.

Nous avons placé dans le tableau les résultats traduits en pourcentages. Ainsi par exemple dans le premier modèle, un homme peu diplômé (diplôme inférieur au bac) âgé de 65 à 69 ans a 74% de chances d'être grand-père. Une femme avec les mêmes caractéristiques a 78,6% de chances. Par contre, si on considère un homme sans diplôme, toutes choses égales par ailleurs, la probabilité perd alors 3 points (71,1%).

Tableau 1.2 : Régression sur le risque d'être grand-parent
(ensemble des individus)

Variables indépendantes	<i>Hommes</i>		<i>Femmes</i>	
	Modèle 1	Modèle 2	Modèle 1	Modèle 2
<i>Modèle</i>	74,0 ***	77,7 ***	78,6 ***	82,0 ***
<i>Diplôme</i>				
. Aucun	71,1 ***	76,8 ns	82,0 ***	84,6 ***
. Inférieur au bac	(-)	(-)	(-)	(-)
. Bac	74,0 ns	77,4 ns	72,6 ***	77,2 ***
. Supérieur	72,3 *	75,4 **	66,3 ***	74,2 ***
<i>Groupe d'âge</i>				
. 60-64 ans	65,8 ***	69,5 ***	75,4 ***	78,1 ***
. 65-69 ans	(-)	(-)	(-)	(-)
. 70-74 ans	76,8 ***	80,4 ***	77,2 **	81,1 ns
. 75-79 ans	76,6 **	80,2 ***	76,7 ***	80,4 ***
. 80 et plus	75,6 *	78,5 ns	70,9 ***	75,2 ***
<i>Vie de couple</i>				
. Aucune union		5,3 ***		16,2 ***
. Une union		(-)		(-)
. Plus d'une union		74,9 ***		80,2 ***
Degré de significativité : * 0,05 ; ** 0,01 ; *** 0,001				
(-) Modalité de référence				

Nous retrouvons l'effet inverse du diplôme analysé en première partie : les hommes non diplômés sont moins souvent grands-parents alors que les femmes non diplômées le sont plus souvent. Ainsi, si nous faisons uniquement varier ce critère, un homme de 65-69 ans

non diplômé a 71,1% de chances d'être grand-père, alors qu'une femme ayant les mêmes caractéristiques a 82% de chances d'être grand-mère. La prise en compte de la vie de couple rend l'influence du diplôme plus faible pour les hommes, mais pour les femmes, les différences selon le diplôme restent importantes.

La probabilité augmente avec l'âge, mais dans une moindre mesure pour les hommes les plus âgés. La probabilité décroît avec les générations pour les femmes âgées de plus de 70 ans, et lorsque nous introduisons la vie de couple cet effet persiste. Il s'explique sans doute en partie par la fréquence des femmes restées sans enfant, plus forte pour les générations nées au début du siècle, comme nous l'avons vu.

Le fait d'introduire la vie de couple et de prendre comme référence une personne âgée de 65-69 ans, peu diplômée et qui a eu un conjoint, augmente la probabilité de plus de trois points (on passe de 74% à 77,7% pour les hommes et de 78,6 à 82% pour les femmes).

Les deux autres modèles portent sur la population des parents.

Le modèle 3 fait apparaître les différences liées au nombre d'enfants. Quant aux autres paramètres, ils ont toujours une influence. Le fait de prendre en compte le nombre d'enfants et l'âge au premier enfant réduit l'influence négative observée aux âges élevés, sans doute parce que ces générations avaient eu leurs enfants plus tard (modèle 4). Plus on a d'enfants et plus on les a tôt, plus on a de chances d'être grand-parent, mais le fait d'avoir une famille nombreuse, même si la venue du premier enfant n'était pas précoce, renforce la probabilité d'être grand-parent.

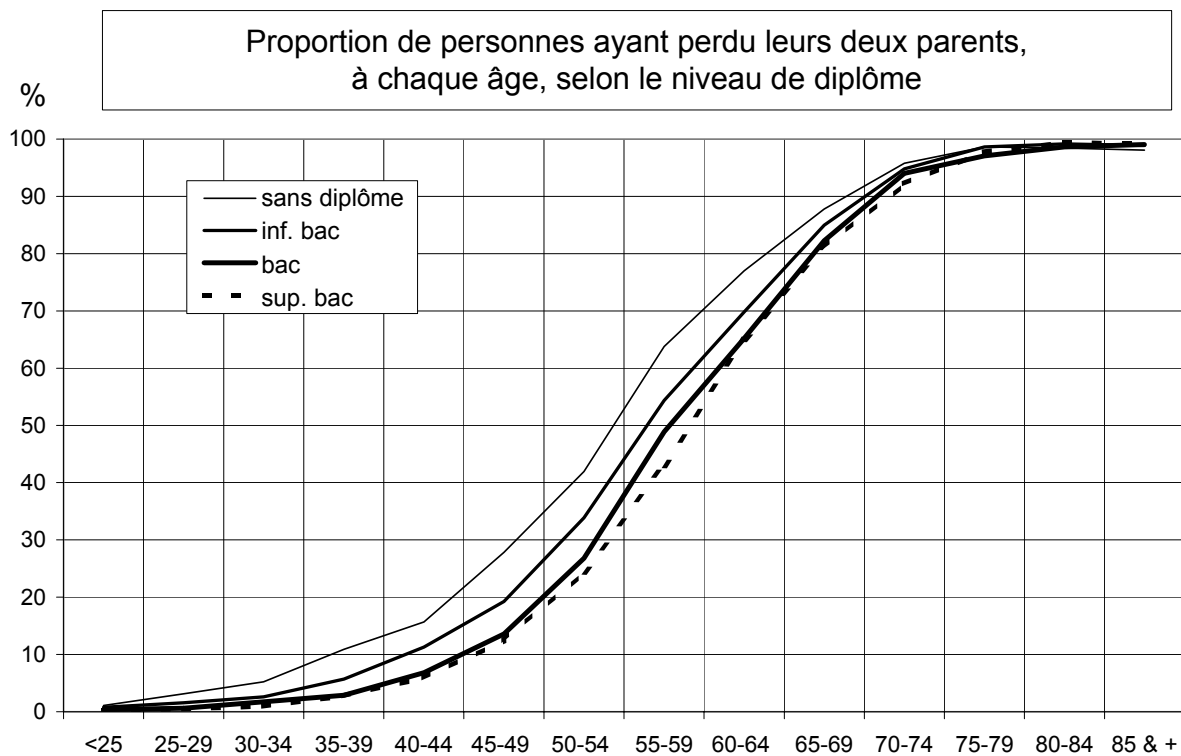
Tableau 1.3 : Régression sur le risque d'être grand-parent
(population des parents)

Variables indépendantes	<i>Hommes</i>		<i>Femmes</i>	
	Modèle 3	Modèle 4	Modèle 3	Modèle 4
<i>Modèle</i>	85,8 ***	92,5 ***	88,9 ***	92,5 ***
<i>Diplôme</i>				
. Aucun	85,2 ns	92,3 ns	89,7 *	93,0 *
. Inférieur au bac	(-)	(-)	(-)	
. Bac	85,0 ns	92,2 ns	86,7 ***	91,4 **
. Supérieur	80,5 ***	90,0 ***	83,1 ***	89,5 ***
<i>Groupe d'âge</i>				
. 60-64 ans	74,8 ***	85,8 ***	82,9 ***	88,0 ***
. 65-69 ans	(-)	(-)	(-)	
. 70-74 ans	88,8 ***	94,2 ***	89,4 ns	92,8 ns
. 75-79 ans	89,5 ***	94,6 ***	89,9 *	93,3 **
. 80 et plus	87,8 **	93,9 ***	85,9 ***	90,4 ***
<i>Nombre d'enfants</i>				
. Un enfant	68,7 ***		74,6 ***	
. Deux enfants	(-)		(-)	
. Trois enfants	91,7 ***		93,4 ***	
. Plus de trois enfants	93,1 ***		94,2 ***	
<i>Nombre d'enfants et âge au premier enfant</i>				
. 1 enfant (<25)		80,2 ***		81,3 ***
. 2 enfants (<25)		(-)		(-)
. 3 enfants (<25)		95,1 ***		94,7 ***
. +3 enfants (<25)		94,5 **		94,1 ***
. 1 enfant (>=25)		66,2 ***		70,0 ***
. 2 enfants (>=25)		83,7 ***		85,1 ***
. 3 enfants (>=25)		90,3 ***		91,3 **
. +3 enfants (>=25)		92,2 ns		94,5 ***
Degré de significativité : * 0,05 ; ** 0,01 ; *** 0,001				
(-) Modalité de référence				

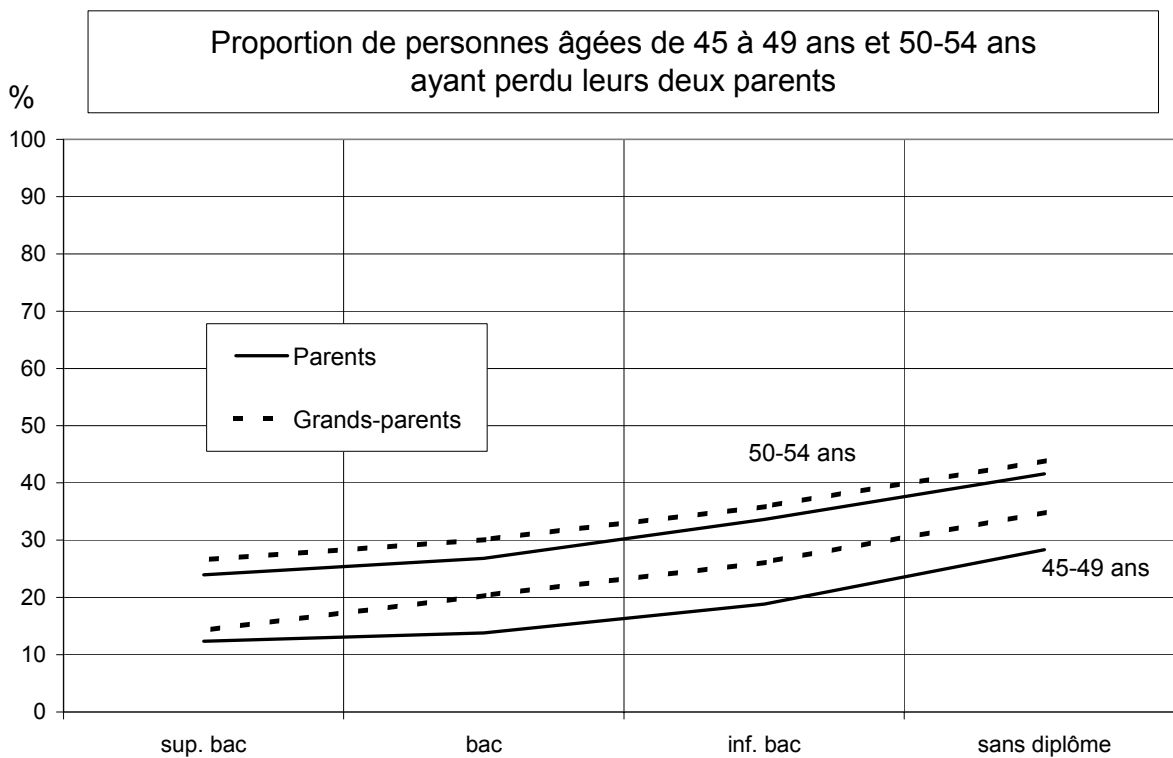
1.6 – Situations familiales, niveau d'études et perte des parents

Les précédents résultats avaient mis en évidence une perte plus fréquente des parents parmi les personnes âgées de moins de 55 ans, lorsque ces derniers étaient grands-parents (graphiques 1.9 et 1.10, p.16 du *Dossiers et Recherches N°93*). Nous avons supposé alors que cette perte plus précoce des parents lorsque l'on était un jeune grand-parent s'expliquait par la mortalité différentielle : les jeunes grands-parents appartiendraient à des milieux plus défavorisés. Le graphique 1.21 montre effectivement une perte plus précoce des parents parmi les moins diplômés. Cependant, si l'on croise diplôme et situation parentale, il persiste une différence, à niveau de diplôme égal (graphique 1.22). Cette différence s'explique sans doute en partie par un effet de structure de ces deux populations et peut-être aussi par le fait que le diplôme n'est pas un indicateur suffisant pour expliquer des différentiels sociaux. Le fait d'être ou non diplômé ne renseigne pas directement sur le milieu social des parents.

Graphique 1.21



Graphique 1.22



2 – Analyse de la fécondité : milieu social et sexe du locuteur

Les données de l'enquête EHF nous permettent de produire des résultats sur la fécondité différentielle selon le sexe de la personne et selon sa propre profession et catégorie sociale, ce qui est encore peu courant à l'heure actuelle. Elles nous permettent aussi de produire des résultats sur le nombre moyen d'enfants des hommes selon la P.C.S. de leur conjointe, ce qui est encore moins courant. Ainsi ces nombres moyens vont présenter plus ou moins d'hétérogénéité selon le type de témoignage recueilli. Les contrastes les plus forts étant ceux de la fécondité féminine selon leur propre P.C.S..

Nous avons cette fois-ci utilisé la P.C.S., cet indicateur étant plus concret en terme de milieu de vie que ne l'est le diplôme : il renseigne plus sur le moment d'entrée dans la vie active. D'autre part, à niveau d'étude égal, le type d'activité peut varier énormément. De plus, des résultats en terme de P.C.S. permettent une comparaison avec la littérature existante sur fécondité et milieu social.

Nous présentons tout d'abord les nombres moyens d'enfants des hommes et des femmes selon leur propre P.C.S.. Les proportions de personnes qui n'ont jamais vécu en couple peuvent varier fortement, nous avons donc produit des résultats sur la fécondité pour les hommes et les femmes en général, puis parmi ceux qui ont vécu en couple.

Dans un deuxième temps, nous avons analysé la fécondité selon la P.C.S. du conjoint, ce qui est fait traditionnellement pour les femmes.

2.1 – Fécondité des hommes et des femmes selon leur propre P.C.S.

Les graphiques ci-dessous montrent plusieurs choses : une hétérogénéité plus forte de la fécondité des femmes que de la fécondité des hommes (graphiques 1.23 et 1.24). Aux âges jeunes, si les inactifs ont peu d'enfants c'est parce qu'ils ne vivent pas en couple (graphique 1.25).

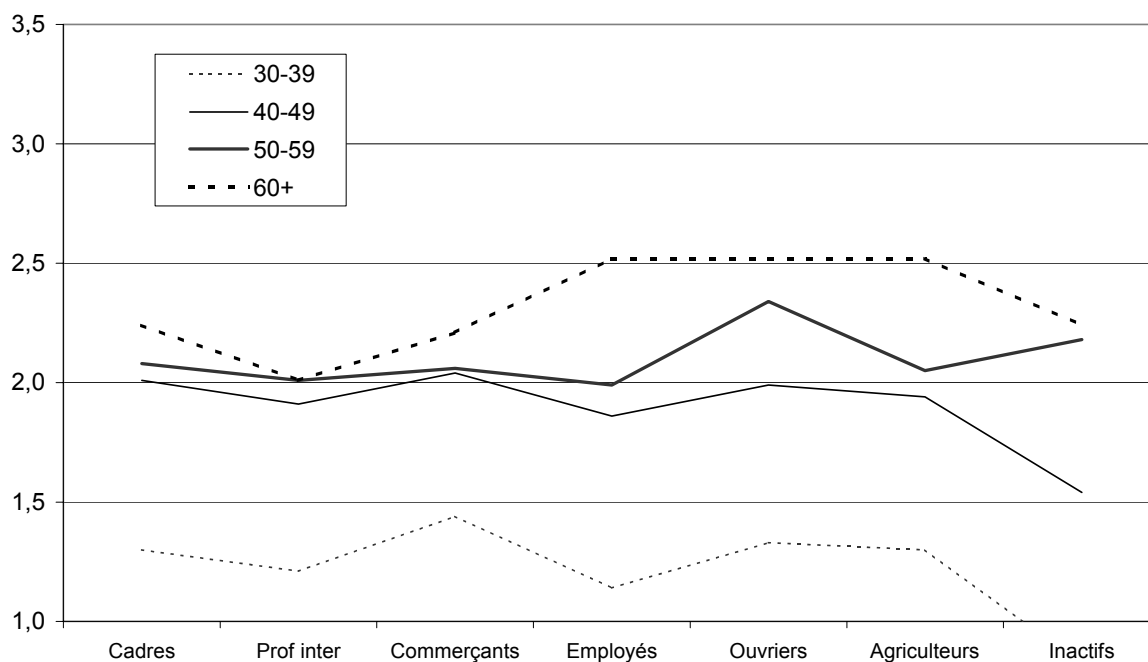
Ces données montrent le célibat connu des hommes agriculteurs. Ce phénomène concerne fortement les jeunes générations. Les femmes agricultrices, elles, sont rarement célibataires, tout comme les inactives et les commerçantes. Trois statuts qu'il serait difficile d'avoir sans conjoint.

L'étude de la fécondité des personnes ayant vécu en union montre très peu de différences selon la P.C.S. pour les hommes. La fécondité des inactifs est beaucoup plus élevée que précédemment. Pour les femmes, en revanche, les différences liées à la P.C.S. ne changent pas, seule l'intensité de la fécondité est sensiblement augmentée.

Nous avons choisi de représenter les données par des courbes par souci de lisibilité, il va de soi qu'il n'y a aucune hiérarchie dans ce classement.

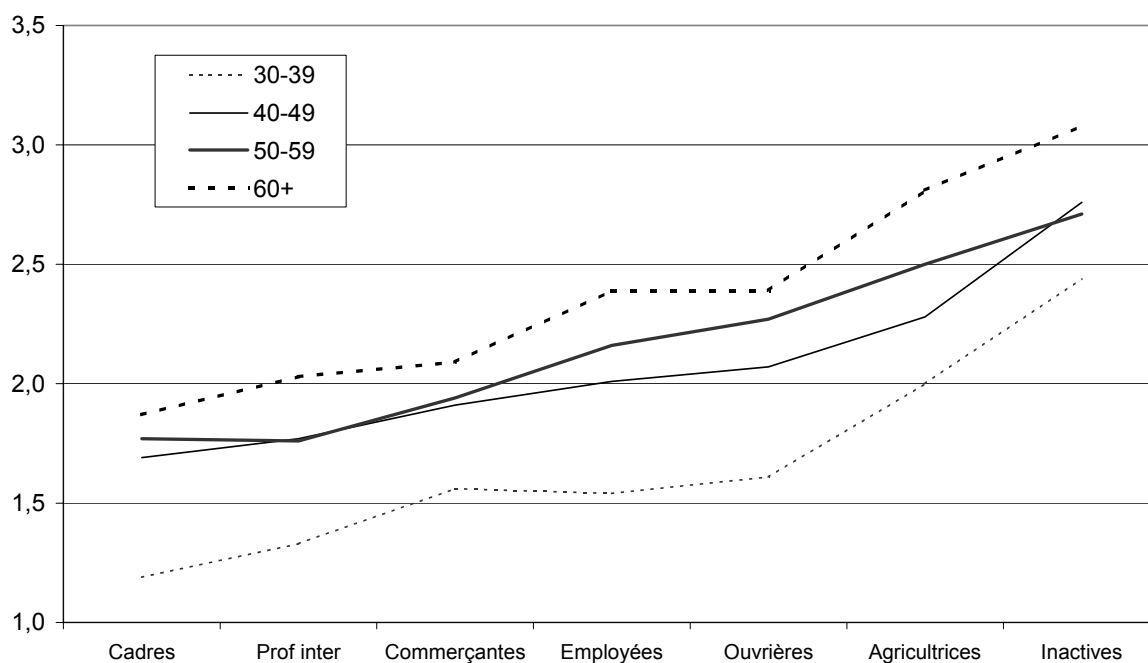
Graphique 1.23

Nombre moyen d'enfants des hommes selon leur âge à l'enquête et leur P.C.S

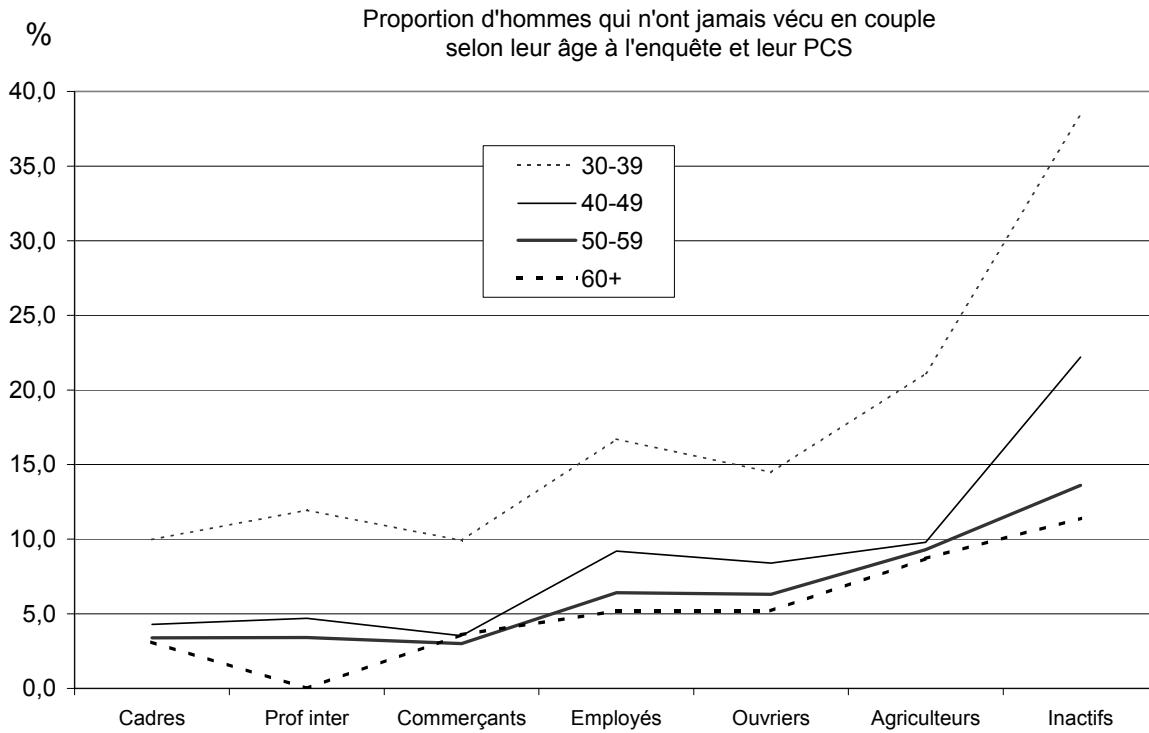


Graphique 1.24

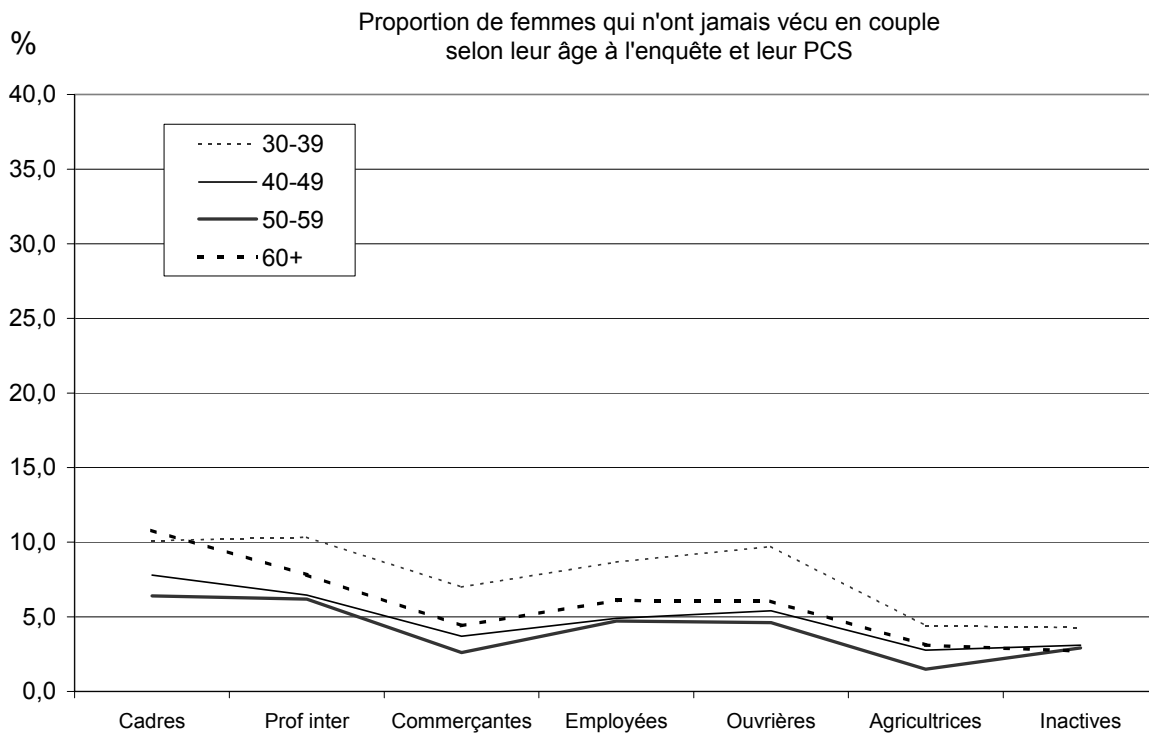
Nombre moyen d'enfants des femmes selon leur âge à l'enquête et leur P.C.S



Graphique 1.25

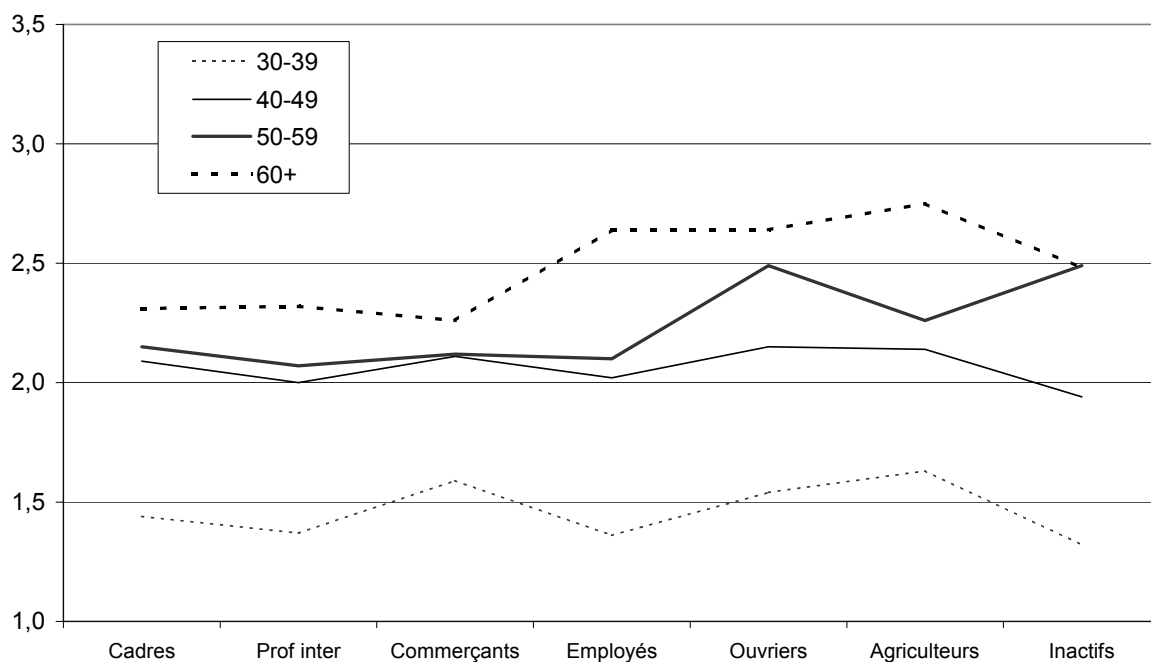


Graphique 1.26



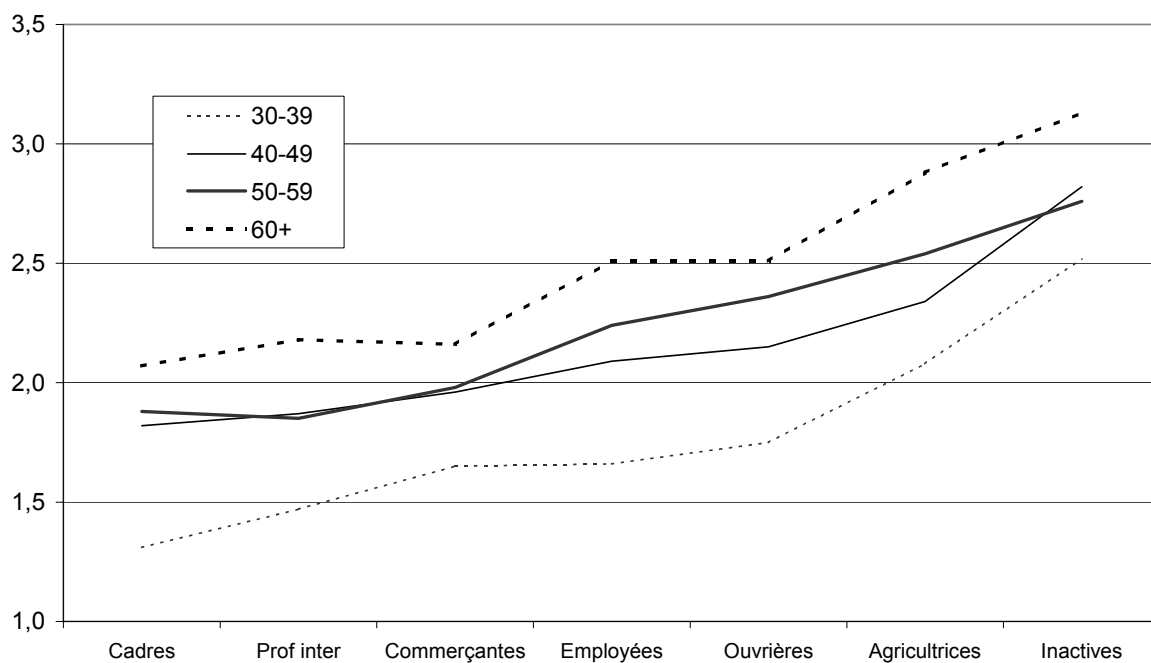
Graphique 1.27

Nombre moyen d'enfants des hommes en union (rompue ou en cours)
selon leur âge à l'enquête et leur P.C.S



Graphique 1.28

Nombre moyen d'enfants des femmes en union (rompue ou en cours)
selon leur âge à l'enquête et leur P.C.S



2.2 – Fécondité des hommes et des femmes selon la P.C.S. de leur conjoint

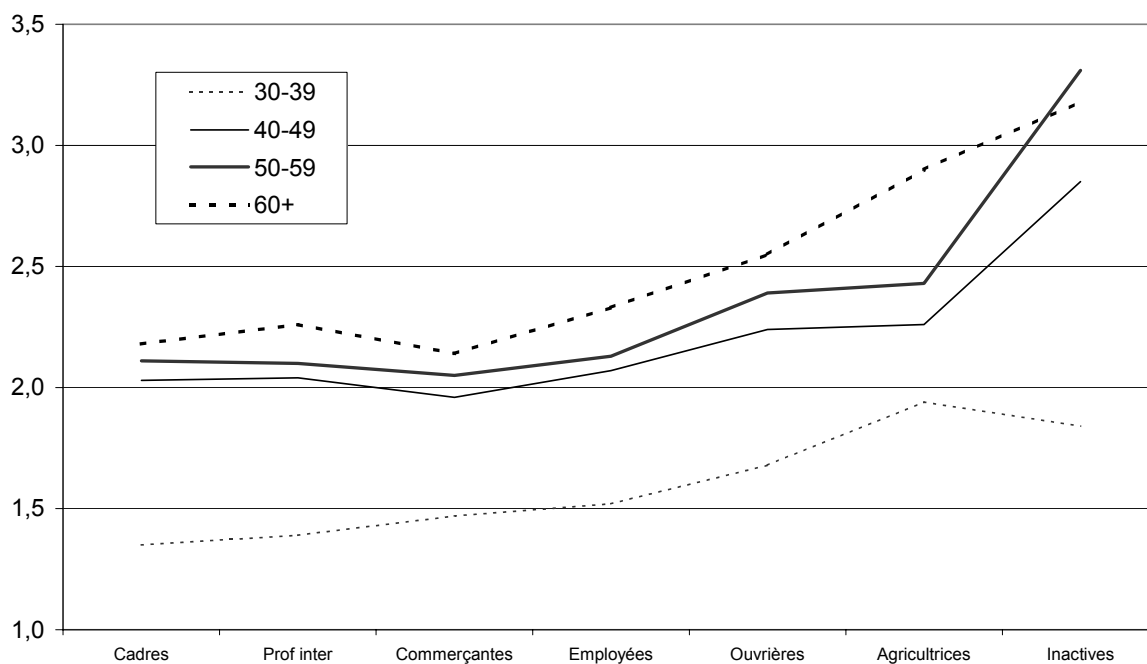
Les résultats présentés ci-dessous sont classiques pour les femmes alors qu'ils sont inhabituels pour les hommes (graphiques 1.30 et 1.29).

Si l'on observe le dernier graphique, il donne l'impression que les différentiels sociaux disparaissent, la courbe oscillant pour les 30-39 ans entre 1,6 et 2 enfants selon le milieu social de référence. Pourtant le graphique 1.28 faisait apparaître de nettes différences selon la P.C.S. des femmes. D'ailleurs, la fécondité des hommes étudiée selon la P.C.S. de leur conjointe (graphique 1.29) est plus hétérogène que lorsqu'on la rapproche de leur propre P.C.S. (graphique 1.27).

Ces deux types de production de données montrent que le sexe du locuteur et le type d'approche privilégiée conduisent à des résultats non contradictoires mais en tout cas différents. Les différentiels observés pour les femmes en tenant compte de leur propre P.C.S. donnent une vision assez différente de la fécondité féminine.

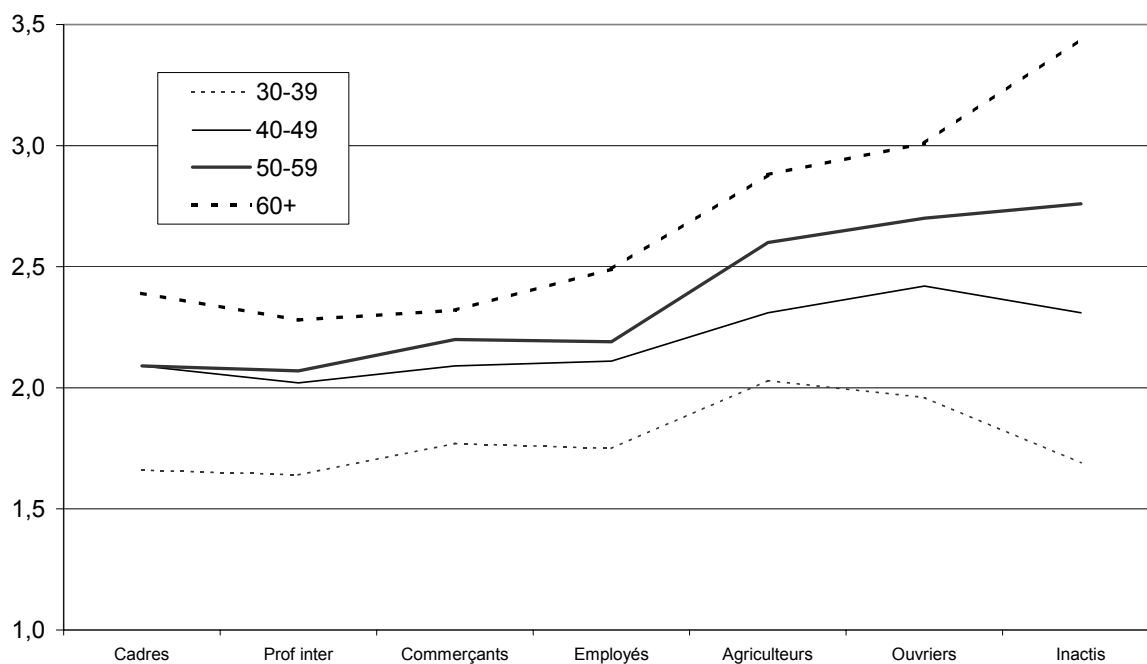
Graphique 1.29

Nombre moyen d'enfants des hommes selon leur âge à l'enquête
et la P.C.S de leur conjointe



Graphique 1.30

Nombre moyen d'enfants des femmes selon leur âge à l'enquête
et la P.C.S de leur conjoint



Seconde partie

Disponibilité et qualité des données

Les résultats présentés dans ce rapport et le précédent sont issus de versions provisoires du fichier de l'enquête. Des travaux de corrections, de validation et d'enrichissement ont été effectués depuis 1999, et sont depuis le mois de juillet 2002 disponibles dans une version définitive du fichier.

Les principaux changements depuis la dernière version concernent les pondérations. Plusieurs types de pondérations ont été élaborés, comme nous l'expliquons ci-dessous.

Tout au long de nos travaux sur la parentalité s'est posée la question de la validité des données. Dans le fichier enfants en particulier, nous avons ôté certains enregistrements qui figuraient en double, comme nous l'avons expliqué dans la seconde partie du *Dossiers et Recherches N°93*. Nous avons également mis en évidence que les résultats en terme de nombre d'enfants sont plus éloignés de la statistique d'état civil quand on les rapproche des déclarations masculines que des déclarations féminines. Cette estimation de la qualité de la source a été poursuivie : en parallèle avec l'exploitation des données, nous avons mené un travail comparatif sur les déclarations relatives aux enfants de la personne, recueillies dans le cadre de l'enquête EHF et de l'enquête Biographies et entourage. Au terme de cette comparaison, on compte moins de 5% de discordance quant aux déclarations sur le nombre d'enfants biologiques, et une part encore plus faible d'information manquante. Les omissions relèvent de trois logiques : la sous-déclaration des enfants mort-nés ou décédés en bas âge, la parentalité biologique sans contacts ultérieurs avec le ou les enfant(s), et les conflits familiaux entraînant une rupture des contacts.

1 – Déclarations relatives aux enfants et protocole d'enquête¹

Par le biais d'un échantillon apparié issu de l'enquête EHF et d'une enquête ultérieure dont le protocole était complètement différent, il a été possible de confronter les déclarations relatives aux enfants. On dispose, d'une part, des réponses au questionnaire EHF (qui était, rappelons le, sous la forme d'un quatre pages auto-administré) et, d'autre part, des données de l'Enquête Biographies et Entourage, une source riche en informations, collectées par le biais d'un enquêteur lors d'entretiens en face à face auprès d'un échantillon représentatif à un niveau local.

Les confrontations peuvent être faites sur des thèmes communs aux deux questionnaires, tels que l'histoire matrimoniale, les enfants et beaux-enfants, les professions, la taille de la fratrie etc. Sont présentés ici les résultats comparatifs portant sur les enfants biologiques de la personne. Ils mettent en évidence une forte concordance des déclarations relatives aux enfants ; quant aux déclarations discordantes, elles résultent de logiques clairement repérables.

Ce travail est issu d'une réflexion commune, menée avec Eva Lelièvre *et al.*, visant à comparer des informations recueillies par l'intermédiaire de deux protocoles d'enquête complètement différents. Il s'inscrit également dans la perspective méthodologique qui vise à estimer la qualité des données du fichier EHF. Par ailleurs, cette comparaison a permis un contrôle de cohérence du travail de remaniement des données, présenté dans la seconde partie du rapport intermédiaire, consacrée à la validation des données.

L'estimation de la qualité des données peut être faite soit en utilisant la seule source de manière interne (par le croisement de données individuelles), soit en confrontant les résultats avec d'autres sources de données ou avec la même source à deux dates. L'originalité de ce travail consiste en un croisement avec une autre source de données, au niveau individuel, grâce à un échantillon apparié.

L'enquête EHF, dont le but est de reconstituer « l'histoire familiale des individus », et dont les bulletins furent remplis par les personnes elles-mêmes en même temps que les formulaires du recensement en l'absence d'enquêteur, peut souffrir de certains biais. Dans le cadre d'un questionnaire auto-administré, l'enquêté est seul arbitre de ce qu'il va répondre, du tri qu'il va devoir effectuer parmi les multiples événements qui ont jalonné sa vie.

¹ Ce travail a été présenté lors du XII^e colloque national de démographie : « Collecte des données et connaissance des populations », organisé par la Conférence universitaire de démographie et d'étude des populations – Amiens, mai 2002.

L'enquête « Biographies et entourage » a été réalisée ultérieurement à l'Enquête Famille, elle porte sur un échantillon de 2 830 personnes, représentatif des résidents en île de France, nés entre 1930 et 1950. Elle a pour objectif de retracer l'histoire familiale, résidentielle et professionnelle des enquêtés et de leur entourage. Les principaux thèmes abordés sont : la trajectoire familiale, la mobilité résidentielle, la géographie de la parenté, la carrière professionnelle des enquêtés et de leurs proches.

Les deux protocoles

L'enquête EHF s'appuie sur un questionnaire auto-administré, qui synthétise en quatre pages l'histoire familiale des individus. Les districts hommes et femmes sélectionnés « Enquête Famille » sont des districts disjoints. Ainsi pour un district sélectionné tous les hommes d'un ménage devaient répondre à l'enquête. Nous n'avons donc aucun renseignement sur les conditions de remplissage du questionnaire (la personne enquêtée était-elle seule ou un tiers était-il présent lors du remplissage du questionnaire, le questionnaire a-t-il été rempli réellement par l'enquêté en personne ? ...). L'enquête Biographies et entourage a au contraire été réalisée par des entretiens en face à face, avec une durée moyenne de passation du questionnaire d'une heure et quarante minutes. L'information a été collectée de la manière la plus exhaustive possible, afin que l'individu n'ait pas à faire de « tri » dans sa vie, ceci reposant sur le postulat selon lequel tout dire est plus conforme à la remémoration. D'autre part, il nous a été possible de faire un retour sur les questionnaires papiers qui contiennent, en plus des informations sur l'enquêté, d'autres renseignements relatifs au déroulement de l'entretien (l'enquêté était-il seul au moment de l'enquête, était-il réticent dans ses réponses ? *etc.*).

Les individus communs à Biographies et entourage et EHF

Le tirage de l'échantillon de l'Enquête Biographies et entourage a été effectué en prenant comme base de sondage l'enquête EHF. On pouvait ainsi facilement cibler un groupe d'âges particulier. L'Insee a permis cette opération, réalisée par J.M. Grosbras, qui offre la possibilité de comparer les réponses faites par un même enquêté dans le cadre de deux protocoles radicalement différents. Ce tirage, initialement prévu uniquement dans EHF, a été complété par un tirage dans les bulletins individuels du recensement, l'Enquête Famille n'étant pas représentative à un niveau fin. On dispose donc au total de 1306 individus communs, et de plusieurs thèmes communs.

Le travail de comparaison a débuté sur diverses pistes : professions des parents et du dernier ou de l'actuel conjoint, histoire des unions, histoire génésique. Est présenté ici le travail de comparaison au sujet des enfants des enquêtés. La tranche d'âges des personnes

enquêtées (50-70 ans) est d'autant plus intéressante qu'elles ont eu le temps de constituer leur descendance. Les différences que l'on peut s'attendre à observer entre Biographies et entourage et EHF sont dues au protocole d'enquête, après contrôle des rares changements intervenus entre les deux dates d'enquêtes.

Un premier constat : les corrections effectuées dans EHF étaient pertinentes

Sur les 1306 personnes du fichier apparié, huit avaient fait l'objet de corrections dans EHF, concernant les données relatives à leurs enfants. En effet, le questionnaire distingue les enfants biologiques de l'enquêté de ses beaux-enfants. Ainsi sont présentées dans un premier tableau les questions relatives aux enfants d'ego, et dans un second tableau les questions relatives aux enfants du conjoint d'ego, élevés par ego.

Rappel des questions préliminaires aux deux tableaux-enfants et des erreurs repérées et corrigées:

Les deux questions préliminaires aux tableaux enfants étaient les suivantes :

« Avez-vous eu des enfants ? (En comptant tous les enfants que vous avez eus ou adoptés y compris ceux qui seraient décédés) »

« En dehors de vos propres enfants, avez-vous élevé ou élevez-vous encore des enfants de votre conjoint (marié ou non) ou d'un ex-conjoint ? »

L'ambiguïté de certaines situations implique qu'un certain nombre d'enfants ont été, comme nous l'avions vu, déclarés deux fois. Il s'agissait, soit d'enfants qui ont été déclarés également comme enfants d'un autre conjoint que le conjoint actuel par confusion, soit de beaux-enfants considérés par l'enquêté comme ses propres enfants, figurant alors dans le tableau des enfants et dans celui des beaux-enfants. Les corrections effectuées ont replacé chaque enfant saisi deux fois, soit en enfant biologique, soit en bel-enfant (cf. tableau II.1.1 du Dossiers et Recherches N°93).

Sur les huit cas corrigés de l'échantillon, tous les enfants avaient été recodés comme enfants de la personne. Le retour à Biographies et entourage a confirmé que le diagnostic avait été correct et conforte une des explications à ce double remplissage, que nous comptions comme hypothèse auparavant : les personnes qui ont eu des enfants avec plusieurs conjoints remettent en « enfants élevés d'un conjoint » les enfants qu'elles ont eu avec un conjoint autre que le conjoint actuel. Par ailleurs, cette correction faite, les données

concernant les enfants des ces huit personnes dans EHF sont concordantes avec les données de Biographies et entourage.

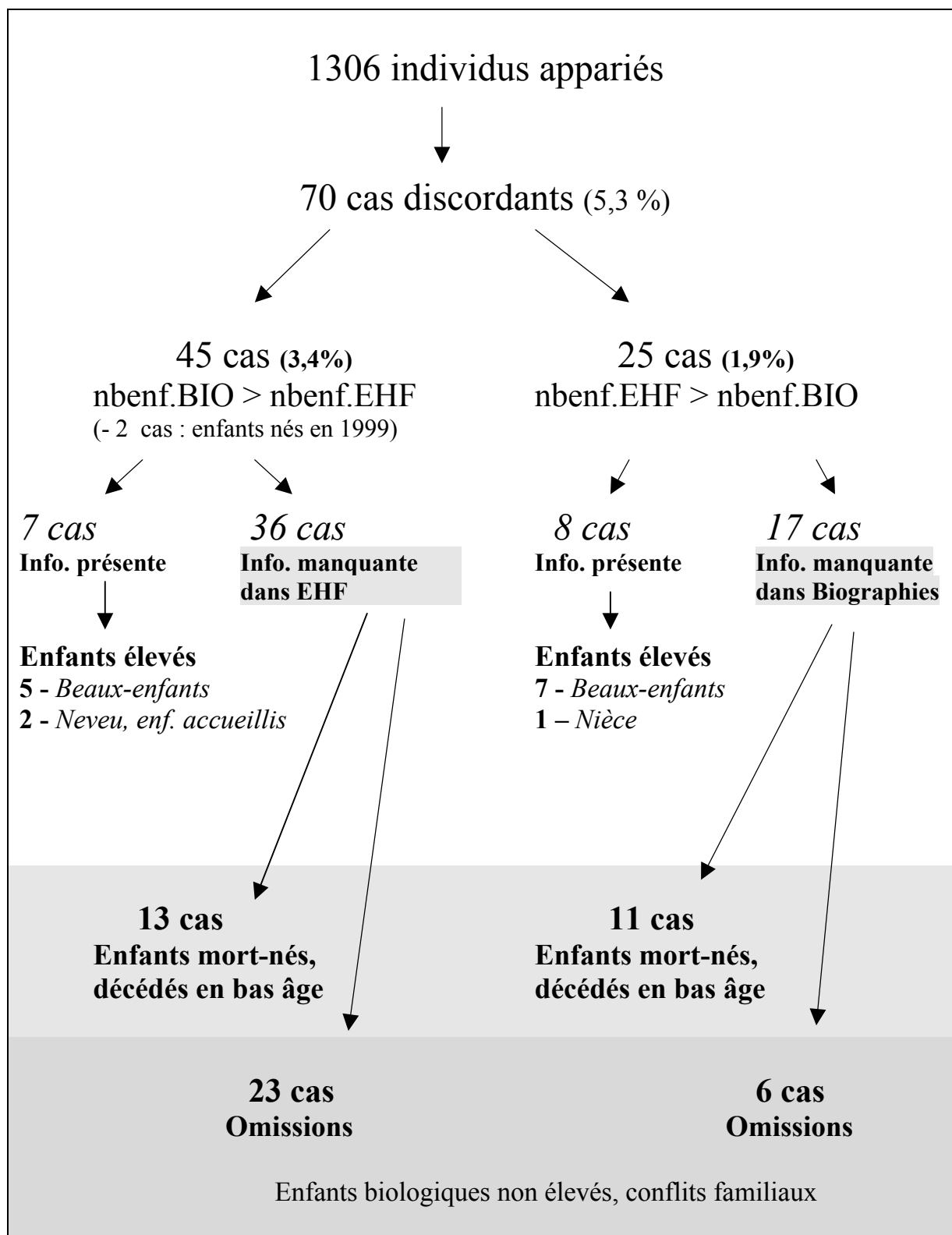
Les réponses discordantes relatives aux enfants biologiques :

Sur l'ensemble des individus appariés, on note 68 personnes pour lesquelles le nombre d'enfants ne correspond pas strictement dans les deux enquêtes. Cela représente environ 5% de l'échantillon. Certaines personnes cumulent plusieurs cas de discordance, ce qui représente au total 70 cas (figure 2.1).

Deux cas de figure sont possibles : l'information est présente dans les deux enquêtes mais a été enregistrée de manière différente ou bien l'information est manquante dans l'une des deux enquêtes.

Dans près de deux tiers des cas, on note un nombre d'enfants déclarés dans Biographies et entourage supérieur au nombre d'enfants déclarés dans EHF, et pour un tiers des cas la situation inverse.

**Les discordances portant sur les déclarations
relatives aux enfants biologiques d'Ego**



Les quarante-cinq cas pour lesquels le nombre d'enfants dans Biographies et Entourage est supérieur au nombre d'enfants dans EHF :

On peut ôter de ce chiffre les deux cas d'enfants nés en 1999, puisqu'il ne s'agit pas de discordance.

► Information présente : 7 cas

Les enfants élevés, qui sont considérés comme enfants biologiques : sept enquêtés dont quatre hommes. Il s'agit le plus souvent des enfants d'un conjoint, mais aussi de neveux, nièces, ou d'autres enfants élevés, qui ont co-résidé dans le ménage.

► Information manquante : 36 cas

Les enfants mort-nés ou décédés en bas âge : treize cas

Neuf de ces enquêtés sont des femmes. On peut penser qu'elles ont omis d'évoquer ces épisodes de leur vie dans EHF alors qu'elles en ont parlé dans Biographies du fait d'un questionnement plus précis.

Les omissions : vingt-trois cas (dont 11 hommes)

Les omissions concernent des personnes ayant eu des itinéraires un peu compliqués, des unions multiples, des conflits familiaux...

Ego est le parent biologique des enfants mais ne les a pas élevés : l'enquêté va déclarer les enfants dans le cadre d'un entretien parce-qu'on le questionne, alors que dans le questionnaire auto-administré il omet de déclarer les enfants (ou il mentionne tout au plus un nombre d'enfants mais ne donne pas d'autre information). Il s'agit le plus souvent de personnes aux unions multiples, qui ont eu des enfants mais n'ont jamais vécu avec eux et n'ont jamais vécu très en longtemps en couple.

Le deuxième cas de figure s'explique par les conflits familiaux : lorsque le parent a rompu le contact avec les enfants, lorsqu'il ne veut plus entendre parler d'ex-conjoints.

Les vingt-cinq cas pour lesquels le nombre d'enfants déclarés dans EHF est supérieur à celui déclaré dans Biographies et Entourage :

► Information présente : 8 cas

Les enfants élevés déclarés comme enfants biologiques (beaux-enfants le plus souvent)

► Information manquante : 17 cas

Les enfants mort-nés ou décédés en bas âge : onze cas

Il s'agit huit fois sur onze du témoignage d'une femme. La proximité avec l'enquêteur et l'explicitation possible des épisodes de la vie peut impliquer que l'enquêté développe plus les épisodes heureux et passe sous silence les épisodes les plus douloureux. Dans EHF, le travail de mémoire est peut-être plus neutre, puisqu'il s'agit juste de noter une date. D'autre part, mettant l'accent dans Biographies et entourage sur ce qui a vraiment façonné la vie de l'enquêté, on s'attache peut-être plus à expliciter les événements qui ont des implications sur le long terme, alors que la perte d'un enfant est un épisode certes malheureux, mais ponctuel dans la vie d'un individu.

Les omissions : six cas dont quatre hommes

Il s'agit d'enfants d'autres unions elles-mêmes complètement passées sous silence dans biographies et entourage : cela concerne 6 enquêtés dont quatre hommes. En général il s'agit des premiers enfants d'ego.

Conclusions :

Les redressements effectués sont corrects

Le problème principal que nous avons au départ était la saisie double de certains enfants, qui obligeait à un remaniement des données. Cet appariement a permis un contrôle du bien-fondé des redressements.

On déclare très bien ses enfants

On s'aperçoit à travers cette mise en regard des données que lorsqu'il s'agit de déclarer ses enfants les réponses sont très proches, quel que soit le protocole d'enquête. Pour comparer les nombres d'enfants, nous sommes partis du postulat selon lequel les enquêtés « n'inventent pas des enfants ». Ce qui nous a conduit à « reconstruire la vérité » et à mettre en évidence les logiques expliquant les discordances : la sous-déclaration des enfants mort-nés ou décédés en bas âge, la déclaration des beaux-enfants comme enfants, les omissions dues au protocole d'enquête et la part de non réponse qu'il y aura toujours, aussi faible soit-elle, dans une enquête. Si les deux enquêtes souffrent de la sous-déclaration des enfants mort-nés, Biographies et entourage fournit très peu d'omissions, étant plus complète et réalisée par des entretiens, et EHF un peu plus d'omissions, comme on pouvait s'y attendre. Mais la différence de qualité entre ces deux enquêtes est très faible au total.

Finally, what is the missing information in EHF ?

The main bias concerns stillborn or deceased children. Under-reporting of stillborn children is a known phenomenon, and it is interesting to have been able to estimate it by cross-referencing these two sources. It is certainly a minimal estimate, as we do not know in which measure the information is missing in the two surveys.

As for omissions, they are relatively infrequent and explainable, either because the parent is the « biological parent », or because of conflictual relationships or a refusal to evoke the past. These omissions concern about one percent of the total children. However, there is no complete loss of information since a child not declared by the father (for example fathers who have never lived with their children), can be declared by the mother. On the other hand, one can think that these children not declared are perhaps also not recognized. Conversely, it happens, in the case where biological children are declared as children of the person, that some of them have been recognized or adopted by the respondent: one can therefore discuss the idea that they have been « placed » as children of the person.

The concept of child, of biological child, and of multi-parentality

All these results highlight the weak ambiguity of the notion of child. It is strongly linked in France to biological filiation. Situations of parenthood without biological filiation are quite restricted (adopted children, recognized children, children raised since childhood) and tend to broaden parenthood. In the same way, biological parenthood does not make parenthood. This work of comparison shows that the notion of child is quite clear, and the rationality of the respondents. The fact of declaring a biological child as child is not in itself an error, especially if the parent has recognized or adopted it.

2 – Disponibilité et pondérations du fichier

Tous les résultats présentés dans les trois rapports sont issus de versions provisoires du fichier EHF. Une version définitive est disponible depuis le mois de juillet, avec en particulier de nouvelles pondérations. Est reproduite ci-dessous une note explicative relative aux pondérations, disponible dans la dernière version de la documentation du fichier EHF.

2.1 – À quoi servent les pondérations ?

Les 380 000 bulletins de l'enquête EHF constituent un échantillon représentatif de la population de la France métropolitaine âgée de 18 ans ou davantage au 1er janvier 1999. La pondération du fichier répond à trois objectifs.

Premier objectif, la pondération restitue la structure de la population, en tenant compte du plan de sondage. Par exemple, on a constitué l'échantillon en sélectionnant une femme sur 80 et un homme sur 130. C'est ainsi que l'échantillon de l'enquête comprend 145 000 hommes et 235 000 femmes âgés de 18 ans ou plus, soit 62% de femmes, alors que la population des adultes comprend 52% de femmes. Le poids des hommes (0,13) est plus important que celui des femmes (0,08) de manière à ce que l'échantillon pondéré respecte la structure par sexe de la population générale. De même, les régions les moins peuplées sont surreprésentées dans l'échantillon pour améliorer la précision des résultats régionaux, et les pondérations y sont plus faibles pour rétablir leur part dans la population.

Deuxième objectif, les effectifs pondérés représentent la population dans son ensemble. L'enquête ayant été réalisée auprès d'un individu sur 100, chaque individu est affecté d'un poids, de l'ordre de 100, qui permet de fournir des estimations pondérées correspondant à l'effectif total concerné dans la population. Ainsi les 380 000 personnes interrogées représentent les 44 millions d'adultes vivant en France métropolitaine. Pour des raisons de lisibilité, les poids sont divisés par 1000 de manière à ce que les estimations soient exprimées en milliers, et l'effectif pondéré de l'enquête s'élève finalement à 44 067 milliers de personnes.

Troisième objectif, la pondération corrige les biais dus à la non participation à l'enquête. Par exemple les personnes vivant en couple ont plus souvent rempli un bulletin de l'enquête EHF que les personnes ne vivant pas en couple. Il en résulte une surreprésentation dans l'échantillon des personnes en couple. On peut corriger cette distorsion en attribuant un poids plus important aux personnes non en couple qui ont participé à l'enquête. C'est la post-stratification de l'échantillon.

Il est donc tout à fait indispensable d'utiliser les pondérations pour que les résultats obtenus à partir de l'échantillon interrogé s'appliquent à la population dans son ensemble. La pondération poids^{m7} est la plus utile, et elle est présente dans le fichier sous le nom de POIDS. Le fichier comprend plusieurs autres pondérations, qui correspondent à différentes post-stratifications.

2.2 – Le plan de sondage

La collecte du recensement est organisée à partir d'un découpage spécifique du territoire. Au sein de chaque région sont constituées des zones de délégués (3 600 au total), elles-mêmes divisées en secteurs d'agents recenseurs (110 000 au total, soit 30 en moyenne par zone de délégué).

L'échantillon de l'enquête EHF correspond à un secteur d'agent recenseur sur cinquante. Les agents recenseurs participant à l'enquête EHF distribuaient, en plus des bulletins du recensement, des bulletins de l'enquête, soit à tous les hommes, soit à toutes les femmes de leur secteur. Pour ne pas compliquer la collecte du recensement, il a été décidé de limiter la collecte de l'enquête EHF à un délégué sur dix. Puis, dans chaque zone de délégué sélectionnée, on a choisi un agent recenseur sur cinq, soit huit sur quarante, dont cinq pour distribuer des bulletins aux femmes et trois pour en distribuer aux hommes.

Ce schéma général a été affiné dans trois directions : l'échantillon a été augmenté dans les régions les moins peuplées, des extensions ont été constituées pour l'étude des langues parlées en France, et l'enquête a été couplée avec l'enquête Vie quotidienne et santé.

Pour rendre possible des études démographiques à l'échelle de la région, il a été décidé de fixer des tailles minimales pour l'échantillon de chaque région. Pour cela on a sélectionné une proportion plus importante de zones de délégué dans les régions les moins peuplées. De plus, on a augmenté le nombre de zones choisies dans les régions périphériques où les langues régionales sont ou étaient parlées en famille : Nord, Bretagne, Pays basque, Pays catalan, Corse, Alsace. Enfin, l'enquête était couplée avec l'enquête Vie quotidienne et santé (VQS), enquête qui a servi de filtre à l'enquête Handicaps-incapacités-dépendance. Dans certaines régions et certains départements, l'enquête VQS a été réalisée auprès d'un échantillon plus important pour permettre des études locales. Dans ce cas, on a multiplié les zones de délégué et, dans chacune de ces zones, on a sélectionné moins d'agents recenseurs pour l'enquête EHF. Par ailleurs un échantillon national de collectivités (cités universitaires et foyers de jeunes travailleurs) a été inclus dans l'échantillon, tandis qu'une opération spécifique était réalisée pour les hommes détenus.

2.3 – Les non-réponses

Le champ de l'enquête EHF en ménages ordinaires est composé de tous les hommes adultes recensés dans les secteurs d'agent recenseur désignés pour l'enquête auprès des hommes et de toutes les femmes des secteurs « femmes ». Parmi les 465 000 personnes éligibles pour l'enquête, 369 000, soit 79,4%, ont rempli également un bulletin de l'enquête EHF. Le taux d'échec de l'enquête s'élève donc à 20,6%, ce qui est comparable au taux d'échec d'autres enquêtes non obligatoires mais très supérieur au taux d'échec de l'enquête Famille de 1990 (3%). Les 96 000 personnes qui ont rempli leur bulletin de recensement mais n'ont pas souhaité remplir également le bulletin de l'enquête EHF ne se répartissent pas « au hasard » dans la population. Ainsi les hommes ont moins souvent répondu que les femmes (78,8% contre 79,8%). Comme dans toutes les enquêtes, les personnes de moins de 25 ans ou de plus de 85 ans, les immigrés arrivés en France après 1975, les personnes vivant seules, les célibataires, ont moins participé que les autres. Les personnes n'ayant pas complètement rempli leur bulletin de recensement (pas d'activité professionnelle ni de diplôme déclaré) ont également plus souvent refusé l'enquête.

Une modélisation de la probabilité de non-réponse à l'enquête selon différents critères confirme la plus faible participation des personnes âgées, des célibataires ne vivant pas en couple, des personnes nées à l'étranger, n'ayant pas déclaré de diplôme. Par contre, elle montre que les jeunes ont autant participé que les autres, « toutes choses égales par ailleurs ». Elle confirme également des disparités importantes d'une région à l'autre, de 16% à 31%. Ces probabilités différentes de participer à l'enquête sont prises en compte dans les pondérations finales, grâce aux méthodes de post-stratification.

2.4 – La post-stratification

La population recensée dans les secteurs d'agents recenseurs sélectionnés pour l'enquête EHF peut différer de la population générale, en particulier à l'échelle d'une région, puisque l'échantillon est concentré dans certaines zones géographiques. De plus, la probabilité de participer à l'enquête n'est pas la même pour tous, ce qui « déforme » la structure de l'échantillon. Pour corriger ces biais, on opère une post-stratification selon deux types de critères : ceux qui définissent les secteurs choisis, et ceux pour lesquels on a repéré une variation du taux de participation. Les variables utilisées pour la post-stratification sont donc la tranche d'unité urbaine, le sexe, l'âge, la situation de couple, le niveau de diplôme, le pays de naissance et la date d'arrivée en France métropolitaine.

La post-stratification consiste à imposer à la structure de l'échantillon de respecter les répartitions de l'ensemble de la population recensée selon chacune de ces variables. La procédure CALMAR mise au point par l'Insee a été utilisée.

2.5 – Les pondérations finales

Le fichier contient un certain nombre de variables de pondération. Toutes sont exprimées en milliers de personnes. Les variables POIDSM4 à POIDSM7 correspondent à quatre post-stratifications, nationales ou régionales, par sexe et âge ou selon un plus grand nombre de variables.

POIDSM1 représente l'inverse du taux de sondage estimé, divisé par 1000.

POIDSM2 résulte de la troncature de POIDSM1. Pour chaque région, les poids peuvent différer assez fortement en cas d'extension VQS ou Langues. De plus, la probabilité d'inclusion est proportionnelle à la population de la zone de délégué. Or des erreurs dans l'appariement entre secteurs d'agent recenseur et zone de délégué ont pu être commises, qui conduisent à surestimer leur taille (fusion de zones) ou à la sous-estimer (erreur de saisie). Le poids étant inversement proportionnel à la taille de la zone de délégué, ces erreurs peuvent entraîner des écarts de poids importants. On a donc limité la variation de poids à un rapport de 1 à 10, entre trois fois moins et trois fois plus que le mode de distribution de Poidsm1 dans chaque région, en tenant compte des extensions VQS ou langues. On a ramené les poids extrêmes à la moyenne des poids non-extrêmes de la région. 3,7% des individus EHF ont ainsi vu leur pondération modifiée.

POIDSM3 est la pondération provisoire utilisée dans les premières analyses, résultat d'une post-stratification par sexe et âge exact (croisés) à partir du fichier pondéré par POIDSM1. Les personnes vivant en ménage collectif ont un poids nul.

POIDSM4 reprend cette logique de pondération par sexe et âge exact (croisés), mais à partir de la pondération tronquée POIDSM2. Cette pondération permet d'obtenir une pyramide des âges conforme à la pyramide du recensement, y compris les personnes vivant en foyers de jeunes travailleurs et en cité universitaire.

POIDSM5 résulte d'une post-stratification sur la répartition nationale selon les 7 variables suivantes : sexe et groupe d'âges, situation de couple, diplôme, pays de naissance, date d'arrivée en France, région de résidence, tranche de l'unité urbaine. Les personnes vivant en ménage collectif ont un poids. Le détail des variables et des modalités utilisées figure dans la note jointe.

POIDSM6 correspond à une post-stratification ajustée, pour chaque région, sur la répartition régionale par sexe et âge détaillé. Les personnes vivant en ménage collectif n'ont pas de poids.

POIDSM7, enfin, est obtenu après une post-stratification, pour chaque région, sur la répartition régionale selon les 6 variables suivantes : sexe et groupe d'âges, situation de couple, diplôme, pays de naissance, date d'arrivée en France, tranche d'unité urbaine. Les personnes vivant en ménage collectif n'ont pas de poids.

Les post-stratifications conduisent parfois à des poids très dispersés, en particulier quand une tranche d'unité urbaine n'est pas ou peu représentée. D'une part, cette variable a été parfois regroupée et, d'autre part, les poids POIDSM4, POIDSM5, POIDSM6 et POIDSM7 ont été tronqués : les poids finals ont été ramenés à un rapport de 1 à 10 pour chaque combinaison sexe*région. Quand ce rapport n'était pas dépassé, les poids ont été tronqués aux 1^{er} et 99^e percentiles.

Tous ces poids sont disponibles dans le fichier.

2.6 – Quelle pondération utiliser ?

La variable POIDSM7, obtenue par post-stratification, pour chaque région, selon les principales variables repérées comme associées à la non-réponse ou aux aléas de tirage, est la plus générale. La variable POIDS qui est égale à POIDSM7 pour les adultes vivant en ménage ordinaire et POIDSM5 pour les autres est conseillée par défaut.

Les poids POIDSM7 sont cependant très dispersés : de 1 à 39, coefficient de variation CV=0,69. Il n'y a pas de pondération idéale, et le choix de la pondération dépend de plusieurs critères : la post-stratification diminue les biais mais augmente la variance. De plus, elle repose sur des hypothèses sur les comportements. Le tableau ci-dessous présente quelques statistiques sur ces pondérations, pour les individus vivant en ménage ordinaire.

Variable	N	Moyenne	Minimum	Maximum	Sigma	Max/min	CV
POIDSM1	378374	0,087	0,018	0,526	0,053	30	0,609
POIDSM2	378374	0,084	0,018	0,322	0,046	18	0,546
POIDSM3	378374	0,116	0,018	0,782	0,071	44	0,616
POIDSM4	378374	0,116	0,023	0,345	0,063	15	0,549
POIDSM4V	378374	0,116	0,023	0,519	0,064	23	0,554
POIDSM5	378374	0,116	0,029	0,407	0,069	14	0,598
POIDSM5V	378374	0,116	0,020	1,293	0,072	64	0,619
POIDSM6	378374	0,116	0,026	0,594	0,068	23	0,584
POIDSM6V	378374	0,116	0,022	1,354	0,068	60	0,586
POIDSM7	378374	0,116	0,018	0,697	0,079	39	0,686
POIDSM7V	378374	0,116	0,010	1,506	0,081	151	0,704

sigma : écart-type
CV : coefficient de variation (sigma/moyenne)

Pour les études à l'échelle nationale, on peut préférer la pondération POIDSM5, moins dispersée (CV=0,60), sous l'hypothèse que les distorsions sont de même nature dans toutes les régions.

L'utilisation des pondérations non tronquées n'est pas recommandée. Elles ne sont dans le fichier qu'à titre informatif, pour permettre à l'utilisateur de mesurer la troncature.

Enfin, la post-stratification repose sur l'hypothèse que, dans chaque groupe, les personnes ayant répondu sont similaires à celles qui n'ont pas répondu. Si l'on ne croit pas à cette hypothèse, on pourra utiliser les pondérations calées sur la pyramide des âges, soit pour chaque région séparément (POIDSM6, CV=0,58), soit pour l'ensemble (POIDSM4, CV=0,55).

2.7 – Les pondérations des enfants

Il y a deux manières d'utiliser le fichier des enfants. La première consiste à « remonter » les informations au niveau de l'individu « parent ». Par exemple on cherche à classer les pères selon le nombre d'enfants et l'âge du dernier enfant. Dans ce cas, l'« individu statistique » reste l'adulte et on utilise sa pondération d'adulte.

Le fichier des enfants peut également servir à produire des statistiques sur les enfants. On peut, par exemple, voir comment se répartissent les enfants selon la taille de la famille. Dans ce cas on doit disposer d'un échantillon représentatif des enfants dûment pondéré. Plusieurs échantillons sont possibles :

L'échantillon des enfants des femmes. Le champ est celui des enfants dont la mère est vivante et réside en France métropolitaine.

L'échantillon des enfants des hommes ou des femmes. Le champ est alors celui des enfants dont un des parents est vivant et réside en France. La probabilité d'inclusion est alors proportionnelle au nombre de parents vivants en France métropolitaine.

L'échantillon des enfants recensés dans le ménage. Le champ est alors celui des enfants vivant avec au moins un adulte. La probabilité d'inclusion est alors proportionnelle au nombre d'adultes du ménage.

L'échantillon a. est, de beaucoup, le plus utile. Les enfants orphelins de mère et ceux dont la mère réside hors de France manquent, mais ils sont très peu nombreux, et les autres ont la même probabilité d'apparition que leur mère. On peut donc utiliser ce sous-ensemble des enfants comme un échantillon représentatif des enfants vivant en France, les informations sur leur mère étant issues du fichier des adultes. Dans le fichier, seuls les enfants des femmes ont une valeur pour la variable POIDS, la pondération de leur mère. Les beaux-enfants et les enfants des hommes n'ont pas de poids par défaut. On peut restreindre cet échantillon, par exemple, aux enfants résidant en France métropolitaine.

L'échantillon b. est en théorie plus complet, mais beaucoup plus délicat à utiliser. Les enfants dont les deux parents sont éligibles pour l'enquête ont une probabilité double d'apparition, et leur poids doit donc être celui de leur parent adulte interrogé, divisé par deux si les deux parents sont éligibles. Mais le nombre de parents éligibles n'est pas toujours connu : la fécondité apparente des hommes est plus faible que celle des femmes, et il semble que 6% des enfants ne sont pas déclarés par leur père biologique ; en cas de séparation des parents, l'information sur le lieu de résidence de l'autre parent ne figure pas dans l'enquête.

Enfin, l'échantillon c. est encore plus délicat, car dans un ménage où un adulte EHF a été interrogé les enfants dont aucun parent n'a rempli un bulletin n'apparaissent pas explicitement. La probabilité d'apparition des enfants est proportionnelle au nombre d'adultes éligibles dans le ménage.

Pour ce type d'utilisation, les beaux-enfants – repérés dans le fichier des enfants par « ENR= '6' » – ne doivent pas être considérés : en principe, tout bel-enfant est d'abord un enfant et le même enfant peut apparaître comme bel-enfant de plusieurs adultes. Pour certaines études particulières, on peut étudier l'ensemble des enfants qui vivent avec un parent et un beau-parent au moment de l'enquête. On leur donne alors le poids de leur parent, homme ou femme.

Bibliographie

Références citées dans le texte

DAGUET Fabienne « L'évolution de la fécondité des générations nées de 1917 à 1949 : analyse par rang de naissance et niveau de diplôme », *Population*, N°6, novembre-décembre 2000, pp.1021-1034

TOULEMON Laurent and LAPIERRE-ADAMCYK Evelyne, "Demographic Patterns of Motherhood and Fatherhood in France", *Fertility and the male life-cycle in the era of fertility decline*, Oxford University Press, 2000, pp.293-330

MAZUY Magali, TOULEMON Laurent, " Etudes de l'histoire familiale. Premiers résultats de l'enquête en ménages ", INED, coll. *Dossiers et recherches*, N°93, février 2001

Bibliographie de l'enquête Famille 1999

*Bibliographie disponible sur le site de l'Enquête Famille : <http://www-ehf.ined.fr> . On y trouvera également un exemplaire du questionnaire de l'enquête, et l'article paru dans le *Courrier des Statistiques*, présentant l'enquête.*

- CASSAN Francine, HÉRAN François, TOULEMON Laurent, " Etude de l'histoire familiale : l'édition 1999 de l'enquête Famille ", *Courrier des Statistiques*, n°93, mars 2000.
- CASSAN Francine, HÉRAN François, TOULEMON Laurent, " Study of family history : France's 1999 Family Survey, *Courrier des Statistiques*, n°6, 2000
- CASSAN Francine (coord.), *L'histoire familiale des hommes détenus*, Collection *Synthèses*, N°59, INSEE, janvier 2002.
- CASSAN Francine, KENSEY Annie, TOULEMON Laurent, " L'histoire familiale des hommes détenus ", *Insee Première*, n°706, avril 2000
- CASSAN Francine, MAZUY Magali, TOULEMON Laurent, " Douze millions et demi de grands-parents ", *Insee première* N°776, mai 2001
- CASSAN Francine, MAZUY Magali, CLANCHÉ François, " Refaire sa vie de couple est plus fréquent pour les hommes ", *Insee première* N° 797, juillet 2001
- CASSAN Francine, MARY-PORTAS France-Line, "Précocité et instabilité familiale des hommes détenus ", *Insee Première*, n° 828, février 2002
- CLANCHÉ François, "Langues régionale, langues étrangères de l'héritage à la pratique", *Insee Première* n°830, février 2002
- HÉRAN François, FILHON Alexandra, DEPRESZ Christine, " La dynamique des langues en France au fil du XXème siècle ", *Population et Sociétés*, n°376, février 2002
- MAZUY Magali, TOULEMON Laurent, " Etudes de l'histoire familiale. Premiers résultats de l'enquête en ménages ", INED, coll. *Dossiers et recherches*, N°93, février 2001
- TOULEMON Laurent, " Combien d'enfants, combien de frères et soeurs depuis cent ans ? ", *Population et Sociétés*, n°374, décembre 2001
- TOULEMON Laurent, MAZUY Magali, " Les naissances sont retardées mais la fécondité est stable ", *Population*, N°4, juillet-août 2001, pp.611-644
- TOULEMON Laurent, MAZUY Magali, " Cinq projections de fécondité fondées sur une hypothèse de stabilité des comportements ", *Population* N°4, juillet - août 2001, pp.647-656

Annexe : documentation du fichier

On dispose depuis le mois de juillet 2002 de la version définitive du fichier de l'enquête Famille avec sa documentation. Cette nouvelle version intègre l'ensemble des travaux d'enrichissement du fichier, effectués depuis 1999, ainsi que les nouvelles pondérations.

Contacts pour l'exploitation de l'enquête :

Toute personne souhaitant travailler sur l'enquête EHF est priée de contacter une des personnes ci-dessous. Dans tous les cas, merci de faire une copie à Cécile Lefèvre.

Coordination générale : Cécile Lefèvre. lefevre@ined.fr

Responsable de l'enquête à l'Insee : François Clanché. francois.clanche@insee.fr

Comparaisons internationales : Laurent Toulemon. toulemon@ined.fr

Adresses postales

Ined, 133, Boulevard DAVOUT

75 980 PARIS Cedex 20

Insee, division des enquêtes et études démographiques,

Timbre F170,

18, Boulevard Adolphe PINARD

75 675 PARIS Cedex 14