

Qui gravit l'échelle des revenus par rapport à ses parents ?

Une analyse de la mobilité intergénérationnelle à partir de l'Échantillon Démographique Permanent

Documents de travail

N° 2022 - F04 - Mai 2022





Institut national de la statistique et des études économiques

2022/04

**Qui gravit l'échelle des revenus par rapport à ses parents ?
Une analyse de la mobilité intergénérationnelle à partir de
l'Échantillon Démographique Permanent**

Hicham ABBAS *, Michaël SICSIK **

Mai 2022

Direction des Statistiques Démographiques et Sociales - Timbre F001
88, avenue Verdier - CS 70058 - 92541 MONTROUGE CEDEX - France
Tél. : 33 (1) 87 69 62 82 - E-mail : dg75-f001@insee.fr - Site Web Insee : <http://www.insee.fr>

Ces documents de travail ne reflètent pas la position de l'Insee et n'engagent que leurs auteurs.

Working papers do not reflect the position of INSEE but only their author's views.

* Insee au moment de l'étude, et Crest.

** Insee, et CRED/TEPP.

Les auteurs remercient Aliocha Accardo, Valérie Albouy, Christel Colin, Julie Labarthe, Raphaël Lardeux, Isabelle Robert-Bobée et Alain Trannoy pour leurs commentaires sur une version précédente de cet article. Nous remercions également Clément Dherbécourt, Lino Galiana, Bertrand Garbinti, Sébastien Grobon, Gustave Kenedi, Stéphane Legleye, Louis Sirugue et les participants au séminaire de la DSDS d'octobre 2021 pour des discussions stimulantes sur ce travail. Nous remercions enfin Céline Dennevault et Pierre Lamarche pour leur aide sur les données de l'EDP et de Fidéli, ainsi qu'Erwan Pouliquen pour une extraction de données de l'enquête logement.

Qui gravit l'échelle des revenus par rapport à ses parents ? Une analyse de la mobilité intergénérationnelle à partir de l'Échantillon Démographique Permanent

Résumé :

Nous étudions la mobilité intergénérationnelle en nous intéressant au lien entre les revenus d'un échantillon de jeunes adultes et les revenus de leurs parents, à partir des données fiscales de l'Échantillon Démographique Permanent. Nous mesurons la position relative des parents dans la distribution des revenus puis la comparons pour la première fois directement avec la position relative de leurs enfants autour de 28 ans.

La relation entre la position des parents et celle des enfants à 28 ans est positive et linéaire, avec une pente de 0,24. Selon cet indicateur, la mobilité positionnelle serait plus élevée en France qu'aux États-Unis, mais plus faible que dans les pays scandinaves. En séparant les parents et les enfants en cinquièmes de population, les enfants des familles aux plus hauts revenus ont trois fois plus de chances d'être classés parmi les 20 % les plus aisés que ceux issus des familles des 20 % les plus pauvres : les inégalités se transmettent donc partiellement à la génération suivante. Cependant, pour un même niveau de revenu des parents, les revenus des enfants varient fortement à l'âge de 28 ans. En 2018, le taux de mobilité ascendante du plus bas au plus haut cinquième est de 12 %. Toutes choses égales par ailleurs, la mobilité ascendante est d'autant plus forte que les parents ont des revenus du capital élevés, sont diplômés du supérieur, sont immigrés, sont mobiles géographiquement, ou résident en Île-de-France à la majorité de leurs enfants. À l'inverse, être une femme, avoir vécu dans une famille monoparentale, avoir des parents ouvriers ou employés, ou avoir vécu dans les Hauts-de-France sont des facteurs qui influent négativement sur la mobilité ascendante.

Mots-clés : mobilité intergénérationnelle, revenus, inégalités, variations géographiques.

Who Moves Up the Income Ladder Relative to their Parents ? An Analysis of Intergenerational Income Mobility in France

Abstract:

We study intergenerational income mobility in France using tax data from a large panel. We measure the rank of parents in their income distribution and compare it directly for the first time with the rank of their children around age 28.

The rank-rank correlation is 0.24. According to this indicator, intergenerational mobility is higher in France than in the United States, but lower than in the Scandinavian countries. Children born to parents in the bottom 20% of their income distribution are three times more likely to be among the top 20 % as adults than those from poorest families, and so inequality is partially transmitted to the next generation. However, for the same level of parental income, children's incomes vary greatly when they are young adults. In 2018, the rate of upward mobility from the bottom 20 % to the top 20 % is 12 %. Upward mobility is stronger the higher the parents' capital income and education, when parents are immigrants, are more geographically mobile, or live in the Île-de-France region when their children become adults. Conversely, being a woman, having lived in a single-parent family, having parents who are manual workers, or having lived in the Hauts-de-France region are all factors that have a negative impact on upward mobility.

Keywords: intergenerational mobility, income, inequality, spatial variations, France

Classification JEL : J62, J61, D31, H0, R1.

Table des matières

1 Introduction.....	4
2 Données, traitements et champ.....	9
2.1 Les données.....	9
2.2 Constitution de la base d'étude et traitements.....	10
2.3 Le champ de l'étude.....	12
2.4 Définition des variables utilisées dans l'analyse.....	14
3 Statistiques descriptives.....	17
4 Méthodes.....	19
4.1 Indicateurs de mobilité.....	19
4.2 Discussion sur les biais de cycle de vie et d'atténuation.....	20
4.3 Analyse « toutes choses égales par ailleurs ».....	21
5 Résultats.....	23
5.1 La mobilité au niveau national.....	23
5.1.1 Résultats principaux.....	23
5.1.2 Tests de robustesse.....	29
5.2 La mobilité selon les caractéristiques individuelles ou celles des ménages.....	32
5.3 La mobilité au niveau des territoires.....	35
5.3.1 Régions et départements.....	35
5.3.2 Densité et catégories de communes.....	38
5.4 Analyse de la mobilité « toutes choses égales par ailleurs ».....	39
5.4.1 Analyse de la mobilité ascendante.....	39
5.4.2 Tests de robustesse.....	41
5.4.3 Résultats sur la mobilité descendante.....	44
5.5 Quel rôle de l'environnement familial pour expliquer le rang des enfants ?.....	45
5.5.1 Méthode par régression.....	46
5.5.2 Méthode par utilisation des fratries.....	48
Bibliographie.....	50
Annexes.....	53
Annexe A. Traitement des individus multi-identifiés et des enfants en garde alternée.....	53
Annexe B. Traitement des effectifs changeants selon l'année fiscale.....	55
Annexe C. Traitement des données du recensement de la population de 1999 et des enquêtes annuelles de recensement.....	57
Annexe D. Statistiques descriptives complémentaires et effets de sélection.....	61
Annexe E : Résultats complémentaires.....	65
Annexe F : Calcul mathématique de l'espérance de la différence de rang.....	86

1 Introduction

Diminuer la reproduction des inégalités entre générations est devenu un objectif important des politiques publiques ainsi qu'un consensus de politique économique. La mobilité intergénérationnelle des individus permet une croissance plus inclusive où chacun a les mêmes chances de prospérer (OCDE, 2018), et peut stimuler l'innovation (Aghion et al., 2019). En effet, une faible mobilité dans le bas de la distribution des revenus conduit à la perte de nombreux talents et entrepreneurs potentiels (Bell et al., 2019) et, selon Jaravel (2021), la France pourrait gagner 0,2 point de PIB supplémentaire par an grâce à une meilleure inclusion sociale. Une absence de mobilité en haut de la distribution des revenus engendre également des coûts d'efficience élevés via des rentes persistantes (OCDE, 2018). Diminuer la reproduction des inégalités entre générations permet de tendre vers une égalité des opportunités (Roemer et Trannoy, 2015). Pourtant, s'il y a consensus sur l'objectif d'encourager la mobilité entre les générations, des divergences existent sur le diagnostic de la mobilité relative¹ en termes de revenus en France (Dherbécourt, 2020) : selon les publications, la France apparaît comme un pays à niveau de mobilité intermédiaire (Lefranc et Trannoy, 2005), faible (OCDE, 2018) ou élevée (Alesina et al. 2018). La mesure de la mobilité intergénérationnelle en termes de revenus est peu documentée en France, à l'inverse de la mobilité sociale en termes de catégorie socioprofessionnelle (ou encore de la mobilité en termes d'éducation), du fait de l'absence jusqu'il y a peu de bases de données permettant de lier directement les revenus des parents à ceux de leurs enfants (contrairement aux États-Unis et aux pays scandinaves où cela est possible depuis plus longtemps et sur plus longue période).

Dans cet article, nous étudions la mobilité intergénérationnelle en reliant pour la première fois en France les revenus des enfants (quand ils sont jeunes adultes) à ceux de leurs parents, avec un très grand échantillon de paires enfant-parents, représentatif de l'ensemble de la population. Nous mobilisons pour cela l'Échantillon Démographique Permanent (EDP), en particulier sa partie panel de données fiscales et sociales depuis 2010, qui permet de suivre environ 2,5 millions de personnes chaque année. À partir de l'EDP, nous pouvons comparer directement les revenus d'un échantillon de jeunes adultes, en 2018, à ceux de leurs parents, en 2010, lorsque parents et enfants vivaient dans le même foyer fiscal, huit ans plus tôt. Les revenus des parents et des enfants n'étant pas mesurés au même âge, c'est la position relative des parents en 2010 (dans la distribution des revenus des parents) qui est comparée à celle de leur enfant (dans la distribution des revenus des enfants), âgé de 28 ans environ en 2018, à un âge où presque tous les « enfants » sont sur le marché du travail (en effet, à 28 ans, seul 1 % des individus est en études initiales, Bernard, 2021). La pente, dite « rang-rang », de la relation entre ces deux variables de position permet d'obtenir une mesure de la mobilité positionnelle relative des jeunes adultes, à l'instar de Chetty et al. (2014). Notre analyse principale porte sur les enfants âgés de 28 ans, ce qui représente environ 14 000 paires enfants-parents (après toutes les restrictions de champ). Observer les revenus des enfants à l'âge de 28 ans conduit à mesurer une pente légèrement plus faible qu'à 35 ans². L'intérêt d'étudier la mobilité intergénérationnelle en comparant les positions relatives des parents et des enfants dans leurs distributions respectives des revenus réside en ce que cette stratégie ne rend pas indispensable l'observation des revenus des parents et des enfants au même âge. De plus, cette approche est beaucoup plus robuste que la

1 Nous étudions ici la mobilité relative qui a gagné en importance et en pertinence dans le débat public par rapport à la mobilité absolue des revenus. Tandis que la première mesure le lien entre les revenus des enfants et des parents ou leur classement respectif, la deuxième s'intéresse à la question de savoir quelle est la part des enfants qui gagnent plus que leurs parents. Cette dernière fait l'objet de moins de débat, et est très directement liée à l'évolution de la croissance et des inégalités dans le temps (voir Berman, 2021).

2 Dans la revue de littérature de cette introduction, nous montrons que l'écart entre la pente à 28 ans et la pente à 35 ans est d'environ 10 %.

comparaison traditionnelle des logarithmes des revenus³, comme l'ont montré les récentes études sur la mobilité (voir notamment Chetty et al., 2014 et Nybom et Stuhler, 2017). Une mesure alternative de la mobilité consiste à construire des matrices de transition entre cinquièmes de la population pour mesurer les probabilités de mobilité ascendante, de mobilité descendante, de plancher et de plafond collant. Nous analysons ensuite la mobilité grâce à de nombreuses variables socio-démographiques définies au niveau de l'individu (disponibles grâce à l'appariement de l'EDP avec le recensement de la population) et géographiques (la région, le département, ou d'autres zonages géographiques). Cette analyse permet notamment de mesurer la mobilité en fonction du diplôme ou du statut migratoire des parents. L'EDP regroupant également les données fiscales, nous pouvons étudier la mobilité intergénérationnelle sous l'angle des revenus du capital, ce qui n'a encore jamais été fait à notre connaissance. Pour produire des estimations robustes selon l'ensemble des caractéristiques évoquées précédemment, nous étendons dans cette partie notre population aux « enfants » âgés de 26 à 29 ans, soit un échantillon de 60 000 paires enfants-parents. Ensuite, les analyses portant sur les territoires et les caractéristiques individuelles sont principalement commentées à partir des indicateurs de mobilité ascendante, qui sont moins sensibles au cycle de vie et se stabilisent autour de 26-27 ans⁴, ce qui devrait réduire l'éventualité d'un biais. Nous mettons en œuvre des régressions de Poisson modifiées pour expliquer la mobilité ascendante et la mobilité descendante dans le cadre d'une analyse « toutes choses égales par ailleurs ». Dans une dernière partie, nous essayons de mesurer la part de la variation des revenus des enfants liée à la famille. Pour cela, nous identifions les fratries dans notre échantillon et comparons les revenus des frères et sœurs. Nous déterminons également les poids respectifs des revenus des parents et des autres variables socio-démographiques pour expliquer le rang des enfants.

Résultats. Une première mesure de la mobilité est donnée en régressant le rang moyen des enfants (au sein de leur distribution) selon le centième de revenu des parents. Le coefficient de régression entre les deux variables de classement, qui mesure la persistance des revenus, est de 0,24. Les inégalités se transmettent en partie entre générations : mieux les parents sont classés dans l'échelle des revenus, mieux le sont en moyenne leurs enfants par rapport aux jeunes adultes de leur génération. Ce résultat peut s'expliquer, par exemple, par une probabilité plus forte des jeunes des milieux favorisés de suivre des études supérieures comme l'ont montré récemment Chetty et al. (2020) aux États-Unis et Bonneau et Grobon (2022) en France. La mobilité positionnelle serait significativement plus élevée qu'aux États-Unis (coefficient de 0,32 entre 28 ans et 30 ans selon Chetty et al. (2014)⁵). La mobilité serait proche (ou un peu plus élevée) de celle observée en Italie (coefficient de 0,25 selon Acciari et al., 2021) et en Australie (coefficient de 0,26 selon Murray et al., 2018)⁶ en corrigeant par le biais de cycle de vie. La mobilité serait cependant plus faible qu'en Suisse, en Suède, au Danemark et au Canada où les pentes rang-rang seraient inférieures ou égales à 0,2⁷. De plus, d'après nos résultats, la persistance des revenus est moins forte en France entre les

3 Cette dernière, l'élasticité intergénérationnelle des revenus, a au moins quatre inconvénients par rapport à la mesure de la mobilité positionnelle : (i) elle est basée sur des logarithmes et donc empêche de prendre en compte les revenus nuls ou négatifs, (ii) elle est très non linéaire contrairement au coefficient en termes de rang ; (iii) elle est beaucoup plus sensible aux biais d'atténuation et de cycle de vie ; (iv) elle ne peut pas être utilisée pour faire des analyses géographiques.

4 Autour de 26 ans, selon nos estimations, pour la mobilité ascendante (voir partie 5.1.2), et vers 27 ans selon les résultats de Kenedi et Sirugue (2021, Figure 4) pour la mobilité ascendante relative.

5 Résultats estimés à partir des revenus au niveau ménage pour les enfants. En considérant les revenus individuels, les auteurs obtiennent une pente rang-rang de 0,29, également très supérieure à celle que nous obtenons.

6 Nous considérons ici la mesure la plus proche de la nôtre, en utilisant le revenu individuel des enfants. Au niveau du revenu total du ménage, le coefficient obtenu par les auteurs est de 0,27. Même en corrigeant nos coefficients de 10 à 15 % pour refléter le résultat à l'âge de 35 ans, la pente rang-rang en Australie serait plus élevée qu'en France. À noter qu'en Italie, la pente de 0,25 est obtenue après correction de l'évasion fiscale. Elle était de 0,25 dans Acciari et al. (2020) et est de 0,22 avec de nouvelles données non corrigées de l'évasion.

revenus des pères et ceux des enfants (filles ou fils), très faible entre les revenus des mères et ceux des filles et non significative entre les revenus des mères et ceux des fils. En outre, le rang des enfants est davantage lié au rang du parent dont le revenu est le plus élevé, la relation étant ainsi plus forte qu'avec la moyenne des revenus des parents (0,26 contre 0,24). Les résultats sont très proches même en utilisant différentes définitions des revenus des parents (revenus individuels moyens ou niveau de vie notamment). Cependant, si la relation entre la position des parents et la moyenne de la position des enfants est positive et quasi-linéaire, la position des enfants varie beaucoup, même à centième de revenu des parents fixé : même dans le bas de la distribution des parents, 10 % des enfants ont des revenus qui dépassent le 80^e centième alors que, tout en haut de la distribution des parents, 10 % des enfants ont des revenus inférieurs à ceux du 10^e centième. L'écart inter-quartile à rang des parents donné est ainsi d'environ 5 déciles, et l'écart interdécile de 8 déciles. Les nombreuses situations de mobilité ascendante et de mobilité descendante suggèrent que les revenus des parents ne déterminent pas à eux seuls les revenus des enfants, qui dépendent de nombreux autres facteurs.

L'étude des matrices de transition entre cinquièmes de population permet de compléter notre état des lieux de la mobilité intergénérationnelle en France. Elle confirme d'abord que les inégalités se transmettent entre génération : les enfants des familles dont les revenus sont élevés ont trois fois plus de chances d'être classés parmi les 20 % les plus aisés que ceux issus des familles modestes. On observe aussi que le pourcentage d'enfants restant favorisés est plus élevé que le pourcentage d'enfants restant défavorisés. Mais cette analyse fait aussi apparaître une mobilité certaine. 72 % des jeunes adultes appartiennent à un cinquième de revenu différent de celui de leurs parents. Le taux de mobilité ascendante des jeunes adultes de 28 ans, du plus bas au plus haut cinquième de la distribution des revenus, est de 12 %, un taux plus élevé qu'aux États-Unis (7,5 % selon Chetty et al., 2014, et 8 % selon Alesina et al., 2018) et qu'en Allemagne (Schnitzlein, 2016) mais moins élevé qu'au Canada (13,4 % selon Corak et Heisz, 1999)⁸. Par conséquent, la mesure de la mobilité ascendante confirme le diagnostic établi à partir de l'étude de la persistance des revenus en France. Or, cette réalité est peu connue dans notre pays, probablement faute de données permettant de la documenter. Ainsi, Alesina et al. (2018) montrent que les Français (mais aussi les Anglais et les Suédois) ressentent une mobilité ascendante des enfants de familles défavorisées plus faible que celle qui est mesurée (à l'inverse des Américains qui en ont une perception par trop optimiste).

Le lieu de résidence pendant l'enfance a une forte influence sur la mobilité des individus. La mobilité ascendante entre cinquièmes extrêmes est la plus élevée en Île-de-France (21 %) et la plus faible dans les Hauts-de-France (7 %) et en Normandie (8 %). Il en est de même pour la mobilité ascendante absolue : les enfants de familles modestes peuvent en moyenne espérer atteindre le 49^e centième des revenus en Île-de-France, contre le 41^e dans les Hauts-de-France. La persistance des revenus est la plus faible en Île-de-France, en

7 La pente rang-rang serait de 0,14 en Suisse (Chuard-Keller et Grassi, 2021, Figure 1), 0,20 en Suède (Heidrich, 2017, Table 2), et 0,17 au Canada (Corak et Heisz, 1999, reporté par Chetty et al., 2014, Figure II). Au Danemark, Helsø (2021) obtient une pente de 0,22 et Boserup, Kopczuk, et Kreiner (2017, Figure 4), une pente de 0,20. En Suède, Nybom et Stuhler (2017) obtiennent une pente inférieure à 0,2 en mesurant la pente à l'âge de 28-29 ans. Toutes ces études considèrent les revenus des enfants au niveau individuel, de la même manière que dans notre étude, ce qui permet d'assurer la comparabilité des résultats. La comparaison est plus incertaine dans les pays où les résultats portent sur les revenus au niveau de la famille, notamment en Allemagne où la pente rang-rang est estimée à 0,245, soit légèrement supérieure à la nôtre. Malgré cette différence conceptuelle également présente en Norvège, on peut conclure que la mobilité en Norvège est plus forte qu'en France, étant donné que la pente rang-rang estimée par Bratberg et al. (2017) serait de 0,22 (soit inférieure à la nôtre) et que les estimations sur les revenus au niveau du ménage sont plus élevées qu'au niveau individuel.

8 Données sur le Canada reportées par Chetty et al. (2014). La comparaison sur l'Allemagne porte sur le taux de mobilité ascendante par quartile (17 % selon nos résultats pour la France contre 15 % en Allemagne). Le pourcentage de mobilité ascendante serait plus proche au Danemark, de 11,7 %, selon Boserup, Kopczuk, et Kreiner (2013), mais les différences de données et méthodes ne permettent pas de trancher.

Nouvelle-Aquitaine, et dans les Pays de la Loire, et la plus élevée dans les Hauts-de-France et dans le Grand Est. Des différences existent également selon la densité ou le caractère urbain ou rural des territoires mais celles-ci sont plus faibles qu'au niveau des régions.

À partir de régressions de Poisson, nous analysons les facteurs associés à une plus forte mobilité ascendante. Nous privilégions alors une mesure de la mobilité élargie, à savoir la probabilité de passer des 40 % les plus modestes aux 40 % les plus aisés. Toutes choses égales par ailleurs, cette mobilité ascendante élargie est d'autant plus forte que : les parents ont des revenus du capital élevés; la personne de référence (i.e., le parent au plus haut revenu) a le baccalauréat ou un diplôme du supérieur, et est immigrée; la famille a été mobile géographiquement pendant l'enfance de l'individu, et que l'enfant réside en Île-de-France à sa majorité. À l'inverse, être une femme, avoir vécu dans une famille monoparentale ou dans une famille dans laquelle la personne de référence est ouvrier ou employée et vivre dans les Hauts-de-France influent négativement sur la mobilité ascendante. Enfin, résider en Île-de-France favorise surtout la mobilité ascendante « extrême » entre le plus bas et le plus haut cinquième de la distribution des revenus. Cet effet est lié à l'attractivité et aux opportunités d'études supérieures et d'emplois qu'offre l'Île-de-France. Il existe plus généralement une corrélation positive entre revenus médians des territoires et mobilité, mais pas de corrélation négative entre inégalités territoriales et mobilité intergénérationnelle comme dans certains pays. L'effet propre des territoires demeure malgré la prise en compte du niveau de richesse du territoire, sauf pour l'Île-de-France. Le fait d'être immigré conduit également à un risque plus élevé de mobilité descendante. Afin de déterminer le poids de chaque facteur dans la mobilité ascendante, on mobilise le concept de « fraction attribuable à la population ». Dans cette analyse, le rôle de la région d'origine est prépondérant, devant le sexe et les diplômes des parents.

Ajouter d'autres variables explicatives (sexe, configuration familiale, catégories socioprofessionnelles et diplômes des parents, région d'origine, revenus du capital) pour expliquer le rang des enfants conduit à diminuer le coefficient du rang des parents (de moitié environ), et le rapport de chance d'atteindre les 20 % les plus aisés entre les enfants des familles aisées et les enfants des familles modestes (de 3 à 2 environ). Si les revenus des parents demeurent le principal déterminant de la probabilité d'atteindre les 20 % les plus aisés, la majeure partie de la variabilité des revenus des enfants n'est pas expliquée. Ainsi, étudier la mobilité intergénérationnelle des revenus ne permet de saisir qu'une partie de l'importance du milieu familial dans la distribution des revenus : on la complète donc, dans la dernière partie de cet article, en comparant les revenus des frères et sœurs. Étant donné que la différence de rang moyenne entre ces derniers est d'un peu plus de 2 déciles et que, par ailleurs, la différence théorique de rang entre deux enfants est de 0,33 déciles, on en déduit qu'environ 30 % de la variation de leurs revenus serait liée à la famille.

Revue de littérature. Notre travail s'inscrit dans le prolongement des études antérieures sur la mobilité intergénérationnelle. Il suit pour une large part le travail de Chetty *et al.* (2014). Celui-ci se fonde lui-même sur les travaux de Solon (1999) et Black et Devereux (2011) qui mesurent la mobilité en reliant les revenus des parents et ceux des enfants. Chetty *et al.* (2014) sont les premiers à présenter une analyse en termes de rang et montrent que la mobilité intergénérationnelle du revenu est faible aux États-Unis, et que celle-ci varie considérablement d'une région à l'autre. Corak et Heisz (1999), Schnitzlein (2016), Boserup *et al.* (2017), Nybom et Stuhler (2017), Muray *et al.* (2018), Helsø (2021) et Acciari *et al.* (2021) mettent en œuvre des méthodologies similaires sur des données produites au Canada, en Allemagne, au Danemark, en Suède, en Australie et en Italie. Par rapport à ces articles, l'une des contributions originales de notre travail consiste d'abord à documenter la situation d'un nouveau pays, la France. Notre approche permet également de décrire la mobilité intergénérationnelle en mobilisant de nombreuses variables socio-démographiques au niveau individuel, ainsi que des variables géographiques. C'est notamment la première

fois, à notre connaissance, que la mobilité est mesurée en fonction des revenus du capital des parents, ce qui permet d'étudier la manière dont la détention de capital influence la mobilité.

Si la mobilité sociale intergénérationnelle en termes de catégories sociales a fait l'objet de nombreuses études en France (voir Vallet, 1999, 2015 ou plus récemment Collet et Pénicaud, 2019) grâce aux enquêtes Formation et qualification professionnelle (FQP), la mobilité en termes de revenus a été peu abordée jusqu'à présent. En effet, les enquêtes FQP ne permettent pas d'observer le niveau de revenu des parents. Or, du fait de l'évolution de la structure des emplois dans le temps, les mesures s'appuyant sur les professions sont difficilement comparables dans le temps et entre pays. Pour contourner ce problème de données, deux catégories de travaux ont été menées. D'une part, une méthode mixte a été mise en œuvre par Lefranc et al. (2014) et Dherbécourt (2018), en reliant les revenus des enfants à la catégorie socioprofessionnelle des parents. Dherbécourt (2018) montre que l'origine sociale a un effet très discriminant sur l'accès à un niveau de vie élevé, et que cela est en grande partie lié aux différences de diplômes. Dherbécourt et Kenedi (2020) poursuivent ce travail en intégrant également la dimension géographique. Dans une deuxième catégorie de travaux se basant sur des études internationales, Trannoy et Lefranc (2005), dans un article séminal publié en France, puis Lefranc (2018) et Alesina et al. (2018), imputent le revenu moyen des pères par des méthodes d'inférence par variables instrumentales (popularisées par Bjorklund et Jantti, 1997) à partir des enquêtes FQP, et estiment l'élasticité intergénérationnelle des revenus (IGE). Kenedi et Sirugue (2021) mettent également en œuvre cette méthode pour calculer différents indicateurs de mobilité (notamment en utilisant le rang) en ayant cette fois-ci recours à l'EDP, ce qui leur permet de décrire la mobilité au niveau local. Nos estimations, obtenues à partir du même concept de revenu et d'un échantillon d'enfants d'âges comparables, sont très proches des leurs, mais elles sont plus faibles que celles des auteurs qui ont recours à la notion de revenus au niveau du ménage pour les enfants et non aux revenus individuels. En effet, les effets d'homogamie augmentent la persistance des revenus et, en prenant les revenus individuels, nos résultats convergent avec les leurs⁹. À partir des résultats de Lefranc (2018) sur l'IGE, l'OCDE (2018) calcule un indicateur de mobilité sur plusieurs générations. Bien que ces auteurs utilisent des méthodes similaires, les estimations de l'IGE divergent selon ces contributions, du fait de différences d'hypothèses, de méthodes et de données. Par rapport à toutes ces publications, l'originalité de notre contribution réside en ce qu'elle relie directement les revenus des parents aux revenus de leurs enfants. Cette démarche a l'avantage de fournir des statistiques de mobilité semblables aux premières études citées, tout en n'étant pas tributaire des hypothèses d'imputation et de restriction de population¹⁰ des méthodes évoquées précédemment (celles-ci induisent un aléa sur les résultats). Par ailleurs, la méthode d'imputation aurait tendance à biaiser à la hausse l'inertie des revenus, de 0,1 point ou plus¹¹ (Björklund et Jantti, 1997, Acciari et al. 2021, Kenedi et Sirugue, 2021). C'est, à notre connaissance, la première fois que les revenus des parents et ceux des enfants sont directement reliés en France. Cela nous permet également de présenter de nouveaux résultats au niveau des territoires, et de les décliner selon d'autres variables

9 Les auteurs obtiennent également une pente rang-rang de 0,24 en mesurant les revenus des enfants au niveau individuel à l'âge de 28 ans.

10 En effet, ces méthodes d'inférence nécessitent de faire des hypothèses d'imputation et d'exclure une partie des professions des parents (notamment les indépendants et une partie des fonctionnaires dans l'EDP).

11 Tous ces auteurs comparent l'estimation de l'IGE (Intergenerational income elasticity) entre les revenus des enfants et ceux des parents, en utilisant la méthode par imputation et des données permettant de lier directement enfants et parents. Björklund et Jantti (1997) montrent un biais de 33 % (0,13 point) à la hausse, du fait de cette méthode (qui donne un IGE de 0,52 contre 0,39 pour les données réelles). Jerrim et al. (2016) obtiennent toujours un biais haussier mais plus faible, de 13 % (0,64 contre 0,57). Acciari et al. (2020, partie 4.1, page 17) trouvent un IGE de 0,4 en Italie avec la méthode par imputation (cohérent avec les autres estimations qui sont situées entre 0,35 et 0,55), contre 0,25 avec la méthode directe, soit un biais plus important (de 60 %, ou 0,15 pt). Enfin Kenedi et Sirugue (2021) mettent en évidence un biais de 22 % (0,08 pt) en France.

spécifiques aux parents comme le sexe, le diplôme, le statut d'immigré et le niveau de revenu du capital, en utilisant un modèle économétrique au niveau individuel.

La principale limite de notre approche réside en ce que les résultats obtenus ne concernent que des adultes âgés de moins de 30 ans. La littérature montre cependant que la pente rang-rang se stabilise autour de 30 ans, et que la différence entre les pentes rang-rang à 28 ans et à 35 ans est d'environ 10 %¹². De plus, certains travaux montrent que la mobilité intragénérationnelle au cours de la vie est relativement faible¹³. Cela suggère que nos résultats sur les enfants de 28 ans ne souffrent que légèrement du biais de cycle de vie. Par ailleurs, mesurer les revenus à ces âges n'est pas atypique dans la littérature : Chetty et al. (2014) utilisent les revenus à 29-30 ans pour étudier la mobilité intergénérationnelle, Chetty et al. (2011) les revenus des jeunes adultes à 27 ans pour mesurer l'effet d'une réforme éducative sur les revenus des enfants, et Chetty et al. (2016) les revenus entre 24 ans et 28 ans pour mesurer l'effet de l'exposition à un meilleur voisinage sur les revenus des enfants à l'âge adulte.

L'article est structuré comme suit. La deuxième partie est consacrée à la description des données, des traitements effectués, et du champ adopté. La troisième partie présente des statistiques descriptives sur la base de données spécifiquement construite pour cette étude et la quatrième partie introduit la méthode et les concepts utilisés par la suite. La partie 5 présente les résultats obtenus, d'abord au niveau national, puis en les déclinant selon différentes caractéristiques socio-démographiques et par territoires et, enfin, selon les termes d'une analyse toutes choses égales par ailleurs. La dernière section cherche à expliquer la part de la variation des revenus des enfants liée à l'environnement familial.

2 Données, traitements et champ

Nous commençons par décrire dans cette section la source de données utilisée, puis détaillons les traitements opérés afin de répondre à notre question de recherche. Enfin, nous décrivons le champ et les variables de revenus utilisés dans le cadre de cette étude.

2.1 Les données

Constitué à partir de 1968, l'Échantillon Démographique Permanent (EDP) est un panel d'individus construit par l'Insee qui recueille des informations statistiques sur l'ensemble des personnes nées l'un des « seize jours EDP »¹⁴ de chaque année civile (du 2 au 5 janvier, puis les quatre premiers jours du premier mois de chaque trimestre). Si l'EDP s'est historiquement construit à partir de l'état civil (qui enregistre notamment les naissances en France) et du recensement de la population, il rassemble aujourd'hui, outre ces sources statistiques, des données issues des déclarations annuelles de données sociales (DADS), du fichier électoral et, depuis 2015, des données fiscales et sociales (Fidéli et Filosofi) (Robert-Bobée et Gualbert, 2021). Les données fiscales du millésime 2019 de l'EDP couvrent l'ensemble des revenus des années 2010 à 2018 et permettent de disposer de données annuelles sur tous les « individus EDP » (soit pour un peu plus de 4 % de la population française chaque année depuis 2008).

Les données fiscales de l'EDP offrent de nombreuses informations de nature fiscale et démographique qui permettent de suivre l'évolution des revenus individuels détaillés de l'ensemble des individus apparaissant sur la déclaration fiscale de « l'individu EDP » (type

12 En France, Kenedi et Sirugue (2021) obtiennent un écart de 8 %. Aux États-Unis, Chetty et al. (2014) observent un écart d'environ 10 %, voir partie 4.3.

13 Deux tiers des personnes restent dans le plus haut cinquième de la population sur une décennie selon l'OCDE (2018). Nous observons également que la pente rang-rang a tendance à se stabiliser à 28 ans en France.

14 Depuis 2008. Avant cette date, les individus EDP étaient identifiés seulement les 4 premiers jours d'octobre.

de revenu perçu par chaque individu, montant de chaque type de revenu) ainsi que celle de variables agrégées au niveau du ménage de chaque individu EDP, grâce à l'appariement des données fiscales avec les données sociales dans le cadre de la source Filosofi (niveau de vie et revenus du ménage fiscal de l'individu EDP, montants de diverses prestations sociales, montants des impôts payés...). En outre, elles permettent de localiser l'individu EDP à son lieu de résidence chaque année entre 2010 et 2018.

L'EDP intègre des données des recensements exhaustifs de 1968 à 1999 et, depuis 2004, des enquêtes annuelles de recensement. Ces sources contiennent des variables qui permettent de caractériser la situation socio-démographique de l'individu EDP comme celle de l'ensemble des personnes qui résident habituellement avec lui (son ménage) : on connaît en particulier le diplôme et la catégorie socioprofessionnelle de l'individu EDP et de chaque individu appartenant au même ménage que celui-ci. Enfin, les DADS contiennent des informations annuelles sur l'emploi du seul individu EDP (type d'emploi, nature du contrat de travail, rémunération, quotité de travail...) mais ne concernent que les salariés et ne sont disponibles que jusqu'en 2017.

2.2 Constitution de la base d'étude et traitements

a) Constitution de la base d'étude

Une base de données idéale permettant d'étudier la mobilité intergénérationnelle des revenus en France offrirait une profondeur temporelle importante (au moins une trentaine d'années) afin de pouvoir observer les revenus des parents et ceux de leurs enfants au même âge, par exemple entre 30 et 40 ans. Malheureusement, l'EDP ne couvre actuellement que neuf années de données fiscales¹⁵. Ce manque relatif de profondeur temporelle nous conduit à développer la stratégie suivante pour étudier la mobilité intergénérationnelle des revenus. Elle s'appuie sur la présence des enfants sur la déclaration fiscale de leurs parents (ou d'au moins un de leurs parents) pour récupérer les revenus individuels des parents, et le revenu du ménage, en 2010 (c'est-à-dire au cours de l'année fiscale 2011, première année de disponibilité des données fiscales dans l'EDP¹⁶), lorsque leurs enfants ont autour de 18-20 ans. Elle consiste ensuite à comparer ces revenus à ceux de leurs enfants huit ans plus tard en 2018, lorsque ceux-ci sont entrés sur le marché du travail. L'objectif est de reconstituer les informations individuelles (revenus individuels, sexe, statut conjugal, lieu de résidence, ...) et les informations au niveau du ménage de l'ensemble de ces « enfants EDP » chaque année entre 2010 et 2018¹⁷. Pour cela, nous sélectionnons les enfants EDP qui, en 2010, sont dans le foyer fiscal au sens de l'impôt sur le revenu¹⁸ de leurs parents. L'année 2018, la dernière année de disponibilité des données fiscales, nous offre ainsi la possibilité de comparer les revenus des parents des enfants identifiés en 2010 avec ceux de leurs enfants à l'entrée sur le marché du travail, entre 26 et 29 ans (i.e., nés entre 1989 et 1992).

15 Il serait théoriquement possible d'utiliser les données DADS sur longue période, mais cela pose différents problèmes : d'abord le champ est biaisé car ces données n'intègrent pas les indépendants (et les fonctionnaires avant les années 2000) ; ensuite, l'échantillon serait de petite taille car il ne concernerait que les enfants EDP qui ont des parents EDP ; et, enfin, avant 2008, l'échantillon de l'EDP était beaucoup plus réduit car il intégrait seulement les personnes nées l'un des quatre jours historiques de l'EDP et non celles nées lors de l'un des seize jours EDP, ce qui accentuerait le problème lié à la taille de l'échantillon.

16 Pour faciliter la lecture, nous évoquerons principalement des années de revenu et non des années fiscales par la suite. À noter cependant que les informations socio-démographiques individuelles sont relatives à l'année fiscale.

17 C'est-à-dire l'année fiscale 2019, dernière année de disponibilité des données fiscales dans l'EDP.

18 Rappelons que le foyer fiscal correspond à l'ensemble des individus appartenant à la même déclaration de revenus. Différents foyers fiscaux peuvent vivre dans le même logement et constituer un ménage fiscal. La notion de ménage fiscal, au niveau du logement, est donc plus large que celle de foyer fiscal.

Une fois les informations fiscales sur ces différentes cohortes d'enfants obtenues, nous cherchons à produire les informations fiscales des parents des individus EDP (à la fois les caractéristiques individuelles et leurs revenus individuels). Ces informations ne sont pas produites directement dans l'EDP : pour l'année 2010 (déclaration en 2011), on connaît dans l'EDP seulement les revenus des déclarants et des conjoints de déclarants dans les déclarations fiscales dans lesquelles figurent les enfants EDP. Nous comparons alors les informations individuelles des parents issues de la table générale de la source fiscale dans l'EDP (table FISC_INDIVIDU) à celles des déclarants et conjoints de déclarants issues de la table détaillée des revenus de cette même source (table FISC_REVDET) pour pouvoir déterminer qui du déclarant et du conjoint du déclarant sont respectivement le père et la mère de l'individu EDP¹⁹. Dans les déclarations fiscales sur lesquelles les enfants EDP sont déclarés, les « déclarants » et « conjoints de déclarants » constituent les parents potentiels de ces enfants. Il peut s'agir parfois de beaux-parents. En effet, en cas de séparation antérieure des parents biologiques de l'enfant EDP et de remise en couple de l'un des deux parents, le déclarant et le conjoint de déclarant figurant sur la déclaration fiscale sont alors un parent et un beau-parent, et c'est le revenu de ce beau-parent, et non celui du parent biologique, qui tient lieu de revenu du père ou de la mère pour notre analyse (voir annexe A pour le traitement de ces cas). Cette approche reste pertinente : les familles recomposées ne sont pas une situation majoritaire (en 2018, 11 % des 15-17 ans vivent avec un parent et un beau-parent – Algava et al., 2020), et nous cherchons davantage à mesurer la reproduction des inégalités liées au niveau de vie et à l'environnement pendant l'enfance qu'à documenter la situation des parents biologiques²⁰.

b) Pondérations et sélection des enfants dans l'EDP

L'intégration d'individus dans l'EDP repose sur un certain nombre d'informations identifiantes (en particulier, leurs dates de naissance). Or, ces informations, qui sont contenues dans les données fiscales, sont moins accessibles et de moins bonne qualité pour les enfants que pour les adultes. L'identification et l'intégration des enfants dans la base s'avère donc souvent plus difficile, surtout pour la première année de données fiscales incluse dans l'EDP (l'année 2010). En effet, le nombre d'enfants présents dans les données fiscales augmente après 2010 à mesure que les enfants deviennent adultes et que les problèmes d'identification s'estompent (Annexe B). Pour pallier le défaut d'identification des données fiscales et augmenter la taille de nos cohortes, nous intégrons aux cohortes d'enfants identifiés dès l'année 2010 (année fiscale 2011) les enfants qui n'apparaissent qu'à partir des années 2011 ou 2012. On considère ainsi que la situation familiale de ces enfants observée au cours des années 2011 et 2012 (en corrigeant les revenus de l'inflation) correspond à celle de l'année 2010 dès lors qu'ils apparaissent encore comme enfants à charge dans la déclaration fiscale de leurs parents. Cette approche permet de reconstituer en grande partie, mais pas totalement, les différentes cohortes qui sont étudiées dans le cadre de cet article.

Afin de constituer un échantillon d'enfants représentatif de la population française et atténuer les problèmes d'identification évoqués précédemment, les données sont ensuite pondérées. L'EDP fournit par défaut plusieurs jeux de pondérations pour chacune des sources qui l'alimentent. Étant donné que les données fiscales constituent la base qui nous permet d'identifier les enfants suivis entre 2010 et 2018, nous utilisons la variable POIDS_FIDELI

19 On procède de même pour les années 2011 et 2012, pour les individus qui n'apparaissent qu'à partir de ces deux années (cf. infra, partie suivante).

20 Par exemple, si les parents divorcent au cours de l'enfance d'un individu et que, après la remise en couple de l'un de ses parents, le beau-parent a des revenus plus élevés que ceux de l'un de ses parents, cette situation peut influencer positivement sur le niveau de formation de l'enfant (possibilité d'accéder à des formations post bac onéreuses, par exemple). Dans le cas où le beau parent est moins aisé, nous incluons les pensions alimentaires versées dans le niveau de vie des parents (et nous prenons ainsi en compte le niveau de vie du parent biologique).

pour la construction de nos statistiques. Cette variable est fournie dans l'EDP pour chacun des millésimes des données fiscales. Elle cale les effectifs intégrés à l'EDP sur les marges de l'exhaustif de Fidéli. Elle ne permet cependant pas de reconstituer avec une grande précision les effectifs de la population française par âge et par sexe. Par conséquent, nous repondérons cette variable en utilisant les fichiers détaillés de l'Insee des effectifs de la population française par âge, par sexe et par état matrimonial pour estimer plus précisément les effectifs de la population française selon les modalités de ces variables.

c) Intégration de variables socio-démographiques à notre base de données

Nous ajoutons également à notre base de données les informations issues du recensement de la population de 1999 afin de décrire le milieu social d'origine des enfants EDP identifiés dans les données fiscales. Ces informations décrivent le milieu d'origine des enfants, au moment où ces derniers étaient âgés de 7 à 10 ans. Cependant, si le recensement de la population de 1999 est exhaustif, les données du recensement de la population de 1999 intégrées à l'EDP ne portent que sur des personnes EDP nées lors d'un des quatre jours historiques de l'EDP (et non lors de l'un des seize jours EDP comme c'est le cas actuellement). Jusqu'à l'année 2008, en effet, seuls les individus nés les quatre premiers jours d'octobre (lors d'un des quatre jours « historiques » de l'EDP) étaient intégrés à l'EDP. Ainsi, environ un quart seulement des enfants figurant dans les données fiscales sont présents dans le recensement de l'année 1999. Pour augmenter les effectifs d'enfants dont il est possible de décrire le milieu familial d'origine, nous utilisons les enquêtes annuelles de recensement après 2008. Par exemple, pour les enfants nés en 1992, nous enrichissons nos effectifs d'enfants couverts par le recensement de 1999 en utilisant les enquêtes annuelles de recensement entre 2008 et 2016 : les informations sur le milieu social des enfants concernés ont alors été obtenues à un âge inférieur ou égal à 18 ans, à un âge auquel ces enfants vivent encore très fréquemment chez leurs parents. Grâce à cette opération, le taux de couverture de nos enfants par les données du recensement de 1999 ou une enquête annuelle de recensement passe de 25 % à plus de 70 % en moyenne par génération. Ce taux de couverture étant différent selon la cohorte, une repondération a été mise en œuvre dans les analyses mobilisant ces variables. L'ensemble des traitements sont détaillés dans l'annexe C.

2.3 Le champ de l'étude

Dans cette étude, nous nous intéressons aux seuls enfants repérés dans les données fiscales qui vivent dans le foyer fiscal de leurs parents en 2010, 2011 ou 2012, déclarent seuls leurs revenus en 2018, ont des revenus positifs ou nuls (en 2018), et dont les parents perçoivent également des revenus positifs ou nuls en 2010. Nos résultats portent sur la France métropolitaine²¹. Les effets des différentes restrictions sont indiqués dans l'Annexe D.

Nous utilisons deux échantillons principaux dans l'analyse qui suit. Les résultats qui portent sur la population nationale mobilisent les données relatives à la génération d'enfants nés en 1990 : ces enfants ont 28 ans en 2018 (dernière année pour laquelle nous disposons des données fiscales). En effet, l'étude de cette population permet de minimiser le biais de cycle de vie pour certains indicateurs comme la pente rang-rang (cf. infra). Ensuite, les analyses portant sur les liens entre la mobilité intergénérationnelle, d'une part, les caractéristiques socio-démographiques des enfants et des parents ainsi que les territoires, d'autre part, sont réalisées à partir d'un échantillon d'enfants plus large afin de produire des résultats plus robustes : la base de données construite à cet effet consiste en un empilement des

21 Seules les données sur la Réunion et la Martinique sont disponibles dans l'EDP, mais ces données sont trop peu nombreuses pour être analysées dans cette étude. Les études sur l'EDP portent en général sur ce même champ de la France métropolitaine.

générations 1989, 1990, 1991 et 1992. Ainsi, les revenus des enfants sont observés en 2018 entre les âges de 26 et 29 ans.

Nous focalisons notre analyse de la mobilité intergénérationnelle sur les générations 1989 à 1992 pour deux raisons. D'abord, nous n'analysons pas les résultats relatifs aux générations 1993 à 1996 (i.e., âgées de 22 à 25 ans en 2018) car les indicateurs ne sont pas robustes pour ces cohortes, une trop grande proportion d'enfants étant encore scolarisés dans le supérieur en 2018. De même, nous ne livrerons pas non plus de résultats concernant les générations nées avant 1989 car peu d'« enfants » nés avant 1989 vivent encore dans le foyer fiscal de leurs parents en 2010 (à 22 ans ou plus). Rappelons en effet ici qu'une des restrictions de champ nécessaires à notre étude (cf. supra) consiste à exclure les individus qui déclarent seuls leurs revenus à l'âge de 18 ans afin de récupérer l'information relative aux revenus de leurs parents. Ces individus sont considérés comme déclarants ou conjoints de déclarants dans leur déclaration fiscale et, comme ils déclarent seuls leurs revenus, il n'est pas possible d'observer les revenus de leurs parents²². Ils représentent ainsi environ 10 % des individus nés en 1992, près d'un quart des individus nés en 1991 et environ 40 % de ceux nés en 1990, et plus encore pour les générations les plus âgées, notamment en intégrant les données fiscales 2011 et 2012 à l'analyse (cf. annexe D). Ils représentent une proportion trop importante de la génération 1988 qui a 30 ans en 2018 (plus de 70 %), ce qui nous conduit à ne pas inclure la génération 1988 dans nos analyses. Les revenus des individus rattachés à leurs parents dans la déclaration fiscale sont un peu plus élevés (de 5 % pour la génération 1990) mais restent relativement proches des revenus de l'ensemble des enfants (Annexe D). Comme, de plus, la sélection conduit à une proportion plus élevée de parents aisés dans l'échantillon que dans la population totale (davantage de cadres et de diplômés du supérieur, voir tableaux de l'annexe B), l'effet du biais sur les indicateurs de mobilité est incertain et devrait être relativement faible²³. Face à ce biais de sélection qui affecte également potentiellement leur estimation de la mobilité en Italie, Acciari et al. (2021) montrent finalement que la sélection a peu d'effet sur les revenus et que celle-ci devrait peu altérer les résultats.

Notre approche pose enfin la question des potentiels étudiants figurant encore dans notre échantillon : leur niveau de vie observé ne correspond pas à celui dont ils bénéficieront une fois entrés sur le marché du travail. Ce problème ne paraît pas en mesure de biaiser notre analyse car, à 28 ans, seul 1,0 % des individus sont encore en études initiales (Bernard, 2021). De plus, nous utilisons également l'enquête annuelle de recensement de 2018 pour repérer et supprimer les étudiants de notre base de données²⁴. Cette enquête ne couvre cependant que 14 % de l'échantillon en 2019, ce qui ne règle qu'en partie cet éventuel problème. Nous réalisons également des tests de robustesse en utilisant un seuil minimum de revenus du travail afin d'éliminer de notre base les individus qui cumulent emploi et études. Deux seuils sont retenus pour ces tests : un seuil de 7 200 euros annuels, qui correspond à un peu plus de la moitié du Smic à temps plein²⁵, et un seuil plus élevé, égal à 10 000 euros, qui correspond à un peu plus de 70 % du Smic à temps plein. Dans notre

22 Ces individus perçoivent a priori des revenus plus faibles, en moyenne, que les autres individus de la base, ce qui pourrait biaiser les résultats. Cependant, les statistiques descriptives montrent que les revenus, au sein de chaque génération, ne sont pas très différents. Nous classons les individus par génération, ce qui doit minimiser l'ampleur de ce problème.

23 Il pourrait néanmoins affecter plus certainement la constante C de l'équation rang-rang.

24 Les étudiants sont repérés grâce à la question du recensement portant sur l'inscription dans un établissement d'enseignement.

25 Ce seuil est supérieur au revenu moyen reçu par les étudiants selon l'enquête ENRJ : les étudiants perçoivent en moyenne environ 350 euros par mois (soit 4 200 euros annuels) de revenus du travail, et au maximum 500 euros (soit 6 000 euros annuels) pour les étudiants en écoles d'ingénieurs selon Lièvre (2018). Les étudiants perçoivent également des revenus de la part de leurs parents ou des revenus sociaux, mais ces revenus n'apparaissent pas dans nos données (nous ne les considérons donc pas). Le seuil choisi devrait donc permettre d'exclure la majorité des étudiants qui cumulent études et emploi.

analyse principale, nous ne retenons pas ces seuils pour tenir compte des personnes qui ont des difficultés d'insertion sur le marché du travail et qui peuvent percevoir des revenus du travail très faibles (inactivité, chômage ou temps partiel subi). En effet, exclure ces personnes risque de biaiser les estimations car les revenus de leurs parents sont en moyenne significativement plus faibles que ceux des autres parents (la médiane de leurs revenus est inférieure de 4 000 à 5 000 euros par an)²⁶. Nous avons également essayé d'utiliser des variables sur le marché du travail issues des DADS (ces données sont présentes dans l'EDP), mais cette voie s'est révélée infructueuse car les DADS couvrent une année de moins que les données fiscales dans l'EDP. Un filtre sur les personnes en stage ou apprenties en 2017 conduirait à supprimer de notre étude des personnes qui ont des revenus significativement plus élevés en 2018, probablement parce qu'elles ne sont plus en stage ou apprenties et ont trouvé un travail.

2.4 Définition des variables utilisées dans l'analyse

a. Les variables de revenus

Les variables de revenus utilisées dans cet article sont décrites dans les paragraphes suivants. Elles sont toutes corrigées de l'inflation (base 100 en 2018).

Les revenus individuels (d'activité et de remplacement) des parents

Nous mobilisons la base Fideli pour connaître les revenus individuels des parents. Nous calculons les revenus individuels des parents, en additionnant les revenus du travail, les allocations chômage et les pensions de retraites, des pères et des mères. Il s'agit des revenus reportés dans les déclarations d'impôts sur le revenu, nets de cotisations et de CSG déductible. Cette approche permet de calculer le revenu individuel des parents avant toute redistribution. Par la suite, nous ferons référence aux « revenus individuels » sans précision supplémentaire. Il faut ensuite rapporter ce revenu à la taille des ménages, ce qui est fait de deux façons différentes : (i) en le rapportant au nombre d'adultes du ménage, (ii) en le rapportant au nombre d'unités de consommation. Ainsi, ce qu'on appelle dans la suite de cette étude les « revenus des parents » consiste en un revenu commun des parents et c'est cette variable (définie donc au niveau du couple) qui servira à classer les parents entre eux dans la distribution des revenus des parents. Ces revenus sont observés en 2010.

Une précision importante doit être apportée sur le traitement du statut conjugal. Dans les données fiscales de Fideli, la disponibilité des informations relatives aux revenus individuels des parents (leurs types comme leurs montants) dépend du statut conjugal de ces derniers. Ainsi, dans le cas où les parents des enfants sont séparés, un enfant peut alors apparaître sur la déclaration fiscale d'un seul des deux parents si celui-ci en a la garde exclusive (auquel cas les revenus de l'un des deux parents sont inconnus), ou sur les déclarations fiscales de chacun de ses deux parents s'il est en garde alternée (auquel cas, les revenus des deux parents sont disponibles, mais font l'objet de deux observations différentes dans l'EDP, voir annexe A). Ensuite, dans le cas où les deux parents cohabitent, les revenus détaillés des deux parents ne sont disponibles dans l'EDP que lorsque ces derniers sont mariés ou pacsés (une seule déclaration fiscale dans laquelle figurent les informations de chacun des membres du couple). Mais l'information relative aux revenus de l'un des deux parents manque dans le cas où les parents vivent en union libre car l'enfant n'apparaît sur la déclaration de revenus que d'un seul de ses deux parents (chaque parent remplit alors sa propre déclaration fiscale).

Le niveau de vie des parents. Cette variable correspond au niveau de vie du ménage auquel est rattaché l'enfant en 2010. Le niveau de vie est calculé comme le revenu

26 De plus, le fait de choisir un seuil de revenu proche d'un demi Smic diminue de 12 % le nombre d'individus, ce qui serait un inconvénient notable pour les estimations.

disponible (i.e., revenus du travail et du capital, diminués des impôts, auxquels on ajoute les prestations sociales monétaires) rapporté aux nombres d'unités de consommation du ménage. Il prend également en compte les pensions alimentaires versées en cas de séparation. Il est issu de la base Filosofi, qui mesure les niveaux de vie des ménages fiscaux (i.e., au niveau du logement). Toutes les déclarations fiscales rattachées au logement sont alors prises en compte. On retranche les éventuels revenus de l'enfant pour calculer le niveau de vie des parents. Étant donné que nous souhaitons disposer d'une mesure des revenus cohérente entre les enfants et les parents et que les niveaux de vie ne sont pas disponibles pour les enfants (cf. infra), le niveau de vie des parents ne sera utilisé que dans le cadre d'un test de robustesse.

Les revenus individuels d'activité des enfants. Observés en 2018, ils sont tirés de la base Fideli intégrée à l'EDP et sont calculés en sommant les revenus du travail et les revenus du chômage. Ces revenus sont directement tirés des informations déclarées à l'administration fiscale. Ils sont préremplis pour une grande partie de la population et sont donc très fiables.

Les variables de revenus, issues de Filosofi (un appariement entre données fiscales et données sociales), offrent des informations statistiques au niveau des logements des individus EDP. Il est donc théoriquement possible de comparer, dans l'EDP, la position dans l'échelle des niveaux de vie des parents en 2010 avec celle des enfants en 2018. Cependant, le recul temporel offert par la source ne permet pas à ce jour de travailler sur les niveaux de vie des enfants dans le cadre de la présente étude. S'il est possible, avec neuf années de données fiscales, de distinguer clairement les revenus individuels des parents de ceux des enfants, il est difficile de séparer le niveau de vie des enfants de celui des parents car les enfants peuvent vivre chez leurs parents longtemps après la majorité. Dans ce cas, le niveau de vie des parents et celui des enfants sont les mêmes : il s'agit en effet dans les deux cas du niveau de vie de toute la famille, car celui-ci tient compte à la fois des revenus des parents et de ceux des enfants. D'après Pouliquen (2018), 20 % des jeunes âgés de 25 à 29 ans vivent tout ou partie de l'année chez leurs parents. Si ce taux décroît avec l'âge²⁷, il reste néanmoins encore supérieur à 6 % entre 30 et 35 ans. Il ressort de ces éléments qu'il n'est pas possible de travailler proprement sur la comparaison entre les niveaux de vie des parents et ceux des enfants. Par conséquent, seuls les revenus individuels permettront de décrire les enfants. Cependant, malgré ces contraintes, une analyse spécifique restreinte aux enfants qui ont changé de logement entre 2010 et 2018 (et qui ne vivent donc en principe plus chez leurs parents et disposent donc de niveaux de vie distincts de ceux de leurs parents) sera proposée dans la partie 5.1.

La variable de classement. Les enfants sont classés en fonction de leurs revenus par rapport aux revenus des enfants de leur génération. D'autre part, les parents sont classés en fonction de leurs revenus par rapport aux parents qui ont des enfants de la même génération. Cette stratégie permet de ne pas biaiser les résultats, malgré le fait que les revenus des parents sont observés à des âges différents (ce qui est classique dans la littérature, Chetty et al., 2014).

b. Les variables de localisation

Régions et Départements. Les variables de localisation sont issues des données fiscales (via la base Fidéli). Elles permettent de localiser les enfants EDP chaque année entre 2010 et 2018. Nous utilisons l'information de cette variable qui correspond à l'année 2010, une année au cours de laquelle les enfants ont un âge proche de 18 ans et sont le plus souvent

²⁷ Le taux de cohabitation chez les parents est égal à près de 33 % à l'âge de 25 ans et décroît progressivement pour atteindre 13,9 % à l'âge de 29 ans d'après des données de l'enquête logement préparées spécialement pour cette étude.

rattachés fiscalement à leurs parents. La variable relative au département correspond au code départemental de l'adresse fiscale de taxation et comprend 98 modalités (une pour chaque département) dont une pour les foyers fiscaux résidant à l'étranger.

Nous utilisons également le département de naissance, issu des données fiscales, mais cette variable est moins bien remplie que celle qui concerne le département de résidence en 2010.

La mobilité géographique. En comparant les départements où résident les enfants vers l'âge de 18 ans à leurs départements de naissance²⁸, il est possible de constituer un indicateur de mobilité géographique des familles. Dans cette étude, une famille est donc considérée comme mobile si un déménagement impliquant un changement de département s'est produit pendant l'enfance des individus observés.

Aire d'attraction d'une ville (AAV). L'aire d'attraction d'une ville²⁹ définit l'étendue de son influence sur les communes environnantes. Une aire est composée d'un pôle, défini à partir de critères de population et d'emploi, et d'une couronne, constituée des communes dont au moins 15% des actifs travaillent dans ce pôle. Au sein de celui-ci, la commune la plus peuplée est appelée commune-centre. Pour construire cette variable dans l'EDP, nous utilisons une table de passage de l'Insee entre les communes et les aires d'attraction des villes. Comme le zonage correspond à la géographie 2020, nous avons préalablement recalculé les communes à l'année 2020, pour ensuite appliquer la grille de passage commune/AAV.

Tranche d'Aire d'attraction d'une ville. À partir de la variable AAV précédemment construite, nous utilisons un découpage des AAV par nombre d'habitants en six modalités, en distinguant l'aire de Paris et les communes hors attraction des villes.

Densité des communes. Pour prendre en compte la population communale et sa répartition dans l'espace, nous utilisons la nouvelle grille communale de densité de l'Insee. Celle-ci s'appuie sur la distribution de la population dans la commune en découpant le territoire en carreaux de 1 kilomètre de côté. Elle permet ainsi de repérer des zones agglomérées. C'est l'importance de ces zones agglomérées au sein des communes qui permet de les caractériser (et non la densité communale habituelle). Les données de population sont issues des fichiers démographiques sur les logements et les individus ([Fideli 2018](#)).

Pour construire cette variable dans l'EDP, nous utilisons une table de passage de l'Insee entre les communes et la densité des communes. Comme le nouveau zonage des densités de communes est produit en géographie 2020, nous avons préalablement recalculé les communes à l'année 2020, pour ensuite appliquer la grille de passage commune/densité.

c. Les variables du recensement de la population sur le milieu d'origine des enfants

Nous utilisons le recensement de la population de 1999 et les enquêtes annuelles de recensement de 2008 à 2012³⁰ (voir infra, partie 2.2 et annexe C) pour décrire le milieu d'origine des enfants grâce aux variables suivantes : diplômes du père et de la mère, nationalités et pays de naissance (du père et de la mère), ascendance migratoire, catégories socioprofessionnelles (du père et de la mère), type d'activité (du père et de la mère), structure familiale du ménage, type de ménage, statut du père de famille, statut de la mère de famille, conditions d'emploi (du père, de la mère), activité (du père, de la mère). L'annexe C décrit les modalités de chaque variable ainsi que les regroupements auxquels nous avons procédé pour obtenir des effectifs suffisants au sein de chaque modalité. Cette annexe

28 Le département de naissance n'est pas forcément le département où la famille habite à la naissance (selon la proximité des maternités), mais ces cas sont a priori peu nombreux.

29 <https://www.insee.fr/fr/information/4808607>

30 Les EAR sont utilisées jusqu'à 2016, mais les individus sont très peu nombreux à partir de 2012.

présente également des statistiques descriptives pour les variables les plus importantes. Dans les analyses qui suivent, nous utilisons la variable de la personne de référence, définie comme le parent ayant le plus haut revenu.

d. Autres variables

Les revenus du capital. Les revenus du capital sont issus de la base Filosofi et sont observés en 2010. Ces revenus incluent tous les revenus fiscalisés via les données fiscales, mais aussi les revenus exonérés de l'épargne réglementée (livrets, assurance-vie, PEA) par des imputations³¹. Nous classons ensuite les ménages selon leurs revenus du patrimoine en trois catégories : ceux dont les revenus du patrimoine sont inférieurs à la médiane, ceux dont les revenus du patrimoine sont compris entre la médiane et le neuvième décile et, enfin, ceux dont les revenus du patrimoine sont supérieurs au neuvième décile.

Le statut d'occupation du logement. Cette variable indique si le logement est occupé par un propriétaire, un locataire, ou un accédant. Elle est relative à l'année 2010 et est issue des données fiscales.

Le type de ménage. Le type de ménage est obtenu à partir de la base Filosofi. La variable est initialement codée en 9 modalités mais nous la recodons en 4 modalités dans le cadre de cette étude : couple avec un enfant, couple avec deux enfants ou plus, famille monoparentale, et ménages complexes. La variable est relative à l'année 2010.

3 Statistiques descriptives

En cumulant les cohortes 1989 à 1992, nous disposons d'environ 60 000 paires enfants-parents qui représentent 2,1 millions d'enfants (en pondérant les données) après les traitements et les restrictions de champ (cf. supra). Chaque génération comprend environ 10 000 à 18 000 paires enfants-parents (tableau 1), un nombre croissant avec les années de naissance (en raison du fait que les enfants doivent être retrouvés sur la déclaration de leurs parents en 2010, 2011 ou 2012), jusqu'à la génération 1993 (à partir de cette année de naissance, on retrouve moins d'individus dans les données fiscales en 2010 car les mineurs sont moins bien identifiés dans l'EDP). 13 730 paires enfants-parents sont présentes pour la génération 1990, celle qui est utilisée dans l'analyse principale de cet article au niveau national. Des statistiques et analyses supplémentaires sur les effectifs sont proposées dans l'annexe D.

Les filles représentent environ la moitié de notre échantillon. Les parents sont en moyenne âgés d'environ 50 ans (51 ans pour les pères des cohortes 1989 à 1992 et 48 ans pour les mères) et environ 90 % des parents ont un âge compris entre 40 et 60 ans en 2010 (moins de 2 % des mères ont plus de 60 ans). Notre source permet également de documenter le type de familles dans lesquelles vivent les enfants identifiés en 2010 : près de la moitié des enfants de l'échantillon vivent au sein de couples ayant un ou deux enfants tandis qu'un quart d'entre eux vivent au sein de couples ayant au moins trois enfants. Enfin, 21 % vivent au sein de familles monoparentales. Ces pourcentages varient peu selon la génération considérée.

Les revenus individuels médians des enfants en 2018 sont compris entre 17 000 et 22 000 euros selon les générations³² et ceux des parents, observés en 2010, entre 32 000 et 38 000 euros (tableau 2). Les revenus sont en général plus élevés à mesure que les enfants et les parents sont plus âgés (on rappelle que les variables de revenus sont toutes corrigées de l'inflation). C'est pour cela que nous classons les individus par génération. Il faut noter que

31 Les revenus imputés aux propriétaires occupants ne sont pas inclus.

32 Dans les tests de robustesse, le fait de choisir un seuil de revenu proche d'un demi Smic conduit à des revenus moyens environ 5 000 euros plus élevés.

nous n'étudions pas la génération 1993 comme indiqué précédemment, cette génération ayant également les revenus les plus faibles (17 000 et 32 000 euros au niveau de la médiane pour les enfants et les parents respectivement). Le niveau du deuxième décile de revenu des enfants est de 11 100 euros pour le cumul des cohortes 1989-1992 et de 27 100 pour le huitième décile. L'écart entre le seuil des 20 % les plus aisés et celui des 20 % les plus modestes est de 16 000 euros annuels, soit 1 330 euros mensuels. Ce même écart inter-quintile est plus élevé pour les parents : il est de 3 260 euros mensuels. Le pourcentage d'enfants dont les revenus sont positifs ou nuls augmente chaque année pour atteindre près de 100 % en 2018 (figure D1, annexe D). 5 % à 6 % des enfants ne perçoivent pas de revenus selon les générations.

La figure D2 en annexe D exprime le revenu moyen des parents en fonction du centile de revenu des parents : on observe une courbe en forme de tilde avec une quasi droite entre les 20^e et le 70^e centiles, puis une forte hausse pour les derniers centiles (le revenu moyen est d'environ 70 000 euros au 99^e centile et de plus de 100 000 euros au dernier centile). La figure D3 de l'annexe D exprime cette fois le revenu moyen des enfants en fonction des revenus des parents : on observe une forme logarithmique, avec une forte hausse jusqu'au neuvième décile, puis un ralentissement en haut de la distribution.

Enfin des statistiques socio-démographiques sur les parents établies à partir des données du recensement sont présentées dans l'annexe C (tableau C3). 12 % des parents (identifiés dans le recensement) sont immigrés, une proportion similaire pour les pères et les mères. Plus d'un tiers des pères ou mères sont titulaires du BEP, du CAP ou du brevet des collèges ; respectivement 11 % et 17 % des pères et des mères sont titulaires du baccalauréat ; et environ un quart des parents sont diplômés de l'enseignement supérieur. La composition en termes de catégories socioprofessionnelles est très différente selon le sexe des parents. Ainsi, les pères sont plus souvent ouvriers et cadres que les mères tandis que 40 % des mères sont employées.

Tableau 1 : Statistiques descriptives sur la base de données, par génération

Génération	Effectif		Age des parents		Composition familiale, en %		
	Non pondéré	Pondéré	Père	Mère	Couple 1 ou 2 enfant	Couple 3 enfants ou +	Famille monoparentale
1989 (29 ans)	9 709	304 011	53	50	51	23	21
1990 (28 ans)	13 727	423 836	52	49	50	24	21
1991 (27 ans)	17 707	561 655	51	48	49	25	21
1992 (26 ans)	18 496	785 305	50	47	48	26	22
1993 (25 ans)	17 846	709 907	49	46	46	27	22
89-90-91-92 (26-29 ans)	59 639	2 074 809	51	48	49	25	21

Champ : France métropolitaine. Individus rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019 .

Tableau 2 : Quantiles de revenus individuels des enfants et des parents par génération

Génération	Revenus des enfants					Revenus des parents				
	D2	D4	D5	D6	D8	D2	D4	D5	D6	D8
1989 (29 ans)	12 884	19 404	21 823	24 329	32 148	19 682	31 530	37 636	44 397	64 792
1990 (28 ans)	12 167	18 196	20 461	22 772	29 084	19 012	29 718	35 082	40 809	58 703
1991 (27 ans)	11 450	17 307	19 281	21 466	26 828	17 814	28 345	33 700	39 483	55 910
1992 (26 ans)	9 689	15 943	18 000	19 864	24 724	17 042	27 165	32 484	37 898	53 892
1993 (25 ans)	9 188	15 153	17 117	18 948	23 476	16 526	26 281	31 631	37 125	52 307
89-90-91-92 (26-29 ans)	11 063	17 223	19 279	21 472	27 088	18 011	28 570	34 017	39 755	57 130

Champ : France métropolitaine. Individus rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

4 Méthodes

Dans cette section, différentes définitions de la mobilité sont étudiées. Quel que soit l'indicateur retenu, les enfants sont classés en 2018 en fonction de leurs revenus individuels par rapport aux autres enfants de la même cohorte de naissance. Les parents de ces enfants sont ensuite classés en fonction de leurs revenus en 2010 par rapport aux autres parents ayant des enfants de la même cohorte de naissance. Cela permet de comparer les positions des uns et des autres dans leurs distributions respectives des revenus: on parle ainsi de mobilité « positionnelle ». Trois définitions du « revenu des parents » sont utilisées : les revenus individuels moyens des parents rapportés au nombre d'adultes, la somme des revenus des parents rapportés aux unités de consommation, ou le niveau de vie du ménage (cf. *infra*).

4.1 Indicateurs de mobilité

1. La pente rang-rang

Une première façon d'étudier la mobilité intergénérationnelle consiste à comparer les rangs des enfants avec ceux de leurs parents dans leurs distributions respectives des revenus.

Soit R_i^E le rang (en centième) de l'enfant i dans la distribution des revenus des enfants et R_i^P le rang (en centième) du parent i dans la distribution des revenus des parents. On peut régresser le rang des enfants sur le rang de leurs parents comme dans l'équation (1) :

$$R_i^E = C + \beta R_i^P + \varepsilon_i \quad (1)$$

C est la constante de l'équation, elle mesure le rang moyen espéré des enfants des parents les plus modestes. Le coefficient de régression β_{PR} obtenu est parfois appelé dans la suite de l'étude "pente rang-rang". Ce coefficient est égal au coefficient de corrélation entre les deux variables car les distributions des rangs des parents et des rangs des enfants suivent une loi uniforme dont les écarts-types sont identiques par construction ($\beta_{PR} = \text{Corr}(R_i^E, R_i^P)$). Il peut également être appelé « corrélation rang-rang » (« rank-rank correlation » ou RRC en anglais). Si la pente est égale à zéro, le revenu d'activité d'un enfant à l'âge adulte est sans aucun rapport avec la situation de ses parents : son rang est le rang médian, et la mobilité relative des revenus est maximale. Inversement, si la pente est de 1, les revenus d'activité sont uniquement déterminés par ceux des parents et la mobilité est au plus bas. La pente rang-rang mesure ainsi l'association entre la position d'un enfant dans la distribution des revenus (lorsqu'il est jeune adulte) et le rang de ses parents. Elle constitue donc une mesure de la persistance des revenus entre générations. Nous utiliserons ce terme par la suite pour désigner cet indicateur.

La pente rang-rang est une mesure de la mobilité relative car elle mesure la différence de rang entre les enfants de parents à hauts revenus et les enfants de parents à faibles revenus. Le coefficient de régression est étroitement lié à l'élasticité intergénérationnelle des revenus (« Intergenerational income elasticity », IGE), souvent étudiée dans la littérature³³. L'IGE est estimée en régressant le logarithme du revenu ($\log Y_i$) sur le logarithme du revenu des parents ($\log X_i$). Nous n'estimons pas ce coefficient ici car les revenus des parents et les revenus des enfants ne sont pas mesurés au même âge. De plus, la mesure de l'IGE fait l'objet de nombreuses critiques (Chetty et al., 2014). L'IGE est en effet extrêmement sensible aux hypothèses sur les valeurs extrêmes (Acciari et al., 2020), et aux biais d'atténuation et de cycle de vie, contrairement à la pente rang-rang (Nybom and

33 L'IGE diffère de la pente rang-rang dans la mesure où le degré d'inégalité diffère selon les générations. Une augmentation donnée des revenus des parents a un effet plus important sur le niveau des revenus des enfants lorsque le taux d'inégalité est plus élevé entre les enfants qu'entre les parents.

Stuhler, 2017). Enfin, la pente rang-rang est un indicateur de mobilité indépendant du niveau d'inégalité du pays à un instant donné, alors que l'IGE dépend à la fois de la corrélation des revenus parent/enfant et du niveau d'inégalité monétaire relatif entre la génération des parents et celle des enfants. Inversement, la pente rang-rang peut être plus difficile à interpréter en termes de distance sociale parcourue par rapport aux parents car le gain monétaire n'est pas le même pour des enfants de familles pauvres ou pour des enfants de familles dont les revenus sont proches de la médiane (cf. supra). De plus, les comparaisons internationales peuvent être biaisées entre des pays qui ont des niveaux d'inégalités très différents (Acciari et al. (2020) : dans les pays les moins inégalitaires, le gain de revenu nécessaire pour s'élever dans les rangs est plus faible, en termes absolus, que dans les pays les plus inégalitaires car les revenus y sont davantage dispersés.

2. Matrice de transition et mobilité ascendante

Une deuxième façon d'étudier la mobilité intergénérationnelle consiste à mesurer la probabilité, pour un enfant dont les parents appartiennent au premier (c'est-à-dire le plus bas) cinquième de la distribution des revenus des parents, d'appartenir au plus haut cinquième de la distribution des enfants. On appellera cet indicateur « mobilité ascendante », et nous ferons référence au « ratio B20/T20 » ultérieurement (Bottom 20 %/Top 20 %). Plus cet indicateur est élevé, plus la mobilité est forte. Cet indicateur est calculé à partir d'une matrice de transition entre rang des parents et rang des enfants. Celle-ci peut être constituée par cinquième de revenu, dixième de revenu ou même centième de revenu. On peut également mesurer la probabilité, pour un enfant dont les parents appartiennent aux deux premiers cinquièmes (soit parmi les 40 % les plus modestes), d'appartenir aux deux derniers cinquièmes (soit parmi les 40 % les plus aisés). On appellera cet indicateur « mobilité ascendante élargie » (B40/T40). Cet indicateur permet d'intégrer davantage d'individus dans les analyses économétriques ultérieures. Les matrices de transition ont l'avantage d'offrir des représentations graphiques simples et lisibles. Il faut cependant noter que des précautions doivent être prises dans l'interprétation des résultats car les intervalles (dixièmes, ou cinquièmes de revenu) regroupent le même nombre d'individus mais ne couvrent pas la même amplitude en euros. Les intervalles extrêmes sont ainsi plus larges qu'au milieu de la distribution, ce qui implique une mobilité plus forte dans les cinquièmes du milieu qu'au niveau de ses extrémités.

3. La mobilité ascendante absolue : le rang moyen des enfants des parents au premier quartile de revenu

À partir de la spécification économétrique (1), il est possible de proposer un autre indicateur, introduit initialement par Chetty *et al.*, 2014, qui consiste à se demander : "Quel est le rang moyen espéré des jeunes adultes dont les parents appartiennent au 25^e centième de la distribution ? ». La figure E1 (annexe E) montre comment cet indicateur est construit. Si, au niveau national, cet indicateur est lié au précédent (il est égal à $C + 0,25\beta$) et apporte seulement une information plus explicite, il donne au niveau local une information sur la mobilité ascendante absolue. Cet indicateur peut être utilisé au niveau local pour donner une information sur la mobilité absolue (les revenus dans un territoire donné contribuant peu au niveau national, la relation établie au niveau national est utilisable au niveau local). Nous appellerons cet indicateur « mobilité ascendante absolue » (ou parfois mobilité absolue par abus de langage) dans la suite de l'étude, à l'instar de Chetty *et al.*, 2014.

4.2 Discussion sur les biais de cycle de vie et d'atténuation

Le cycle de vie des enfants. La question de la date à laquelle les revenus des enfants sont mesurés est importante étant donné que les données ne permettent de disposer que de neuf années de revenus, comme expliqué plus haut. La littérature montre que les résultats sur la

mobilité sont sensibles à l'âge des enfants à la date à laquelle les revenus sont mesurés (Solon, 1992, Mazumder, 2005, Chetty et al., 2014, Nybom et Stuhler, 2017). Les résultats de ces études montrent que l'IGE est particulièrement sensible à ce biais, tandis que la pente rang-rang et les statistiques sur les matrices de transition le sont beaucoup moins (mais à condition que les revenus des enfants soient mesurés aux environs de 30 ans). Chetty et al. (2014) montrent que la pente rang-rang augmente très fortement entre 22 et 25 ans aux États-Unis, mais se stabilise entre 27 et 30 ans ou un peu après selon les données. En mobilisant les données les plus exhaustives, qui ont néanmoins l'inconvénient de ne pas comprendre d'individus plus âgés que 32-33 ans, les auteurs montrent que l'écart entre les pentes à 28 ans et à 32 ans est d'environ 10 %; tandis qu'à partir des données de la division des statistiques sur les revenus de l'IRS (Internal Revenue Service), l'écart est d'environ 10 % entre 28 et 35 ans et la pente rang-rang est même quasiment égale à 27 ans et à 37 ans. Les auteurs utilisent donc dans leur analyse principale des données relatives à des enfants âgés de 29-30 ans. En France, Kenedi et Sirugue (2021) observent également une forme très concave de la relation entre la pente rang-rang et l'âge, avec une hausse faible après 28 ans et une stabilisation peu après 30 ans, conduisant à un écart de 0,02, soit 8 %, entre les pentes à 28 ans et 35 ans en utilisant les revenus individuels (et de 12 % au niveau des revenus des ménages). Ces auteurs constatent une stabilisation de la mobilité ascendante entre le plus bas et le plus haut cinquième de revenu à 27 ans. De même, Nybom et Stuhler (2017) observent ce résultat un peu après 30 ans en Suède (avec un écart de 20 % entre les pentes à 28 ans et 35 ans entre les pères et les fils) et montrent que la distribution des rangs moyens des enfants à 30 ans en fonction de ceux des parents est proche de celle obtenue en utilisant les revenus des enfants sur l'ensemble de la vie (i.e., leurs revenus permanents). Enfin, Acciari et al. (2020) constatent un effet supérieur du cycle de vie sur les indicateurs de mobilité, ce qui est probablement lié au fait que les jeunes Italiens restent plus longtemps chez leurs parents et entrent plus tard sur le marché du travail. Ces observations laissent penser que nos résultats sur les enfants de 28 ans ne souffrent que légèrement du biais de cycle de vie et que les résultats à 35 ans (ou obtenus à partir du concept de revenu permanent) devraient être plus élevés d'environ 10 %.

Le cycle de vie des parents. En ce qui concerne les revenus des parents, un biais théorique de cycle de vie existe aussi selon l'âge des parents pris en compte. En pratique, Acciari et al. (2021) montrent que la pente rang-rang varie très peu selon l'âge des parents, sauf à partir de 60 ans, un âge à partir duquel la pente décroît fortement, tandis que Kenedi et Sirugue (2021) mettent en évidence un biais à la baisse de la pente rang-rang avant 37 ans. Nous présentons ultérieurement les résultats de tests de robustesse avec des restrictions sur l'âge des parents.

Le biais d'atténuation. Enfin, un biais d'atténuation existe selon le nombre d'années de revenus retenu, pour les parents comme pour les enfants. Théoriquement, ce biais peut résulter d'erreurs de mesure sur les revenus et conduire à une sous-estimation de la mobilité intergénérationnelle (Solon, 1992, Mazumder, 2005, Nybom et Stuhler, 2017). Nybom et Stuhler (2017) et Kenedi et Sirugue (2021) montrent que ce biais est important pour la mesure de l'IGE mais n'affecte quasiment pas la pente rang-rang ou les statistiques des matrices de transition.

4.3 Analyse « toutes choses égales par ailleurs »

Afin de mesurer l'effet de chaque variable indépendamment des autres, nous mettons en œuvre des régressions de Poisson modifiées et des régressions logistiques (cf. infra), dans lesquelles on fait dépendre la variable Mob_i de $\sum \beta_{km} Var_{kmi}$, avec :

- Mob_i , la mobilité ascendante (ou descendante). Il s'agit de la variable à expliquer. C'est une variable binaire 0-1. Plusieurs variables seront utilisées dans cette partie, notamment la probabilité de passer des deux cinquièmes les plus bas aux deux cinquièmes les plus hauts de la distribution des revenus (B40/T40) ou la probabilité de passer du premier cinquième au dernier cinquième (B20/T20).

- Var_{kmi} sont les k variables explicatives, parmi celles décrites dans la partie 2.3. Nous utilisons plusieurs variables, toutes qualitatives, avec plusieurs modalités m . Une modalité de référence doit être fixée pour chaque variable. Nous choisissons en général celle qui est la plus « faible » quand la variable est ordonnée. Dans le cas contraire, nous utilisons la modalité qui correspond au plus grand nombre d'observations (« non Immigré » pour la variable relative à l'ascendance migratoire, les « couples avec un enfant » pour le type de famille, ou « les familles non mobiles » pour la mobilité géographique).

Dans le cas des variables binaires comme la mobilité ascendante, il est usuel d'utiliser des régressions logistiques et de présenter les rapports des cotes, appelés odds ratio, qui sont la simple transformation exponentielle des coefficients estimés par le modèle : $e^{\beta_{km}}$ (voir Afsa, 2016). Cependant, l'interprétation de ces odds ratio est complexe et souvent trompeuse. En effet, la cote correspond à une chance relative mesurée par un rapport de probabilités ($r/(1-r)$ où r est la fréquence de l'événement) et non par une simple probabilité, et l'odds ratio est donc un rapport de probabilités : l'individu de la modalité m de la variable k a $e^{\beta_{km}}$ fois plus de chance de suivre une mobilité ascendante (i.e. $Mob_i=1$) que de ne pas en faire qu'un individu appartenant à la modalité de référence de la variable k . Or, on préférerait connaître plus simplement la probabilité/chance d'effectuer une mobilité ascendante par rapport à la modalité de référence, c'est-à-dire le rapport de probabilité/chance. Dans la littérature, ce rapport est appelé le risque relatif (« relative risk ») et peut être estimé par une régression de Poisson. Nous utilisons la procédure de Zou (2004) permettant de produire une régression de Poisson modifiée robuste à l'erreur de variance. Nous comparerons aussi les résultats à des odds ratios obtenus par des régressions logistiques.

Outre sa simplicité, le risque relatif permet de calculer la fraction de la « population attribuable » (« population attributable fraction » ou PAF) qui correspond, en épidémiologie, à la réduction du nombre de cas (de maladie rare par exemple) si le risque est celui associé à une modalité de référence (probabilité la plus faible d'être atteint par la maladie). Dans notre cas, elle peut s'interpréter, pour une certaine modalité, comme la réduction de la mobilité si celle-ci était celle de la modalité de référence. La formule de la PAF est la suivante :

$PAF(variable) = \sum_i pd_i \frac{RR_i - 1}{RR_i}$, où pd_i représente la proportion de cas totaux dans la population provenant de la $i^{\text{ème}}$ catégorie et RR_i représente le risque relatif pour la $i^{\text{ème}}$ catégorie (par rapport à la catégorie de référence).

Nous mettons également en œuvre un modèle logit polytomique ordonné (avec odds ratios proportionnels). La variable à expliquer est une variable ayant plusieurs modalités (K), comme la mobilité ascendante du premier cinquième aux autres cinquièmes de la distribution, ou les cinquièmes de revenu des enfants. Nous avons alors $(K-1)$ logits et comparons une modalité avec toutes les modalités qui lui sont inférieures. Il faut alors introduire l'hypothèse de pentes identiques des logits, c'est-à-dire que l'effet d'une variable est le même quel que soit le niveau de la variable Y et que les coefficients des variables sont les mêmes dans tous logits, seule la constante étant propre au niveau. Cette hypothèse facilite l'interprétation des résultats.

5 Résultats

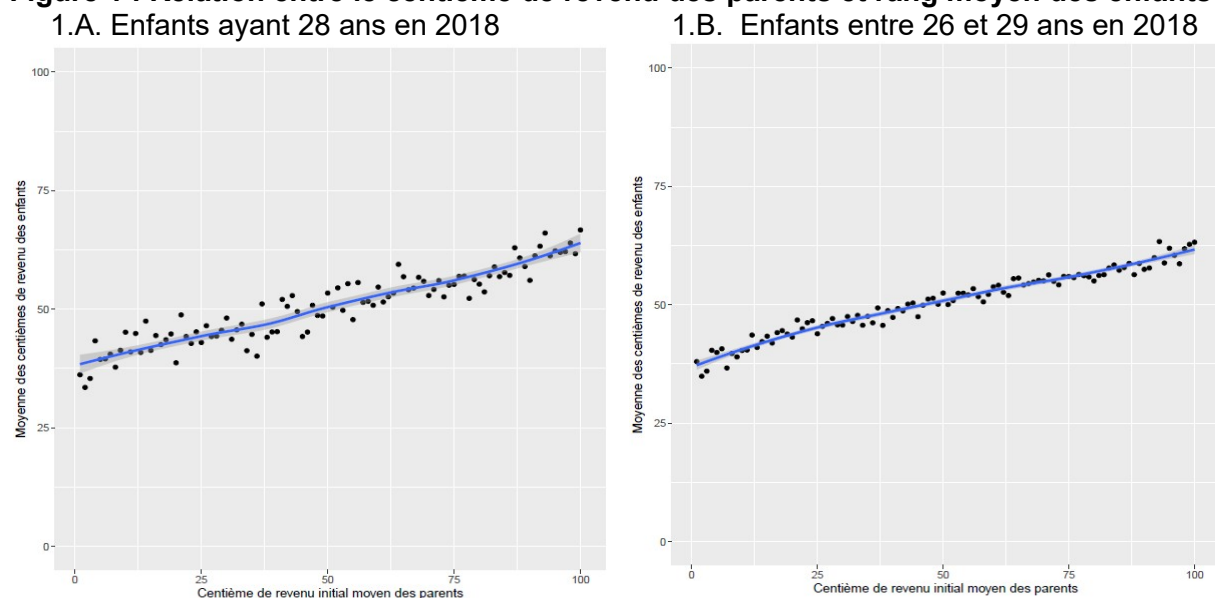
5.1 La mobilité au niveau national

5.1.1 Résultats principaux

Les résultats de cette section décrivent la relation entre les rangs des enfants et les rangs des parents dans leurs distributions respectives des revenus au niveau national. Nous nous intéressons à la génération des enfants nés en 1990 et qui sont âgés de 28 ans en 2018 (l'année d'observation de leurs revenus). Nous commentons d'abord les résultats relatifs à la pente rang-rang, puis ceux qui concernent les matrices de transition.

La relation entre les rangs respectifs des enfants et des parents. La figure 1 montre la relation entre le centième de revenu individuel des parents et le centième moyen de revenu des enfants de 28 ans. Une régression polynomiale locale entre les deux variables peut être estimée. Cette régression non paramétrique, illustrée sur la figure 1, permet d'obtenir une représentation graphique de la relation entre les deux variables. La zone grisée représente un intervalle de confiance de 95 %. La relation est positive et remarquablement quasi-linéaire. Elle n'est cependant pas totalement linéaire : la pente est légèrement plus forte quand les revenus des parents sont très faibles ou parmi les plus élevés. La persistance des revenus est donc légèrement plus forte dans les premiers et dans les derniers centiles de la distribution³⁴. La figure de droite (1.B.) montre la même régression non paramétrique mais sur une population plus large, celle des enfants de 26 à 29 ans. Celle-ci apparaît alors plus nettement du fait de la taille plus grande de l'échantillon mais la pente est très proche de celle obtenue à partir de l'échantillon restreint des individus âgés de 28 ans (cf. infra). La relation est légèrement plus forte en considérant le revenu maximum des parents et non le revenu moyen (figure E2, annexe E).

Figure 1 : Relation entre le centième de revenu des parents et rang moyen des enfants



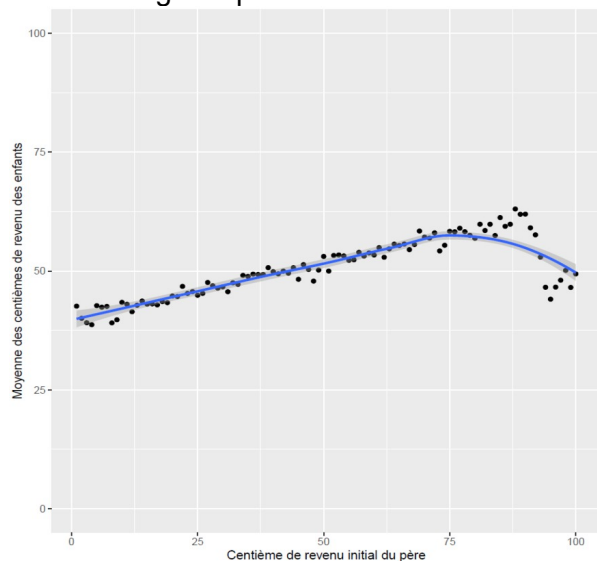
Note : la courbe bleue est une régression polynomiale locale entre les rangs des parents et des enfants.

Champ : France métropolitaine. Individus rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et des revenus positifs ou nuls en 2018.

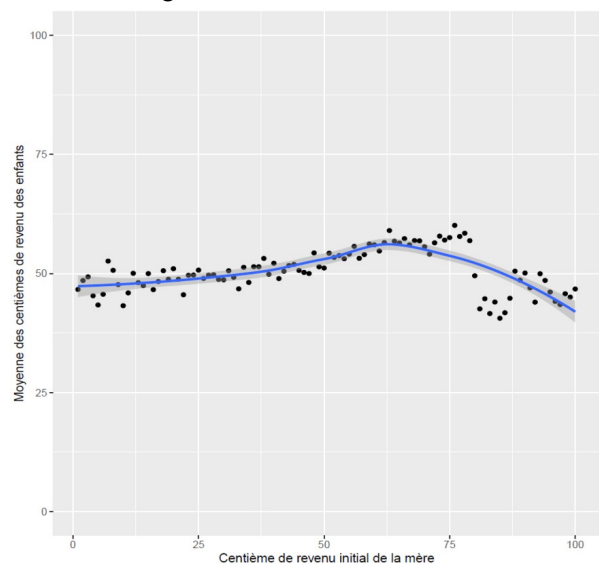
Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, *échantillon démographique permanent 2019*

³⁴ Cela est en partie cohérent avec le fait que l'OCDE (2018) obtient des élasticités IGE qui augmentent avec le niveau des quantiles des revenus des parents en France.

Figure 2 : Relation entre le rang des parents et celui des enfants entre 26 et 29 ans
 2.A. Rang des pères



2.B. Rang des mères



Note : la courbe bleue est une régression polynomiale locale entre les rangs des parents et des enfants. La zone grisée correspond à l'intervalle de confiance.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

La pente rang-rang. La régression linéaire entre les deux variables de la figure 1 donne une pente (ou corrélation) rang-rang de 0,24 : un enfant dont les parents sont classés 10 centièmes de revenu plus haut que ceux d'un autre a en moyenne un classement plus élevé de 2,4 centièmes de revenu. Autrement dit, un enfant de 28 ans dont les parents sont classés tout en haut de la distribution est en moyenne classé entre 2 à 3 déciles plus haut qu'un enfant dont les parents sont situés tout en bas de celle-ci. Ce résultat peut s'expliquer, par exemple, par une probabilité plus élevée de suivre des études supérieures pour les enfants de parents aisés, comme l'ont montré récemment Chetty et al. (2020) aux États-Unis et Bonneau et Grobon (2022) en France. La pente de 0,24 estimée à partir de nos données est très proche de celle que Kenedi et Sirugue (2021) ont obtenue pour la France en mesurant le revenu des enfants au niveau individuel et à l'âge de 28 ans (0,235). La comparaison de la pente rang-rang avec d'autres pays n'est pas évidente car les résultats diffèrent selon le type de revenu pris en compte et l'âge des enfants. D'abord, les concepts de revenus utilisés doivent être similaires aux nôtres (les revenus individuels des enfants) et, en outre, il est nécessaire d'appliquer une hausse d'environ 10 % à nos estimations pour obtenir un chiffre comparable pour les « enfants » âgés d'environ 35 ans (qui est en général l'âge de référence dans la littérature). Ainsi, la pente rang-rang mesurée au niveau du revenu individuel des enfants serait supérieure (et donc la mobilité inférieure) à celle obtenue en Suisse (0,14 selon Chuard-Keller et Grassi, 2021), Suède (0,2 selon Heidrich, 2017 et même inférieure à 0,2 à l'âge de 28-29 ans selon Nybom et Stuhler, 2017), au Danemark (0,20 selon Boserup, Kopczuk, et Kreiner, 2017, et 0,22 selon Helsø, 2021) et au Canada (0,17 selon Corak et Heisz, 1999)³⁵. Notre estimation serait proche de celle obtenue en Italie par Acciari et al. (2021) sur un échantillon d'« enfants » de 36 ans (entre 0,22 et 0,25 selon les hypothèses³⁶), mais serait plus élevée autour de 28 ans. La pente rang-rang serait

35 Les pentes observées dans ces pays sont estimées à partir d'un concept de revenu individuel similaire au nôtre. Mais celles-ci sont mesurées à des âges supérieurs à 30 ans, et seraient donc probablement plus faibles en-dessous de 30 ans.

36 0,25 avec correction de l'évasion fiscale et de données manquantes (et 0,22 sans corrections).

proche de celle obtenue en Australie par Murray et al. (2018) de 0,26³⁷, en rehaussant notre estimation de 10 %. Les résultats sont plus difficilement comparables avec les études définissant le revenu des enfants au niveau de la famille car cela conduit à des estimations plus élevées de la persistance des revenus (cf. infra). C'est le cas de l'Allemagne (Bratberg et al., 2017) et du Danemark (Landersø et Heckman, 2017) où les estimations sont légèrement supérieures à la nôtre. Il fait cependant peu de doute que la persistance des revenus serait plus faible (et donc la mobilité plus élevée) en France qu'aux États-Unis, selon les résultats de Chetty et al. (2014) qui obtenaient un coefficient de 0,31 à 28 ans (voire même 0,34 avec des données de l'IRS). Ce résultat est obtenu par les auteurs en utilisant un concept de revenu familial, mais l'écart est tel qu'il ne semble pas pouvoir être expliqué par la différence de concept de revenu³⁸.

La constante de la régression estimée, qui correspond à l'ordonnée du graphique 2, est de 38,2. Ainsi, le rang moyen des enfants dont les parents sont les plus défavorisés est le 38^e centième. Il est également possible de calculer notre troisième indicateur de mobilité, le rang moyen des enfants dont les parents appartiennent au 25^e centième de la distribution. Nous observons que ces enfants sont, à l'âge de 28 ans, dans le 44^e centième de revenu d'activité en moyenne (voir Figure E1 pour une illustration entre 26 et 29 ans). Cette mobilité ascendante absolue serait proche de celle obtenue par Acciari et al. (2021) en Italie (44), et supérieure à celle obtenue par Chetty et al. (2014) aux États-Unis (41).

La relation entre revenus des parents et revenus des enfants est relativement proche lorsqu'on classe les parents en fonction d'autres définitions du revenu que le revenu individuel moyen. Ainsi, la pente rang-rang est légèrement plus élevée (0,25) lorsqu'on classe les parents par revenu d'activité par UC, et non par adulte, et en termes de niveau de vie (tableau 3). Elle est encore un peu plus élevée en prenant le revenu maximum des parents, et non le revenu moyen (0,26, voir aussi figure E2, annexe E), et un peu plus faible en prenant le revenu déclaré moyen (0,22). Cependant, la relation est beaucoup moins forte en prenant les revenus du père ou de la mère seulement, surtout dans le haut de la distribution (Figure 2). Cela est lié aux situations dans lesquelles, d'une part, les revenus des parents sont très différents et, d'autre part, les revenus des enfants sont en moyenne plus proches du plus haut revenu de leurs parents.

L'analyse économétrique confirme ces résultats (tableau 3) : la pente rang-rang est plus faible quand on utilise le revenu du père (0,18) et non significative avec le revenu de la mère. Décliner cette analyse selon le sexe des enfants est également intéressant : la pente est systématiquement légèrement plus forte pour les filles, quel que soit le type de revenu des parents (notamment par rapport au rang du revenu maximum des parents), tandis que celle entre le rang des mères et le rang des fils est quasi nulle alors qu'elle est positive (mais faible) et significative pour les filles. Notons enfin que si la pente entre le rang des pères et celui des filles et des fils est la même, la constante de l'équation ne l'est pas : le rang espéré des filles dont les parents sont défavorisés est donc plus faible que pour les fils (41 contre 48, au 25^e centième de la distribution). Nous verrons que la mobilité ascendante est également beaucoup plus faible pour les filles que pour les fils : seules 9 % d'entre elles passent du premier au plus haut cinquième de revenu, contre 15 % pour les fils. Les filles sont à l'inverse plus nombreuses à subir une mobilité descendante et à rester dans le premier cinquième (cf. infra).

37 Ces deux estimations sont obtenues sur la base d'un concept de revenu individuel proche du nôtre. Sur un concept de revenu prenant en compte les revenus du capital et les transferts publics, la pente rang-rang en Australie serait plus faible et égale à 0,214. Elle serait de 0,273 en considérant le revenu total au niveau de la famille.

38 Les auteurs obtiennent une pente de 0,28 en utilisant une variable de revenu individuel aux environs de 30 ans.

Tableau 3 : Régression rang-rang selon le type de revenu des parents pris en compte et le sexe des enfants (à 28 ans)

	Dépendant variable:											
	Ensemble des enfants			Revenu initial des enfants (centième)				Fils		Fille		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Revenu initial moyen des parents	0.243*** (0.008)											
Revenu initial des parents par UC		0.245*** (0.008)										
Niveau de vie des parents			0.247*** (0.008)									
Revenu initial max des parents				0.258*** (0.008)			0.244*** (0.012)			0.272*** (0.011)		
Revenu initial du père					0.182*** (0.008)			0.163*** (0.012)			0.200*** (0.012)	
Revenu initial de la mère						-0.005 (0.009)			-0.030** (0.012)			0.021* (0.012)
Constant	38.236*** (0.485)	38.145*** (0.485)	38.017*** (0.485)	37.469*** (0.485)	41.306*** (0.492)	50.760*** (0.500)	41.016*** (0.701)	45.067*** (0.716)	54.886*** (0.723)	33.774*** (0.657)	37.480*** (0.666)	46.438*** (0.683)
Observations	13,523	13,523	13,523	13,523	13,523	13,523	6,747	6,747	6,747	6,776	6,776	6,776
R ²	0.059	0.060	0.061	0.067	0.033	0.00003	0.058	0.026	0.001	0.079	0.043	0.0005
Adjusted R ²	0.059	0.060	0.061	0.066	0.033	-0.00005	0.058	0.026	0.001	0.079	0.042	0.0003
Residual Std. Error	156.782 (df= 13521)	156.707 (df= 13521)	156.605 (df= 13521)	156.147 (df= 13521)	158.918 (df= 13521)	161.615 (df= 13521)	162.427 (df= 6745)	165.162 (df= 6745)	167.253 (df= 6745)	147.836 (df= 6774)	150.708 (df= 6774)	153.987 (df= 6774)
F Statistic	846.894*** (df= 1; 13521)	860.526*** (df= 1; 13521)	879.328*** (df= 1; 13521)	963.932*** (df= 1; 13521)	463.162*** (df= 1; 13521)	0.371 (df= 1; 13521)	413.021*** (df= 1; 6745)	177.878*** (df= 1; 6745)	5.917** (df= 1; 6745)	579.050*** (df= 1; 6774)	301.411*** (df= 1; 6774)	3.297* (df= 1; 6774)

Note:

*p<0.1; **p<0.05; ***p<0.01

Note : Estimation du coefficient β de la pente rang-rang de l'équation (1).

Champ : France métropolitaine. Individus nés en 1990, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFiP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Les résultats sont ainsi comparables, quel que soit le concept de revenu retenu pour les parents, tant que les revenus de chacun des deux conjoints sont intégrés à l'analyse. Par contre, nous ne pouvons pas mesurer les revenus des enfants au niveau du ménage (car même à 28-30 ans une partie des enfants vivent encore chez leur parent, cf. supra). La littérature indique que la pente rang-rang mesurée à partir d'une variable de revenus du ménage pourrait être significativement plus élevée, d'un quart environ (de 18 % selon Chetty et al., 2014, de 28 % selon Murray et al., 2017 et de 21 % selon Kenedi et Sirugue, 2021³⁹). Cela serait lié à une certaine homogamie sociale lors de la mise en couple, un phénomène qui augmente la persistance des revenus (Chetty et al., 2014).

Quantiles de pente rang-rang. L'analyse précédente a été entièrement réalisée en exprimant la moyenne des rangs des enfants en fonction du rang des parents. Cependant, les rangs des enfants sont très hétérogènes à rang des parents fixé. La figure 3 montre différents quantiles de centièmes de revenu des enfants à centième de revenu des parents fixé. Ainsi, la médiane est plus pentue que la moyenne, notamment en haut de la distribution (figure 3.A.). Cette différence est liée au fait que des valeurs extrêmes affectent la moyenne : les cas de mobilités ascendantes extrêmes font augmenter le rang moyen des enfants en bas de la distribution des parents alors que l'essentiel des enfants sont également situés en bas de la distribution. Ce résultat avait déjà été observé par Acciari et al. (2021) en Italie. En régressant la médiane des centièmes de rang des enfants sur les rangs des parents, on obtient une pente de 0,38, c'est-à-dire largement plus élevée qu'en prenant la moyenne des rangs (voir aussi tableau 5 de la partie suivante). Ainsi, la moitié des enfants dont les parents sont classés tout en haut de la distribution sont classés quatre déciles plus haut que la moitié inférieure des enfants dont les parents sont situés tout en bas de la distribution.

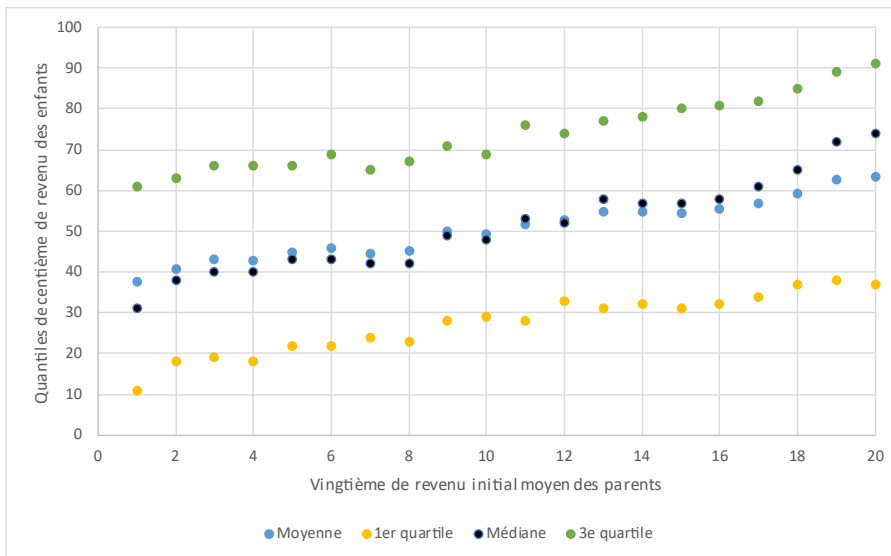
L'étude des autres quantiles de revenus des enfants montre une forte hétérogénéité des rangs des enfants à rangs des parents fixés. L'écart interquartile de la distribution conditionnelle des rangs des enfants, à revenu des parents donné, est de 47 centièmes (égal à celui obtenu par Acciari et al., 2021), et l'écart interdécile de plus de 80 centièmes en haut de la distribution (figure 3.B.). Ainsi, même dans le bas de la distribution des parents, un

39 Chetty et al. (2014) obtiennent une pente rang-rang de 0,339 avec les revenus des enfants au niveau des ménages contre 0,287 au niveau individuel. Ces deux estimations sont de 0,273 et 0,214 selon Murray et al. (2018) en Australie, et de 0,34 et 0,28 selon Kenedi et Sirugue (2021) en France (dans leur principale spécification, et de 0,32 à 0,25 à 35 ans).

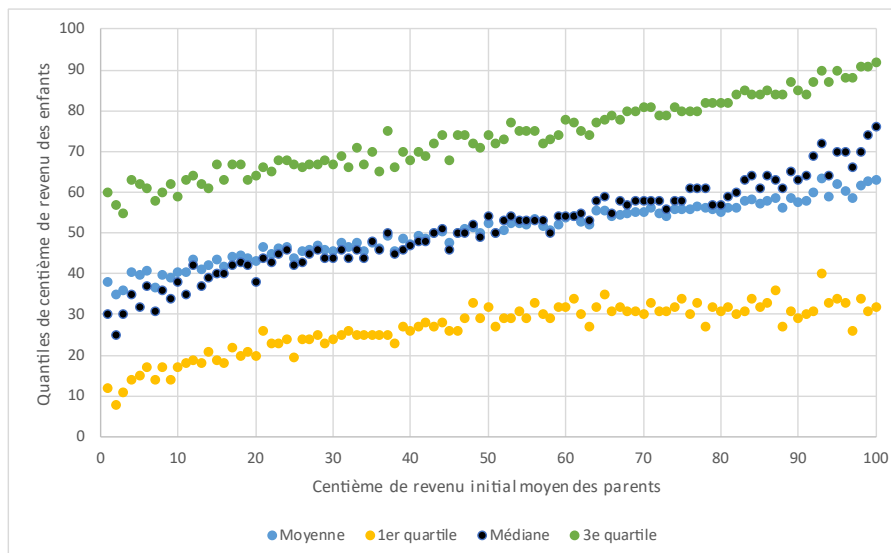
quart des individus dépassent le sixième décile de la distribution des revenus des enfants (et 10 % dépassent le huitième), alors que tout en haut de la distribution des parents, un quart des enfants ont des revenus inférieurs à ceux du troisième décile (et 10 % ont des revenus inférieurs à ceux du premier décile). Cette hétérogénéité des positions et des revenus des enfants avait déjà été observée au niveau de la catégorie sociale des parents par Lefranc et al. (2004). De nombreuses situations de mobilité ascendante et de mobilité descendante sont donc visibles et suggèrent que les revenus des parents ne déterminent pas entièrement les revenus des enfants. Ceux-ci dépendent de nombreux autres facteurs (voir dernière partie pour une quantification de la part liée à l'environnement familial). Notons aussi que les deuxième et quatrième quartiles ont des profils très croissants, tandis que les premier et neuvième déciles ont des profils plus faiblement croissants.

Figure 3 : Relation entre le rang des parents et les quantiles de rang des enfants

3.A. Enfants ayant 28 ans



3.B. Enfants entre 26 et 29 ans



Champ : France métropolitaine. Individus rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Matrice de transition. 72 % des enfants appartiennent à un cinquième de revenu (de la distribution des revenus des enfants) différent de celui de leurs parents. Ce chiffre montre qu'il existe une certaine mobilité dans l'échelle des revenus en France. Parmi les enfants dont les parents appartiennent au premier (plus bas) cinquième de la distribution des revenus initiaux⁴⁰, 31 % figurent également dans le premier cinquième (ce qui crée un « plancher adhérent » en bas de l'échelle des revenus), et 12 % dans le dernier cinquième (Figure 4). Ce taux de 12 % de mobilité ascendante est légèrement plus élevé que celui obtenu par Alesina et al. (2018) et Kenedi et Sirugue (2021) en France, qui l'évaluent à 11 % et 10 % respectivement (pour les enfants ayant plus de 30 ans, i.e. une population différente de la nôtre et à partir de définitions de revenus différentes, cf. infra pour résultats obtenus avec des définitions différentes). Le taux de mobilité ascendante est significativement plus élevé qu'aux États-Unis (7,5 % selon Chetty et al., 2014, et 8 % selon Alesina et al., 2018) mais moins élevé qu'au Canada (13,4 % selon Corak et Heisz, 1999)⁴¹ et qu'en Suède (15,7 % selon Heidrich, 2017). À l'inverse, 34 % des enfants dont les parents appartiennent au dernier cinquième sont eux aussi classés dans le dernier cinquième (ce qui génère un « plafond adhérent » en haut de la distribution), contre un pourcentage de 20 % dans une situation théorique d'égalité parfaite. Ainsi, les enfants issus des familles appartenant aux 20 % les plus aisés ont trois fois plus de chances d'atteindre eux aussi les 20 % les plus aisés (de leur génération) que les enfants des familles appartenant aux 20 % les plus pauvres. Cela montre, d'une part, que les inégalités se reproduisent en partie d'une génération à l'autre et, d'autre part, que les plafonds sont plus adhérents que les planchers (34 % contre 31 %). De façon équivalente, la probabilité de ne pas figurer dans le même cinquième de revenu que celui de ses parents est plus forte dans le premier cinquième que dans le dernier cinquième, et cette proportion est encore plus forte dans le troisième cinquième dans lequel cette proportion atteint 76 %. 15 % des enfants effectuent une mobilité descendante vers le premier cinquième. Ce chiffre est cohérent avec les estimations de Kenedi et Sirugue (2021) qui mesurent les revenus au niveau individuel et à l'âge de 28 ans. Ce pourcentage serait plus faible en utilisant le revenu du ménage et non le revenu individuel de l'enfant, en partie en raison de l'arrêt temporaire de l'activité de l'un des deux conjoints au moment de la naissance d'un enfant au sein du couple⁴².

Enfin, en choisissant une définition plus large de la mobilité ascendante, des 40 % les plus modestes aux 40 % les plus aisés, le taux de mobilité est de 29 %, et en retenant une mobilité définie par quart de population (c'est-à-dire entre les 25 % du bas et les 25 % du haut de la distribution), le taux de mobilité ascendante est de 16,5 %⁴³. Ce dernier chiffre est légèrement supérieur à celui qui est obtenu en Allemagne (15%) par Schnitzlein (2016). À l'inverse, la mobilité entre les 10 % les plus modestes et les 10 % les plus aisés est mécaniquement plus faible, mais reste significative : 5 % des enfants des parents appartenant aux 10 % les plus modestes sont classés parmi les 10 % les plus aisés. 8 % suivent une mobilité descendante extrême en « tombant » de neuf déciles (figure E3), tandis que 24 % restent dans le plus haut dixième de revenu.

40 On considère ici les revenus individuels moyens des parents. D'autres définitions des revenus des parents sont proposées par la suite.

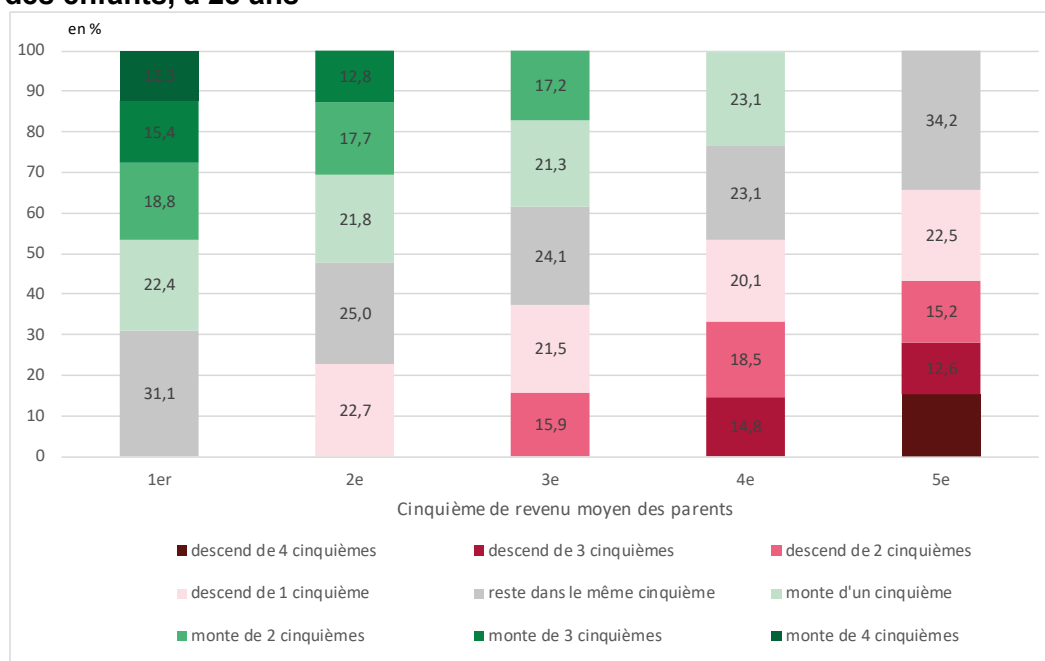
41 Données sur le Canada reportées par Chetty et al. (2014). La comparaison sur l'Allemagne porte sur le taux de mobilité ascendante par quartile (17 % selon nos résultats pour la France contre 15 % en Allemagne). Le pourcentage de mobilité ascendante serait plus proche au Danemark (de 11,7 %) selon Boserup, Kopczuk, et Kreiner (2013), mais les différences de données ne permettent pas de trancher.

42 Le fait que la mobilité descendante soit plus forte pour les femmes va dans le sens de cette explication.

Kenedi et Sirugue (2021) observent par exemple une mobilité descendante de 10 % en retenant une définition des revenus au niveau du ménage.

43 L'OCDE (2018) obtient avec cette définition un taux de mobilité ascendante d'environ 14 %, en utilisant une méthode par imputation qui, selon Acciari et al. (2021), biaise à la baisse la mobilité (voir d'autres références sur le sujet en introduction).

Figure 4 : Mobilité entre cinquièmes de revenus des parents et cinquièmes de revenus des enfants, à 28 ans



Lecture : parmi les enfants dont les parents étaient dans le plus bas cinquième de la distribution de revenus de leur classe d'âge en 2018, 31 % sont dans le plus bas cinquième de la distribution de revenus de leur classe d'âge en 2018, et 12 % dans le plus haut cinquième

Champ : France métropolitaine. Individus nés en 1990, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Ces chiffres restent très proches lorsqu'on considère d'autres variables de revenus des parents à la place des revenus initiaux moyens des parents (tableau 4), ce qui n'affecte donc pas le diagnostic. La mobilité ascendante est cependant légèrement plus faible (11 %) en utilisant le revenu maximum des parents et non le revenu moyen. Elle est également de 11 % en utilisant les niveaux de vie des parents pour les enfants âgés de 26 à 29 ans (cf. partie 5.1.2).

Tableau 4 : Transition entre des cinquièmes de revenus des parents et cinquièmes de revenus des enfants selon la définition des revenus des parents retenue

	Plancher collant (B20/B20)	Mobilité ascendante (B20/T20)	Plafond collant (T20/T20)	Mobilité descendante (T20/B20)
Revenu moyen des parents	31%	12%	34%	15%
Revenu des parents par UC	31%	12%	34%	15%
Niveau de vie	31%	12%	35%	15%
Revenu max des parents	32%	11%	35%	15%

Champ : France métropolitaine. Individus nés en 1990, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

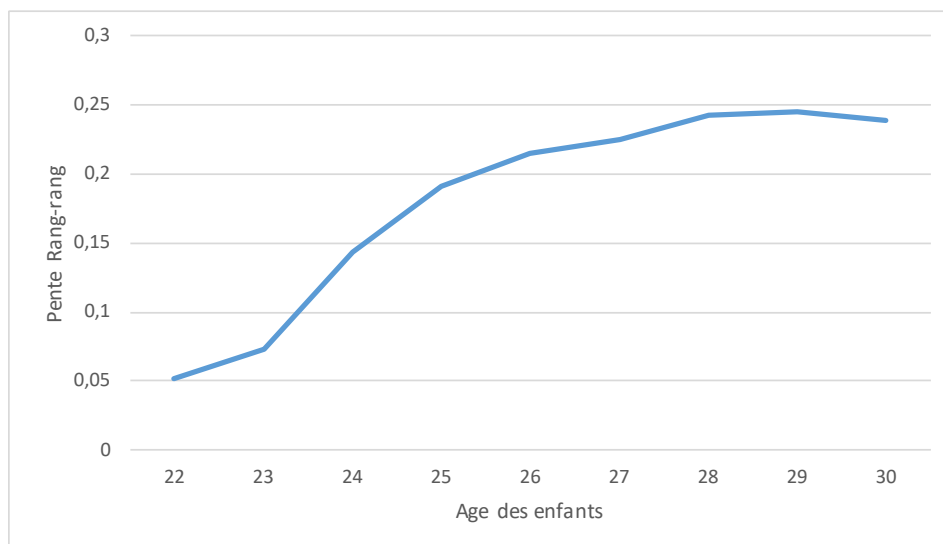
5.1.2 Tests de robustesse

Biais lié au cycle de vie des enfants. Nous commentons dans cette partie la relation entre la pente rang-rang (obtenue par régression) et l'âge des enfants qui apparaît dans la figure 5. La pente rang-rang augmente fortement entre 22 et 25 ans, puis un peu moins entre 26

ans (génération 1992) et 28 ans et se stabilise à partir de 28 ans (génération 1990). Cela conforte notre choix de focaliser notre analyse sur les « enfants » de 28 ans. Il faut noter que les résultats portant sur les « enfants » de 30 ans peuvent être affectés par des effets de sélection et qu'ils doivent être interprétés avec précaution (Annexe D). Les résultats sont très comparables à ceux de la figure 5 en choisissant d'autres définitions des revenus des parents (tableau E1, annexe E). Nos données permettent seulement d'estimer une relation jusque l'âge de 30 ans⁴⁴, mais différentes études montrent un écart d'environ 10 % (voire 20 %) entre les pentes rang-rang à 28 ans et 35 ans (voir partie 4.2), ce qui permet de donner un ordre de grandeur du biais de cycle de vie (cf. supra).

Concernant la mobilité ascendante absolue, le biais de cycle de vie est également visible, mais plus faible que pour la pente rang-rang. En effet, l'effet de l'âge des enfants sur la constante de l'équation va dans le sens inverse de celui observé pour la pente rang-rang. Le rang moyen des jeunes adultes dont les parents appartiennent au 25^e centième de la distribution est de 44 entre 28 et 30 ans et augmente ensuite à mesure que les enfants sont plus jeunes (tableau E1).

Figure 5 : pente rang-rang selon l'âge des enfants



Champ : France métropolitaine. Individus rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Enfin, nous étudions l'effet de ce biais sur les statistiques des matrices de transition. Nous comparons dans le tableau E2 de l'annexe E ces transitions selon la génération (et donc l'âge des enfants), ainsi que pour les générations empilées 1989-1990-1991-1992. Les résultats relatifs aux trois premiers indicateurs sont proches pour les générations 1989 à 1993. Selon les années de naissance, 30 % à 31 % des enfants dont les parents appartiennent au premier cinquième de la distribution des revenus initiaux figurent eux aussi dans le premier cinquième (première colonne), tandis que 11 % à 12 % d'entre eux appartiennent au plus haut cinquième (deuxième colonne). Le pourcentage d'enfants dont les parents appartiennent au plus haut cinquième et qui y restent (plafond adhérent) est également très peu affecté par l'âge des enfants, avec un taux de 33 à 35 % (colonne 3). Par contre, le pourcentage de mobilité descendante est davantage influencé par l'âge des enfants : il baisse avec l'âge en passant de 20 % à 25 ans, à 18 % à 26 ans et 15 % à 28

44 Avec une qualité moindre pour cette génération du fait d'un effet de sélection important (annexe D).

ans. Ces chiffres restent très proches en considérant les niveaux de vie à la place des revenus moyens des parents (tableau E3, annexe E).

Les résultats des parties suivantes, qui utilisent l'empilement des générations 1989 à 1992 entre 26 ans et 29 ans (afin de constituer un échantillon d'individus plus nombreux), concernent principalement la mobilité ascendante absolue et les matrices de transition, qui sont les moins sensibles au biais de cycle de vie à partir de 26 ans comme cela a été expliqué précédemment.

Le cycle de vie des parents. Acciari et al. (2021) montrent que la pente rang-rang varie très peu avec l'âge des parents sauf à partir de 60 ans, un âge à partir duquel la pente décroît fortement. De leur côté, Kenedi et Sirugue (2021) mettent en évidence un biais à la baisse de la pente rang-rang avant 37 ans.

D'après nos données, seuls 1,5 % et 6,9 % des mères et des pères sont respectivement âgés de plus de 60 ans. N'inclure que les parents âgés de 40 à 60 ans augmente très légèrement la pente rang-rang (à 0,25), qui plus est de façon non significative (tableau 5). Nous montrons aussi que contrôler les régressions par l'âge des parents ne modifie pas les résultats. Cela est logique car la corrélation entre l'âge des parents et leur rang en termes de revenu n'est pas significative.

Les revenus des enfants. Comme indiqué dans la partie 2.4, on considère dans cette analyse le revenu individuel et non le niveau de vie des enfants car une proportion importante d'enfants vivent encore chez leurs parents à 28 ans et il n'est donc pas possible de distinguer leurs niveaux de vie de ceux de leurs parents. On peut toutefois restreindre les données aux enfants qui ne vivent plus dans le même logement qu'en 2010 (i.e., celui de leurs parents). Cependant, cette approche ne permet pas d'exclure ceux qui ont déménagé entre 2010 et 2018 et qui résident encore avec leurs parents. Ce faisant, on supprime ainsi 20 % des individus. Sur cette population restreinte, la pente rang-rang est de 0,255 en considérant les revenus individuels des enfants et des parents (en raison de l'effet de sélection évoqué précédemment), et de 0,276 en considérant les niveaux de vie des parents et des enfants. La pente plus élevée (0,02) est liée à l'homogamie sociale lors de la mise en couple. Ces effets d'homogamie devraient être encore plus élevés pour des enfants plus âgés (comme le montrent Kenedi et Sirugue, 2021) car ces derniers devraient être plus souvent en couple qu'à 28 ans.

Les pondérations. Les régressions sont pondérées, ne pas les pondérer ne change quasiment pas les résultats (tableau 5).

Les fratries. 4,5 % des enfants de notre échantillon appartiennent à une fratrie dont au moins deux enfants sont dans notre base de données. La présence de frères et sœurs dans notre échantillon est un aspect original de notre source permettant d'étudier l'effet de la famille sur la mobilité intergénérationnelle (comme nous le montrerons dans la dernière partie). Cependant, celle-ci risque de biaiser les intervalles de confiance de nos estimations car ces observations ne peuvent être traitées comme indépendantes. Pour contrôler ce biais, nous « clusterisons » les régressions au niveau du foyer fiscal en 2010. Comme attendu, nous obtenons les mêmes coefficients dans nos régressions mais, cette fois-ci, les intervalles de confiance sont plus précis. Par ailleurs, nos résultats ne sont pas modifiés par le retrait des fratries de notre échantillon (tableau 5).

Le choix du seuil de revenu minimum. En prenant en compte des seuils minimum de revenus (7 200 euros ou 10 000 euros, cf. partie 2.3), les résultats sont globalement comparables. Le taux de mobilité ascendante et la mobilité absolue au 25^e centième sont peu modifiés par les seuils. Cependant, en introduisant les seuils de revenus dans l'analyse, la pente rang-rang apparaît un peu plus élevée alors que le taux de mobilité descendante est plus faible de deux à trois points.

Tableau 5 : Test de robustesse sur la régression entre le rang des enfants et les rangs des parents

	constante	pente rang-rang
Régression principale (baseline)	38,24	0,243
Régression non pondérée	38,34	0,242
Filtre sur age des parents	38,28	0,247
Régression de la moyenne des centiles des enfants selon le centile des parents	38,23	0,243
Régression de la médiane des centiles des enfants selon le centile des parents	31,71	0,383
Régression principale pour les enfants n'étant plus chez leur parent	40,24	0,255
Régression principale avec le niveau de vie des enfants et des parents pour les enfants n'étant plus chez leur parent	34,66	0,276

Champ : France métropolitaine. Individus nés en 1990, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : *Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019*

5.2 La mobilité selon les caractéristiques individuelles ou celles des ménages

Nous mesurons dans cette partie les indicateurs de mobilité selon différentes caractéristiques, celles qui se rapportent aux individus ou celles qui concernent les ménages. À chaque fois, on utilise les générations empiées 1989-1992 avec un classement par génération, et le revenu moyen des parents pour classer les parents. Les résultats sont détaillés dans le tableau 6. L'analyse de la partie 5.4 rassemblera toutes ces variables au sein d'un même modèle pour mesurer l'effet de chaque variable toutes choses égales par ailleurs.

Sexe. Les hommes ont une probabilité beaucoup plus forte de réaliser une mobilité ascendante et moins forte de faire une mobilité descendante: 15 % des fils dont les parents figurent dans le premier cinquième de la population appartiennent au plus haut cinquième au moment de leur entrée dans la vie active contre 8 % des filles, alors que 34 % des filles restent dans le plus bas cinquième, contre 27 % des fils.

Le patrimoine. La mobilité est-elle plus ou moins forte en fonction du niveau de patrimoine des ménages ? Cette question n'a, à notre connaissance, jamais été traitée empiriquement dans la littérature. Théoriquement, l'effet du patrimoine est ambigu : d'une part, un patrimoine important des parents peut réduire les incitations des enfants à entreprendre de longues études ou à travailler; d'autre part, un patrimoine élevé permet d'accéder à des formations onéreuses (auxquelles n'ont pas ou peu accès les moins aisés) et à un capital social plus important. Nous approximons ici le patrimoine par les revenus du patrimoine et classons ainsi les ménages selon cette variable en trois catégories (supra).

Il apparaît que la mobilité vers le haut de l'échelle des revenus est beaucoup plus importante pour les plus hauts revenus du patrimoine : 24 % des enfants dont les parents ont les plus hauts patrimoines effectuent une mobilité ascendante, contre une proportion de 10 % pour les enfants dont les familles perçoivent des revenus du patrimoine inférieurs à la médiane (et 6 % pour les 40 % les moins bien dotés). Ils sont globalement moins nombreux à rester dans le même cinquième de revenus que celui de leurs parents (69 % contre 74 %). L'espérance de rang des jeunes adultes dont les parents sont au 25^e centième de la distribution est beaucoup plus élevée pour les ménages dotés des patrimoines les plus élevés (52 % contre 43 %). La persistance des revenus est également plus faible pour les hauts revenus du patrimoine. Les résultats sont similaires en utilisant le niveau de vie des parents plutôt que leurs revenus individuels.

Ces chiffres sont à relier aux résultats qui découlent de l'analyse du **statut d'occupation du logement**. Les propriétaires ont ainsi une plus forte probabilité de réaliser une mobilité ascendante (15 %, contre 8 % pour les locataires du secteur social) et plus faible de rester en bas de la distribution et de réaliser une mobilité descendante.

La mobilité géographique. Il est possible d'étudier le lien entre mobilité intergénérationnelle et mobilité géographique en comparant les départements de résidence des familles à la naissance et à la majorité des enfants de notre échantillon. Dans cette étude, on considère comme mobiles les familles qui ont déménagé au moins une fois entre ces deux moments de la vie des enfants. Il apparaît que le pourcentage de mobilité ascendante est plus élevé pour les enfants des familles mobiles (16 %), que la persistance des revenus est plus faible et que la mobilité ascendante absolue espérée des enfants de parents défavorisés est plus élevée. Ces conclusions sont cohérentes avec les résultats de Dherbécourt (2015) qui observe un effet positif de la mobilité géographique (mais au niveau régional seulement) sur la mobilité ascendante en termes de catégories sociales.

L'ascendance migratoire des parents. L'ascendance migratoire des parents joue également sur la mobilité. A priori, le fait d'être descendant d'immigré peut impliquer deux effets contraires : (i) d'un côté, les enfants d'immigrés peuvent souffrir de discriminations et vivent parfois dans des territoires pauvres, ce qui peut réduire leur mobilité. Mais (ii), d'un autre côté, plusieurs facteurs peuvent favoriser une plus forte mobilité : une forte volonté d'élévation sociale, un effet de composition lié à la localisation (ils résident souvent dans les plus grandes villes et les territoires les plus dynamiques) et un effet lié à l'inadéquation entre la rémunération de leurs parents et leurs compétences réelles qui peut leur donner plus de chances de gravir l'échelle des revenus, par exemple. En France, le deuxième groupe de facteurs semble l'emporter : les enfants d'immigrés ont en moyenne une probabilité plus forte de réaliser une mobilité ascendante par rapport à leurs parents (15 %, contre 10 % pour les enfants qui ne sont pas issus de l'immigration). La persistance des revenus des enfants d'immigrés est plus faible et leur mobilité plus élevée, mais la mobilité ascendante absolue espérée des enfants dont les parents appartiennent au premier quartile de la population est proche de celle des enfants qui ne sont pas issus de l'immigration. De même, Abramitzky et al. (2021) montrent que la mobilité ascendante est plus forte pour les immigrés que pour les non immigrés aux États-Unis depuis plus d'un siècle. Selon ces auteurs, récemment, la plus forte mobilité des immigrés vient pour moitié de leur localisation. Ils montrent ainsi que, après avoir contrôlé la localisation, la plus forte mobilité des immigrés de deuxième génération est liée principalement à l'inadéquation entre la rémunération et les compétences des pères (notamment pour ceux qui ont émigré plus tardivement que dans la petite enfance, du fait d'une absence de maîtrise de la langue qui les empêche de trouver un travail adapté à leurs qualifications). Ce facteur peut également être lié au manque de réseau social. Le rôle de l'éducation pourrait aussi jouer mais n'a pas été démontré. Les enfants d'immigrés ont cependant une probabilité plus forte de rester dans le plus bas cinquième des revenus (33 % contre 27 %) et d'effectuer une mobilité descendante (24 % contre 17 %).

Le continent de naissance des parents. Parmi les enfants d'immigrés, les enfants dont la personne de référence de la famille (très souvent les pères) est née en Asie ont une plus forte probabilité de mobilité ascendante (16 % quand on considère le continent des « personnes de référence », et 17 % en considérant le continent du père). La mobilité ascendante élargie B40/T40 est la plus élevée pour les enfants de parents venant d'Asie, d'Amérique et d'Europe (38 % en moyenne, contre 27 % en Afrique Sub-Saharienne). Le rang espéré des enfants dont les parents appartiennent au premier quartile est ainsi très différent selon les continents : 49 pour l'Europe et l'Amérique, 47 pour l'Asie, 43 pour l'Afrique du Nord et 39 pour l'Afrique sub-saharienne. Ces résultats sont cohérents avec ceux obtenus sur les données américaines d'Abramitzky et al. (2021).

Les diplômes des parents. La mobilité varie beaucoup selon le diplôme de la personne de référence du ménage (définie comme celui des deux parents ayant le plus haut revenu, les résultats étant comparables lorsqu'on utilise la variable de diplôme du père) : la mobilité ascendante est de 17 % pour les enfants dont les parents sont titulaires de diplômes de l'enseignement supérieur et titulaires du baccalauréat (et 21 % pour les mères titulaires de

diplômes du supérieur), contre 11 % pour ceux dont les parents sont titulaires de diplômes inférieurs au baccalauréat et 10 % pour ceux dont les parents n'ont pas de diplômes. La mobilité ascendante absolue est également plus forte pour les plus diplômés, tandis qu'on observe une plus grande persistance des revenus pour les enfants de non diplômés. Ces résultats sont également cohérents avec Dherbécourt (2015). La probabilité de mobilité ascendante est encore plus élevée pour les enfants d'immigrés diplômés du supérieur : elle est de 20 %, contre 9 % pour les enfants d'immigrés non diplômés.

Tableau 6 : Indicateurs de mobilité selon différentes variables

	Nombre d'individus	Mobilité asc. B20/T20 (en %)	Mobilité asc. B40/T40 (%)	Mobilité desc. T20/B20 (%)	Plancher collant B20 (%)	Même cinquième (%)	Ordonnée à l'origine (C)	Pente rang-rang (β)	Mobilité absolue
Sexe (2010)									
Homme	29463	15,1	35,6	17,7	27,4	73,7	43,2	0,20	48,2
Femme	29351	8,3	23,7	16,7	33,6	72,3	34,9	0,25	41,2
Revenus du capital (2010)									
Inf. a D5	28389	9,6	26,5	18,0	32,8	73,2	36,8	0,23	42,5
D5-D9	24316	15,4	36,1	17,2	24,0	73,6	45,1	0,16	49,0
Sup. a D9	6109	24,2	41,9	16,8	24,6	69,3	48,6	0,14	52,1
Statut d'occupation du logement (2010)									
Propriétaire	42473	14,5	33,6	16,9	24,4	73,6	43,1	0,18	47,6
Locataire secteur social	7196	8,2	23,0	19,5	37,9	70,9	33,1	0,26	39,6
Locataire secteur privé	9090	9,8	26,2	19,7	34,3	72,2	36,4	0,22	42,0
Mobilité géographique pendant l'enfance									
Immuable	26594	11,8	30,8	16,8	28,7	73,4	40,1	0,22	45,6
Mobile	13736	16,0	33,9	17,6	29,1	72,7	41,6	0,19	46,4
Statut migratoire (pers. de ref., 1999-2012)									
Non immigré	27370	10,5	30,0	16,7	27,4	72,8	40,2	0,22	45,6
Immigré	2877	15,0	33,4	24,4	33,4	73,2	40,9	0,18	45,4
Non renseigné	28567	11,7	28,5	16,7	32,3	73,0	37,6	0,24	43,7
Continent de naissance des immigrés (pers. de ref., 1999-2012)									
Afrique du Nord	790	14,3	29,8	31,7	38,8	68,7	37,8	0,20	42,7
Afrique sub	178	12,7	27,1	31,7	33,7	73,5	36,8	0,10	39,3
Asie	346	16,3	38,1	26,7	29,9	72,4	42,6	0,17	46,9
Europe & Amérique	733	15,1	38,4	23,9	21,7	78,0	46,3	0,10	48,8
Autre immigré	812	16,0	34,4	18,7	31,3	73,7	42,7	0,19	47,5
Diplôme (pers. de ref., 1999-2012)									
Sans diplôme	4658	9,8	26,0	15,2	34,3	72,1	36,6	0,22	42,0
Diplôme inf. au bac.	12454	10,6	30,0	15,0	25,7	73,8	41,3	0,20	46,2
Bac. ou équivalent	4404	16,5	38,3	16,4	23,2	74,9	46,4	0,13	49,7
Diplôme sup. au bac.	8628	17,1	37,9	18,0	29,4	71,1	42,9	0,19	47,6
Non renseigné	28670	11,7	28,4	16,7	32,4	73,0	37,6	0,24	43,6
Catégorie socio-professionnelle (pers. de ref., 1999-2012)									
agriculteurs	721	13,5	35,9	13,3	18,5	78,7	47,3	0,16	51,3
indépendants	2193	16,0	34,8	15,2	22,8	74,3	44,6	0,15	48,4
cadres	5981	23,4	41,4	17,5	21,8	68,8	46,3	0,15	50,1
prof. int.	7485	11,6	33,4	17,5	24,5	74,8	43,2	0,18	47,6
employés	4901	11,6	28,7	18,0	29,0	73,6	41,1	0,17	45,3
ouvriers	7897	9,7	28,6	14,3	31,2	72,7	37,8	0,26	44,2
autres	966	10,3	27,6	18,5	37,7	73,4	34,9	0,22	40,3
Non renseigné	28670	11,7	28,4	16,7	32,4	73,0	37,6	0,24	43,6
Type de ménage (2010)									
Couple 1 ou 2 enfants	29761	12,7	32,0	16,8	27,0	73,4	41,4	0,21	46,6
Couple 3 enfants ou plus	14792	12,2	31,1	15,5	28,6	71,5	39,4	0,24	45,5
Famille monoparentale	11260	10,0	25,0	20,5	35,3	73,8	35,9	0,21	41,1
Ménages complexe	3001	10,6	25,4	19,4	36,9	73,1	36,8	0,20	41,9
TOTAL	58814	11,6	29,6	17,2	30,6	73,2	38,8	0,23	44,6

Lecture : 15 % des hommes font une mobilité ascendante, 74 % restent dans le même cinquième que leur parent. Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

La catégorie socioprofessionnelle des parents. La mobilité ascendante est la plus élevée pour les enfants de cadres⁴⁵ et la plus faible pour les enfants d'ouvriers (23 % contre 10 %). Les enfants de parents qui sont indépendants ont également une probabilité de mobilité ascendante plus élevée que la moyenne. Le rang espéré des enfants des parents appartenant au premier quartile est très élevé pour les cadres (50).

45 Ces cas pourraient être liés en partie à des chocs temporaires sur le revenu des parents cadres.

Le type de famille. Enfin, les indicateurs de mobilité ne sont pas très différents selon le type de famille de l'enfant, sauf pour les familles monoparentales. La mobilité ascendante est légèrement plus faible pour les familles monoparentales (et les ménages complexes), de même que la mobilité ascendante absolue espérée, tandis que la probabilité de mobilité descendante et la probabilité de rester dans le plus bas cinquième sont plus élevées pour ces familles. Enfin, la persistance des revenus est plus élevée pour les enfants des familles nombreuses (plus de 3 enfants). Les familles monoparentales (et les familles nombreuses, dans une moindre mesure) sont susceptibles de connaître plus de difficultés financières que les autres, ce qui peut entraver le cursus scolaire des enfants et, par conséquent, leurs revenus futurs.

5.3 La mobilité au niveau des territoires

5.3.1 Régions et départements

On décline dans cette section une série de résultats aux niveaux des régions et des départements de résidence des familles des enfants en 2010, c'est-à-dire lorsque ceux-ci étaient âgés d'environ 18-20 ans. On étudie également la mobilité en fonction du territoire de naissance (ces données sont néanmoins moins bien renseignées).

Tableau 7 : Indicateurs de mobilité et d'inégalité par région

	Nombre d'individus	Mobilité asc. B20/T20 (en %)	Mobilité asc. B40/T40 (%)	Mobilité desc. T20/B20 (%)	Plancher collant B20 (%)	Même cinquième (%)	Ordonnée à l'origine (C)	Pente rang-rang (β)	Mobilité absolue
Auvergne-Rhône-Alpes	6925	10,8	32,1	15,9	30,4	73,5	32,1	0,23	46,0
Bourgogne-Franche-Comté	2526	11,1	33,5	14,8	27,7	73,7	33,5	0,21	47,8
Bretagne	3205	10,4	29,4	16,6	22,8	72,4	29,4	0,20	46,3
Centre-Val de Loire	2240	9,9	28,5	14,0	27,6	75,5	28,5	0,22	45,1
Grand Est	5375	8,8	27,5	17,9	35,5	72,4	27,5	0,27	42,0
Hauts-de-France	6307	7,3	23,8	15,6	35,1	70,8	23,8	0,28	41,2
Île-de-France	11831	20,8	37,9	17,2	30,1	72,1	37,9	0,19	48,5
Normandie	3192	8,4	29,0	18,4	29,8	74,5	29,0	0,25	44,1
Nouvelle-Aquitaine	4686	9,9	27,4	20,0	26,6	74,5	27,4	0,15	44,5
Occitanie	4718	9,9	25,6	20,2	32,0	73,1	25,6	0,22	42,3
Pays de la Loire	3502	12,4	32,8	14,7	18,3	74,6	32,8	0,16	47,9
Provence-Alpes-Côte d'Azur	4107	11,0	27,0	19,1	35,0	73,5	27,0	0,24	42,7

Note : Données pas assez significatives pour la Corse.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

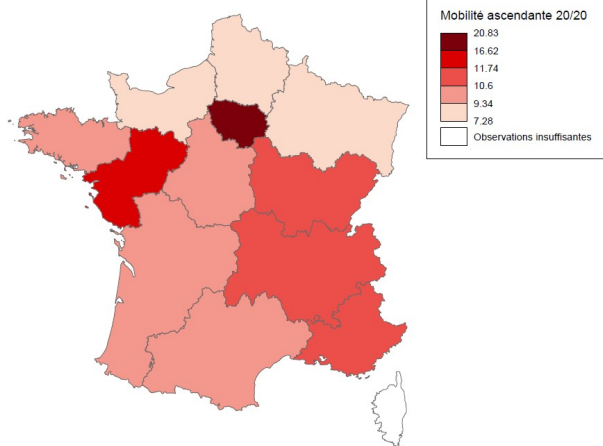
La mobilité ascendante est la plus élevée en Île-de-France (21 %, tableau 7) et la plus faible dans les régions du Nord de la France : les Hauts-de-France (7 %), la Normandie (8 %) et le Grand Est (9 %). Au niveau des départements, la mobilité ascendante est la plus élevée dans les Yvelines, à Paris, en Haute-Savoie et en Seine-et-Marne, et la plus faible dans les Deux-Sèvres, le Puy-de-Dôme, l'Eure, l'Allier et le Pas-de-Calais (Figure 6.B). Les résultats sont similaires à ceux qui sont énoncés ci-dessus lorsqu'on considère la mobilité ascendante étendue B40/T40, à part que la Normandie est mieux classée et l'Occitanie moins bien classée. Les résultats relatifs à la mobilité ascendante absolue des enfants des familles modestes sont légèrement différents : les personnes dont les parents sont au 25^e centième peuvent en moyenne espérer atteindre le 49^e centième de revenu en Île-de-France, contre le 41^e dans les Hauts-de-France. Cet indicateur montre que le Grand Est et l'Occitanie sont également moins bien classés que les autres régions (Figure 6.C)⁴⁶. À l'inverse, les régions Bourgogne-Franche-Comté et Pays de la Loire sont, avec l'Île-de-France, les régions où la mobilité absolue est la plus forte. Dans ces régions, la mobilité absolue est la plus élevée en Haute-Savoie, suivie du Val d'Oise et de la Saône-et-Loire. Ces résultats confirment ceux de

46 Cela est surtout lié à la faiblesse de la constante estimée dans l'équation 1 : une personne dont les parents sont au tout en bas de la distribution a en moyenne moins de chance d'évoluer vers le haut dans ces régions.

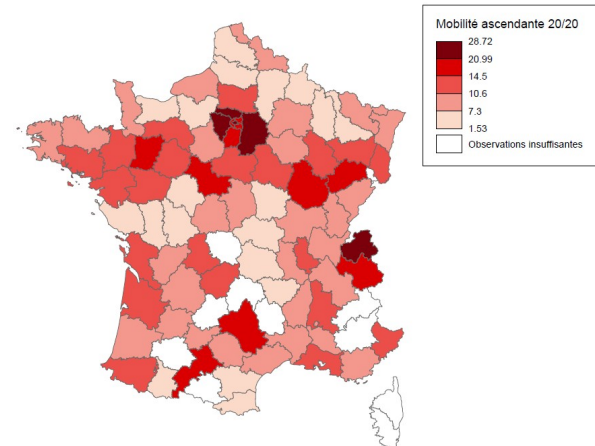
Dherbécourt et Kenedi (2020) et Poncelet *et al.* (2015) qui montrent que les perspectives de niveau de vie des enfants d'ouvriers ou d'employés sont plus élevées en Île-de-France. En considérant l'indicateur de mobilité ascendante élargie (B40/T40), l'Île-de-France reste la région la plus mobile (mais dans une ampleur moins marquée), tandis que l'Occitanie et la Nouvelle-Aquitaine rejoignent les Hauts-de-France comme régions les moins mobiles.

Figure 6 : Mobilité par région et par département

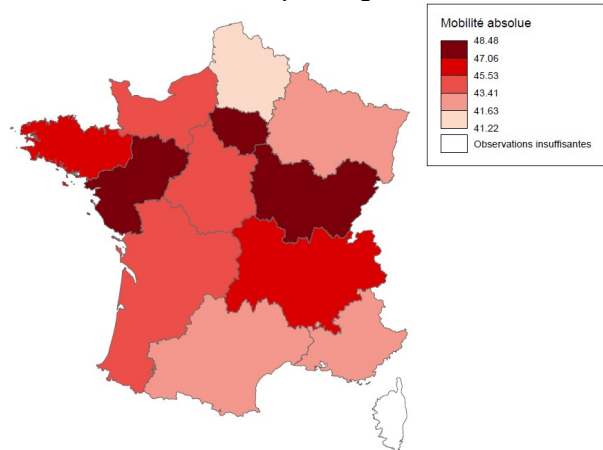
6.A. Mobilité ascendante par région



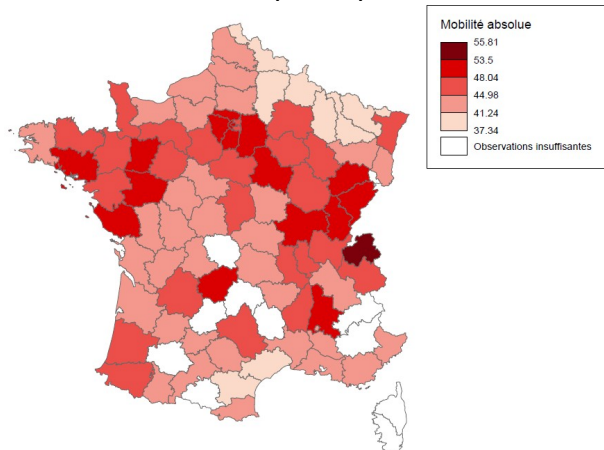
6.B. Mobilité ascendante par département



6.C. Mobilité absolue par région



6.D. Mobilité absolue par département



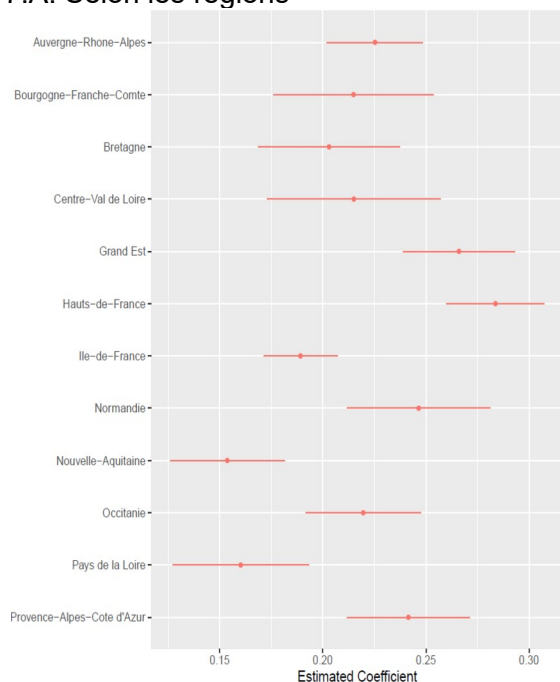
Note : la mobilité en 6.C. et 6.D. correspond à la mobilité ascendante absolue.
 Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.
 Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

On observe également que la persistance des revenus est plus faible (inférieure à 0,20) en Île-de-France, en Nouvelle-Aquitaine, et dans les Pays de la Loire (tableau 7). Elle est en revanche la plus élevée dans les Hauts-de-France (0,28) et dans le Grand Est (0,27). La persistance des revenus est significativement plus élevée dans ces deux régions, en Normandie et en PACA, qu'en l'Île-de-France (région où la persistance des revenus est la plus faible, cf. figure 7.A). Les intervalles de confiance sont en général plus larges par département, du fait que les échantillons d'individus dont nous disposons sont plus petits : nous sélectionnons alors les départements dans lesquels le nombre de paires parents-enfants est supérieur à 120 et pour lesquels les intervalles de confiance sont inférieurs à 0,2. La pente rang-rang est relativement faible en Haute-Savoie, en Charente-Maritime, dans le Morbihan et en Seine et Marne, et relativement forte dans les départements de l'Aisne, de

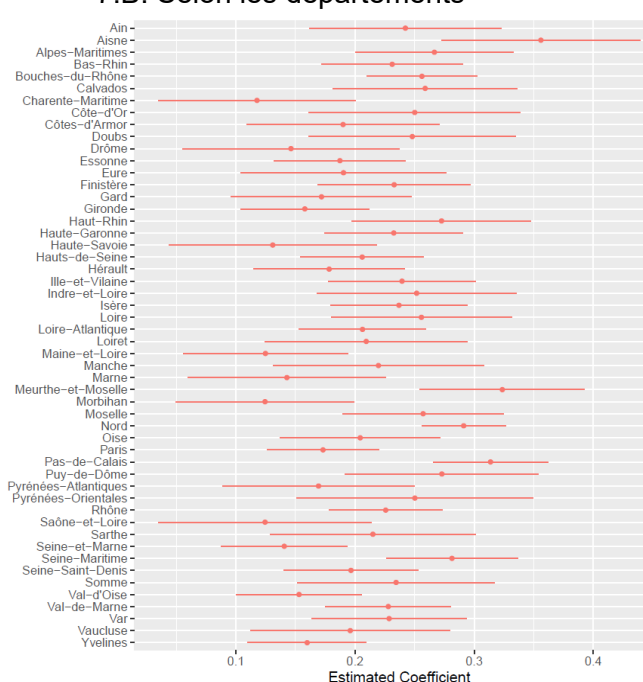
Meurthe-et-Moselle, du Pas-de-Calais et du Nord (figure 7.B). La pente rang-rang est de même niveau en Seine-Maritime, dans les Bouches du Rhône et dans le Haut-Rhin.

Figure 7 : Pente rang-rang par région et par département

7.A. Selon les régions



7.B. Selon les départements



Note : Les barres représentent l'intervalle de confiance de la pente rang-rang. Données pas assez significatives pour la Corse et certains départements (effectifs inférieures à 120 ou intervalles de confiance supérieures à 0,2).
 Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

En considérant la région de naissance, et non plus la région de résidence en 2010, le constat reste le même à quelques différences près (voir tableau E4 annexe E). La mobilité est toujours la plus élevée en Île-de-France mais l'écart se resserre avec les régions Pays de la Loire et Auvergne-Rhône-Alpes. De plus, la mobilité ascendante élargie est désormais supérieure à la moyenne dans le Grand Est et la persistance des revenus est moins forte que dans les résultats obtenus avec la région de résidence en 2010. En outre, la mobilité absolue des enfants des familles modestes est la plus élevée en Auvergne-Rhône-Alpes.

Il est enfin intéressant d'étudier la relation potentielle entre mobilité et inégalité monétaire. Chetty et al. (2014) et Corak (2013) observent une relation négative entre mobilité et inégalité monétaire, dite « courbe de Gatsby » (au niveau des États américains pour les premiers auteurs et au niveau des pays pour Corak). Cependant, aucune corrélation entre mobilité et inégalité monétaire⁴⁷ n'est visible selon les découpages administratifs au niveau des régions et des départements (voire même selon d'autres découpages territoriaux, cf. infra), comme l'indique la figure E4 de l'annexe E. Cela est cohérent avec la revue de littérature de Durlauf et al. (2021) sur la courbe de Gatsby. Au contraire, on observe que certaines des zones qui ont les taux de mobilité les plus forts sont celles où les inégalités sont les plus élevées (communes denses, très peuplées et Paris). La forte mobilité en Île-de-France est surtout liée au niveau de richesse du territoire, comme le montre la partie suivante à partir d'un modèle explicatif de la mobilité ascendante. Ainsi Kenedi et Sirugue

47 Telles que mesurées à partir de l'EDP. Mais nous avons vérifié que les résultats étaient qualitativement inchangés en utilisant les données de la base Filosofi (dans sa version exhaustive).

(2021) mettent en évidence une relation négative entre inégalité et mobilité en utilisant la mobilité absolue et en contrôlant d'autres caractéristiques territoriales, comme le revenu. Une corrélation positive plus nette apparaît entre mobilité ascendante et revenu médian du territoire (voir figure E5 annexe E). Cette corrélation positive est également forte entre la mobilité ascendante et le niveau du neuvième décile de niveau de vie⁴⁸. Cela peut être lié au fait que les territoires les plus riches sont les plus attractifs et offrent davantage d'opportunités d'emploi et que la mobilité ascendante y trouve un terrain favorable. Cela est cohérent avec le fait que les perspectives de revenus à l'âge adulte augmentent avec le niveau de richesse du territoire d'origine (Dherbécourt et Kenedi, 2020).

5.3.2 Densité et catégories de communes

Le tableau 8 décrit des indicateurs de mobilité selon différents zonages géographiques (densité de la commune d'habitation, tranche d'aire d'attraction des villes et catégories d'aires). Ces deux dernières variables sont issues des nouveaux découpages en aires d'attraction des villes.

La mobilité ascendante est plus forte dans les aires les plus peuplées, notamment dans l'aire de Paris. Ainsi, c'est dans l'aire de Paris que les personnes dont les parents sont situés au 25^e centième de la distribution des revenus peuvent en moyenne espérer atteindre le centième le plus élevé. Comme cela a été expliqué dans la partie précédente, ce phénomène peut être lié au fait que les opportunités d'emploi sont plus nombreuses dans les grandes agglomérations. À l'inverse la mobilité ascendante est plus faible dans les aires les moins peuplées. On observe dans le même temps que la persistance des revenus est plus faible dans les communes denses et dans l'aire de Paris (voir aussi figure E6, annexe E). Si le coefficient de Paris est significatif par rapport aux autres tranches d'aires urbaines, on ne peut rien dire sur la persistance des revenus dans les communes « hors attraction des villes » étant donné l'étendue de l'intervalle de confiance. Ainsi, si des différences existent selon la densité ou le type de commune, elles sont plus faibles (en dehors de l'effet lié à l'aire de Paris) qu'au niveau des régions (voir aussi Dherbécourt et Kenedi, 2020).

A l'intérieur des zones urbaines, la mobilité ascendante est un peu plus faible dans les « communes-centres » et la persistance des revenus y est plus forte, notamment par rapport aux communes situées dans la couronne de ces villes (figure E6). En lien avec la persistance des revenus, la proportion de personnes restant dans le cinquième du bas est plus élevée dans les communes-centre.

Tableau 8 : Indicateurs de mobilité selon le territoire d'appartenance

	Nombre d'individus	Mobilité asc. B20/T20 (en %)	Mobilité asc. B40/T40 (%)	Mobilité desc. T20/B20 (%)	Plancher collant B20 (%)	Même cinquième (%)	Ordonnée à l'origine (C)	Pente rang-rang (bêta)	Mobilité absolue
Tranche d'aire d'attraction des villes									
Commune hors attraction des villes	3339	10	32	16	25	75	32	0,20	46
Aire de moins de 50 000 habitants	6640	9	26	18	31	73	26	0,24	43
Aire de 50 000 à moins de 200 000 habitants	10783	9	29	16	29	73	29	0,22	44
Aire de 200 000 à moins de 700 000 habitants	13195	10	28	19	32	73	28	0,24	44
Aire 700 000 habitants ou plus (hors Paris)	12173	11	28	17	33	73	28	0,24	43
Aire de Paris	12684	20	37	17	30	72	37	0,19	48
Catégories d'aires d'attraction des villes									
Commune-centre	13534	10	26	20	36	71	26	0,26	42
Autre commune d'un pôle	13990	15	32	17	33	72	32	0,23	45
Commune de la couronne	27349	11	30	16	26	74	30	0,20	46
Commune hors attraction des pôles	3339	10	32	16	25	75	32	0,20	46
Densité de la commune									
Communes denses	19303	12	31	15	24	75	31	0,18	47
Communes de densité intermédiaire	17834	9	28	17	33	73	28	0,26	43
Communes peu denses	21677	14	30	18	34	72	30	0,24	44

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

48 C'est en Île-de-France que la mobilité, le revenu médian et le neuvième décile sont les plus élevés, puis en Auvergne-Rhône-Alpes et en Provence-Alpes-Côte d'Azur.

Enfin, on observe une corrélation positive entre mobilité ascendante et revenu médian du territoire, comme dans la partie précédente (figure E7). De même que par département, il ne semble pas exister de courbe de Gatsby décroissante en France selon ces découpages géographiques et, par conséquent, pas de corrélation entre mobilité et inégalité.

5.4 Analyse de la mobilité « toutes choses égales par ailleurs »

Les résultats sur la mobilité ascendante déclinés selon les variables exposées dans les parties précédentes (par territoire, et selon les caractéristiques individuelles et celles des ménages) sont-ils liés à des effets de composition ? Pour répondre à cette question, il est nécessaire de mener une analyse « toutes choses égales par ailleurs ».

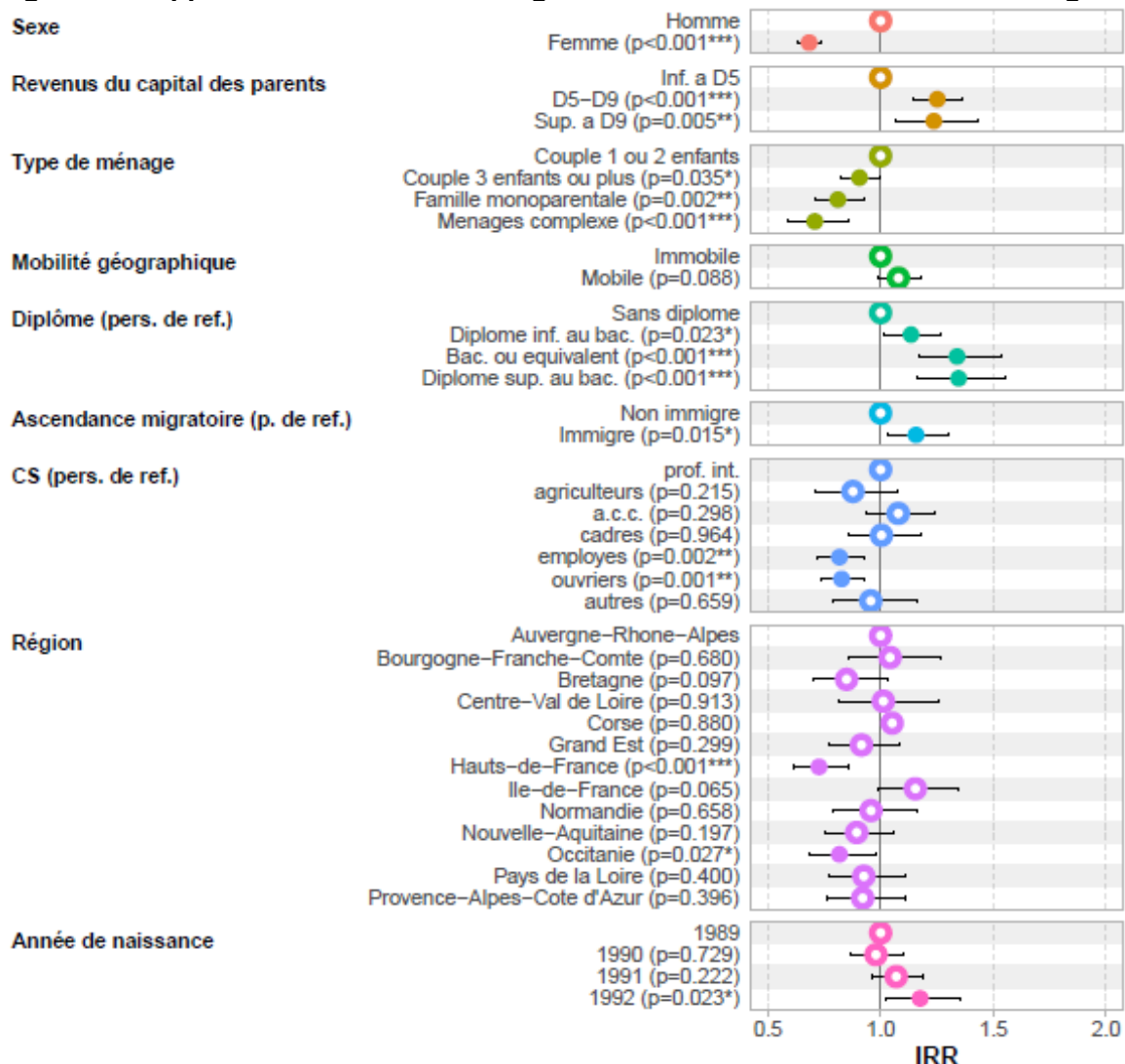
5.4.1 Analyse de la mobilité ascendante

Nous mettons en œuvre une régression (équation 2, partie 4.3) expliquant la mobilité ascendante à partir des principales variables explicatives présentées précédemment. Pour améliorer la robustesse de notre analyse, nous considérons une définition de la mobilité ascendante un peu plus large que celle présentée précédemment (la mobilité dite « élargie ») : la probabilité de réaliser une mobilité des 40 % les plus modestes aux 40 % les plus aisés. Cette définition plus souple nous permet de constituer un échantillon plus important d'enfants ayant réalisé des mobilités ascendantes (plus de 5 600 observations, contre 2 400 dans le cas de l'indicateur B20/T20) et de réduire les intervalles de confiance. Nous nous restreignons aux individus ayant des informations dans le recensement de la population ou les enquêtes annuelles de recensement (EAR, voir annexe B). La partie suivante présente des résultats complémentaires en fonction des différentes définitions de la mobilité, des restrictions de population, du type de régression et de la pondération retenus.

La figure 8 présente les rapports de probabilités (ou de chances) de mobilité ascendante (appelés aussi risques relatifs, voir partie 4.3) obtenus par régression, ainsi que les intervalles de confiance associés. La modalité de référence est à chaque fois la première modalité de la figure. Ainsi, les femmes ont 1,5 fois moins de chance de faire une mobilité ascendante que les hommes (et près de deux fois moins en choisissant l'indicateur B20/T20 ou une régression logistique, voir partie suivante). De même, la probabilité de mobilité ascendante est plus faible dans les familles monoparentales et les ménages complexes (et dans une moindre mesure dans les familles nombreuses) que dans les couples avec un ou deux enfants. À l'inverse, le fait que les parents disposent de revenus du capital élevés favorise la mobilité : être parmi les 10 % les plus riches en termes de revenu du capital augmente de 1,3 fois la probabilité de mobilité ascendante élargie (et 1,5 fois pour l'indicateur B20/T20). Ce résultat est plus faible que celui de l'analyse descriptive mais reste très significatif. Les enfants dont le parent au plus haut revenu (souvent le père) est titulaire du baccalauréat ou d'un diplôme de l'enseignement supérieur ont significativement plus de chances de réaliser une mobilité ascendante (environ 1,3/1,4 fois). En outre, les enfants dont la famille a été mobile géographiquement dans l'enfance ont également plus de chances (mais dans une moindre mesure). Les enfants dont le parent a le plus haut revenu est immigré ont significativement plus de chances de gravir l'échelle des revenus que les non immigrés, mais aussi un risque plus élevé de faire une mobilité descendante (1,4 fois), ce qui est conforme aux résultats de l'analyse descriptive. Les effets des catégories socio-professionnelles sur la mobilité sont cependant différents de ceux présentés dans l'analyse descriptive : le fait d'être cadre n'est pas significativement associé à une mobilité ascendante plus élevée par rapport aux professions intermédiaires. En effet, la forte probabilité des cadres observée dans l'analyse descriptive était liée en partie à d'autres variables corrélées avec la catégorie socioprofessionnelle. À l'inverse, la mobilité ascendante est significativement plus faible pour les enfants d'ouvriers et d'employés que pour les enfants

de parents exerçant des professions intermédiaires dans l'analyse principale. Finalement, ces résultats confirment l'analyse descriptive précédente sur les déterminants de la mobilité ascendante mais atténuent l'influence de la catégorie sociale.

Figure 8 : Rapport des chances de la régression de la mobilité ascendante élargie



Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité ascendante des 40 % du bas de la distribution des revenus des parents vers les 40 % du haut de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{er} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018, présents dans les EAR ou le RP.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Les résultats relatifs aux régions de résidence sont cohérents mais légèrement moins polarisés que dans l'analyse descriptive. La probabilité de mobilité ascendante est significativement plus élevée en Île-de-France qu'en Auvergne-Rhône-Alpes, et significativement plus faible dans les Hauts-de-France et en Occitanie (toujours par rapport à la région Auvergne-Rhône-Alpes). Le rapport des chances de l'Île-de-France est plus faible que dans l'analyse descriptive, ce qui est lié à l'indicateur utilisé : avec l'indicateur de mobilité ascendante B20/T20 le rapport des chances est de près de 2, c'est-à-dire proche de ce qu'on observe dans l'analyse descriptive (et avec un intervalle de confiance plus faible

pour la population totale, voir partie suivante). Résider en Île-de-France favorise en fait davantage la mobilité extrême, du plus bas au plus haut cinquième. Cet effet peut être lié au fait que la région Île-de-France est la plus riche : l'attractivité et les opportunités d'études supérieures et d'emplois suscitent en moyenne une mobilité ascendante plus élevée. On a en effet observé dans la partie descriptive précédente une corrélation positive entre richesse du territoire et mobilité (mais pas de corrélation négative entre inégalité et mobilité).

Pour déterminer si les différences de mobilité par régions sont des effets propres ou des effets tirés par le revenu moyen de la région (par un phénomène d'attractivité et d'opportunité), on ajoute à la régression le cinquième de revenu médian de la commune ou de l'aire urbaine dans laquelle réside l'enfant en 2010. En ajoutant cette variable, l'Île-de-France ne se distingue plus significativement mais la probabilité de mobilité apparaît toujours plus faible dans les Hauts-de-France (figure E8 de l'annexe E) et dans d'autres régions sur l'ensemble de la population (figure E9). Les signes des coefficients des autres régions sont également proches de ceux obtenus sans variable de contrôle du revenu de la zone d'habitation. Ainsi, l'effet propre des territoires demeure malgré le contrôle par le niveau de richesse, sauf pour la région Île-de-France dont les résultats positifs seraient uniquement liés à son niveau de richesse, toutes choses égales par ailleurs.

Enfin, nous remplaçons dans l'équation les indicatrices de régions par d'autres variables pour caractériser les territoires : les tailles d'aires d'attraction des villes, le PIB de la région et le taux de diplômés du supérieur du département. Il en résulte que le rapport des chances est supérieur à 1, et qu'il existe donc une plus forte probabilité de mobilité dans les territoires où les taux de diplômés et le PIB par habitant sont les plus élevés, tandis que les autres modalités et la variable de type d'aire d'attraction des villes ne sont pas significatives (figure E10 de l'annexe E).

Le poids des différents facteurs dans l'explication de la mobilité ascendante est ensuite dérivé des fractions attribuables à la population (FAP), sur la base des estimations de la figure 8 (en modifiant les modalités de référence pour obtenir des RR positifs). Le rôle de la région d'origine est prépondérant, puisque la mobilité ascendante diminuerait de 25 % si tous les enfants vivaient dans les Hauts-de-France (situation de référence). Les rôles du sexe et des diplômes des parents seraient également importants avec des FAP d'environ 20 %.

5.4.2 Tests de robustesse

Dans cette partie, nous présentons les résultats de différents tests de robustesse sur la mobilité ascendante (les tests de robustesse sur la mobilité descendante de la partie suivante conduisent aux mêmes résultats).

Les variables d'intérêt. Le tableau E8 (annexe E) présente les résultats de plusieurs régressions comprenant différents indicateurs de mobilité. Dans la première colonne figurent les coefficients d'une régression de la variable de mobilité ascendante mesurant la probabilité de passer du cinquième le plus bas au cinquième le plus haut de la distribution (voir figure E12 de l'annexe E également). La variable expliquée dans la deuxième colonne du tableau correspond à la probabilité de passer des 40 % les plus modestes aux 40 % les plus aisés (mobilité dite « élargie », cf l'analyse économétrique ci-dessus). Dans la troisième colonne, on cherche à expliquer la probabilité de monter de deux cinquièmes ou plus dans l'échelle des revenus : du premier cinquième au troisième ou plus, du deuxième cinquième au quatrième ou plus, et du troisième cinquième au dernier cinquième (voir figure E13 de l'annexe E). Enfin, dans la quatrième colonne, on propose un modèle multinomial ordonné restreint aux enfants dont les parents appartiennent au premier cinquième et dans lequel le score est d'autant plus haut que les enfants sont dans un cinquième élevé (voir figure E14 de l'annexe E).

Les principaux enseignements de ce tableau sont les suivants. Appartenir aux 10 % les plus riches en termes de revenus du capital a un impact plus important sur le premier indicateur de mobilité (B20 /T20) que sur les autres. L'effet négatif lié au fait de vivre dans une famille monoparentale est également plus élevé en ce qui concerne le premier indicateur et est toujours significatif, sauf pour le troisième indicateur. Le fait que la personne de référence ait un diplôme supérieur au baccalauréat est un facteur positif et significatif partout, sauf pour le premier indicateur (non significatif). Avoir fait une mobilité géographique dans l'enfance est un facteur significatif, sauf pour le troisième indicateur. Le coefficient associé à la modalité « être enfant de cadres » n'est significatif nulle part, tandis que le fait d'avoir un parent indépendant a une incidence significativement positive sur la mobilité pour les deux derniers indicateurs. Enfin, résider en Île-de-France a un effet près de deux fois plus important sur les premier et quatrième indicateurs que sur le deuxième. Ainsi, le coefficient de ce facteur est toujours significativement positif et cette variable de localisation s'avère encore plus déterminante en matière de mobilité extrême.

Nous avons également effectué une régression linéaire de la variable de mobilité construite à partir de la différence entre les rangs des parents et les rangs des enfants (en cinquièmes et en dixièmes de revenu, voir tableau E5). Cette régression contrôle le rang des parents pour prendre en compte la plus forte probabilité de réaliser une mobilité ascendante quand les revenus des parents appartiennent à la partie basse de la distribution. Les résultats sont proches de ceux indiqués précédemment, à ceci près que le fait d'être descendant d'immigré et que vivre dans une famille de trois enfants ou plus ne sont plus des facteurs significatifs. Le R^2 de l'équation est de 0,4 (i.e., 40 % de la variance de la mobilité est expliquée par le modèle).

Champ et traitement des données manquantes. Les résultats sont également légèrement différents si on prend en compte l'ensemble de la population, et non plus seulement les individus pour lesquels les variables du recensement de la population sont renseignées. Dans ce cas, il faut ajouter des modalités « manquantes » pour les variables issues du RP et des EAR, ce que nous faisons dans la figure E15 pour la mobilité ascendante élargie et la figure 9 ci-dessous pour l'indicateur B20/T20⁴⁹. Par rapport à la figure 8, les différences sont les suivantes : les rapports des chances des modalités relatives aux hauts revenus du capital et aux pères diplômés du supérieur sont un peu plus élevés en tenant compte de la population totale (sur l'indicateur B40/T40) ; le coefficient associé à la modalité « indépendant » du père devient significativement supérieur à 1 ; le rapport des chances de la région Île-de-France est un peu plus élevé. Pour l'indicateur B40/T40, les coefficients de la Nouvelle-Aquitaine, de l'Occitanie et de la Provence-Alpes-Côtes-d'Azur deviennent significativement inférieurs à 1 (vivre dans ces régions diminue la probabilité de mobilité par rapport à la région Auvergne-Rhône-Alpes) et sont très légèrement supérieure à 1 avec l'indicateur B20/T20.

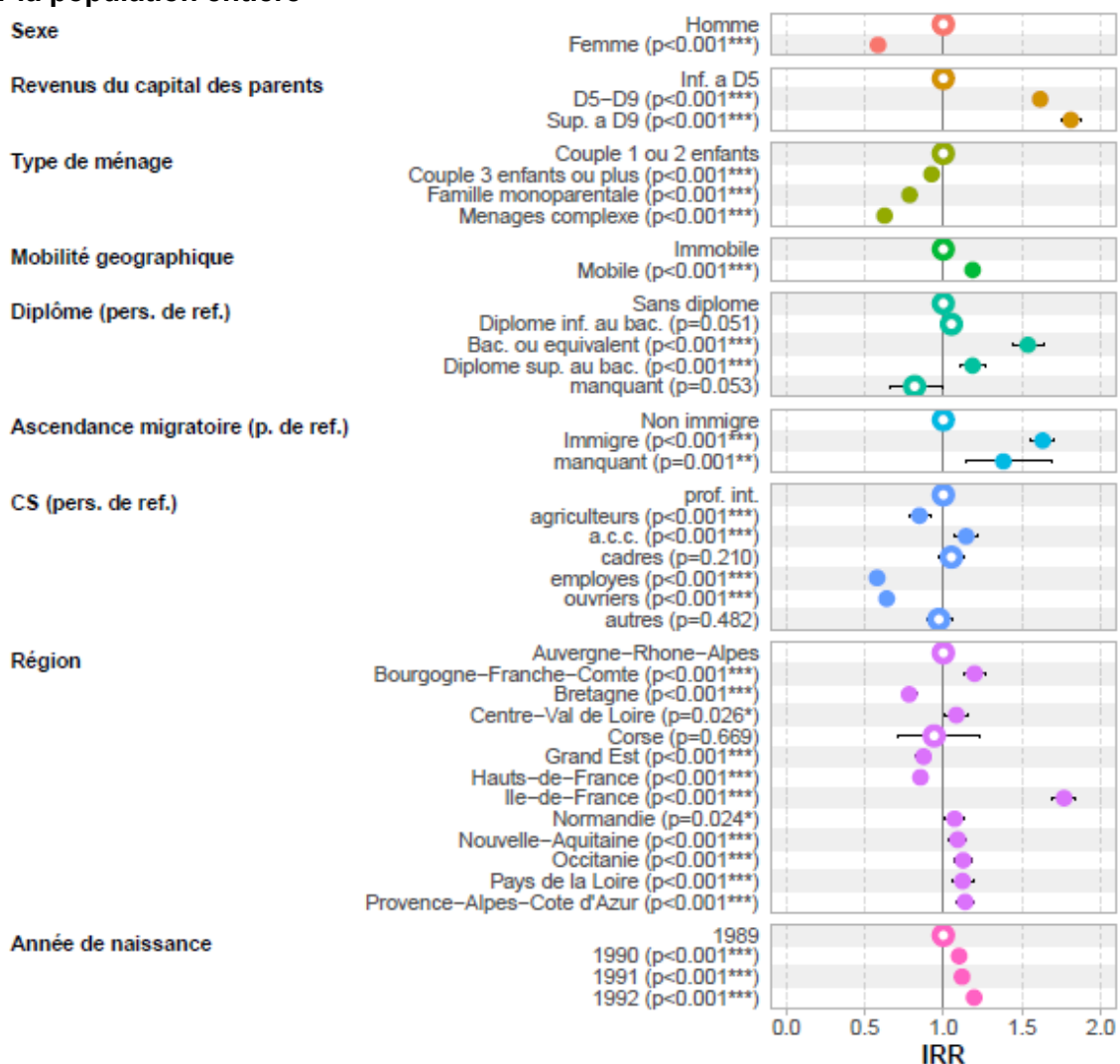
Régressions logistiques et régressions de Poisson. Comme exposé dans la partie 4.3, il est classique d'estimer des régressions logistiques et de commenter des odds ratio dans ce type d'analyse. Nous présentons principalement les coefficients issus de régressions de Poisson pour faciliter l'interprétation des résultats mais nous montrons dans les figures E16 et E17 comment les résultats sont modifiés en utilisant des régressions logistiques pour les indicateurs B20/T20 et B40/T40. Ainsi, les odds ratios sont systématiquement plus élevés que les rapports de chance issus des régressions de Poisson, et logiquement d'autant plus avec l'indicateur B40/T40. Avec cet indicateur, par exemple, les enfants de parents diplômés du supérieur ont un rapport des cotes de 1,6 de réaliser une mobilité ascendante par rapport aux non diplômés, contre un rapport de chance (risque relatif) de 1,3 avec la régression de Poisson.

49 Qui souffre le plus de la sélection liée au RP.

Régressions pondérées ou non pondérées. Faut-il pondérer les régressions ? Cette question classique n'a malheureusement pas de réponse claire. Davezies et D'Haultfoeuille (2009) montrent que la robustesse et la précision de l'estimation dépendent des variables utilisées dans le calage et le plan de sondage de la base, et de la sélection des observations. Lorsque ces informations ne sont pas totalement disponibles, il est recommandé de comparer les estimations avec et sans pondération. C'est ce qui est fait dans la figure E18 de l'annexe E, où on teste ces deux choix (la spécification de base intègre les pondérations). Celle-ci montre que les résultats sont très proches, et même égaux, en arrondissant à un chiffre après la virgule. Le degré de significativité des coefficients est le même.

Âge des parents. Dans la figure E19, on contrôle l'équation par l'âge des parents : les résultats sont quasiment inchangés.

Figure 9. Rapport des chances de la régression de la mobilité ascendante (B20/T20) sur la population entière



Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité ascendante des 20 % du bas de la distribution des revenus des parents vers les 20 % du haut de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{er} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

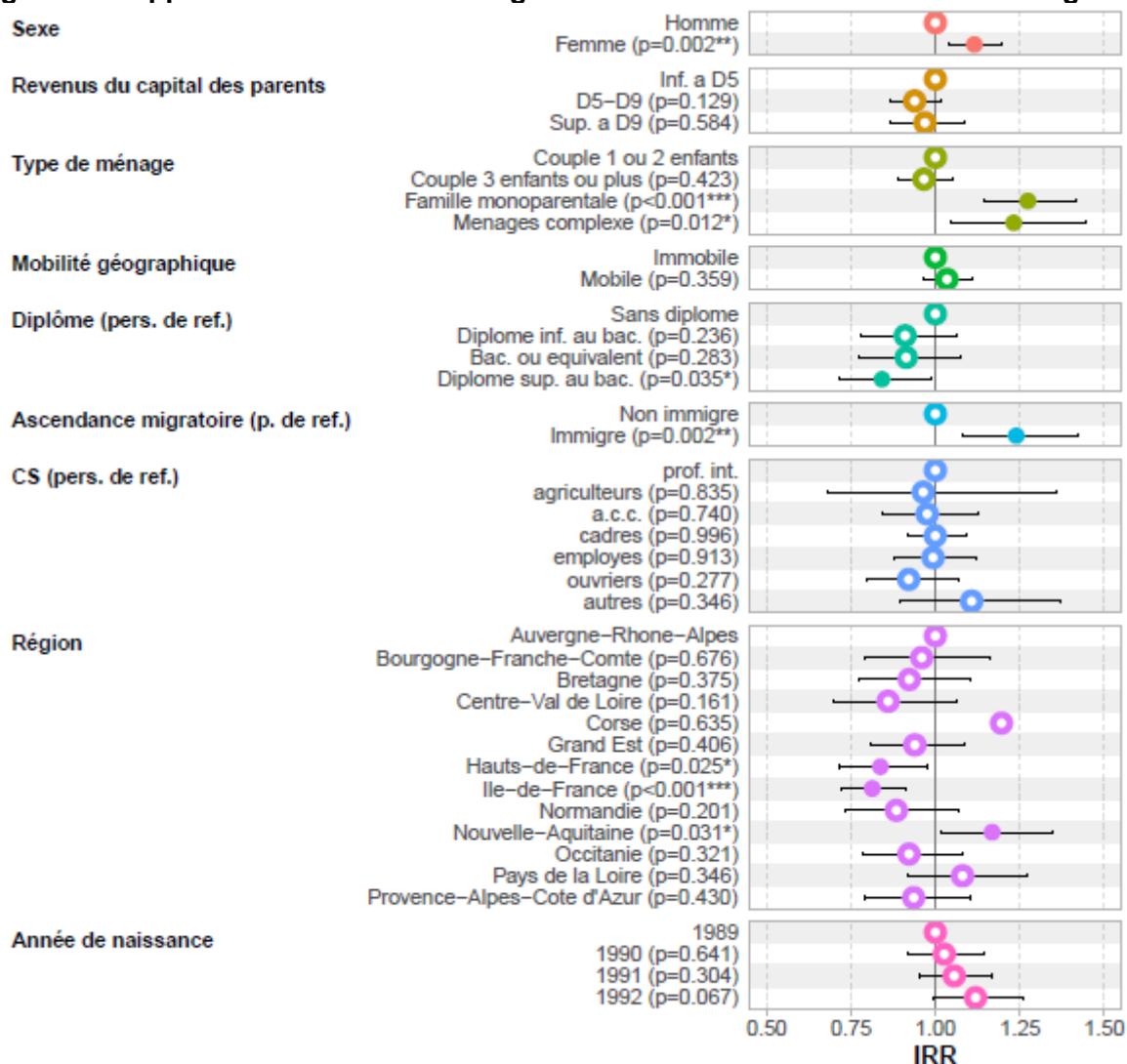
On peut conclure de cette partie que les résultats présentés dans la partie précédente sont robustes. Ils sont proches et plus précis (écarts-types plus faibles) en prenant en compte la population totale et supérieurs et significativement différents de 1 en ce qui concerne l'indicateur B20/T20. Cependant, les conclusions restent qualitativement les mêmes quant à l'importance de chaque variable.

5.4.3 Résultats sur la mobilité descendante

Nous mettons en œuvre une régression de Poisson de même type que la précédente pour expliquer la mobilité descendante (des 40 centiles les plus hauts aux 40 centiles les plus bas), à partir des principales variables explicatives présentées précédemment. Les conclusions sont, en général, le reflet des résultats relatifs à la mobilité ascendante, à quelques différences près.

Les femmes ont 1,2 fois plus de chances que les hommes de réaliser une mobilité descendante. De même, la probabilité de descendre dans l'échelle des niveaux de vie est plus forte dans les familles monoparentales et les ménages complexes (et, dans une moindre mesure, dans les familles nombreuses). À l'inverse, le fait que les parents disposent de revenus du capital élevés diminue un peu la mobilité, mais cet effet est non significatif et beaucoup moins fort que dans le cadre de la mobilité ascendante. Les enfants dont le parent aux revenus les plus élevés (souvent le père) est titulaire d'un diplôme de l'enseignement supérieur ont significativement moins de risque de faire une mobilité descendante, tandis que l'effet des autres diplômes est non significatif. Contrairement à la mobilité ascendante, une mobilité géographique pendant l'enfance n'affecte pas la mobilité descendante. Les enfants dont le parent doté du plus haut revenu est immigré ont une probabilité significativement plus élevée de descendre dans l'échelle des revenus. Ce résultat n'est donc pas le symétrique de celui sur la mobilité ascendante. Les coefficients des modalités relatives aux catégories socioprofessionnelles sont tous non significatifs. En ce qui concerne les régions de résidence, la probabilité de descendre dans l'échelle des revenus est significativement plus élevée en Nouvelle-Aquitaine qu'en Auvergne-Rhône-Alpes, et significativement moins élevée en Île-de-France et dans les Hauts-de-France. La figure E20 expose les résultats obtenus sur la population totale (c'est-à-dire non restreinte aux individus qui figurent dans les données du recensement de la population) et présente l'indicateur de mobilité descendante T20/B20.

Figure 9 : Rapport des chances de la régression de la mobilité descendante élargie



Note : La figure présente un rapport de risques (RR) de réaliser une mobilité descendante des 40 % du haut de la distribution des revenus des parents vers les 40 % du bas de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{ère} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018, présents dans les EAR ou le RP.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

5.5 Quel rôle de l'environnement familial pour expliquer le rang des enfants ?

Le but de cette dernière partie est de mieux comprendre le rôle du contexte familial dans la formation des revenus des enfants. Dans un premier temps, on compare l'importance respective des revenus des parents, d'un côté, et des autres caractéristiques socio-démographiques, de l'autre, pour expliquer le revenu des enfants. Dans un deuxième temps, on mesure la part des revenus liée au milieu familial au sens large en utilisant les fratries.

5.5.1 Méthode par régression

Une première façon de répondre à ces questions importantes consiste à mesurer l'effet respectif de plusieurs variables sur le rang des enfants, notamment les variables relatives aux caractéristiques familiales comme les revenus des parents, leurs diplômes, la région de résidence, etc.

Régression linéaire. Dans un premier temps, nous ajoutons les caractéristiques socio-démographiques et les caractéristiques des territoires à la régression du rang des enfants sur le rang des parents (estimée sur la population des enfants âgés de 26 à 29 ans). La pente rang-rang diminue en incluant ces variables. Elle passe de 0,23 à 0,17 en intégrant le sexe, les revenus du capital, le type de ménage et la région. Elle atteint 0,13 en prenant en compte les caractéristiques individuelles issues du recensement de la population (tableau 9). Ainsi, les caractéristiques socio-démographiques font diminuer de près de la moitié la pente rang-rang, et donc l'importance des revenus des parents.

Cependant, les coefficients liés au rang, une variable quasi continue, sont difficilement comparables aux coefficients des autres caractéristiques qui sont discrètes avec seulement quelques modalités. Nous discrétisons donc la variable de revenu des parents dans le tableau E5. On observe d'abord que les coefficients des cinquièmes de revenu des parents sont divisés par 1,5 à 2 selon les modalités, après la prise en compte des autres variables (deuxième colonne), par rapport à une régression où les cinquièmes de revenus sont les seules variables explicatives (première colonne). Ensuite, les coefficients les plus importants sont obtenus pour les modalités des cinquièmes de revenus des parents (10,5 pour le plus haut cinquième, cf. tableau E5, annexe E).

Tableau 9 : Régression du rang des enfants sur différentes variables

	Baseline (1)	(1)+contrôles individuels (2)	(2) + contrôles régions (3)	(3) + contrôles var. RP (4)
C	39.00 *** (0.234)	45.087 *** (0.539)	45.669 *** (0.620)	44.423 *** (1.193)
β_{PR}	0.228 *** (0.004)	0.171 *** (0.005)	0.158 *** (0.005)	0.131 *** (0.007)
Contrôles var. individuelles		X	X	X
Contrôles régions d'appartenance			X	X
Contrôles diplôme et CS parents				X
Observations	58814	58736	40280	16681

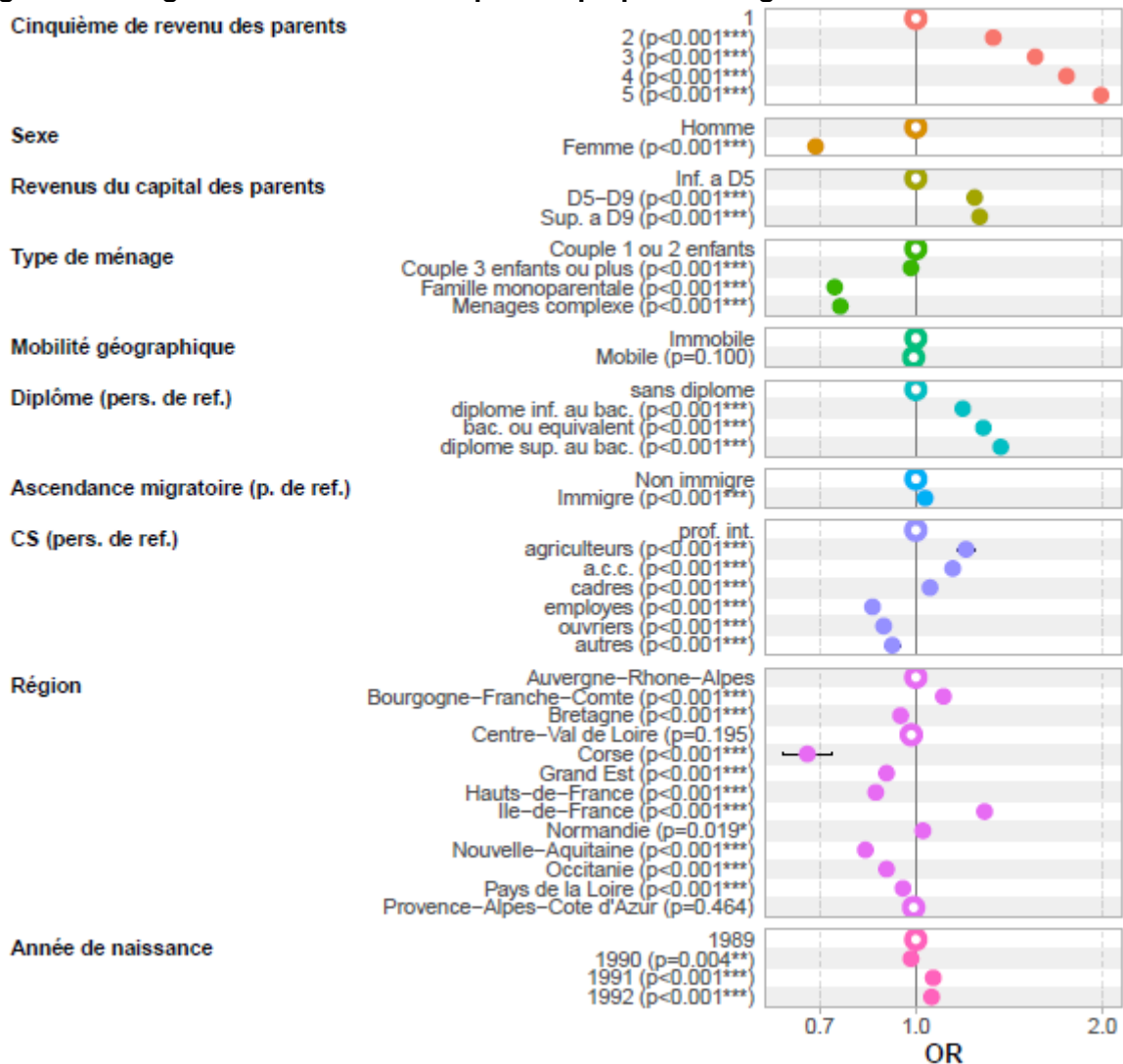
Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Régression logistique. Afin de mieux comparer les effets de chaque variable, nous mettons en œuvre une régression logistique polytomique à variable dépendante ordinale (logit cumulatif avec odds proportionnels). La variable à expliquer est le cinquième de revenus des enfants, en fonction du cinquième de revenus des parents et d'autres variables. La figure 10 présente les odds ratios issus de cette régression. Les odds ratios relatifs aux modalités des cinquièmes de revenus croissent avec les cinquièmes de revenus des parents et atteignent 2 pour le plus haut cinquième. C'est plus que les odds ratios des autres variables, ce qui

signifie que les revenus des parents ont plus d'effet que les autres variables dans la formation des revenus des enfants.

Figure 10 : Régression multinomiale pour expliquer le rang des enfants



Note : La figure présente les odds ratio associés à une régression multinomiale expliquant le rang des enfants en cinquième de revenu. Les ronds sans poids blanc à l'intérieur indiquent un OR significativement différent de 1. Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Effectuer une régression de Poisson ou une régression logistique sur la probabilité d'atteindre le plus haut cinquième des revenus conduit à des résultats proches (Figure E21 et E22) : les enfants dont les parents appartiennent au plus haut cinquième de revenus ont environ deux fois⁵⁰ plus de chances d'atteindre le dernier cinquième de revenus que les enfants des familles défavorisées, à autres caractéristiques identiques. L'effet est également plus élevé que celui des autres variables. Celui-ci est encore plus élevé si on considère la probabilité d'appartenir au plus haut cinquième plutôt qu'au plus bas cinquième et en calculant un odds ratio rapportant la probabilité de faire partie des ménages les plus aisés à

50 1,7 fois pour la régression de Poisson sur la population des individus recensés par le RP, 1,9 pour la régression de Poisson sur la population totale, 2,1 pour une régression logistique sur la population restreinte et 2,3 sur la population totale.

la probabilité de faire partie des ménages les plus modestes : les enfants des familles aisées ont 2,7 fois plus de chances que les enfants des familles modestes d'atteindre le plus haut cinquième de la distribution que d'intégrer le plus bas cinquième.

5.5.2 Méthode par utilisation des fratries

La méthode précédente permet de comparer l'effet de chaque variable sur les revenus des enfants, mais pas de quantifier l'importance du milieu familial dans la formation des revenus. L'effet de la famille sur les revenus peut être mesuré en comparant les rangs ou les revenus des frères et sœurs à l'intérieur des fratries. En effet, étudier la mobilité intergénérationnelle comme nous l'avons fait jusqu'ici ne permet de saisir qu'une partie de l'importance de l'environnement familial dans la distribution des revenus. Il est possible de compléter l'analyse de l'importance du milieu familial en comparant les revenus ou les rangs des frères et sœurs. En effet, les fratries ont en commun les revenus de leur parent, mais pas seulement : elles partagent aussi un capital culturel, un patrimoine génétique, et l'influence du milieu social et du voisinage (école, amis, religion, etc.).

Dans un premier temps, nous étudions les revenus des jumeaux nés en 1990 à leurs 28 ans, comme dans l'analyse principale (ce qui permet de minimiser le biais de cycle de vie et l'effet de sélection). A 28 ans, les frères et sœurs ont une différence de rang moyenne de 26 centiles. Cependant, une partie de cette différence s'explique par le fait que les femmes ont en moyenne un revenu plus faible et il peut donc être plus pertinent de comparer les frères entre eux et les sœurs entre elles. La différence est de 23 centiles entre frères jumeaux et de 24 centiles entre sœurs jumelles (tableau 10). Étant donné que la différence théorique de rang entre deux individus est de 33,3 centiles (voir Annexe F pour le calcul théorique), on peut estimer qu'environ 30 % (31 % pour les frères et 29 % pour les sœurs) de la différence de rang serait liée au milieu familial au sens large, le reste étant lié à d'autres facteurs.

L'échantillon des jumeaux ayant 28 ans en 2018 est cependant de petite taille (307 individus) et ces estimations doivent être prises avec précaution. En élargissant l'échantillon et en considérant les fratries nées entre 1989 et 1992 (c'est-à-dire qui ont entre 26 et 29 ans en 2018), et qui ont moins de 3 ans de différence (afin que l'effet de génération soit peu marqué), on obtient un échantillon d'environ 1640 frères et sœurs. Les différences de rang entre frères, d'une part, et entre sœurs, d'autre part, sont respectivement de 23 et 24 centiles. Ainsi, sur cette population, la part de la variation des revenus liée à la famille serait de 30 % pour les frères et de 27 % pour les sœurs. Elle serait de 24 % en prenant les frères et les sœurs ensemble. Ce résultat est proche, mais un peu plus faible, que celui obtenu par Boutchenik et al. (2015) qui montrent qu'un peu plus d'un tiers de la variabilité des niveaux de profession est lié à l'influence familiale en France. Cette proportion estimée à partir de nos données serait légèrement plus élevée que celle obtenue en Italie par Acciari et al. (2020), même s'il est difficile de comparer directement ces estimations du fait que, dans cette dernière étude, les « enfants » sont plus âgés que dans nos données (ce qui, a priori, a pour effet de diminuer la part de la variation de rang liée à la famille).

Une autre façon (et plus courante dans la littérature) de juger de l'importance relative du facteur familial consiste à mesurer la corrélation entre les revenus des frères et sœurs. Le coefficient de corrélation (de Pearson) des revenus est alors compris entre 30 % et 40 % selon les années et le sexe (il est plus élevé pour les frères et les enfants âgés)⁵¹. Cependant, en général, c'est le logarithme des revenus qui est considéré dans la littérature. Utiliser le logarithme des revenus pose le problème des revenus négatifs et nuls, et la manière de traiter ces revenus a une forte influence sur les résultats obtenus, comme cela a été montré en ce qui concerne l'indicateur IGE. En supprimant les revenus négatifs ou nuls, la corrélation du logarithme des revenus est plus faible (de 0,16 pour les frères). Ce résultat

51 À noter également que la corrélation au niveau des rangs de revenus donne des résultats un peu plus élevés.

est relativement proche des chiffres obtenus dans les pays scandinaves pour les frères (Norvège, Finlande, Suède), et beaucoup plus faible que ceux obtenus aux États-Unis, en Allemagne et en Chine, selon la compilation des estimations de la littérature réalisée par Jäntti et Jenkins (2015). Ces estimations sont effectuées sur des échantillons d'enfants plus âgés que ceux figurant dans nos données⁵², et sont obtenues avec des techniques statistiques différentes : les estimations ne sont pas directement comparables et doivent être prises avec précaution pour la France. Si nous ne pouvons pas comparer, à âge donné, nos résultats à ceux obtenus en Allemagne et dans les pays scandinaves, la corrélation des revenus entre fratries serait a priori plus faible en France qu'aux États-Unis et en Chine, et la reproduction des inégalités de revenus liée à la famille y serait donc plus faible.

Tableau 10 : Différence de rang et corrélation des revenus au sein des fratries, selon la génération

		1989 (29 ans)	1990 (28 ans)	1991 (27 ans)	1992 (26 ans)	89-90-91-92
Nombres de frères et sœurs		174	307	411	379	1638
Différence de rang	fratrie	26,8	25,6	24,7	24,7	25,4
	frères	24,2	23,0	24,8	21,1	23,3
	sœurs	25,6	23,7	23,3	24,0	24,3
Part de la variation de rang liée à la famille	fratrie	20%	23%	26%	26%	24%
	frères	27%	31%	26%	37%	30%
	sœurs	23%	29%	30%	28%	27%
Corrélation entre les revenus	fratrie	7%	39%	29%	22%	25%
	frères	43%	39%	39%	30%	36%
	sœurs	13%	34%	42%	24%	31%

Champ : France métropolitaine. Individus rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

52 Nos estimations doivent être considérées comme des bornes hautes par rapport aux autres études.

Bibliographie

- Alesina, A., Stantcheva, S. et Teso, E. (2018). Intergenerational Mobility and Preferences for Redistribution, *American Economic Review*, vol. 108(2), p. 521-554.
- Acciari, P., Polo, A. et Violante, G.L. (2020). 'And Yet It Moves': Intergenerational Mobility in Italy. NBER Working Paper 25732. <https://doi.org/10.3386/w25732>
- Acciari, P., Polo, A. et Violante, G.L. (2021). 'And Yet It Moves': Intergenerational Mobility in Italy. *American Economic Journal: Applied Economics* (forthcoming).
- Algava E., Bloch, K., Vallès V. (2020). En 2018, 4 millions d'enfants mineurs vivent avec un seul de leurs parents au domicile, *Insee Première*, n°1788. <https://www.insee.fr/fr/statistiques/4285341>
- Abramitzky, R., Boustan, L., Jacome, L., et S. Perez. 2021. Intergenerational Mobility of Immigrants in the United States over Two Centuries. *American Economic Review*, 111 (2): 580-608. <https://www.aeaweb.org/articles?id=10.1257/aer.20191586>
- Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., & D. Hemous. (2019). Innovation and Top Income Inequality, *The Review of Economic Studies*, Volume 86, Issue 1, Pages 1–45, <https://doi.org/10.1093/restud/rdy027>
- Abbas, H., Garbinti, B. (2019). De la rupture conjugale à une éventuelle remise en couple : l'évolution des niveaux de vie des familles monoparentales entre 2010 et 2015, in *France Portrait Social*, édition 2019.
- Algava E., Penant S. et L. Yankan (2019). En 2016, 400 000 enfants alternent entre les deux domiciles de leurs parents séparés, *Insee Première* 1728.
- Asfa C (2016). Le modèle Logit : Théorie et application. *Document de travail* n°2016-01 de l'Insee.
- Bell, B., Chetty, R., Jaravel, X., Petkova, N., Van Reenen, J. (2019). Who Becomes an Inventor in America? The Importance of Exposure to Innovation, *The Quarterly Journal of Economics*, Vol. 134, Issue 2, p. 647–713
- Berman, Y. (2021). The Long Run Evolution of Absolute Intergenerational Mobility. *American Economic Journal: Applied Economics* (forthcoming).
- Bernard, J. (2021). Les jeunes ni en emploi, ni en études, ni en formation : jusqu'à 21 ans, moins nombreux parmi les femmes que parmi les hommes. *Insee Focus*, no 229 <https://www.insee.fr/fr/statistiques/5346969>
- Bjorklund, A. & Jantti, M., (1997). [Intergenerational Income Mobility in Sweden Compared to the United States](https://www.jstor.org/stable/2951338), *American Economic Review*, 87, issue 5, p. 1009-18. <https://www.jstor.org/stable/2951338>
- Bonneau, C., Grobon, S. (2022). Unequal access to higher education based on parental income: evidence from France. WID.world Working Paper, N° 2022/01. <https://wid.world/news-article/unequal-access-to-higher-education/>
- Boserup, S., Kopczuk, W. et Kreiner, C. (2014). Stability and persistence of intergenerational wealth formation: Evidence from Danish wealth records of three generations. University of Copenhagen.
- Bratberg, E., Davis, J., Mazumder, B., Nybom, M., Schnitzlein, D., and K. Vaage (2017). A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US. *The Scandinavian Journal of Economics*, 119 (1), 72–101.
- Boutchenik, B, Coron, C., Grobon, S. Goffette, C. et Vallet L-A (2015). Quantifier l'influence totale de la famille d'origine sur le devenir scolaire et professionnel des individus. *Economie et statistique* N° 477.

- Chetty, R., Friedman, J., Hilger, N., Saez, E., Schanzenbach, D. & Yagan D. (2011). [How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR](http://www.rajchetty.com/chettyfiles/STAR.pdf). *Quarterly Journal of Economics* 126(4): 1593-1660. <http://www.rajchetty.com/chettyfiles/STAR.pdf>
- Chetty, R., Hendren, N., Kline, P. and Saez, E. (2014a). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, 129(4), pp.1553-1623. <https://academic.oup.com/qje/article/129/4/1553/1853754>
- Chetty, R., Hendren, N., Kline, P., Saez, E. & Turner, N. (2014b). Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. *American Economic Review*, 104(5), pp.141-147. http://www.equality-of-opportunity.org/files/mobility_trends.pdf
- Chetty, R., Hendren, N. & L. Katz (2016). [The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment](https://opportunityinsights.org/wp-content/uploads/2018/03/mto_paper.pdf), *American Economic Review* 106(4): 855-902, 2016. https://opportunityinsights.org/wp-content/uploads/2018/03/mto_paper.pdf
- Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020). Income segregation and intergenerational mobility across colleges in the united states. *The Quarterly Journal of Economics*, 135 (3), 1567–1633
- Clanché F. (2014). Le nouveau recensement en continu : une méthode originale, des résultats nombreux et faciles d'accès", *Informations sociales* 2014/3 (n°183), pages 22 à 25
- Collet, M., Pénicaud, E. (2019). La mobilité sociale des femmes et des hommes : évolutions de 1977 à 2015, in *France, Portrait Social*, Insee. <https://www.insee.fr/fr/statistiques/4238441?sommaire=4238781>
- Corak, M. (2013). Income Inequality, Equality of Opportunity, and Intergenerational Mobility. *Journal of Economic Perspectives*, 27(3), pp.79-102.
- Corak, M, et A. Heisz (1999). The Intergenerational Earnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data, *Journal of Human Resources*, 34, no. 3, 504–533.
- Costemalle, V. (2017). Les données fiscales de l'EDP : une nouvelle source d'informations sur les couples et les familles ? *Document de travail*, Insee.
- Davezies L. et X. D'Haultfoeuille (2009). Faut-il pondérer ?... Ou l'éternelle question de l'économètre confronté à des données d'enquête. *Insee working paper* n°2009-06.
- Dherbécourt C. (2015). La géographie de l'ascension sociale, *La Note d'analyse*, n° 36, France Stratégie.
- Dherbécourt, C. (2018), Nés sous a même étoile ? Origine sociale et niveau de vie. *La Note d'analyse*, n°68, *France Stratégie*. <https://www.strategie.gouv.fr/publications/nes-meme-etoile-origine-sociale-niveau-de-vie>
- Dherbécourt C. et Kenedi G. (2020), Quelle influence du lieu d'origine sur le niveau de vie ?, *France Stratégie*, *La Note d'analyse*, n° 91, juin. <https://www.strategie.gouv.fr/publications/influence-lieu-dorigine-niveau-de-vie>
- Dherbécourt (2020). La mobilité sociale en France : que sait-on vraiment ?, *France Stratégie*. <https://www.strategie.gouv.fr/point-de-vue/mobilite-sociale-france-sait-vraiment>
- Heidrich, S. (2017). Intergenerational Mobility in Sweden: A Regional Perspective. *Journal of Population Economics*, 30(4), 1241–1280. <https://link.springer.com/article/10.1007/s00148-017-0648-x>
- Helsø, A.-L. (2021). Intergenerational Income Mobility in Denmark and the United States. *Scand. J. of Economics*, 123: 508-531. <https://doi.org/10.1111/sjoe.12420>
- Jaravel X. (2021). Les inégalités constituent un puissant frein à l'innovation. <https://www.lesechos.fr/economie-france/social/les-inegalites-constituent-un-puissant-frein-a-linnovation-1331142>

- Jäntti M., Jenkins S.P. (2015). Income mobility. in *Handbook of income distribution*, vol 2A, eds A.B. Atkinson, F. Bourguignon
- Kenedi G. et Sirugue L. (2021). The Anatomy of Intergenerational Income Mobility in France and its Spatial Variations. *PSE working paper*. <https://halshs.archives-ouvertes.fr/halshs-03455282/>
- Landersø, R. and J. J. Heckman, (2017). The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US. *The Scandinavian Journal of Economics*, 119 (1), 178–230. <https://onlinelibrary.wiley.com/doi/full/10.1111/sjoe.12219>
- Lapinte, A. (2013). Un enfant sur dix vit dans une famille recomposée. *Insee Première* n°1470. <https://www.insee.fr/fr/statistiques/1281054>
- Lefranc A. (2018). Intergenerational earnings persistence and economic inequality in the long-run : Evidence from French cohorts, 1931-1976. *Economica*, vol. 85(340), pages 808-845 <https://doi.org/10.1111/ecca.12269>
- Lefranc, A., Pistolesi, N., Trannoy, A. (2004). Le revenu selon l'origine sociale. *Economie et statistique*, 371, 49-82.
- Lefranc A. et A. Trannoy (2005). Intergenerational Earnings Mobility in France: Is France More Mobile than the US? *Annales d'Économie et de Statistique*, (78), pp.57-77.
- Lièvre, A. (2018). Les ressources des étudiants selon la formation suivie. *Note d'information du SIES* n°18-05.
- Mazumder, B. (2005). Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data, *Review of Economics and Statistics*, 87 (2), 235–255
- Murray, C., Clark, R.G., Mendolia, S. and Siminski, P. (2018). Direct Measures of Intergenerational Income Mobility for Australia. *Econ Rec*, 94: 445-468. <https://doi.org/10.1111/1475-4932.12445>
- Nybom, M., et J. Stuhler (2017): Biases in Standard Measures of Intergenerational Income Dependence, *Journal of Human Resources*, 52(3), 800–825. doi:10.3368/jhr.52.3.0715-7290R
- OCDE, 2018, *A Broken Social Elevator? How to Promote Social Mobility*, OCDE, Paris <https://www.oecd.org/social/broken-elevator-how-to-promote-social-mobility-9789264301085-en.htm>
- Poncelet T., Trigano, L., et M. Sagot (2015). Gravir l'échelle sociale est plus aisé en Île-de-France qu'en province, *Insee analyses Ile-de-France* n°50. <https://www.insee.fr/fr/statistiques/2529731>
- Pouliquen, E. (2018). Depuis 2000, la part des 18-29 ans habitant chez leurs parents augmente à nouveau. *Insee Première* N° 1686. <https://www.insee.fr/fr/statistiques/3315412>
- Robert-Bobée I., Gualbert, N. (2021). L'échantillon démographique permanent : en 50 ans, l'EDP a bien grandi !. *Courrier des Statistiques* n°6, 2021. <https://www.insee.fr/fr/information/5398685?sommaire=5398695>
- Roemer, J. E., et A. Trannoy. (2016). Equality of Opportunity: Theory and Measurement. *Journal of Economic Literature*, 54 (4): 1288-1332. <https://www.aeaweb.org/articles?id=10.1257/jel.20151206>
- Solon, G. (1999). Intergenerational Mobility in the Labor Market. *Handbook of Labor Economics* Vol. 3, pp.1761-1800.
- Vallet, L. (1999). Quarante années de mobilité sociale en France. *Revue française de sociologie*, 40(1), pp.5-64.
- Zou G. (2004). A modified poisson regression approach to prospective studies with binary data. *American Journal of Epidemiology*. 1;159(7):702-6.

Annexes

Annexe A. Traitement des individus multi-identifiés et des enfants en garde alternée.

Les individus EDP, comme tous les Français, peuvent être rattachés à plusieurs déclarations fiscales selon la configuration familiale ou leur âge. Par exemple, les enfants en garde alternée sont déclarés à la fois par leur père et par leur mère et sont donc rattachés à deux logements. Ils sont donc représentés par deux observations dans les données fiscales ; il en est de même pour les étudiants qui sont souvent rattachés à la déclaration fiscale de leurs parents mais ont leur propre déclaration pour la taxe d'habitation dans le cas où ils vivent dans un logement distinct de celui de leurs parents (dans ce cas aussi, ces individus sont censés apparaître deux fois dans l'EDP). Il est donc nécessaire de traiter de manière spécifique ces individus. On remarquera que les autres configurations familiales ne donnent lieu qu'à l'établissement d'une seule déclaration fiscale pour un individu donné. Au total, on considère dans cette étude quatre configurations familiales :

a) un enfant est rattaché à des parents mariés ou pacsés. Dans ce cas, il n'y a qu'un seul foyer fiscal et qu'une seule déclaration (et un seul logement fiscal). On retrouve par conséquent directement les revenus détaillés de chacun des deux parents dans les données fiscales de l'EDP.

b) un enfant vit dans une famille dont les parents vivent en union libre. Dans ce cas, l'enfant n'est rattaché qu'à la déclaration fiscale d'un seul de ses parents (théoriquement) et il n'est pas possible de récupérer dans l'EDP⁵³ les revenus détaillés du second parent car la table des revenus détaillés ne concerne que les revenus de l'ensemble des individus qui appartiennent au même foyer fiscal que celui de l'individu EDP. En revanche, il est possible d'identifier ces ménages en utilisant la variable TYPMEN9 (ces enfants vivent dans des logements dans lesquels leurs parents vivent en couple, modalités 30, 41, 42 et 43). En conséquence, considérer l'ensemble des enfants qui ne sont rattachés qu'à un seul de leurs parents revient à surestimer le nombre d'enfants en familles monoparentales car, parmi ces enfants, certains vivent aussi avec leurs deux parents quand ceux-ci sont en union libre. Pour distinguer ces deux cas, nous utiliserons donc la variable TYPMEN9 (dont deux modalités 21 et 22 permettent de distinguer les familles monoparentales). Nous ne faisons pas de traitement particulier pour ces enfants.

c) si les parents d'un enfant sont séparés, l'enfant n'est rattaché qu'à la déclaration fiscale du parent chez lequel il réside (cas des familles monoparentales). Il n'est possible d'observer que le revenu détaillé du parent chez lequel il réside, pas celui de l'autre parent chez lequel il ne réside pas ni n'est déclaré.

d) si un enfant vit en résidence alternée (i.e., chez ses deux parents qui ont chacun leur propre logement), il fait l'objet de deux déclarations fiscales et est censé pouvoir être repéré par la case T_charge=H. D'après les données existantes sur la garde alternée (cf Algava, Penant et Yankan, 2019), les enfants en garde alternée et âgés de plus de 14 ans

53 Le traitement est possible directement dans les bases « brutes » de Fidéli (grâce à un algorithme permettant d'imputer l'information) mais cette information n'est pas conservée dans l'EDP.

représentent moins de 2 % des enfants de moins de 18 ans en 2010 (d'après les données Fidéli, i.e. les données utilisées dans la présente étude). Dans la pratique, tous les individus déclarés en garde alternée par leurs parents (environ 1,5 % des enfants de 14 à 17 ans dans notre échantillon) ne font pas l'objet de deux déclarations fiscales. On ne retrouve deux observations que pour 0,9 % des enfants de 14 à 17 ans de notre échantillon.

Nous traitons les cas a), b) et c) de la même manière : si l'individu EDP est rattaché à plusieurs déclarations fiscales, on ne garde que les informations issues de la déclaration d'impôts sur le revenu et la taxe d'habitation (TYPE_DECL=1), de préférence à une déclaration issue d'une seule déclaration de revenus (TYPE_DECL=2). On élimine les observations qui découlent de la seule taxe d'habitation (TYPE_DECL=3) car elles ne permettent pas de récupérer les revenus des parents.

Pour le cas d), nous gardons les observations des deux parents que nous fusionnons pour ne garder qu'une seule observation permettant de disposer sur une même ligne des revenus détaillés du père et de ceux de la mère. Ces observations sont repérables grâce à la variable T_CHARGE (modalité égale à H). Nous ne gardons ensuite que les informations relatives à celui des deux logements dans lequel le niveau de vie est le plus élevé. Attention, le traitement de ces observations ne concerne que les enfants de moins de 18 ans (i.e., qui ont au plus 17 ans) : il concerne une très petite proportion des individus de notre échantillon, moins de 1 %.

Nous détaillons à présent le cas des enfants alternants (les enfants de moins de 18 ans, en 2010, qui résident chez chacun de leurs deux parents séparés) et présentons quelques statistiques permettant de mieux comprendre leur identification dans l'EDP.

Il existe deux manières de repérer les enfants alternants : soit ils apparaissent à la fois dans deux logements et dans deux foyers fiscaux différents (et ils font donc l'objet de deux observations), soit leurs parents déclarent l'alternance (case T_CHARGE== « H »). En principe, les deux informations devraient coïncider, ce qui n'est pas le cas, comme l'illustre l'exploitation de quelques variables permettant de décrire leur situation (nombre de logements auxquels les enfants alternants sont rattachés, nombre de foyers fiscaux auxquels ils sont rattachés, et la variable de déclaration de la résidence alternée T_CHARGE). Nous utilisons dans les statistiques descriptives suivantes le champ des enfants nés entre 1993 et 1996 qui apparaissent dès 2010 dans les données fiscales (70 698 individus) et non le champ complet des enfants de notre base qui est constitué non seulement des enfants nés entre 1993 et 1997 identifiés dès 2010, mais également des enfants nés au cours de ces mêmes années mais apparaissant dans les données fiscales à partir de 2011 et 2012 (79 089 individus) car nous ne disposons pas d'une information suffisante pour ces enfants (nous ne connaissons, pour ces enfants, que la nature de leur statut fiscal à un âge au moins égal à 18 ans ; or, le statut d'alternant ne s'applique pas aux enfants adultes) :

- parmi les 70 698 enfants âgés de 14 à 17 ans figurant dans notre base en 2010 (année fiscale 2011), 1,5 % ne résident pas dans un logement identifié et 0,9 % sont rattachés à deux logements (et ont donc deux observations dans la base),

- parmi ces 70 698 enfants, 1 % sont rattachés à deux foyers fiscaux. Les enfants qui sont à la fois rattachés à deux logements et deux foyers fiscaux représentent 0,9 % de tous les enfants ;

- parmi les 70 698 enfants, 1,51 % sont déclarés comme alternants, soit plus que l'ensemble des enfants qui sont rattachés à deux logements et deux foyers fiscaux. Ainsi, seuls 23,67 % des enfants déclarés comme alternants sont rattachés à deux logements, les enfants restants étant rattachés soit à un logement (74,7%) soit à aucun logement (1,6%)

Nous en déduisons qu'une bonne partie des enfants déclarés comme alternants ne peuvent être rattachés à deux déclarations fiscales différentes et qu'il n'est pas possible de récupérer à la fois les revenus du père, et ceux de la mère (cela n'est possible que pour près de 24 % d'entre eux environ) Nous en déduisons également que bon nombre d'enfants qui sont déclarés dans deux logements (0,89 % au total) ne sont pas déclarés comme alternants dans la déclaration fiscale de leurs parents (40,29 % seulement, 41,8 % étant déclarés comme enfants à charge exclusive).

Tableau A1. Croisement du nombre de logements auxquels les enfants de moins de 18 ans sont rattachés et du statut des enfants dans la déclaration fiscale (T_CHARGE)

Nombre de logements	Catégorie d'enfant à charge					
	Inconnu	Mineur à temps plein	Mineur alternant	Majeur célibataire	Majeur marié	Personne invalide
0	73	968	17	1	0	0
1	343	67 859	799	8	1	1
2	116	258	253	1	0	0

Lecture: 799 enfants nés entre 1993 et 1997 sont rattachés à un seul logement dans les données fiscales et sont en même temps déclarés comme alternants par leurs parents.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Nous décidons de traiter les enfants de moins de 18 ans de la manière suivante : pour ceux qui apparaissent dans deux logements différents (0,9%, i.e. qui ont deux observations), nous cherchons à récupérer les informations de l'autre conjoint (y compris les revenus détaillés) et nous ne gardons les informations relatives qu'au logement dont le niveau de vie est le plus élevé des deux. Pour les autres, nous ne procédons à aucun traitement particulier, nous gardons l'observation qui les concerne (il n'en existe qu'une, car il n'y a qu'un seul logement).

Annexe B. Traitement des effectifs changeants selon l'année fiscale

Dans les données fiscales de l'EDP, qui servent à repérer les enfants suivis entre 2010 et 2018 dans notre étude, les effectifs d'individus nés une année donnée fluctuent selon l'année considérée, surtout pour les cohortes d'enfants ayant 18 ans ou moins en 2010. Sans pondération, le nombre d'individus âgés de 18 ans issus des données fiscales dans l'EDP est ainsi toujours inférieur au nombre attendu d'après les naissances qui ont eu lieu 18 ans plus tôt. Par exemple, pour l'année fiscale 2011 (revenus 2010), on trouve 17 479 (chiffre non pondéré) enfants EDP âgés de 18 ans, alors qu'on s'attend à en trouver de l'ordre de 29 720 (4 % des 743 000 naissances en France en 1992). Plusieurs explications peuvent être apportées pour expliquer cette différence. Premièrement, l'année fiscale 2011 de Fideli n'est, d'après le producteur de la source, pas d'une qualité parfaite : il s'agit du premier millésime, il est exploratoire. D'autre part, dans les bases fiscales, la date de naissance au

jour près n'est connue que pour les enfants majeurs, ce qui peut générer des problèmes d'identification pour les personnes âgées de 18 ans une année donnée. Sélectionner des individus EDP à partir des bases fiscales peut donc amener à sous-estimer le nombre d'enfants mineurs. Pour chaque année fiscale, nous observons ainsi davantage d'individus au moment du passage à la majorité (nés en 1991 vs 1992 en 2011, nés en 1992 vs 1993 en 2012, etc.). Ainsi, nous retrouvons dans les données fiscales de l'EDP en 2010 17 479 individus nés en 1992, mais 23 496 pour l'année 2011 et environ 30 000 à partir de 2016 : le nombre d'enfants nés en 1992 et présents dans les données fiscales ne cesse d'augmenter au cours des années suivantes, à mesure que les enfants deviennent adultes et que les problèmes d'identification s'estompent, et atteint le nombre attendu en 2019.

De manière plus générale, la difficulté principale semble être de suivre les enfants au moment de la décohabitation, qui doit souvent survenir autour des âges qui nous intéressent dans le cadre de cette étude. Pour atténuer ce problème, nous décidons d'intégrer aux cohortes d'individus nés entre 1988 et 1996 et observés en 2010 les individus nés la même année mais qui apparaissent seulement en 2011 ou en 2012 pourvu que ces individus soient rattachés en 2011 ou en 2012 au foyer fiscal de leurs parents (ce qui devient moins probable à mesure que l'on s'éloigne de l'âge de 18 ans). Finalement, cette méthode nous permet de travailler sur des cohortes constituées d'individus plus nombreux (ligne 2013 du tableau ci-dessous et cf. annexe D).

Les données sont ensuite pondérées en utilisant les poids introduits dans l'EDP (calage sur les données fiscales exhaustives), recalés sur la structure par sexe, âge et état matrimonial de la population. Les pondérations viennent ainsi corriger les défauts de représentativité de notre échantillon, en donnant aux enfants identifiés un poids pour qu'ils représentent, une fois les statistiques pondérées, l'ensemble des enfants français.

Tableau B1. Tableau croisé des années fiscales et des effectifs « d'enfants » nés une année donnée entre 1988 et 1996 (effectifs bruts)

année fiscale	Année de naissance								
	1988	1989	1990	1991	1992	1993	1994	1995	1996
2011	27875	27520	26896	24844	17479	15805	15900	17893	19687
2012	29053	28957	28826	27643	23496	17217	15589	17570	19374
2013	30109	30057	30234	29379	25750	22276	17170	17540	19180
2014	30473	30670	30983	30378	27132	24641	21854	19389	19349
2015	31395	31850	32418	32106	28954	26735	24812	24592	21601
2016	31808	31962	32794	32675	29595	27578	26069	27654	28270
2017	32138	32520	32795	32995	30016	28121	26636	28414	30538
2018	32317	32794	33172	32773	30068	28254	26733	28440	30483
2019	32335	32921	33297	32955	29878	28245	26795	28427	30576

Champ : individus EDP dont le logement est identifié

Lecture : en 2010, 17 479 individus apparaissent dans les données fiscales de l'EDP

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Annexe C. Traitement des données du recensement de la population de 1999 et des enquêtes annuelles de recensement

a) Méthode et traitements

Dans cette étude, nous exploitons les données du recensement de la population de 1999. Lors de ce recensement, les enfants de notre population d'intérêt (l'ensemble des enfants nés entre 1988 et 1996 qui figurent dans l'EDP) sont âgés de 3 à 11 ans, des âges auxquels les données permettent de décrire leur milieu d'origine, en particulier les caractéristiques socio-démographiques de leurs parents.

Cependant, si le recensement de la population de 1999 est exhaustif, il ne permet pas de récupérer les variables caractérisant le milieu d'origine de tous les enfants de la base d'étude en raison du taux d'échantillonnage du recensement dans l'EDP. En effet, jusque 2007, le recensement de la population comme, ensuite, les enquêtes annuelles de recensement (EAR) jusqu'à cette date, ne concernent, dans l'EDP, que les individus nés l'un des quatre jours d'octobre (le taux d'échantillonnage de l'EDP est alors de 1 % de la population française). Ce n'est que lors de l'enquête annuelle de recensement de 2008 que la source couvre les « seize jours EDP ». Par conséquent, nous utilisons le recensement de la population de 1999 pour récupérer les caractéristiques sociales de tous les enfants de notre base nés au mois d'octobre de l'une de nos années d'intérêt (1988 à 1996, ce qui représente 25 % des individus ainsi identifiés) puis, pour compléter cette information, nous appariions notre base avec les enquêtes annuelles de recensement de 2008 à 2016 pour récupérer les variables socio-démographiques d'un maximum d'individus nés en dehors d'un mois d'octobre.

Cette approche nous conduit à inclure dans notre base des informations sur les caractéristiques du milieu d'origine des enfants à un âge plus avancé que pour ceux nés un mois d'octobre. Afin de ne pas fausser l'analyse, nous décidons d'inclure, pour chaque individu de la base dont nous retrouvons au moins une observation dans une EAR entre 2008 et 2016 (les individus nés en 1996 ont au plus 20 ans au moment de l'enquête annuelle de recensement de 2016):

- l'observation la plus ancienne si un individu apparaît dans plusieurs enquêtes annuelles de recensement. Ainsi, on essaie de caractériser le milieu social d'origine au plus tôt dans la vie des enfants de la base ;
- nous décidons de n'inclure dans la base finale que les observations des EAR telles que l'âge de l'enfant dans une EAR donnée est inférieur ou égal à 20 ans. Nous nous assurons ainsi que les informations sociales récupérées correspondent à une situation dans laquelle l'enfant observé vit bien chez au moins l'un de ses parents.

Dans la pratique, cette approche conduit à récupérer les caractéristiques sociales de 72,51 % de la population d'enfants. Ce pourcentage est légèrement inférieur au taux d'appariement théorique : si 25 % de la population d'intérêt est présente dans le recensement de 1999 et que le taux de sondage des enquêtes annuelles de recensement est environ de 14 %, il doit être possible de récupérer les informations socio-démographiques d'environ 77,5 % d'individus au bout d'un cycle de 5 ans d'enquêtes annuelles de recensement ($0,25 + 5 * 0,14 * 0,75 = 0,775$) Ce taux est inférieur au taux théorique car il peut arriver que les mêmes individus soient interrogés plus d'une fois au

cours d'un cycle de 5 ans dans les communes de plus de 10 000 habitants (Clanché, 2014). Le tableau C1 suivant indique le pourcentage d'individus dont on retrouve au moins une observation entre 1999 et 2016 : le taux d'individus dont on extrait l'information dans au moins une EAR baisse au cours du temps du fait de la méthode retenue (nous apparions les individus avec l'observation la plus ancienne), et est très faible après 2012. Au total, il n'a pas été possible de retrouver une observation dans le recensement ou dans une enquête annuelle de recensement pour 27,5 % des enfants de la base.

Tableau C1. Taux d'appariement des enfants avec les caractéristiques sociales de leurs parents issues du recensement de la population en 1999 ou de l'une des enquêtes annuelles de recensement entre 2008 et 2016.

Année	Effectifs	Pourcentages bruts	Pourcentages sur les observations présentes	Pourcentages cumulés
1999	43564	26,97	37,19	37,19
2008	16121	9,98	13,76	50,96
2009	16587	10,27	14,16	65,12
2010	14785	9,15	12,62	77,74
2011	12605	7,8	10,76	88,5
2012	10051	6,22	8,58	97,08
2013	2259	1,4	1,93	99,01
2014	1116	0,69	0,95	99,96
2015	31	0,02	0,03	99,99
2016	10	0,01	0,01	100
Observations non appariées	44417	27,49		

Lecture : 9.98% des enfants figurant dans les données ont été appariés avec les données de l'enquête annuelle de recensement de 2008.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Si l'utilisation du seul recensement de la population de 1999 conduit à une pondération identique des différentes années de naissance au sein de l'échantillon, recourir aux EAR à partir de 2008 pour compléter l'information socio-démographique du RP 1999 engendre une déformation du poids des différentes années de naissance au sein des individus dont nous retrouvons une observation dans au moins l'une des deux sources. En effet, la colonne 2 du tableau C2 suivant montre que ce pourcentage des différentes années de naissance est différent de la part de chaque génération au sens de la table d'étude finale (dernière colonne du tableau). Intuitivement, on comprend qu'il n'est possible d'utiliser que l'EAR de 2008 pour la génération d'enfants nés en 1988 (année au cours de laquelle ces « enfants » ont au plus 20 ans). Inversement, les enfants de la génération 1996 peuvent être appariés avec des observations appartenant aux EAR de 2008 à 2016 (année au cours de laquelle ces « enfants » ont au plus 20 ans). La probabilité de retrouver une observation idoine dans l'une des EAR entre 2008 et 2016 est donc plus forte pour les enfants nés au cours des dernières années d'observation que pour ceux qui sont nés au cours d'une année proche de 1988.

Pour pallier ce problème, on repondère les données de manière à retomber sur les proportions de la dernière colonne du Tableau C2 suivant.

Tableau C2. Effectifs d'enfants retrouvés dans le recensement de la population de 1999 ou dans au moins l'une des enquêtes annuelles de recensement entre 2008 et 2016 selon l'année de naissance.

Année	Effectifs	Pourcentages	Pourcentages cumulés	Part de chaque génération au sein de la table d'étude
1988	2494	2,13	2,13	5
1989	4653	3,97	6,1	7,4
1990	8214	7,01	13,11	10,5
1991	12847	10,97	24,08	13,6
1992	16811	14,35	38,44	14,5
1993	17650	15,07	53,5	13,8
1994	16261	13,88	67,39	10,7
1995	18210	15,55	82,93	11,7
1996	19989	17,07	100	12,8

Lecture: 16 811 enfants nés en 1992 ont pu être retrouvés dans le recensement de la population de 1999 ou dans au moins une enquête annuelle de recensement.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

b) Les variables du recensement de la population de 1999 et des enquêtes annuelles de recensement utilisées dans l'étude.

Le recensement de la population de 1999 et les enquêtes annuelles de recensement (principalement entre 2008 et 2012) permettent de retrouver les caractéristiques du milieu social des « enfants » identifiés dans les données fiscales. Elles permettent en particulier de connaître les diplômes respectifs des deux parents, ainsi que leurs catégories sociales. Dans les analyses, nous prenons la variable de la personne de référence, définie comme le parent ayant le plus haut revenu.

Les diplômes des parents

La nomenclature harmonisée des diplômes depuis 1968 regroupe neuf modalités : aucun diplôme, certificat d'études primaires, brevet des collèges, certificat d'études professionnelles, brevet d'études professionnelles, baccalauréat général, baccalauréat technologique ou professionnel, diplôme universitaire de premier cycle (y compris BTS et DUT), diplôme universitaire de second cycle ou de troisième cycle (y compris les grandes écoles).

Depuis l'EAR de 2004, le BEP et le CAP font l'objet de modalités ad hoc. Depuis 2015, une nouvelle modalité est apparue pour distinguer les titulaires d'un doctorat (hors santé).

Dans cette étude, en raison du nombre relativement limité d'observations, l'ensemble de ces modalités ont été regroupées en quatre groupes : les individus qui n'ont aucun diplôme et/ou dont la scolarité s'arrête au niveau du collège (a), les titulaires du brevet des collèges, d'un BEP ou d'un CAP (b), les titulaires d'un baccalauréat général, technologique ou professionnel (c), les titulaires d'un diplôme du supérieur.

Les catégories socioprofessionnelles des parents

Les variables indiquant les catégories socioprofessionnelles des deux parents ont été établies sur la base de la nomenclature harmonisée des catégories sociales depuis 1968. Dans cette étude, à nouveau du fait du nombre relativement limité des observations, les 42

postes de cette nomenclature ont été regroupés en sept catégories : les agriculteurs (a), les artisans et les commerçants (b), les cadres (c), les professions intermédiaires (d), les employés (e), les ouvriers (f) et les anciens salariés ou indépendants (g).

Par ailleurs, nous utilisons les variables indicatrices d'immigration, de nationalité, de type d'activité, de condition d'emploi, de secteur d'activité et de pays de naissance. Certaines de ces variables ont été recodées en raison de changements de nomenclatures entre le recensement de 1999 et les enquêtes annuelles de recensement (notamment pour les nationalités, les catégories socioprofessionnelles et les diplômes).

Le tableau C3 suivant indique que les femmes sont légèrement plus souvent diplômées de l'enseignement supérieur que les hommes (26,4% pour les femmes, 23,7% pour les hommes). Ces statistiques sont compatibles avec celles émanant d'autres sources, comme celles de l'enquête emploi qui estime que les taux de diplômés de l'enseignement supérieur sont plus élevés pour les femmes que pour les hommes⁵⁴.

Tableau C3. Tableau des statistiques socio-démographiques des parents (pondérées) des enfants nés en 1990

Caractéristiques socio-démographiques	... du père (en %)	... de la mère (en %)
Diplôme		
Aucun diplôme	13,6	16,6
BEP, CAP ou brevet des collèges	36,4	36,9
Baccalauréat	11,3	17,2
Diplôme de l'enseignement supérieur	23,7	26,4
Inconnu	14,4	3,9
Catégorie socio-professionnelle		
Agriculteurs	3,1	1,2
Artisans et commerçants	9,7	2,9
Cadres	22,0	8,6
Professions intermédiaires	22,9	21,7
Employés	10,6	37,9
Ouvriers	29,9	7,5
Autres	1,9	20,1
Statut d'immigré		
Non immigré	88,4	88,6
Immigré	11,6	11,4

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

54 Voir par exemple les données de l'enquête emploi dans l'ouvrage France Portrait Social 2019: <https://www.insee.fr/fr/statistiques/4238409?sommaire=4238781>. 16,2 % des femmes âgées de 45 à 54 ans sont titulaires d'un bac+2 ou bac+3 contre 13,2 % des hommes, et 18,3 % d'entre elles sont titulaires d'un diplôme plus élevé contre 17,9 % des hommes.

Annexe D. Statistiques descriptives complémentaires et effets de sélection

Dans cette annexe, nous détaillons l'effet des différentes étapes de la constitution de la base de données sur les effectifs retenus, et l'effet de sélection induit par nos traitements. Nous cherchons également à savoir quel est le sens de l'effet de sélection sur les indicateurs de mobilité en étudiant les revenus des enfants et les caractéristiques des parents.

Taille des cohortes. Le tableau D1 suivant indique les effectifs d'individus obtenus pour chaque année de naissance, après différentes étapes dans la construction de la base de données. Les deux premières colonnes montrent l'effet de chacune des étapes présentées dans l'annexe B. Avec les données brutes (mais après les traitements mis en oeuvre pour obtenir la forme de la base voulue), et donc sans aucune restriction, les effectifs décroissent légèrement avec les années de naissance, surtout après la génération 1992 (qui a 18 ans en 2010, première colonne du tableau) : ils varient de 28 000 environ pour les enfants nés en 1988 à 16 000 pour les plus jeunes enfants. Cela est lié au fait que les enfants qui ont 18 ans ou moins en 2010 (les générations les plus récentes) sont mal identifiés dans l'EDP (cf. supra). L'intégration dans la base de données des enfants nés au cours de ces mêmes années mais n'apparaissant dans l'EDP qu'en 2011 et 2012 corrige un peu ce biais, mais les effectifs décroissent toujours avec les années de naissance (colonne 2). La sélection en colonne 3 des seuls enfants qui sont rattachés à la déclaration fiscale de leurs parents (ou de l'un de leurs parents) en 2010, 2011 ou 2012 (afin de récupérer l'information sur les revenus de leurs parents) entraîne une réduction plus importante des effectifs d'enfants les plus âgés : ainsi, plus de 98 % des enfants nés après 1993 figurent toujours dans la colonne 3, contre 85 % des enfants nés en 1992, un peu plus de 55 % des individus nés en 1990 et 26 % des enfants nés en 1988. Nous excluons donc la génération 1988 de nos analyses (tableau D2) car l'effet de sélection est trop fort (perte de 74 % des effectifs).

La taille des effectifs diminue encore avec de nouvelles restrictions : les parents ont des revenus non nuls (colonne 4), les enfants sont présents dans les données fiscales en 2018, avec des revenus positifs ou nuls et ne sont pas rattachés au foyer fiscal des parents (colonne 5), La Réunion et la Martinique sont exclues des données (colonne 6), et les étudiants repérés dans l'EAR 2019 sont exclus des données (colonne 7). Au final, il y a près de 14 000 paires enfants-parents pour la génération 1990 (représentant 424 000 paires après pondération) et près de 60 000 paires dans l'analyse principale cumulant les cohortes 1989 à 1992 (représentant près de 2 millions de personnes).

Tableau D1. Taille (non pondérée) des cohortes d'individus par année de naissance selon différentes restrictions

Génération	(1) Données brutes - présents en 2010	(2)=(1)+ présents en 2011 ou 2012	(3)=(2)+ enfants dans le foyer fiscal des parents en 2010-11-12	(4)=(3)+ revenus des parents non nuls	(5)=(4) + enfants ayant des revenus, et déclarants ou conjoints en 2018	(6) = (5) + hors DOM	(7) = (6) + hors étudiants repérés dans l'EAR 2019	Nombre d'individus pondérés après tout filtre
1988 (30 ans)	27 875	30565	8090	8084	6580	6459	6456	203 412
1989 (29 ans)	27 520	30596	12002	11993	9908	9715	9709	304 011
1990 (28 ans)	26 896	30939	16970	16956	13997	13735	13727	423 836
1991 (27 ans)	24 844	30247	21869	21865	18076	17743	17707	561 655
1992 (26 ans)	17 479	27075	23298	23297	18940	18546	18496	785 305
1993 (25 ans)	15 805	22624	22212	22212	18246	17916	17846	709 907
1994 (24 ans)	15 900	17298	17235	17235	14173	13916	13806	549 519
89-90-91-92	96 739	118 857	74 139	74 111	60 921	59 739	59 639	2 074 809

Lecture: la cohorte des individus nés en 1990 est constituée de 26 896 individus dans les données brutes.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Effet de sélection. Quel est l'impact de cet effet de sélection ? Pour répondre à cette question, on compare (i) le revenu des enfants sans et avec la sélection (tableau D2), et (ii) les caractéristiques des parents sans et avec la sélection (tableaux D3 et D4).

(i) On mesure les revenus des enfants en 2018⁵⁵, d'une part, dans la population totale (les deux premières colonnes du tableau D2) et, d'autre part, pour les seuls individus rattachés à leurs parents dans les déclarations fiscales (colonnes 3 et 4). Les revenus des individus rattachés aux déclarations fiscales de leurs parents sont un peu plus élevés que ceux de l'ensemble des enfants (sans restriction liée au foyer fiscal), surtout pour les générations les plus âgées (1988 et 1989) : cet effet décroît avec les années de naissance (il est de 5 % pour la génération 1990, de 2 % pour la génération 1992, et nul après 1993). Il semble donc surtout exister un biais à la hausse sur les revenus des générations 1988-1989. Ce biais est beaucoup plus faible pour les enfants nés entre 1990 et 1993.

Tableau D2. Les revenus des enfants en 2018 selon le critère de sélection

Génération	Population totale (sans restrictions)		Enfants dans le foyer fiscal des parents		Différence en %	
	moyenne	mediane	moyenne	mediane	moyenne	mediane
1988	21112	19872	24352	23179	15%	17%
1989	20851	19800	22810	21797	9%	10%
1990	20236	19446	20995	20502	4%	5%
1991	19351	19170	19566	19571	1%	2%
1992	17611	18111	17490	18000	-1%	-1%
1993	15754	16752	15735	16743	0%	0%
1994	13293	14601	13295	14608	0%	0%
1995	10605	11350	10605	11353	0%	0%
1996	8185	6592	8186	6592	0%	0%

Champ : Enfants observés en 2010, sans restriction particulière

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

(ii) Pour se faire une idée du biais de sélection sur les indicateurs de mobilité, il faut mesurer le biais induit par nos traitements sur les revenus des parents. Cette approche n'est malheureusement pas réalisable car nous ignorons les revenus de ceux dont les enfants ne sont plus dans le foyer fiscal. Néanmoins, il est possible d'étudier l'effet du biais sur la répartition des parents par CS et par diplôme, qui sont des variables proxy des revenus. Les tableaux D3 et D4 montrent la répartition en pourcentages des parents par catégories sociales (tableau D3) et selon le diplôme (tableau D4), d'abord sur la population totale (colonnes 1 et 2), puis sur le seul champ des enfants rattachés au foyer fiscal de leurs parents en 2010 (colonnes 3 et 4). La différence entre cette dernière population (celle de l'étude) et la population totale (colonnes 5 et 6) montre que la sélection conduit à légèrement sur-représenter les cadres (environ un point de pourcentage) et les diplômés du supérieur (un peu plus d'un point) dans nos données, et à réduire les parts des ouvriers et des non diplômés. Les différences sont cependant faibles, ce qui est rassurant : la sélection ne déforme pas fortement les caractéristiques des parents. Elle conduit donc en principe à conserver des parents qui sont un peu plus aisés que dans l'ensemble de la population. Comme la sélection sur-représente également les enfants un peu plus aisés, l'effet du biais sur les indicateurs de mobilité est incertain. En effet, dans le cas où la sélection contribuait à augmenter les revenus des enfants et à abaisser ceux des parents, elle entraînerait alors une sur-estimation de la mobilité. Mais comme elle biaise dans le même sens les revenus des enfants comme ceux des parents, l'effet de la sélection sur la mobilité est incertain et devrait même être faible.

55 Observés en 2010 et avant réintégration des années 2011 et 2012 (opération détaillée dans l'annexe B).

Tableau D3. Catégorie socioprofessionnelle des pères et des mères selon le critère de sélection

	Proportion en %				Différence (en pt)	
	population totale		foyer fiscal parents		mère	père
	mère	père	mère	père		
agriculteurs	1,2	3,5	1,1	3,4	-0,1	-0,1
acc	2,9	9,5	2,9	9,7	0,0	0,2
cadres	7,8	17,8	8,3	19,0	0,5	1,2
prof. Int.	19,6	21,2	20,3	21,5	0,7	0,4
employes	39,6	10,1	39,0	10,0	-0,6	0,0
ouvriers	8,7	35,9	8,1	34,2	-0,5	-1,7
autres	20,2	2,2	20,2	2,2	0,0	0,1

Tableau D4. Diplômes des pères et des mères selon le critère de sélection

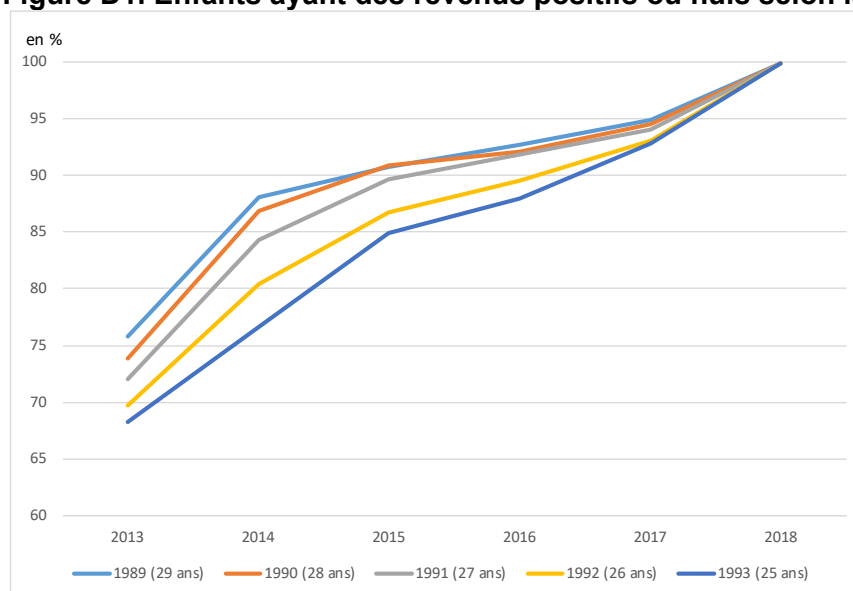
	proportion en %				Différence (en pt)	
	population totale		foyer fiscal parents		mère	père
	mère	père	mère	père		
Sans diplôme	18,4	16,0	17,8	15,3	-0,6	-0,7
Diplôme inf. au bac.	36,6	38,5	35,3	37,1	-1,3	-1,4
Bac. ou équivalent	16,6	11,3	16,9	11,5	0,3	0,1
Diplôme sup. au bac.	25,0	20,8	26,1	22,0	1,2	1,1

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Autres statistiques

La figure D1 montre le pourcentage d'enfants ayant des revenus positifs ou nuls entre 2013 et 2018 selon leur génération. Le champ est celui des enfants déclarants ou conjoints de déclarants en 2018. Le taux d'enfant ayant des revenus augmente au cours du temps. Il atteint près de 100 % en 2018.

Figure D1. Enfants ayant des revenus positifs ou nuls selon la génération



Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Les figures D2 et D3 présentent des statistiques sur les revenus des parents et les revenus des enfants (ayant entre 26 et 29 ans en 2018) en fonction du rang ou des revenus des parents. Ces figures sont commentées dans la partie 3 de l'article.

Figure D2. Revenus des parents en fonction du centile de revenu des parents

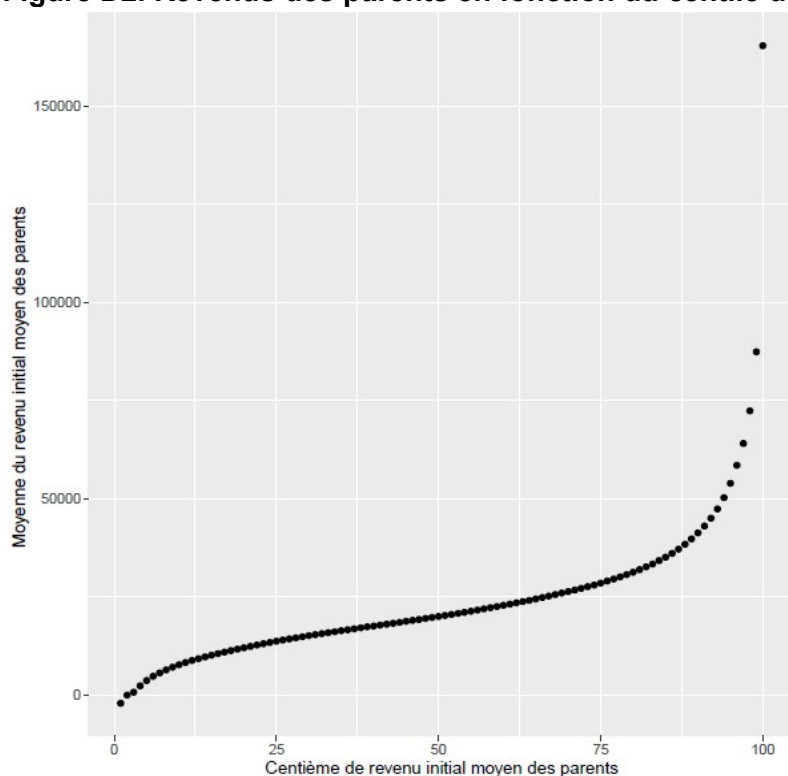
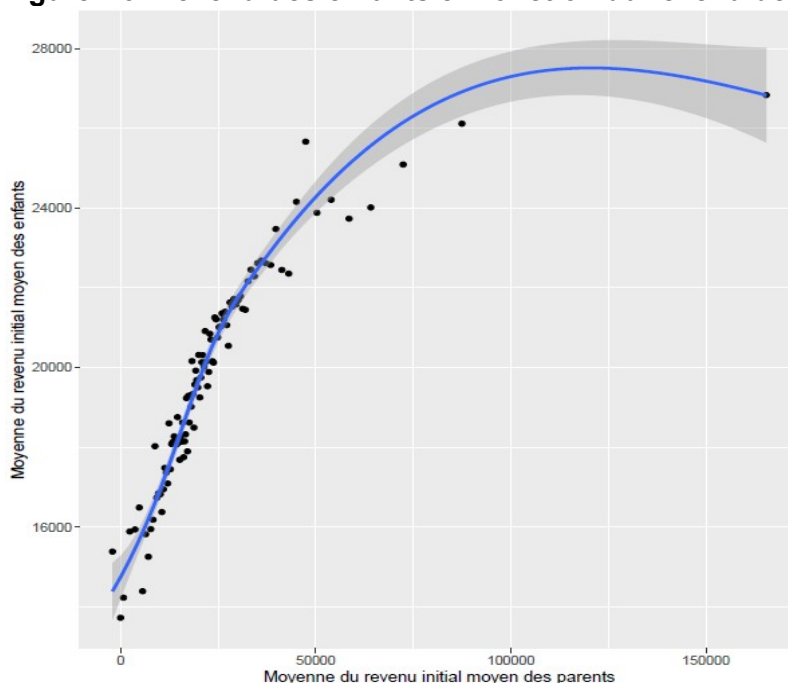


Figure D3. Revenu des enfants en fonction du revenu des parents

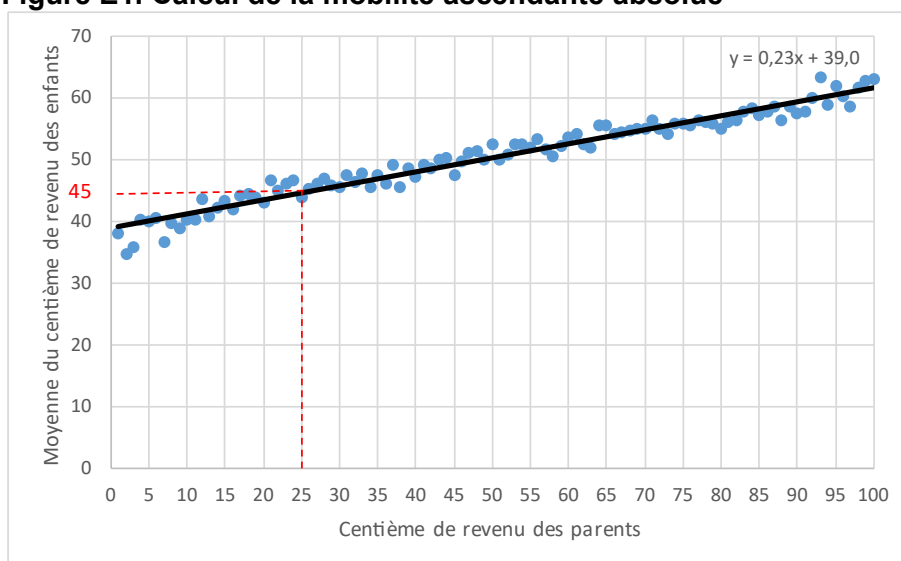


Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Annexe E : Résultats complémentaires

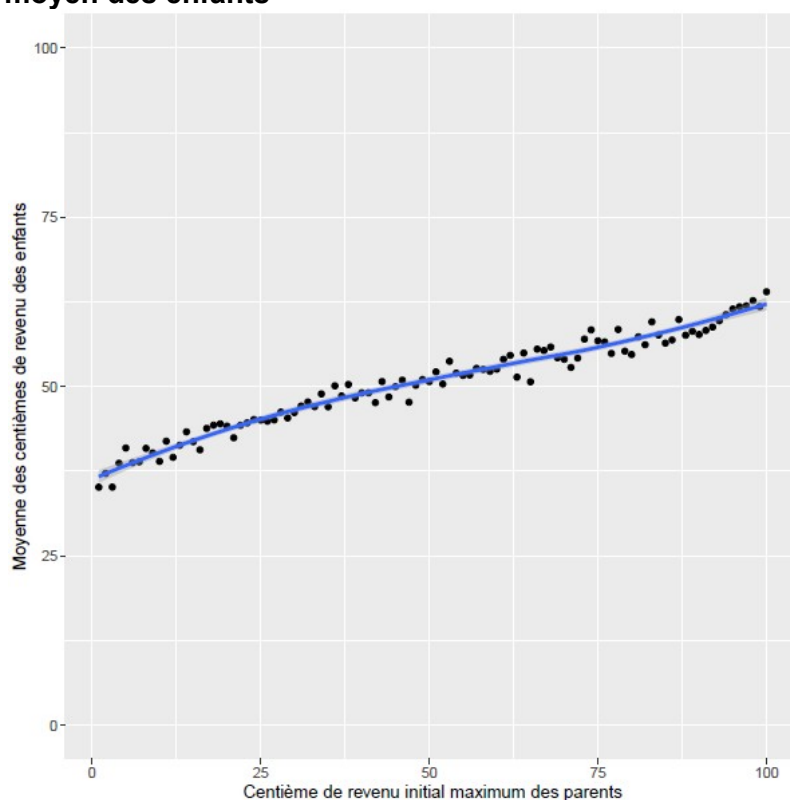
Figure E1. Calcul de la mobilité ascendante absolue



Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E2. Relation entre le centième de revenu maximum des parents et le rang moyen des enfants



Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

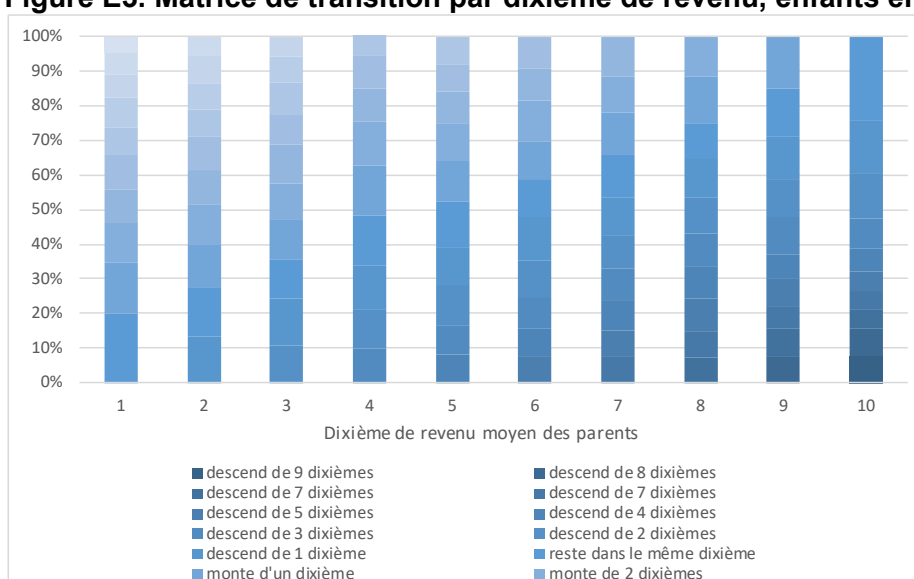
Tableau E1. Pente rang-rang et mobilité ascendante absolue selon la génération et le revenu des parents pris en compte

	revenu moyen		Niveau de vie		Revenu maximum	
	Pente rang-rang	mobilité absolue	Pente rang-rang	mobilité absolue	Pente rang-rang	mobilité absolue
1988 (30 ans)	0,24	44,4	0,23	38,9	0,24	38,1
1989 (29 ans)	0,25	44,2	0,25	38,1	0,25	38,1
1990 (28 ans)	0,24	44,3	0,25	38,0	0,26	37,5
1991 (27 ans)	0,22	44,8	0,23	39,0	0,23	38,8
1992 (26 ans)	0,21	45,0	0,22	39,3	0,22	39,6
1993 (25 ans)	0,19	45,6	0,20	40,5	0,19	40,7
1994 (24 ans)	0,14	46,8	0,14	43,3	0,15	43,0
1995 (23 ans)	0,07	48,6	0,08	46,6	0,07	46,8
89-90-91-92	0,23	44,7	0,23	38,8	0,23	38,7

Champ : France métropolitaine. Individus rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E3. Matrice de transition par dixième de revenu, enfants entre 26 et 29 ans



Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Tableau E2. Statistiques des matrices de transition selon la génération

générations	bas20/bas20	bas20/haut20	haut20/haut20	haut20/bas20	Stabilité dans le même cinquième
	plancher collant	mobilité ascendante	plafond collant	mobilité descendante	
1988 (30 ans)	29%	15%	34%	15%	73%
1989 (29 ans)	31%	12%	35%	15%	73%
1990 (28 ans)	31%	12%	34%	15%	72%
1991 (27 ans)	30%	11%	34%	18%	73%
1992 (26 ans)	31%	12%	33%	18%	73%
1993 (25 ans)	30%	12%	33%	20%	74%
1994 (24 ans)	28%	11%	29%	24%	75%
1995 (23 ans)	28%	14%	23%	25%	77%
1996 (22 ans)	27%	16%	20%	25%	78%
89-90-91-92 (26-29 ans)	31%	12%	34%	17%	73%

Champ : France métropolitaine. Individus rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Tableau E3. Matrice de transition avec la variable de niveau de vie des parents

	bas20/bas20 plancher collant	bas20/haut20 mobilité ascendante	haut20/haut20 plafond collant	haut20/bas20 mobilité descendante
1989	30%	11%	34%	16%
1990	31%	12%	35%	15%
1991	30%	11%	33%	17%
1992	30%	12%	34%	17%
89-90-91-92	30%	11%	34%	16%

Champ : France métropolitaine. Individus rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

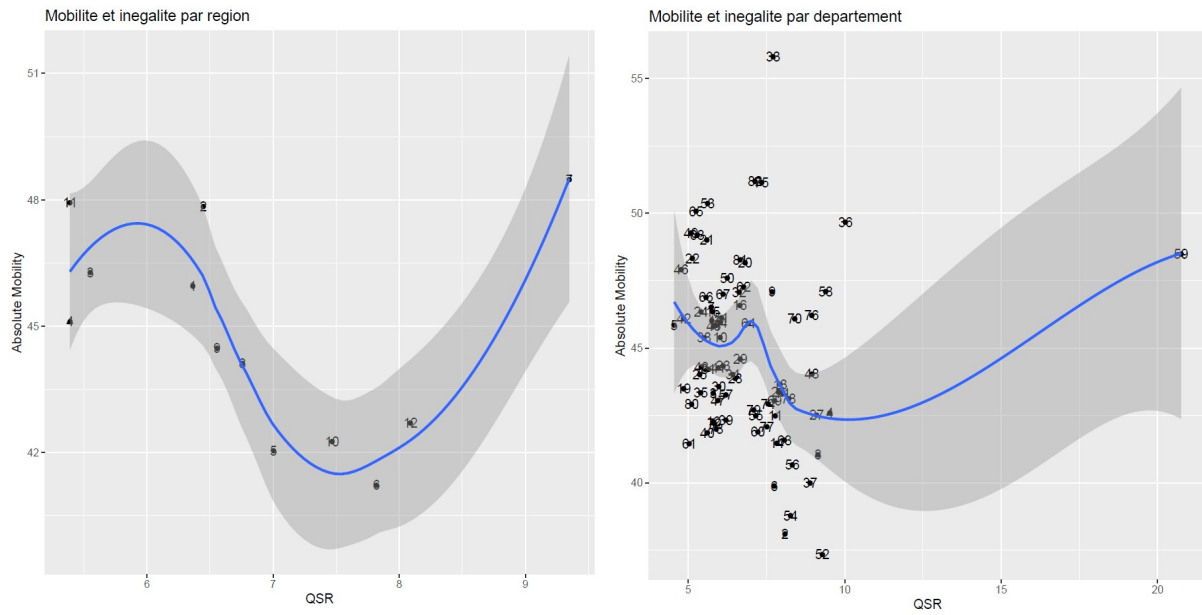
Tableau E4. Indicateurs de mobilité selon la région de naissance

	Nombre d'individus	Mobilité bas20/haut20	Mobilité bas40/haut40	Mobilité haut20/bas20	Stable bas20	Stable par cinquième	Ordonnée à l'origine	Pente rang- rang	Mobilité absolue
Auvergne-Rhone-Alpes	4276	14%	36%	15%	25%	73%	0,4	0,20	48,1
Bourgogne-Franche-Comte	1633	13%	35%	17%	26%	74%	0,3	0,20	48,1
Bretagne	1977	10%	30%	16%	20%	72%	0,3	0,20	47,3
Centre-Val de Loire	1393	13%	32%	17%	24%	75%	0,3	0,16	47,3
Grand Est	3620	10%	31%	18%	32%	74%	0,3	0,23	43,9
Hauts-de-France	4106	8%	25%	16%	33%	72%	0,3	0,28	42,3
Ile-de-France	9748	22%	39%	18%	27%	72%	0,4	0,17	49,4
Normandie	2108	12%	32%	16%	29%	74%	0,3	0,21	45,6
Nouvelle-Aquitaine	2654	11%	30%	18%	29%	74%	0,3	0,17	44,9
Occitanie	2476	13%	29%	18%	29%	73%	0,3	0,21	44,0
Pays de la Loire	2108	14%	33%	13%	16%	75%	0,3	0,17	48,6
Provence-Alpes-Cote d'Azur	2627	11%	29%	17%	32%	74%	0,3	0,23	44,1

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

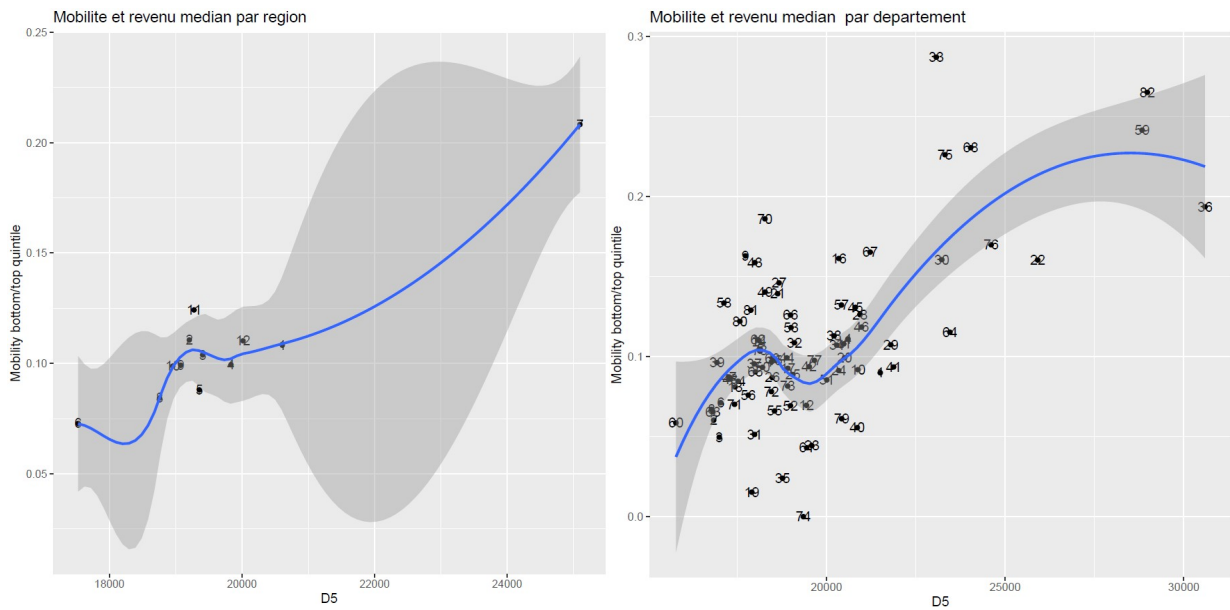
Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E4. Courbe de Gatsby par région et par département



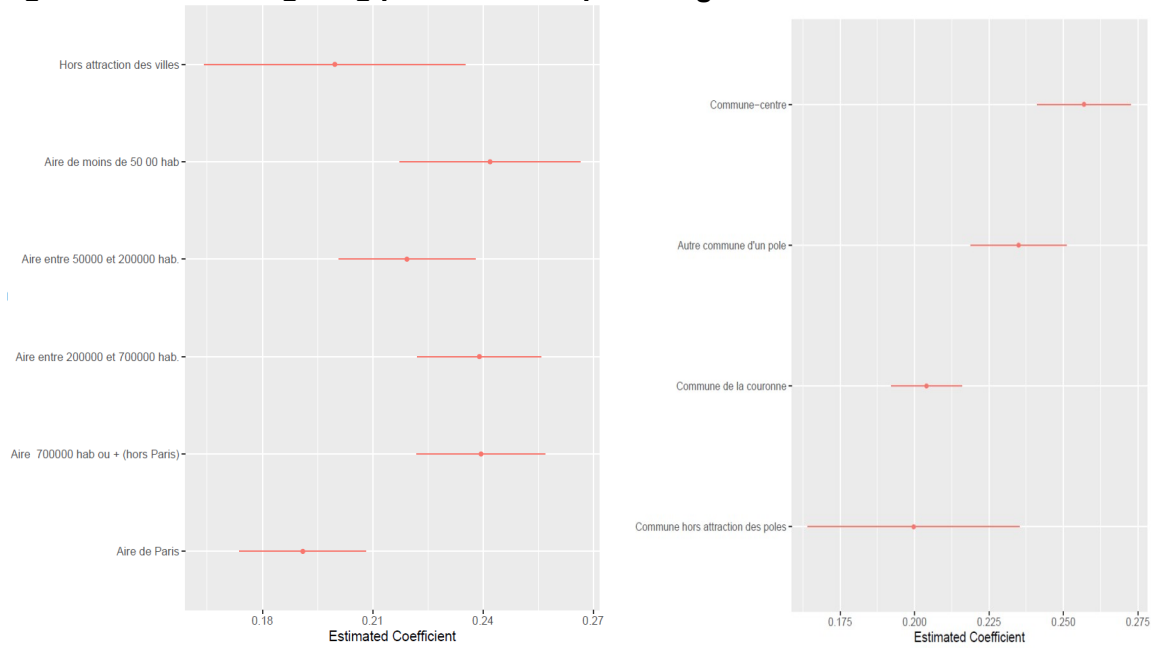
Note : Mobilité ascendante absolue (rang espéré des enfants de famille au 25 centième) et indice d'inégalité quintile share ratio (QSR) selon les régions et département.
 Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.
 Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E5. Mobilité ascendante et niveau de vie médian par région et par département
 7.A. Selon les régions 7.B. Selon les départements



Note : Mobilité ascendante (B20/T20) et revenu médian selon les régions et département.
 Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.
 Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

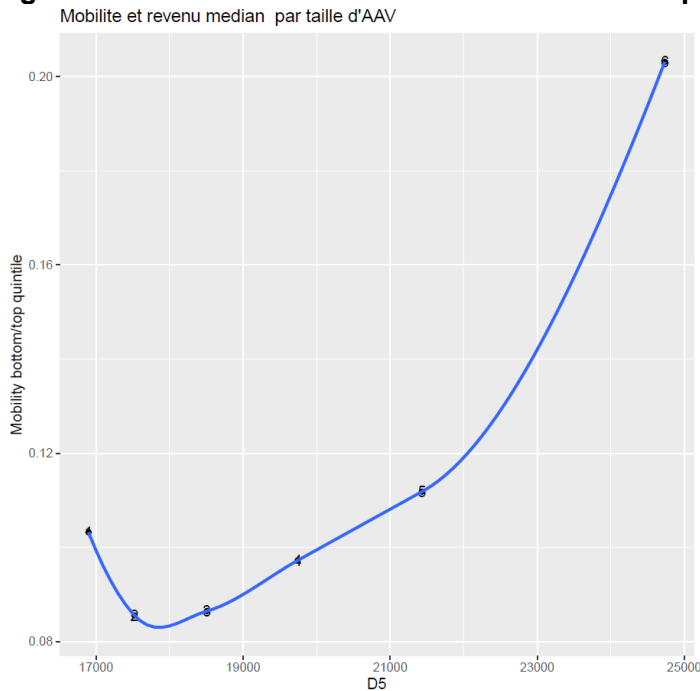
Figure E6. Pente rang-rang par tranche et par catégorie d'aire d'attraction des villes



Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

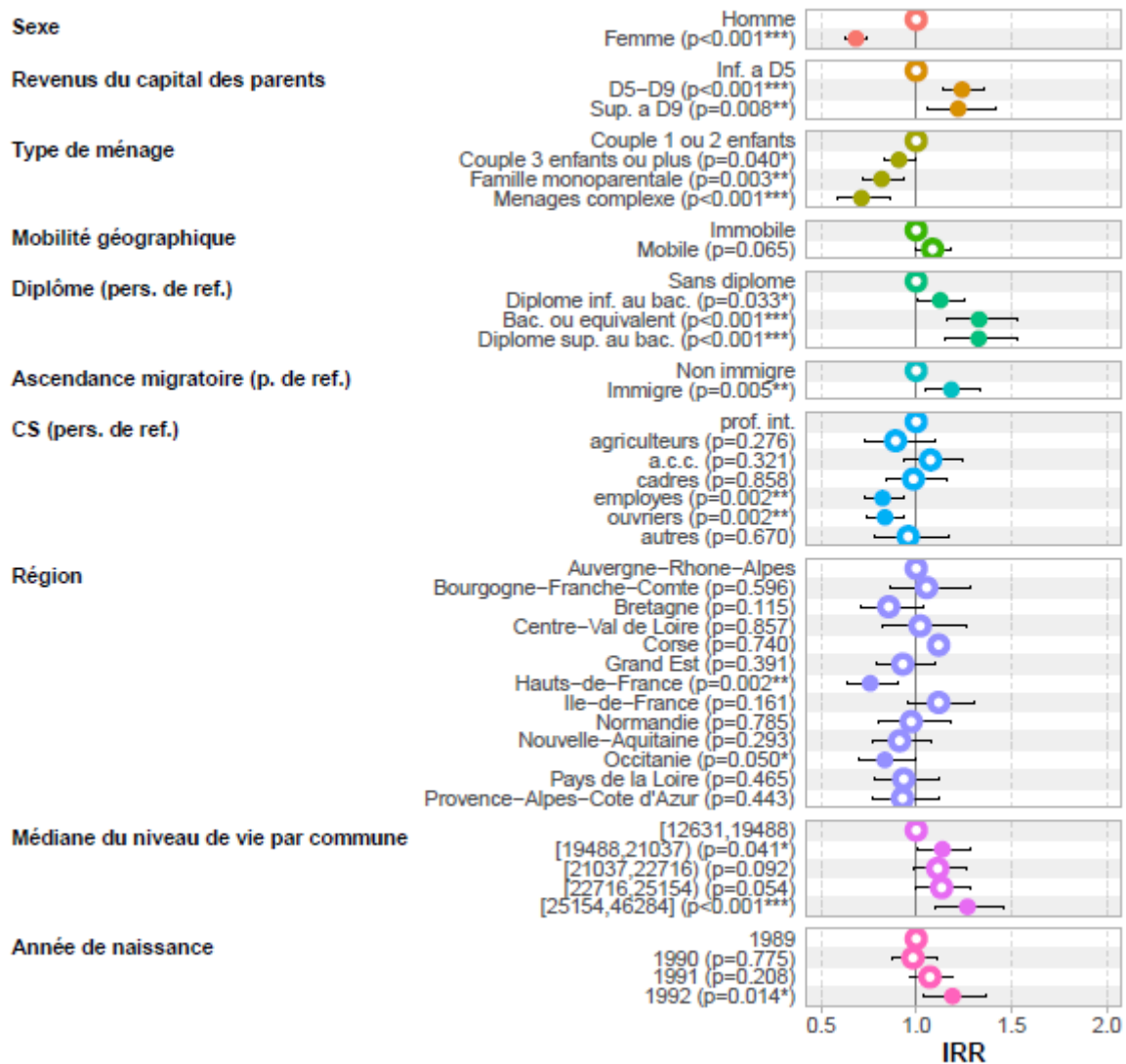
Figure E7. Mobilité ascendante et revenu médian par taille d'aire d'attraction des villes



Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E8. Régression de Poisson de la mobilité ascendante élargie avec contrôle sur le revenu de la commune

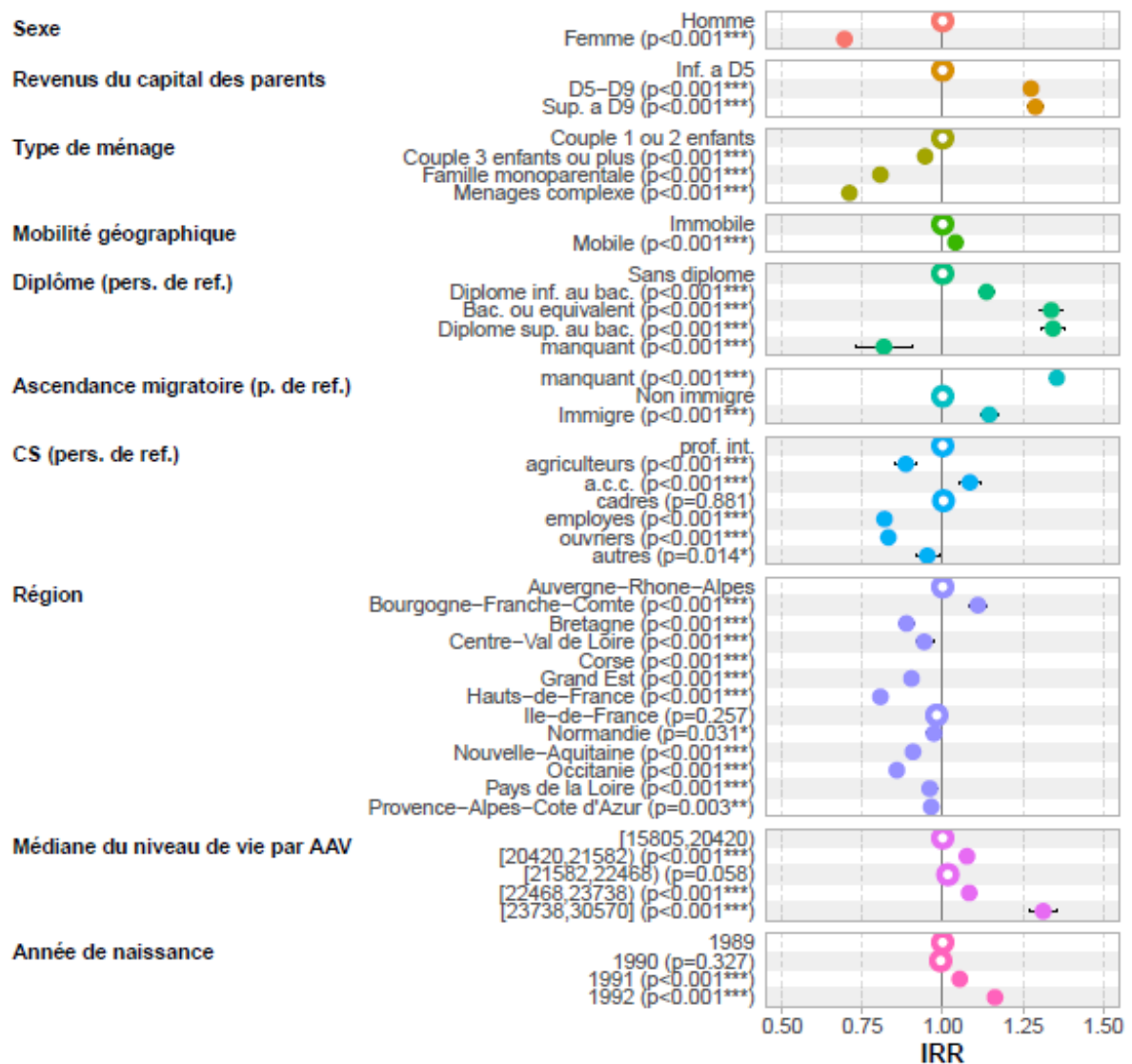


Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité ascendante des 40 % du bas de la distribution des revenus des parents vers les 40 % du haut de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{er} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018. Individus présents dans les EAR ou le RP.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E9. Régression de Poisson de la mobilité ascendante élargie avec contrôle sur le revenu de l'AAV : population totale

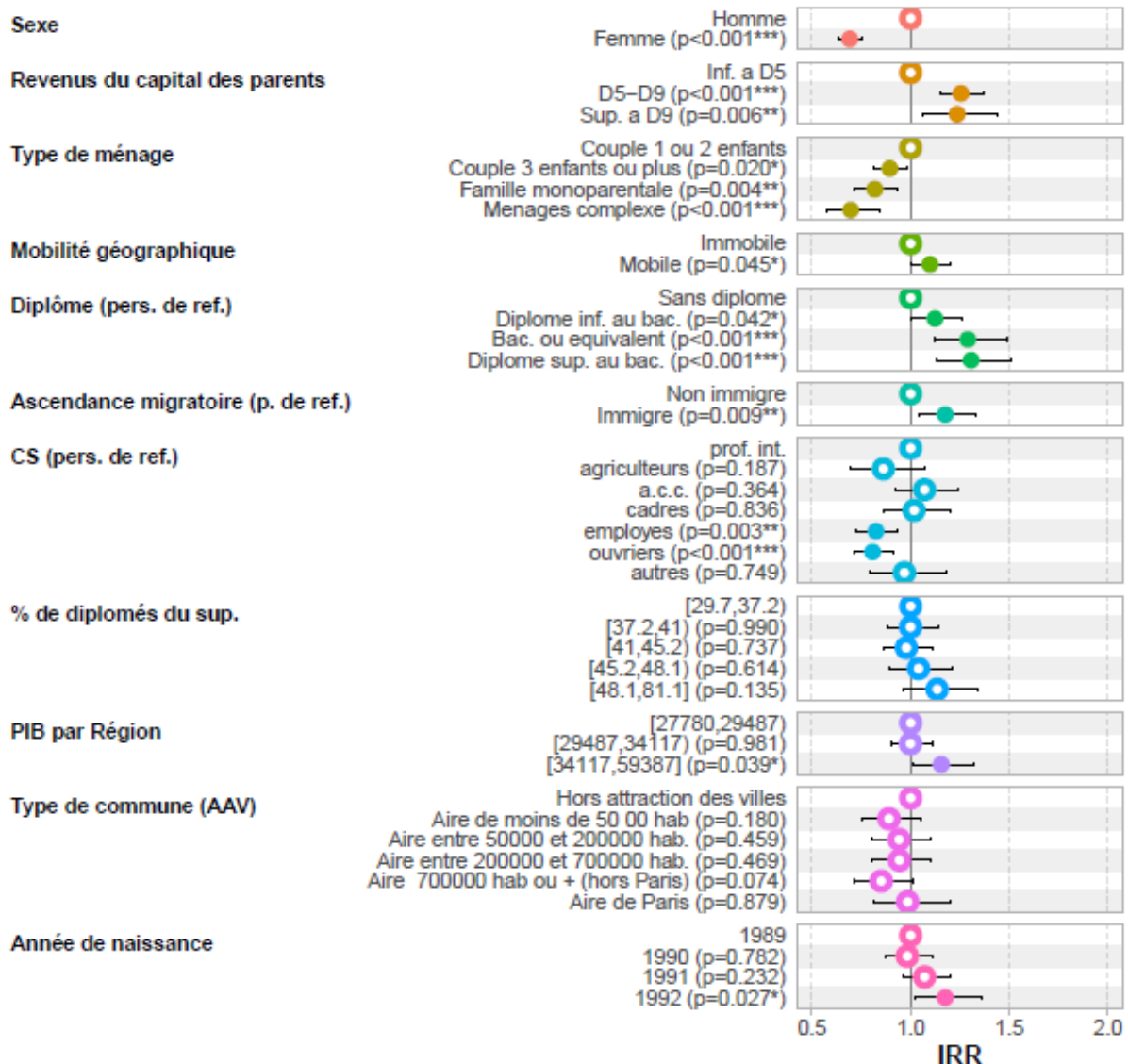


Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité ascendante des 40 % du bas de la distribution des revenus des parents vers les 40 % du haut de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{er} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E10. Régression de Poisson de la mobilité ascendante élargie selon différentes variables sur les territoires



Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité ascendante des 40 % du bas de la distribution des revenus des parents vers les 40 % du haut de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{er} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018. Individus présents dans les EAR ou le RP.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E11. Comparaison des régressions selon l'indicateur de mobilité utilisé

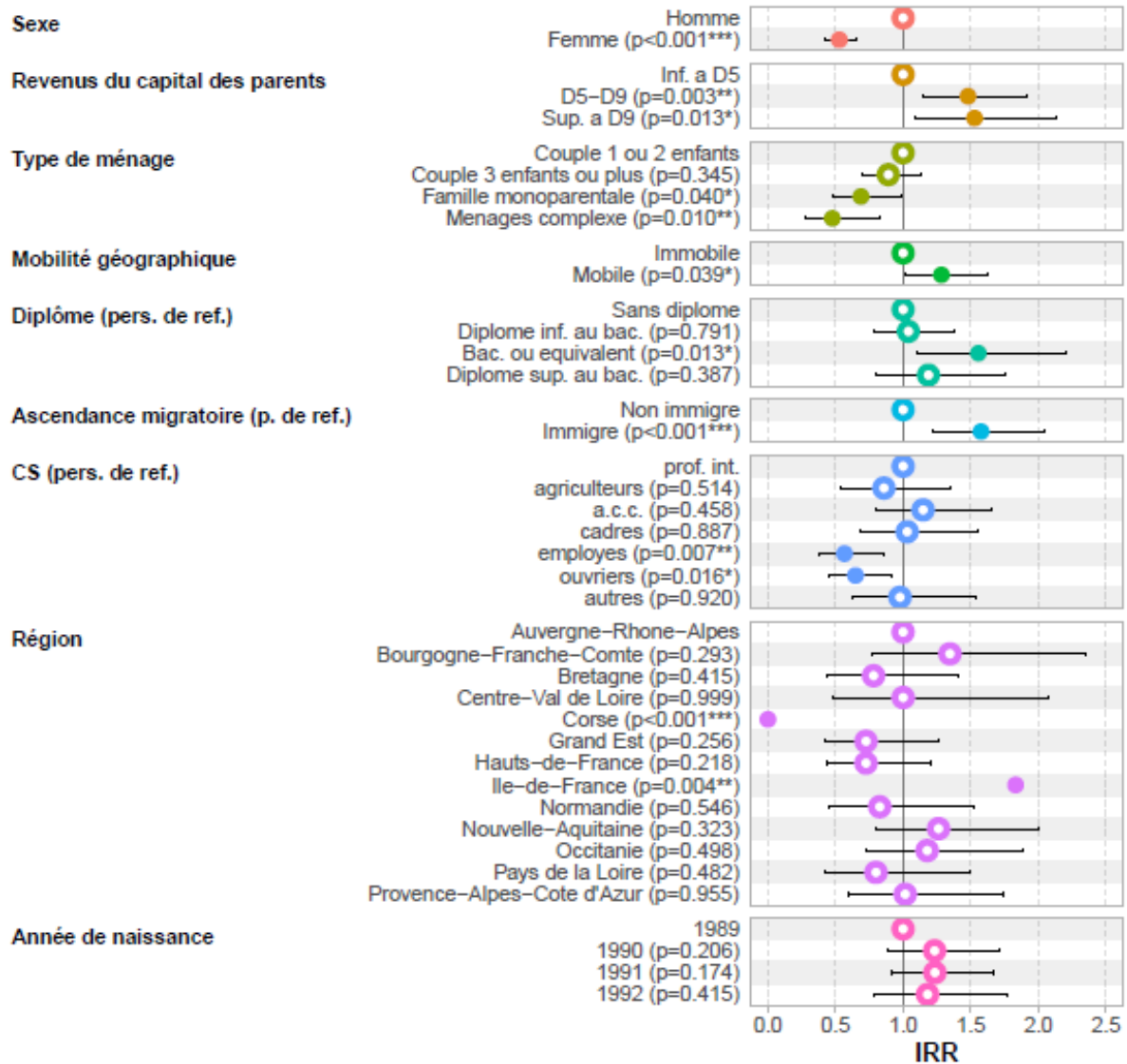
	B20/T20 <i>logistic</i>	B40/T40 <i>logistic</i>	Ascendant binaire <i>logistic</i>	Ascendant ordinale <i>cumulative link</i>					
	(1)	(2)	(3)	(4)					
Femme	-0.763*** (0.135)	-0.571*** (0.060)	-0.521*** (0.046)	-0.448*** (0.012)	Bourgogne-Franche-Comte	0.336 (0.338)	0.069 (0.158)	0.042 (0.123)	0.363*** (0.032)
Capital D5-D9	0.486*** (0.154)	0.330*** (0.068)	0.148*** (0.052)	0.497*** (0.015)	Bretagne	-0.289 (0.352)	-0.264* (0.147)	-0.129 (0.112)	-0.042 (0.030)
Capital >D9	0.601*** (0.220)	0.370*** (0.123)	0.428*** (0.098)	0.331*** (0.025)	Centre-Val de Loire	-0.014 (0.417)	0.048 (0.172)	-0.133 (0.133)	-0.114*** (0.039)
Couple 3 enfants ou plus	-0.106 (0.152)	-0.138* (0.070)	0.061 (0.054)	0.017 (0.014)	Corse	-12.800 (425.927)	0.169 (0.549)	-0.411 (0.481)	0.082 (0.165)
Famille monoparentale	-0.443** (0.210)	-0.292*** (0.097)	-0.031 (0.076)	-0.280*** (0.018)	Grand Est	-0.325 (0.317)	-0.128 (0.129)	-0.042 (0.099)	-0.127*** (0.026)
Menages complexe	-1.020*** (0.330)	-0.485*** (0.138)	-0.276** (0.109)	-0.806*** (0.026)	Hauts-de-France	-0.416 (0.292)	-0.449*** (0.124)	-0.206** (0.096)	-0.261*** (0.024)
Mobilité géographique	0.273* (0.147)	0.119* (0.069)	0.045 (0.053)	0.154*** (0.014)	Ile-de-France	0.801*** (0.258)	0.292** (0.126)	0.243** (0.095)	0.514*** (0.026)
diplome inf. au bac.	0.038 (0.176)	0.190** (0.079)	0.020 (0.064)	0.096*** (0.015)	Normandie	-0.223 (0.359)	-0.062 (0.148)	-0.003 (0.113)	0.029 (0.030)
bac. ou equivalent	0.536** (0.220)	0.450*** (0.108)	0.190** (0.085)	0.373*** (0.022)	Nouvelle-Aquitaine	0.267 (0.285)	-0.197 (0.133)	-0.182* (0.103)	0.092*** (0.027)
diplome sup. au bac.	0.124 (0.237)	0.455*** (0.113)	0.254*** (0.087)	0.167*** (0.023)	Occitanie	0.170 (0.285)	-0.308** (0.135)	-0.144 (0.104)	0.047* (0.027)
parent immigré	0.485*** (0.175)	0.163* (0.094)	0.417*** (0.076)	0.142*** (0.017)	Pays de la Loire	-0.341 (0.371)	-0.104 (0.143)	-0.058 (0.109)	0.210*** (0.030)
agriculteurs	-0.241 (0.282)	-0.199 (0.159)	0.855*** (0.133)	-0.067** (0.027)	Provence-Alpes-Cote d'Azur	-0.040 (0.325)	-0.109 (0.149)	-0.014 (0.114)	-0.050* (0.030)
Independants	0.196 (0.242)	0.152 (0.123)	0.434*** (0.097)	0.180*** (0.026)	generation 1990	0.262 (0.203)	-0.030 (0.092)	-0.114 (0.071)	0.057*** (0.016)
cadres	0.032 (0.292)	0.013 (0.138)	0.012 (0.095)	-0.048 (0.033)	generation 1991	0.269 (0.190)	0.101 (0.085)	-0.014 (0.065)	0.085*** (0.015)
employés	-0.758*** (0.246)	-0.310*** (0.097)	-0.120* (0.073)	-0.481*** (0.023)	generation 1992	0.217 (0.251)	0.253** (0.112)	0.025 (0.086)	0.012 (0.024)
ouvriers	-0.537** (0.214)	-0.286*** (0.090)	0.058 (0.069)	-0.359*** (0.020)	Constant	-1.953*** (0.351)	-0.583*** (0.153)	-0.742*** (0.117)	
autres CS	-0.070 (0.287)	-0.081 (0.157)	0.244* (0.127)	-0.171*** (0.029)	Observations	2,402	5,613	9,207	92,853
					Log Likelihood	-806.798	-3,339.150	-5,545.058	-141,835.200
					Akaike Inf. Crit.	1,679.596	6,744.300	11,156.120	

Note: *p<0.1; **p<0.05; ***p<0.01

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018. Individus présents dans les EAR ou le RP.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E12. Régression de Poisson de la mobilité ascendante (variable B20/T20)

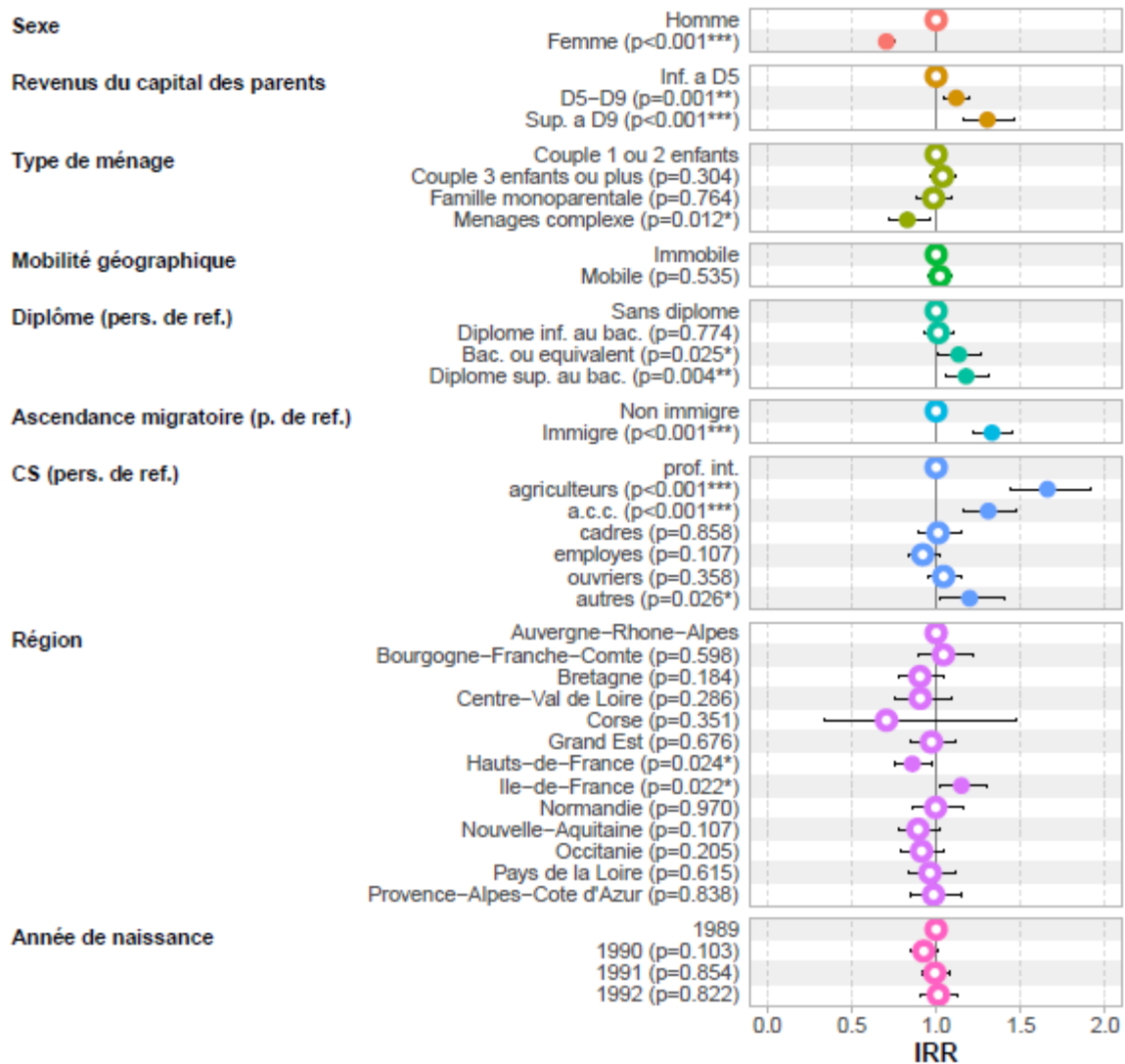


Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité ascendante des 20 % du bas de la distribution des revenus des parents vers les 20 % du haut de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{er} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018. Individus présents dans les EAR ou le RP.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E13. Régression de Poisson sur la mobilité « saut de cinquième »

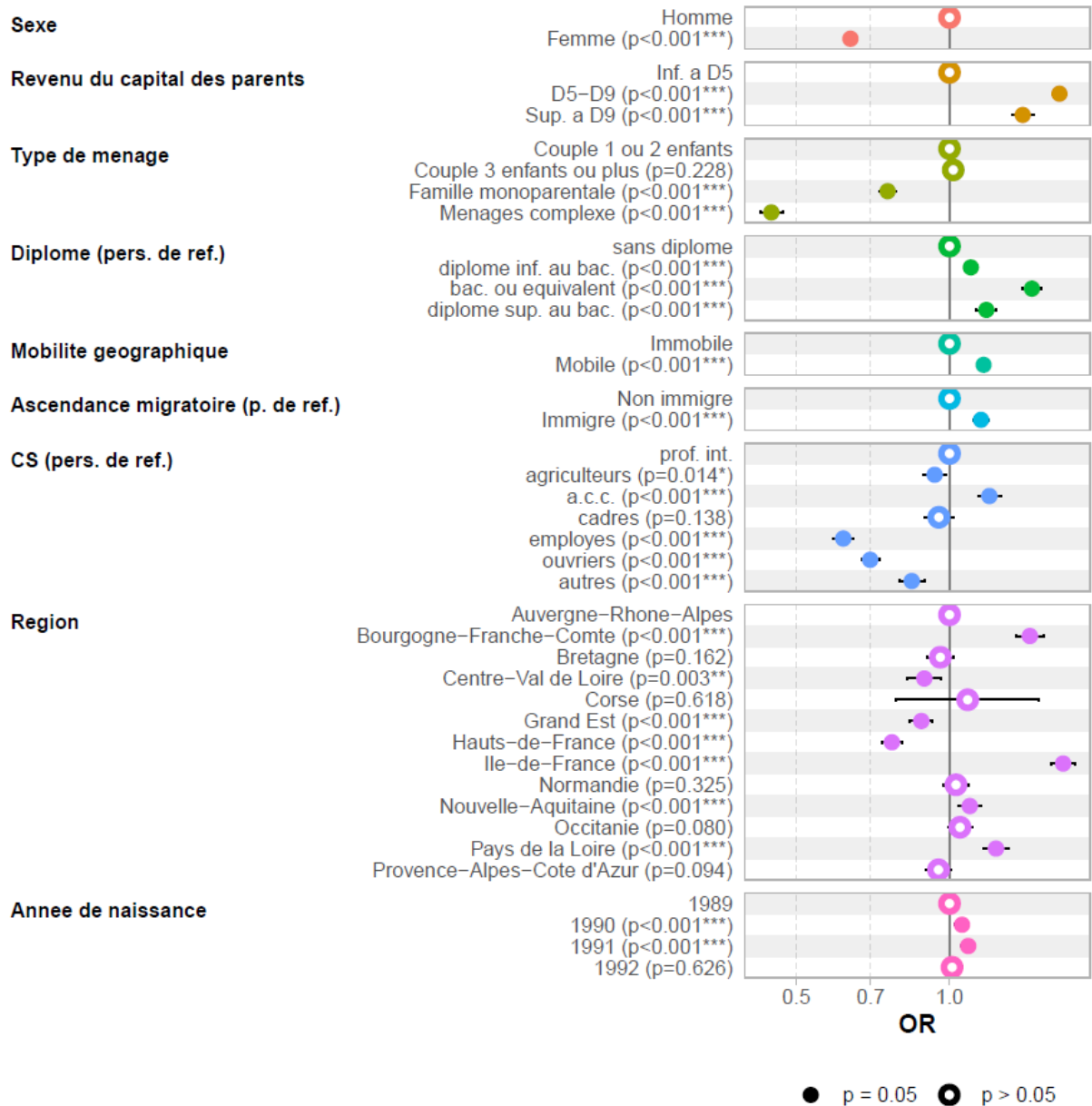


Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité ascendante en « sautant » un cinquième, c'est-à-dire en passant du 1^{er} au 3^e, du 2^e ou 4 ou du 3^e ou 5^e cinquième. Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018. Individus présents dans les EAR ou le RP.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

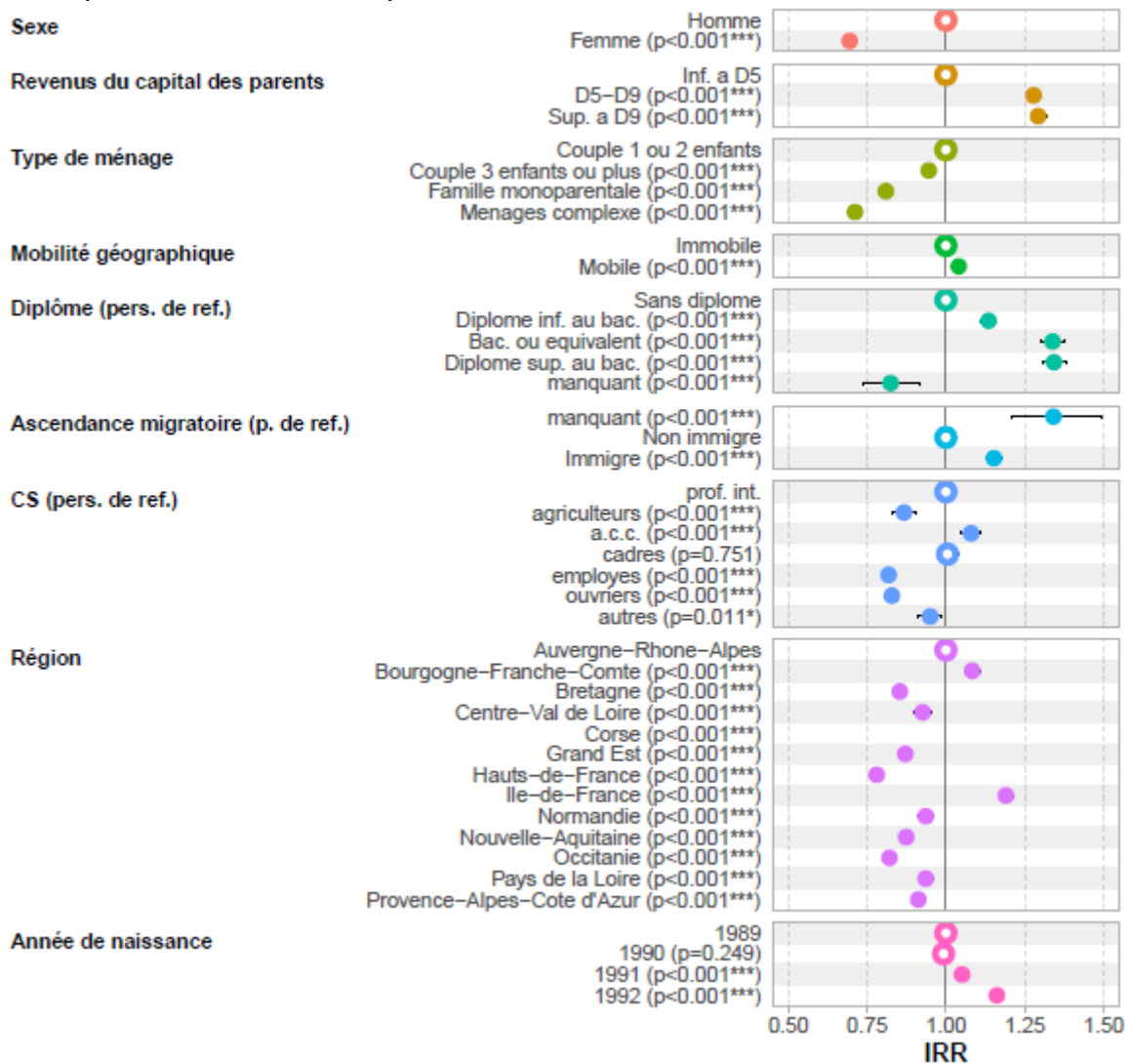
Figure E14. Régression multinomiale ordonnée sur la mobilité



Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018. Individus présents dans les EAR ou le RP.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E15. Régression de Poisson de la mobilité ascendante élargie sur la population totale (variable bas40/haut40)



Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité ascendante des 40 % du bas de la distribution des revenus des parents vers les 40 % du haut de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{er} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E16. Comparaison régression logistique et régression de Poisson, variable B20/T20

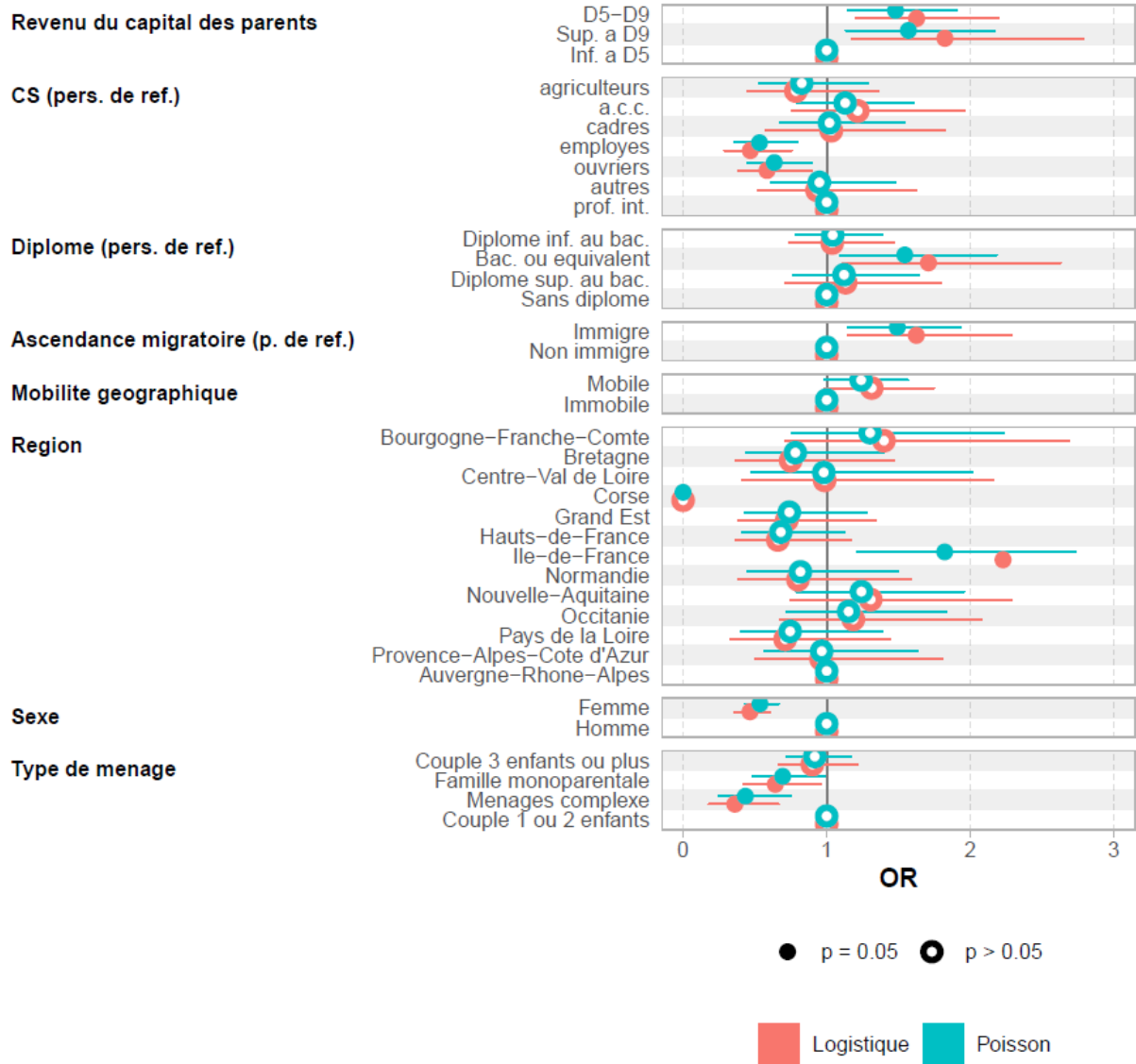


Figure E17. Comparaison régression logistique et régression de Poisson de la mobilité ascendante élargie

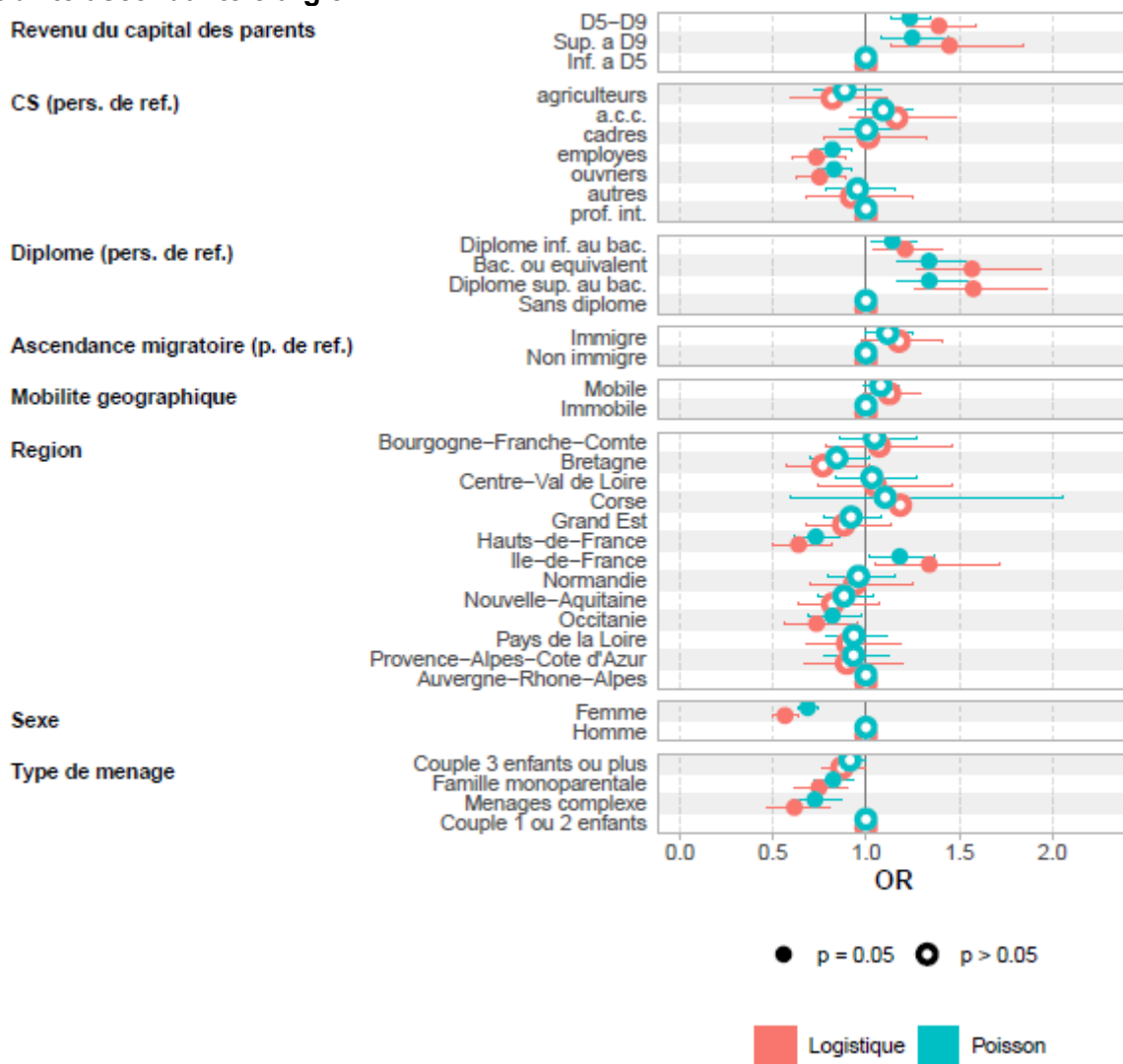


Figure E18. Régression de Poisson de la mobilité ascendante élargie (pondérée et non pondérée)

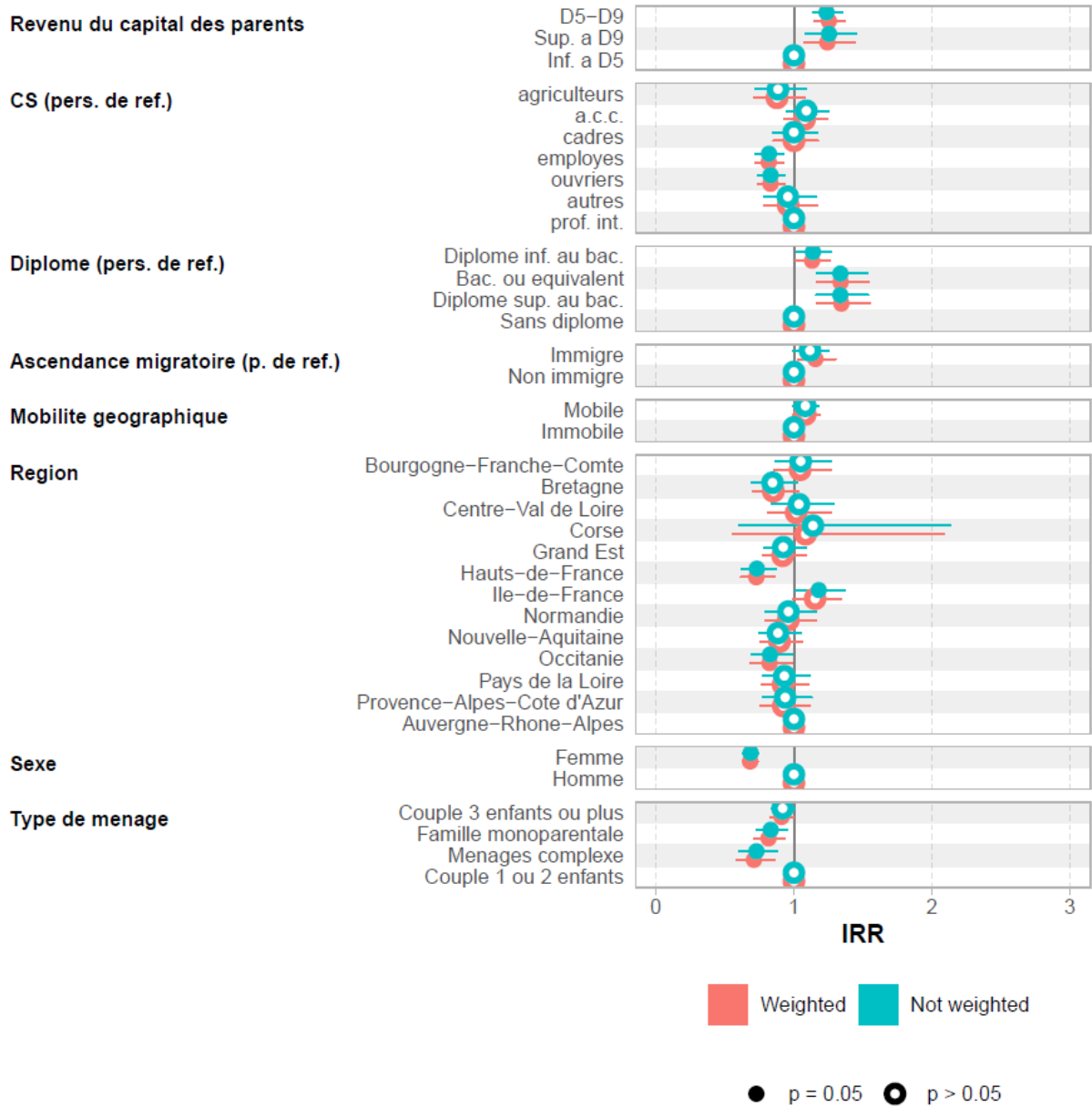
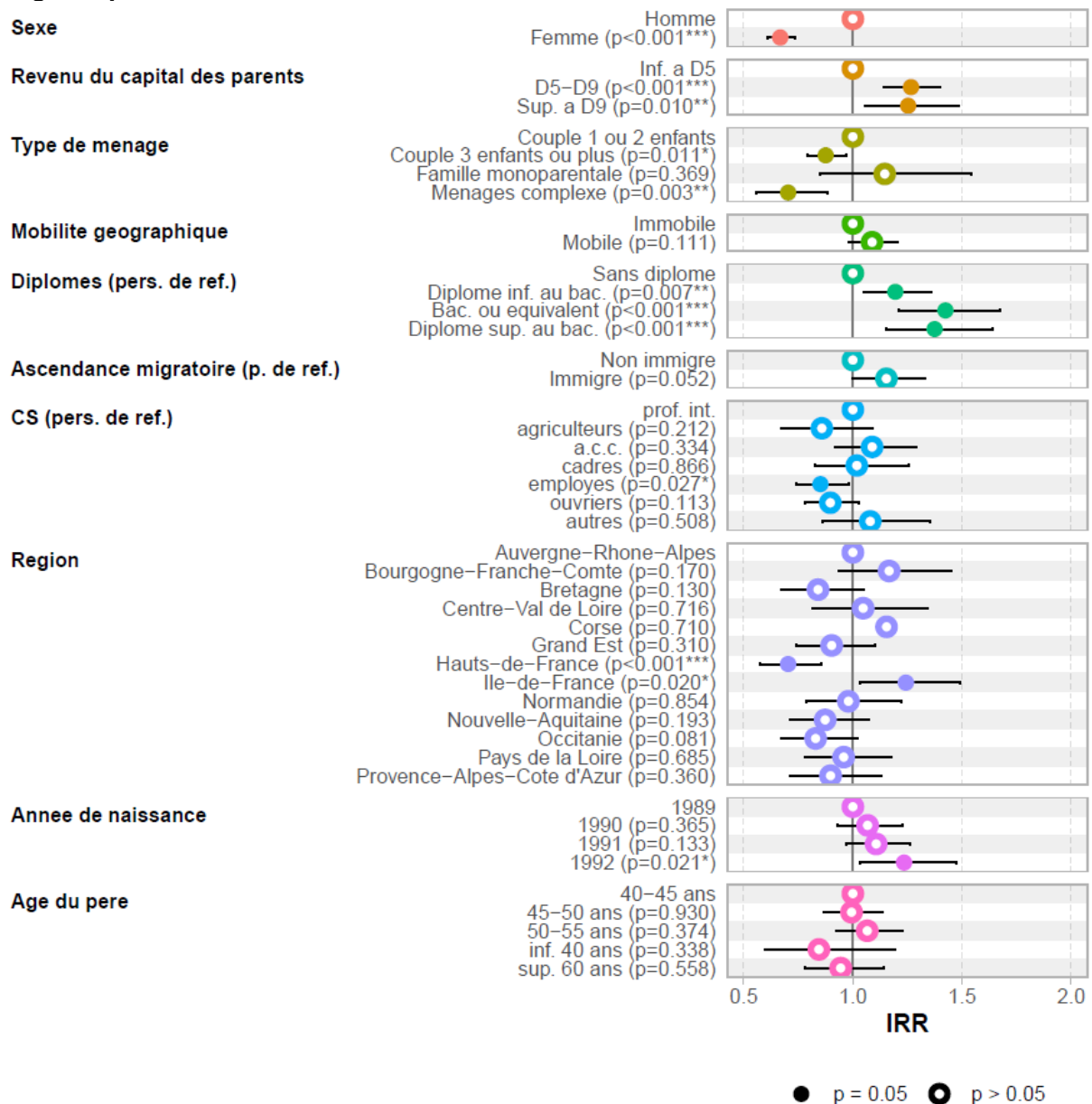


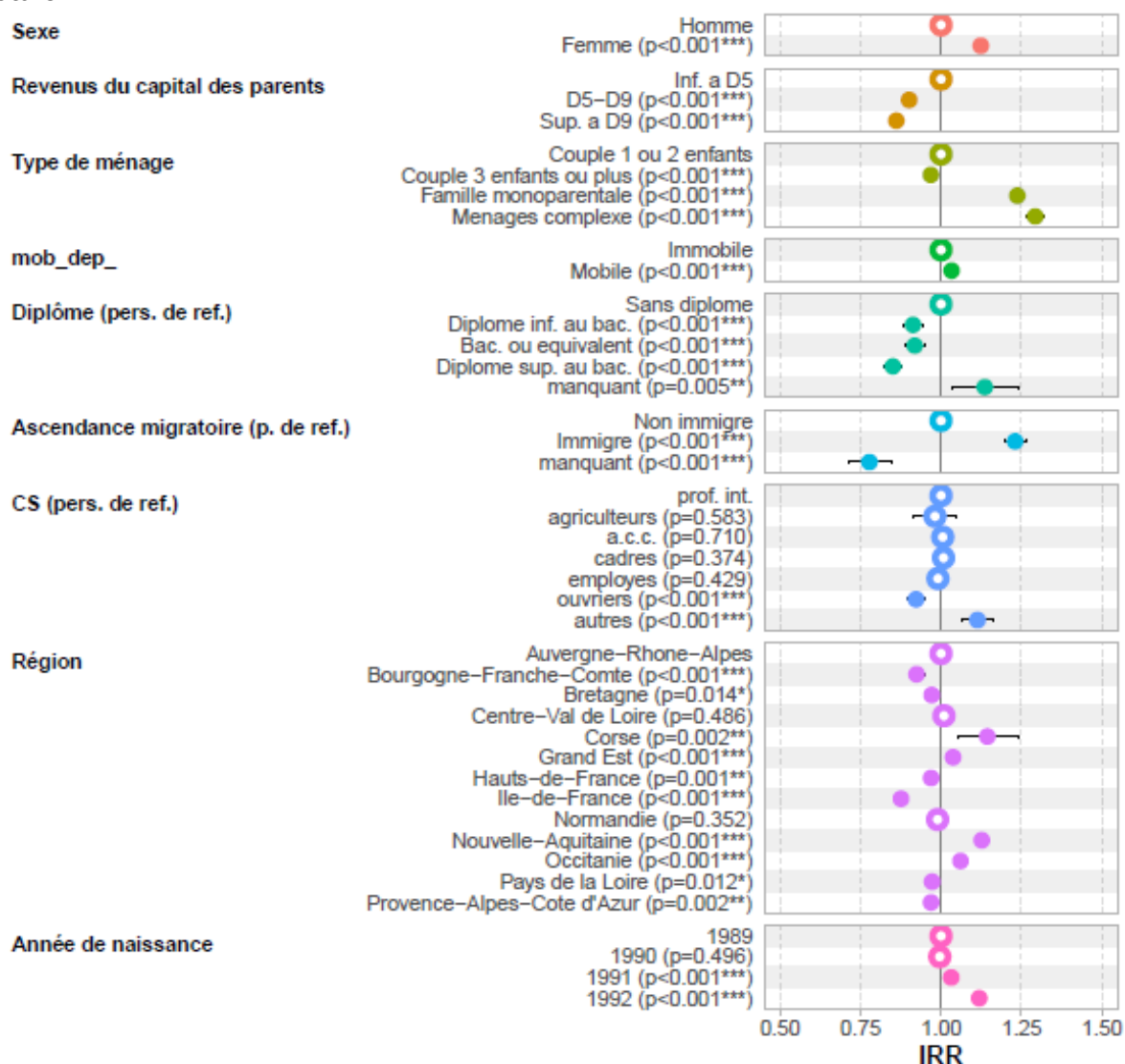
Figure E19. Régression de Poisson de la mobilité ascendante élargie avec contrôle de l'âge du père



Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité ascendante des 40 % du bas de la distribution des revenus des parents vers les 40 % du haut de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{er} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, ayant des revenus positifs ou nuls en 2018, et présents dans les EAR ou le RP.
Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E20. Régression de Poisson de la mobilité descendante T40/B40, population totale



Note : La figure présente un rapport de risques (RR) ou de chances de réaliser une mobilité descendante des 40 % du haut de la distribution des revenus des parents vers les 40 % du bas de la distribution des enfants, entre une certaine modalité et la modalité de référence (1^{er} modalité de chaque variable). Les ronds sans poids blanc à l'intérieur indiquent un RR significativement différent de 1.

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

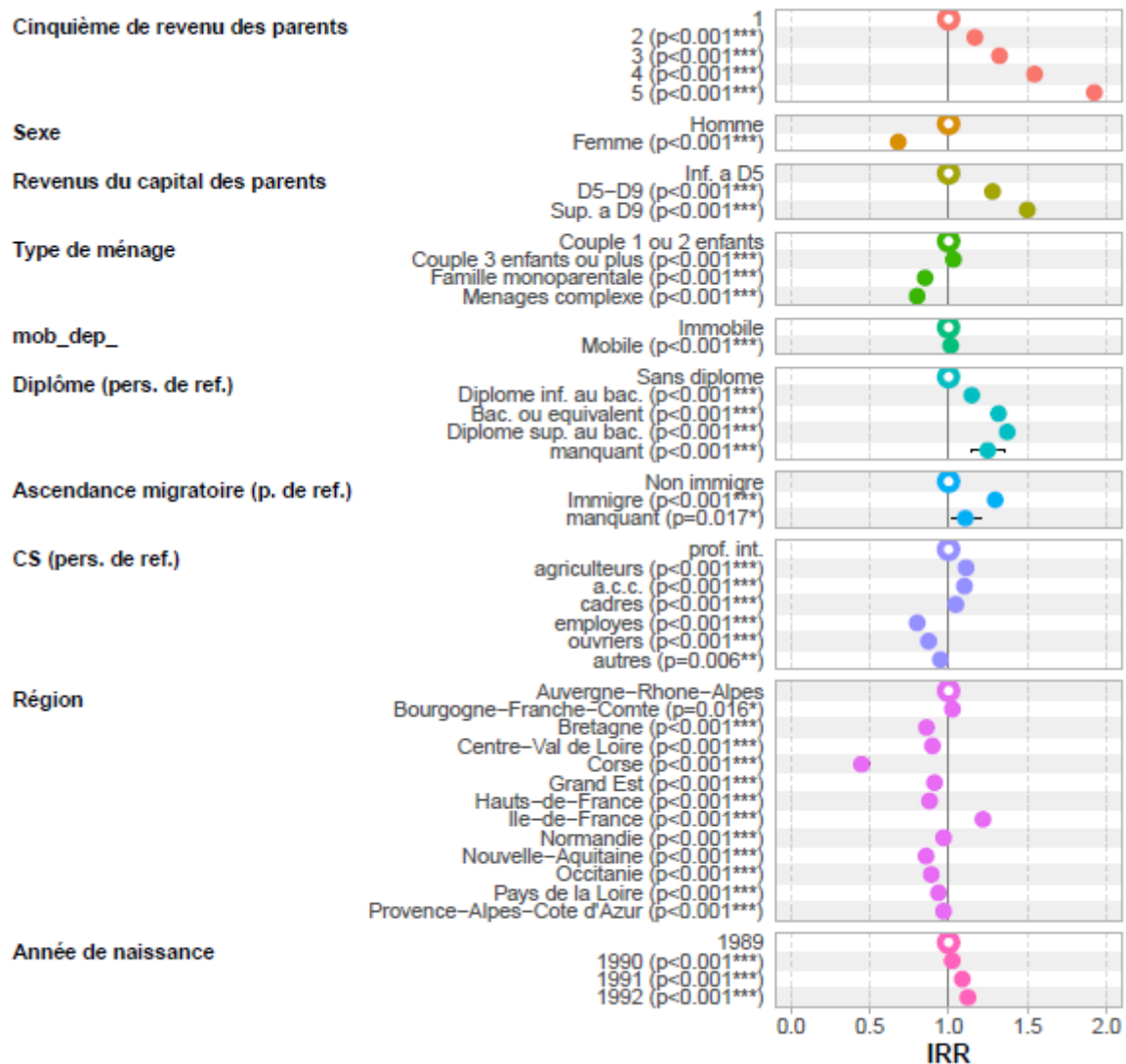
Tableau E5. Régressions linéaires des centiles de revenus des enfants sur les revenus des parents (en 5 modalités) et différentes autres variables

	Baseline (1)	(1)+contrôles individuels (2)	(2) + contrôles régions (3)	(3) + contrôles var. RP (4)
C	40.695 *** (0.259)	45.730 *** (0.559)	46.388 *** (0.660)	44.308 *** (1.248)
$\beta_{\text{cinquième 2}}$	5.851 *** (0.367)	4.996 *** (0.365)	4.635 *** (0.479)	4.574 *** (0.755)
$\beta_{\text{cinquième 3}}$	10.338 *** (0.367)	8.492 *** (0.369)	7.360 *** (0.474)	6.948 *** (0.757)
$\beta_{\text{cinquième 4}}$	14.319 *** (0.367)	11.317 *** (0.379)	10.249 *** (0.474)	9.118 *** (0.790)
$\beta_{\text{cinquième 5}}$	18.502 *** (0.367)	13.580 *** (0.401)	12.783 *** (0.491)	10.449 *** (0.861)
Contrôles var. individuelles		X	X	X
Contrôles régions d'appartenance			X	X
Contrôles diplôme et CS parents				X
Observations	58814	58736	40280	16681

Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, ayant des revenus positifs ou nuls en 2018, et présents dans les EAR ou le RP.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

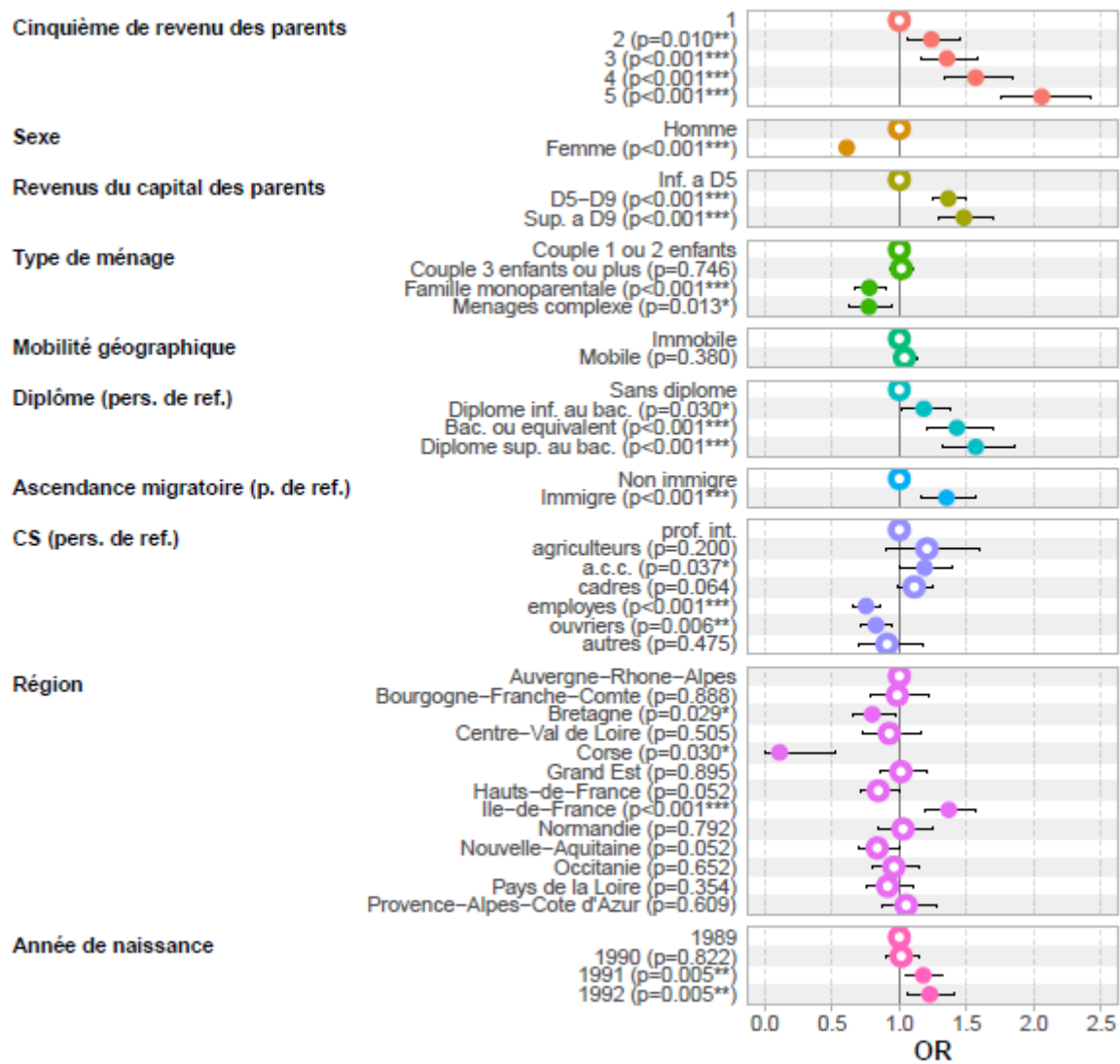
Figure E21. Régression de Poisson d'être dans le plus haut cinquième de revenu, population totale



Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, et ayant des revenus positifs ou nuls en 2018.

Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Figure E22. Régression logistique d'être dans le plus haut cinquième de revenu



Champ : France métropolitaine. Individus nés entre 1989 et 1992, rattachés à la déclaration fiscale de leurs parents en 2010, 2011 ou 2012, ayant des revenus positifs ou nuls en 2018, et présents dans les EAR ou le RP..
 Source : Insee-DGFIP-Cnaf-Cnav-CCMSA, échantillon démographique permanent 2019

Annexe F : Calcul mathématique de l'espérance de la différence de rang

On cherche la valeur moyenne théorique de la valeur absolue de la différence de centiles de revenu entre deux personnes tirées au hasard dans la population.

Les centiles suivent une distribution uniforme entre 0 et 100 mais pour simplifier les calculs, nous considérons des distributions entre 0 et 1 (avec donc des centiles allant de 0,01 à 0,99) et nous multiplierons le résultat par 100.

On considère deux variables aléatoires X et Y qui suivent une loi uniforme sur $[0,1]$ et qui sont indépendantes. En termes statistiques, on cherche $E[|X-Y|]$: l'espérance de la valeur absolue de la différence. Celle-ci est obtenue en cherchant, au préalable, la fonction de répartition et densité de $Y-X$, puis de $|Y-X|$.

Soit $Z=Y-X$ et $Z(\Omega)=[-1,1]$

On calcule alors la fonction de répartition de Z :

$$p(Z \leq z) = \int_0^1 p(y \leq Y \leq y+dy \cap -X \leq z-y)$$

$$p(Z \leq z) = \int_0^1 p(y \leq Y \leq y+dy) [1 - p(X \leq y-z)]$$

Examinons d'abord le cas où $z \in [0,1]$:

$$p(y \leq Y \leq y+dy) = dy \text{ et}$$

$$p(X \leq y-z) = y-z \text{ si } y-z \in [z,1] \text{ ou } p(X \leq y-z) = 0 \text{ si } y \in [0,z]$$

$$\text{Donc : } p(Z \leq z) = \int_0^1 [1 - p(X \leq y-z)] dy$$

$$p(Z \leq z) = \int_0^z dy + \int_z^1 [1 - p(X \leq y-z)] dy$$

$$p(Z \leq z) = y \Big|_0^z + \frac{-y^2}{2} + y(1+z) \Big|_z^1$$

$$p(Z \leq z) = z + \left(\frac{-1}{2} + (1+z) - \frac{z^2 + 2z}{2} \right)$$

et donc la fonction de répartition de Z pour $z \in [0,1]$ est :

$$p(Z \leq z) = 1 - \frac{(1-z)^2}{2}$$

Par des calculs similaires, si $z \in [-1,0]$ on a :

$$p(Z \leq z) = \frac{(1+z)^2}{2}$$

On définit alors W , la valeur absolue de Z : $W = |Y-X|$ avec $W(\Omega) = [0,1]$

$$p(W \leq w) = p(-w \leq W \leq w) = p(W \leq w) - p(W \leq -w) = 1 - \frac{(1-w)^2}{2} - \frac{(1+(-w))^2}{2}$$

$$p(W \leq w) = 2w - w^2$$

On dérive cette expression pour obtenir la fonction de densité, qu'on note $f(x)$

$f(w) = 2(1-w)$ pour $w \in [0, 1]$ et nulle sinon

Puis on calcule l'espérance de cette fonction de densité :

$$E(W) = \int_0^1 f(w) dw$$

$$E(W) = 2(1-w)w \Big|_0^1$$

$$E(W) = 2 * \left(\frac{1}{2} - \frac{1}{3} \right) = \frac{1}{3}$$

En remultipliant par 100, on obtient que l'espérance de la valeur absolue de la différence entre deux valeurs indépendantes de loi uniforme entre 1 et 100, qui peuvent correspondre à des centiles, est de 33,3. **Ainsi, deux personnes prises aléatoirement auront en moyenne 33,3 centiles d'écart de revenu.**