

S1 – Données et traitements

1. Traitements et construction du panel

Le panel a été construit en plusieurs étapes à partir des données individuelles.

Identifiants fiscaux manquants. Chaque année, les identifiants fiscaux individuels (spis) de certains individus apparaissent en valeur manquante dans les données fiscales. La première étape de traitement des données consiste donc à retrouver ces identifiants, lorsque cela est possible, afin de maximiser la taille du panel. Les identifiants sont imputés à partir des données des années adjacentes et d'une clé d'appariement. La clé d'appariement repose sur des variables identifiantes renseignées chaque année : année de naissance, 8 premiers caractères du nom, code commune, situation familiale, sexe, etc. Après imputation, la proportion de spis manquants est inférieure à 1 %, en moyenne, sur l'ensemble de la période (excepté en début de période, en 2002 et 2003, où elle est inférieure à 3 %).

Doublons. Lorsqu'un individu apparaît plusieurs fois avec des âges distincts, seul l'âge maximal est conservé. En général, les âges en question sont très proches ou contiennent une erreur de dizaine.

Changements de situation. Si un individu a effectué plusieurs déclarations, ses revenus sont sommés. Par exemple, dans le cas d'un décès, le conjoint survivant doit effectuer deux déclarations en distinguant ses revenus jusqu'à la date de décès et ses revenus ultérieurs. Dans le cas d'une union pré-2011, trois déclarations étaient requises : les deux personnes devaient déclarer leurs revenus de manière distincte jusqu'à la date d'union puis ensemble à partir de celle-ci. En ce qui concerne la situation familiale (célibataire, marié/pacsé, veuf, divorcé/séparé), nous prenons en compte les modifications de situation en cours d'année.

Trous. Un individu peut disparaître des données une année pour plusieurs raisons : un décès, une sortie du territoire fiscal français, un traitement de déclaration dans une émission ultérieure à celle considérée dans cette étude¹, ou encore si une personne en perte d'autonomie devient personne à charge d'un autre foyer. Le problème n'est cependant pas aussi prégnant que dans les données DADS (Magnac & Roux 2009), dans lesquelles certaines absences sont structurelles (changements de statut d'activité) ou dues à l'absence de données certaines années.

Enfin, certaines informations étant manquantes en 2002, cette année n'a pas été utilisée dans l'analyse et les résultats, mais a servi pour imputer des identifiants manquants en 2003.

2. Construction des variables géographiques et de sexe

Variables géographiques. Les variables de localisation géographique ont été construites à partir des identifiants de communes disponibles dans les données fiscales. L'identifiant du département correspond au code départemental de l'adresse d'imposition et comprend 100 modalités (une pour chaque département).

Variable de sexe. Dans les fichiers POTE en panel, la variable de sexe n'est pas renseignée pour le déclarant 2. Pour les déclarants 2, nous imputons donc la valeur de cette variable en deux étapes. Premièrement, pour les individus qui sont déclarants 1 au moins une année, la valeur est connue et peut être complétée les autres années. Deuxièmement, les autres valeurs sont imputées en faisant l'hypothèse que tous les couples mariés ou

¹ Cette probabilité est plus forte avant 2006 car avant cette date les 4^e ou 5^e émissions étaient utilisées, et non la 6^e comme après 2006.

pacsés sont composés d'individus de sexes différents en 2003². Une fois la variable de sexe reconstituée pour tous les déclarants, les femmes représentent 51 % du panel.

S2 – Comparaisons internationales : compléments

Cette section apporte des détails et compléments sur les comparaisons internationales présentées dans l'article. Si le champ, la période ou le revenu considéré peuvent être différents de ceux de notre étude, de nombreuses comparaisons confortent l'idée d'une mobilité plus faible en France qu'aux États-Unis.

Corrélation rang-rang. Bradbury (2011) obtient une corrélation de 0,60 à 0,65 entre 1970 et 1995 aux États-Unis sur une période d'une dizaine d'années (à partir d'un revenu après redistribution au niveau du ménage, sur la population active). Sur 20 ans³, Kopczuk *et al.* (2010) obtiennent une corrélation rang-rang de 0,57 aux États-Unis sur la population des salariés travaillant dans le commerce et l'industrie. L'écart avec nos résultats peut être en partie lié à la différence de champ (notre champ est plus large), et également au fait que les auteurs considèrent une fenêtre plus longue (20 ans contre 18 ans), ce qui conduit mécaniquement à une baisse de la corrélation. Sur une fenêtre de 5 ans, ils obtiennent une corrélation rang-rang d'environ 0,9, contre 0,82 sur notre panel.

Plus récemment, Guvenen *et al.* (2022) calculent la corrélation rang-rang dans de très nombreux pays et sur une période plus récente. À partir de la période 1997-2007, ils calculent une mobilité entre t et $t+5$: ils obtiennent une corrélation rang-rang de 0,83 en France (contre 0,80 sur notre panel et sur la période 2003-2008 sur l'ensemble du champ et 0,81 sur un champ comparable, c'est-à-dire hors indépendants), plus élevée qu'aux États-Unis (0,75) et dans les pays nordiques (0,68 au Danemark, 0,70 en Norvège et 0,67 en Suède). Parmi les pays les plus développés, le Royaume-Uni, l'Italie et la France sont ceux où la mobilité telle que mesurée dans leur étude est la plus faible.

Matrices de transition. La mobilité mesurée dans notre étude est inférieure à celle documentée par Auten, Gee & Turner (2013a) aux États-Unis. Ces auteurs étudient la mobilité à partir des revenus avant redistribution, rapportés à des unités de consommation⁴. Un peu moins de la moitié des individus âgés de 35 à 40 ans et appartenant au plus haut cinquième restent dans ce cinquième 20 ans plus tard⁵, 40 % restent dans le dixième le plus élevé et 24 % dans le centième le plus élevé⁶. Sur nos données, en considérant une population restreinte à la même tranche d'âge (35-40 ans) et en considérant les mêmes revenus, nous obtenons des proportions respectives de 59 %, 52 % et 32 % sur une période de longueur comparable (18 ans). Auten *et al.* (2013a) obtiennent également une mobilité très ascendante supérieure (5 %, contre 3 % ici). Ces comparaisons suggèrent là encore que la mobilité est plus élevée aux États-Unis qu'en France.

La comparaison avec d'autres études américaines, portant sur des périodes ou des concepts de revenus différents, va encore dans le sens d'une mobilité plus faible en France. Hungerford (2011) trouve que la

² Cette hypothèse est peu restrictive, car peu de personnes de même sexe partagent leur déclaration en 2003. En effet, en 2003, le mariage entre personnes de même sexe n'existe pas encore. Et seuls 19 251 Pacs entre personnes de même sexe ont été célébrés en France hors Mayotte depuis sa création en 1999 (cf. <https://www.insee.fr/fr/statistiques/2381498>).

³ En faisant une moyenne sur 5 ans en début et fin de période.

⁴ Ces dernières sont calculées en prenant la racine carrée du nombre de personnes dans le ménage, ce qui donne des résultats très proches de la méthode usuelle à l'Insee.

⁵ Auten & Gee (2009) montrent également, à partir des mêmes types de revenus, que la moitié des individus des cinquièmes extrêmes y restent 10 ans plus tard, et que la mobilité est stable entre les périodes 1987-1996 et 1996-2005.

⁶ Ils analysent la persistance dans le top 1 % sur différentes fenêtres de temps et montrent que si 42 % des individus du top 1 % en 2000 y restent en 2004, seuls 26 % y appartiennent tous les ans entre 2000 et 2004.

proportion d'individus restant dans le plus haut cinquième (mesurée à partir d'un revenu après redistribution au niveau du ménage) est de 50 % à 54 % sur 10 ans (respectivement entre 1980 et 1989, et entre 1990 et 1999). Sur une période de 10 ans, nous obtenons une proportion plus élevée, de 70 % sur notre panel.

L'OCDE (2018) montre que la persistance dans le plus bas cinquième, mesurée sur quatre ans et à partir des niveaux de vie, est, en France, largement supérieure à la moyenne des pays de l'OCDE. La France se situe en revanche dans la moyenne des pays de l'OCDE en ce qui concerne la persistance dans le plus haut cinquième.

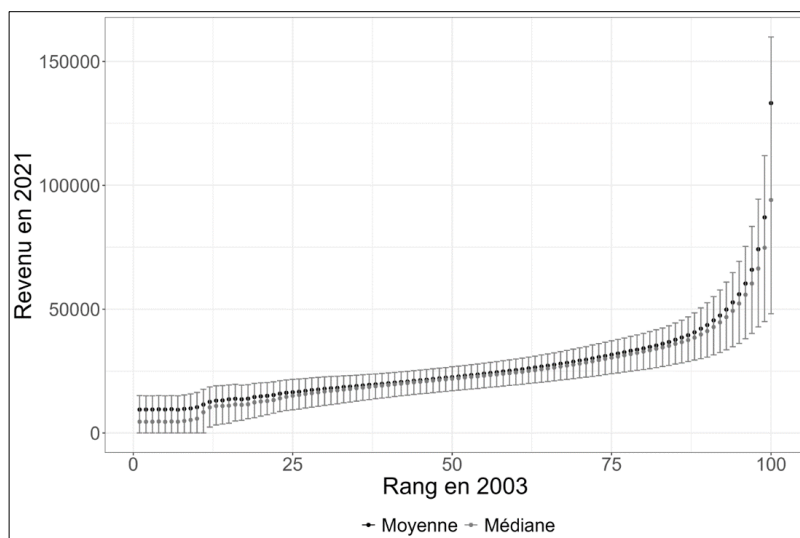
Indice de Shorrocks. Pour cet indicateur également, les comparaisons ne sont pas faciles à mener en raison de différences de champ. Aux États-Unis, Kopczuk *et al.* (2010) obtiennent un indice de mobilité de 4 % sur une période de 5 ans et sur une population restreinte aux salariés (qui exclut notamment les indépendants), soit un niveau un peu plus élevé que ce nous obtenons sur une période de 5 ans également (mais à une date plus récente)⁷. Jäntti & Jenkins (2015), en reprenant les données de Bradbury (2011), obtiennent, quant à eux, un indice de 13-14 % sur une période de 11 ans entre 1970 et 1995, supérieur à celui que nous obtenons sur une période de longueur identique de 11 ans (5 %), ce qui suggère encore que la mobilité serait supérieure aux États-Unis. Buchinsky & Hunt (1999) trouvent que les inégalités de salaires et revenus individuels sont réduites de 12 à 26 % par la mobilité. Ils obtiennent ainsi une mobilité plus élevée que ce que l'on obtient en France, même si cette comparaison doit être prise avec précaution car les auteurs se focalisent sur une population plus jeune (a priori plus mobile) mais aussi plus stable (hors indépendants et salariés), et une période plus courte (ce qui va dans le sens d'une moindre mobilité). Enfin, l'OCDE (2018) ne calcule pas d'indice de Shorrocks mais produit des indicateurs d'inégalités issus de revenus pris en compte sur 4 ans ou 9 ans, et montre que l'indice de Gini diminue davantage en France en prenant en compte les revenus permanents, et que le classement des pays ne change globalement pas lorsque les revenus sont évalués sur quatre ans.

⁷ Nous obtenons un coefficient de 3,4% sur une période de 5 ans (entre t et $t+4$) sur l'ensemble de notre champ, et 3,3% en excluant les revenus des indépendants pour être sur le même champ que Kopczuk *et al.* (2010).

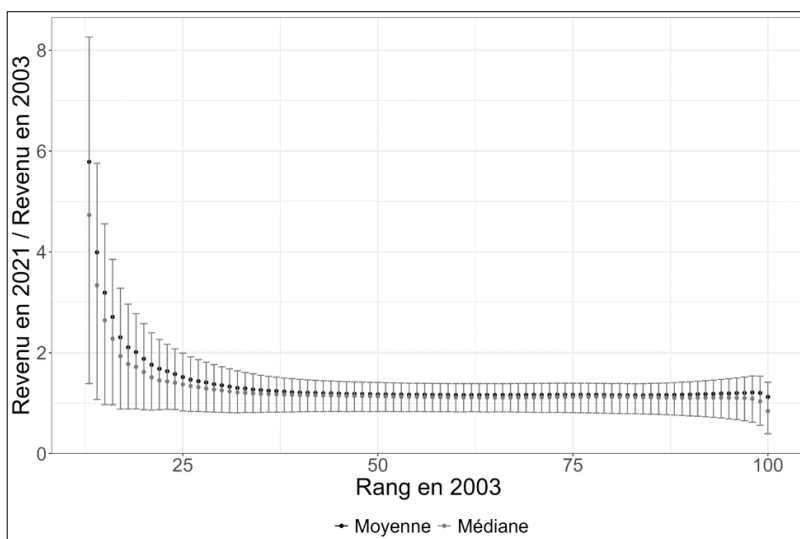
S3 – Figures complémentaires

Figure S3-I – Revenus des individus en 2003 et 2021 en fonction du rang en 2003

A) Quantiles en 2021



B) Quantiles du ratio des revenus entre le début et la fin de la période



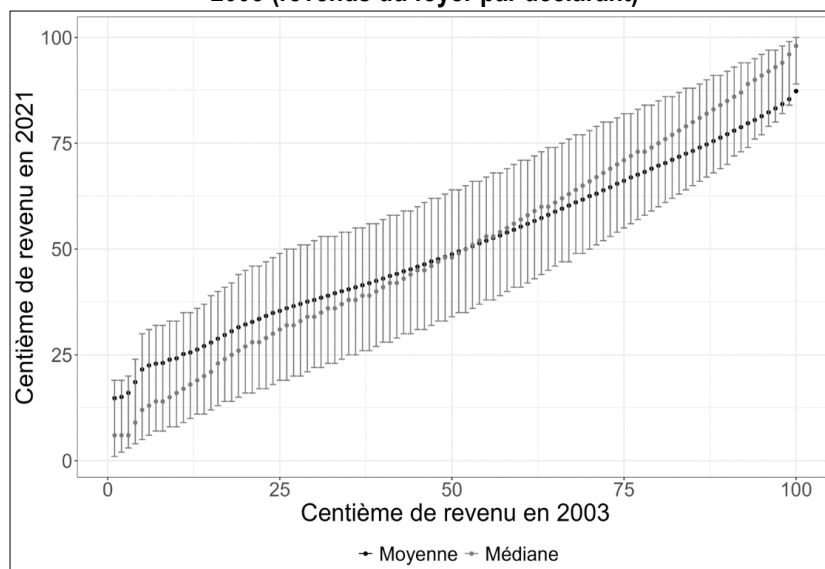
Note : les intervalles représentent l'écart entre le 1^{er} et le 3^e quantiles. Pour plus de lisibilité, sur la figure B, les 12 premiers centièmes ne sont pas représentés ; euros constants.

Lecture : Figure A : Les individus du 25^e centième de revenus en 2003 ont, en 2021, un revenu moyen de 16 500 euros, un revenu médian de 15 000 euros. Le premier quartile de leur revenu en 2021 est de 9 300 euros et le troisième quartile de 21 600 euros ; Figure B : les individus du 25^e centième de revenus en 2003 ont un ratio moyen de revenu entre le début et la fin de la période de 1,5 et un ratio médian de 1,4. Le premier quartile de leur ratio est de 0,8 et le troisième quartile de 2,0.

Champ : France hors Mayotte, individus présents dans les données POTE en 2003 et 2021 ayant entre 25 et 49 ans en 2003 (hors premier dixième de revenus en 2003).

Sources : Insee-DGFiP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-II – Moyenne et quartiles (Q1, médiane, Q3) des rangs des individus en 2021 en fonction du rang en 2003 (revenus du foyer par déclarant)



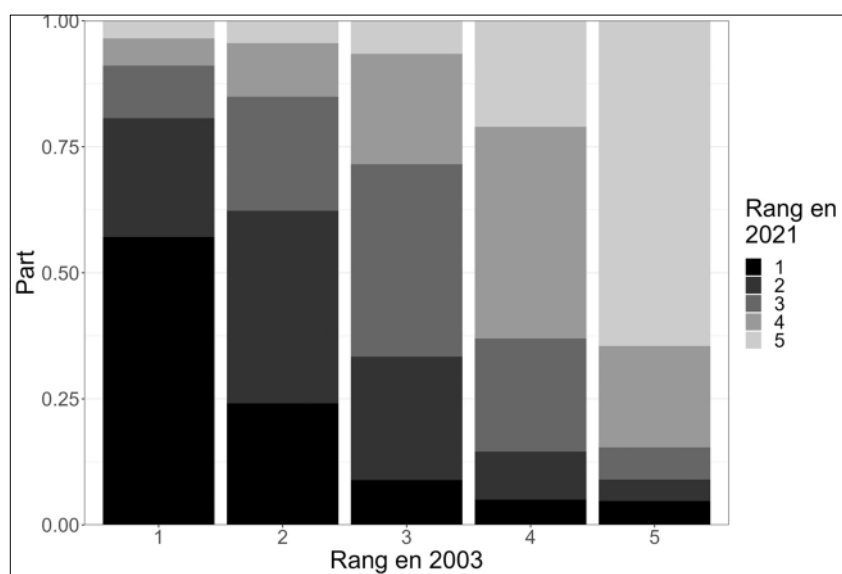
Note : les intervalles représentent l'écart entre le 1^{er} et le 3^e quartiles.

Lecture : Les individus du 25^e centième de revenus en 2003 ont, en 2021, un rang moyen de 35 et un rang médian de 31.

Champ : France hors Mayotte, individus présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003.

Sources : Insee-DGFiP, POTE panéliste 2003-2021, calculs des auteurs.

Figure S3-III – Transitions entre cinquièmes de revenus entre 2003 et 2021

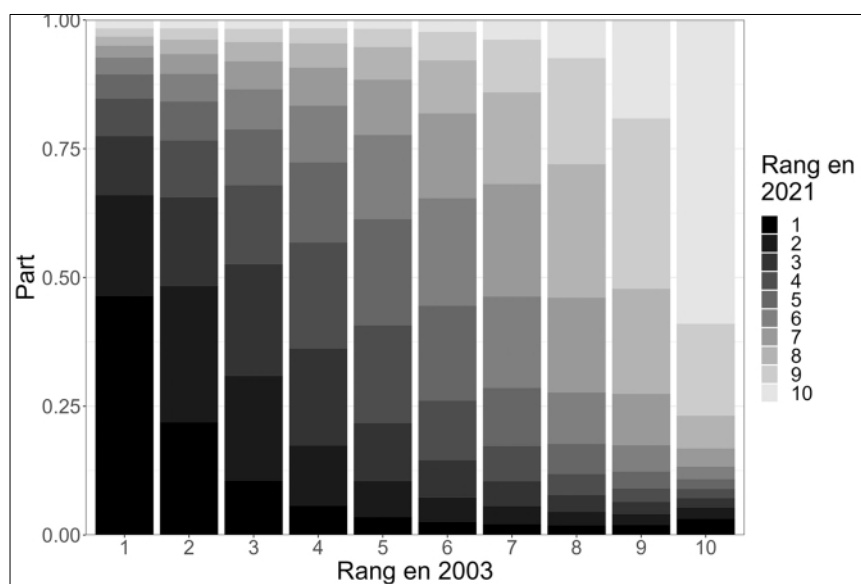


Lecture : 4 % des individus situés dans le premier cinquième de revenus en 2003 sont dans le dernier cinquième en 2021.

Champ : France hors Mayotte, individus présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003.

Sources : Insee-DGFiP, POTE panéliste 2003-2021, calculs des auteurs.

Figure S3-IV – Transitions entre dixièmes de revenus entre 2003 et 2021

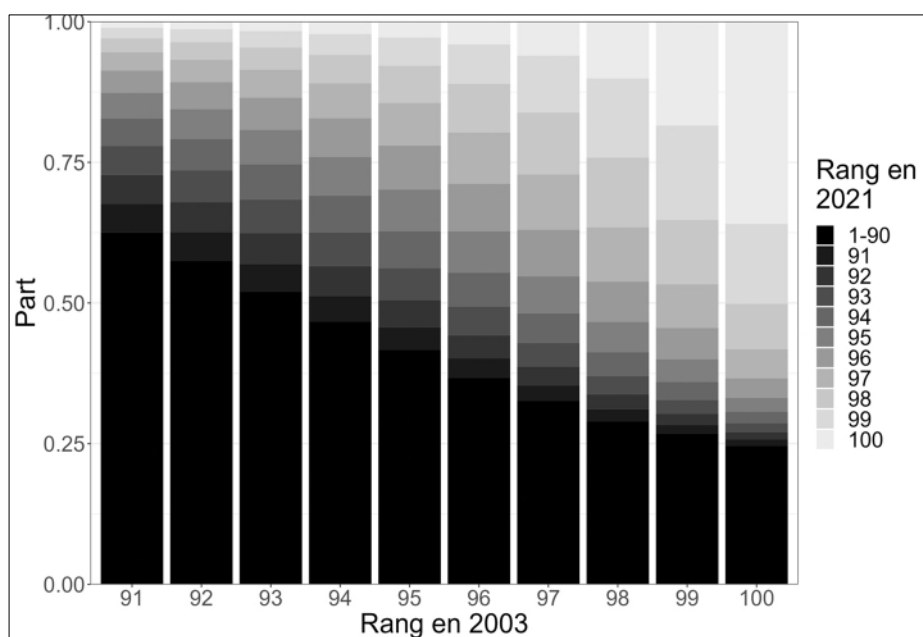


Lecture : parmi les individus dans le plus bas dixième en 2003, 46 % y restent en 2021.

Champ : France hors Mayotte, individus présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003.

Sources : Insee-DGFiP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-V – Transitions entre centièmes de revenus entre 2003 et 2021 dans le haut de la distribution

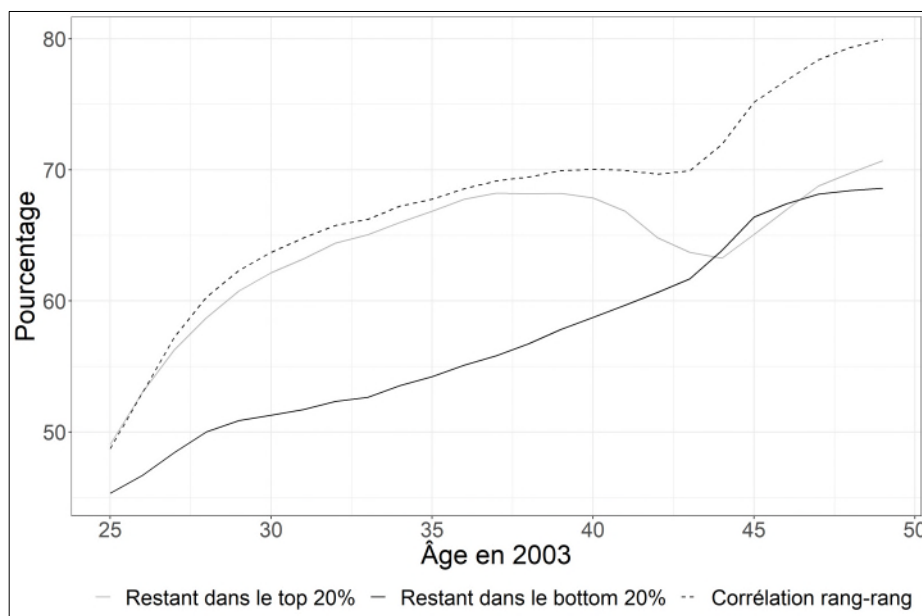


Lecture : Les individus parmi les 1 % les plus aisés en 2003 ont 36 % de chances d'en faire encore partie en 2021.

Champ : France hors Mayotte, individus présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003.

Sources : Insee-DGFiP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-VI – Indicateurs d'inertie selon l'âge en 2003



Note : Les individus sont classés selon les quantiles intragénérationnels (au sein de leur cohorte).

Lecture : Parmi les individus ayant 30 ans et comptant parmi les 20% les plus aisés en 2003, 62 % sont encore parmi les 20 % les plus aisés en 2021.

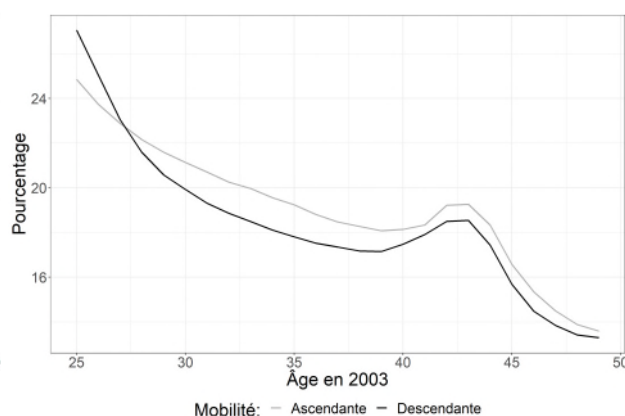
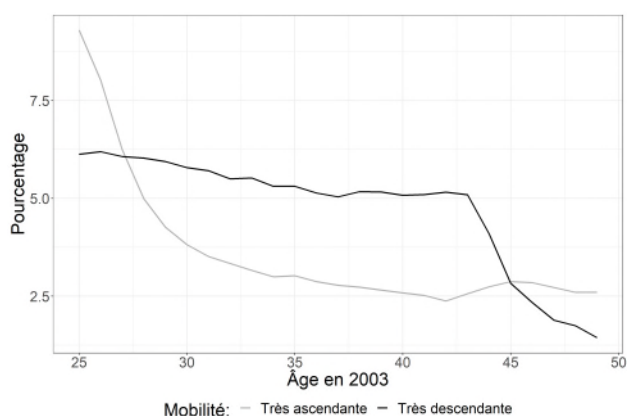
Champ : France hors Mayotte, individus parmi les 20% les plus aisés en 2003 et présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003.

Sources : Insee-DGFIP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-VII – Indicateurs de mobilité selon l'âge en 2003

A. Très ascendante/descendante

B. Ascendante/Descendante



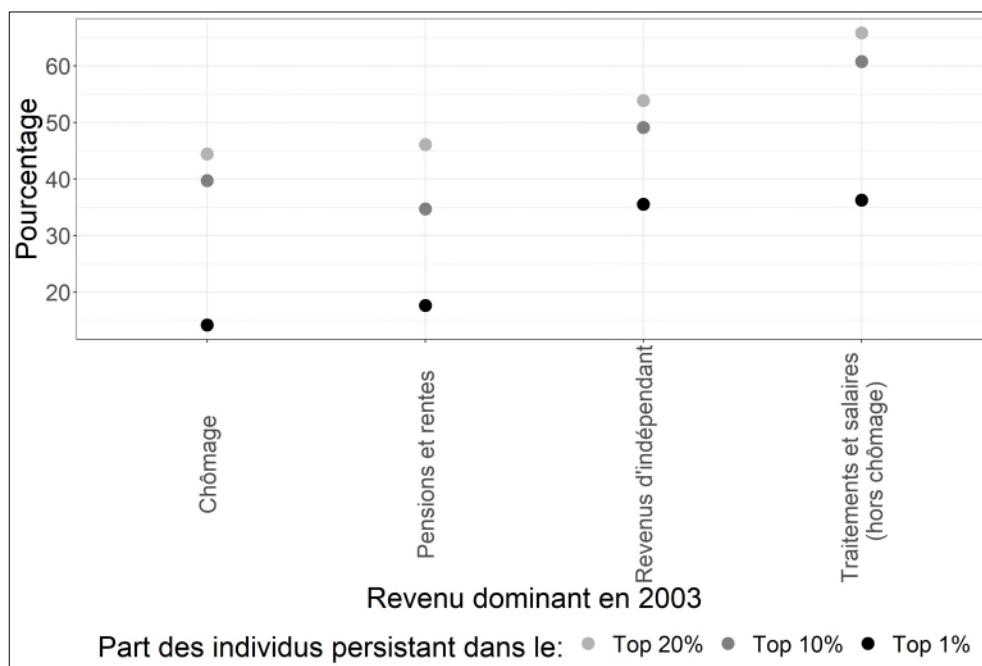
Note : Les individus sont classés selon les quantiles intragénérationnels (au sein de leur cohorte).

Lecture : Figure A : Parmi les individus ayant 30 ans en 2003 et comptant parmi les 20 % les plus aisés en 2003, 5,8 % ont connu une mobilité très ascendante entre 2003 et 2021 ; Figure B. : Parmi les individus ayant 30 ans en 2003, 21 % connaissent une mobilité ascendante entre 2003 et 2021.

Champ : France hors Mayotte, individus présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003.

Sources : Insee-DGFIP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-VIII – Indicateurs d'inertie selon la composition du revenu en 2003

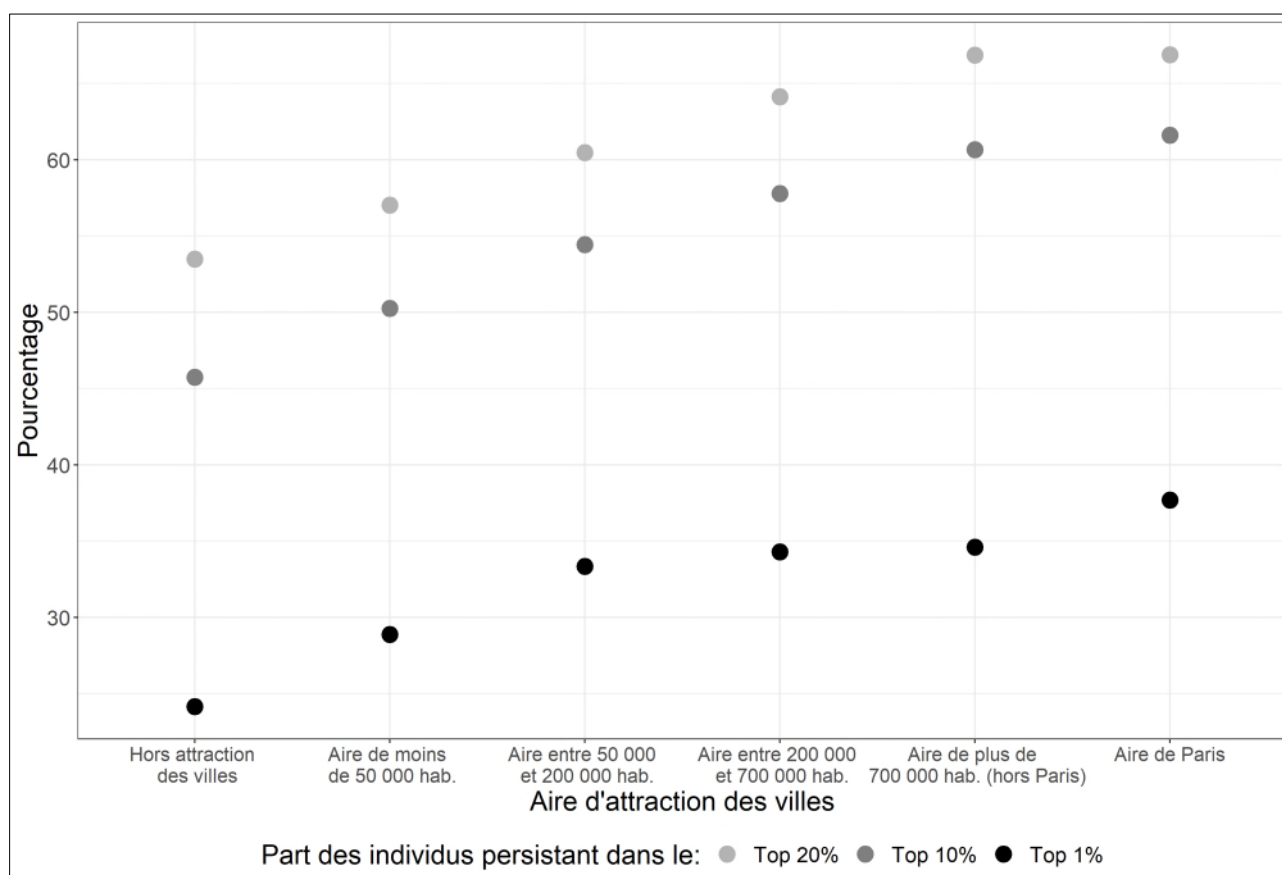


Lecture : Parmi les individus ayant pour revenu principal « Revenus d'indépendant » en 2003, et comptant parmi les 20 % les plus aisés en 2003, 54 % sont encore parmi les 20 % les plus aisés en 2021.

Note : Une personne dont le revenu est nul n'est classée dans aucun des statuts d'activité.

Champ : France hors Mayotte, individus parmi les 20 % les plus aisés en 2003 et présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003. Sources : Insee-DGFIP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-IX – Indicateurs d'inertie selon la taille d'aire d'attraction des villes (à partir de la commune en 2003 et du zonage de 2020)

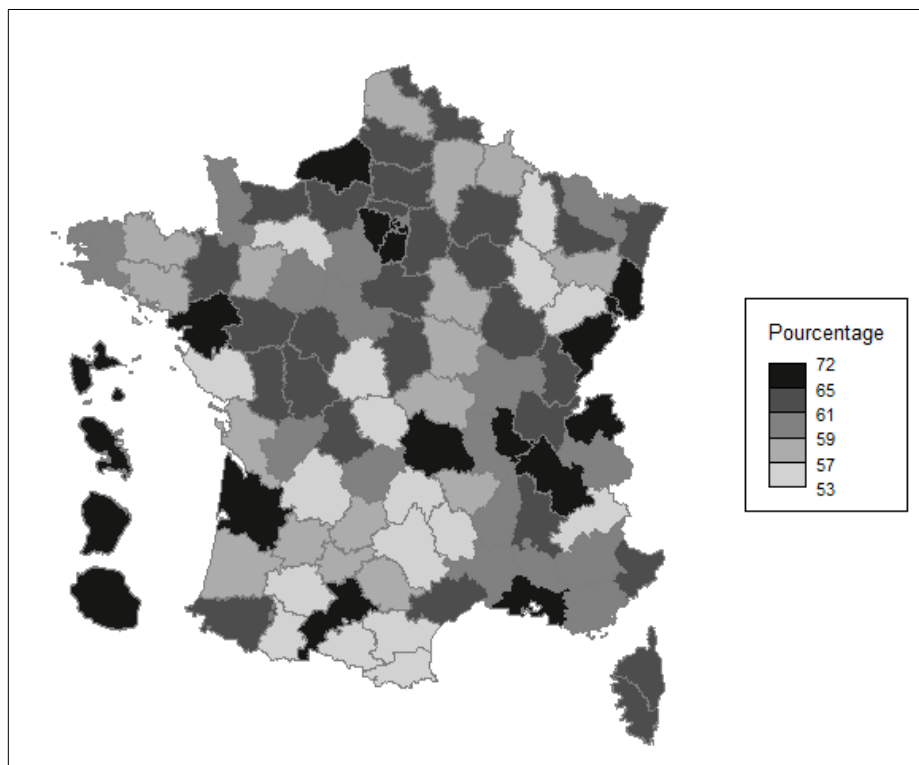


Note : l'aire d'attraction des villes (zonage et taille de 2020) est déterminée à partir du lieu de vie pris en 2003.

Lecture : Parmi les individus hors attraction des villes, et comptant parmi les 20 % les plus aisés en 2003, 53 % sont encore parmi les 20 % les plus aisés en 2021.

Champ : France hors Mayotte, individus parmi les 20 % les plus aisés en 2003 et présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003. Sources : Insee-DGFIP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-X – Proportion d'individus restant parmi les 20 % les plus aisés par département

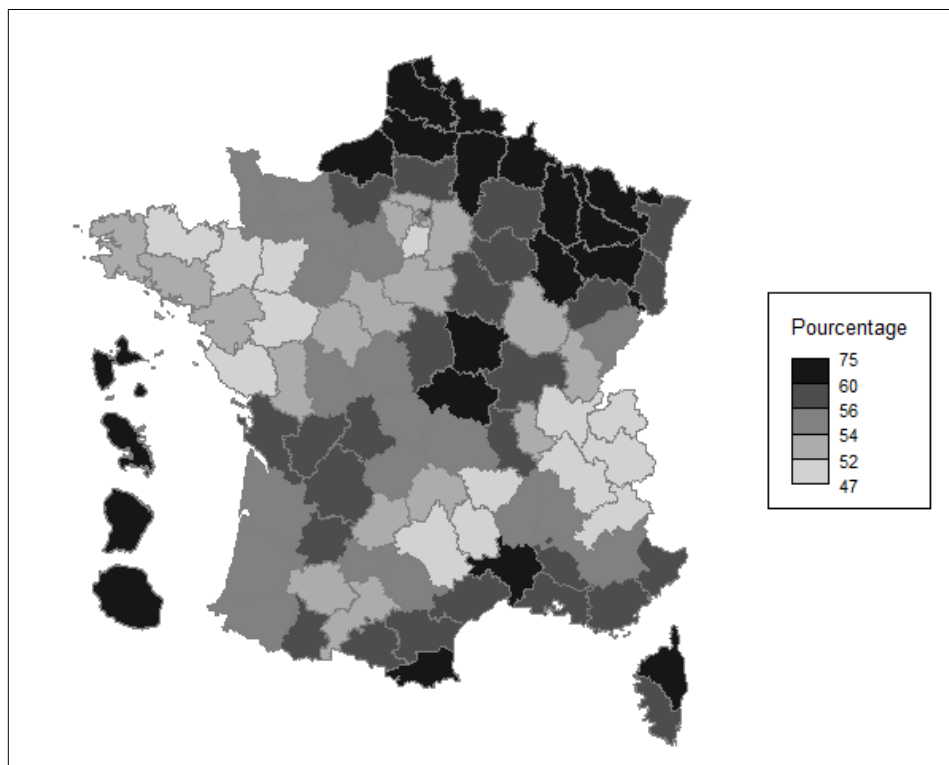


Lecture : Parmi les 20 % les plus aisés en 2003, à Paris, la proportion d'individus comptant parmi les 20 % les plus aisés en 2021 est comprise entre 65 % et 72 %.

Champ : France hors Mayotte, individus parmi les 20 % les plus aisés en 2003, présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003.

Sources : Insee-DGFiP, POTE panéalisé 2003-2021, calculs des auteurs.

Figure S3-XI – Proportion d'individus restant parmi les 20% les plus modestes par département

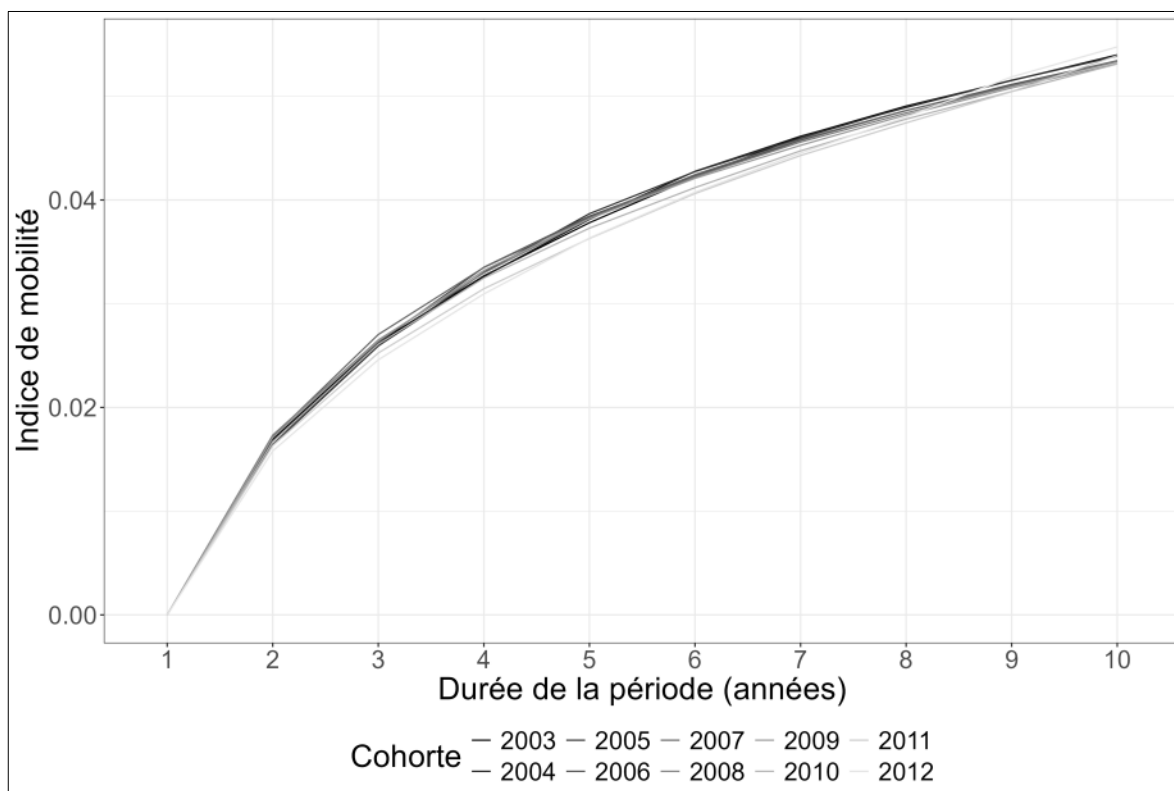


Lecture : Parmi les 20 % les moins aisés en 2003, à Paris, la proportion d'individus comptant parmi les 20 % les moins aisés en 2021 est comprise entre 54 % et 56 %.

Champ : France hors Mayotte, individus parmi les 20 % les moins aisés en 2003, présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003.

Sources : Insee-DGFiP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-XII – Indice de mobilité de Shorrocks sur dix ans par cohorte



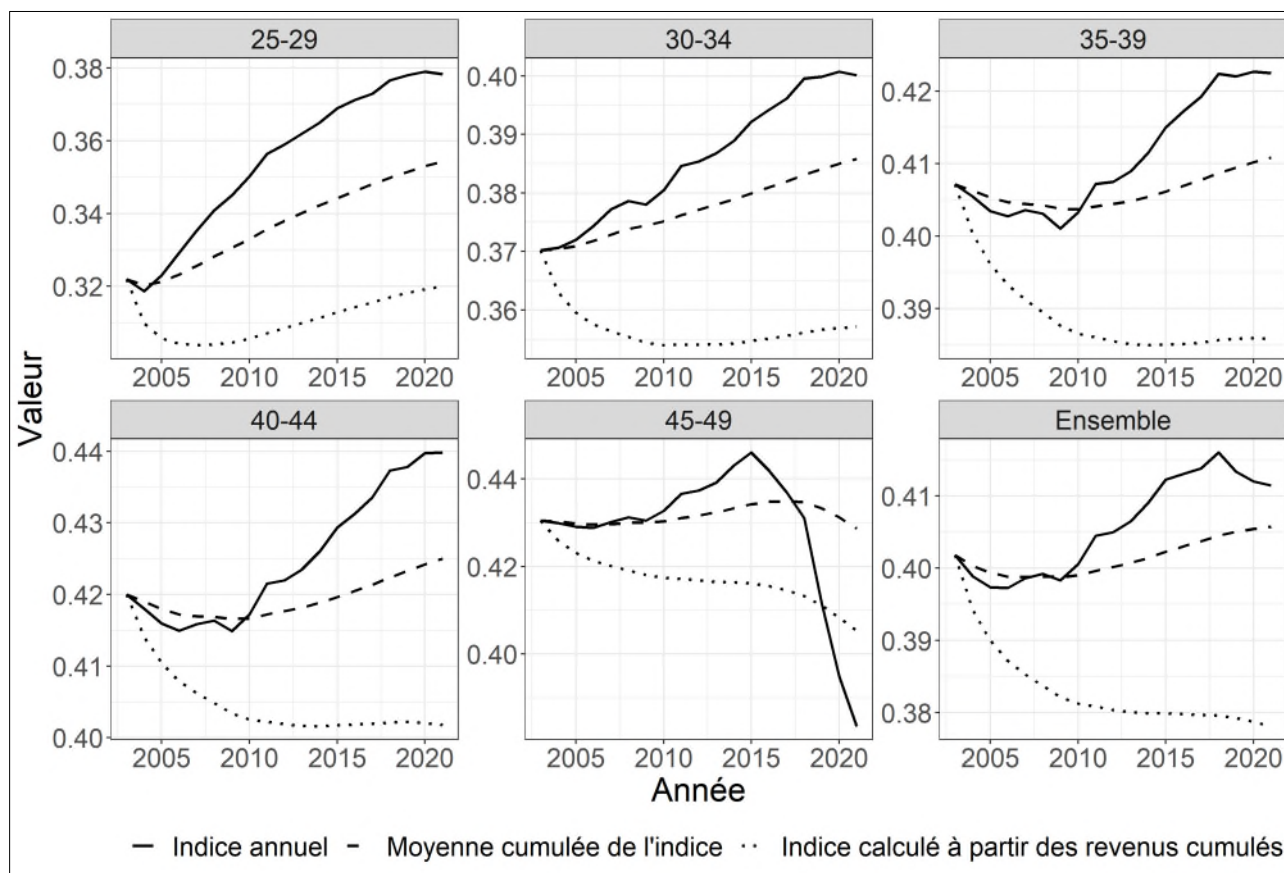
Note : La cohorte n est composée des individus âgés de 35 ans l'année n et présents dans les données POTE de l'année n à l'année $n+9$.

Lecture : Les individus âgés de 35 ans en 2003 ont un indice de mobilité de 5,4 % sur la période 2003-2012.

Champ : individus âgés de 35 ans une année n (entre 2003 et 2011) et présents dans les données POTE de l'année n à l'année $n+9$.

Sources : Insee-DGFiP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-XIII – Indice de Gini selon la tranche d'âge



Note : les courbes correspondent aux éléments de l'équation (1) (partie méthode).

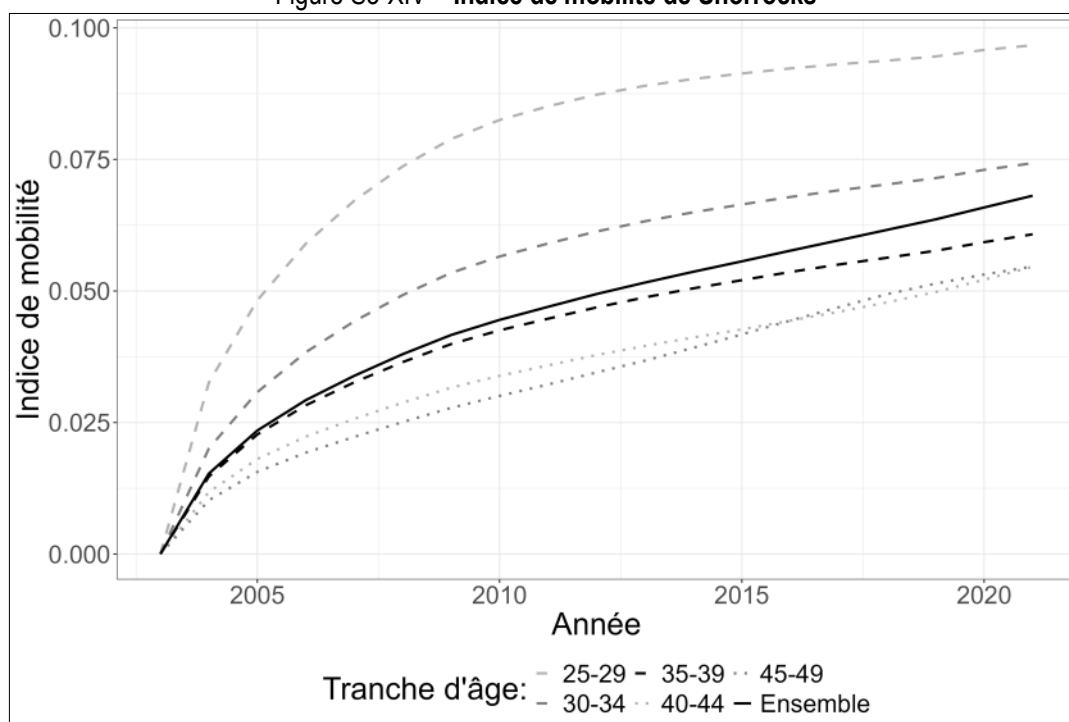
Lecture : En 2021, parmi les 25-29 ans, l'indice de Gini est de 0,378. En moyenne sur la période 2003-2021, il est de 0,354.

En revenu permanent, il est de 0,320.

Champ : France hors Mayotte, individus présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003.

Sources : Insee-DGFIP, POTE panéliste 2003-2021, calculs des auteurs.

Figure S3-XIV – Indice de mobilité de Shorrocks

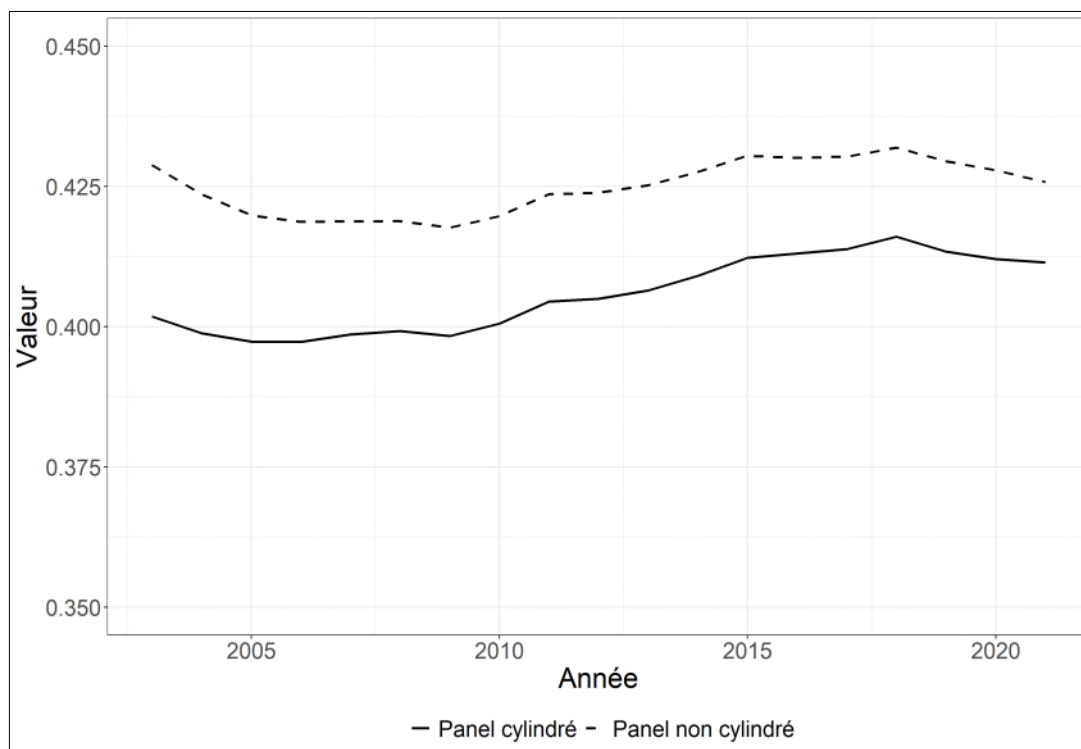


Note : l'indice de mobilité est défini à partir de l'équation (2).

Champ : France hors Mayotte, individus présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003

Sources : Insee-DGFiP, POTE panélisté 2003-2021, calculs des auteurs.

Figure S3-XV – Indice de Gini calculé en coupe sur le panel non cylindré



Note : Le panel cylindré fait référence au panel mobilisé dans cette étude uniquement pour le calcul de l'indice de Shorrocks. Le panel non cylindré fait référence au panel mobilisé dans tout le reste de l'article.

Champ : France hors Mayotte, individus présents dans les données POTE en 2003 et 2021 et âgés de 25 à 49 ans en 2003

Sources : Insee-DGFiP, POTE panélisté 2003-2021, calculs des auteurs.