



HAL
open science

La proximité géographique compte-t-elle toujours ?

Olivier Bouba-Olga, Marie Ferru

► **To cite this version:**

Olivier Bouba-Olga, Marie Ferru. La proximité géographique compte-t-elle toujours ? : Une analyse empirique de deux bases de données relationnelles de longue période. 2012. hal-00725073v1

HAL Id: hal-00725073

<https://hal.science/hal-00725073v1>

Preprint submitted on 23 Aug 2012 (v1), last revised 24 Aug 2012 (v2)

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.

La proximité géographique compte-t-elle toujours ? Une analyse empirique de deux bases de données relationnelles de longue période

Olivier Bouba-Olga & Marie Ferru

CRIEF, Université de Poitiers

Résumé : L'objectif de cet article est d'apporter des éléments de réponse, à un niveau empirique, à la question "la proximité géographique compte-t-elle toujours dans le cadre des collaborations pour l'innovation ?", le manque de données disponibles n'ayant permis d'y répondre véritablement jusqu'alors. En nous appuyant sur deux véritables bases de données relationnelles de longue période relatives aux collaborations science-industrie en France et sur deux indicateurs complémentaires de proximité géographique, nous montrons que la proximité continue de compter (nombre de collaborations intra-départementales importantes et distance moyenne entre les partenaires faible). L'analyse dynamique permet de préciser toutefois que ce sont conjointement les partenariats les plus proches et les plus distants qui augmentent le plus sur la période d'étude. Nous montrons également que le rôle de la proximité géographique et son évolution au cours du temps diffèrent selon le type de contrat science-industrie et la spécialisation sectorielle des partenaires.

Mots clés: proximité géographique, collaborations, recherche, science-industrie

Jel Codes: O31, R12, O39

Remerciements : Les auteurs remercient Michel Grossetti pour avoir mis à notre disposition la base CNRS, Nadine Massard pour nous avoir fourni la base CIFRE et Koen Frenken pour ses remarques sur une première version de ce texte.

1. Introduction

Comme le souligne Morgan (2004, p.3) « globalization and digitization have been presented as ineluctable forces which signal the death of geography ». Le développement des TIC et des moyens de transport entrainerait la construction d'un monde économique global où la distance ne compte plus (Castells, 1996 ; Cairncross, 1997). Les activités d'innovation sembleraient faire exception à ce processus selon bon nombre d'auteurs. Leur forte polarisation observée dès la fin des années 1990 (Puga, 1999 ; Caniels, 1997 ; Lallement et al., 2007, etc.) et la persistance des effets de proximité géographique qui en découle proviendraient du caractère tacite des connaissances échangées, de l'existence de réseaux sociaux entre les partenaires développés localement, de l'intervention des politiques, etc.

D'un point de vue empirique, "the death of distance hypothesis has not been proven" (Frenken et al., 2010) dans le cadre des collaborations de recherche et les travaux réalisés souffrent de limites importantes compte-tenu de la faible disponibilité des données. En effet, non seulement, « most of existing contribution on this matter lacks data at the individual level » (Giuri et Mariani, 2007) mais également et surtout « many studies on proximity and innovation have been analysed from a static perspective » (Boschma, 2005). Ainsi, bien que la question de la proximité géographique dans le cadre des collaborations de recherche ait déjà été posée dans de nombreux travaux (voir notamment Oinas, 2000 ou Lublinski, 2003), aucun ne permet véritablement de donner une réponse à la question « does geographical proximity *still* matter ? ».

Nous cherchons dans cet article à donner des éléments de réponse à cette question grâce à un travail empirique basé sur l'exploration de deux véritables bases de données relationnelles sur longue période. Ces dernières intègrent de nombreuses informations relatives à plus de 30 000 couples de collaborations science-industrie réalisées en France entre 1986 et 2005, telle que la localisation des partenaires scientifique et industriel, leur spécialisation et la date du projet. Ces données nous permettent de réaliser une analyse de l'importance et de l'évolution de la proximité géographique pour deux types de collaborations, une partie des partenariats étudiés relevant de contrats CIFRE, l'autre partie des contrats CNRS (voir plus loin pour des précisions sur ces deux types de contrats). Cette analyse comparative apparaît d'autant plus intéressante que nous pouvons nous attendre à des résultats divergents selon la nature du contrat : le besoin de proximité géographique pourrait être plus important dans le cadre des collaborations CIFRE puisque ces contrats reposent sur le travail de recherche d'un doctorant alternant des périodes en entreprise et au laboratoire.

Nous proposons de mener une analyse empirique de ces deux bases de données en adoptant deux stratégies complémentaires de définition de la proximité géographique : i) nous la mesurons d'abord grâce à un indicateur de distance relative, le temps de déplacement en train entre les deux capitales des régions où sont localisés les deux partenaires, ii) nous définissons ensuite une typologie des échelles spatiales des collaborations permettant de définir différents degrés de proximité géographique. Nous distinguons plus précisément les collaborations au sein d'un même département (DEP), celles entre départements d'une même région (REG), les collaborations entre régions limitrophes (LIMITROPHE), celles avec la région capitale (ICAP) et enfin celles entre régions non limitrophes, hors région capitale (IREG).

Dans le premier temps de notre analyse empirique, nous mesurons l'évolution sur la période des distances-temps et de chaque classe de notre typologie et testons l'existence d'un trend en nous appuyant sur des tests non paramétriques (tau de Kendall), en distinguant selon la nature du contrat

(CIFRE vs. CNRS). Dans un second temps, nous testons l'influence de différents déterminants potentiels (l'année de réalisation du contrat, le secteur d'activité de l'établissement ou le domaine de recherche du laboratoire) sur les effets de proximité géographique grâce à deux modèles économétriques complémentaires : le premier modèle (logit multinomial) appréhende les effets de proximité grâce à la typologie des échelles spatiales alors que le second (sample selection model, tow-step Heckman) mesure l'influence des mêmes déterminants sur la distance-temps entre les partenaires.

La suite de l'article est organisée de la manière suivante. Nous réalisons une revue de la littérature dans une seconde section en insistant sur les limites inhérentes aux travaux empiriques réalisés jusqu'alors. Nous présentons dans une troisième section les données avant de préciser la méthodologie mobilisée dans l'exploration de nos deux bases de données. Dans une quatrième section, nous donnons des premières statistiques descriptives et les résultats clés issus des traitements économétriques réalisés. Nous concluons et discutons ces résultats dans une dernière section.

2. Revue de la littérature

Il existe depuis de nombreuses années, un éternel débat autour de l'importance de la proximité dans le cadre des collaborations de recherche : « the relevance of proximity is one of the most controversially discussed topics in the context of innovative linkages and networks » (Sternberg, 1999, p.553). Knoblen et Oerlemans (2006), dans une revue de la littérature sur le sujet soulignent l'intérêt particulièrement important accordé à la proximité géographique (« geographical proximity (...) is the most frequently used dimension of proximity in the literature » (p. 74)) et recensent 19 travaux liant proximité géographique et collaborations.

Ce débat repose essentiellement sur la thèse évolutionniste développée dans les années 1980 selon laquelle les connaissances tacites échangées lors du processus de collaboration, ne sont, par définition, pas codifiables et ne peuvent donc être diffusées facilement dans l'espace. Ceci a conduit certains auteurs (Verspagen, 1999 ; Patel et Pavitt, 1994 notamment) à considérer que seuls des rapports de face à face permettent d'en bénéficier. Foray (1995), par exemple, affirme que « l'aptitude à assimiler et à transférer les connaissances scientifiques et technologiques qui ne sont pas totalement codifiées est largement tributaire des opportunités de contact personnel direct entre les parties concernées » (p.19). Bien qu'il ait ainsi été communément accepté que la proximité géographique favorise le développement des collaborations (grâce à la réduction de l'ensemble des coûts de transaction), « popular belief holds that geography no longer matters in scientific collaborations with the arrival of cheap air travel, English as a global language and the internet » (Frenken and al., 2010). La proximité géographique constituerait alors de moins en moins une condition nécessaire ni suffisante pour la réalisation de partenariats de recherche (Boschma, 2005 ; Gilly and Torre, 2000 ; Bunnell et Coe, 2001 ; Breschi and Lissoni, 2003 ; Gertler, 2003).

À un niveau empirique, malgré le développement de travaux cherchant à mesurer l'importance de la proximité géographique, « little is known about the role of geographical proximity in science collaborations » (Ponds et al., 2007) et « the death of distance hypothesis has not been proven » (Frenken et al., 2010). Les résultats obtenus divergent fortement selon les travaux : alors que certains attestent de l'augmentation dans le temps des partenariats de longue distance (Frenken et

al., 2009 ; Narin et al., 1991 ; Luukkonen et al., 1993 ; Frenken, 2002 ; Magioni et Uberti, 2007 ; Wagner et Leydesdorff, 2005) d'autres, au contraire, montrent que la distance géographique joue encore aujourd'hui un rôle significatif dans la création de collaborations européennes de recherche (Autant-Bernard et al. 2007 ; Maggioni et Uberti 2009 ; Scherngell et Barber 2009). Egalement et surtout, les travaux empiriques existants souffrent de trois grandes limites.

1) Une première limite provient de la qualité des données et plus particulièrement des indicateurs de collaborations mobilisés. Les modèles économétriques relevant de la « géographie de l'innovation » (Feldman, 1994) - et notamment ceux ayant initialement rendu compte de la forte polarisation des activités d'innovation - s'appuient sur des données agrégées tels que les brevets, le volume des dépenses, le nombre de salariés en R&D, etc. Dans ces travaux, les interactions entre acteurs ne sont donc pas véritablement observées mais seulement supposées et les modèles rendent alors compte de la co-présence d'acteurs et non de véritables collaborations. D'autres travaux, plus récents, mesurent les collaborations au travers des citations ou de la participation à des projets communs (Autant-Bernard et al., 2007 ; Scherngell and Barber, 2009 ; D'Este and Iammarino, 2010). Bien que ces données permettent de mieux appréhender les interactions entre acteurs que les précédentes, la réalisation d'une véritable collaboration entre les partenaires n'est que supposée : dans certains types de collaborations, comme c'est le cas des partenariats européens notamment (EU Framework Program), le nombre de partenaires est tel qu'il n'est pas possible que l'ensemble des partenaires ait véritablement collaboré lors du projet. Il apparaît ainsi que « very few papers use data on direct collaborative agreements on R&D » (Massard et Mehier, 2009). Seuls les travaux reposant sur des données de co-brevets (Maggioni et al., 2007) ou co-publications (Katz 1994; Ponds et al. 2007 ; Frenken et al., 2010 ; Hoekman et al., 2010 ; Scherngell and Hu, 2011) ou plus encore ceux portant sur des relations dyadiques de collaboration (Laursen et al., 2011 ; Levy et al., 2009 ; Grossetti and Nguyen, 2001) peuvent véritablement rendre compte d'une relation de collaboration entre les acteurs et constituent à notre sens de véritables bases de données relationnelles.

2) Un second problème apparaît quant à l'indicateur de proximité géographique mobilisé. « The definition of this dimension of proximity differs slightly between different authors (...) with regard to the scale at which geographical proximity is defined » (Knoben and Oerlemans, 2006, p.74). Certains auteurs évaluent la proximité grâce à une variable binaire distinguant les collaborations nationales des collaborations internationales (Ponds et al., 2007) ; d'autres considèrent que la proximité géographique existe au delà des frontières régionales, et l'évaluent au niveau NUTS 2 (Scherngell et Barber, 2009 ; Hoekman et al. (2010)) ou grâce à une mesure ordinale (Drejer et Vinding (2007) distinguent par exemple les collaborations intra-régionales, intra-nationales et internationales). Appréhendée comme telle, la proximité géographique apparaît trop large pour pouvoir vérifier l'hypothèse selon laquelle la proximité rend les interactions de face à face plus aisées. Le niveau régional (niveau NUTS 3) de même que l'utilisation d'un indicateur de distance kilométrique maximum (inférieure à 100 miles pour Laursen et al., 2011 par exemple) apparaissent plus pertinents bien qu'ils peuvent s'avérer encore trop large pour certains territoires. Rares sont pourtant les travaux qui à notre connaissance permettent une analyse de la proximité à un niveau plus fin (département, communes). De la même manière, certains travaux ont montré l'importance de la proximité géographique en testant le rôle des frontières administratives (Hoekman et al., 2010) mais très peu ont pris en considération les spécificités relatives aux territoires d'étude tels que les « effets capitales » ou les « effets régions limitrophes », à l'exception de Grossetti and Nguyen (2001) qui soulignent l'hypertrophie de la région Ile-de-France dans les collaborations françaises. Enfin et

surtout, compte tenu du développement des moyens de transport, il nous semble important de tester des indicateurs de distance relative, en tenant compte par exemple du temps ou du coût de transport. Finalement, il apparaît que « the geographical dimension is in complex ways related to other mechanisms that are still barely identified and that take place at different geographical levels » (cf. Autant-Bernard et al., 2007), montrant la nécessité de confronter différents indicateurs de proximité géographique.

3) Enfin et surtout, « many studies on proximity and innovation have been analysed from a static perspective » (Boschma, 2005). Très peu de travaux reposent sur des données longitudinales et ils ne permettent donc généralement pas de rendre compte de l'évolution de la proximité géographique à l'exception de Hoekman et al. (2010) et Grossetti et Nguyen (2001). Les premiers soulignent pour 33 pays européens à la fois la persistance des collaborations locales (appréhendées au niveau NUTS 2) « and collaborations within territorial borders did decrease » entre 2000 et 2007 ; les seconds montrent l'augmentation des collaborations locales entre 1986 et 1998 pour la France.

Sur la base de ces différents travaux et des limites qui leur sont inhérentes, nous cherchons à compléter la littérature existante grâce à l'exploration de deux véritables bases de données relationnelles sur longue période et de la confrontation de différents indicateurs de proximité géographique.

3. Données et méthode

3.1 données

Nous mobilisons dans cet article deux bases de données complémentaires : la base CIFRE et la base CNRS. La première rassemble l'ensemble des contrats science-industrie réalisés dans le cadre de Conventions Industrielles de Formation par la Recherche (CIFRE) entre 1981 et 2005. Le dispositif CIFRE subventionne depuis 1981 toute entreprise de droit français qui embauche un doctorant pour le placer au cœur d'une collaboration de recherche avec un laboratoire public. Ces contrats de recherche sont en grande partie financés par le ministère de l'Enseignement supérieur et de la Recherche. La seconde base de données rassemble différentes informations relatives aux contrats de recherche réalisés sur la période 1986-2006 entre des firmes françaises et des laboratoires du Centre National de la recherche Scientifique (CNRS), la principale agence de recherche gouvernementale en France.

[TABLEAU 1]

Pour chaque base, nous disposons d'informations sur l'année de signature du contrat (ANNEE), la localisation régionale du laboratoire impliqué (REGLAB), la localisation régionale de l'entreprise (REGETAB), le secteur d'appartenance de l'entreprise (SECTEUR) et le domaine scientifique du laboratoire (DOMAINE). S'agissant de ces deux dernières variables, un problème important se pose car les nomenclatures mobilisées par l'ANRT ne sont pas celles habituellement utilisées. Nous avons donc dû agréger les catégories disponibles dans la base CNRS et construire une matrice de passage entre les nomenclatures ANRT et CNRS. Au final, nous pouvons ventiler les contrats en trois secteurs et six domaines scientifiques. Nous disposons d'une base agrégée de 32 764 observations totalement renseignées, dont 23 954 contrats CNRS et 8 810 contrats CIFRE, pour la période 1985-2006. Le tableau 1 présente les variables qualitatives disponibles dans les deux bases de données mobilisées.

3.2 Méthode

Afin d'analyser l'évolution de la proximité géographique des collaborateurs, nous proposons d'adopter deux stratégies complémentaires.

La première stratégie consiste à définir une typologie (TYPO) des échelles spatiales mobilisées, composée de cinq classes :

- * TYPO=DEP lorsque le contrat concerne deux partenaires localisés dans le même département,
- * TYPO=REG quand les deux partenaires sont localisés dans deux départements différents de la même région,
- * TYPO=LIMITROPHE si les deux partenaires sont localisés dans deux régions limitrophes,
- * TYPO=ICAP quand l'un des partenaires est localisé dans la région capitale (l'Ile-de-France), l'autre partenaire est localisé dans une autre région,
- * TYPO=IREG quand les deux partenaires sont localisés dans deux régions différentes hors Ile-de-France.

Cette typologie permet selon nous de mieux appréhender la proximité géographique que cela n'est généralement fait dans les études empiriques. Quatre des cinq classes permettent de définir des degrés de moins en moins fort de proximité géographique mesurée en termes de distance physique : les relations au sein d'un même département sont les plus proches géographiquement (DEP), suivies par celles au sein d'une même région (REG), puis celles entre régions limitrophes (LIMITROPHE) et enfin celles entre régions différentes hors capitale (IREG). La dernière classe de la typologie (ICAP) permet de tenir compte du caractère spécifique de la géographie française, la région capitale étant au cœur du réseau d'infrastructure de transport du pays, qu'il s'agisse du réseau routier, ferré ou aérien.

La deuxième stratégie, plus classique, consiste à se doter d'une mesure relative de la proximité géographique entre acteurs, à savoir le temps de déplacement, en train, entre les deux capitales des régions où sont localisés les acteurs (variable TRAIN). Cette mesure nous semble plus satisfaisante que la distance kilométrique entre les capitales régionales, souvent utilisée. Une alternative encore plus satisfaisante aurait bien sûr été de mesurer la distance non pas entre les capitales régionales mais entre les localisations précises des acteurs. Les données mobilisées ne sont malheureusement pas géo-localisées, ce qui nous interdit de procéder ainsi.

Méthodologiquement, nous proposons de présenter, dans un premier temps, différentes statistiques descriptives permettant de capturer l'évolution temporelle de la proximité géographique, en distinguant l'ensemble des contrats CNRS des contrats CIFRE. Nous chercherons à vérifier l'existence ou l'absence de trends temporels, en réalisant des tests non paramétriques (tau de Kendall) entre chaque variable et le temps t. Nous proposons ensuite d'estimer deux modèles économétriques complémentaires. Le premier modèle est un logit multinomial avec pour variable expliquée la variable TYPO. Le deuxième modèle est un modèle de Heckman avec équation de sélection : la première équation (équation de sélection) estime la probabilité que la collaboration entre deux acteurs soit interrégionale (INTER=1) plutôt qu'infrarégionale (INTER=0), la deuxième équation (équation d'output) estime quant à elle la distance-temps, par train, entre les deux collaborateurs.

Dans les deux modèles, les variables explicatives sont les mêmes : i) l'année du contrat (ANNEE), ii) la nature du contrat (CIFRE vs. CNRS), iii) le secteur d'appartenance de l'entreprise (INDUSTRIE,

SERVICES, AUTRES), iv) le domaine scientifique d'appartenance du laboratoire (PHYSIQUE-MATHEMATIQUES, CHIMIE, SCIENCES DE L'UNIVERS, SCIENCES DE LA VIE, SCIENCES HUMAINES ET SOCIALES, SCIENCES POUR L'INGENIEUR, SCIENCES DE L'INFORMATION).

4. Résultats

4.1. Statistiques descriptives

Afin de mesurer l'évolution de la proximité géographique sur la période d'étude, nous avons dans un premier temps construit trois graphiques complémentaires (cf. graphique 1a, 1b, 1c). Le premier représente, pour l'ensemble des contrats CNRS, l'évolution de la part de chaque modalité de notre typologie. Le deuxième graphique procède de même pour l'ensemble des contrats CIFRE. Le troisième mesure l'évolution de la moyenne de la distance-temps en les deux capitales régionales où sont localisés les partenaires.

[GRAPHIQUE 1a, 1b, 1c]

Plusieurs résultats méritent d'être soulignés.

S'agissant des contrats CNRS (graphique 1a), le résultat le plus remarquable est la baisse importante de la part des contrats interrégionaux impliquant la région capitale (ICAP) : elle passe, sur la période, de près de 60% à environ 40%, soit une baisse de vingt points en une vingtaine d'années. Comment expliquer cette forte baisse ? Une hypothèse possible, selon nous, tiendrait à la déconcentration spatiale des chercheurs CNRS, qui aurait conduit à un rééquilibrage au profit des régions de province et au détriment de la région capitale. Nous ne disposons hélas pas des données nécessaires à la validation de cette hypothèse, elle est cependant conforme aux résultats obtenus, sur le début de notre période d'observation, par Grossetti et Nguyen (2001). On remarque également un accroissement de la part des contrats CNRS au sein d'un même département (DEP), d'environ 10% à plus de 20%. Les autres modalités de la typologie semblent plus stables.

Comparativement, l'évolution observée pour les contrats CIFRE (graphique 1b) témoigne d'une très forte inertie de la géographie des collaborations, la part de chaque modalité semblant particulièrement stable. La comparaison des deux graphiques (1a et 1b) montre donc que l'évolution de la géographie des collaborations est très différente selon la nature des contrats science-industrie (Cifre vs. CNRS). Toutefois, si l'on se centre sur les données de fin de période, on constate que le classement des différentes modalités de la typologie est le même pour les deux types de contrats : ce sont les relations interrégionales avec la région capitale qui dominant, suivies des relations intra-départementales puis des relations infrarégionales entre départements.

Le graphique (1c) reprenant l'évolution des moyennes des distances-temps apporte également des éclairages intéressants. Si l'on se concentre tout d'abord sur l'ensemble des contrats (courbes CIFRE_TOTAL et CNRS_TOTAL), on observe que les contrats CIFRE sont mis en œuvre avec des partenaires plus proches géographiquement que les contrats CNRS. Ceci peut être vu comme une confirmation de l'hypothèse formulée en introduction : dans le cadre des contrats CIFRE, le doctorant doit alterner des périodes en entreprise et des périodes au sein du laboratoire, une plus grande proximité géographique entre acteurs peut être recherchée afin de réduire les coûts de déplacement - et plus généralement l'ensemble des coûts de transaction - que le doctorant doit supporter. S'agissant ensuite du sous-ensemble des contrats interrégionaux, les distances-moyennes ne diffèrent qu'assez peu entre les deux types de contrats. On note cependant une différence

importante : la distance moyenne semble stable pour les contrats CIFRE alors qu'elle est croissante pour les contrats CNRS.

Afin d'éprouver la validité des évolutions observés, nous avons cherché à identifier l'existence de trends croissants ou décroissants. Compte-tenu du faible nombre de points de notre échantillon (20 années), nous avons opté pour des tests non paramétriques de corrélation de rang (tau de Kendall).

[TABLEAU 2]

Les résultats observés confirment la lecture des graphiques. On observe une forte inertie pour les contrats CIFRE puisque seule la modalité REG présentant un trend faiblement significatif. S'agissant des contrats CNRS, toutes les modalités présentent des trends significatifs. Les coefficients les plus forts sont observés pour les relations intra-départementales (trend croissant) et pour les relations avec la région capitale (trend décroissant). L'analyse des distances-temps sur la période d'étude rend compte également d'un trend croissant pour les relations interrégionales dans le cadre de partenariats CNRS. Les tests non paramétriques de corrélation de rang confirment ainsi l'existence d'une dynamique spatiale différente selon la nature du contrat : une relative stabilité de la géographie des contrats CIFRE vs. Une évolution significative pour les contrats CNRS.

4.2. Estimation des modèles économétriques

Afin de confirmer ou d'infirmer ces premiers résultats en intégrant d'éventuels effets sectoriels (secteurs d'activité côté entreprise et domaines scientifiques côté laboratoire), nous avons procédé à l'estimation des deux modèles économétriques présentés plus haut. Le premier modèle est un logit multinomial, avec comme variable expliquée la variable TYPO. La modalité de référence du logit multinomial est la variable DEP, qui prend la valeur 1 si le contrat est intra-départemental et la valeur 0 sinon.

[TABLEAU 3]

S'agissant de l'évolution de la géographie des collaborations, on remarque que la variable ANNEE est significative pour trois des échelles spatiales (REG, ICAP et IREG), avec un coefficient de signe négatif pour les deux premières et positif pour la dernière. Ceci confirme donc la double évolution de la géographie des collaborations : accroissement de la part des collaborations locales et accroissement de la part des collaborations interrégionales les plus lointaines. S'agissant de la nature des collaborations, la variable CIFRE est toujours significative et les coefficients sont systématiquement négatifs. Ces résultats démontrent le caractère beaucoup plus local des contrats CIFRE, ce que l'on peut expliquer, comme nous l'avons déjà dit, par la contrainte qui pèse sur le doctorant d'alterner périodes en entreprise et périodes en laboratoire.

On observe également des effets sectoriels. S'agissant des domaines scientifiques, les sciences de l'univers ont des coefficients systématiquement négatifs et significatifs, ce qui signifie qu'il s'agit d'un domaine auquel on peut associer des contrats beaucoup plus locaux que pour le domaine de référence (les mathématiques et la physique). A contrario, les sciences chimiques ont des coefficients positifs pour trois des quatre modalités, ce qui témoigne d'une géographie moins locale. Autre constat : les sciences humaines et sociales ont un coefficient négatif très significatif pour les modalités IREG et ICAP, autrement dit pour les collaborations les plus lointaines. Ce domaine scientifique est donc favorable aux collaborations plutôt locales, soit à l'échelle des départements, de la région ou des régions limitrophes. Pour les sciences de la vie, les résultats sont plus contrastés : coefficients négatifs pour REG et ICAP, coefficients positifs pour IREG et LIMITROPHE, signes de

tendances contradictoires en termes d'évolution de la géographie des collaborations pour ce domaine scientifique. Les deux derniers domaines (sciences pour l'ingénieur et sciences de l'information) présentent peu de coefficients significatifs.

S'agissant des secteurs d'activité, le résultat essentiel concerne le secteur des services, comparativement au secteur de référence (l'industrie). En effet, le secteur des services présente des coefficients systématiquement significatifs (au seuil de 1%) et négatifs. Par rapport à l'industrie, ce secteur influe donc très fortement sur la probabilité que le contrat soit local (à l'échelle du département). Pour l'ensemble plus hétérogène des autres secteurs, les coefficients sont également significatifs, trois sont négatifs (REG, IREG, LIMITROPHE) et l'un est positif (ICAP).

Le deuxième modèle est un *sample selection model* (two-step Heckman), avec comme variable expliquée, dans l'équation de sélection, le caractère interrégionale vs. Infrarégionale de la collaboration et, dans l'équation d'output, la distance-temps par train entre les capitales des régions où sont localisés les acteurs.

[TABLEAU 4]

Le modèle apparaît tout d'abord globalement satisfaisant. Notons également que le ratio de Mills est significatif, ce qui valide le choix du modèle économétrique adopté. Plusieurs résultats intéressants plus précis sont à souligner. S'agissant des effets temporels, d'abord, on remarque un effet négatif de la variable ANNEE sur la probabilité que le contrat soit interrégional, preuve de la tendance au renforcement du rôle de la proximité géographique sur la période, ce que nous avons déjà rendu compte précédemment. Toutefois, dans le même temps, on observe grâce à ce modèle que cette variable exerce un effet positif sur la distance entre les partenaires pour les contrats interrégionaux. On retrouve donc un résultat essentiel qui traverse notre article : l'évolution de la géographie des collaborations semble se traduire conjointement par un renforcement à la fois des relations les plus locales et les plus distantes.

La variable CIFRE exerce la même influence que la variable ANNEE : une influence négative très significative sur la probabilité que le contrat soit interrégional et dans le même temps une influence positive, mais faiblement significative, sur la distance entre les partenaires interrégionaux.

S'agissant des effets sectoriels, ils sont peu nombreux et faiblement significatifs dans l'équation d'output, à l'exception du secteur d'activité « autres », dont le résultat est cependant difficile à interpréter. En revanche, plusieurs effets très significatifs sont à noter dans l'équation de sélection : les sciences humaines et sociales, côté domaine scientifique, et les services, côté secteur d'activité, exercent une influence négative très significative sur la probabilité que le contrat soit interrégional ; la chimie et les sciences pour l'ingénieur, à l'inverse, exerce une influence positive très significative sur cette même probabilité.

5. Conclusion

L'intérêt de cet article était d'apporter de nouveaux résultats empiriques sur le rôle de la proximité géographique dans le cadre des collaborations science-industrie et sur l'évolution de son rôle au cours du temps. En nous appuyant sur deux bases de données relationnelles de longue période et

grâce à la construction de deux indicateurs complémentaires de proximité géographique, nous avons pu réaliser une analyse comparative et dynamique des collaborations science-industrie.

Nous avons ainsi pu répondre à la question « la proximité géographique compte-t-elle toujours? » et apporter un résultat véritablement nouveau : les tests économétriques réalisés montrent en effet que la proximité continue de compter encore aujourd'hui, ce qui se traduit notamment par le poids non négligeable des collaborations intra-départementales (supérieur à 20%), stable pour les contrats CIFRE et croissant pour les contrats CNRS. La perspective dynamique nous permet de préciser ce résultat et révèle que dans le même temps les collaborations interrégionales les plus distantes croissent sur la période d'étude. On a donc finalement la conjonction d'un renforcement des collaborations les plus locales et des collaborations les plus distantes. Une telle dynamique spatiale des collaborations avec des évolutions dans des sens *a priori* antagonistes avait déjà été observée par Hoekman et al. (2010) au niveau européen.

Un second résultat important apparaît : on observe des évolutions significativement différentes selon la nature des contrats, le rôle de la proximité géographique apparaît plus stable et plus significatif dans le cadre des contrats CIFRE que pour les contrats CNRS. Ce résultat montre la nécessité de distinguer les collaborations selon leurs caractéristiques. Non seulement le rôle de la proximité géographique diffère selon la nature (science-industrie vs. inter-entreprises) de la collaboration comme l'ont montré Ponds et al. (2007) mais également, à l'intérieur même des collaborations science-industrie, selon le type de contrat (CIFRE vs. CNRS). Notre travail confirme enfin l'existence de différences sectorielles, déjà observées par Ponds et al. (2007) pour les domaines scientifiques.

Au final, l'analyse de la proximité géographique sur longue période révèle des évolutions plus contrastées que ce que l'on s'imaginait généralement. Ceci a des implications fortes en termes de politiques publiques : ces dernières ont généralement tendance à s'appuyer sur des régularités supposées transversales, ce qui n'est pas le cas. Le calibrage des politiques doit donc être particulièrement fin, car les résultats dépendent des caractéristiques précises des collaborations comme nous avons pu le montrer. Il convient toutefois de rappeler que les résultats obtenus concernent la France et qu'il conviendrait évidemment de les reproduire pour d'autres pays, afin de d'évaluer leur degré de généralité.

6. Bibliographie

- AUTANT-BERNARD C., BILLAND P. and MASSARD N. (2007) Social distance versus spatial distance in R&D cooperation: empirical evidence from european collaboration choices in micro and nanotechnologies, *Papers in Regional Science* 86 (3), 495-519 ;
- BOSCHMA R. (2005) Proximity and innovation: a critical assessment, *Regional studies* 39 (1), 61-74;
- BUNNELL T. and COE N. (2001) Spaces and scales of innovation, *Progress in Human Geography* 25, 569-589 ;
- BRESCHI S. and LISSONI F. (2003) Mobility and social networks: Localized knowledge spillovers revisited, CESPRI working paper nr. 142, March
- CAIRNCROSS F. (1997) *The Death of Distance: How the Communications Revolution will Change Our Lives*, Harvard Business School Publishing, Cambridge.

- CANIELS M. (1997) The geographic distribution of patents and value added accross european regions, *Working Paper, Merit, aout*, 10 ;
- CASTELLS M. (1996) *The Rise of the Network Society*, Blackwell Publishers, Oxford.
- D'ESTE P. and IAMMARINO S. (2010) The spatial profile of university-business research partnerships, *Papers in Regional Science* 86, 423-443;
- DREJER I. and VINDING A. (2007) Searching Near and Far: Determinants of Innovative Firms' Propensity to Collaborate Across Geographical Distance, *Industry and Innovation* 14(3), 259-275;
- FELDMAN M. (1994) *The geography of innovation*, Kluwer Academics Publishers, Dordrecht.
- FORAY D. (1995) Distribution et expansion de la base de connaissances scientifiques et technologiques, *Revue STI, OCDE*, 14-71 ;
- FRENKEN K. (2002) A new indicator of european integration and an application to collaboration in scientific research, *Economic Systems Research* 14(4), 345-361 ;
- FRENKEN, K., PONDS, R., VAN OORT, F.G. (2010) The citation impact of research collaboration in science-based industries: a spatial-institutional analysis, *Papers in Regional Science* 89(2), 351-371 ;
- FRENKEN K., HARDEMAN S. and HOEKMAN J. (2009) Spatial scientometrics: towards a cumulative research program, *Journal of Informetrics* 3 (3), 222–232;
- GERTLER M.S. (2003) Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there), *Journal of Economic Geography*, 3, 75-99;
- GILLY JP. and TORRE A. (2000) *Dynamique de proximité*, Ed. L'Harmattan, Paris.
- GIURI P. and MARIANI M. (2007) Inventors and the geographical breadth of knowledge spillovers, *druid working paper 08-01*.
- GROSSETTI M. and 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 Régionale et Urbaine* 2, 311-328 ;
- HOEKMAN J., FRENKEN K. and TIJSSEN R. (2010) Research collaboration at a distance: changing spatial patterns of scientific collaboration within Europe, *Research Policy* 39 (5), 662-673;
- LAURSEN K., REICHSTEIN T. and SALTER A. (2011) Exploring the Effect of Geographical Proximity and University Quality on University-Industry Collaboration in the United Kingdom, *Regional Studies*, 45: 4, 507-523;
- KATZ J.S. (1994) Geographic proximity and scientific collaboration, *Scientometrics*, 31 (1), 31-43 ;
- KNOBEN J. and OERLEMANS L. (2006) Proximity and inter-organizational collaboration: a literature review, *International Journal of Management Review* 8 (2), 71–89;
- LALLEMENT R., MOUHOUD EM. and PAILLARD S. (2007) Polarisation et internationalisation des activités d'innovation : incidences sur la spécialisation technologique des nations », In A. Rallet and A. Torre (dir), *Quelle proximité pour innover ?*, Ed. L'Harmattan.

- LAURSEN, K., REICHSTEIN, T. and SALTER A. (2010) Exploring the effect of geographical proximity and university quality on university-industry collaboration in the United Kingdom, *Regional Studies* 45 (4), 507-523;
- LEVY R., ROUX P. and WOLFF S. (2009) Study of science-industry collaborative patterns in a large european university, *Journal of Technology Transfer* 34 (1), 1-23;
- LUBLINSKI E. (2003) Does geographic proximity matter? Evidence from clustered and non-clustered aeronautic firms in Germany, *Regional Studies*, 37(5), 453-467
- LUUKKONEN T., TIJSSEN R.J.W., PERSSON O., SIVERTSEN G. (1993) The measurement of international scientific collaboration, *Scientometrics* 28, 15-36 ;
- MAGGIONI A. and UBERTI T. (2009) Knowledge networks across europe: which distance matters?, *The Annals of Regional Science* 43(3), 691-720 ;
- MAGGIONI M., NOSVELLI M. and UBERTI T. (2007) Space versus networks in the geography of innovation, *Papers in Regional Science* 86 (3), 471-493;
- MAGGIONI M. and 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.
- MASSARD N. and MEHIER C. (2009) Proximity and innovation through an « accessibility to knowledge » lens », *Regional Studies* 43 (1), 77-88 ;
- MORGAN K. (2004) The exaggerated death of geography: learning, proximity and territorial innovation systems, *Journal of Economic Geography*, 4(1), 3-21;
- NARIN, F., STEVENS, K., WHITLOW, E.S. (1991) Scientific cooperation in europe and the citation of multinationally authored papers, *Scientometrics* 21 (3), 313-323 ;
- OINAS P. (2000) Distance and learning: does proximity matter? In s. bakkers, R.P.J.H. RUTTEN, K. MORGAN, AND F. BOEKEMA (EDS), *Learning regions, theory, policy and practice*, ALDERSHOT, 57-69 ;
- PATEL, E and K. PAVITT (1994) Uneven (and divergent) technological accumulation among advanced countries: evidence and a framework for explanation, *Industrial and Corporate Change* 3, 739-787 ;
- PUGA D. (1999) The rise and fall of regional inequalities, *European Economic Review* 43 (2), 303-334;
- PONDS R., VAN OORT F. and FRENKEN K. (2007) The geographical and institutional proximity of research collaboration, *Paper in Regional Science* 86 (3), 423-443;
- SCHERNGELL T. and HU Y. (2011) Collaborative Knowledge Production in China: Regional Evidence from a Gravity Model Approach, *Regional Studies*, 45(6), 755-772;
- STERNBERG R. (1999) Innovative Linkages and Proximity: Empirical Results from Recent Surveys of Small and Medium Sized Firms in German Regions, *Regional Studies* 33 (6), 529-540;
- VERSPAGEN B. (1999) Europe : one or several systems of innovation ?, IN J. FAGERBERG, P. GUERRIERI and B. VERSPAGEN (eds), *The economic challenge to europe*.

WAGNER C. and LEYDESDORFF L. (2005) Network structure, self-organization, and the growth of international collaboration in science, *Research Policy* 34 (10), 1608-1618;

Figure 1a : Evolution de la part des différentes échelles spatiales (contrats CNRS)

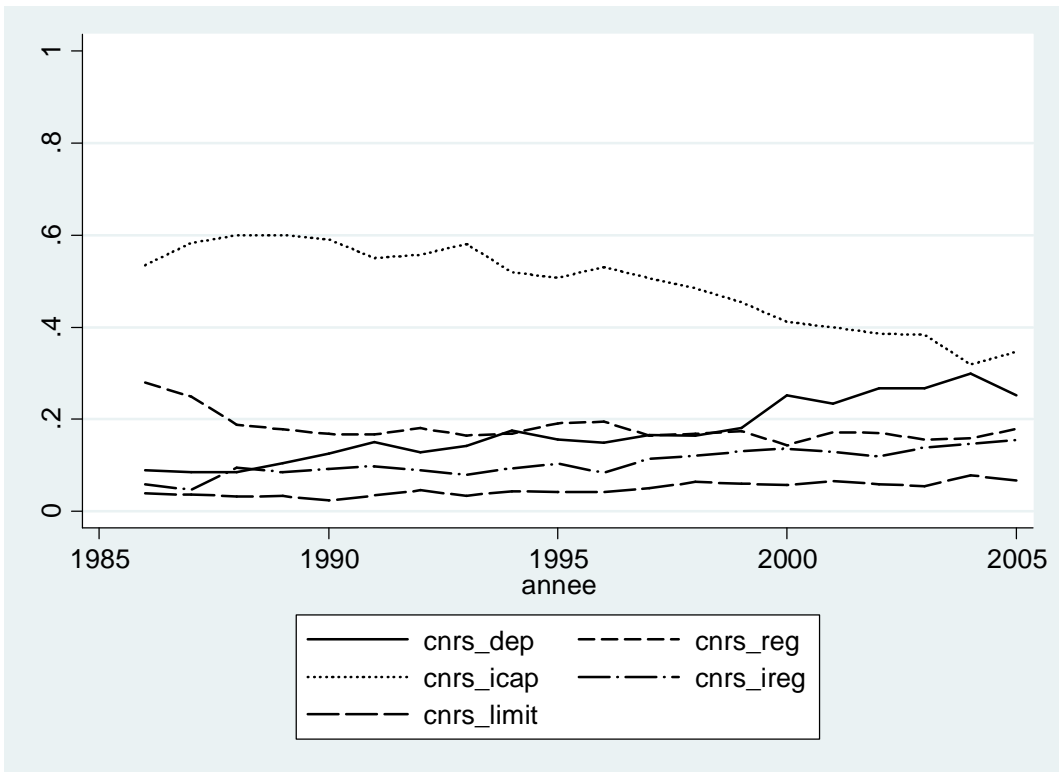


Figure 1b : Evolution de la part des différentes échelles spatiales (contrats CIFRE)

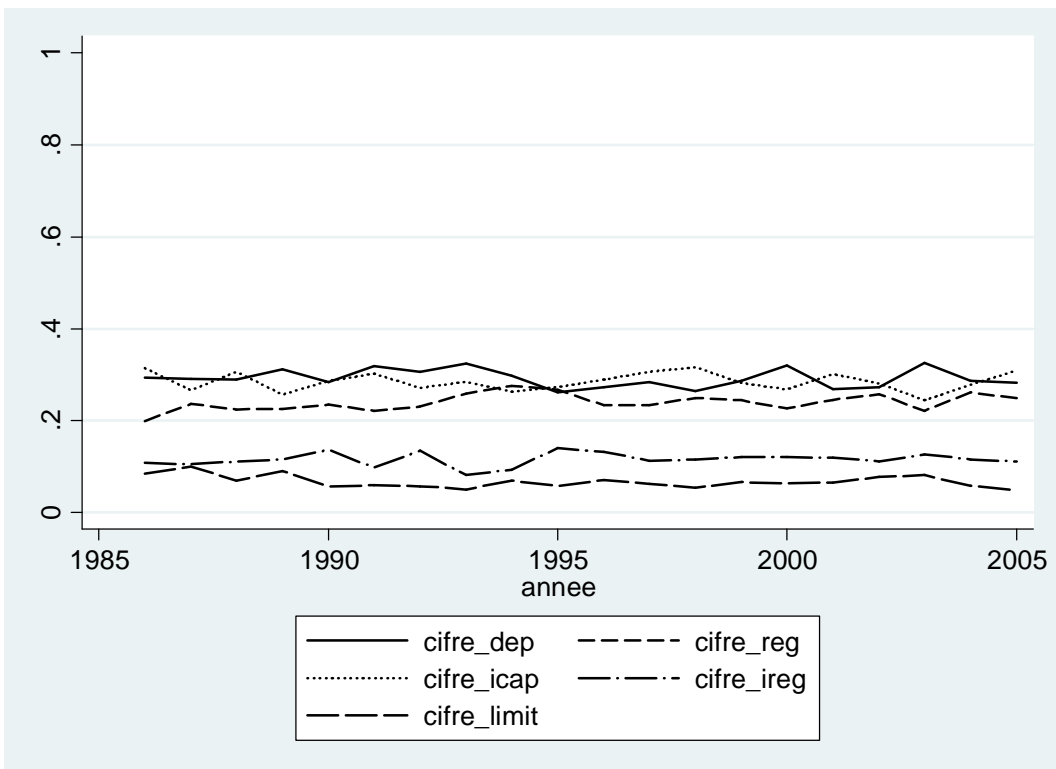


Figure 1c : évolution des distances-temps par train

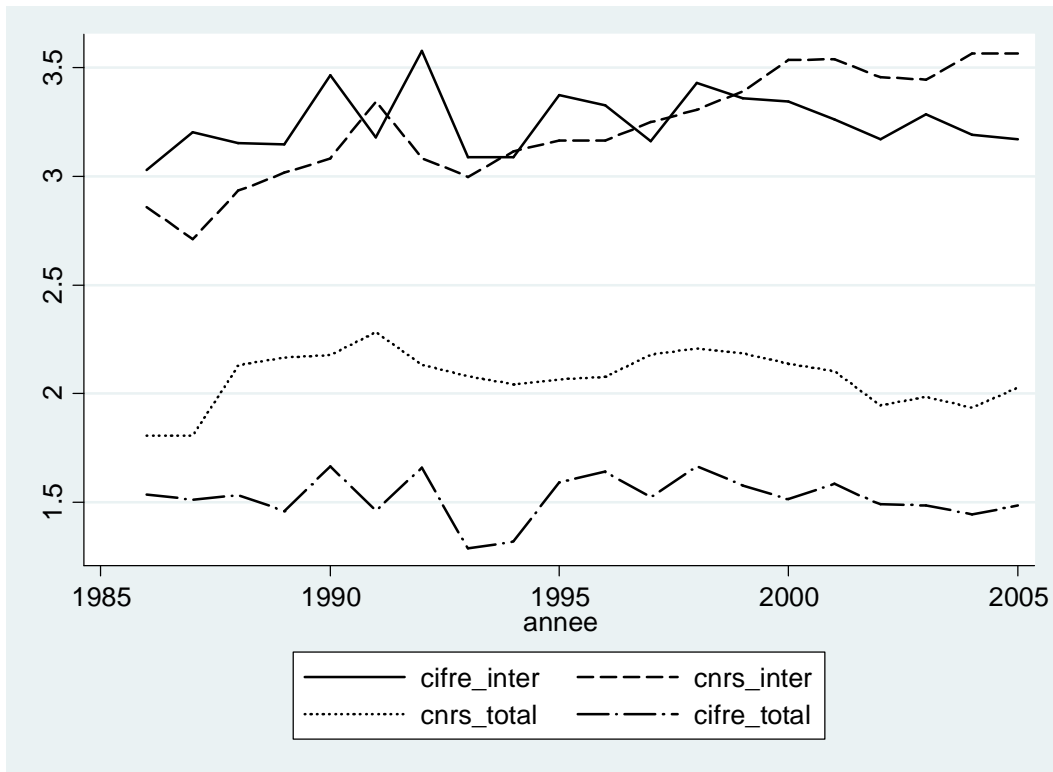


Tableau 1 : présentation des données disponibles

Variable	Modalités	effectifs	part
Nature du Contrat	CNRS	23954	73,1%
	CIFRE	8810	26,9%
	Total	32764	100,0%
Domaine scientifique	Physique, maths (pnc)	3330	10,2%
	chimie (sch)	8588	26,2%
	Sciences de l'univers (sdu)	2122	6,5%
	Sciences de la vie (sdv)	5512	16,8%
	Sciences humaines et sociales (shs)	2179	6,7%
	Sciences pour l'ingénieur (spi)	6334	19,3%
	Sciences de l'information (sti)	4699	14,3%
	Total	32764	100,0%
Secteur d'activité	Industrie	16142	49,3%
	Services	11816	36,1%
	Autres	4806	14,7%
	Total	32764	100,0%
Typologie des échelles spatiales	contrats intra-départementaux (DEP)	7277	22,2%
	contrats infrarégionaux (REG)	6302	19,2%
	Contrats entre régions limitrophes (LIMITROPHE)	1846	5,6%
	Contrats avec la région capitale (ICAP)	13576	41,4%
	Autres contrats interrégionaux (IREG)	3763	11,5%
	total	32764	100,0%

Tableau 2 : identification des trends temporels

variable	niveau	Kendall's tau
cnrs	dep	0,842 ***
cnrs	reg	-0,358 **
cnrs	limitrophe	0,653 ***
cnrs	icap	-0,821 ***
cnrs	ireg	0,726 ***
cifre	dep	-0,147
cifre	reg	0,295 *
cifre	limitrophe	-0,174
cifre	icap	-0,021
cifre	ireg	0,111
Cnrs (total)	Temps de train	-0,095
Cifre (total)	Temps de train	-0,126
Cnrs (inter)	Temps de train	0,832 ***
Cifre (inter)	Temps de train	0,105

Tableau 3 : résultats du logit multinomial (base outcome : DEP)

		Coef.	Std. Err.	z	p> z	signif.
REG	annee	-0,034	0,004	-9,58	0,00	***
	cifre	-0,378	0,042	-9,08	0,00	***
	sch	-0,096	0,065	-1,47	0,14	
	sdu	-0,634	0,091	-6,98	0,00	***
	sdv	-0,171	0,068	-2,54	0,01	**
	shs	-0,011	0,079	-0,14	0,89	
	spi	-0,590	0,067	-8,76	0,00	***
	sti	-0,070	0,068	-1,04	0,30	
	services	-0,662	0,041	-15,99	0,00	***
	autres	0,365	0,059	6,23	0,00	***
	_cons	67,941	7,048	9,64	0,00	***
LIMITROPHE	annee	-0,006	0,005	-1,15	0,25	
	cifre	-0,172	0,063	-2,72	0,01	***
	sch	0,241	0,103	2,35	0,02	**
	sdu	-0,499	0,148	-3,36	0,00	***
	sdv	0,324	0,103	3,14	0,00	***
	shs	-0,160	0,129	-1,24	0,22	
	spi	-0,070	0,105	-0,67	0,50	
	sti	-0,051	0,110	-0,47	0,64	
	services	-0,210	0,060	-3,49	0,00	***
	autres	-0,170	0,102	-1,67	0,10	*
	_cons	11,104	10,715	1,04	0,30	
ICAP	annee	-0,047	0,003	-15,42	0,00	***
	cifre	-1,186	0,038	-31,2	0,00	***
	sch	0,104	0,059	1,75	0,08	*
	sdu	-0,278	0,076	-3,67	0,00	***
	sdv	-0,181	0,062	-2,9	0,00	***
	shs	-0,214	0,076	-2,82	0,01	***
	spi	0,058	0,059	0,97	0,33	
	sti	0,096	0,062	1,54	0,12	
	services	-0,725	0,035	-20,46	0,00	***
	autres	0,755	0,051	14,78	0,00	***
	_cons	95,878	6,143	15,61	0,00	***
IREG	annee	0,009	0,004	2,13	0,03	**
	cifre	-0,421	0,050	-8,42	0,00	***
	sch	0,336	0,079	4,26	0,00	***
	sdu	-0,229	0,107	-2,14	0,03	**
	sdv	0,258	0,081	3,18	0,00	***
	shs	-0,638	0,116	-5,51	0,00	***
	spi	0,060	0,080	0,76	0,45	
	sti	-0,248	0,088	-2,82	0,01	***
	services	-0,665	0,047	-14,08	0,00	***
	autres	-0,339	0,077	-4,38	0,00	***
	_cons	-18,135	8,369	-2,17	0,03	**

Number of obs	32764	LR chi2(40)	3660,910
Log likelihood	-44921,268	Prob > chi2	0,000 ***

Tableau 4 : résultat du sample selection model (two-step Heckman)

	Coef.	Std. Err.	z	p> z	Signif.
<i>Output equation (explained variable : train)</i>					
annee	0,065	0,021	3,030	0,002	***
cifre	1,945	1,008	1,930	0,054	*
sch	-0,609	0,336	-1,810	0,070	*
sdu	0,289	0,233	1,240	0,215	
sdv	0,266	0,211	1,260	0,207	
shs	0,606	0,487	1,240	0,213	
spi	-0,720	0,474	-1,520	0,129	
sti	-0,364	0,203	-1,790	0,073	*
services	0,930	0,499	1,860	0,062	*
autres	-1,107	0,417	-2,650	0,008	***
_cons	-121,817	40,673	-3,000	0,003	***
<i>Selection equation (explained binary variable : inter)</i>					
annee	-0,009	0,001	-6,570	0,000	***
cifre	-0,452	0,018	-25,800	0,000	***
sch	0,133	0,027	4,900	0,000	***
sdu	-0,004	0,037	-0,110	0,915	
sdv	0,044	0,029	1,520	0,128	
shs	-0,168	0,036	-4,670	0,000	***
spi	0,205	0,028	7,420	0,000	***
sti	0,035	0,029	1,210	0,226	
services	-0,224	0,017	-13,320	0,000	***
autres	0,200	0,022	9,180	0,000	***
_cons	18,915	2,831	6,680	0,000	***
mills ratio	-7,073	3,735	-1,890	0,058	*
Number of obs.	32764		Wald chi2(10)	49,190	
Censored obs.	13579		Prob > chi2	0,000	***
uncensored obs.	19185				