

La mobilité géographique des enseignants-chercheurs : une analyse sur données françaises de longue période

BERNELA Bastien, BOUBA-OLGA Olivier, FERRU Marie
bastien.bernela@univ-poitiers.fr ; obouba@univ-poitiers.fr ; marie.ferru@univ-poitiers.fr

CRIEF EA 2249, UFR Sciences économiques
2 rue Jean Carbonnier, Bât.A1, BP 623, 86022 POITIERS Cedex, FRANCE

Résumé : Nous proposons dans cet article une analyse empirique de la mobilité géographique des chercheurs au début de leur trajectoire professionnelle, en nous appuyant sur une base de données de longue période (1970-2000) portant sur la France. Nous mettons d'abord en évidence la faible mobilité des chercheurs, puisque plus de 60% d'entre eux dirigent leur première thèse là où ils ont soutenu la leur. A l'aide d'une seconde base de données, nous montrons que ce niveau de mobilité n'est pas plus faible que celui des docteurs s'insérant hors université. Nous testons alors l'impact de déterminants structurels à l'aide d'un modèle gravitaire et montrons notamment l'influence positive de la taille scientifique des régions et de la proximité géographique sur la mobilité des chercheurs.

Mots-clefs : Mobilité géographique, enseignants-chercheurs, effets structurels

Abstract: The aim of this paper is to provide empirical analysis of geographical researchers' mobility during the beginning of their career, by using long term data (1970-2000) about France. First of all, we highlight the low mobility of researchers: more than 60% of them direct their first PhD where they defended their own one. Then, thanks to a second database, we show that this mobility level is not lower than non-university PhD one's. We test the impact of structural determinants with a gravity model (sample selection model) and we observe in particular the positive influence of the scientific size of regions and geographical proximity on the researchers' mobility.

Keywords: Geographical mobility, scholars, structural effects

1. Introduction

Les politiques publiques de recherche et d'innovation s'intéressent fortement à la population scientifique. Des rapports sont régulièrement commandés sur le système d'enseignement supérieur et de recherche, que ce soit au niveau français ou européen. L'Union Européenne, engagée dans la stratégie de Lisbonne et dans l'ère de la connaissance, prête une attention particulière à la mise en place d'un espace commun de recherche, et la mobilité fait partie intégrante de ces réflexions (CEC, 2001, 2008). Les travaux sur la mobilité des chercheurs sont d'autant plus importants qu'ils permettent d'avancer sur divers champs de recherche.

Parmi eux, un ensemble de travaux couple la question de la mobilité des chercheurs à celle de la diffusion spatiale des connaissances. Associés au courant de la géographie de l'innovation, ils cherchent à comprendre le rôle de la mobilité en tant que vecteur de circulation des connaissances entre les régions et les pays (Jaffe et al., 1993 ; Almeida et Kogut, 1999 ; Zellner, 2003 ; Criscuolo, 2005 ; Agrawal et al., 2006 ; Zucker et Darby, 2008 ; Breschi et Lissoni, 2009 ; Edler et al., 2011 ; Trippel, 2011 ; Franzoni et al., 2012). La mobilité des chercheurs est également abordée par la littérature florissante et controversée sur les créatifs. Si Florida (2002) considère que les chercheurs présentent, de la même manière que l'ensemble des créatifs, une mobilité géographique intense, se

déplaçant vers les villes jugées "attractives", plusieurs auteurs (Shearmur, 2007 ; Martin-Brelot et al., 2010) mettent en doute ce résultat. Un autre ensemble de travaux, principalement initié par les chercheurs américains, analyse l'impact de la mobilité sur la productivité des chercheurs et l'avancement de carrière (Allison et Long, 1987 ; Debackere et Rappa, 1995 ; Bozeman et Mangematin, 2004 ; Hoisl, 2007 ; De Filippo et al., 2009 ; Cruz-Castro et Sanz-Menendez, 2010 ; Latham et al., 2011 ; Barrufaldi et Landoni, 2012). Certains abordent alors la question de la mobilité en interrogeant les pratiques de recrutement à l'université et le fonctionnement du marché du travail des universitaires (Gaughan et Robin, 2004 ; Combes et al., 2008 ; Musselin, 2009 ; Bonnal et Giret, 2010 ; Pezzoni et al., 2012).

L'étude de la mobilité est souvent rendue difficile par le manque de données disponibles. En effet, la communauté scientifique dispose rarement de données de long terme sur la trajectoire des chercheurs, et notamment leur localisation. Lorsqu'on voit la diversité des sujets connexes à la mobilité, on comprend que ce déficit de données empêche de progresser dans la compréhension du processus de mobilité, de ses déterminants et de ses impacts. Cet article a pour vocation d'interroger empiriquement la mobilité spatiale des enseignants-chercheurs français et ses déterminants structurels, en mobilisant une base de données originale de longue période (1970-2000) disponible pour la France, qui recense les transitions entre soutenance de thèse et première direction de thèse de plus de 12 000 individus. Connaissant la localisation de la soutenance et de la direction de thèse, nous pouvons nous focaliser sur ces étapes particulières de la trajectoire professionnelle des chercheurs et nous interroger sur sa dimension géographique : dirigent-ils leur première thèse là où ils ont soutenu la leur ? Dans quelle proportion ? S'ils vont dans une autre université, où se situe-t-elle ? Peut-on identifier certains déterminants de l'éventuelle mobilité observée ? Dans cette perspective, nous insistons sur la nécessité de distinguer deux échelles d'observation – l'échelle de l'établissement et l'échelle de la région - afin de vérifier l'existence d'effets de proximité spatiale (mobilités infrarégionales). Notre article analyse également les déterminants de la mobilité interrégionale des chercheurs, pour évaluer plus précisément le jeu de certaines caractéristiques structurelles des régions : alors que de nombreux auteurs font des modalités du processus de recrutement le déterminant essentiel de la (non-)mobilité des chercheurs, nous posons l'hypothèse que la taille des régions et les distances interrégionales (géographique et disciplinaire) influent potentiellement sur les mobilités observées. Nous testons pour cela un modèle gravitaire (modèle de Heckman avec équation de sélection), dont les résultats montrent notamment l'incidence forte de la taille scientifique des régions et de la distance géographique qui les sépare sur la trajectoire spatiale des chercheurs.

Dans un souci de comparaison, nous nous sommes dotés d'une base de données complémentaire. Cette dernière est le résultat d'une enquête, menée par le CEREQ¹ en 1998, et donne de l'information sur la trajectoire sur dix ans, suite à la sortie du système scolaire, de près de 10 000 individus, tous niveaux de diplôme confondus. Cette source de données nous permet de positionner le comportement de mobilité des enseignants-chercheurs français relativement à d'autres types de population, afin de vérifier leur spécificité ou non.

L'article est organisé comme suit. Dans la section 2, nous ferons une revue des travaux empiriques sur la mobilité des chercheurs, et reviendrons notamment sur les avantages et les inconvénients des méthodes d'identification de la mobilité. Dans la section 3, nous présenterons les données mobilisées et la méthode utilisée. Dans la section 4, nous donnerons les principaux résultats. L'article conclut en s'interrogeant sur les implications de nos résultats en matière de politiques de recherche.

2. Revue de la littérature

La recension des travaux empiriques sur la mobilité des chercheurs révèle tout d'abord le faible nombre d'études portant exclusivement sur la mobilité spatiale des chercheurs et leurs déterminants.

¹ Centre d'Etudes et de Recherches sur l'Emploi et les Qualifications

Elle permet par ailleurs de repérer deux grandes méthodes d'identification de la mobilité, dont nous présentons les avantages et les limites.

Un premier ensemble de travaux empiriques étudie la mobilité des chercheurs à partir de données sur les brevets ou publications, dans la lignée de Jaffe et al. (1993). Il s'agit de regarder l'affiliation des inventeurs/auteurs aux différentes dates de dépôt du brevet/publication, et de déduire s'il y a eu mobilité ou non en comparant les localisations lors de ces événements. Cette méthode a l'avantage d'offrir un grand nombre d'informations périphériques à la mobilité du chercheur et permet d'étudier simultanément, à partir d'une seule source de données, leur trajectoire spatiale (*via* l'affiliation à un établissement), les processus collaboratifs (*via* les co-productions) et les effets de réseau (*via* les citations). Ces travaux couplent donc généralement la question de la mobilité à celle de la diffusion spatiale des connaissances (Almeida et Kogut, 1999 ; Agrawal et al., 2006 ; Breschi et Lissoni, 2009 ; Latham et al., 2011). La taille importante de ces bases de données en fait un outil de recherche très puissant.

En revanche, cette méthode a comme principal inconvénient d'être approximative et de supposer la mobilité. L'entrée de l'analyse étant le brevet ou la publication et non le chercheur lui-même, les données mobilisées ne permettent pas de représenter de façon fiable la trajectoire spatiale des chercheurs. Le dépôt de brevet ou la publication constitue un moment de la trajectoire du chercheur et permet de connaître sa localisation pour chacun des moments identifiés, sans pouvoir intégrer les changements intervenants entre ces deux moments. Ainsi, ce type de données occulte les périodes de « non-production » dans la trajectoire des individus. Par ailleurs, Katz et Martin (1997) soulignent le problème de la double affiliation institutionnelle de nombreux chercheurs : un universitaire peut être affilié à deux laboratoires de villes différentes sans que l'on ne puisse discriminer entre ces deux affiliations, l'affiliation principale n'étant pas nécessairement celle où il réside. Enfin, ces données ont l'inconvénient de ne porter que sur certains domaines scientifiques : les données sur brevets excluent les domaines dont la valorisation ne passe pas par le dépôt d'un brevet (Griliches, 1990), ce qui est principalement le cas des sciences humaines et sociales. On retrouve le même type de biais dans les études portant sur les publications : « bibliometric indicators (...) are mostly obtained from bibliographic databases that cover journal articles; thus measuring only one aspect of scientific activity. » (De Filippo et al., 2009, p.192). Repérer les chercheurs *via* leurs brevets ou publications revient à opérer un tri parmi la communauté scientifique, introduisant un biais de sélection. Les « producteurs » peuvent présenter des caractéristiques individuelles et un comportement de mobilité spécifiques, ce qui remet en cause la représentativité des résultats et leur possible systématisation.

Conscients des limites inhérentes à l'utilisation de ce premier type de données, un second groupe de travaux s'est développé (Allison et Long, 1987 ; Debackere et Rappa, 1995 ; De Filippo et al., 2009 ; Cruz-Castro et Sanz-Menendez, 2010 ; Martin-Brelot et al., 2010 ; Canibano et al., 2011 ; Edler et al., 2011). Ces travaux se concentrent sur les suivis de trajectoires de chercheurs, reconstruites à partir de CV (Dietz et al., 2000), d'enquêtes en ligne ou d'entretiens. Ils ont comme qualité intrinsèque de collecter des données suffisamment précises sur les carrières « rather than quantifying particular knowledge products at particular times » (Canibano et al., 2011, p.655) et permettent ainsi d'étudier « the researchers' career trajectories, spatial mobility and mapping of collective capacity » (ibid.). Le champ de l'étude, et plus précisément les individus ciblés, sont cette fois-ci davantage sous le contrôle du chercheur, qui va pouvoir définir la population étudiée (limite géographique et/ou disciplinaire). La création de données portant directement sur la trajectoire des chercheurs a donc l'avantage incontestable de générer des données représentatives de la réalité étudiée. Cependant, la collecte de données nécessite un investissement relativement important d'un point de vue temporel, ce qui peut restreindre la taille de l'échantillon (principalement lorsqu'il s'agit d'entretiens semi-directifs). Cette limite s'évanouit lorsque ce sont les institutions (ministères, Union Européenne) qui mettent en place des dispositifs de collecte de données, à condition que les chercheurs puissent y accéder.

Indépendamment de la méthode d'identification de la mobilité, nous adressons un ensemble de critiques aux travaux existants. Premièrement, si la mobilité est nécessairement étudiée en dynamique (localisations à un moment t et à un moment $t+1$), elle est souvent mesurée sur des données de courte

période. Notre base de données couvre trente années (1970-2000) et rend compte d'une réalité de long terme. Les résultats obtenus dépendent alors moins d'éléments de contexte. Une seconde limite des études existantes sur la mobilité est le manque de comparaison des populations : certaines recherches se concentrent sur les chercheurs universitaires, d'autres sur les chercheurs industriels, mais à notre connaissance, aucune étude ne s'attache à comparer les comportements de mobilité de ces deux catégories d'individus. Nous chercherons donc à comparer le comportement de mobilité des docteurs s'insérant dans l'enseignement supérieur et la recherche avec celui des autres docteurs. Enfin, les travaux existants se focalisent davantage sur des caractéristiques individuelles pour expliquer la mobilité et occultent les déterminants d'ordre structurel. Des modèles économétriques (type *probit* ou *logit*) testent l'impact de déterminants individuels de la mobilité des chercheurs, tels que l'âge, le sexe ou la position académique, mais l'influence de déterminants structurels apparaît trop rarement mesurée, alors même qu'elle est largement testée (et qu'elle apparaît significative) dans le cas des collaborations de recherche. Des effets liés à la taille des régions ont par exemple été soulignés par Grossetti et Nguyen (2001), Hoekman et al. (2010) et Bouba-Olga et al. (2012). De même, Maggioni et al. (2007), Ponds et al. (2007), Frenken et al. (2009) ou encore Bouba-Olga et al. (2012), montrent le rôle significatif des distances entre les territoires, en distinguant distance géographique et distance cognitive (Boschma, 2005). Dans la continuité de ces travaux, nous cherchons à savoir si la mobilité des chercheurs dépend de ces différents déterminants structurels, de la même manière que pour les collaborations de recherche : nous nous attendons à ce que la mobilité interrégionale des chercheurs dépende positivement de la taille de la région d'appartenance des individus, et négativement de la distance entre les régions. Autrement dit, un individu aurait plus de chance de se déplacer d'une région à une autre que ces deux régions sont proches géographiquement ou cognitivement. Nous pouvons également penser que plus une région est de taille importante (ie. rassemble un grand nombre de chercheurs) et plus le nombre de départs de chercheurs (ie. les mobilités) serait important.

Sur la base de cette recension de la littérature et des limites observées, nous cherchons à compléter les travaux empiriques existants, grâce i) à l'exploration de bases de données françaises de longue période (dans la lignée du groupe de travaux sur le suivi de trajectoire), ii) à la comparaison des comportements de mobilité entre différentes catégories d'individus et iii) à l'intégration, dans un modèle économétrique issu du commerce international, de différents indicateurs structurels peu testés jusqu'alors. Chacun de ces trois points sera successivement développé dans les deux sections suivantes.

3. Données et méthode

3.1. La base de données DOCTHESE : repérer la transition doctorat-direction de thèse

L'étude de la trajectoire spatiale des chercheurs est un exercice que le manque de données sur longue période rend difficile. Dans cet article, nous proposons de mobiliser une base de données originale, DOCTHESE, qui compile de manière quasi exhaustive les thèses soutenues entre 1970 et 2000 (soit plus de 200 000 thèses). Sur chaque ligne, sont renseignés les nom et prénom du docteur qui soutient sa thèse, les nom et prénom de son (ses) directeur(s) de thèse, l'année, l'université et la discipline de soutenance. Le fait d'avoir des données sur longue période permet de repérer des docteurs qui deviennent quelques années plus tard directeurs de thèse : c'est le cas pour 12 261 individus². Pour chacun d'eux, nous avons alors pu appareiller les données relatives à la soutenance de leur propre thèse avec les données relatives à la première thèse qu'ils ont dirigée. Bien que notre base de données ne nous permette pas de qualifier de façon exhaustive la trajectoire spatiale des chercheurs, on peut considérer que ce travail d'appariement relève de la méthode de suivi de trajectoire. On reconstitue, à l'aide de deux événements que sont la soutenance de thèse et la direction de thèse, une partie de la vie professionnelle d'un individu.

² Pour plus d'informations sur la construction de la base de données, cf. Annexe 1

Nous disposons ainsi d'une base de données composée de 12 261 lignes qui correspondent à une transition doctorat-direction de thèse. Sur chaque ligne, sont renseignés les nom et prénom de l'individu dont on a repéré la transition, les noms et prénoms de son directeur de thèse et de son premier doctorant, les années, universités et disciplines de sa propre thèse et de sa première thèse dirigée. Nous déduisons de l'université de soutenance et de première direction, la région de soutenance et de première direction, ce qui permet d'observer leur éventuelle mobilité à différentes échelles. Notons que la durée d'accession à la direction de thèse est de onze ans en moyenne après l'obtention du grade de docteur.

Godechot et Louvet (2008) ont mobilisé cette base pour mesurer le poids du recrutement local dans les universités françaises, qu'ils définissent comme « le fait de soutenir et de diriger sa première thèse (...) dans la même université et dans la même discipline » (p. 9). Ils montrent son importance quantitative, avec des variations non négligeables selon les disciplines et les régions, les universités parisiennes étant significativement moins « localistes ». Face à de tels résultats, ils plaident pour une réforme du processus de recrutement en France, avec comme objectif principal l'interdiction du recrutement local. L'analyse proposée par ces auteurs souffre cependant de plusieurs limites, soulignées par Bouba-Olga et al. (2008)³. Se prononcer sur la question du recrutement dans l'enseignement supérieur à partir de ces données n'est pas adapté, car aucune variable ne permet de saisir la nature des recrutements. Nous nous focaliserons donc sur la problématique de la mobilité géographique des chercheurs, au début de leur trajectoire professionnelle.

Pour caractériser cette mobilité, nous proposons de regarder, pour chaque région française, la destination des docteurs qui ont soutenu leur thèse dans cette région : combien d'entre eux dirigent leur première thèse dans la même région, en Ile-de-France, ou encore dans une région limitrophe ? Lorsqu'il n'y a pas mobilité, nous distinguons deux échelles d'observations :

- celle de l'établissement : l'individu soutient sa première thèse et dirige sa première thèse dans le même établissement, c'est-à-dire dans la même université ;
- celle de la région : l'individu soutient sa première thèse et dirige sa première thèse dans la même région.

Nous pensons que cette distinction est essentielle puisqu'elle permet de mettre en évidence des effets de proximité spatiale. En effet, dans les régions composées d'un grand nombre d'universités, on peut avoir simultanément des niveaux élevés de mobilité à l'échelle des établissements, mais faibles au niveau régional. Ce type de configuration révèle l'existence de mobilités infrarégionales : les docteurs d'une université dirigent leur première thèse dans une autre université de la même région. Nous verrons que cette configuration correspond typiquement au cas de l'Ile de France.

Pour étudier la dimension spatiale du système universitaire, une autre échelle possible est celle des académies. Le découpage académique du territoire français correspond au découpage régional, hormis trois cas : la région Ile-de-France est divisée en trois académies (Créteil, Paris et Versailles), la région Provence-Alpes-Côte-d'Azur en deux académies (Marseille et Nice) et la région Rhône-Alpes en deux académies (Grenoble et Lyon). Pour ces deux dernières régions, l'échelle de l'académie nous paraît plus adaptée dans la mesure où il existe deux villes universitaires éloignées au sein de la même région, l'étude des mobilités entre académies ne doit pas être occultée. A l'inverse, la région Ile-de-France est caractérisée par une forte concentration d'établissements dans un même continuum géographique : migrer d'une académie à l'autre n'implique pas de changement d'agglomération et ne constitue pas une mobilité spatialement significative. Autrement dit, dans la suite de l'article, le terme de région réfère à la nomenclature des régions administratives, à l'exception de Rhône-Alpes et Provence-Alpes-Côte-d'Azur pour lesquelles nous retenons l'échelle académique.

Il s'agit de mettre en évidence la matrice des flux de chercheurs français observés *via* les données DOCTHESE, traitement qui à notre connaissance, n'a pas été réalisé parmi les travaux existants sur la

³ Les échanges entre ces auteurs sont visibles sur le site de la Vie des Idées : <http://www.laviedesidees.fr/Le-localisme-universitaire,1566.html>.

mobilité des chercheurs. On peut, à partir de cette matrice, dégager de premiers enseignements sur le niveau de mobilité des chercheurs et le déploiement de ces mobilités sur le territoire français.

3.2. La base de données CEREQ : comparer le niveau de mobilité des docteurs

La base de données DOCTHESE souffre de deux limites principales, que nous proposons de dépasser en mobilisant une base de données complémentaire.

i) La trajectoire des chercheurs est reconstituée à partir de l'accès à la direction de thèse. Nous ne savons donc pas ce qui se passe entre la soutenance et la direction de thèse, période de onze ans en moyenne. Autrement dit, nous n'avons pas d'informations sur les dix premières années qui suivent l'obtention du doctorat. Ne sachant pas quand les mobilités interviennent le plus souvent dans la vie des individus, nous ne pouvons prétendre que la mobilité repérée est représentative de la trajectoire de l'individu.

ii) *Via* DOCTHESE, nous ne touchons que des docteurs qui accèdent à la direction. Ceux qui n'ont pas fait carrière à l'université, qui n'ont pas voulu diriger de thèse, ou qui n'ont pas encore obtenu leur habilitation à diriger des recherches échappent à cette base de données. Nous ne touchons donc qu'une partie des docteurs. De plus, et comme nous l'avons souligné plus haut, les études portant sur la mobilité souffrent bien souvent de données comparatives entre groupes d'individus. Or, lorsque l'on étudie la mobilité géographique d'une population donnée, on peut s'interroger sur la spécificité de son comportement. Il nous a donc semblé important d'apporter dans cet article des éléments de comparaison entre la mobilité des docteurs et celle de populations proches. Plusieurs auteurs (Greenwood, 1975 ; Schwartz, 1976 ; Yankow, 2003) mettent en évidence la relation positive entre niveau de diplôme et mobilité des individus : plus un individu est diplômé, plus il a tendance à être mobile. Nous avons donc voulu vérifier cette hypothèse, selon laquelle les docteurs seraient les individus les plus mobiles. Deuxièmement, nous avons voulu établir une distinction essentielle au sein des docteurs entre ceux qui s'engagent dans l'enseignement supérieur et la recherche et les autres. A niveau de diplôme équivalent, le monde professionnel dans lequel on s'insère impacte-t-il significativement les degrés de mobilité ?

Dans cette perspective, nous nous sommes dotés d'une base de données construite par le CEREQ dans le cadre des « enquêtes génération », qui proposent de suivre une cohorte de diplômés de 1998 durant les dix premières années de leur vie professionnelle. Ce suivi sur dix ans permet d'avoir de l'information sur la période manquante avec les données DOCTHESE. Autrement dit, on peut considérer qu'il correspond à la transition entre doctorat et direction de thèse (onze ans en moyenne).

Au sein de la base de données CEREQ, composée de 9121 individus, nous définissons quatre types de populations : les individus de diplôme inférieur ou égal au baccalauréat (qui représentent 49.5% des individus), ceux de diplôme compris entre le baccalauréat et le master (40.7%), les individus titulaires d'un master (dits Masters) (6.8%) et les docteurs (3%). Au sein de cette dernière catégorie, nous distinguons les docteurs s'insérant dans le monde académique (1%) des docteurs s'insérant hors monde académique (2%). Dans la suite de l'article, nous parlerons de docteurs ESR et de docteurs non ESR. Pour chaque individu, nous disposons d'informations sur sa localisation (à l'échelle régionale), permettant ainsi de reconstituer sa trajectoire spatiale : lieu d'obtention du diplôme en 1998, puis lieux de travail en 2001 et en 2008. A partir de ces informations, nous construisons deux variables binaires permettant de capturer la trajectoire spatiale des individus :

- Mob98-01=0 si l'individu est localisé dans la même région en 1998 et en 2001, Mob98-01=1 sinon ;
- Mob98-08=0 si l'individu n'a connu aucune mobilité entre 1998 et 2008, Mob98-08=1 s'il a connu au moins une mobilité.

Nous estimons alors la probabilité d'observer une mobilité entre 1998 et 2001 (Mob98-01) puis entre 1998 et 2008 (Mob98-08) à partir d'un modèle *probit* simple. Les variables d'intérêt pour expliquer les mobilités sont le niveau de diplôme et le monde professionnel des docteurs (ESR vs. non ESR), décrits dans le paragraphe ci-dessus. Nous intégrons également trois variables de contrôle d'ordre

sociologique (âge, genre et enfant) et des dummies régionales (région de formation en 1998). Les variables de contrôle sont des variables binaires construites de la façon suivante :

- La variable « âge »⁴ prend la valeur 1 quand l'individu est plus âgé que la moyenne de son groupe de niveau, et 0 sinon.
- La variable « homme » prend la valeur 1 quand l'individu est un homme, et 0 sinon.
- La variable « enfant » prend la valeur 1 si l'individu a eu un enfant avant l'obtention du diplôme, et 0 sinon.

Nous montrons que les docteurs ESR ne sont pas moins mobiles que les Masters et les docteurs non ESR et qu'il semble donc difficile d'imputer la faible mobilité des chercheurs à l'organisation du système universitaire. Ce résultat vient justifier l'intérêt de ne pas nous focaliser sur les procédures de recrutement, mais d'expliquer la mobilité spatiale par des déterminants d'ordre structurels, peu étudiés dans la littérature.

3.3. Tester l'impact de déterminants d'ordre structurel sur la mobilité des universitaires

Après s'être doté d'un point de comparaison avec les données CEREQ, nous revenons aux données DOCTHESE pour regarder si les comportements de mobilité répondent à des variables structurelles. Nous cherchons à expliquer la probabilité d'observer des mobilités entre deux régions puis le nombre de mobilités entre ces couples de régions. Pour cela, nous testons un modèle gravitaire avec équation de sélection (Heckman, 1979) afin d'évaluer l'existence d'effets de taille et de distance dans les choix de mobilité des chercheurs, en nous focalisant uniquement sur les mobilités interrégionales. Les individus qui ont soutenu et dirigé dans la même région sont donc exclus du modèle. Le modèle testé, communément utilisé dans les recherches sur le commerce international, a récemment été mobilisé pour expliquer la géographie des collaborations (Bouba-Olga et al., 2012 ; Ponds et al., 2007).

Plus précisément, nous définissons une typologie de régions telles que $r1$ correspond aux régions de soutenance et $r2$ aux régions de direction. Le couple $r1/r2$ ne correspond donc pas au couple $r2/r1$: il existe un sens de la mobilité, de la région de départ vers la région d'arrivée. Rappelons que les couples pour lesquels $r1 = r2$ sont exclus du modèle. Nous supposons d'abord que l'existence d'une mobilité pour chaque paire (i) de régions $r1$ et $r2$, dépend d'une première variable latente, non observable et continue, combinaison linéaire d'un vecteur de variables exogènes Z_i qui sont ici la taille scientifique de la région de soutenance de thèse et le différentiel de taille avec la région de première direction, d'une part, et de différentes formes de distance (spatiale et profil disciplinaire), d'autre part :

$$d_i^* = Z_i\gamma + \mu_{i1}$$

Quand la variable d_i^* est supérieure à un certain seuil, on observe au moins une mobilité entre les deux régions et lorsqu'elle reste inférieure à cette limite, on n'observe pas de mobilité. D'où :

$$d_i = \begin{cases} 1 & \text{si } d_i^* \geq 0 \\ 0 & \text{si } d_i^* < 0 \end{cases}$$

La variable d_i nous permet de voir quand il est possible d'observer le nombre de mobilité entre deux régions, nombre représenté par une deuxième variable latente m_i^* expliqué par les mêmes variables exogènes, auxquelles nous ajoutons une variable notée « Ile-de-France » pour capturer l'effet de la région capitale :

⁴ Concernant l'effet d'âge, l'âge absolu est colinéaire avec le niveau de diplôme, nous ne pouvons donc pas l'intégrer en tant que tel dans le modèle. C'est pourquoi nous testons un effet d'âge relatif : indépendamment de l'effet de diplôme, le fait d'être plus ou moins âgé que la moyenne de son groupe impacte-t-il significativement la probabilité d'être mobile ?

$$m_i^* = W_i \theta + \mu_{i2}$$

Le nombre de mobilité m_i pour la paire i de régions considérées est obtenu comme suit :

$$m_i = \begin{cases} m_i^* & \text{si } d_i^* \geq 0 \\ 0 & \text{si } d_i^* < 0 \end{cases}$$

Nous supposons que les termes d'erreurs μ_{i1} et μ_{i2} sont normaux.

Concernant les variables explicatives du modèle que sont les indicateurs de taille des régions et de distance entre les régions, nous les avons en partie construits à partir de la base de données DOCTHESE.

La taille scientifique des régions

Pour approximer la taille scientifique des régions, nous avons construit un indicateur à partir de la répartition régionale des thèses référencées dans la base DOCTHESE d'origine : il s'agit en d'autres termes de la contribution de chaque région à la production nationale de docteurs⁵. Nous considérons donc qu'une région est d'autant plus grande scientifiquement qu'un grand nombre de docteurs ont soutenu dans cette région. Dans le modèle, nous intégrons la variable de taille scientifique de la région de départ, afin de savoir si le fait d'avoir soutenu dans une grande région impacte significativement la probabilité d'être mobile. Nous testons également l'effet du différentiel de taille entre les régions⁶ pour déterminer l'éventuelle tendance des docteurs à aller d'une petite région vers une grande région, ou l'inverse.

La distance spatiale entre les régions

Pour vérifier l'existence éventuelle d'effets de distance spatiale, nous intégrons une matrice des distances interrégionales, mesurée en heures de train entre les capitales régionales.

Pour cette variable, nous avons testé plusieurs indicateurs : temps de train entre les capitales régionales, distance en kilomètres entre les capitales régionales et nombre de frontières à traverser pour atteindre l'autre région. La corrélation étant très forte (tau de Kendall) entre tous ces indicateurs, nous avons retenu la première variable comme indicateur de distance spatiale interrégionale.

La distance en termes de profil disciplinaire des régions

Cette variable permet de tester l'existence d'un effet lié à la ressemblance ou non des régions en termes de structure disciplinaire. Elle a été créée à partir des données d'origine relatives aux disciplines⁷. D'abord, nous avons calculé, pour chaque région i , la part p que représente chacune des disciplines j ⁸. L'indicateur de distance des profils disciplinaires entre les régions 1 et 2 est calculé comme suit :

$$\text{Distance des profils disciplinaires (r1; r2)} = \frac{\sum_{j=1}^{19} |p_{r1,j} - p_{r2,j}|}{2}$$

Cet indicateur est construit de telle façon que, plus il est proche de 1, plus les profils disciplinaires des deux régions considérées sont différents. Ce que nous approximons *via* cet indicateur est l'impact des structures disciplinaires sur les mobilités observées entre deux régions : est-ce que les docteurs migrent davantage vers des régions disciplinairement proches ? On peut penser que cet élément joue

⁵ Pour plus d'informations sur la répartition régionale des docteurs, cf. Annexe 2.

⁶ Il s'agit plus précisément du rapport entre la taille de la région de départ et celle d'arrivée. Nos variables étant exprimées en logarithme népérien, le différentiel de taille entre les régions $r1$ et $r2$ correspond à la différence entre $\ln(\text{taille de } r1)$ et $\ln(\text{taille de } r2)$. Un différentiel positif signifie donc que le docteur migre vers une région plus petite.

⁷ Pour plus d'informations sur la répartition des docteurs par discipline, cf. Annexe 3.

⁸ Par exemple, la part de docteurs en économie (discipline 10) dans la région Poitou-Charentes (région 20) est nommée $p_{20,10}$.

un certain rôle, puisque si une discipline pèse fort dans une région, l'offre de postes dans cette discipline est élevée. Les docteurs peuvent donc avoir plus de chances d'être recrutés dans une région spécialisée dans leur discipline.

Afin de dégager des résultats par disciplines, il semble pertinent de décliner ce modèle gravitaire. Tester un modèle gravitaire par cohorte de docteurs pour chaque discipline n'est pas possible étant donné le faible nombre d'observations, c'est pourquoi nous procédons à un regroupement en deux groupes : les sciences fondamentales et appliquées et les sciences humaines et sociales. Pour vérifier si cette décomposition améliore la qualité du modèle, nous avons réalisé un test de Chow, dit de stabilité des coefficients. À partir de régressions simples entre la variable expliquée et les variables explicatives du modèle, le test de Chow montre qu'il existe une différence significative entre la somme des carrés des résidus de l'ensemble de la population et l'addition de la somme des carrés des résidus calculée à partir des deux sous-populations, ce qui nous conduit à tester un modèle gravitaire pour les deux groupes de disciplines.

Toutes les variables sont exprimées en logarithme népérien.

4. Résultats

4.1. La « faible » mobilité des enseignants-chercheurs français

Nous proposons d'analyser les mobilités observées entre les régions françaises. Pour cela, nous avons construit une matrice des flux afin de caractériser les tendances de mobilité (ou de non-mobilité) des enseignants-chercheurs français (tableau 1). En moyenne, 64.0% restent dans leur région de soutenance, 8.9% partent vers l'Ile-de-France, 7.8% vers des régions limitrophes et 19.3% vers des régions non limitrophes. Autrement dit, le taux de mobilité interrégionale des docteurs qui deviennent directeur de thèse est de seulement 36.0%. On descend à moins d'un tiers des individus si on enlève les mobilités vers les régions limitrophes.

La différence entre la première et la deuxième colonne correspond au taux de mobilité inter-établissements au sein d'une même région. Cette différence est logiquement nulle pour les régions dotées d'une seule université. On constate que la variation entre les deux mesures est particulièrement sensible. Le taux passe de 43.9% en moyenne pour l'ensemble des universités françaises à 64.0% à l'échelle régionale. Autrement dit, un cinquième des mobilités observées dans la base, soient 2 465 individus, se font entre deux universités d'une même région.

L'Ile-de-France est principalement concernée par ces mobilités infrarégionales : seulement 29.6% des docteurs restent dans le même établissement (score très largement inférieur à celui des autres régions) alors même que 61.2% restent en région (score dans la moyenne des autres régions). Ce résultat met en évidence l'existence d'une forte circulation des docteurs au sein de la région capitale : si les docteurs parisiens dirigent peu souvent leur première thèse dans l'université dans laquelle ils ont soutenu leur propre thèse, une grande partie d'entre eux reste tout de même en région Ile-de-France.

L'Ile-de-France exerce un fort pouvoir d'attractivité, plus marqué chez les régions limitrophes, dont les docteurs sont « aspirés » par la capitale : c'est particulièrement le cas pour les régions Centre et Champagne-Ardenne, dont le taux de mobilité régionale est supérieur à la moyenne des autres régions. On peut qualifier ce phénomène d'effet d'ombre de la capitale (Brouillat et Lung, 2010), ce qui laisse penser que la distance spatiale explique partiellement les mobilités interrégionales.

Tableau 1 : de la région de soutenance à la région de direction de thèse

	Même Région	<i>Dont : Même université</i>	<i>Dont : Autre université de la région</i>	Ile-de-France	Régions limitrophes (hors IDF)	Autres Régions	Total général
Académie de Grenoble	75.6%	53.9%	21.7%	7.7%	4.4%	12.3%	100.0%
Académie de Lyon	70.3%	50.7%	19.6%	8.5%	8.9%	12.3%	100.0%
Académie de Marseille	68.5%	53.4%	15.1%	6.8%	10.3%	14.4%	100.0%
Académie de Nice	64.8%	57.1%	7.7%	10.0%	7.8%	17.4%	100.0%
Alsace	63.9%	60.0%	3.9%	9.7%	6.5%	19.9%	100.0%
Aquitaine	60.3%	53.6%	6.7%	8.0%	6.9%	24.8%	100.0%
Auvergne	57.0%	57.0%	-	10.3%	23.0%	9.7%	100.0%
Bourgogne *	65.0%	65.0%	-	7.9%	11.5%	15.6%	100.0%
Bretagne	64.3%	58.0%	6.3%	12.3%	6.6%	16.8%	100.0%
Centre *	48.5%	45.5%	3.0%	17.9%	10.4%	23.2%	100.0%
Champagne-Ardenne *	43.5%	43.5%	-	22.2%	13.3%	21.0%	100.0%
Franche-Comté	50.9%	50.9%	-	10.2%	24.1%	14.8%	100.0%
Ile-de-France	61.2%	29.6%	31.6%	-	5.9%	32.9%	100.0%
Languedoc-Roussillon	60.3%	54.0%	6.3%	10.5%	16.0%	13.3%	100.0%
Limousin	67.9%	67.9%	-	1.8%	12.5%	17.8%	100.0%
Lorraine	62.5%	44.4%	18.1%	8.2%	6.0%	23.2%	100.0%
Midi-Pyrénées	70.6%	59.9%	10.7%	8.5%	6.8%	14.0%	100.0%
Nord-Pas-De-Calais	73.7%	65.5%	8.2%	6.7%	1.7%	17.9%	100.0%
Basse-Normandie	52.9%	52.9%	-	8.8%	21.6%	16.7%	100.0%
Haute-Normandie *	74.2%	69.7%	4.5%	7.9%	0.0%	17.9%	100.0%
Pays-De-Loire	67.8%	67.1%	0.7%	9.2%	7.2%	15.8%	100.0%
Picardie *	68.9%	66.7%	2.2%	15.6%	2.2%	13.3%	100.0%
Poitou-Charentes	65.8%	65.8%	-	7.7%	10.7%	15.8%	100.0%
France	64.0%	43.9%	20.1%	8.9%	7.8%	19.3%	100.00%

* régions limitrophes à l'Ile-de-France. NB : Lorsqu'il n'y a pas de valeur dans la troisième colonne, cela signifie que la région ne comporte qu'une seule université. Les taux inscrits dans les première et seconde colonnes sont donc identiques.

4.2. Analyse comparée de la trajectoire des docteurs

Afin de se doter d'un point de comparaison, nous avons mobilisé une base de données complémentaire pour positionner le comportement de mobilité des docteurs par rapport à d'autres populations. Avant de présenter les résultats du *probit*, quelques statistiques descriptives sont données dans le tableau 2.

Tableau 2 : Niveau de diplôme, âge et mobilité

	Nb. Obs.	Age moyen (années)	Mob98-01 (%)	Mob98-08 (%)
≤ Baccalauréat	4516	20.3	16.1	23.8
Baccalauréat < niveau < Master	3710	23.1	24.5	32.0
Master	622	24.6	45.7	51.8
Doctorat	273	28.7	33.0	39.6
Docteur ESR	81	29.0	38.3	40.7
Docteur non ESR	192	28.5	30.7	39.1
Total	9 121	22.0	22.0	29.5

Les résultats obtenus avec la base de données CEREQ confirment ceux de la section précédente. A partir des données DOCTHESE, on obtenait un taux de mobilité de 36.0%. A horizon de trois ans, il est ici de 33.0% pour les docteurs, et 38.3% plus précisément pour les docteurs ESR. La mobilisation de cette seconde base de données vient donc valider nos premiers résultats.

Globalement, il apparaît que les individus de niveau inférieur au baccalauréat sont les moins mobiles, puisque seul un cinquième change de région durant les trois années suivant l'obtention du diplôme. Les Masters sont la catégorie la plus mobile avec un taux de mobilité d'environ 50%. Les docteurs sont quant à eux moins mobiles que les Masters, ce qui vient infirmer l'existence d'une relation linéaire et positive entre niveau de diplôme et mobilité. En effet, l'âge influe positivement sur le niveau de mobilité jusqu'au master, au-delà duquel il influe négativement sur le niveau de mobilité. Ce résultat peut s'expliquer par des effets d'ordre sociologique liés aux trajectoires individuelles (travail du conjoint, enfants scolarisés, etc.). A partir d'un certain âge (moyenne de 28.7 ans en fin de doctorat), les individus peuvent être plus fortement ancrés dans leur région et le comportement de mobilité devient moins automatique.

Les écarts entre les taux de mobilité 1998-2001 et 1998-2008 sont relativement faibles, ce qui signifie que la majorité des mobilités s'effectue en début de carrière. Plus précisément, sur les 2694 personnes ayant connu au moins une mobilité entre 1998 et 2008 (soit 29.5% de l'ensemble de la base), 2009 personnes l'ont connu entre 1998 et 2001, soit près de 75%. Ce constat vient renforcer la validité de la méthode utilisée pour DOCTHESE, puisque la trajectoire sur dix ans est relativement stable, une fois passé le premier recrutement. C'est d'ailleurs pour les docteurs ESR que cette tendance est la plus forte, avec un écart de seulement 2.4 points entre les deux taux de mobilité. Ce résultat peut s'expliquer par la stabilité du premier emploi des docteurs ESR, à savoir les postes de maître de conférences et ingénieur/chargé de recherche, qui engagent souvent l'individu sur plusieurs dizaines d'années. En revanche, l'écart le plus élevé entre les deux taux de mobilité concerne les docteurs non ESR : 8.4% d'entre eux connaissent une mobilité entre 2001 et 2008.

Le tableau 3 présente les résultats des modèles *probit*, la partie a donnant les résultats pour la mobilité à trois ans, et la partie b à dix ans. Des statistiques descriptives sur les variables explicatives sont données en Annexe 4.

Tableau 3 : probabilité de mobilité (effets marginaux)

	Mob98-01		Mob98-08	
	Modèle 1a	Modèle 2a	Modèle 1b	Modèle 2b
≤ Baccalauréat	-27.4 ***	-27.4 ***	-27.8 ***	-27.8 ***
> Baccalauréat & < Master	-17.1 ***	-17.1 ***	-17.7 ***	-17.7 ***

Master	ref.		ref.		ref.		ref.
Doctorat	-6.6	***			-7.2	**	
Doctorat ESR			-2.3				-5.7
Doctorat Non ESR			-8.3	***			-7.9 **
Age	+1.4		+1.4		+0.6		+0.6
Homme	+3.6	***	+3.6	***	+5.9	***	+5.9 ***
Enfant	-10.4	***	-10.4	***	-12.2	***	-12.2 ***

Note : *P<0.1 **P<0.05 ***P<0.01

Lecture : Le fait d'avoir un doctorat plutôt qu'un master diminue de 6.6 points de pourcentage la probabilité d'être localisé dans la même région en 1998 et 2001, étant entendu que la part des Masters non mobiles est de 54.3% en 2001 et de 48.2% en 2008.

Si la mobilité est logiquement plus forte pour les Masters que pour les titulaires d'un diplôme inférieur au master, elle est aussi plus forte que pour ceux titulaires d'un doctorat, à court ou à long terme (modèle 1a). S'agissant des effets d'âge, le fait d'être plus âgé que la moyenne de son groupe de niveau n'a pas d'effet significatif sur la mobilité relativement à ceux qui sont moins âgés. Comme nous l'avons vu plus haut, l'âge, fortement corrélé au niveau de diplôme, impacte la mobilité, mais l'écart à l'âge moyen n'a pas d'effet. En revanche, le genre a un effet significatif au seuil de 1%, les femmes étant moins mobiles que les hommes. Le fait d'avoir eu un enfant avant l'obtention du diplôme diminue également significativement la probabilité d'être mobile. Ces résultats confirment l'idée selon laquelle la mobilité est dépendante de la temporalité de la trajectoire et qu'elle s'explique en partie par des déterminants d'ordre sociologique (Bonney et Love, 1991 ; Shauman et Xie, 1996).

La décomposition des docteurs dans le modèle 2a indique que les docteurs ESR ne sont pas significativement moins mobiles que les Masters, contrairement aux docteurs non ESR. L'insertion des docteurs hors université diminue donc la probabilité de changer de région entre l'obtention du diplôme en 1998 et le lieu de travail en 2001, relativement aux docteurs ESR.

En comparant les mobilités à trois ans et à dix ans, on observe très peu de différences, résultat peu surprenant du fait que les mobilités interviennent essentiellement en début de vie professionnelle. Les docteurs non ESR ont une mobilité significativement plus faible que les Masters, aux deux horizons temporels.

Au terme de ces traitements CEREQ, le résultat le plus surprenant est la plus faible mobilité des docteurs non ESR relativement aux docteurs ESR. Pour l'expliquer, nous proposons plusieurs hypothèses interprétatives. Les postes se font plus rares à l'université qu'ailleurs : pour un individu ne souhaitant pas faire carrière dans l'ESR, il est donc plus facile de trouver un emploi localement. En effet, pour être recruté dans l'université de soutenance, il faut d'abord qu'un poste soit ouvert l'année qui suit la thèse. Les logiques de recrutement hors université (recrutement dans le privé, création d'entreprise) dépendent moins de contraintes de calendrier : elles offrent davantage de souplesse et permettent aux docteurs non ESR de trouver plus facilement un travail en région. De plus, le recrutement local peut simplement être interdit dans certaines universités, impliquant un changement d'établissement si le docteur souhaite continuer dans l'ESR. Enfin, avant d'accéder à un poste académique permanent, notamment celui de maître de conférences, de plus en plus de docteurs passent par des statuts transitoires, comme le post-doctorat. Ces transitions ne se font pas forcément localement et génèrent donc des mobilités.

4.3. Les déterminants de la mobilité interrégionale des universitaires

Dans cette partie, les estimations ont été obtenues à partir d'un modèle de correction de biais d'Heckman en deux étapes (Heckman, 1979). Les tableaux 4 et 5 présentent les résultats de l'équation de sélection, obtenue à l'aide d'un modèle probit (étape 1), et de l'équation d'output des modèles gravitaires réalisés par groupe de disciplines (étape 2). On dénombre 145 couples de régions pour

lesquels on n'observe pas de mobilité interrégionale en sciences fondamentales et appliquées, et 199 en sciences humaines et sociales. La significativité du ratio de Mills au seuil de 5% indique la présence d'un biais de sélection, contrôlé par notre modèle. Des statistiques descriptives sur les variables explicatives sont données en Annexe 4.

Tableau 4 : déterminants de la mobilité interrégionale des chercheurs

Equation de sélection	Mobilités en sciences fondamentales et appliquées			Mobilités en sciences humaines et sociales		
	n=506			n=506		
	Coefficient	Ecart-type		Coefficient	Ecart-type	
Taille scientifique région de départ	1.235	0.143	***	1.378	0.143	***
Différentiel de taille scientifique	-0.454	0.083	***	-0.497	0.081	***
Distance spatiale	-0.443	0.191	**	-0.509	0.177	***
Distance des profils disciplinaires	-0.221	0.259		0.146	0.244	
constante	-0.196	0.476		0.333	0.438	
Equation d'output	n=361			n=307		
	Coefficient	Ecart-type		Coefficient	Ecart-type	
Taille scientifique région départ	1.832	0.367	***	2.072	0.583	***
Différentiel de taille scientifique	-0.730	0.162	***	-0.746	0.256	***
Distance spatiale	-0.710	0.206	***	-0.947	0.304	***
Distance des profils disciplinaires	-0.115	0.286		0.138	0.435	
Ile-de-France	-0,297	0.446		0.060	0.635	
constante	-0.311	0.681		-0.726	0.975	
Terme d'erreur	Coefficient	Ecart-type		Coefficient	Ecart-type	
Ratio de mills	1.485	0.643	**	2.094	0.822	**

Plusieurs résultats méritent d'être soulignés. S'agissant de l'équation de sélection, d'abord, c'est-à-dire l'estimation de la probabilité d'observer des mobilités entre deux régions, on remarque l'existence i) d'un effet taille des régions d'une part, et ii) d'un effet de distance géographique d'autre part. La taille scientifique de la région de départ influe positivement sur la probabilité estimée. Plus une région produit de docteurs, plus on s'attend à ce que certains d'entre eux migrent vers d'autres régions. Le différentiel de taille entre région de départ et région d'arrivée est également significatif au seuil de 1% : la probabilité d'observer des mobilités entre deux régions est d'autant plus grande que la région de départ est plus petite que la région d'arrivée. Le fait que deux régions soient proches géographiquement augmente également la probabilité que des docteurs migrent d'une région à l'autre.

S'agissant de l'équation d'output, qui estime le nombre de mobilités entre deux régions, on retrouve l'influence significative au seuil de 1% de la taille scientifique de la région de départ. Le différentiel de taille entre région de départ et région d'arrivée est toujours significatif au seuil de 1% : on observe d'autant plus de mobilités entre deux régions que la région de départ est plus petite que la région d'arrivée. La distance joue un rôle très significatif sur le nombre de mobilités observées, qui sera d'autant plus élevé que les régions sont proches. Ce résultat laisse penser que la proximité spatiale est un déterminant essentiel des choix de mobilité des chercheurs, les chercheurs se déplaçant vers des régions proches de celle où ils ont soutenu leur thèse. Le travail réalisé vient montrer le rôle de la proximité spatiale comme déterminant du comportement de mobilité des chercheurs. Notons que la distance des profils disciplinaires ne joue toujours pas. La variable muette « Ile-de-France » n'a pas non plus d'effet significatif sur le nombre de mobilités observées entre deux régions.

On observe très peu de différences entre les résultats des deux modèles. Il semble donc que les éléments d'explication des mobilités ne diffèrent pas d'un groupe de disciplines à l'autre. Cependant, on peut signaler que, si les variables explicatives impactent la mobilité dans le même sens dans les

deux groupes de disciplines, les coefficients sont plus forts pour le modèle sciences humaines et sociales : les variables testées ont donc un impact plus fort dans ces disciplines.

5. Conclusion/Discussion

Cet article met en évidence le faible niveau de mobilité des enseignants-chercheurs français. En identifiant la mobilité à partir de deux événements de leur trajectoire, que sont les soutenances de leur propre thèse et de celle de leur premier doctorant, nous montrons que seuls 36% des individus ont changé de région. Les résultats diffèrent peu entre les régions. De plus, il apparaît clairement que cette faible mobilité n'est pas spécifique aux docteurs, et encore moins au monde de l'enseignement supérieur et de la recherche. A niveau de diplôme équivalent, les docteurs continuant leur carrière à l'université présentent un degré de mobilité plus fort que les autres docteurs.

Pour comprendre ces comportements de mobilité, nous avons testé l'impact de plusieurs déterminants. Les déterminants d'ordre sociologique ont une incidence forte sur la mobilité des individus : l'effet d'âge et la temporalité de la trajectoire personnelle impactent directement la probabilité d'observer une mobilité, tout comme les contraintes familiales. Un deuxième ensemble de déterminants, d'ordre structurel, vient expliquer les flux interrégionaux. La taille scientifique des régions, d'abord, augmente significativement la probabilité d'observer des mobilités. La proximité spatiale, ensuite, structure sensiblement la carte des mobilités. Les chercheurs semblent intégrer la distance dans leurs arbitrages, se déplaçant ainsi vers des régions proches de leur région de soutenance. Cet argument peut rejoindre le volet sociologique expliquant le système des mobilités. Non seulement, à un certain âge, l'ancrage territorial semble limiter les comportements de mobilité, mais les individus, parfois contraints de migrer pour obtenir un poste, cherchent à minimiser la distance qui les sépare de leur environnement personnel.

Les conclusions de notre article peuvent alimenter deux débats très actuels. Le premier concerne les travaux de Richard Florida sur la classe créative. Cette dernière, qui inclut les scientifiques, est présentée comme une valeur ajoutée stratégique pour les territoires, qui ont intérêt à les attirer. Pour cela, ils doivent développer les *soft factors* auxquels les créatifs seraient particulièrement sensibles. La mobilité spatiale de ces créatifs devient un paramètre essentiel, puisque du degré et des déterminants de leur mobilité dépend la réussite des stratégies de développement économique local qui misent sur l'attractivité de cette catégorie de population. Au final, la mobilité des chercheurs se révèle plutôt faible et, lorsqu'on l'observe, elle s'explique en partie par des déterminants structurels et individuels, sur lesquels les politiques locales n'ont que peu de prise, rejoignant ainsi les résultats de Martin-Brelot et al. (2010). A l'aune de ces résultats, la stratégie d'attractivité vis-à-vis des créatifs peut sans doute être réinterrogée.

Le second débat est relatif à la problématique du localisme universitaire. Elle est particulièrement prégnante dans le paysage académique français (Godechot et Louvet, 2008 ; Bouba-Olga et al., 2008), et nombreux sont ceux qui font du système de recrutement à l'œuvre le principal responsable de la faible mobilité des chercheurs, venant perturber le *matching* naturel sur le marché du travail académique. Indépendamment de cette question, nous montrons que ce sont des déterminants structurels et individuels qui expliquent les choix de mobilité, sans oublier que la faible mobilité n'est pas spécifique à l'ESR.

Enfin, la question du localisme soulève celle de la diffusion spatiale des connaissances, puisque ceux qui dénoncent les recrutements locaux mettent en avant le risque de sclérose scientifique. Conclure à l'inertie cognitive sur la base de la faible mobilité spatiale des chercheurs est cependant excessif : la mobilité géographique d'un chercheur peut être temporaire, *via* l'inscription dans des réseaux de recherche, la participation à des colloques, l'implication dans des programmes de recherche collaboratifs, la réalisation de séjours dans d'autres universités, etc. Dans le prolongement de notre travail, il serait donc intéressant d'étudier les mobilités temporaires des chercheurs, difficilement observables (cf. Rallet et Torre, 2005, sur la notion de proximité temporaire ; Canibano et al. 2011 ; Barrufaldi et Landoni, 2012), et d'analyser leur impact sur les processus collaboratifs.

Bibliographie

- Agrawal, A., Cockburn, I., Michale, J., 2006. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships, *Journal of Economic Geography*, 6(5), 571-591.
- Allison, P.D., Long, J.S., 1987. Interuniversity mobility of academic scientists, *American Sociological Review*, 52(5), 643-652.
- Almeida, P., Kogut, B., 1999. Localization of knowledge and the mobility of engineers in regional networks, *Management Science*, 45(7), 905-917.
- Barrufaldi, S.H., Landoni, P., 2012. Return mobility and scientific productivity of researchers working abroad: the role of home country linkages, *Research Policy*, 41, 1655-1665.
- Bonnal, L., Giret, J.F., 2010. Determinants of access to academic careers in France, *Economics of Innovation and New Technology*, 19(5), 437-458.
- Bonney, N., Love, J., 1991. Gender and migration: geographical mobility and the wife's sacrifice, *The Sociological Review*, 39(2), 335-348.
- Boschma, R.A., 2005. Proximity and Innovation: a critical assessment, *Regional Studies*, 39(1), 61-74.
- Bouba-Olga, O., Grossetti, M., Lavigne, A., 2008. Le localisme dans le monde académique : une autre approche, *La Vie des Idées*, 12 mai.
- Bouba-Olga, O., Ferru, M., Pepin, D., 2012. Exploring spatial features of science-industry partnerships: a study on French data, *Papers in Regional Science*, 91(2), 355-375.
- Bozeman, B., Mangematin, V., 2004. Editor's introduction: building and deploying scientific and technical human capital, *Research Policy*, 33(4), 565-568.
- Breschi, S., Lissoni, F., 2009. Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows, *Journal of Economic Geography*, 9, 439-468.
- Brouillat, E., Lung, Y., 2010. Spatial distribution of innovative activities and economic performances: a geographical-friendly model, *Druid Summer conference*, London, 16-18 June.
- Canibano, C., Otamendi, F.J., Solis, F., 2011. International temporary mobility of researchers: a cross-discipline study, *Scientometrics*, 89, 653-675.
- CEC, 2001. A mobility strategy for the European research area, *Communication from the Commission to the Council and the European Parliament*.
- CEC, 2008. Realising a single labour market for researchers, *Report of the ERA expert group*.
- Combes, P.P., Linnemer, L., Visser, M., 2008. Publish or peer-rich? The role of skills and networks in hiring economics professors, *Labour Economics*, 15, 423-441.
- Criscuolo, P., 2005. On the road again: researcher mobility inside the R&D network, *Research Policy*, 34, 1350-1365.
- Cruz-Castro, L., Sanz-Menendez, L., 2010. Mobility versus job stability, assessing tenure and productivity outcomes, *Research Policy*, 39, 27-38.
- Debackere, K., Rappa, M.A., 1995. Scientists at major and minor universities: mobility along the prestige continuum, *Research Policy*, 24, 137-150.

- De Filippo, D., Sanz-Casado, E., Gomez, I., 2009. Quantitative and qualitative approaches to the study of mobility and scientific performance: a case study of a Spanish university, *Research Evaluation*, 18(3), 191-200.
- Dietz, J.S., Chompalov, I., Bozeman, B., Lane, E.O., Park, J., 2000. Using the curriculum vitae to study the career paths of scientists and engineers: an exploratory assessment, *Scientometrics*, 49(3), 419-442.
- Edler, J., Fier, H., Grimpe, C., 2011. International scientist mobility and the locus of knowledge and technology transfer, *Research Policy*, 40, 791-805.
- Florida, R., 2002. *The Rise of the Creative Class, and How It's Transforming Work, Leisure, Community And Everyday Life*, New York, Basic.
- Franzoni, C., Scellato, G., Stephan, P., 2012. Foreign born scientists: mobility patterns for sixteen countries, NBER Working Paper, No.18067.
- Frenken, K., Hardeman, S., Hoekman, J., 2009. Spatial scientometrics: towards a cumulative research program, *Journal of Informetrics*, 3(3), 222-232.
- Gaughan, M., Robin, S., 2004. National science training policy and early scientific careers in France and the United States, *Research Policy*, 33, 569-581.
- Génération 1998-2008, 2008. CEREQ, Centre Maurice Halbwachs (CMH).
- Godechot, O., Louvet, A., 2008. Le localisme dans le monde académique : un essai d'évaluation, *La Vie des Idées*, 22 avril.
- Greenwood, M.J., 1975. Research on internal migration in the United States: survey, *Journal of Economic Literature*, 13(2), 397-433.
- Griliches, Z., 1990. Patent statistics as economic indicators: a survey, *Journal of Economic Literature*, 28, 1661-1707.
- Grossetti, M., Milard, B., 2011. La concentration spatiale des activités scientifiques en question, *Communication pour le troisième colloque international du RESUP, « L'enseignement supérieur et la recherche en réformes »*, Paris, 27, 28 et 29 janvier 2011.
- Grossetti, M., Nguyen, D., 2001. La structure spatiale des relations science-industrie en France : l'exemple des contrats entre les entreprises et les laboratoires CNRS, *Revue d'Economie Rurale et Urbaine*, 101, 23-31.
- Heckman, J., 1979. Sample selection bias as a specification error, *Econometrica*, 47, 153-161.
- Hoekman, J., Frenken, K., Tijssen, R.J.W., 2010. Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe, *Research Policy*, 39(5), 662-673.
- Hoisl, K., 2007. Tracing mobile inventors, the causality between inventor mobility and inventor productivity, *Research Policy*, 36, 619-636.
- Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, *The Quarterly Journal of Economics*, 108(3), 577-598.
- Katz, J.S., Martin, B.R., 1997. What is research collaboration? *Research Policy*, 26, 1-18.
- Latham, W., Le Bas, C., Bouklia-Hassane, J.R., Volodin, D., 2011. Interregional mobility, productivity and the value of patents for prolific inventors in France, Germany and the UK, *International Journal of Foresight and Innovation Policy*, 7, 92-113.

- Maggioni, M., Uberti, T., 2007. Inter-regional knowledge flows in Europe: an econometric analysis, in Frenken K. (eds) *Applied evolutionary, economics and economic geography*, Edward Elgar, Cheltenham, UK.
- Martin-Brelot, H., Grossetti, M., Eckert, D., Gritsai, O., Kovacs, Z., 2010. The Spatial Mobility of the 'Creative Class': A European Perspective, *International Journal of Urban and Regional Research*, 34(4), 854-870.
- Musselin, C., 2009. *The Markets for Academics*, translation of « Le marché des universitaires » published in Presses de Sciences Po in 2005, New York, Routledge.
- Pezzoni, M., Sterzi, V., Lissoni, F., 2012. Career progress in centralized academics systems: social capital and institutions in France and Italy, *Research Policy*, 41, 704-719.
- Ponds, R., Van Oort, F., Frenken, K., 2007. The geographical and institutional proximity of research collaboration, *Regional Science*, 86(3), 423-443.
- Rallet, A., Torre, A., 2010. Proximity and localization, *Regional studies*, 39, 47-59.
- Schwartz, A., 1976. Migration, Age and Education, *Journal of Political Economy*, 84(4), 701-719.
- Shearmur, R., 2007. The new knowledge aristocracy: a few thoughts on the creative class, mobility and urban growth, *Work Labour and Globalization*, 1(1), 31-47.
- Shauman, K.A., Xie, Y., 1996. Geographic mobility of scientists: sex differences and family constraints, *Demography*, 33(4), 455-468.
- Trippl, M., 2011. Scientific Mobility and Knowledge Transfer at the Interregional and Intraregional Level, *Regional Studies*, 1-15.
- Yankow, J., 2003. Migration, job change, and wage growth: a new perspective on the pecuniary return to geographic mobility, *Journal of Regional Science*, 43(3), 486-516.
- Zellner, C., 2003. The economic effects of basic research: evidence for embodied knowledge transfer via scientists' migration, *Research Policy*, 32, 1881-1895.
- Zucker, L.G., Darby, M.R., 2008. Star scientists, innovation and regional and national immigration, NBER Working Paper, No.13547.

Annexe 1 – Construction de la base de données

La base de données DOCTHESE recense environ 218 000 thèses soutenues en France entre 1970 et 2000. Pour chacune des lignes qui correspondent à une thèse, nous disposons du nom et prénom de l’auteur de la thèse et du directeur de thèse. A partir de ces informations, il est possible de repérer des individus qui soutiennent une thèse et que l’on retrouve plus tard comme directeurs de thèse. Les tableaux ci-dessous illustrent la méthode de repérage des transitions et de transformation de la base de données :

Base d’origine : 217 994 lignes				
Docteur	Directeur de thèse	Année de soutenance	Université de soutenance	Discipline de soutenance
Individu A	Individu B	1979	Lyon	Economie
Individu C	Individu A	1990	Poitiers	Economie

Base après repérage des transitions : 12 261 lignes								
	Directeur de thèse	Année de soutenance	Université de soutenance	Discipline de soutenance	Première direction	Année de direction	Université de direction	Discipline de direction
Individu A	Individu B	1979	Lyon	Economie	Individu C	1990	Poitiers	Economie

Dans les cas d’homonymie du nom de famille, nous avons fait les appariements manuellement : le principal obstacle des données a été l’irrégularité du remplissage du prénom de l’auteur (parfois l’initiale seulement). Lorsqu’un individu apparaissait plusieurs fois en tant qu’auteur, c’est-à-dire qu’il a réalisé plusieurs thèses, nous avons gardé la première thèse. De la même façon, lorsqu’un individu apparaissait plusieurs fois en tant que directeur, nous avons gardé la première thèse qu’il a dirigée. Un moyen de contrôle de l’appariement a été de vérifier qu’il y ait au minimum trois ans d’écart entre soutenance et première direction, de façon à ne pas avoir de cas erronés. En prenant les précautions nécessaires, nous avons finalement obtenus une base de 12 261 individus, pour lesquels nous avons donc de l’information sur leur propre thèse et sur la première thèse dirigée. Ce nombre peut paraître faible au regard du nombre de lignes initiales, cependant plusieurs éléments peuvent expliquer cet écart :

- La perte des appariements en début et en fin de base : logiquement, il est peu probable que l’on retrouve dans cette base la thèse du directeur d’un docteur qui a soutenu au début des années 1970. Symétriquement, un docteur qui soutient dans les années 1990 n’aura probablement accès à la direction de thèse qu’après 2000.
- Des fautes d’orthographe, et des noms et prénoms mal renseignés ont inévitablement généré des appariements manqués.
- Plus un individu dirige de thèses, plus il « consomme » de lignes. Par exemple, lorsqu’un individu apparaît dix fois en tant que directeur de thèse, on passe de onze lignes dans la base originale à une seule ligne après appariement. La ligne sur laquelle l’individu est auteur et les dix lignes où il apparaît comme directeur ne deviennent qu’une seule ligne.
- Enfin, il existe des docteurs qui ne dirigent jamais de thèse.

Annexe 2 – Taille scientifique des régions

Le tableau ci-dessous donne la répartition régionale des docteurs à partir des 217 994 thèses recensées (1^è colonne), des 12 261 individus côté soutenance de thèse (2nde colonne) et des 12 261 individus côté direction de thèse (3^è colonne). Cette répartition montre l’hypertrophie de l’Ile-de-France qui produit 43.3% des docteurs, résultat proche des travaux empiriques sur la géographie de la science. Lorsque l’on compare la répartition des docteurs par région de soutenance et par région de direction, le poids de la région francilienne passe de la moitié (46.3%) des effectifs au tiers (33.3%), tandis que le poids de toutes les autres régions augmente. Cette tendance semble illustrer un processus de déconcentration de la recherche, déjà repéré par d’autres auteurs (Grossetti et Milard, 2011).

Région	Base d’origine	Région de Soutenance	Région de Direction
Académie de Grenoble	5.3%	5.4%	6.6%
Académie de Lyon	5.4%	5.5%	6.4%
Académie de Marseille	4.6%	4.7%	5.6%
Académie de Nice	1.9%	1.9%	2.3%
Alsace	4.0%	3.3%	3.6%
Aquitaine	3.8%	3.8%	3.9%
Auvergne	1.2%	1.0%	1.4%
Bourgogne	1.3%	1.1%	1.6%
Bretagne	3.0%	2.9%	3.8%
Centre	1.3%	1.1%	2.1%
Champagne-Ardenne	0.4%	0.4%	0.7%
Franche-Comté	0.9%	0.9%	1.2%
Ile-de-France	43.3%	46.3%	33.3%
Languedoc-Roussillon	4.5%	3.5%	4.4%
Limousin	0.5%	0.5%	0.7%
Lorraine	3.3%	3.0%	3.5%
Midi-Pyrénées	6.3%	6.3%	6.5%
Nord-Pas-De-Calais	3.0%	3.3%	3.9%
Basse-Normandie	1.0%	0.8%	1.3%
Haute-Normandie	0.9%	0.7%	1.3%
Pays-de-la-Loire	1.6%	1.2%	2.6%
Picardie	0.9%	0.7%	1.3%
Poitou-Charentes	1.6%	1.6%	1.9%
France	100.0%	100.0%	100.0%
Nb. Obs. :	217 994	12 261	12 261

Annexe 3 – Disciplines

Le tableau ci-dessous donne la répartition des docteurs par discipline, en fonction desquelles on observe des différences fortes en matière de mobilité. Globalement, les docteurs en sciences humaines et sociales changent davantage de région entre soutenance et première direction que les docteurs en sciences fondamentales et appliquées. Les nombres d'observations peuvent différer d'une colonne à l'autre, la discipline n'étant pas renseignée pour certains individus.

Discipline	Base d'origine	Discipline de soutenance	Discipline de direction	Taux de mobilité régionale
<i>Sciences humaines et sociales</i>	41.7%	40.9%	35.4%	47.4%
Droit, Sciences politiques	6.7%	4.5%	4.1%	42.9%
Littérature	3.5%	3.8%	3.2%	62.1%
Langues	7.0%	8.9%	7.7%	52.3%
Gestion, Communication	2.1%	1.6%	2.2%	40.5%
Histoire	5.1%	5.2%	4.5%	47.4%
Géographie	2.6%	2.8%	2.3%	51.8%
Philosophie	2.0%	2.3%	1.4%	51.5%
Psychologie	2.5%	3.2%	2.7%	37.0%
Sociologie, Sciences de l'éducation	3.7%	3.0%	2.7%	44.7%
Economie	4.5%	4.0%	2.9%	34.4%
Arts	2.0%	1.6%	1.7%	47.5%
<i>Sciences fondamentales et appliquées</i>	58.2%	59.1%	64.6%	29.3%
Biologie	14.4%	14.9%	17.4%	29.2%
Chimie	5.7%	4.4%	4.3%	32.0%
Mathématiques	2.8%	3.9%	3.7%	42.8%
Sciences médicales et pharma.	3.4%	3.1%	3.7%	24.4%
Physique	10.7%	12.8%	12.8%	25.7%
Sciences de l'ingénieur	12.7%	10.4%	13.1%	26.4%
Informatique	4.2%	5.2%	5.3%	28.1%
Sciences de la terre	4.3%	4.5%	4.3%	36.5%
Total	100.0%	100.0%	100.0%	36.0%
Nb. Obs. :	208 623	11 248	11 885	11 248

Annexe 4 – Statistiques descriptives

4a. Variables explicatives du modèle probit (données CEREQ)

	Obs.	Mean	Std. Dev.	Min.	Max.
≤ Baccalauréat	9121	0.49	0.50	0	1
> Baccalauréat & < Master	9121	0.41	0.49	0	1
Master	9121	0.07	0.25	0	1
Doctorat	9121	0.03	0.17	0	1
Docteur ESR	9121	0.01	0.09	0	1
Docteur non ESR	9121	0.02	0.14	0	1
Age	9121	0.38	0.49	0	1
Homme	9121	0.52	0.50	0	1
Enfant	9121	0.04	0.18	0	1

4b. Variables explicatives du modèle gravitaire (données DOCTHESE)

	Obs.	Mean	Std. Dev.	Min.	Max.
Taille scientifique région de départ	506	0.80	1.00	-0.93	3.77
Différentiel de taille	506	0.00	1.45	-4.70	4.70
Distance spatiale	506	1.55	0.43	0.00	2.25
Distance des profils disciplinaires	506	-1.40	0.29	-2.25	-0.79
Ile-de-France	506	0.09	0.28	0	1