

Méthodologie et apports de l'analyse de fratrie

Béatrice Boutchenik, Clotilde Coron, Sébastien Grobon,¹
Céline Goffette² & Louis-André Vallet³

¹ ENSAE

² CEREQ, CREST/LSQ

³ CNRS, CREST/LSQ

Résumé. Nous souhaitons montrer la pertinence des dispositifs d'enquête dans lesquels certaines variables sont renseignées pour plusieurs membres d'une même fratrie. Les données de cette forme sont rarement produites par la statistique publique française malgré leur intérêt pour l'étude de l'importance de l'origine sociale, ou plus largement de l'influence familiale. Notre étude utilise les enquêtes « Formation et qualification professionnelle », 2003 (INSEE) et « Niveau intellectuel des enfants d'âge scolaire », 1965 (INED). La structure de ces données nous permet d'y appliquer les modèles de fratrie issus de la sociologie quantitative américaine de la mobilité sociale (Sewell *et al.*, 2004). Ces modèles, que ce soient des modèles multiniveaux ou des modèles Multiples Indicateurs Multiples Causes (MIMIC), permettent l'estimation de ce que nous appelons l'influence familiale cohésive, qui prend en compte les facteurs observés mais également inobservés de l'influence familiale, à travers la mesure de ce qui est commun aux parcours des différents membres de la fratrie - en termes de niveau d'études atteint par ceux-ci par exemple. Nous articulons ces modèles avec une classification des fratries prenant également en compte la ressemblance entre parents et enfants en termes de parcours, pour établir un lien entre l'influence familiale cohésive et la mobilité sociale.

L'influence familiale cohésive telle que nous la mesurons à travers les modèles de fratrie explique 55% de la variabilité des niveaux d'études, et 33% lorsque nous utilisons une mesure du statut professionnel atteint. Dans les deux cas, on montre que l'utilisation des caractéristiques familiales observées pour l'enquête seulement (pourtant très complètes) conduit à sous-estimer de moitié l'influence familiale : les caractères inobservables entrent donc fortement en jeu dans la socialisation et contribuent à rendre similaires les parcours des membres d'une même fratrie. L'utilisation de données de fratrie nous permet ainsi de montrer la force de la socialisation familiale, qu'une analyse par classification permet de lier à la mobilité sociale et donc à la question de l'inégalité des chances. Nous nous concentrons ici sur la présentation des méthodes d'analyse de fratrie, en indiquant leurs apports et certaines des questions théoriques qui y sont liées.

Plan de l'article

Introduction.....	3
L'intérêt des données de fratrie.....	3
Enquête FQP 2003.....	4
Enquête sur le niveau intellectuel des enfants d'âge scolaire.....	5
Transformation des variables catégorielles en échelles numériques.....	5
Échelle d'éducation et de profession dans l'enquête FQP.....	5
Score de compétences scolaires dans l'enquête INED.....	6
Méthodologie de l'analyse de fratrie.....	7
Modèles multiniveaux.....	7
Modèles MIMIC.....	7
Classification.....	8
Principaux résultats.....	8
Estimation de l'influence familiale cohésive sur le niveau d'éducation.....	9
L'influence familiale sur la profession.....	10
L'influence familiale sur les compétences dans le jeune âge.....	10
L'effet propre de l'éducation.....	11
Prendre en compte le changement du contexte scolaire et mesurer l'évolution de l'influence familiale cohésive au cours du temps..	12
Influence familiale cohésive et mobilité sociale.....	13
Questions soulevées par notre étude.....	14
Sur le plan théorique : quels liens peut-on établir entre notre modélisation et les différentes formes de socialisation familiale ?	14
Sur le plan technique : les difficultés d'une prise en compte du caractère discret des variables d'intérêt.....	15
Conclusion.....	16
Tableau de conclusion :.....	17
Prolongements méthodologiques.....	16
Questions sociologiques.....	16
Annexes.....	17
Bibliographie.....	21

Introduction

L'analyse de fratrie a été créée pour répondre au problème de la sous-estimation de l'influence familiale dans la sociologie américaine de la mobilité sociale. Elle s'appuie sur l'emploi de données renseignées pour plusieurs membres d'une même fratrie. Des travaux américains mais également français aident à compléter cette approche à travers l'analyse des différentes manières dont s'exercent les influences des parents vers leurs enfants, ou au sein de la fratrie. Nous présentons ici l'intérêt d'une articulation entre l'utilisation des modèles de fratrie, principalement les modèles MIMIC, et celle de techniques de classification, pour enrichir l'analyse de la socialisation.

Ce travail nous permet de poser plusieurs questions théoriques et méthodologiques liées à l'estimation d'une influence « totale » de la famille sur les parcours scolaires et professionnels, différente de celle que l'on observe dans les modèles de régression traditionnels, et qui évite la sous-estimation de l'influence du milieu d'origine.

Nous tentons de montrer l'intérêt des enquêtes qui renseignent des informations sur plusieurs membres d'une même fratrie dans les enquêtes de la statistique publique, et indiquons quelles sont les données de cette forme qui existent en France. Cela nous amène à la question de la formalisation des parcours des individus. Pour pouvoir tirer parti des données de fratrie, nous devons transformer les variables de diplôme et de profession en échelles numériques dont nous détaillons la construction. Nous présentons les modèles multiniveaux et MIMIC, que nous utilisons pour mesurer l'influence familiale cohésive à l'aide de ces échelles, et nous montrons pourquoi il est sociologiquement fécond d'articuler cette mesure avec une méthode de classification. Nous évoquons enfin les difficultés d'interprétation liées à la complexité de l'étude de l'influence familiale, en même temps qu'à des limites méthodologiques de notre analyse.

L'intérêt des données de fratrie

Comme le montrent les analyses sociologiques, la transmission familiale se fait par une imprégnation continue de caractères « secondaires » (Lahire, 1995) tels que le langage, la façon de se tenir, le rapport au temps et au scolaire. L'influence du milieu social d'origine met également en jeu des relations sociales et de multiples autres caractéristiques qui ne peuvent être intégrées de manière exhaustive dans les enquêtes statistiques cherchant à décrire l'origine sociale. Le caractère nécessairement partiel des attributs renseignés conduit donc à sous-estimer l'influence de la famille.

Dès 1972, Duncan, Featherman et Duncan, dans leur étude de la corrélation entre origine sociale, niveau scolaire et capacité à obtenir un emploi prestigieux, notent l'existence d'une corrélation substantielle entre les niveaux scolaires des frères et des sœurs. Ces travaux parmi d'autres (notamment Jencks, 1972) conduisent à l'introduction de variables relatives aux frères et sœurs dans la *Wisconsin Longitudinal Study*, qui suit sur le long terme une cohorte de 10 000 personnes ayant terminé leur éducation secondaire en 1957 : l'âge, le genre, le niveau d'études final, puis les résultats obtenus à des tests d'intelligence. Hauser, Sewell et Springer, qui sont les principaux acteurs du développement des études de ressemblance au sein de la fratrie, estiment en effet que l'emploi de ces données de fratrie rend possible une meilleure analyse de l'influence familiale, et donc des inégalités sociales :

« The resemblance of siblings raised together is, of course, a fundamental indicator of the force with which the family functions to create and maintain systems of social differentiation and inequality. Sibling resemblance captures the effects of social and economic background, of family structure, and of other commonalities of social and

psychological functioning of the family. »

(Sewell, Hauser, Springer et Hauser, 2004, p. 46)

Le fait que les différents membres d'une même fratrie aient connu une socialisation similaire au sein de la famille permet de saisir ce qui leur a été transmis en utilisant les ressemblances existant entre leurs parcours. Dans l'ensemble de la société, des parcours de frères et sœurs corrélés sont le signe d'une importante influence de la famille plutôt que d'une différenciation individuelle. On mesure ainsi l'influence « totale » de la famille à travers la ressemblance des frères et sœurs pour des caractéristiques données, par exemple l'éducation ou la profession. L'influence de la famille ainsi mesurée n'inclut pas seulement les caractéristiques familiales mesurables, mais également toutes les caractéristiques inobservables que les frères et sœurs peuvent partager. Parmi ces caractéristiques, on trouve tout d'abord des facteurs propres aux parents : style d'éducation et capital culturel informel, caractères génétiques... On mesure aussi dans ce facteur familial l'influence réciproque entre frères et sœurs. Enfin, il mesure l'influence de l'environnement commun aux frères et sœurs mais dépassant l'action socialisatrice de la famille : quartier, qualité de l'école fréquentée, influence des pairs. On peut toutefois parler d'influence familiale dans la mesure où le choix du quartier, par exemple, relève de déterminants familiaux. Les analyses de fratrie permettent donc de mesurer de manière beaucoup plus complète l'influence familiale.

Cette tradition est reprise par des sociologues néerlandais tels que De Graaf (1986) et Sieben (2001), qui systématisent l'utilisation de modèles de fratrie en Europe pour mesurer l'influence « totale » de la famille et comparer l'importance du milieu d'origine dans différents pays, dont la France ne fait pas partie. L'absence de telles approches en France, auxquelles sont souvent préférées les études portant sur données individuelles, tient à l'absence de questions sur les parcours de plusieurs membres de la même famille dans la plupart des enquêtes de la statistique publique française. Nous mettons à profit pour notre étude des données provenant d'une part de l'enquête « Formation et Qualification Professionnelle », dont nous faisons une utilisation originale en exploitant les informations portant sur un membre de la fratrie, collectées pour la première fois en 2003 ; et d'autre part de l'enquête « Niveau intellectuel des enfants d'âge scolaire » (1965) de l'INED, qui permet également d'identifier les membres d'une même fratrie. Leur structure spécifique nous permet de mettre en œuvre deux types de modèles statistiques de ressemblance dans la fratrie, les modèles multiniveaux d'une part qui sont des modèles hiérarchiques, et les modèles MIMIC (Multiples indicateurs multiples causes) d'autre part, qui font plus largement partie des modèles à équations structurelles.

Enquête FQP 2003

L'enquête FQP est une enquête qui est réalisée régulièrement par l'INSEE depuis 1964, pour reconstituer les parcours scolaires et professionnels des individus interrogés et documenter leur origine sociale. L'enquête de 2003 intègre pour la première fois des données sur un frère ou une sœur tiré(e) au hasard dans la fratrie pour lequel ou laquelle l'enquêté a décrit le diplôme maximal et la profession. Elle comprend également des informations sur les parents et les grands-parents de l'individu. Le champ de l'enquête FQP 2003 est constitué de 39 312 personnes ayant entre 18 et 65 ans au moment de l'enquête (c'est-à-dire étant nées entre 1938 et 1985 inclus pour l'enquête de 2003), tirées au sort dans l'échantillon de logements de l'INSEE. Nous avons restreint l'échantillon aux individus de plus de 25 ans ayant des parents résidant en France au moment de la fin de leurs études, de façon à ne garder que des individus ayant fini leurs études et qui ont été dans le système scolaire français. Nous obtenons un échantillon de 29 954 individus. Nous nous intéressons exclusivement aux individus qui ont un frère ou une sœur, puisque les enfants uniques ne peuvent pas être inclus dans l'estimation de l'influence familiale cohésive.

Enquête sur le niveau intellectuel des enfants d'âge scolaire

L'enquête sur le niveau intellectuel des enfants d'âge scolaire a été réalisée en 1965 auprès de 120 000 élèves âgés de 6 à 14 ans. L'objectif de cette enquête est de « dresser un tableau général du niveau intellectuel des jeunes Français », mais aussi d'identifier et de dénombrer les élèves ayant des difficultés à s'adapter à l'enseignement scolaire, et enfin de mettre au jour les facteurs individuels, familiaux et sociodémographiques propres à influencer le niveau intellectuel et la réussite scolaire des élèves.

L'échantillon que nous utilisons nous permet d'identifier 6 608 familles de deux individus ou plus, dont 358 couples de jumeaux. On n'a pu prendre en compte que les fratries au sein d'un même établissement à partir du rang alphabétique du patronyme (la variable de nom n'a pas été conservée pour protéger l'anonymat des répondants). Dans le système d'enseignement de 1965, la mixité dans l'enseignement n'était pas encore généralisée, ce qui implique que nous ne pouvons pas identifier un nombre important de fratries mixtes. Or l'identité de genre a vraisemblablement une influence importante sur la proximité des parcours scolaires, aussi nous serons sûrement amenés à surestimer l'influence de la famille. Plus largement, les évolutions du système scolaire depuis les années 1960 imposent de considérer les résultats avec plus de précautions. Nous présentons plus en détail dans la section suivante la façon dont nous interprétons les scores de QI de cette enquête.

Transformation des variables catégorielles en échelles numériques

Échelle d'éducation et de profession dans l'enquête FQP

Dans l'enquête FQP, deux aspects de la proximité des parcours de membres d'une même fratrie nous intéressent : proximité en termes d'éducation, et proximité en termes de profession. Dans le cadre des modèles que nous mettons en œuvre, il est important de disposer de variables numériques pour ces deux dimensions.

Concernant l'éducation, le diplôme est converti en nombre d'années d'études. Idéalement, il serait souhaitable de pouvoir associer à chaque diplôme le nombre « normal » d'années nécessaires à son obtention -pour un élève n'ayant jamais redoublé ni sauté une classe. C'est la stratégie utilisée en particulier par Ganzeboom et Treiman (1994). Cela est toutefois impossible dans notre cas, d'une part car certaines catégories de diplôme regroupent plusieurs nombres d'années requises différents, d'autre part parce que ces nombres « normaux » ont pu varier pendant la longue période que couvre l'enquête. On a donc associé, à chaque catégorie détaillée de diplôme maximal obtenu, le mode du nombre d'années d'études effectivement suivies par les individus possédant ce diplôme, c'est-à-dire la différence entre l'âge à la fin des études, et l'âge de six ans. Le mode peut être peu représentatif pour les diplômes plutôt rares : on a donc construit une seconde échelle s'appuyant cette fois sur la médiane du nombre d'années d'études pour les individus possédant au maximum un diplôme donné. On les appellera, par commodité, échelle ou variable modale et échelle ou variable médiane.

Le premier inconvénient de notre échelle est qu'elle tend à réduire l'écart entre les niveaux d'éducation les plus élevés et les plus bas. En effet, ce sont probablement les mêmes élèves qui à la fois arrêtent relativement tôt leur scolarité et prennent du retard dans celle-ci, alors que d'autres seront à la fois en avance ou « à l'heure » et obtiendront les diplômes les plus élevés. Ainsi, on associe aux diplômes les plus bas des nombres d'années d'études effectifs artificiellement élevés, qui ne reflètent pas nécessairement un réel avantage éducatif. Cela explique en partie le fait que le nombre d'années d'études soit plus élevé pour la catégorie « Aucun diplôme » que pour la catégorie CEP, ceci étant également dû au fait que le CEP existait à une époque où le nombre d'années d'études était de manière générale plus faible. Le second et principal inconvénient tient au fait que l'on peut attribuer une même grandeur à des diplômes pourtant très différents, par exemple dans le

cas du baccalauréat général et du baccalauréat technologique ou professionnel (14 années d'études à chaque fois). On assimile ainsi des individus appartenant à des filières dont les débouchés, le contenu et le prestige sont inégaux. Cependant, ces échelles permettent d'utiliser une modélisation originale et peu utilisée dans l'étude de la mobilité sociale et de l'influence familiale : les modèles multiniveaux et surtout les modèles MIMIC, qui ne peuvent pas être utilisés sur des variables modales.

Pour la profession, nous reprenons l'échelle élaborée par Chambaz, Maurin et Torelli (1998) associant aux différentes PCS des scores de statut social, obtenus en s'appuyant sur les classements subjectifs de 122 professions effectués par 3600 individus environ. Cette échelle existe sous deux formes chez les auteurs. L'échelle non-linéaire part de l'hypothèse que les scores subjectifs attribués aux professions sont de simples indications qualitatives permettant seulement de restituer l'ordre des professions en termes de statut. Dans le cas de l'échelle linéaire, ces scores sont considérés comme des quantités reflétant la valeur sociale associée à chaque profession. D'après les auteurs, le modèle linéaire donne des résultats très proches de ceux du modèle non-linéaire.

L'utilisation de telles échelles permet de raisonner en termes de continuum et non de classes. Cela est beaucoup plus facilement maniable pour les modèles qui sont mis en œuvre. Toutefois, il faut souligner que le passage à une mesure linéaire appauvrit la représentation de l'espace social, représenté comme multidimensionnel dans une grande partie des travaux sociologiques français. Par exemple, bien que l'échelle de Chambaz, Maurin et Torelli synthétise plusieurs éléments tels que le salaire ou les conditions de travail, ceux-ci sont réduits à une unique hiérarchie et on perd ainsi plusieurs distinctions qui peuvent être importantes dans l'analyse de l'influence de l'origine familiale, par exemple celle entre salariés et indépendants.

Ces transformations en variables quantitatives ont donc leurs limites ; toutefois, elles sont nécessaires pour utiliser les modèles MIMIC et donc introduire de l'analyse de fratrie dans l'étude de l'influence familiale, ce qui permet de renouveler ce champ de recherche et d'exploiter de façon plus approfondie la structure des données de l'enquête FQP 2003.

Score de compétences scolaires dans l'enquête INED

L'enquête sur le niveau intellectuel des enfants d'âge scolaire a nécessité la construction de tests permettant de comparer l'ensemble des élèves de 6 à 14 ans. L'enjeu était de proposer des cahiers de tests de difficulté variable selon la classe suivie par les élèves, puis de les étalonner pour les rendre comparables, d'une part en tenant compte de la difficulté relative de chaque question, d'autre part en prenant pour référence les résultats des autres enfants du même âge. Nous considérons ces scores de QI comme une approximation du niveau scolaire, qui n'a de sens que lorsqu'on les utilise à un niveau agrégé¹.

¹ D'une part, P. Benedetto, le concepteur de cette échelle, indique que les résultats doivent être utilisés avec prudence en raison des erreurs de mesure dues au caractère collectif du test, qui s'appliquent particulièrement aux enfants proches de la normale. Selon le chercheur, c'est le cas des enquêtés puisqu'ils sont tous scolarisés. Nous retenons que ces scores dits d'« intelligence » sont moins informatifs pour les enfants les plus en difficultés, ou à l'inverse pour ceux qui sont le plus à l'aise dans l'institution scolaire. D'autre part, on observe une variation ordonnée des scores moyens selon la catégorie socioprofessionnelle du chef de famille : de 100 pour les ouvriers jusqu'à 110 pour les cadres. Le score de QI correspond bien au jugement du maître sur la réussite de l'élève, l'ordre est le même et l'écart de QI entre les différentes catégories d'appréciations est à peu près constant (il est un peu plus élevé entre les élèves réussissant bien et les élèves jugés excellents, ainsi qu'entre les jugés "mauvais" et ceux que l'on considère comme "médiocres").

Méthodologie de l'analyse de fratrie

Modèles multiniveaux

Nous avons utilisé pour cette enquête des modèles multiniveaux, ou modèles linéaires hiérarchiques, qui permettent de mesurer l'importance respective de différents niveaux imbriqués les uns dans les autres. Dans le cadre des modèles de fratrie, le premier niveau est le niveau individuel, et le deuxième niveau est le niveau familial. Les modèles multiniveaux permettent alors, au moyen des parts de variance expliquées, de déterminer l'influence de chaque niveau sur une variable d'intérêt, ici, le score obtenu au test. Ces modèles sont décrits notamment par Snijders et Bosker (1999).

Le modèle vide à deux niveaux est modélisé par deux équations :

Niveau 1 (individuel) : $Y_{ij} = \alpha_j + u_{ij}$

Niveau 2 (familial) : $\alpha_j = \beta + v_j$

α_j représente la moyenne pour la variable Y au sein de la fratrie j et u_{ij} les variations individuelles autour de cette moyenne. β représente alors la moyenne de la variable Y parmi l'ensemble des fratries considérées, v_j reflétant les variations de chaque fratrie autour de cette moyenne. La variance des scores obtenus peut ainsi être décomposée en deux parties : une partie intra-familles (*within*) et une partie inter-familles (*between*). Le rapport entre la variance inter-familles et la variance totale reflète alors l'impact total de la famille sur les scores obtenus. Ce rapport est appelé coefficient de corrélation intra-classe (ICC).

Modèles MIMIC

Les modèles MIMIC (Multiples Indicateurs Multiples Causes), qui sont un type de modèles à équations structurelles, reposent sur le concept de variable latente. Dans le cadre des modèles de fratrie, cette variable latente se comprend comme un « facteur familial ». Elle est non-observable mais peut être décrite et mesurée à partir de variables observées appelées variables manifestes, qui comprennent ici nos variables d'intérêt qui sont le niveau d'éducation ou le statut professionnel, et des variables explicatives. Dans un schéma causal, ce facteur familial est ainsi déterminé par des caractéristiques observables mais également inobservables de la famille, et influe sur une variable d'intérêt donnée, renseignée pour les différents membres de la fratrie, et qui peut également être influencée par des caractéristiques individuelles.

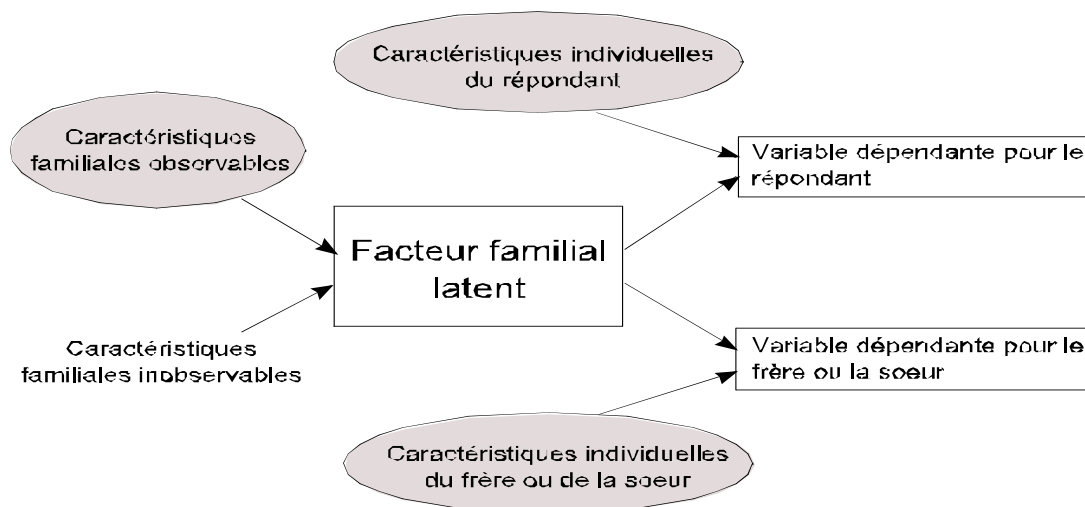


Figure 1 : Modèle MIMIC simple

Lecture : Les flèches représentent les relations de causalité supposées entre les différentes variables, présentes dans l'enquête ou estimées. La variable dépendante, comme elle est renseignée pour le répondant et un membre de sa fratrie, permet l'estimation du facteur familial latent. L'influence familiale cohésive est égale à la valeur du coefficient qui représente l'effet du facteur familial latent sur la variable dépendante. À partir des caractéristiques familiales présentes dans l'enquête, on estime l'ampleur de l'influence des caractéristiques familiales inobservables.

Le schéma met donc en évidence l'absence d'effet direct des caractéristiques familiales sur le niveau d'éducation ou le statut professionnel des enfants : cette influence passe exclusivement par la variable latente inférée en s'appuyant sur la ressemblance entre frères et sœurs. Empiriquement, l'influence cohésive de la famille est alors définie comme la part inter-familles de la variance des parcours, la part de variance intra-familles reflétant au contraire la différenciation au sein de la fratrie. C'est donc sur cette décomposition de la variabilité des parcours entre fratries et au sein des fratries que repose la mesure de l'influence familiale cohésive.

Jöreskog et Goldberger (1975) détaillent la spécification et l'estimation d'un modèle MIMIC comprenant une seule variable latente. L'estimation d'un tel modèle MIMIC requiert de spécifier l'ensemble des influences existant entre les N variables, que celles-ci soient manifestes ou latentes. Le modèle est ainsi décrit dans notre cas par les équations suivantes :

$$Y_L = \alpha_L + \beta_L \cdot X_L + \varepsilon_L$$

$$Y_R = \alpha_R + \beta_R \cdot X_R + \gamma_R \cdot Y_L + \varepsilon_R$$

$$Y_{FS} = \alpha_{FS} + \beta_{FS} \cdot X_{FS} + \gamma_{FS} \cdot Y_L + \varepsilon_{FS}$$

Avec Y_L la variable latente, X_L les caractéristiques observables communes à la fratrie, X_R et X_{FS} les caractéristiques individuelles du répondant et de son frère ou sa sœur tiré(e) au hasard respectivement, et les différents α et ε des constantes et termes d'erreur associés à chaque équation. On impose par ailleurs dans notre spécification l'égalité des coefficients α_R et α_{FS} , celle de β_R et β_{FS} , de γ_R et γ_{FS} , et enfin l'égalité de la variance de ε_R et ε_{FS} , car les répondants et frères/sœurs jouent ici des rôles symétriques.

Classification

Pour préciser l'interprétation de l'influence familiale, et en particulier les différents types d'articulation existant entre proximité verticale et proximité horizontale dans la famille, nous réalisons une classification ascendante hiérarchique avec consolidation par la méthode des centres mobiles. Nous utilisons les six variables actives suivantes : les écarts de niveau d'étude et de score de profession dans la fratrie, ainsi que les mêmes écarts avec la situation du père et avec celle de la mère. Nous avons choisi de construire trois classes en fonction de la règle du coude (décrochage de l'inertie interclasse à partir de 4 classes). Cela nous permet de dresser une typologie des différents parcours familiaux, que nous présentons dans ce qui suit. Nous utilisons comme variable supplémentaire une variable de mobilité subjective qui renseigne si la personne a le sentiment d'avoir accédé à un statut plus élevé, équivalent ou plus bas que son père pour décrire ces classes en termes de mobilité ascendante ou descendante.

Principaux résultats

Estimation de l'influence familiale cohésive sur le niveau d'éducation

L'analyse de la variance peut être effectuée, de manière ici équivalente, par un modèle multiniveaux ou un modèle MIMIC². Un modèle multiniveaux vide, à deux niveaux (niveau individuel emboîté dans un niveau familial) donne un coefficient de corrélation intraclasse (ICC) de 54,1% quand on utilise l'échelle modale du nombre d'années d'études, 55,4% quand on utilise l'échelle médiane. Dans le cas de la variable modale, 54,1% de la variabilité est donc attribuable au niveau famille et 45,9% au niveau individuel : la variabilité inter-familles est supérieure à la variabilité intra-famille.

Sieben (2001) obtenait le même ordre de grandeur sur un échantillon de 11 pays, ce pourcentage variant alors entre 50 et 60%. L'influence familiale au sens de la proximité dans la fratrie permet ainsi d'expliquer plus de la moitié de la diversité des parcours scolaires, moins de la moitié de celle-ci étant due à la différenciation au sein de la fratrie. L'inclusion de caractéristiques individuelles dans le modèle MIMIC permet d'augmenter très légèrement la part expliquée de la variabilité du nombre d'années d'études, de 54,3 à 54,8% : l'inclusion de ces caractéristiques individuelles, telles que le rang dans la fratrie, ne permet d'expliquer qu'une partie mineure de la variance du nombre d'années d'études comparativement au facteur familial. Enfin, l'inclusion dans le modèle d'un ensemble extensif de caractéristiques familiales observables disponibles dans l'enquête FQP comme la PCS du père, de la mère, ou le statut d'activité du père ou de la mère (tableau 2 en annexe) conduit à distinguer ce qui dans le facteur familial latent dépend de caractéristiques familiales observables et ce qui correspond à des caractéristiques familiales inobservables agissant de la même manière sur l'éducation des deux membres de la fratrie : 49,3% de la variance du facteur familial latent est expliquée par les caractéristiques familiales observables. Ainsi, même l'inclusion d'un ensemble très complet de variables familiales observables ne permet de rendre compte que de la moitié de l'influence cohésive exercée par l'appartenance familiale sur le nombre d'années d'études. On retrouve ici encore l'ordre de grandeur obtenu par Sieben (2001), qui montre que 45% environ de la variance du facteur familial est expliquée par les caractéristiques observables. On observe donc que les caractéristiques observables n'expliquent que pour moitié le facteur familial, qui lui-même explique plus de la moitié des variations des niveaux d'éducation.

Il est par ailleurs important de constater que les coefficients standardisés du modèle complet (donnés dans le tableau 3 en annexe) sont tout à fait cohérents avec ce qui serait attendu au regard de la littérature existante. En effet les déterminants les plus importants du facteur familial sont le score de la PCS et le nombre d'années d'études du père, puis le nombre d'enfants dans la fratrie, puis le nombre d'années d'études et le score de la PCS de la mère. Le nombre d'enfants dans la fratrie exerce un effet négatif sur le facteur familial, mais cela est d'autant moins vrai que la catégorie sociale est élevée, comme le montrent les termes d'interaction du nombre d'enfants dans la fratrie avec la catégorie sociale du père. Cet effet a été observé par exemple par Vallot (1973) dans le cas de la performance scolaire. Quant aux caractéristiques individuelles, le fait d'être une fille augmente

² Il est utile de préciser ici les liens existant entre ces deux modélisations. L'estimation d'un modèle MIMIC vide (représenté au graphique 1) permet de retrouver la part d'influence familiale estimée grâce au coefficient de corrélation intraclasse (ICC) avec le modèle multiniveaux. L'ICC est défini comme la part de la variance du nombre d'années d'études expliquée par le facteur familial, qui correspond dans le cadre du modèle MIMIC au coefficient de détermination (R^2) de la régression du nombre d'années d'études sur le facteur familial. Dans notre cas, les chiffres de l'ICC et du R^2 diffèrent légèrement car l'échantillon est réduit de 26305 fratries pour le multiniveaux à 25259 pour le modèle MIMIC. En effet, on élimine les individus pour lesquels il y a des valeurs manquantes parmi les variables explicatives, celles-ci étant incluses dans les modèles MIMIC suivants, plus complets.

par exemple significativement le nombre d'années d'études toutes choses égales par ailleurs. Il existe par ailleurs un avantage au fait d'être l'aîné(e), mais également un léger avantage au fait d'occuper le dernier rang dans la fratrie (mais la différenciation selon le rang dans la fratrie reste faible par rapport à l'influence familiale commune à tous les frères et sœurs).

L'influence familiale sur la profession

La même analyse, menée dans le cas où la variable d'intérêt est le score de la profession, attribue au facteur familial 31,9% des variations du statut de la profession. L'estimation est alors réalisée sur un échantillon de 23181 fratries. L'influence familiale commune aux membres de la fratrie exercée sur le statut de la profession est donc plus réduite que celle exercée sur le nombre d'années d'études : plusieurs facteurs interviennent après la fin des études et différencient les carrières des membres d'une même fratrie. C'est la tendance présente dans l'analyse de Sieben (2001), pour qui le facteur familial explique 37% de la variance du statut professionnel. L'inclusion des caractéristiques individuelles dans le modèle complet n'apporte là encore que peu de pouvoir explicatif supplémentaire : la part expliquée de la variabilité du statut de la profession vaut alors 33,0%. Le facteur familial, qui explique donc le statut professionnel des enfants à hauteur de près d'un tiers, est lui-même déterminé à 51,8% par les caractéristiques familiales observables lorsqu'on inclut celles-ci dans le modèle. De même que pour le nombre d'années d'études, le facteur familial n'est ici expliqué qu'à moitié par les caractéristiques familiales observables, qui ne décrivent donc que partiellement l'influence familiale commune aux membres de la fratrie.

Les coefficients de ce modèle complet sont donnés en annexe (tableau 4). Les statuts de la profession des parents gagnent en importance pour expliquer les scores de la profession des enfants, tandis que le nombre d'années d'études des parents devient moins déterminant, en particulier celui de la mère. Concernant les caractéristiques individuelles, si le fait d'être une fille exerçait un effet positif dans le cas du nombre d'années d'études, il exerce au contraire ici un effet significativement négatif.

L'influence familiale sur les compétences dans le jeune âge

51,1% de la variance des résultats aux tests de compétence sont expliqués par le niveau familial. Cette proximité familiale est intéressante puisque nous nous intéressons à un plus jeune âge (6 à 14 ans, c'est-à-dire concernant des enfants nés entre 1951 et 1959), et à un groupe appartenant à une même cohorte (ces personnes ont entre 44 à 52 ans au moment de la passation de l'enquête FQP de 2003). L'ordre de grandeur de la part expliquée par le niveau familial est comparable à ce qu'on observe pour le nombre d'années d'études.

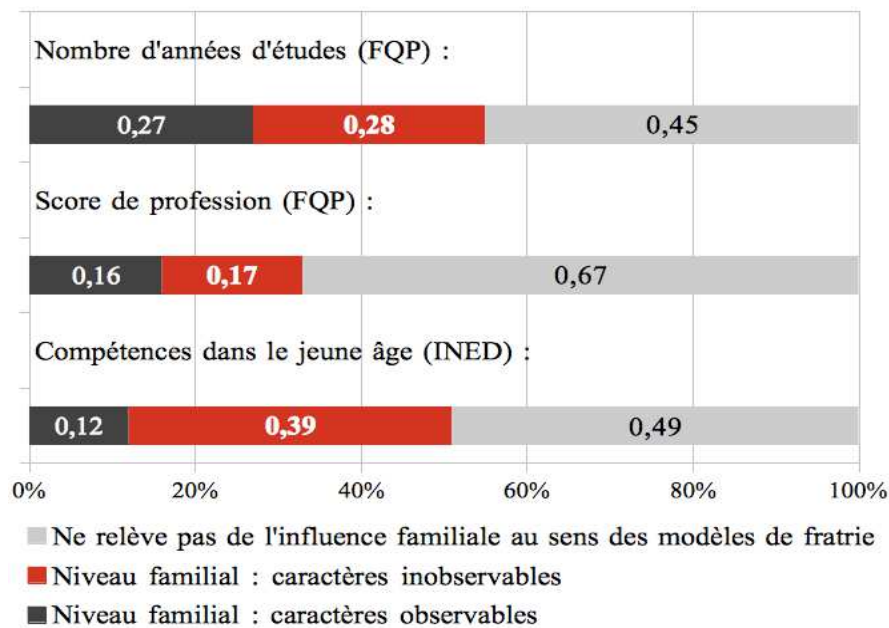


Figure 2 : L'influence familiale cohésive décomposée entre l'effet des caractéristiques familiales observables dans l'enquête FQP et la part inobservable de l'influence familiale cohésive
 Lecture : 55% des variations des parcours individuels en termes de niveau d'études obtenu sont expliquées par le facteur familial latent, et relèvent à ce titre de l'influence familiale cohésive. La moitié de cette influence, à savoir 27% de la variance totale, est expliquée par les variables renseignées dans l'enquête FQP. L'autre moitié, 28%, renvoie à des caractères inobservables.

L'effet propre de l'éducation

Les modèles MIMIC permettent de complexifier les influences à l'œuvre entre les différentes variables, par rapport à celles qui ont été spécifiées en première analyse. Il est ainsi possible d'étudier l'effet de l'éducation sur la profession tout en contrôlant par l'influence cohésive de la famille (et non pas simplement par des caractéristiques observables). Une partie du biais de variable omise est ainsi évacuée, qui correspond aux caractéristiques inobservées qui sont captées par le facteur familial latent. Ces caractéristiques peuvent avoir trait aux compétences individuelles mais également à des éléments liés à l'origine sociale sans relever des compétences : en effet, à compétences données, un individu issu d'une famille plus favorisée pourra à la fois être incité à poursuivre plus longtemps ses études, et bénéficier du capital social de ses parents pour s'établir plus facilement dans la vie professionnelle.

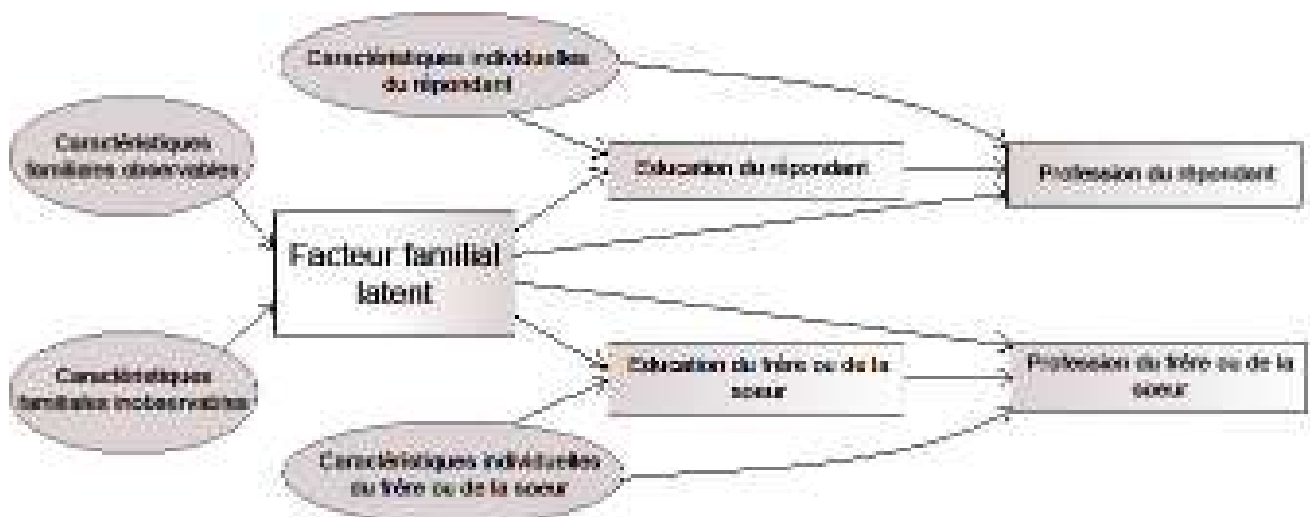


Figure 3 : Modèle MIMIC permettant de mesurer l'effet du niveau d'études sur la profession
Lecture : Les flèches représentent des relations de causalité modélisées entre des variables présentes dans l'enquête ou estimées.

Le modèle inclut les influences représentées sur la figure 3. L'échelle modale est retenue pour l'éducation et l'échelle non-linéaire de statut social pour la profession. Les caractéristiques familiales et individuelles observables sont les mêmes que dans les modèles MIMIC précédents. Le paramètre d'intérêt est le coefficient non-standardisé correspondant à l'influence de l'éducation du répondant sur la profession du répondant : celui-ci, significatif au seuil de 1%, est de 0,127. Une année d'études supplémentaire fait ainsi augmenter de 0,127 le score de statut social associé à la profession exercée (sur une échelle allant de -1,52 à 2,19). Le coefficient équivalent dans une régression linéaire « naïve » contrôlant par toutes les caractéristiques familiales et individuelles observables est 0,164, également significatif. Le fait de contrôler uniquement par les caractéristiques observables conduit ainsi à surestimer l'influence de l'éducation sur le score de la profession de plus de 25%.

Prendre en compte le changement du contexte scolaire et mesurer l'évolution de l'influence familiale cohésive au cours du temps

Les estimations qui ont été réalisées rassemblent dans un même échantillon des individus ayant grandi à des périodes pouvant être très éloignées dans le temps. En couplant ce fait avec l'allongement de la durée des études au fil des générations, on pourrait mesurer une variance inter-familles du nombre d'années d'études ne correspondant pas au phénomène que nous souhaitons approcher (l'influence cohésive de la famille au sein d'une génération donnée), artificiellement gonflée par les écarts d'éducation entre des enfants appartenant à des fratries ayant grandi à des périodes éloignées de plus de quarante ans. Cela conduit à estimer les modèles MIMIC cohorte par cohorte : il s'agit donc d'éviter de capter une trop grande partie d'évolutions structurelles dans la variance inter-familles. Dans le cas de la variable médiane pour le nombre d'années d'études, 54,8% de la variance est expliquée par le facteur familial pour les 34 ans et moins, 53,5% pour les 35 à 44 ans, 53,1% pour les 45 à 54 ans et enfin 51,3% pour les 55 ans et plus³. L'estimation du modèle MIMIC sur l'ensemble de la population ne conduit donc pas à fortement surestimer l'influence familiale (la part expliquée par le facteur familial étant dans ce cas de 55,6%).

Cette analyse permet par ailleurs de mettre en évidence une éventuelle évolution de l'influence familiale et de l'effet propre de l'éducation au cours du temps.

Pour ce qui est de l'influence familiale cohésive, il apparaît qu'aucune des différences de coefficients entre deux cohortes consécutives (dont la précision est estimée à travers un *bootstrap* comptant 500 itérations) n'est significative. Seul l'écart entre la cohorte la plus jeune et la plus ancienne est significatif au seuil de 5%. Il est toutefois à peine significatif à ce seuil, et ne l'est plus dès lors qu'on utilise la variable modale. La tendance temporelle observée, qui est celle d'une influence familiale commune aux frères et sœurs de plus en plus élevée dans les générations récentes, ne semble donc pas assez marquée pour que l'on puisse en tirer de conclusions fortes. Elle va en tout cas à l'encontre d'une éventuelle hypothèse d'individualisation liée aux transformations de

³ Pour concilier la prise en compte des cohortes avec l'analyse de fratrie (notamment dans le modèle multiniveaux), nous avons besoin que les deux membres d'une fratrie appartiennent à la même cohorte. Nous avons pour cela attribué aux deux frères et sœurs la cohorte du plus jeune. Quatre cohortes (25-34 ans, 35-44 ans, 45-54 ans, 55-65 ans en 2003) ont ainsi été créées, de taille suffisante pour pouvoir mener des analyses sur chacune d'elles : elles contiennent respectivement 7 800, 9 130, 7 890 et 5 134 individus.

la famille. Pour ce qui est du score de la profession, aucune tendance globale significative n'apparaît. Les chiffres sont toutefois plus difficiles à interpréter dans ce cas, du fait des effets d'âge qui viennent se mêler aux effets de génération.

Dans le cas de l'effet propre de l'éducation, la progression observée, bien que monotone et semblant indiquer une influence moins forte de l'éducation sur la profession dans les générations anciennes (voir le tableau 1 ci-dessous) ne peut elle non plus être interprétée de manière conclusive, puisqu'elle mêle un effet d'âge (celui-ci augmentant, le temps écoulé entre la fin des études et le dernier emploi occupé augmente, et l'influence directe de l'éducation sur l'activité professionnelle diminue) et un effet de génération (réelle augmentation de l'effet propre au fil des générations). La cohorte la plus jeune, pour laquelle la profession renseignée est la plus proche temporellement de la première profession exercée à la sortie des études, connaît un écart entre les deux coefficients (effet pur et effet capté par régression linéaire avec contrôles) plus réduit que pour l'ensemble de la population. Même dans ce cas, on surestime de plus de 20% l'effet de l'éducation sur le statut de la profession si l'on ne tient pas compte de l'influence cohésive de la famille d'origine.

Cohorte	Modèle MIMIC	Régression linéaire
34 ans et moins	0,160	0,197
35 à 44 ans	0,146	0,198
45 à 54 ans	0,135	0,178
55 ans et plus	0,116	0,156

Tableau 1 – Effet du nombre d'années d'études sur le score de la profession NB : Pour chaque cohorte, on donne l'estimation de l'effet pur et de l'effet capté par régression linéaire avec contrôles

Lecture : Parmi les individus de notre échantillon âgés de 34 ans et moins, on observe qu'une année supplémentaire d'études fait augmenter de 0,160 le score de statut social.

Influence familiale cohésive et mobilité sociale

Nous réalisons comme indiqué plus haut une classification pour déterminer les différents types d'articulation existant entre proximité verticale et proximité horizontale dans la famille :

- Classe 1 (57% de la population) : reproduction sociale marquée. Cette classe est caractérisée par une forte proximité à la fois dans la fratrie et avec les parents, pour le niveau d'études et le score de profession. En décrivant la classe, on voit que les grandes fratries (4 individus ou plus) et les familles divorcées sont surreprésentées. C'est aussi le cas des parents agriculteurs ou sans diplôme. La part des individus estimant avoir atteint un statut "bien plus élevé" que leur père est la plus faible dans ce groupe, alors que la part des personnes estimant avoir atteint "à peu près le même statut" est la plus élevée.
- Classe 2 (23% de la population) : forte différenciation des parcours dans la fratrie. Les écarts dans la fratrie sont supérieurs à la moyenne, ainsi que les écarts de niveaux d'études avec les parents. Les écarts avec les parents du score de la profession sont égaux à la moyenne. Cette classe est la plus hétérogène. Y sont surreprésentées les fratries de deux individus, et les familles n'ayant pas connu de divorce pendant la scolarité. Les parents artisans, commerçants ou chefs d'entreprise, les mères exerçant une profession intermédiaire, et les pères cadres le sont également. Enfin, les mères titulaires du seul BEPC ou bachelières et les pères diplômés du supérieur sont surreprésentés.
- Classe 3 (20% de la population) : forte mobilité familiale. L'écart dans la fratrie est inférieur

à la moyenne pour le score de profession. L'écart avec les parents est très supérieur à la moyenne pour le niveau d'études et le score de la profession. Les modalités surreprésentées sont les fratries de trois membres, les familles n'ayant pas connu de divorce pendant la scolarité, les pères cadres, employés ou profession intermédiaire et les mères employées. Cette classe contient la plus forte proportion d'individus estimant avoir un statut "bien plus haut" que leur père, et les proportions d'individus de même statut ou de statut inférieur à celui de leur père les plus basses.

Ces résultats montrent que l'existence d'une proximité dans la fratrie, et donc d'une influence familiale forte estimée par les modèles employés dans la partie précédente, peut masquer des parcours familiaux très différents. D'un côté la classe 1 pour laquelle la mobilité sociale faible semble due au manque de ressources des familles (absence de qualification des parents, positions socio-économiques basses), et de l'autre la classe 3, c'est-à-dire les plus favorisés, caractérisés par une mobilité ascendante plus importante. La proximité dans la fratrie montre dans ces deux cas une influence familiale certes liée à l'origine sociale, mais qui ne se résume pas à ce qui est mesuré par les variables usuelles. Les caractères inobservés partagés par les membres de la fratrie montrent la pertinence du niveau familial pour expliquer les situations de mobilité.

L'analyse de notre deuxième classe nous invite à plus de prudence au sujet de notre mesure de l'influence familiale. Le fait d'exclure par construction les différences entre frères et sœurs de cette mesure conduit à écarter des formes d'influence familiale qui passeraient précisément par ce biais, ce qui nous invite à tenter d'explicitier les limites inhérentes à notre modélisation.

Questions soulevées par notre étude

Sur le plan théorique : quels liens peut-on établir entre notre modélisation et les différentes formes de socialisation familiale ?

La classe 2 de la classification qui précède pose la question du lien entre différenciation des parcours dans la fratrie et influence familiale. Notre analyse, qui montre que 23 % de la population observée sont concernés par de fortes discordances, à la fois entre membres de la même fratrie, mais aussi en comparaison avec les parents, est cohérente avec l'analyse de C. Peugny, selon lequel malgré l'existence d'une convergence des destins professionnels dans les fratries, le phénomène d'hétérogénéité de réussite au sein des fratries est loin d'être négligeable (Peugny, 2012).

Peut-on considérer que l'influence familiale est faible lorsque les parcours sont différents dans la fratrie ? Bernard Zarka (1995) fournit une autre explication, selon laquelle le statut professionnel du père peut être transmis inégalement entre les différents enfants (par exemple si l'on favorise l'aîné). Sont d'ailleurs surreprésentés dans cette classe les enfants d'artisans, commerçants ou chefs d'entreprise : cela rejoint l'analyse de S. Gollac (2009) qui montre que la différenciation de transmission des ressources familiales au sein de la fratrie est prononcée chez les indépendants.

Nous pouvons aussi considérer l'existence d'une forme de socialisation favorisant pour chacun le développement d'une identité personnelle distincte de celle des autres membres de la famille, parents comme enfants (Singly (de), 1996). Cette dernière hypothèse est aussi confirmée par les études des styles éducatifs des familles (Kellerhals et Montandon, 1991), qui montrent l'existence d'une pluralité de fonctionnements de la cohésion familiale, dont celui d'« association ». Ce dernier est fondé sur une forte individuation, une négociation des rôles au sein de la famille et une ouverture sur l'extérieur. Cette forme familiale est selon les auteurs davantage représentée quand on monte dans l'échelle sociale, ce qui est cohérent avec nos résultats. Ainsi la

variance des parcours qui n'est pas considérée comme relevant de l'influence familiale cohésive (partie représentée en gris sur la figure 2, ou encore classe 2 représentant la différenciation dans la fratrie) n'est pas nécessairement le signe d'une influence familiale faible, qui impliquerait d'invoquer l'influence de caractéristiques individuelles, ou encore de choix personnels. Elle indique également l'importance des formes non cohésives de l'influence familiale que nous venons de présenter.

Comme on inclut dans l'influence familiale tout ce qui est commun aux membres de la fratrie, cette mesure peut aussi apparaître *a contrario* comme trop large. Nous captions également des influences de l'environnement communes aux deux membres, par exemple dues au quartier de résidence ou aux pairs. L'existence de pratiques ségrégatives en matière de logement dans les catégories supérieures légitime en partie ce fait, mais il n'est pas exclu que pour une partie de la population on mesure un effet qui n'est pas à proprement parler familial. Il en est de même avec l'influence des pairs. Les parents peuvent contrôler les fréquentations de leurs enfants, ou ces derniers ne choisir que des camarades ayant une position sociale proche de la leur. Les analyses de l'influence des pairs sur les résultats scolaires fondées sur des données de réseaux montrent pourtant qu'au-delà de l'effet de sélection des pairs, il existe un effet propre de ces derniers (Lomi, Snijders *et al.*, 2011). Enfin, les traditions sociologiques françaises et américaines ont étudié l'influence familiale en lien avec la question de l'égalité des chances. Notre classification nous a permis de souligner la complémentarité entre les approches en termes de mobilité sociale, fondées sur une comparaison des caractéristiques des individus et de celles de leurs parents, et celles qui utilisent la ressemblance dans la fratrie. Une fois l'influence familiale cohésive mesurée à l'aide de l'analyse de fratrie, l'approche par la mobilité sociale permet de voir si la ressemblance dans la fratrie constitue une reproduction sociale ou une mobilité dans la fratrie. L'approche par la ressemblance dans la fratrie permet quant à elle de distinguer les situations où les parcours dans la fratrie sont très proches (qu'il s'agisse d'une mobilité de la fratrie ou de reproduction sociale) de celles où l'on observe au contraire une différenciation, qui renvoie à des formes non cohésives d'influence familiale.

Sur le plan technique : les difficultés d'une prise en compte du caractère discret des variables d'intérêt

Nous avons montré plus haut la méthodologie qui nous a servi à transformer les variables de diplôme en échelle numérique, en raison de la construction des modèles. Nous donnons ici des pistes exploratoires pour une méthode d'estimation plus complexe, qui vise à prendre en compte le caractère discret de cette variable d'intérêt. Comment appliquer l'analyse de fratrie, non plus à la variable du nombre d'années d'études (variable quantitative) mais au diplôme (variable qualitative) ? Le principal problème qu'il convient de résoudre est celui de l'estimation du coefficient intraclasse (ICC). La méthode d'approximation la plus simple consiste à considérer que le paramètre de variance du niveau résiduel (qui n'est pas fourni par le logiciel) est de 3,29, c'est-à-dire la variance d'une distribution logistique standard. Goldstein, Browne et Rasbash (2002) montrent, dans une comparaison de différentes méthodes appliquées à un même exemple de régression logistique, que cette façon de procéder peut donner des résultats très différents de ceux obtenus avec les autres. Nos tentatives ont confirmé ces réserves en nous donnant des résultats non seulement incohérents avec les analyses précédentes, mais également très différents selon les changements de spécification⁴. Les autres méthodes d'estimation de l'ICC, qui seraient plus fiables,

⁴ Nous avons utilisé un modèle logistique cumulatif après avoir regroupé les diplômes en trois classes ordonnées (pas de diplôme, diplôme inférieur ou équivalent au baccalauréat, diplôme d'études supérieures), puis en estimant l'ICC à l'aide de l'hypothèse de distribution logistique dans la

font appel à une méthode de linéarisation ou de simulation. La prise en compte du caractère discret des variables d'intérêts nécessite donc un travail technique important pour repenser l'estimation des modèles de fratrie.

Toutefois, il semble que le fait de considérer les variables d'intérêt comme continues dans nos estimations du modèle MIMIC n'induit pas de biais important dans nos estimations. Par exemple, d'après Creech et Johnson (1983), le fait d'utiliser une variable catégorielle ordonnée dans le cadre des modèles à indicateurs multiples introduit un biais trop faible pour modifier significativement les estimations, dans le cas où l'on a cinq catégories ou plus (ce qui est notre cas aussi bien pour le nombre d'années d'études que pour le score de la profession).

Conclusion

Le fait de considérer la socialisation du point de vue de la fratrie nous semble être fécond pour mettre en évidence que l'influence familiale peut conduire à des parcours fraternels plus ou moins diversifiés selon l'efficacité de la transmission de dispositions scolairement rentables et la nature des normes éducatives parentales. L'approfondissement de ces questions nécessiterait l'introduction plus systématique de données de fratrie dans les enquêtes de la statistique publique, ainsi que la conduite d'enquêtes qualitatives permettant d'analyser l'influence familiale au sein de fratries selon leur position dans l'espace social. Cependant, ces données de fratrie posent aussi des questions méthodologiques et sociologiques qu'il convient de prendre en compte. Le tableau ci-dessous nous semble résumer les principaux avantages et limites de notre démarche.

<p>Intérêt des données de fratrie</p> <ul style="list-style-type: none"> - La ressemblance observée au sein d'une fratrie capte l'ampleur de l'influence familiale, qu'elle soit due à des caractères observables ou non (Sewell, Hauser, Springer, 2004). Elle remédie à l'impossibilité de connaître de façon exhaustive les variables qui caractérisent le milieu familial des individus. 	<p>Problèmes posés par le recueil de données de fratrie</p> <ul style="list-style-type: none"> - Dans FQP, la moyenne de niveau d'étude est un peu plus élevée pour les membres de la fratrie ; - On pourrait penser interroger le membre de la fratrie le plus proche en âge, pour éviter les problèmes de mémoire, mais on aurait alors un écart d'âge biaisé vers le bas. - Dans l'enquête INED, du fait du plan de sondage et de l'appariement <i>a posteriori</i>, on ne capte que les frères et sœurs qui étudient dans le même établissement.
<p>Méthodologie</p> <ul style="list-style-type: none"> - Utilisation de modèles de fratrie (MIMIC ou multiniveaux) pour mesurer l'influence familiale <i>cohésive</i> : ce que partagent les membres de la fratrie ; - Articulation avec une classification pour faire le lien entre influence familiale <i>cohésive</i> et mobilité-immobilité sociale. 	<p>Prolongements méthodologiques</p> <ul style="list-style-type: none"> - Meilleure prise en compte du caractère discret des variables d'intérêt ; - Isolation des effets structurels, particulièrement l'allongement de la durée d'études moyenne au cours du temps.
<p>Intérêts sociologiques</p>	<p>Questions sociologiques</p>

Proc HP mixed de SAS. L'influence familiale apparaissait extrêmement faible. Les tentatives d'utilisation d'un modèle logistique simple en étudiant des seuils binaires (indicatrices indiquant un niveau de diplôme minimum) ont donné des résultats encore différents, qui ne nous ont pas semblés davantage robustes.

- mesure de la force de l'influence familiale, qui prend en compte tous les processus informels de la socialisation partagés dans la fratrie, y compris les caractères "secondaires" (Lahire) ; - lien entre cette approche et la sociologie de la mobilité/reproduction sociale.	- L'influence familiale s'applique-t-elle de la même manière pour les enfants uniques ? - Tout ce qui est partagé entre frères et sœurs relève-t-il de l'influence familiale ? - Formes d'influence familiale non incluses dans l'influence familiale cohésive : invitation à une différenciation de chaque membre de la famille (Singly (de), 1996)
--	--

Tableau 2: synthèse des apports de l'analyse de fratrie, et des questions posées par ces dernières

Annexes

Tableau des variables de l'enquête FQP

	Minimum	Moyenne	Maximum
Score de la PCS du père	-1,52	-0,202	2,19
Score de la PCS de la mère	-1,52	-0,447	2,19
Père inactif n'ayant jamais travaillé	0	0,001	1
Mère inactive n'ayant jamais travaillé	0	0,314	1
Score de la PCS du grand-père maternel	-1,52	-0,368	2,19
Score de la PCS du grand-père paternel	-1,52	-0,326	2,19
PCS du grand-père maternel manquante	0	0,308	1
PCS du grand-père paternel manquante	0	0,335	1
Nombre d'années d'études du père	8	10,3	18
Nombre d'années d'études de la mère	8	10,1	18
Nombre d'enfants dans la fratrie	2	4,00	19
Interaction avec une dummy « père cadre / chef d'entreprise »	0	0,325	13
Interaction avec une dummy « père ouvrier ou employé »	0	2,84	19
Divorce ou séparation des parents pendant les études	0	0,067	1
Résidence en Ile-de-France à la fin des études	0	0,138	1
Taux de chômage dans la région des parents à la fin des études	5,5	12,9	22,4
PCS manquante pour le père (hors inactif, jamais travaillé)	0	0,003	1
PCS manquante pour la mère (hors inactive, jamais travaillé)	0	0,013	1
Sexe féminin	0	0,508	1
Aîné	0	0,288	1
Benjamin	0	0,345	1
Âge de la mère à la naissance	11	27,7	58

Tableau 2 – Statistiques descriptives portant sur les variables explicatives des modèles MIMIC dans l'enquête FQP 2003⁵

Estimation d'un modèle MIMIC : explication des étapes

Afin d'estimer un modèle MIMIC, il est tout d'abord nécessaire de spécifier toutes les influences

⁵ Les variables de nombre d'années d'études et de score de la profession pour les parents et grands-pères sont respectivement l'échelle modale et l'échelle non-linéaire.

existant entre les N variables, que celles-ci soient manifestes ou latentes.

Pour chaque variable, on obtient ainsi une équation, qui est de type :

$V_i = u_i$, si V_i est exogène

$V_i = aV_j + bV_k + u_i$, si V_i dépend d'autres variables, ici par exemple V_j et V_k

Le lien que l'on spécifie entre les différentes variables à l'aide de ces équations permet de construire leur matrice de covariance théorique, et d'estimer les paramètres des équations par la méthode du maximum de vraisemblance.

Les modèles MIMIC se distinguent des modèles multiniveaux par plusieurs aspects. D'une part la structure hiérarchique des données, qui est naturelle dans les modèles multiniveaux puisque les individus sont naturellement emboîtés dans les fratries, est artificielle dans les modèles MIMIC : lorsque la fratrie comporte plus de deux individus, toutes les paires possibles de frères et sœurs doivent être formées pour observer la similarité entre ceux-ci. D'autre part, et contrairement aux modèles multiniveaux, la modélisation MIMIC permet d'étudier l'effet "pur" de l'éducation sur une variable donnée, c'est-à-dire l'effet de l'éducation une fois que l'on a contrôlé par l'impact cohésif total de la famille. Lorsque l'on n'effectue pas ce contrôle et que l'on mesure directement l'influence de l'éducation sur le statut professionnel par exemple, il existe un biais de variable omise qui conduit à surestimer cette influence. En effet, comme cela a été souligné dans les *status attainment models* en particulier, l'origine familiale peut avoir un effet à la fois sur le niveau d'éducation et sur le statut atteint. Les modèles MIMIC permettent d'évacuer ce biais. Ils ne permettent par contre pas l'inclusion d'un niveau supplémentaire, ou la prise en compte du caractère discret de la variable dépendante, au contraire des modèles multiniveaux.

Coefficients obtenus dans les modèles MIMIC

	Echelle modale	Echelle médiane
Equation (0) :		
Score de la PCS du père	0,198 (***)	0,200 (***)
Score de la PCS de la mère	0,112 (***)	0,119 (***)
Père inactif n'ayant jamais travaillé	-0,005	-0,005
Mère inactive n'ayant jamais travaillé	-0,050 (***)	-0,050 (***)
Score de la PCS du grand-père maternel	0,064 (***)	0,066 (***)
Score de la PCS du grand-père paternel	0,024 (***)	0,026 (***)
PCS du grand-père maternel manquante	-0,091 (***)	-0,089 (***)
PCS du grand-père paternel manquante	-0,082 (***)	-0,082 (***)
Nombre d'années d'études du père	0,197 (***)	0,202 (***)
Nombre d'années d'études de la mère	0,137 (***)	0,142 (***)
Nombre d'enfants dans la fratrie	-0,176 (***)	-0,179 (***)
Interaction avec une dummy "père cadre ou chef d'entreprise"	0,028 (***)	0,034 (***)
Interaction avec une dummy "père ouvrier ou employé"	-0,057 (***)	-0,049 (***)
Divorce ou séparation des parents pendant les études	-0,018 (**)	-0,021 (***)
Parents résidant en Ile-de-France à la fin des études	0,021 (***)	0,022 (***)
Taux de chômage dans la région des parents à la fin des études	-0,004	-0,005
Score PCS manquant pour le père (hors inactif, jamais travaillé)	-0,011	-0,012 (*)
Score PCS manquant pour la mère (hors inactive, jamais travaillé)	-0,023 (***)	-0,020 (***)
Equation (1') :		
Facteur familial	0,737 (***)	0,745 (***)
Sexe féminin	0,028 (***)	0,026 (***)
Aîné	0,028 (***)	0,031 (***)
Benjamin	0,017 (***)	0,015 (***)
Âge de la mère à la naissance	0,064 (***)	0,065 (***)

Note sur les tableaux présentant l'effet des variables explicatives : Bien que l'on ait imposé l'égalité des coefficients entre répondant et frère/sœur dans les équations (1') et (2') expliquant le nombre d'années d'études, ces contraintes s'appliquent aux coefficients non-standardisés et non aux coefficients standardisés. Les variances des caractéristiques individuelles pouvant varier entre les répondants et l'ensemble des frères ou sœurs tirés au hasard, les coefficients standardisés ne sont pas nécessairement égaux pour (1') et (2'). Toutefois ils varient très peu, au maximum de 0,002 entre les deux équations. Les résultats présentés sont ceux portant sur l'échantillon des répondants.

Significatif au seuil de (*) 5% (**) 1% (***) 0,1% **Source** Enquête FQP 2003

Tableau 3 - Coefficients standardisés d'un modèle MIMIC expliquant le nombre d'années d'études, et incluant caractéristiques familiales et caractéristiques individuelles

	Score non linéaire	Score linéaire
Equation des variables familiales :		
Score de la PCS du père	0,227 (***)	0,226 (***)
Score de la PCS de la mère	0,156 (***)	0,155 (***)
Père inactif n'ayant jamais travaillé	-0,004	-0,005
Mère inactive n'ayant jamais travaillé	-0,001	-0,002
Score de la PCS du grand-père maternel	0,100 (***)	0,100 (***)
Score de la PCS du grand-père paternel	0,060 (***)	0,061 (***)
PCS du grand-père maternel manquante	-0,090 (***)	-0,089 (***)
PCS du grand-père paternel manquante	-0,087 (***)	-0,087 (***)
Nombre d'années d'études du père	0,128 (***)	0,129 (***)
Nombre d'années d'études de la mère	0,047 (***)	0,049 (***)
Nombre d'enfants dans la fratrie	-0,121 (***)	-0,122 (***)
Interaction avec une dummy "père cadre ou chef d'entreprise"	0,040 (***)	0,041 (***)
Interaction avec une dummy "père ouvrier ou employé"	-0,123 (***)	-0,121 (***)
Divorce ou séparation des parents pendant les études	-0,053 (***)	-0,053 (***)
Parents résidant en Ile-de-France à la fin des études	0,129 (***)	0,132 (***)
Taux de chômage dans la région des parents à la fin des études	-0,001	-0,001
Score PCS manquant pour le père (hors inactif, jamais travaillé)	-0,023 (**)	-0,022 (**)
Score PCS manquant pour la mère (hors inactive, jamais travaillé)	-0,011	-0,011
Equation des variables individuelles :		
Facteur familial	0,561 (***)	0,564 (***)
Sexe féminin	-0,103 (***)	-0,109 (***)
Aîné	0,043 (***)	0,043 (***)
Benjamin	-0,008	-0,008
Âge de la mère à la naissance	0,047 (***)	0,047 (***)

Significatif au seuil de (*) 5% (**) 1% (***) 0,1% **Source** Enquête FQP 2003

Tableau 4 - Coefficients standardisés d'un modèle MIMIC expliquant le score de la profession, et incluant caractéristiques familiales et caractéristiques individuelles

Classification

Classe n°	1	2	3
% pop	57	23	20
Fratrie	Ecart de niveau d'études et de score de la profession inférieurs à la moyenne	Ecart supérieurs à la moyenne (niveau d'études et score de la profession)	Ecart de niveau d'études égal à la moyenne, mais écart de score de la profession inférieur
Père	Ecart inférieurs à la moyenne	Niveau d'études : écart supérieur à la moyenne. Mais l'écart est égal à la moyenne pour le score de la profession.	Ecart très supérieur à la moyenne pour le niveau d'études et le score de la profession
Mère	Ecart inférieurs à la moyenne	Niveau d'études : écart supérieur à la moyenne. Mais l'écart est égal à la moyenne pour le score de la profession.	Ecart très supérieur à la moyenne pour le niveau d'études et le score de la profession
Surreprésentés	Grandes fratries (4 et plus), parents sans diplôme, parents agriculteurs, divorce pendant la scolarité.	Fratrie de deux individus. Pas de divorce pendant la scolarité. Parents artisans, commerçants ou chefs d'entreprise, mère profession intermédiaire, mère titulaire du seul BEPC ou bachelière. Père cadre, père diplômé du supérieur.	Fratrie de trois membres. Pas de divorce pendant la scolarité. père cadre, profession intermédiaire ou employé, mère employée,
Sous-représentés	Mère employée, artisan, commerçant ou chef d'entreprise, Père Profession intermédiaire ou cadre, fratrie de 2 individus	écart d'âge nul entre les membres de la fratrie, divorce pendant la scolarité. Mère ouvrière ou agricultrice, parents sans diplôme, père ouvrier.	Fratrie de 4 membres ou plus, écart d'âge important dans la fratrie, père ouvrier, parents agriculteurs, divorce pendant la scolarité, mère diplômée du supérieur,
Part des enquêtés estimant avoir atteint une position "bien plus élevée" que celle de leur père	9,7	16,9	22,6
Part des enquêtés estimant avoir atteint "à peu près la même" position que celle de leur père	30,6	23,8	21,8
Part des enquêtés estimant avoir atteint une position "plus basse" que celle de leur père	20,2	20,2	15,2

Tableau 5 - Caractérisation des classes obtenues avec la classification ascendante hiérarchique et typologie des différentes trajectoires

Bibliographie

Bowles S., "Schooling and Inequality from Generation to Generation", *Journal of Political Economy*, Vol. 80, n°3, 1972.

Chambaz C., Maurin E. et Torelli C., « L'évaluation sociale des professions en France. Construction et analyse d'une échelle des professions », *Revue française de sociologie*, Vol. 39, n°39-1, 1998.

De Graaf P.M., "The Impact of Financial and Cultural Resources on Educational Attainment in the Netherlands", *Sociology of Education*, Vol. 59, n°4, 1986.

Duncan O.D., Featherman D.L., Duncan B., *Socioeconomic background and achievement*, New-York, Seminar Press, 1972.

Ganzeboom H.B.G. et Treiman D.J., "Preliminary Results on Educational Expansion and Educational Opportunity in Comparative Perspective", in *Solidarity of generations: demographic, economic and social change, and its consequences : proceedings of a symposium held on April 7 and 8, 1993 at Utrecht University*, p.467-506, The Netherlands, Thesis Publishers, Amsterdam, 1994.

Gollac S., « Travail indépendant et transmissions patrimoniales : le poids des inégalités au sein des fratries », *Économie et Statistique*, n°417-418, 2009.

Goldstein H., Browne W., Rasbash J., *Partitioning variation in multilevel models*, *Understanding Statistics*, London, 2002.

Jencks C., *Inequality*, New York, Harper Books, 1972.

Jöreskog K.G. et Goldberger A.S., "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable", *Journal of the American Statistical Association*, vol.70, n°351, p.631-639, 1975.

Kellerhals J. et Montandon C., *Les Stratégies éducatives des familles. Milieu social, dynamique familiale et éducation des préadolescents*, Genève, Delachaux et Niestlé, 1991.

Lahire B., *Tableaux de Famille*, Paris, Hautes Etudes, Seuil, 1995.

Lomi A., Snijders T.A.B., Steglich C.E.G., et Torló V.J., "Why Are Some More Peer Than Others? Evidence from a Longitudinal Study of Social Networks and Individual Academic Performance", *Social Science Research*, vol.40, p.1506-1520, 2011.

Peugny C., « L'expérience vécue de la mobilité sociale : le poids de la fratrie », *Informations sociales* n° 173, p. 94-101, 2012.

Singly (de) F., *Le Soi, le couple et la famille*, Paris, Nathan, 1996.

Sewell W. H., Hauser R. M., Springer K. W., and Hauser T. S., "As We Age : The Wisconsin Longitudinal Study, 1957-2001", *Research in Social Stratification and Mobility*, Vol. 20, London, Elsevier Scientific Publishers, 2004.

Sieben I., *Sibling Similarities and Social Stratification*, Université de Nimègue, 2001.

Snijders T.A.B., Bosker R., *Multilevel Analysis : An Introduction to Basic and Advanced Multilevel Modeling*, London, Sage Publishers, 1999.

Vallot F., « Résultats globaux : niveau intellectuel selon le milieu social et scolaire », in *Enquête nationale sur le niveau intellectuel des enfants d'âge scolaire*, INED-INETOP, Paris, PUF, *Travaux et Documents*, cahier n°64, 1973.

Zarca B., « L'Héritage et la mobilité sociale au sein de la fratrie I - L'héritage et la mobilité différentielle des frères », *Population*, Vol. 50, n°2, 1995.