HAL open science

Analyse empirique de la mobilité géographique des enseignants-chercheurs français

Bastien Bernela, Olivier Bouba-Olga, Marie Ferru

To cite this version:

Bastien Bernela, Olivier Bouba-Olga, Marie Ferru. Analyse empirique de la mobilité géographique des enseignants-chercheurs français. 2012. hal-00701750v1

HAL Id: hal-00701750

https://hal.science/hal-00701750v1

Preprint submitted on 26 May 2012 (v1), last revised 6 May 2013 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analyse empirique de la mobilité géographique des enseignants-chercheurs français

BERNELA Bastien, BOUBA-OLGA Olivier, FERRU Marie
CRIEF EA 2249 - Université de Poitiers[1]

Communication aux Septièmes Journées de la Proximité
« Créativité, Innovation et Proximités » - Montréal, 21-23 mai 2012
Version Provisoire

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 montrons d'abord le poids important de l'ancrage régional des chercheurs, puisque plus de 60% d'entre eux dirigent leur première thèse là où ils ont soutenu la leur. Nous testons ensuite un modèle gravitaire pour évaluer l'importance des effets taille et des effets de proximité sur les mobilités interrégionales observées. Nous montrons notamment l'influence positive de la taille scientifique des régions et de la proximité géographique sur la mobilité des chercheurs.

Mots-clefs : Mobilité, géographie de la science, proximités, effet taille

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 show the importance of regional rooting: more than 60% of the researchers direct their first Ph.D where they defended their own one. Then, we test a gravity model (sample selection model) to estimate the importance of size and proximity effects on the observed interregional mobility. We observe in particular the positive influence of the scientific size of regions and geographical proximity on the researchers' mobility.

Keywords: Mobility, geography of science, proximities, size effect

[1] bastien.bernela@univ-poitiers.fr ; obouba@univ-poitiers.fr ; marie.ferru@univ-poitiers.fr

Laboratoire CRIEF, EA 2249, Faculté de Sciences Économiques de Poitiers, 2 rue Jean Carbonnier, Bât.A1 (BP 623), 86022 POITIERS Cedex

1. Introduction

La mobilité spatiale des créatifs est une question essentielle, puisque du degré et des déterminants de leur mobilité dépend la réussite des stratégies de développement économique locale qui misent sur l'attractivité vis-à-vis de cette catégorie de la population. Cette problématique a été abordée ces dernières années dans le cadre des travaux sur la classe créative, initiés par Richard Florida. Selon ce dernier, les enseignants-chercheurs peuvent être considérés comme une catégorie particulière de la classe créative (2002, p.8)[2] et ils présenteraient, de la même manière que l'ensemble des créatifs, une mobilité géographique intense, se déplaçant vers les villes jugées "attractives". Nous proposons dans cet article de nous pencher sur cette catégorie particulière et importante des créatifs[3], et montrer, dans la continuité d'Artin-Brelet et al. (2010) que leur mobilité n'est pas aussi forte que Florida le suggère.

Dans cette perspective, nous mobilisons une base de données originale de longue période (1970-2000) disponible pour la France, qui recense les transitions entre soutenance de la première thèse et direction de la première thèse de plus de 12 000 individus[4]. Connaissant la localisation de la soutenance et de la direction de thèse, nous pouvons nous focaliser sur cette étape particulière de la trajectoire professionnelle des chercheurs et nous interroger sur sa dimension géographique : dirigent-ils leur première thèse là où ils ont soutenu leur thèse ? 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 ?

Cette base a déjà été exploitée par Godechot et Louvet (2008) pour mesurer le « recrutement » local, défini comme « le fait de soutenir et de diriger sa première thèse (…) dans la même Université et dans la même discipline » (p. 9). Ils montrent alors que 46% des docteurs « recrutés » dirigent dans l'université où ils ont débuté et indiquent par ailleurs que « les établissements les moins « localistes » (…) sont les universités parisiennes » (p. 17). Ces résultats feraient du « recrutement » local un phénomène beaucoup plus marqué que ce que d'autres études laissaient penser (cf. le rapport Fréville (2001) notamment). Les auteurs l'expliquent par les modalités spécifiques de recrutement dans l'enseignement supérieur français, qu'ils proposent logiquement de réformer, en interdisant le recrutement local.

L'objectif de cet article est d'approfondir l'analyse de la mobilité géographique des chercheurs français, en mobilisant la même base de données, mais en proposant une démarche et un traitement empirique différents. Nous considérons d'abord que la base de données ne permet pas de traiter directement du recrutement des chercheurs, mais de leur mobilité spatiale au début de leur trajectoire professionnelle. Nous insistons ensuite sur la nécessité de distinguer deux échelles d'observation : l'échelle de l'établissement et l'échelle de la région. Nous montrons alors que i) le

---

[2] « The creative class is comprised of a « super creative core », which consists of a new class of scientists and engineers, *university professors*, poets, actors, novelists entertainers, artists, architects and designers, cultural worthies, think-tank researchers, analysts and opinion formers, whose economic function is to create new ideas, new technology, and/or new creative content » (Florida, 2002, p.8).

[3] Nous utiliserons par la suite le terme de « chercheur » pour englober les chercheurs et enseignants-chercheurs travaillant à l'Université.

[4] Nous remercions vivement Olivier Godechot d'avoir mis à notre disposition la base de données DOCTHESE.

localisme est particulièrement fort à l'échelle régionale, puisqu'en moyenne plus de 60% des chercheurs dirigent leur première thèse dans la région où ils ont soutenu leur thèse, ii) la spécificité des Universités parisiennes disparaît, puisque prises ensemble à l'échelle régionale, elles sont aussi localistes que les autres. Nous réalisons un modèle probit afin d'expliquer le caractère local de la transition doctorat-direction de thèse et testons ainsi différentes variables explicatives - l'année de soutenance, la discipline scientifique et la région de soutenance - afin de repérer l'existence d'effets temporels, de spécialisation et régionaux sur le localisme.

Le deuxième objectif de notre article est de s'interroger sur les déterminants de la mobilité interrégionale des chercheurs, pour évaluer plus précisément le jeu de certaines caractéristiques structurelles des régions : alors que Godechot et Louvet 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 (taille économique, taille scientifique) et leur degré de proximité (géographique, scientifique, sectorielle) 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.

2. Données et méthode

La mesure de la trajectoire spatiale des chercheurs est un exercice que le manque de données sur longue période rend difficile. Nous proposons ici de suivre la stratégie adoptée par Godechot et Louvet (2008) qui ont mobilisé la base de données DOCTHESE. Cette base, qui recense de manière quasi exhaustive les thèses soutenues entre 1970 et 2000 (soit près de 213 000 thèses), permet de repérer les docteurs qui deviennent quelques années plus tard directeurs de thèse : c'est le cas pour 12 585 individus, qui constituent donc la base intermédiaire pour effectuer nos traitements[5]. Pour chacune des lignes de cette nouvelle base, qui correspondent à une transition doctorat-direction de thèse, on dispose des noms et prénoms de l'individu dont on a repéré la transition, de son directeur de thèse et de son premier doctorant, des années, universités et disciplines de sa thèse et de sa première thèse dirigée. Nous déduisons de l'université de soutenance (respectivement de première direction), la région de soutenance (respectivement de première direction), ce qui permet d'observer la mobilité à différentes échelles.

La répartition régionale des individus (cf. Annexe 2) montre l'hypertrophie de l'Ile-de-France qui produit près de la moitié des docteurs, résultat commun à de nombreux travaux empiriques sur la géographie de la science. Lorsque l'on regarde la répartition par région de direction, la région francilienne ne représente plus qu'un tiers, 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, Milard, 2011). Les régions qui suivent en termes d'effectifs sont Rhône-Alpes, Provence-Alpes-Côte-d'Azur et Midi-Pyrénées : la base DOCTHESE donne donc un classement plutôt habituel des régions françaises. D'ailleurs, ce classement est très corrélé avec celui de la taille économique des régions, et encore davantage avec leur taille scientifique (cf. Annexe 3). On retrouve donc à travers cette base de données une distribution régionale représentative des

[5] Pour plus de précisions sur la construction de la base de données, cf. Annexe 1

dotations scientifiques, à savoir une très forte concentration des laboratoires et des chercheurs dans la région Ile-de-France.

Godechot et Louvet ont mobilisé cette base pour mesurer le poids du recrutement local dans les Universités françaises. Ils définissent le recrutement local comme « le fait de soutenir et de diriger sa première thèse (…) dans la même Université et dans la même discipline » (p. 9), et montrent alors 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. Certaines, d'ordre théorique, ont déjà été soulignées par Bouba-Olga, Grossetti et Lavigne (2008)[6]. C'est sur le versant plus empirique que nous souhaitons nous focaliser.

Leur stratégie d'identification du recrutement, d'abord, est contestable : un individu ayant soutenu sa thèse dans une Université x, étant ensuite recruté comme enseignant-chercheur dans une Université y, revenant enfin dans son Université d'origine par mutation et y dirigeant sa première thèse, sera considéré comme recruté localement. A l'inverse, un individu ayant soutenu en x, recruté en x (donc recruté localement), mais poursuivant sa carrière en y où il dirige sa première thèse, ne sera pas considéré comme recruté localement. La base de données mobilisée ne permet pas d'identifier ce type de trajectoire. La stratégie de Godechot et Louvet est sans doute la mieux adaptée pour approcher la géographie du recrutement à partir de DOCTHESE, mais nous préférons pour notre part décaler la problématique : il ne s'agira pas, pour nous, de nous prononcer sur la question du recrutement dans l'enseignement supérieur, trop difficile à saisir avec les données de la base, mais sur la mobilité géographique des chercheurs au début de leur trajectoire professionnelle. Ceci n'interdit pas de parler de « localisme », « d'attachement local », « d'ancrage local » ou encore de « biais local » (en référence au « biais domestique » cher à l'économie de la finance) lorsqu'un individu soutient et dirige sa première thèse dans la même Université, mais nous n'en déduirons rien sur la question du recrutement et de son éventuel localisme.

Une deuxième limite importante relève de l'échelle d'observation retenue par les auteurs. Ces derniers observent le localisme à l'échelle des établissements, nous proposons pour notre part de distinguer deux niveaux d'observation :

- le localisme d'établissement, défini comme le fait de soutenir sa première thèse et de diriger sa première thèse dans le même établissement, c'est-à-dire dans la même université,
- le localisme régional, défini comme le fait de soutenir sa première thèse et de diriger sa première thèse dans la même région administrative.

Cette distinction est essentielle pour certaines régions composées d'un grand nombre d'universités : le localisme d'établissement peut en effet être relativement faible mais s'accompagner d'un fort localisme régional si les docteurs d'une université dirigent leur première thèse dans une autre université de la région. Nous verrons que cette configuration correspond typiquement au cas de l'Ile de France.

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

La troisième limite de l'analyse de Godechot et Louvet, liée aux précédentes, résulte de leur hypothèse implicite selon laquelle la mobilité ou la non-mobilité des chercheurs s'expliquerait largement par la nature du processus de recrutement dans l'enseignement supérieur. Nous considérons pour notre part que la mobilité géographique des chercheurs s'explique aussi, et peut-être surtout, par d'autres caractéristiques qui échappent largement à la question du recrutement. La taille économique et scientifique des régions ainsi que leur degré de proximité géographique, scientifique et sectorielle peuvent expliquer le nombre de mobilités entre les régions. Afin d'évaluer l'importance de ces effets, nous proposons de tester 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.

3. Du localisme d'établissement au localisme régional

Nous proposons dans cette section de nous focaliser sur le localisme régional. Nous avons distingué précédemment le localisme d'établissement et le localisme régional. Par construction, le dernier est supérieur au premier, mais la variation entre les deux mesures est particulièrement importante : le taux de localisme passe de 43,4% en moyenne pour l'ensemble des établissements français à 63,5% à l'échelle régionale (Tableau 1). On observe donc de nombreuses mobilités inter-établissements mais infrarégionales. Ceci concerne principalement l'Ile-de-France, puisque le localisme d'établissement y est de 29,4% (score très largement inférieur à celui des autres régions) alors que le localisme régional monte à 61,1% (score dans la moyenne). On retrouve ce type d'évolution, un peu moins marquée, pour Rhône-Alpes et Provence-Alpes-Côte-d'Azur. D'autres régions, dotées seulement d'une Université, ne voient pas, logiquement, leur score évoluer.

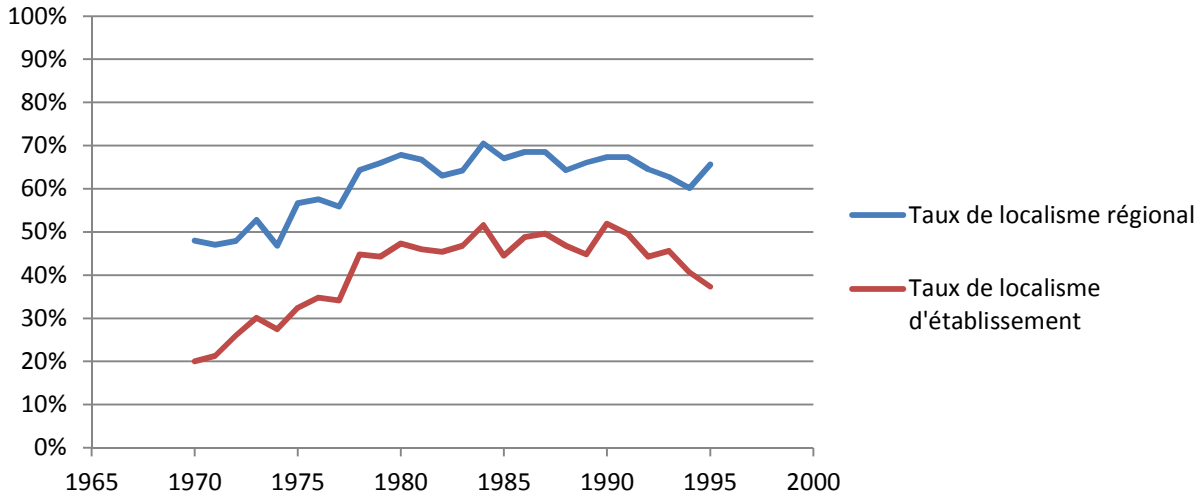
Tableau 1 : Localisme d'établissement vs localisme régional

	Niveau établissement	Niveau régional
Alsace	59,0%	63,1%
Aquitaine	53,9%	60,5%
Auvergne	56,5%	56,5%
Bourgogne	64,8%	64,8%
Bretagne	56,2%	62,3%
Centre	44,2%	47,1%
Champagne-Ardenne	44,4%	44,4%
Franche-Comté	48,7%	48,7%
Ile-de-France	29,4%	61,1%
Languedoc-Roussillon	54,1%	60,5%
Limousin	65,5%	65,5%
Lorraine	44,3%	61,8%
Midi-Pyrénées	58,8%	69,3%
Nord-Pas-De-Calais	64,1%	72,3%
Basse-Normandie	50,9%	50,9%
Haute-Normandie	68,9%	73,3%
Pays-de-la-Loire	65,6%	66,2%
Picardie	66,7%	68,9%
Poitou-Charentes	65,3%	65,3%
Provence-Alpes-Côte-d'Azur	54,6%	67,1%

Rhône-Alpes	51,8%	72,2%
France	**43,4%**	**63,5%**

Si l'on s'intéresse à l'évolution du localisme, on remarque que les taux de localisme d'établissement et régional suivent la même tendance, avec une certaine inertie. Après une hausse progressive durant les années 1970, le taux de localisme régional se stabilise entre 60 et 70%, laissant deviner une relative inertie de la géographie de la science.

Graphique 1 : Evolution des taux de localisme



Afin de confirmer et de préciser ces résultats, nous avons testé un modèle probit binaire, avec LOC_k^* une variable latente et LOC_k la variable binaire associée qui prend la valeur 1 si l'individu k dirige sa première thèse dans l'Université où il a soutenu sa thèse et la valeur 0 sinon.

$$LOC_k = \begin{cases} 1 \; si \; LOC_k^* \geq 0 \\ 0 \; si \; \; LOC_k^* < 0 \end{cases}$$

avec $LOC_k^* = \alpha + Z_k\beta + \mu_k$

Z_k représente le vecteur des variables explicatives et β le vecteur des paramètres associés. Les variables explicatives retenues sont l'année, la région et la discipline[7] de soutenance de thèse du chercheur. Les résultats sont résumés dans le tableau 2.

Tableau 1 : Déterminants du localisme régional (modèle probit)

Variables	**Effet marginal**	**Ecart-type**
Année de soutenance	0,00	0,00
Alsace	référence	
Aquitaine	-0,02	0,03
Auvergne	-0,07	0,05

[7] La répartition des données par discipline est en Annexe 4.

Variables	Effet marginal	Ecart-type	
Bourgogne	0,02	0,05	
Bretagne	0,00	0,04	
Centre	-0,17	0,05	***
Champagne-Ardenne	-0,20	0,08	***
Franche-Comté	-0,12	0,05	**
Ile-de-France	0,02	0,02	
Languedoc-Roussillon	-0,03	0,03	
Limousin	-0,01	0,07	
Lorraine	-0,04	0,04	
Midi-Pyrénées	0,05	0,03	*
Nord-Pas-de-Calais	0,07	0,03	**
Basse-Normandie	-0,11	0,06	**
Haute-Normandie	0,08	0,05	
Pays-de-la-Loire	-0,01	0,05	
Picardie	0,00	0,06	
Poitou-Charentes	0,01	0,04	
Provence-Alpes-Côte-d'Azur	0,04	0,03	
Rhône-Alpes	0,07	0,03	***
Economie	référence		
Arts et archéologie	-0,19	0,04	***
Droit	-0,11	0,03	***
Ethnologie, science des religions	0,00	0,05	
Langues et civilisations	-0,24	0,02	***
Géographie	-0,20	0,03	***
Gestion	-0,12	0,04	***
Histoire	-0,20	0,03	***
Linguistique	-0,15	0,03	***
Littérature	-0,29	0,03	***
Philosophie	-0,21	0,03	***
Psychologie	-0,07	0,03	***
Sociologie	-0,12	0,03	***
Biologie	0,02	0,02	
Chimie	-0,02	0,03	
Mathématiques	-0,12	0,03	***
Sciences médicales	0,08	0,03	**
Physique	0,06	0,02	***
Sciences de l'ingénieur	0,04	0,02	*
Informatique	0,01	0,02	
Terre, océan, géologie	-0,05	0,03	*
Autres	-0,10	0,03	***
Nb. Obs. : 12 585			
LR Chi2(42)=657,71*** - Log likelihood = -7926,87			

Note : *P<0,1, **P<0,05, ***P<0,001

Les résultats des estimations montrent que sept des vingt-et-une régions de France métropolitaine présentent des effets marginaux significatifs. Quatre d'entre elles sont moins localistes que la région de référence, l'Alsace (taux de 63,1%) : le Centre, la Champagne-Ardenne, la Franche-Comté et la Basse-Normandie. Trois sont plus localistes : Midi-Pyrénées, Nord-Pas-de-Calais et Rhône-Alpes, autrement dit des régions de taille importante.

On observe également des effets disciplinaires. La plupart des disciplines de Sciences Humaines sont moins localistes que la discipline de référence (l'économie). Les sciences médicales et la physique sont à l'inverse plus localistes. On note enfin l'absence d'effet statistiquement significatif lié à l'année de soutenance, résultat confirmant la tendance à l'inertie de la géographie de la science.

4. La mobilité interrégionale des enseignants-chercheurs

Nous proposons maintenant d'élargir la perspective, pour analyser les mobilités observées entre les régions françaises. Nous avons commencé par construire une matrice des flux afin de caractériser les tendances de mobilité (ou de non-mobilité) des enseignants-chercheurs français (Tableau 3). En moyenne, 63,54% restent dans leur région de soutenance, 9,65% partent vers l'Ile-de-France, 6,19% vers des régions limitrophes et 20,62% vers des régions non limitrophes.

Tableau 2 : De la région de soutenance à la région de direction de thèse

	Même Région	Ile-de-France	Régions limitrophes (hors IDF)	Autres Régions	Total général
Alsace	63,13%	9,64%	6,27%	20,96%	100,00%
Aquitaine	60,47%	8,67%	6,98%	23,88%	100,00%
Auvergne	56,49%	12,98%	22,13%	8,40%	100,00%
Bourgogne*	64,79%	9,15%	11,27%	14,79%	100,00%
Bretagne	62,33%	13,57%	6,65%	17,45%	100,00%
Centre*	47,10%	18,12%	10,14%	24,64%	100,00%
Champagne-Ardenne*	44,44%	22,22%	13,34%	20,00%	100,00%
Franche-Comté	48,67%	10,62%	23,90%	16,82%	100,00%
Ile-de-France	61,12%	-	5,84%	33,04%	100,00%
Languedoc-Roussillon	60,47%	10,59%	16,00%	12,94%	100,00%
Limousin	65,52%	5,17%	12,07%	17,24%	100,00%
Lorraine	61,80%	8,75%	6,10%	23,35%	100,00%
Midi-Pyrénées	69,30%	9,52%	6,89%	14,29%	100,00%
Nord-Pas-De-Calais	72,29%	7,23%	1,69%	18,79%	100,00%
Basse-Normandie	50,94%	10,38%	20,76%	17,92%	100,00%
Haute-Normandie*	73,33%	7,78%	0,00%	18,89%	100,00%
Pays-De-Loire	66,24%	10,19%	7,01%	16,56%	100,00%
Picardie*	68,89%	15,56%	2,22%	13,33%	100,00%
Poitou-Charentes	65,33%	8,54%	10,55%	15,58%	100,00%
Paca	67,07%	8,81%	9,67%	14,45%	100,00%
Rhône-Alpes	72,16%	8,57%	7,03%	12,24%	100,00%
France	**63,54%**	**9,65%**	**6,19%**	**20,62%**	**100,00%**

* régions limitrophes à l'Ile-de-France.

On remarque que l'attractivité de l'Ile-de-France a un effet plus marqué chez les régions limitrophes, dont les docteurs sont « aspirés » par la capitale. La proximité spatiale explique partiellement les mobilités interrégionales.

Nous avons testé ensuite un modèle gravitaire afin d'évaluer l'existence d'effets taille et d'effets de proximité dans les choix de mobilité des chercheurs, en nous focalisant uniquement sur les mobilités interrégionales. Les cas de localisme régional sont donc exclus du modèle. Le modèle testé est plus précisément un modèle de Heckman avec équation de sélection, communément utilisé dans les recherches sur le commerce international, et invoqué récemment pour expliquer la géographie des collaborations. 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 économique et scientifique de la région de soutenance et de la région de première direction de thèse, d'une part, et de différentes formes de proximité (géographique, scientifique et sectorielle), 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 :

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

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 :

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

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

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

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

Les variables explicatives sont mesurées de la façon suivante :

1) Pour tester l'effet taille des régions, nous introduisons deux variables i) la taille économique mesurée à partir de la part moyenne du PIB régional sur la période 1997-2006 et ii) la taille scientifique, mesurée par la part moyenne de chercheurs publics de la région sur la période 1997-2006 (Source : INSEE).
2) Pour vérifier l'existence éventuelle d'effets de proximité spatiale, nous intégrons une matrice des distances interrégionales, mesurée en nombre de kilomètres entre les capitales régionales.

3) Nous testons d'autres formes de proximité, d'essence non spatiale : il s'agit de proximités scientifique et sectorielle. Pour construire ces variables, nous nous sommes inspirés de l'Aquino index (1978), compris entre 0 et 1 : dans notre cas, plus cet indicateur est proche de 1, plus le profil des deux régions considérées est proche.

Le tableau 4 ci-dessous présente les résultats de l'équation d'output et de l'équation de sélection. La significativité du ratio de Mills au seuil de 1% indique la présence d'un biais de sélection, contrôlé par notre modèle.

Tableau 4 : Déterminants de la mobilité interrégionale des chercheurs (modèle de Heckman)

Probabilité/intensité des mobilités interrégionales			
Equation d'output (346 observations)			
	Coefficient	**Ecart-type**	
Distance	-0,126	0,022	***
Proximité Scientifique	1,268	0,939	
Proximité Sectorielle	0,061	0,757	
Taille économique Région départ	-0,346	0,215	
Taille scientifique Région départ	1,212	0,170	***
Taille économique Région arrivée	0,026	0,181	
Taille scientifique Région arrivée	0,735	0,130	***
Ile-de-France	0,864	0,343	**
constante	0,491	0,335	
Equation de sélection (420 observations)			
	Coefficient	**Ecart-type**	
Distance	-0,482	0,207	**
Proximité Scientifique	0,409	1,767	
Proximité Sectorielle	-0,800	1,583	
Taille économique Région départ	-1,114	0,383	***
Taille scientifique Région départ	1,587	0,262	***
Taille économique Région arrivée	0,233	0,382	
Taille scientifique Région arrivée	0,677	0,244	***
constante	1,256	0,692	
Terme d'erreur			
	Coefficient	**Ecart-type**	
Ratio de mills	0,932	0,311	***
Wald chi2(8) = 323,16	Prob > chi2 = 0,0000		

Note : *P<0,1, **P<0,05, ***P<0,001

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 le rôle négatif de la distance interrégionale, d'une part, ainsi que l'existence de plusieurs effets taille. Plus deux régions sont éloignées géographiquement, plus la probabilité d'observer des mobilités entre elles diminue. La taille scientifique, à la fois de la région de départ et de la région d'arrivée, influent sur la probabilité estimée. Ces résultats sont ceux attendus : plus une région produit de docteurs,

plus on s'attend à ce que certains d'entre eux migrent vers d'autres régions ; plus une région est de taille scientifique importante, plus on s'attend à ce qu'elle accueille des docteurs venus d'autres régions. L'influence négative de la taille économique de la région de départ est en revanche moins facile à interpréter. Les variables résumant les autres formes de proximité ne sont pas statistiquement significatives.

S'agissant de l'équation d'output, qui estime le nombre de mobilités entre deux régions, on observe logiquement l'influence positive de la variable muette « Ile-de-France ». On retrouve également une influence significative de la taille scientifique de la région de départ et de 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 : hors cas de localisme, il apparaît que les chercheurs se déplacent vers des régions proches de celle où ils ont soutenu leur thèse. Notons que les formes de proximité non spatiale ne jouent toujours pas.

5. Conclusion

Nous avons montré dans cet article que le localisme régional des chercheurs est élevé, puisque plus de 60% dirigent leur première thèse là où ils ont soutenu la leur. La région Ile-de-France ne fait pas exception à cette tendance contrairement à ce qu'avaient souligné Godechot et Louvet (2008) dans une analyse précédente. Nous nous sommes focalisés ensuite sur les déterminants de la mobilité interrégionale. Il s'avère que la distance interrégionale mais surtout la taille scientifique des régions exercent une influence certaine sur la probabilité et/ou le nombre de mobilités observées.

Au final, la mobilité des chercheurs se révèle plutôt faible et, lorsqu'on l'observe, elle s'explique pour une part significative par des *hard factors* sur lesquels les politiques locales n'ont que peu de prise. A l'aune de ces résultats, la stratégie d'attractivité vis-à-vis des créatifs peut sans doute être réinterrogée.

Bibliographie

ARTIN-BRELOT H., GROSSETTI M., ECKERT D., GRITSAI O. and KOVÁCS Z. (2010), The Spatial Mobility of the 'Creative Class': A European Perspective. International Journal of Urban and Regional Research, Vol.34, n°4, p. 854–870.

BOUBA-OLGA O., GROSSETTI M., LAVIGNE A. (2008), « Le localisme dans le monde académique : une autre approche », La Vie des Idées, 12 mai 2008.

FLORIDA R. (2002), The Rise of the Creative Class, and How It's Transforming Work, Leisure, Community And Everyday Life.

GODECHOT O., LOUVET A. (2008), « Le localisme dans le monde académique : un essai d'évaluation », La Vie des Idées, 22 avril 2008.

GROSSETTI M., MILARD B., 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

Annexes

Annexe 1 – Construction de la base de données

La base de données DOCTHESE telle que nous l'avons obtenue par Godechot recense 218 875 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. C'est avec le langage VBA que nous avons procédé à l'appariement sur la base du nom et du prénom de chaque auteur. 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. Avec le maximum de précautions, nous sommes arrivés finalement à une base intermédiaire de 12 585 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 le directeur de thèse d'un docteur qui a soutenu au début des années 1970. Symétriquement, un docteur qui soutient à la fin des 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 causé 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 dirigeront jamais de thèse.

Pour chacune des lignes de cette nouvelle base, qui correspond désormais à une transition doctorat-direction de thèse, on dispose : des noms et prénoms de l'individu dont on a repéré la transition, de son directeur de thèse et de son premier doctorant, des années, universités et disciplines de sa soutenance et de sa première direction. La durée d'accession à la direction de thèse est de onze ans en moyenne après l'obtention du grade de docteur.

Variables d'origine
Identifiant de l'individu dont on a repéré la transition
Noms et prénoms de cet individu, de son directeur de thèse et de son premier doctorant
Année de soutenance / de direction
Université de soutenance / de direction
Discipline de soutenance / de direction (3 niveaux de nomenclature)
Variables construites

Région de soutenance / de direction
Tailles économique et scientifique des régions
Distance spatiale par couple de régions
Distance scientifique et sectorielle par couple de régions

Nous déduisons de l'université de soutenance (respectivement de première direction), la région de soutenance (respectivement de première direction) afin de tester le localisme à plusieurs échelles. Nous ajoutons un indicateur de taille économique et scientifique des régions. A chaque individu de la base est associé un couple région de soutenance – région de direction de thèse. Pour chacun de ces couples, on définit la distance géographique (évaluée en nombre de frontières et en temps de train entre les capitales régionales) et la distance cognitive (évaluée à partir de la spécialisation scientifique des régions) qui séparent les deux régions.

Annexe 2 - Répartition régionale des docteurs et directeurs de thèse

Région	Région de Soutenance	Région de Direction
Alsace	3,30%	3,66%
Aquitaine	3,76%	3,93%
Auvergne	1,04%	1,41%
Bourgogne	1,13%	1,59%
Bretagne	2,87%	3,76%
Centre	1,10%	2,07%
Champagne-Ardenne	0,36%	0,69%
Franche-Comté	0,90%	1,23%
Ile-de-France	46,64%	33,66%
Languedoc-Roussillon	3,38%	4,36%
Limousin	0,46%	0,76%
Lorraine	3,00%	3,44%
Midi-Pyrénées	6,34%	6,47%
Nord-Pas-De-Calais	3,30%	3,92%
Basse-Normandie	0,84%	1,29%
Haute-Normandie	0,72%	1,32%
Pays-de-la-Loire	1,25%	2,57%
Picardie	0,72%	1,26%
Poitou-Charentes	1,58%	1,93%
Provence-Alpes-Côte-d'Azur	6,49%	7,87%
Rhône-Alpes	10,85%	12,81%
France	**100,00%**	**100,00%**
Nb. Obs. :	**12585**	**12585**

Annexe 3 – Matrice de corrélation de rang de Kendall

	Région de soutenance	Région de direction	Taille économique	Taille scientifique
Région de soutenance	1	-	-	-

Région de direction	0.8952	1	-	-
Taille économique	0.6190	0.7048	1	-
Taille scientifique	0.8571	0.9429	0.6667	1

Annexe 4 – Répartition par discipline des docteurs et directeurs de thèse

Discipline	Discipline de soutenance	Discipline de direction
Arts et archéologie	1,72%	1,75%
Droit	3,76%	3,35%
Ethnologie, science des religions	0,99%	1,05%
Langues et civilisations	5,78%	5,19%
Géographie	2,81%	2,29%
Gestion	1,45%	1,65%
Histoire	4,40%	3,62%
Linguistique	3,17%	2,65%
Littérature	3,79%	3,14%
Philosophie	2,34%	1,39%
Psychologie	3,27%	2,75%
Economie	3,99%	2,93%
Sociologie	2,29%	1,75%
Biologie	14,66%	16,96%
Chimie	4,41%	4,30%
Mathématiques	3,78%	3,68%
Sciences médicales	2,53%	2,88%
Physique	12,60%	12,70%
Sciences de l'ingénieur	10,40%	13,06%
Informatique	5,21%	5,20%
Terre, océan, géologie	4,41%	4,25%
Autres	2,24%	3,46%
Total	**100,00%**	**100,00%**
Nb. Obs. :	**11567**	**12201**