

Direction des Études et Synthèses Économiques

G 2008 / 09

**Concentration géographique dans l'industrie
manufacturière et dans les services
en France :**

une approche par un indicateur en continu

Muriel BARLET - Anthony BRIANT - Laure CRUSSON

Document de travail



Institut National de la Statistique et des Études Économiques

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

*Série des documents de travail
de la Direction des Études et Synthèses Économiques*

G 2008 / 09

Concentration géographique dans l'industrie manufacturière et dans les services en France :

une approche par un indicateur en continu

Muriel BARLET* - Anthony BRIANT - Laure CRUSSON***

AOÛT 2008

Nous remercions Patrick Sillard pour son aide précieuse dans la géolocalisation d'une partie des données d'établissements et pour avoir accepté de discuter une première version de cet article présentée lors du séminaire D3E du 2 juin 2008. Nous remercions également Gilles Duranton et Laurent Gobillon pour leurs remarques et leurs conseils.

* Département des Études Économiques d'Ensemble - Division « Croissance et Politiques macroéconomiques »
Timbre G220 - 15, bd Gabriel Péri - BP 100 - 92244 MALAKOFF CEDEX

** École d'Économie de Paris / Paris-Jourdan Sciences Économiques (UMR 8545 CNRS-EHESS-ENPC-ENS), 48 Boulevard Jourdan, 75014 Paris, France et INSEE, DESE/DEEE/MSE, 15, Boulevard Gabriel Péri BP100 92244 Malakoff Cedex, anthony.briant@m4x.org

Concentration géographique dans l'industrie manufacturière et dans les services en France : une approche par un indicateur en continu

Résumé

Cet article mesure la concentration géographique dans les secteurs manufacturiers et dans les secteurs de services aux entreprises en France à partir d'une approche en continu développée par Duranton et Overman (2005). Cette approche consiste à comparer la distribution des distances bilatérales entre établissements d'un secteur à une distribution de référence construite pour un secteur avec le même nombre d'établissements répartis aléatoirement sur le territoire national, en tenant compte de la distribution de l'activité économique globale. Nous montrons que les résultats obtenus par ces auteurs, pour les industries manufacturières britanniques, s'étendent à la France. Nous trouvons notamment que, contrairement aux indicateurs discrets (Ellison-Glaeser, Maurel-Sédillot), l'indicateur en continu conclut moins souvent à la concentration des secteurs. Il indique cependant que les secteurs concentrés le sont à courte distance (inférieure à 50km). Comme pour les indicateurs discrets, il apparaît une forte hétérogénéité de la concentration d'un secteur à l'autre. Les secteurs de services sont en général plus concentrés que les secteurs manufacturiers et à plus courte distance. Enfin, sur la décennie 1996-2005, les schémas de concentrations sont relativement stables. Néanmoins, au sein de chaque industrie, les établissements disparaissant tendent à renforcer la concentration, alors les nouveaux entrants ont l'influence inverse.

Mots-clés : concentration sectorielle, K-densité, statistiques spatiales, processus spatiaux ponctuels

Patterns of Localisation in the French Manufacturing and Service Industries: a Distance-Based Approach

Abstract

We explore the patterns of localisation in the French manufacturing and service industries using the distance-based approach developed by Duranton and Overman (2005). The idea of this methodology is to consider the distribution of distances between pairs of plants in an industry and to compare it with that of a hypothetical industry with the same number of plants which are randomly distributed conditional on the distribution of aggregate activity. Relying on the same industrial classification as the one used by these authors, we show that their main conclusions remain valid in the case of French manufacturing industries. First, fewer manufacturing industries are found to be globally localised than usually found with discrete indices (Ellison-Glaeser, Maurel-Sédillot). Second, localisation takes place at small distance (< 50 km). Third, the degree of localisation varies drastically across industries. Furthermore, we find that service industries deviate more often from randomness than manufacturing industries, and that a higher proportion of them show global localisation. This localisation occurs at shorter distances than for manufacturing industries. Finally, these patterns of localisation remain fairly stable over the period 1996-2005. Within industries, we find that exiters are acting to reinforce localisation while entrants are acting in the opposite direction.

Keywords: localisation, K-density, spatial statistics, point-pattern processes

Classification JEL : C19, R12, L70

Table des matières

Motivation	5
1 Mesurer la concentration géographique d'un secteur	7
1.1 Les études sur la concentration géographique des secteurs en France : une revue critique	7
1.2 Illustration des problèmes posés par l'indicateur d'Ellison-Glaeser	10
1.3 L'indicateur en continu de Duranton et Overman (2005) : données géolocalisées et distances bilatérales entre établissements	12
2 Mise en oeuvre de l'indicateur en continu de Duranton et Overman (2005)	16
3 Concentration des activités manufacturières : comparaison France - Royaume-Uni	20
3.1 La géographie de la France	20
3.2 Analyse comparée France/Royaume-Uni	22
4 Analyse comparée de la concentration des secteurs manufacturiers et de services manufacturiers aux entreprises	30
5 Dynamique de concentration entre 1996 et 2005	35
5.1 Comparaison avec les différentes données	35
5.2 Évolution de la concentration entre 1996 et 2005	37
5.3 Évolution de la localisation de l'activité sectorielle	39
Conclusion et perspectives	42
Bibliographie	44
A Indices de concentration discrets	46
B Critères de qualité d'un bon indicateur de concentration	48
C Détails sur les données utilisées	50

Motivation

En France, les débats sur le choix de localisation des activités économiques sont généralement tournés vers des considérations d'aménagement du territoire et d'équité territoriale (Houdebine (1999), Lacour, Delamarre et Thoin (2008)). Jusqu'à une date récente, il était rare d'examiner ces questions en terme de compétitivité. Néanmoins, la réussite médiatisée des exemples de spécialisation locale tels que la Silicon Valley aux États-Unis, Sophia Antipolis en France ou bien les districts du Nord de l'Italie, ainsi que le développement de mesures empiriques plus précises des gains à la concentration, ont incité la sphère politique à s'approprier la question de l'organisation spatiale optimale des secteurs d'activités au travers de politiques industrielles dites de *clusters*.

Ainsi, en 2005, le gouvernement français a consacré 1,5 milliard d'euros (sur trois ans) à la mise en place des *pôles de compétitivité* visant, *sur un territoire donné, à l'association d'entreprises, de centres de recherche et d'organismes de formation, engagés dans une démarche partenariale (stratégie commune de développement), destinée à dégager des synergies autour de projets innovants conduits en commun en direction d'un (ou de) marché(s) donné(s)*.¹

D'un point de vue théorique, la concentration géographique d'un secteur peut s'expliquer soit par l'existence de dotations locales favorables au secteur considéré (proximité à la mer, présence d'eau, ressources minières...), soit par l'existence d'externalités positives de production entre acteurs du secteur. Ces externalités transitent par plusieurs canaux, comme la possibilité de partager un ensemble de fournisseurs, par l'existence d'un marché du travail large permettant une meilleure mutualisation des risques ou bien un meilleur appariement entre employés et employeurs et enfin, par la possibilité d'acquérir des expériences et des innovations produites par d'autres (Duranton et Puga (2004)). L'intervention publique en matière d'organisation spatiale des secteurs d'activité est justifiée dès lors que l'on pense que les forces de marché n'atteignent pas l'optimum social, aboutissant ainsi à trop ou trop peu de concentration (voir la première partie de Duranton, Martin, Mayer et Mayneris (2008) pour une analyse critique des politiques publiques de clusters).

Avant toute intervention, il est donc nécessaire de se questionner sur l'étendue *naturelle* du phénomène de concentration sectorielle. Quels sont les secteurs qui se concentrent ? Comment ces phénomènes de concentration évoluent-ils dans le temps ? Ces phénomènes de rapprochement spatial concernent-ils de manière uniforme toutes les entreprises d'un secteur donné ? Au-delà, quelles sont les zones qui se spécialisent ?

Cette étude a donc pour objectif de mesurer les phénomènes de concentration sectorielle en France à partir d'une méthode originale développée par Duranton et Overman (2005). Le principe est de comparer la distribution des distances bilatérales entre établissements au sein d'un secteur donné à celle au sein d'un secteur fictif ayant le même nombre d'établissements répartis aléatoirement. Ainsi, si la distribution des distances bilatérales observée dévie de la distribution aléatoire à une distance d , on conclut à la concentration à distance d du secteur. Cette méthodologie permet à la mesure de la concentration de s'abstraire du problème de la sensibilité au découpage géographique inhérent aux autres indicateurs (indices d'Ellison-Glaeser ou de

¹Source : www.competitivite.gouv.fr.

Maurel-Sédillot) utilisés dans les études antérieures sur ce sujet en France (Houdebine (1999), Maurel et Sédillot (1999), Duranton, Martin, Mayer et Mayneris (2008)).

À notre connaissance, l'étude de Duranton et Overman (2005) n'a, pour le moment, fait l'objet d'aucune réplique, car elle nécessite de disposer de données précises de localisation des établissements. La géolocalisation mise en oeuvre à l'Insee depuis 2002 répond à ce problème et nous permet de comparer les phénomènes de concentration en France avec ceux, étudiés par Duranton et Overman (2005) pour le Royaume-Uni. La comparaison internationale des phénomènes de concentration sectorielle est riche d'informations. Il est en effet intéressant d'étudier les différences dans l'organisation spatiale d'un même secteur d'un pays à l'autre, ainsi que leurs déterminants. Au delà des particularismes régionaux et des héritages historiques, l'analyse comparée de ces dynamiques de concentration peut permettre au décideur public de mieux cibler les secteurs pour lesquels une intervention est envisageable. Comme le rappellent Duranton, Martin, Mayer et Mayneris (2008), ce n'est pas tant la création *ex-nihilo* de clusters d'activité qui importe qu'une compréhension des défaillances de marché qui empêchent leur éclosion.

Nous étendons ensuite l'analyse aux secteurs des services marchands aux entreprises². Ces secteurs comptent en 2005 pour plus de 25% de la production en valeur et de 22% des emplois salariés intérieurs (en équivalent temps plein)³. Ces secteurs comptent ainsi pour une part importante dans la création de richesses. De plus, le développement des nouvelles technologies d'information et de la communication et, notamment, d'internet, a pu largement influencer le choix de localisation des entreprises et, plus particulièrement, dans les services intensifs en capital informatique. En effet, ces technologies ont contribué à diminuer considérablement les coûts de transfert de certains services. La comptabilité, l'appui juridique ou encore les centres d'appel, par exemple, peuvent maintenant se situer à distance de leurs clients. Les contraintes qui pèsent sur leur choix de localisation sont donc moindres. On peut envisager que les différences de technologie et de relation à la demande entre les secteurs de services et du manufacturier conduisent ceux-ci à des choix de localisation très différents.

La réduction des coûts de transferts consécutive au développement des nouvelles technologies de l'information peut également avoir un impact sur la dynamique de concentration de ces secteurs. C'est pourquoi, dans la dernière partie de cette étude, nous nous intéressons à cette question sur la période 1996-2005. L'absence de données géolocalisées sur l'ensemble de la période nous contraint à retenir une localisation à la commune, qui n'introduit qu'un léger biais dans les mesures.

Nous montrons tout d'abord que les quatre principaux résultats trouvés par Duranton et Overman (2005) pour le Royaume-Uni sont également valables pour la France : (i) l'utilisation d'un indicateur en continu conduit à conclure moins souvent qu'un secteur est concentré ; (ii) les secteurs les plus concentrés le sont le plus souvent à faible distance (entre 0 et 50 km) ; (iii) les valeurs prises par l'indice de concentration sont très variables d'un secteur à l'autre ; (iv) les secteurs appartenant à une même industrie se ressemblent en termes de concentration. En revanche, contrairement au Royaume-Uni, des phénomènes de concentration à moyenne distance (50-150

²Nous entendons par services marchands aux entreprises les secteurs NES K (transports), L (services financiers) et N (services aux entreprises).

³À titre de comparaison, le manufacturier représente 14,7% de l'emploi et 28% de la production.

km) s’observent en France, résultant probablement d’une relocalisation de l’emploi manufacturier en périphérie des grandes agglomérations (voir Gilli (2006) pour la région parisienne).

Les services semblent plus concentrés que les secteurs manufacturiers. De plus, cette concentration a généralement lieu à plus faible distance. Ceci est conforme à l’idée que ces secteurs se localisent principalement dans le cœur des villes et, lorsqu’ils ne sont pas contraints par la localisation de leur demande, dans un petit nombre d’entre elles.

Du point de vue de la dynamique, l’image globale reste celle d’une assez grande immobilité de l’activité économique en France sur la décennie passée. Nous mettons cependant en évidence une légère déconcentration de certains secteurs, consécutive à des entrées et sorties de sites actifs plus dispersées que l’ensemble des sites pérennes. Enfin, un premier exercice de statistique descriptive tend à montrer que les sites actifs n’appartenant pas à un pôle spécialisé sont davantage susceptibles de disparaître.

Cette étude est organisée ainsi. La première partie présente une revue critique des études antérieures sur la concentration sectorielle en France, ainsi que les données utilisées. La seconde détaille la construction de l’indicateur de Duranton et Overman (2005). Les parties suivantes présentent nos résultats : tout d’abord sur la comparaison des phénomènes de concentration dans les secteurs manufacturiers en France et au Royaume-Uni, ensuite sur l’analyse différenciée des secteurs manufacturiers et de services pour la France et, enfin, sur la dynamique sur la période 1996-2005.

1 Mesurer la concentration géographique d’un secteur

Dans cette partie, nous présentons brièvement quelques études antérieures sur la concentration géographique des secteurs d’activité sur données françaises : Houdebine (1999), Maurel et Sédillot (1999) et Duranton, Martin, Mayer et Mayneris (2008). Ces articles seront discutés à l’aune des critères définissant un *bon* indicateur de concentration géographique d’après la littérature (voir notamment le chapitre 10 de Combes, Mayer et Thisse (2006)). La partie suivante sera consacrée à l’introduction de l’indicateur de Duranton et Overman (2005), qui répond à nombre de ces critiques, ainsi qu’à la présentation des données utilisées.

1.1 Les études sur la concentration géographique des secteurs en France : une revue critique

Houdebine (1999) étudie la dynamique de la concentration géographique des activités et de la spécialisation des départements français sur la période 1978-1992. Son étude repose sur les données de l’Enquête Structure des Emplois (ESE). Cette enquête a pour champ tous les établissements du secteur privé, ainsi que les établissements publics à caractère industriel et commercial (EPIC) dont l’effectif au 31 décembre est supérieur à 20 salariés. L’auteur retient la classification NAP100⁴, ce qui lui permet d’étudier à la fois les secteurs industriels et de services. Les décou-

⁴Nomenclature d’Activités et de Produits.

pages géographiques retenus sont les départements français (Corse exclue) et les zones d'emploi⁵. L'auteur analyse trois types d'indicateurs : indices d'Herfindhal, de Gini et d'Ellison et Glaeser⁶. Quelle que soit l'année considérée, Houdebine (1999) montre que les secteurs industriels sont plus concentrés que les secteurs tertiaires au sens de l'indice de Gini (relatif), alors que l'inverse est vrai au regard de l'indice d'Herfindhal. Houdebine (1999) montre aussi que la concentration moyenne des activités économiques en 1992 est plus élevée au niveau des zones d'emploi qu'au niveau des départements, quel que soit l'indicateur choisi (Herfindhal, Gini, Ellison-Glaeser). De manière plus surprenante, il montre que si la concentration géographique moyenne de l'activité, mesurée au niveau des départements, a diminué sur la période 1978-1992, le résultat contraire prévaut lorsque la concentration est mesurée au niveau des zones d'emploi.

Maurel et Sédillot (1999) proposent un indicateur de concentration spatiale dérivé de celui d'Ellison et Glaeser (1997). Il s'interprète comme la corrélation entre les décisions de localisation de deux établissements d'un même secteur. Si cette corrélation est élevée, le secteur est considéré comme concentré. Les auteurs utilisent l'Enquête Annuelle d'Entreprise pour l'industrie manufacturière pour l'année 1993. Cette enquête recouvre l'ensemble des établissements d'entreprises des secteurs manufacturiers de plus de 20 salariés. L'étude est menée au niveau des départements et des régions françaises. Les auteurs identifient trois types d'industries concentrées, les industries extractives, dont le choix de localisation dépend de l'accès aux ressources minières, les industries traditionnelles (notamment du textile et du cuir) dont les choix de localisation, bien qu'anciens, se sont renforcés au cours du temps en raison de l'existence d'externalités de production et, enfin, les industries nouvelles de haute technologie profitant de *spillovers* de connaissance. Enfin, les auteurs concluent à de fortes similitudes entre la France et les États-Unis.

À partir d'un indicateur de Gini, Duranton, Martin, Mayer et Mayneris (2008) fournissent quelques informations sur l'évolution de la concentration sectorielle des secteurs manufacturiers français sur la période 1984-2004. En raison du changement de classification sectorielle en 1993, les auteurs ne présentent des résultats qu'au niveau de la NAF60. De manière surprenante, malgré les résultats de Houdebine (1999), les indicateurs sont présentés uniquement au niveau départemental. Ces choix méthodologiques limitent la portée des résultats. Les auteurs montrent que les secteurs les plus concentrés sont à la fois des secteurs de nouvelles technologies et des secteurs traditionnels. Concernant l'évolution temporelle, aucune conclusion claire n'apparaît. Il semble cependant que les secteurs dont la concentration géographique augmente le plus sont ceux qui perdent des emplois. Une telle affirmation doit être prise avec précaution car l'indice de Gini ne permet pas de séparer les phénomènes de réorganisation industrielle du secteur des phénomènes de réorganisation spatiale (l'indice de Gini ne corrige pas de la concentration industrielle).

⁵Ces dernières sont construites de manière à minimiser les trajets domicile-travail transfrontaliers. Autrement dit, la plupart des salariés français vivent et travaillent dans la même zone d'emploi. L'étude dynamique sur les zones d'emploi est restreinte à la période 1984-1992 car ces zones n'ont été créées par l'Insee qu'en 1984 (modifiées en 1993 puis en 1999).

⁶Ces indicateurs en espace discret sont construits à partir d'un découpage géographique donné du territoire national, par exemple la zone d'emploi. Ils consistent à comparer la répartition de l'emploi d'un secteur donné entre zones d'emploi à la répartition de l'emploi global. L'annexe A fournit plus de détails sur les définitions de ces indicateurs.

Au final, il ressort de ces études des résultats variés et parfois contradictoires. Plusieurs raisons peuvent expliquer ces contradictions :

- Le champ étudié repose sur des données d’enquêtes, il n’est donc pas exhaustif. Or les phénomènes de concentration des petites et grandes structures ne sont pas forcément comparables, comme l’ont d’ores et déjà souligné Holmes et Stevens (2002) sur données américaines, Duranton et Overman (2005) sur données britanniques et, enfin, Lafourcade et Mion (2007) sur données italiennes. À notre connaissance, cette question n’a fait l’objet d’aucune analyse détaillée sur données exhaustives en France.
- Le découpage géographique utilisé repose sur un maillage administratif prédéfini (région, département ou zone d’emploi). Or Briant, Combes et Lafourcade (2007) ont montré que la mesure des phénomènes de concentration à partir d’unités géographiques de tailles et de formes différentes pouvait conduire à des résultats variables, notamment lorsqu’il s’agit de dynamique.
- Ces études apportent peu d’éléments permettant de mesurer la significativité des phénomènes de localisation qu’elles mettent en exergue.
- Enfin, les indicateurs utilisés dans ces études (Herfindhal, Gini, Ellison-Glaeser, Maurel-Sédillot) ont été récemment très critiqués dans la littérature (Duranton et Overman (2005), Combes, Mayer et Thisse (2006) ou Marcon et Puech (2007)).

D’après Combes, Mayer et Thisse (2006), cinq critères de qualité permettent d’évaluer les indicateurs de concentration (pour une discussion détaillée de ces critères, se reporter à l’annexe B). La mesure de la concentration spatiale doit :

1. être effectuée par rapport à une référence clairement établie ;
2. être comparable entre secteurs ;
3. être comparable entre zones géographiques et être insensible à un changement de définition des unités géographiques ;
4. permettre de déterminer si des différences significatives par rapport à la référence ou entre deux situations (zones, périodes ou secteurs) existent ;
5. être insensible à un changement de définition des secteurs.

L’indicateur de Duranton et Overman (2005), tel que calculé ici, répond à un certain nombre des critiques formulées précédemment. Avant de présenter plus formellement cet indicateur, nous montrons dans la partie suivante un exemple pour lequel l’utilisation des indicateurs d’Ellison-Glaeser conduit à une conclusion radicalement différente de celui de Duranton et Overman (2005).

1.2 Illustration des problèmes posés par l'indicateur d'Ellison-Glaeser

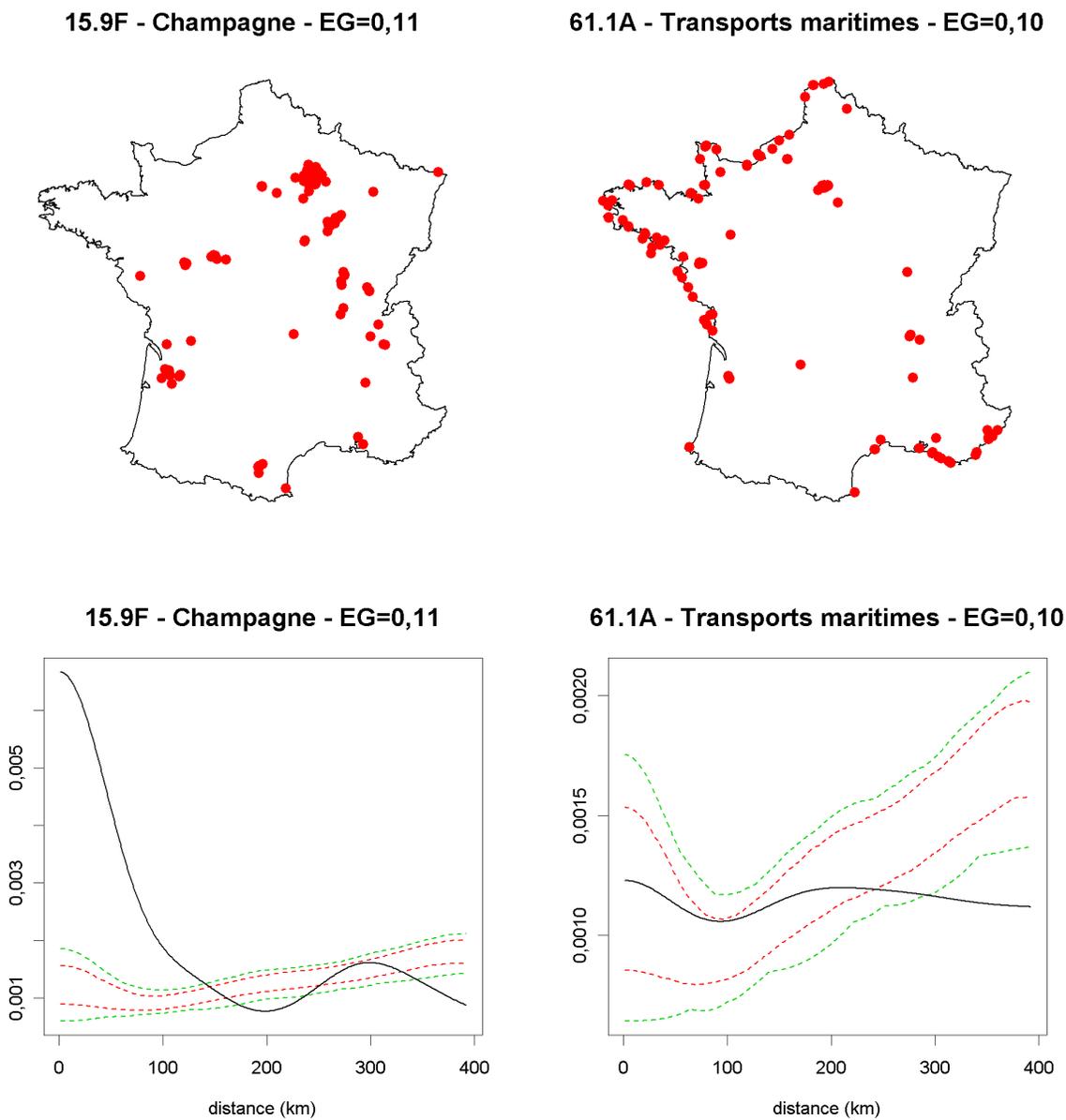
L'indicateur de Duranton et Overman (2005) repose sur la distribution des distances bilatérales entre établissements d'un secteur donné. Il s'abstrait ainsi de toute référence à un quelconque découpage géographique. Or, la principale critique à l'encontre de l'estimateur d'Ellison et Glaeser (ainsi que celui de Maurel et Sédillot) concerne la dépendance à ce découpage (voir annexe B).

En effet, dans le calcul d'un indice d'Ellison et Glaeser, seule compte l'entité géographique de base, par exemple la zone d'emploi, indépendamment de sa position relative par rapport aux autres unités sur le territoire étudié. Autrement dit, une permutation spatiale de ces unités ne change pas la valeur de l'indicateur.

La figure 1 illustre cette remarque. Les deux secteurs d'activité représentés (Champagnisation et Transport Maritime) ont des indices d'Ellison et Glaeser très proches, égaux à 0,11 pour le premier et 0,10 pour le second. Au regard de la distribution des indicateurs d'Ellison et Glaeser (calculés à la commune) pour l'année 2005, ces secteurs sont particulièrement concentrés. Cependant, visuellement, il apparaît très clairement que les distributions spatiales des établissements de ces secteurs sont très différentes. La simple observation des cartes permet de penser que le secteur de la champagnisation est plus concentré que celui du transport maritime. L'indice proposé par Duranton et Overman (2005) tient compte de cette réalité.

Les graphiques de la figure 1 présentent la distribution des distances bilatérales entre les établissements du secteur (courbes en trait plein) et les intervalles de confiance (courbes en pointillés) associés (notions que nous définirons plus précisément par la suite). Pour le secteur de la champagnisation, la distribution des distances présente un « pic » aux courtes distances, absent dans le cas du secteur des transports maritimes. Nous verrons qu'au sens de Duranton et Overman (2005), le secteur de la champagnisation est concentré tandis que le secteur des transports maritimes est dispersé.

FIG. 1 – Impact de la permutation des unités spatiales : comparaison des indices d'Ellison et Glaeser et de Duranton et Overman



Note de lecture : les sites actifs en 2005 des secteurs de la champagnisation et des transports maritimes sont localisés sur les cartes. Les graphiques présentent les distributions des distances bilatérales entre les établissements du secteur (courbes en trait plein) et les intervalles de confiance (courbes en pointillés). Les intervalles de confiance locaux sont représentés par les courbes en pointillés intérieures, les intervalles de confiance globaux sont représentés par les courbes en pointillés extérieures.

Source : Insee - Calculs des auteurs.

1.3 L'indicateur en continu de Duranton et Overman (2005) : données géolocalisées et distances bilatérales entre établissements

Le principe de la méthode

La méthode développée par Duranton et Overman (2005) repose sur la géolocalisation des établissements d'un secteur donné. L'information utile se présente sous forme de coordonnées géographiques (x,y) des établissements ; il ne s'agit donc plus de considérer leur appartenance à des entités géographiques. On passe schématiquement d'une appréhension discrète de l'espace, à une appréhension en continu. Si ces méthodes par points (*point-pattern methods*) sont usuelles dans d'autres domaines scientifiques (notamment l'écologie), elles restent minoritaires en économie⁷. Pourtant, ces méthodes ont l'avantage indéniable de ne reposer sur aucun découpage géographique prédéfini, elles ne sont sensibles ni à la taille, ni à la forme, ni à la position relative des unités géographiques.

Cet indicateur compare la distribution observée des distances bilatérales au sein d'un secteur donné à celle d'un secteur fictif avec le même nombre d'établissements distribués aléatoirement sur l'ensemble de sites actifs en France. Nous entendons par « site actif » l'ensemble des sites effectivement occupés par des établissements actifs, quelque soit son secteur d'activité. Ceci fournit donc un test de la concentration d'un secteur, i.e. de sa tendance à se concentrer relativement plus que le reste de l'activité économique. La construction du contrefactuel, i.e. la distribution aléatoire sur l'ensemble des sites actifs d'un nombre d'établissements égal au nombre effectivement observé dans le secteur, permet de contrôler :

1. la concentration de l'activité économique globale, en choisissant comme univers des possibles l'ensemble des sites effectivement actifs ;
2. la concentration industrielle du secteur, en considérant un secteur hypothétique avec le même nombre d'établissements que celui effectivement observé.

Le contrefactuel correspond à une situation où le secteur ne serait soumis à aucune force de concentration particulière. Aussi, en générant un nombre suffisant de distributions aléatoires, il est possible de construire un intervalle de confiance en dehors duquel le secteur peut-être considéré comme significativement concentré ou dispersé⁸. Au final, cet indicateur vérifie l'ensemble des critères de qualité évoqués précédemment, à l'exception de la non-dépendance à la nomenclature sectorielle⁹.

Une méthode intensive en données

Pour décrire la distribution des distances bilatérales entre établissements, il est nécessaire de localiser très précisément ces établissements. Duranton et Overman (2005) utilisent pour cela une base de données sur les établissements britanniques (*l'Annual Respondent Database*) pour

⁷Marcon et Puech (2007) ont également développé un indicateur de concentration en continu, reposant sur des hypothèses légèrement différentes.

⁸La définition précise de ces deux concepts sera faite dans la partie suivante.

⁹Ce critère n'est vérifié par aucun indicateur.

lesquels on connaît : le code activité (dans la classification SIC¹⁰ à 4 chiffres), le nombre de salariés et le code postal (*postcode*). Duranton et Overman (2005) ne retiennent que les 176 000 établissements du secteur manufacturier (SIC 15 000 à 36 639) enregistrant au moins un salarié pendant l'année 1996. Le *postcode* britannique est une information très précise, correspondant généralement à un petit groupe de bâtiments. Il est possible de connaître les coordonnées géographiques de chacun de ces *postcodes* et, ce faisant, de ne faire qu'une erreur de l'ordre de 100 mètres sur la localisation de l'établissement. Au final, Duranton et Overman (2005) géolocalisent environ 95% de leur échantillon.

Nous utilisons pour la France principalement deux sources de données : les bases issues du projet CLAP (Connaissance Locale de l'Appareil Productif), ainsi que les données géolocalisées du répertoire SIRENE (voir l'annexe C pour plus de détails).

Le projet CLAP constitue une synthèse de sources administratives dont l'objectif est de localiser l'ensemble de l'activité économique au travers des effectifs et des rémunérations. Il fournit donc une photographie du stock d'établissements actifs au 31 décembre de chaque année. Nous utilisons les données CLAP pour l'année 2005. Suivant Duranton et Overman (2005), nous ne retenons dans cette étude que les établissements déclarant au moins un salarié. Nous retenons dans cet ensemble les établissements appartenant aux secteurs industriels (NES B à F) et aux services marchands aux entreprises (NES K, L et N). Nous nous concentrons sur les établissements localisés en France continentale. Notre base initiale est finalement constituée de 518 126 établissements, dont 165 034 appartiennent aux secteurs manufacturiers.

Les données géolocalisées du répertoire Sirene renseignent sur les coordonnées géographiques de chaque établissement au 1er janvier 2006. Elles ne sont cependant disponibles que pour les établissements situés dans les villes de plus de 10 000 habitants¹¹. Ce premier rapprochement de source permet de localiser précisément 46% des établissements.

Pour les établissements restants, le géocodage est effectué à partir des adresses postales. Au final, pour les secteurs retenus, cette technique permet de porter à 89% la part des établissements de l'échantillon qui sont effectivement géolocalisés.

Les derniers établissements sont localisés aux coordonnées de la mairie de leur commune d'appartenance à partir de la base de données du Répertoire Géographique des Communes (RGC) de l'Institut Géographique National (IGN). Dans la conclusion de leur article, Duranton et Overman (2005) préconisent de distribuer aléatoirement les établissements sur le maillage géographique à disposition dans le cas où aucune information géolocalisée n'est disponible. Notons cependant que la superficie médiane d'une commune française est de 10,7 km², équivalent à un rayon médian de moins de 2 km. Ensuite, les établissements localisés au centre de la commune sont localisés dans des communes rurales. Nous faisons donc ici l'hypothèse que l'environnement proche de la mairie est le coeur d'activité. Dans tous les cas, nous pensons donc que la localisation à la

¹⁰Standard Industrial Classification. Cette classification à 4 chiffres n'a pas d'équivalent en France. Cependant, la classification en NAF700 est plus détaillée et chaque NAF700 française correspond à une et une seule classe en SIC4. Il est donc possible de construire facilement une table de passage.

¹¹On compte en France, suivant le recensement de 1999, seulement 888 villes de plus de 10 000 habitants sur plus de 36 200 communes.

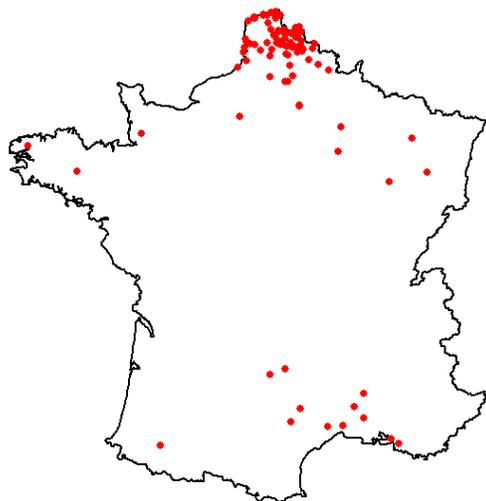
mairie pour les 11% restants de notre échantillon n'introduit qu'une erreur très faible dans les estimations.

La figure 2 présente la distribution des établissements en France pour quatre secteurs manufacturiers. Il s'agit des secteurs de : (a) la transformation et la conservation de pommes de terre (15.3A), (b) l'imprimerie de journaux (22.2A), (c) de la fabrication de pull-overs et articles similaires (17.7C) et (d) de la fabrication de menuiseries et fermetures métalliques (28.1C). Une analyse visuelle laisse penser que le secteur (a) est concentré essentiellement dans le Nord de la France et que le secteur (c) présente trois pôles de concentration autour de Paris, Roanne et Mazamet. À l'inverse, les secteurs (b) et (d) semblent beaucoup plus dispersés sur l'ensemble du territoire national. Nous verrons que cette intuition est confirmée par l'indicateur de Duranton et Overman (2005).

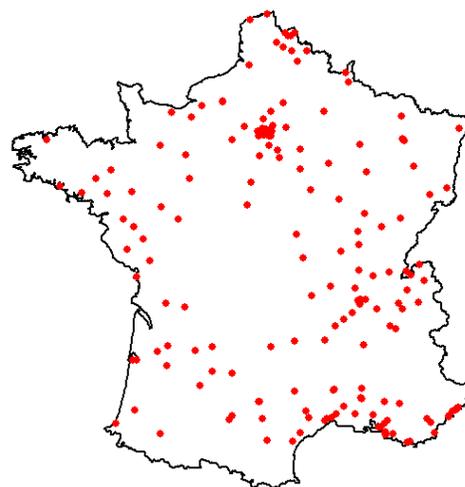
Dans la partie suivante, nous revenons plus longuement sur la technique d'estimation de la distribution des distances bilatérales entre établissements d'un même secteur, ainsi que sur la détermination d'un intervalle de confiance.

FIG. 2 – Distribution spatiale des établissements dans quatre secteurs manufacturiers

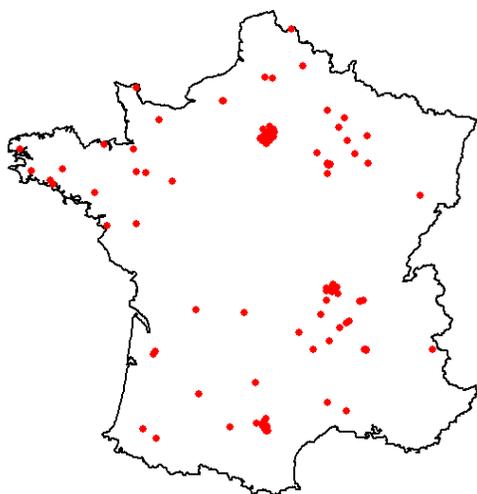
(a) 15.3A - Transformation et conservation de pommes de terre



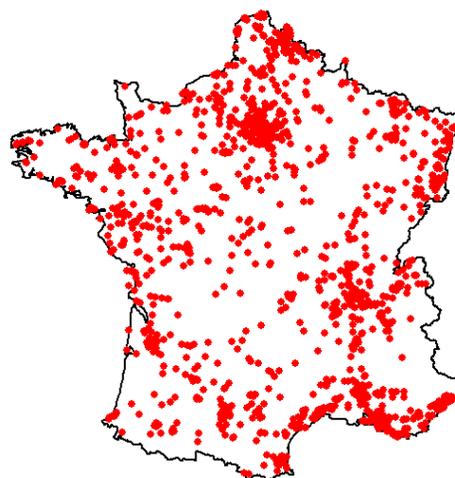
(b) 22.2A - Imprimerie de journaux



(c) 17.7C - Fabrication de pull-overs et articles similaires



(d) 28.1C - Fabrication de menuiseries et fermetures métalliques



Source : Insee.

2 Mise en oeuvre de l'indicateur en continu de Duranton et Overman (2005)

Le calcul de l'estimateur de Duranton et Overman (2005) s'effectue en 5 étapes. Nous ne reprenons ici que les éléments principaux de la présentation détaillée dans Duranton et Overman (2005).

Calculer la distribution observée des distances entre paires d'établissements au sein d'un secteur

La première étape consiste à calculer l'ensemble des distances bilatérales entre les établissements d'un même secteur. Le calcul des distances est soumis essentiellement à deux sources d'erreurs de mesure¹² :

1. On mesure les distances euclidiennes et non les temps de parcours entre deux établissements. D'une part, la France n'étant pas convexe, il se peut que la distance euclidienne traverse la mer ou l'océan (La distance euclidienne entre Brest et Biarritz par exemple). D'autre part, la densité et la congestion des réseaux de transport déterminent la différence entre les distances euclidiennes et les temps de parcours. La question est alors de savoir si cette erreur est spatialement biaisée. Duranton et Overman (2005) remarquent qu'il est difficile d'avoir une idée du sens du biais, car si le tissu routier est plus dense dans les zones d'activité intense, il y est aussi plus congestionné. De plus, sur données françaises, Combes et Lafourcade (2005) ont montré que la corrélation entre la distance euclidienne et le coût de transport généralisé s'élève à 0,97.
2. Certains établissements sont localisés à la commune.

Les distances observées sont donc mesurées avec une certaine erreur. Nous utilisons alors une méthode non paramétrique de lissage par noyau des densités de distribution (*density kernel smoothing*)¹³. Notons que cette procédure de lissage permet de corriger du manque de précision de la mesure des distances mais pas d'un éventuel biais systématique. Nous obtenons ainsi un estimateur de la distribution des distances d entre paire d'établissements au sein du secteur, noté $\hat{K}(d)$.

$$\hat{K}(d) = \frac{1}{n(n-1)h} \sum_{i=1}^{n-1} \sum_{j=i+1}^n f\left(\frac{d-d_{i,j}}{h}\right)$$

où i , et j désignent deux établissements du secteur étudié (qui en compte n). La distance $d_{i,j}$ est la distance effectivement observée entre i et j ¹⁴. Suivant Duranton et Overman (2005), le noyau f choisi est une gaussienne centrée réduite, la fenêtre (*bandwidth*) h choisie correspond à celle proposée par Silverman (1986) (section 3.4.2)¹⁵. Les graphiques de la figure 3 présentent les

¹²Une troisième source très mineure est liée à la courbure de la Terre. Néanmoins, la corrélation entre la vraie distance calculée à partir d'un logiciel de cartographie (Mapinfo 7.8) et la distance euclidienne est supérieure à 0,99.

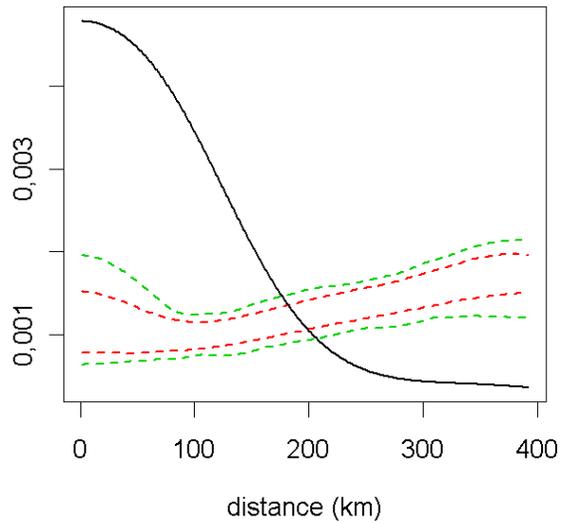
¹³Nous utilisons exactement la même méthode de lissage que Duranton et Overman (2005).

¹⁴le lissage revient à « interpoler » la distribution de densité autour de distances $d_{i,j}$ effectivement observées

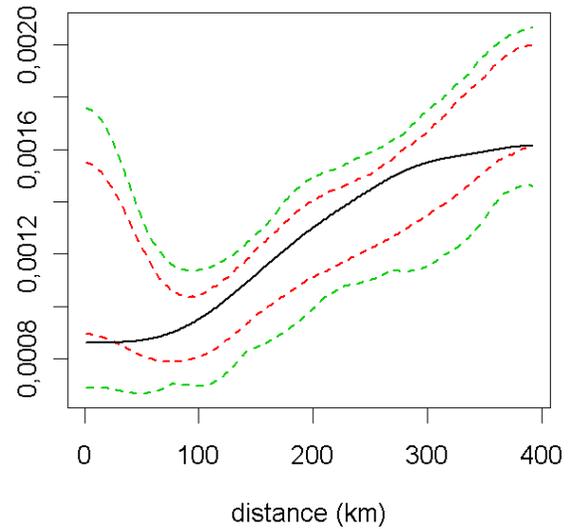
¹⁵La procédure de lissage de la distribution nécessite quelques adaptations du fait de la borne inférieure à zéro sur les distances. Voir Duranton et Overman (2005) pour plus de détails.

FIG. 3 – Distribution des distances dans quatre secteurs manufacturiers

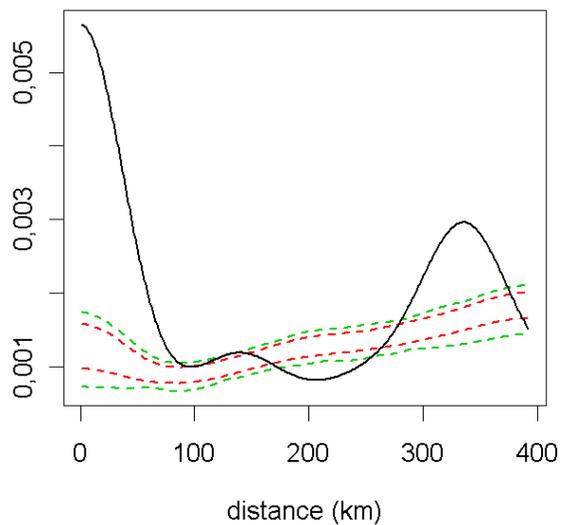
(a) 15.3A Transformation et conservation de pommes de terre



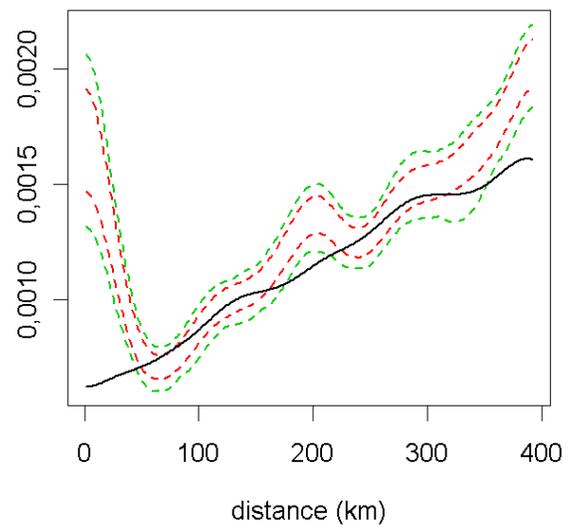
(b) 22.2A Imprimerie de journaux



(c) 17.7C Fabrication de pull-overs et articles similaires



(d) 28.1C Fabrication de menuiseries et fermetures métalliques



Note de lecture : distributions des distances bilatérales entre les établissements des secteurs (courbes en trait plein) et les intervalles de confiance (courbes en pointillés). Les intervalles de confiance locaux sont représentés par les courbes en pointillés intérieures, les intervalles de confiance globaux sont représentés par les courbes en pointillés extérieures.

Source : Insee - Calculs des auteurs.

distributions des distances pour les quatre secteurs pris en exemple précédemment. Ces résultats confirment l'examen visuel des cartes. Ainsi le secteur de la transformation et de la conservation de pommes de terre semble concentré spatialement. En effet, on observe un « pic » dans la distribution pour les faibles distances. Il apparaît que ce « pic » est significatif, au sens où il se trouve en dehors de l'intervalle de confiance représenté par les traits pleins et pointillés. Comment ces intervalles de confiance sont-ils construits ?

Choix du contrefactuel

La seconde étape consiste à définir une distribution de référence permettant de conclure à la significativité de la concentration des secteurs. L'originalité de la méthode de Duranton et Overman (2005) est de redistribuer aléatoirement les établissements du secteur sur un ensemble de sites dits « actifs ». Le choix de cet ensemble des possibles pour la distribution aléatoire correspond aux choix d'un contrefactuel, autrement dit d'un ensemble de points dans l'espace où les établissements du secteur auraient pu se localiser. Duranton et Overman (2005) choisissent comme contrefactuel l'ensemble des sites où est effectivement localisé un établissement manufacturier actif, quel que soit son secteur d'activité.

Une fois les établissements redistribués aléatoirement sur l'ensemble des sites actifs, il est possible, comme précédemment, de reconstruire la distribution des distances entre paires d'établissements. Cette opération est répétée 1000 fois pour obtenir un faisceau de distributions, à partir duquel on déduit un intervalle de confiance¹⁶.

Construction des intervalles de confiance

Une première idée serait de comparer la distribution observée et l'intervalle de confiance à toutes les distances entre 0 et 1 103 km (distance maximale entre deux établissements en France). Or l'intégrale des distributions sommant à 1, il existe une distance à laquelle ces distributions se croisent. Par exemple, un secteur initialement très concentré (i.e. au-dessus de l'intervalle de confiance) passera mécaniquement sous l'intervalle de confiance à plus longue distance. Il n'est donc pas utile de considérer l'information apportée par la distribution aux longues distances, dès lors que l'on étudie la position relative des distributions par rapport aux bornes *supérieures* et *inférieures* de l'intervalle de confiance aux courtes distances.

C'est pourquoi, suivant Duranton et Overman (2005), la comparaison entre la distribution observée et l'intervalle de confiance est réalisée jusqu'à 392 km. Cette distance correspond à la médiane de l'ensemble des distances bilatérales entre tous les établissements, manufacturiers et de services, en France. Cette distance est notée *med*.

Il est possible de construire des intervalles de confiance locaux et globaux. L'intervalle de confiance local est obtenu pour chacune des distances, en retirant de l'échantillon les 5% de valeurs les plus grandes et les plus faibles. On obtient ainsi les bornes de l'intervalle $\bar{K}(d)$ et

¹⁶Notons que dans le cas des données de point, il n'est pas possible de dériver la distribution asymptotique de l'estimateur $\hat{K}(d)$ du fait de l'interdépendance entre les distances issues d'un même point. Autrement dit, même si les localisations sont indépendantes, les distances entre les points ne le sont pas. C'est pourquoi la méthode de Monte-Carlo est utilisée pour définir un intervalle de confiance.

$\underline{K}(d)$ (représentées par les courbes en pointillés intérieures sur les figures 3). Cependant, ne considérer que cet intervalle de confiance conduirait à conclure trop souvent à la concentration ou bien à la dispersion d'un secteur. Même pour une distribution aléatoire pour laquelle la probabilité de sortir de l'intervalle à chaque distance est de 5%, la probabilité de sortir *au moins une fois* sur l'intervalle (0-392) km est très élevée. Pour pallier cette difficulté, on calcule un intervalle de confiance global, de bornes $\overline{\overline{K}}(d)$ et $\underline{\underline{K}}(d)$ (représentées par les courbes en pointillés extérieures sur la figure 3). Ces intervalles sont construits en retirant uniquement les mêmes simulations à toutes les distances, de sorte que l'ensemble retiré ne représente que 5% du total des simulations¹⁷. Pour ce faire, on cherche « par balayage » sur l'ensemble des distances d , le percentile α tel qu'en enlevant pour chaque distance d les $\alpha\%$ simulations extrêmes, le nombre total de simulations supprimées représente 5% de l'ensemble des simulations. Une telle procédure de construction implique évidemment qu'à chaque distance α est plus petit que 5% et donc que l'intervalle de confiance local est inclus dans l'intervalle de confiance global.

Indicateurs de concentration et de dispersion

On définit ensuite, à partir des intervalles de confiance globaux, des indicateurs de concentration $\Gamma(d)$ et de dispersion $\Psi(d)$ à chaque distance :

$$\Gamma(d) = \max(\hat{K}(d) - \overline{\overline{K}}(d), 0)$$

$$\Psi(d) = \begin{cases} \max(\underline{\underline{K}}(d) - \hat{K}(d), 0), & \text{si } \forall d \in [0 : med] \quad \Gamma(d) = 0 \\ 0, & \text{sinon} \end{cases}$$

Notons l'asymétrie dans la définition de ces indicateurs. Un secteur sera concentré à la distance d dès que $\hat{K}(d)$ dépasse la borne supérieure $\overline{\overline{K}}(d)$ de l'intervalle de confiance global. En revanche, un secteur sera dispersé à la distance d si $\hat{K}(d)$ est inférieur à $\underline{\underline{K}}(d)$ et s'il n'est concentré à aucune autre distance. Cette asymétrie peut être justifiée par le fait que, si un secteur est initialement très concentré, il apparaîtra nécessairement dispersé à une certaine distance, l'intégrale des distributions valant 1.

Indicateurs de concentration et de dispersion sectoriels

Les indices sectoriels sont définis ainsi :

$$\Gamma = \int_{d=1}^{med} \Gamma(d), \quad \Psi = \int_{d=1}^{med} \Psi(d)$$

Ils servent à définir le pourcentage de secteurs concentrés (ou dispersés). Notons qu'ils font à nouveau intervenir la distance médiane qui, au total, intervient trois fois : dans le calcul des intervalles de confiance globaux, dans le calcul des indices de dispersion à chaque distance et dans le calcul des indices sectoriels. Le choix de cette distance n'est donc pas anodin. Ces indicateurs permettent d'agréger les résultats obtenus pour chaque distance au prix d'une perte de précision dans l'information fournie.

¹⁷Cette construction nécessite quelques précautions, développées par Duranton et Overman (2005).

Pour le secteur de la transformation et de la conservation de pommes de terre, l'indicateur de concentration Γ vaut 0,355 (l'indicateur de dispersion Ψ est nul), ce secteur est significativement (au seuil de 5%) concentré au sens de Duranton et Overman (2005). Le secteur de l'imprimerie des journaux, dont la distribution ne sort pas de l'intervalle de confiance, n'est ni concentré, ni dispersé ($\Gamma = \Psi = 0$). Le secteur de la fabrication de pull-over est également concentré, avec $\Gamma = 0,222$ et $\Psi = 0$. Enfin, le secteur de la fabrication de menuiserie et de fermetures métalliques apparaît dispersé : $\Psi = 0,023$ et $\Gamma = 0$.

3 Concentration des activités manufacturières : comparaison France - Royaume-Uni

Cette partie est consacrée à la comparaison de la concentration des activités entre la France et le Royaume-Uni. Dans un premier temps nous présentons rapidement la distribution globale de l'activité en France. En effet, une première explication de la différence entre les résultats obtenus pour la France et le Royaume-Uni tient à la géographie de ces pays.

3.1 La géographie de la France

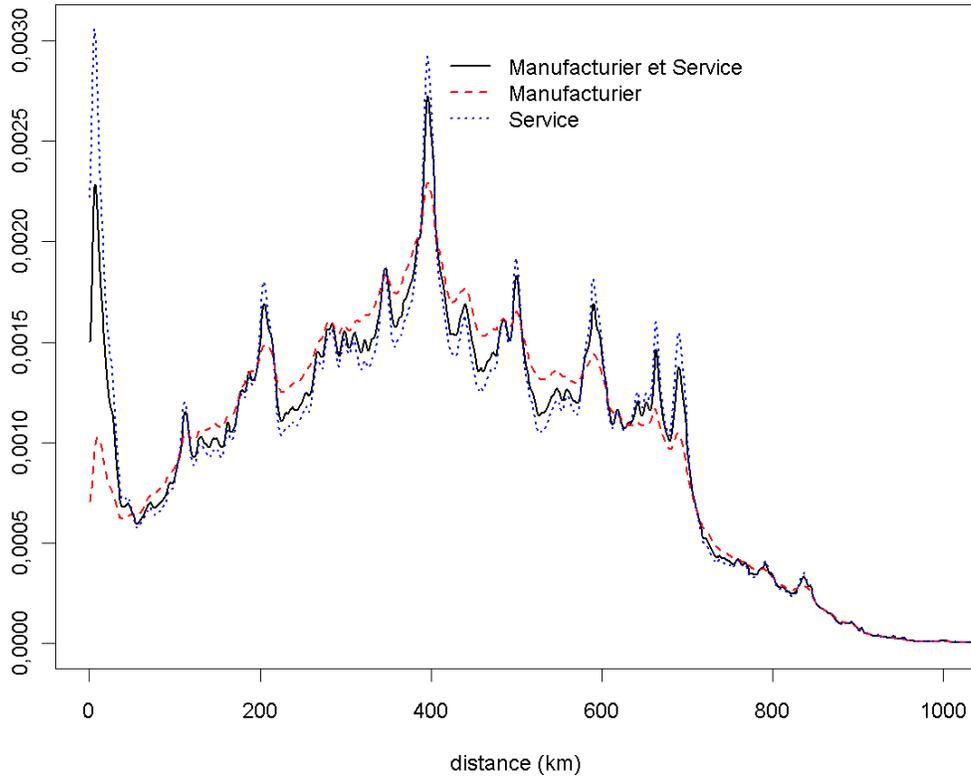
Notre indicateur fait intervenir la distance médiane entre l'ensemble des établissements. Nous considérons successivement l'ensemble des secteurs manufacturiers, l'ensemble des secteurs de service et la réunion de ces deux ensembles. Ces trois ensembles définissent le contour des contrefactuels utilisés dans les parties suivantes. Rappelons que dans la définition des intervalles de confiance interviennent deux informations : l'ensemble des sites actifs du contrefactuel et le nombre d'établissements du secteur considéré. La première information détermine la forme générale de l'intervalle, la seconde influence sa largeur. Ainsi, plus un secteur aura un grand nombre d'établissements, plus l'intervalle de confiance sera serré autour de la distribution observée des distances entre tous les sites actifs du contrefactuel.

Les distributions des distances entre paires d'établissements pour ces trois ensembles ont des formes relativement similaires (figure 4). Elles présentent un premier pic à très faible distance, un second pic moins marqué à 205km et un troisième pic à 396 km. Notons que le premier pic est nettement moins élevé pour l'ensemble des établissements manufacturiers que pour l'ensemble des établissements de service. En toute logique, pour la réunion des deux ensembles, le pic se situe entre les deux.

Le premier pic correspond aux grandes agglomérations françaises, en particulier à l'agglomération parisienne. Comme les agglomérations sont relativement peu étendues spatialement mais regroupent un grand nombre d'établissements, la proportion de distances bilatérales inférieures à 50km est relativement élevée¹⁸. Les établissements manufacturiers ayant moins tendance à se regrouper uniquement dans le centre des villes, le premier pic est beaucoup moins marqué pour ce groupe. Ceci est une différence notable entre les secteurs manufacturiers et de services, non sans conséquence sur la suite de nos résultats.

¹⁸Duranton et Overman (2005) ne présentent pas la distribution globale des distances entre paires d'établissements au Royaume-Uni. Il est cependant fortement probable que le même phénomène soit observé, du fait des grandes villes britanniques.

FIG. 4 – K-estimateur pour l'ensemble des établissements



Source : Insee - Calculs des auteurs.

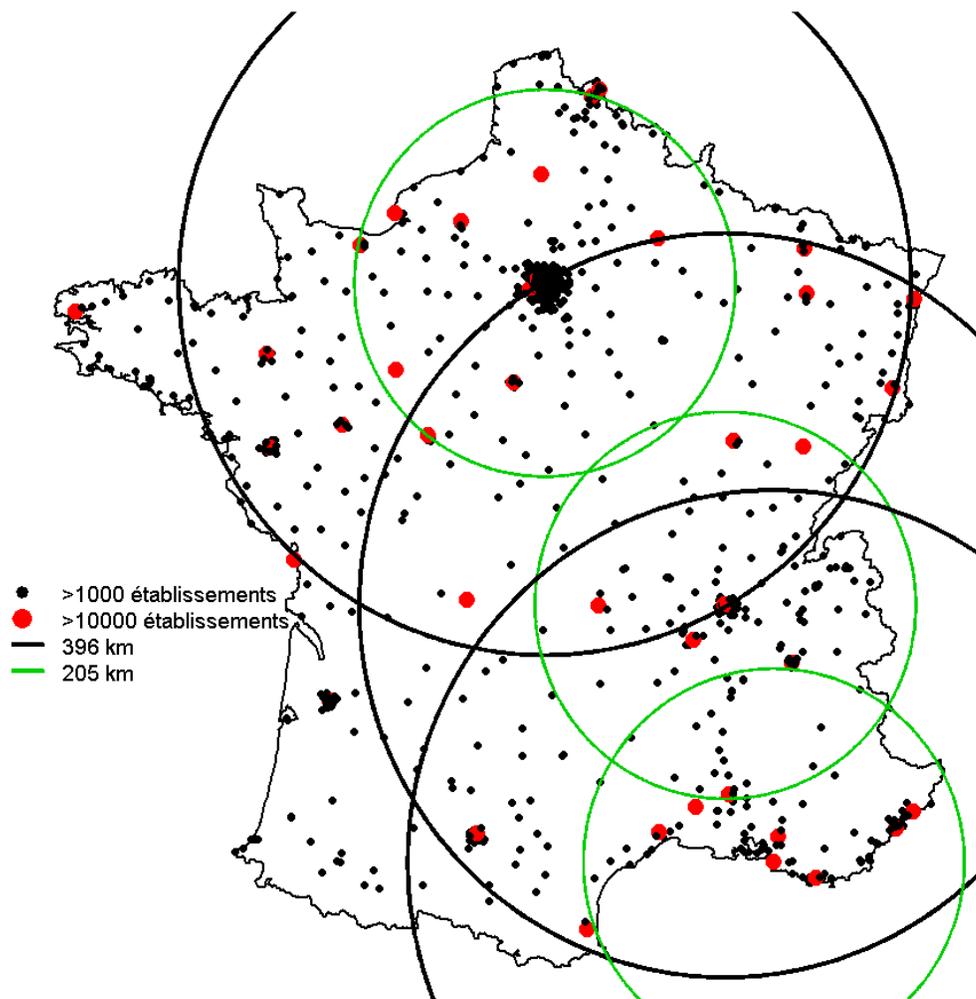
Les pics secondaires correspondent aux distances entre les grandes villes françaises (figure 5). En particulier, il y a 396 km entre Paris et Lyon (à vol d'oiseau), mais aussi entre Paris et Strasbourg ou Paris et La Rochelle. De même, 200 km correspond à la distance entre Paris et Lille ou Paris et Tours, ainsi que Marseille et Montpellier¹⁹.

Sur la fenêtre d'analyse de nos résultats (0 à 392 km), nous devrions donc observer pour un secteur réparti aléatoirement sur le territoire une plus forte proportion des distances bilatérales à moins de 50 km, à 205 et vers 392 km. Notons que la répartition aléatoire est très différente de la répartition uniforme. En effet, la distribution des distances correspondant à une répartition uniforme des établissements sur le territoire serait linéairement croissante avec la distance (aux effets de bord près)²⁰. Choisir comme contrefactuel l'ensemble des sites actifs (et non pas par exemple l'ensemble des coordonnées géographiques accessibles sur le territoire français) permet de tenir compte de cette concentration globale de l'activité économique, prise comme référence.

¹⁹Notons que dans ce dernier cas, l'erreur commise entre la distance euclidienne et la distance par la route évoquée précédemment est flagrante.

²⁰Dans le cas d'une répartition uniforme, pour un établissement donné le nombre de voisins entre r et $r+dr$ km est proportionnel à la surface du disque entre r et $r+dr$. Or cette surface ($2\pi r dr$) dépend linéairement de r .

FIG. 5 – Distances entre grandes villes françaises



Source : Insee - Calculs des auteurs.

3.2 Analyse comparée France/Royaume-Uni

La disponibilité des données françaises permet de réaliser une étude comparée des phénomènes de concentration sectorielle en France et au Royaume-Uni. La comparaison internationale des phénomènes de concentration sectorielle est riche d'informations. Il est en effet intéressant d'étudier les différences dans l'organisation spatiale d'un même secteur d'un pays à l'autre, ainsi que leurs déterminants. Pour comparer directement les résultats, nous utilisons une table de passage qui associe de manière unique un code de la nomenclature NAF700 (révision 1, 2003) à un code de la nomenclature SIC 92 (à 4 chiffres) utilisée par Duranton et Overman (2005).

Combien d'industries dévient de la distribution aléatoire et à quelle distance ?

Nous considérons donc les 229[234] secteurs en SIC 92 à 4 chiffres avec plus de 10 établissements actifs au 31 décembre 2005. Dans ce qui suit les chiffres entre crochets correspondent aux valeurs équivalentes pour Duranton et Overman (2005).

Parmi ces secteurs, 164[177] dévient de la distribution aléatoire, dont 138[122] sont concentrés, contre 26[55] dispersés. On trouve donc 60%[52%] des secteurs *globalement* concentrés contre 11%[24%] *globalement* dispersés. Deux conclusions s'imposent. Il semble tout d'abord y a avoir en France légèrement plus de concentration sectorielle qu'au Royaume-Uni. Cependant, comme pour Duranton et Overman (2005), dans le cas français, l'utilisation du K-estimateur conduit à conclure à un nombre de secteurs concentrés bien inférieur aux études précédentes utilisant des indicateurs en espace discret. En effet, Maurel et Sédillot (1999) concluent à une concentration significativement positive pour 77% des secteurs manufacturiers étudiés.

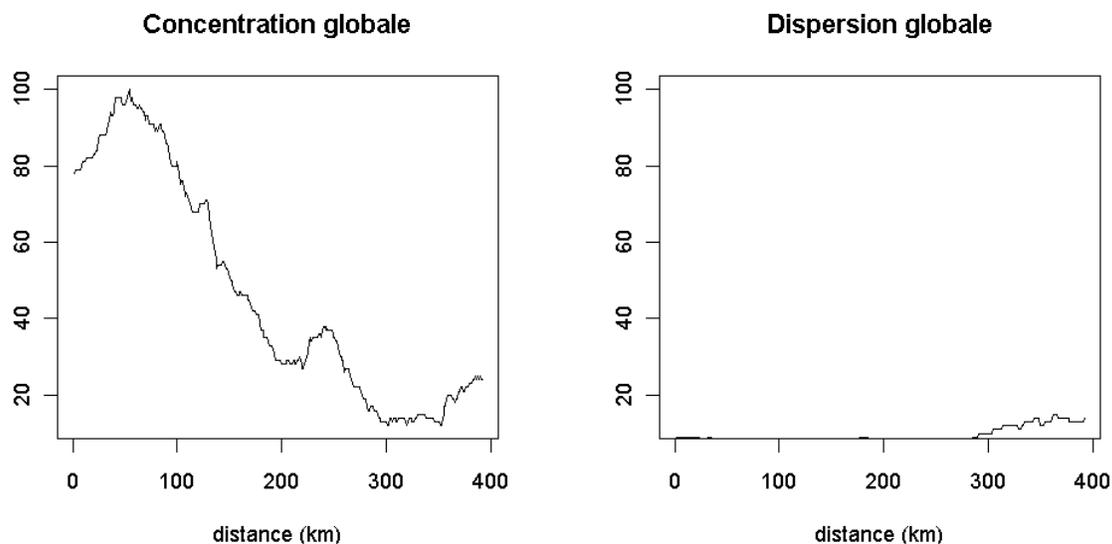
En France, la distance à laquelle le plus grand nombre de secteurs apparaissent concentrés s'établit à 50km. De plus, la majorité des secteurs concentrés le sont sur des distances inférieures à 100-120 km. Cependant, on peut noter une fréquence relativement élevée de secteurs concentrés autour de 245 km au sens de Duranton et Overman (2005) (figure 6). Duranton et Overman (2005) remarquent également sur les données britanniques la présence de deux pics de concentration : le premier entre 5 et 30 km et le second entre 100 et 120 km. Il apparaît donc une sorte d'« homothétie » entre les résultats français et britanniques, que l'on peut certainement attribuer aux différences géographiques entre les pays. Les secteurs concentrés le restent néanmoins à courte distance.

Cependant, les graphiques ci-dessus ne donnent aucune information quantitative sur l'ampleur de ces déviations. Pour en avoir une vision, nous sommions à chaque distance la valeur des $\Gamma(d)$ sur l'ensemble des secteurs. Il apparaît que les phénomènes de concentration sont particulièrement marqués à courte distance (inférieures à 150 km) (figure 7). Cette conclusion est également valable dans le cas du Royaume-Uni, mais les ordres de grandeurs ne sont pas tout-à-fait les mêmes. Duranton et Overman (2005) montrent ainsi que les phénomènes de concentration dépassent rarement 60 km. Il s'agit donc d'expliquer pourquoi, en France, certains secteurs apparaissent encore concentrés au delà de 60 km.

Deux interprétations sont envisageables :

1. les secteurs concentrés au Royaume-Uni et en France ne le sont pas à la même distance. Dans ce cas, les forces de concentration ne sont pas spécifiques au secteur mais bien au pays étudié.
2. ou bien les secteurs concentrés dans les deux pays le sont à la même distance, mais il existe des secteurs présentant en France une concentration à moyenne distance (60-150 km) et pas au Royaume-Uni.

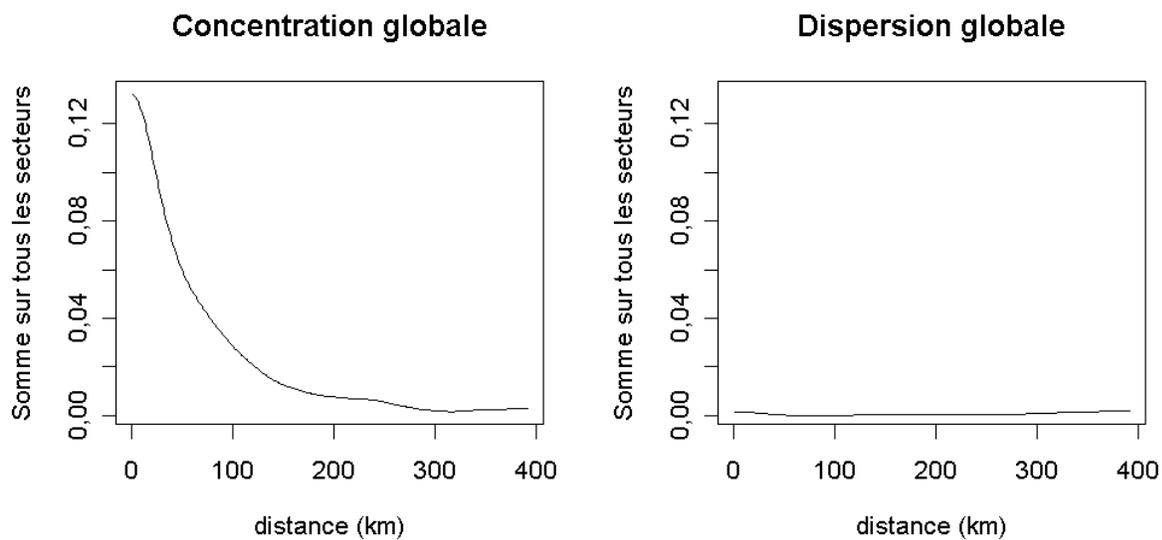
FIG. 6 – Concentration et dispersion par distance en France



Note de lecture : Sont représentés, pour chaque distance, le nombre de secteurs qui apparaissent significativement concentrés (resp. dispersés).

Source : Insee - Calculs des auteurs.

FIG. 7 – Ampleur de la concentration et de la dispersion par distance en France



Note de lecture : Sont sommées, pour chaque distance, les valeurs de $\Gamma(d)$ sur l'ensemble des secteurs.

Source : Insee - Calculs des auteurs.

L'étendue des phénomènes de concentration est-elle la même en France et au Royaume-Uni ?

Duranton et Overman (2005) ne donnent pas de résultats par secteur sur l'étendue spatiale des phénomènes de concentration. Ils reportent cependant deux types d'information : 1) pour chaque distance d , le pourcentage de secteurs qui apparaissent comme concentrés (équivalent de la figure 6), 2) pour les dix secteurs les plus concentrés et dispersés, la valeur des coefficients Γ et Ψ .

Des secteurs parfois concentrés à plus longue distance ...

Le tableau 1 réplique le tableau 1 de l'article de Duranton et Overman (2005) et renseigne à 5, 30, et 120 km le nombre de secteurs concentrés. Les phénomènes de concentration sont relativement persistants spatialement. 34% des secteurs sont concentrés à 5 km, mais 10 % des secteurs sont concentrés à la fois à 5, 30 et 150km. En France, un certain nombre de secteurs ont des formes de concentration à courte distance, inférieure à 30 km, comme le trouvent Duranton et Overman (2005) pour le Royaume-Uni. Par contre, contrairement au Royaume-Uni, en France, une part plus importante de secteurs possèdent des formes de concentration à distance moyenne (entre 30 et 150 km). Ainsi, 10% des secteurs apparaissent concentrés à 150 km uniquement. Cette proportion reste stable à 245 km.

La France a une distribution spatiale des activités économiques plus étendue que celle du Royaume-Uni, comme le suggère la comparaison des distances médianes entre tous les établissements : 395 km dans le premier cas, contre seulement 180 km dans le second. Cette dispersion plus grande de l'activité économique en France se répercute en partie dans les formes de concentration sectorielle. Aussi certains secteurs manufacturiers fuient la pression foncière des grandes agglomérations pour se localiser en lointaine périphérie. Gilli (2006) illustre pour la région parisienne ces phénomènes de périurbanisation des emplois manufacturiers.

TAB. 1 – Part de secteurs manufacturiers concentrés aux distances de 5, 30 et 150 km

5km	5km seul	5 & 30 km seul	5 & 150 km seul	5 & 30 & 150 km
34 %	0 %	24 %	0 %	10 %
30km	30km seul	30 & 150 km seul		
38 %	2 %	2 %		
150km	150km seul			
22 %	10 %			

Note de lecture : 34 % des secteurs manufacturiers sont significativement concentrés à 5 km. Parmi ces secteurs, aucun n'est concentré à 5 km uniquement. En revanche, 24 % des secteurs sont concentrés à la fois à 5 et 30 km. Source : Insee - Calculs des auteurs.

... mais les plus concentrés le restent à très courte distance

Néanmoins, les secteurs les plus concentrés restent en France, comme au Royaume-Uni, concentrés à relativement faible distance. Ceci suggère que les externalités de production qui jouent dans ces secteurs ne s'exercent que dans un périmètre restreint.

Le tableau 2 reprend les 10 secteurs les plus concentrés tels qu'ils sont présentés dans le tableau 2 de Duranton et Overman (2005). Nous reportons la valeur de concentration mesurée par ces auteurs (Γ_{DO}) et calculons la valeur équivalente sur les données françaises. Nous donnons pour chaque secteur la valeur de l'indicateur sur l'intervalle (0-180 km) (Γ_{0-180}), puis sur l'intervalle complet (Γ). Le tableau 2 de Duranton et Overman (2005) ne permet pas de conclure sans équivoque sur la distance de concentration d'un secteur. Il serait intéressant par exemple de connaître la contribution des intervalles (0-5km), (5-30km), (30-100km) et (100-180km) dans la valeur de l'indice de concentration.

TAB. 2 – Les 10 secteurs les plus concentrés au Royaume-Uni

Rg RU	SIC	Secteur	Γ_{DO}	Γ	Γ_{0-180}	Part	Rg FR
1	2214	Édition d'enregistrements sonores	0,470	0,468	0,468	100 %	3
2	1711	Filature de l'industrie cotonnière	0,411	0,086	0,086	100 %	27
3	2231	Reproduction d'enregistrements sonores	0,403	0,150	0,150	100 %	16
4	1760	Fabrication d'étoffes à maille	0,321	0,117	0,117	100 %	22
5	1713	Préparation et filature de l'industrie lainière	0,319	0,075	0,075	100 %	31
6	2861	Fabrication de coutellerie	0,314	0,309	0,225	73 %	8
7	1771	Fabrication de bas et chaussettes	0,290	0,073	0,028	39 %	35
8	1810	Fabrication de vêtements en cuir	0,203	0,245	0,245	100 %	12
9	1822	Fabrication de vêtements de dessus	0,181	0,346	0,346	100 %	5
10	2211	Édition de livres	0,178	0,319	0,319	100 %	7

Note de lecture : Rg RU=Rang du secteur au Royaume-Uni. Γ_{DO} = valeur du Γ reportée par Duranton et Overman (2005), calculée sur l'intervalle (0-180) km. Γ = valeur du Γ calculée sur données françaises, sur l'intervalle (0-392) km. Γ_{0-180} = valeur du Γ calculée sur données françaises, sur l'intervalle (0-180) km. $Part = \frac{\Gamma_{0-180}}{\Gamma} * 100$. Rg FR= Rang du secteur en France.

Source : Insee - Calculs des auteurs.

Nous pouvons cependant voir si, pour les mêmes secteurs en France, l'intervalle (0-180km) contribue pour une part importante de la valeur de l'indicateur de concentration mesurée sur l'intervalle global. Seuls quatre des dix secteurs les plus concentrés au Royaume-Uni sont aussi classés dans les 10 plus concentrés en France. Les autres secteurs restent cependant relativement concentrés, le rang n'excédant pas 35 (sur 229). On remarque que pour l'ensemble de ces secteurs (hormis la fabrication de coutellerie et la fabrication de bas et de chaussettes), la concentration a lieu à une distance inférieure à 180 km (Γ_{0-180} étant égal à Γ). On ne peut donc pas exclure que ces secteurs soient concentrés à une distance aussi faible au Royaume-Uni qu'en France. Dans ce cas, les phénomènes de concentration sont bien propres à chaque industrie et aux forces d'agglomération qui y sont à l'oeuvre plutôt qu'à la géographie économique du pays étudié.

Pour aller plus loin dans cette analyse, nous décomposons pour les 10 secteurs les plus concentrés et les plus dispersés en France la valeur de Γ par tranche de distances : (0-5km), (5-20km),

(20-50km), (50-100km), (100-200km), puis (200-392km). Pour l'ensemble des secteurs les plus concentrés (hormis le secteur de la cidrerie), la concentration a essentiellement lieu dans un rayon inférieur à 50 km (la somme de $\Gamma(d)$ pour les distances inférieures à 50 km représente plus de la moitié de la valeur globale du Γ). Ceci nous permet donc de confirmer que pour les secteurs les plus concentrés, les phénomènes de concentration ont lieu dans un espace inférieur à 50 km.

TAB. 3 – Les 10 secteurs les plus concentrés et les plus dispersés en France

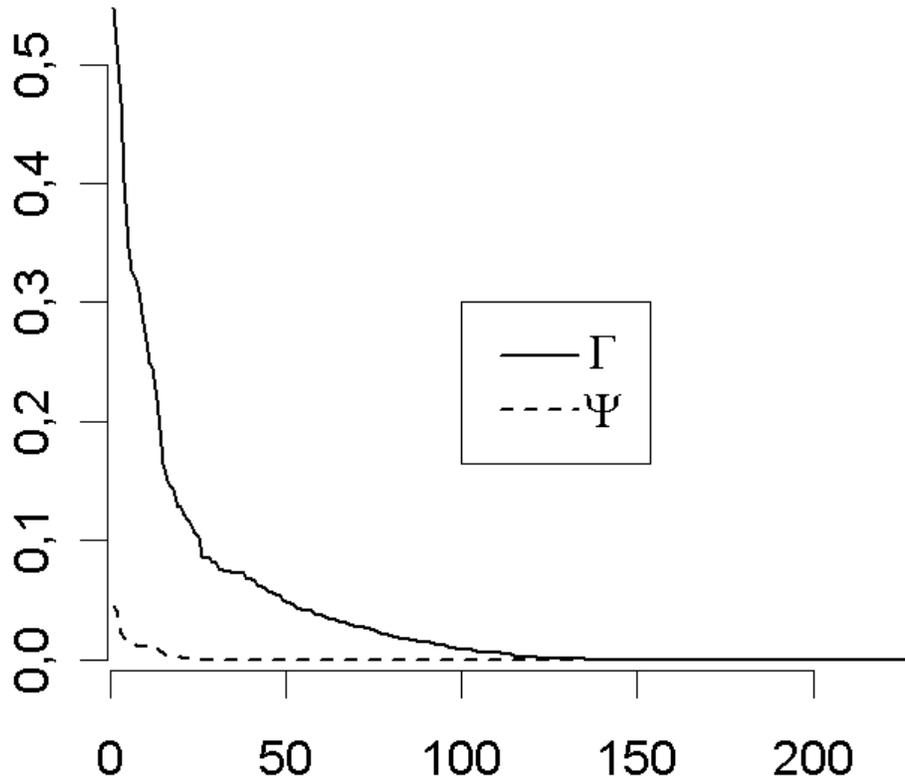
Les plus concentrés									
	SIC	Secteur	Γ	0-5	6-20	21-50	51-100	101-200	201-395
1	1724	Tissage de soieries	0,544	4 %	13 %	27 %	41 %	14 %	0 %
2	1714	Préparation et filature du lin	0,507	2 %	7 %	15 %	24 %	40 %	10 %
3	2214	Édition d'enregistrements sonores	0,468	15 %	39 %	40 %	6 %	0 %	0 %
4	1531	Transformation de pommes de terre	0,404	4 %	13 %	26 %	36 %	21 %	0 %
5	1822	Fabrication de vêtement de dessus	0,346	24 %	54 %	22 %	0 %	0 %	0 %
6	1715	Texturation de la soie	0,327	4 %	13 %	25 %	35 %	23 %	0 %
7	2211	Édition de livres	0,319	16 %	43 %	38 %	3 %	0 %	0 %
8	2861	Fabrication de coutellerie	0,312	8 %	21 %	30 %	10 %	7 %	25 %
9	1594	Cidrerie	0,287	1 %	3 %	9 %	23 %	49 %	15 %
10	2213	Édition de revues et périodiques	0,269	20 %	48 %	31 %	1 %	0 %	0 %
Les plus dispersés									
	SIC	Secteur	Ψ	0-5	6-20	21-50	51-100	101-200	201-395
1	1552	Fabrication de glaces	0,044	0 %	0 %	0 %	0%	6 %	94%
2	1520	Industrie du poisson	0,039	0 %	0 %	0%	0%	0%	100%
3	1582	Biscotterie	0,022	2 %	5 %	3 %	0 %	8 %	83 %
4	2812	Fabrication de menuiseries	0,019	2 %	4 %	0 %	0 %	17 %	76 %
5	2663	Fabrication de béton	0,013	14 %	36 %	35 %	1 %	0 %	15 %
6	1586	Transformation du thé/café	0,012	0 %	0 %	3 %	15 %	0 %	81 %
7	1740	Fabrication d'articles textiles	0,012	0 %	0 %	0 %	1 %	76 %	23 %
8	1584	Chocolaterie	0,011	13 %	37 %	39 %	4 %	5 %	3 %
9	1533	Transformation de fruits et légumes	0,011	0 %	0 %	0 %	0 %	10 %	90 %
10	3613	Fabrication de meubles de cuisine	0,010	6 %	16 %	16 %	0 %	0 %	62 %

Note de lecture : Dans le haut du tableau les secteurs sont classés par ordre décroissant de concentration. Dans le bas du tableau les secteurs sont classés par ordre décroissant de dispersion. Pour le secteur « Tissage de soierie » l'indice de concentration globale vaut 0,544. Les distances de 0 à 5 km contribuent pour 4% à la valeur de l'indice global.

Source : Insee - Calculs des auteurs.

Hétérogénéité sectorielle

La figure 8 présente la distribution des valeurs de Γ (resp. Ψ), en classant les secteurs par valeur croissante. En France, comme au Royaume-Uni, la distribution est particulièrement dissymétrique : peu de secteurs ont de très grandes valeurs et beaucoup de secteurs ont des valeurs faibles. Les indicateurs discrets de type Ellison-Glaeser et Maurel-Sédillot présentent également des distributions particulièrement dissymétriques.

FIG. 8 – Distribution des indices de concentration (Γ) et de dispersion (Ψ) pour la France

Source : Insee - Calculs des auteurs.

Homogénéité au sein des secteurs en SIC à deux chiffres

Le tableau 4 donne le pourcentage de secteurs concentrés au sein de chaque SIC à deux chiffres. La principale différence avec les résultats pour le Royaume-Uni est que la moitié des secteurs de l'industrie alimentaire sont concentrés alors qu'ils sont tous (sauf un) dispersés au Royaume-Uni. En revanche, il apparaît comme pour Duranton et Overman (2005) qu'il y a des tendances par industrie : pour certaines industries l'ensemble des secteurs sont concentrés. Notons que les industries pour lesquelles les tendances sont les plus marquées sont communes à la France et au Royaume-Uni, sauf pour le travail du bois, où 5 des 6 secteurs sont concentrés en France alors qu'aucun ne l'est au Royaume-Uni.

TAB. 4 – Nombre de secteurs concentrés au sein de chaque industrie SIC2 en France

SIC2	Nom	Nb Secteurs	% Concentrés
15	Industries alimentaires	29	52%
16	Industrie du tabac	1	0%
17	Industrie textile	21	81%
18	Industrie de l'habillement et des fourrures	6	83%
19	Industrie du cuir et de la chaussure	3	100%
20	Travail du bois et fabrication d'articles en bois	6	83%
21	Industrie du papier et du carton	7	57%
22	Édition, imprimerie, reproduction	13	92%
24	Industrie chimique	19	37%
25	Industrie du caoutchouc et des plastiques	7	71%
26	Fabrication d'autres produits minéraux non métalliques	24	21%
27	Métallurgie	14	57%
28	Travail des métaux	16	62%
29	Fabrication de machines et d'équipements	22	55%
30	Fabrication de machines de bureau et de matériel informatique	2	100%
31	Fabrication de machines et appareils électriques	7	71%
32	Fabrication d'équipements de radio, télévision et communication	3	100%
33	Fabrication d'instruments médicaux, de précision	5	100%
34	Industrie automobile	3	100%
35	Fabrication d'autres matériels de transport	8	50%
36	Fabrication de meubles ; Industries diverses	13	62%

Source : Insee - Calculs des auteurs.

Conclusions intermédiaires

Que pouvons-nous retenir des similarités et différences entre la France et le Royaume-Uni ? Nous retrouvons tout d'abord les quatre conclusions principales de Duranton et Overman (2005): (i) le nombre de secteurs apparaissant concentrés avec un indicateur en continu est notablement plus faible qu'avec un indicateur discret de type Ellison-Glaeser ou Maurel-Sédillot, (ii) la concentration a lieu à relativement faible distance, (iii) il existe une très grande hétérogénéité dans les valeurs des indicateurs de concentration, (iv) au sein d'un même sur-secteur (SIC à 2 chiffres), les formes de concentration des secteurs en SIC à 4 chiffres sont relativement homogènes.

Néanmoins, les deux pays ne sont pas similaires en tout point. S'il est vrai que pour les secteurs les plus concentrés, la concentration a lieu à relativement courte distance ($\leq 50km$), il existe en France un ensemble de secteurs concentrés à moyenne distance (de 50 à 150 km). Ceci est certainement dû aux différences dans l'organisation géographique des deux pays. De plus, la France a connu ces dernières années un desserrement des pôles manufacturiers, notamment en région parisienne, vers des zones péri-urbaines.

Enfin, les secteurs concentrés en France et au Royaume-Uni ne sont pas toujours identiques. Seuls quatre des 10 secteurs les plus concentrés au Royaume-Uni appartiennent aussi au groupe de tête en France. Que faut-il y voir ? Certaines de ces différences sont la conséquence de l'Histoire et, pour partie, de choix de spécialisation régionale différents. Concernant les autres secteurs, il est envisageable que ceux-ci n'en soient pas au même stade dans leur dynamique de concentration dans les deux pays.

4 Analyse comparée de la concentration des secteurs manufacturiers et de services manufacturiers aux entreprises

Dans cette partie, nous abandonnons la comparaison France-Royaume-Uni pour nous consacrer à la comparaison entre les secteurs manufacturiers et les secteurs de services en France. L'étude de la concentration dans les secteurs de services est assez peu développée jusqu'ici. Houdebine (1999) montre ainsi que les services sont en moyenne plus concentrés que les secteurs manufacturiers sur la période 1978-1992²¹. Or, comme nous l'avons souligné en introduction, les services représentent une part importante de l'activité française²². En outre, le développement des nouvelles technologies de l'information a pu considérablement influencer leur choix de localisation.

Nous utilisons dorénavant la classification NAF700 révision 1 de 2003. Nous disposons de 309 secteurs manufacturiers et de 96 secteurs de services²³. Ces secteurs sont de tailles hétérogènes (en terme de nombres d'établissements) comme le montre le tableau suivant.

TAB. 5 – Nombre de secteurs par tranche de taille

	entre 10 et 100 étab.	entre 100 et 10000 étab.	plus de 10000 étab.
Manufacturiers	125	183	1
Services	7	76	13

Source : Insee - Calculs des auteurs.

Comparaison avec les indicateurs d'Ellison-Glaeser et de Maurel-Sédillot

Nous commençons par comparer les résultats obtenus à l'aide de l'indicateur de Duranton et Overman (2005) avec des indicateurs d'Ellison-Glaeser et de Maurel-Sédillot. Ces derniers sont insensibles à la permutation des zones géographiques les unes par rapport aux autres, comme le montre l'exemple de la partie 1.2. L'indicateur de Duranton et Overman (2005) ne souffre pas de cette limite. Potentiellement, ces deux groupes d'indicateurs donnent donc des images très différentes de la concentration sectorielle.

Le tableau 6 fournit la corrélation de rang entre les indicateurs d'Ellison-Glaeser et de Maurel-Sédillot (calculés à la commune, à la zone d'emploi et au département), en excluant les secteurs à Γ nul (i.e. ceux ne présentant pas de concentration globale au sens de Duranton et Overman (2005)²⁴. Il apparaît clairement que, quel que soit le niveau géographique choisi, la corrélation de rang entre l'indicateur en continu et les indicateurs discrets reste relativement faible (égale à 0,53 dans le meilleur des cas). L'utilisation d'un indicateur en continu modifie donc substantiellement les résultats. En outre, nous retrouvons à travers ces tableaux la sensibilité des indicateurs discrets au choix du découpage géographique (*Modifiable Areal Unit Problem*). Notons que le

²¹Kolko (2000) et Kolko (2007) sont deux études plus extensives sur la concentration des services aux États-Unis.

²²Les services représentent 48% de la production et 47% des emplois salariés intérieurs.

²³Une limite à notre étude tient aux différences dans la finesse de classification entre les activités manufacturières et de services.

²⁴Ces résultats sont robustes à l'introduction des secteurs à Γ nul.

passage d'un espace discret à un espace en continu a un impact plus marqué sur le classement des secteurs que le passage d'un découpage géographique à un autre avec les indicateurs discrets. Autrement dit, le problème de l'insensibilité aux permutations des unités géographiques dont souffrent les indicateurs discrets en comparaison de l'indicateur en continu corrompt plus notre perception des phénomènes de concentration que le problème de la forme et de la taille des unités géographiques. Ceci justifie l'étude de la concentration à l'aide de l'indicateur de Duranton et Overman (2005), qui présente de meilleures propriétés.

TAB. 6 – Corrélation de rang entre l'indicateur en continu (Γ) et les indicateurs d'Ellison-Glaeser (EG) et Maurel-Sédillot (MS) : exclusion des secteurs à Γ nul

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MS (communes) (1)	1,00	0,79	0,81	0,55	0,75	0,51	0,43
EG (communes) (2)		1,00	0,66	0,80	0,67	0,76	0,47
MS (zones d'emploi) (3)			1,00	0,72	0,88	0,62	0,48
EG (zones d'emploi) (4)				1,00	0,73	0,93	0,45
MS (départements) (5)					1,00	0,78	0,53
EG (départements) (6)						1,00	0,48
Gamma (7)							1,00

Nombre d'observations : 252. Sont exclus les secteurs avec $\Gamma = 0$. Toutes les corrélations sont significatives à 1%.
Source : Insee - Calculs des auteurs.

Choix du contrefactuel

Pour comparer directement les secteurs manufacturiers avec les secteurs de services, il est nécessaire de disposer du même contrefactuel, c'est-à-dire du même ensemble des possibles pour la distribution aléatoire des établissements lors du calcul des intervalles de confiance. Dans cette partie, nous changeons donc d'ensemble des sites actifs. Jusqu'à la fin de cette partie, l'ensemble des sites actifs considérés pour construire le contrefactuel est donc l'ensemble des sites *manufacturiers et de services* actifs au 1^{er} janvier 2005²⁵. Le contrefactuel et la classification sectorielle ayant changé, il est possible que les conclusions concernant la concentration des secteurs manufacturiers soient légèrement différentes de celles exposées précédemment.

Le tableau 7 renseigne sur le pourcentage de secteurs manufacturiers et de services qui dévient de la distribution aléatoire. Ce pourcentage est plus grand dans les services (90%) que dans le manufacturier (67%). De plus, les services apparaissent relativement plus concentrés (75%) que les secteurs manufacturiers (56%)²⁶.

²⁵Précédemment, l'ensemble des sites possibles était uniquement l'ensemble des sites manufacturiers

²⁶Rappelons qu'avec la classification SIC92, 60% des secteurs manufacturiers apparaissaient concentrés. Le changement de contrefactuel et de nomenclature sectorielle n'a donc qu'un impact limité sur les résultats.

TAB. 7 – Concentration et dispersion des secteurs manufacturiers et de services

	Dispersion	Aléatoire	Concentration
Manuf.	11 %	32 %	56 %
Services	15 %	10 %	75 %

Note de lecture : 13% des secteurs manufacturiers sont significativement dispersés au sens de Duranton et Puga (2005).

Source : Insee - Calculs des auteurs.

Quelle que soit l'industrie en NES16 considérée, les services dévient plus souvent de la distribution aléatoire en étant majoritairement concentrés, en particulier pour les secteurs appartenant aux activités financières (82%). Parmi les activités manufacturières, la tendance est moins claire : les industries de biens de consommation semblent les plus concentrées (tableau 8). Kolko (2007) observe également ce phénomène sur données américaines en utilisant un indicateur d'Ellison-Glaeser.

TAB. 8 – Concentration et dispersion : détails par NES16

	Disp.	Aleatoire	Conc.
Industries agricoles et alimentaires	21 %	33 %	47 %
Industrie des biens de consommation	6 %	20 %	75 %
Industrie automobile	25 %	0 %	75 %
Industries des biens d'équipement	16 %	34 %	49 %
Industries des biens intermédiaires	8 %	36 %	56 %
Transports	22 %	4 %	74 %
Activités financières	6 %	12 %	82 %
Services aux entreprises	13 %	13 %	73 %

Note de lecture : 23% des secteurs des industries agricoles et alimentaires sont significativement dispersés.

Source : Insee - Calculs des auteurs.

Le tableau 9 présenté pour les secteurs manufacturiers et de services sont comparables à celui donné au niveau de SIC92 dans la partie 1. Les secteurs de services sont concentrés à plus courte distance, 45% le sont à 5 km contre seulement 19% pour les secteurs manufacturiers. De plus, la persistance spatiale des phénomènes de concentration dans les services est moins importante que dans le manufacturier : 15% des secteurs de services sont concentrés à 5 km uniquement. Ensuite, aucun des secteurs de services qui apparaît concentré à 5km ne le reste au delà de 50 km.

TAB. 9 – Part de secteurs concentrés aux distances de 5, 50 et 245 km

Secteurs manufacturiers				
5km	5km S	5 & 50 km S	5 & 245 km S	5 & 50 & 245 km
19 %	1 %	16 %	0 %	2 %
50km	50km S	50 & 245 km S		
24 %	4 %	2 %		
245km	245km S			
22 %	17 %			
Secteurs de services				
5km	5km S	5 & 50 km S	5 & 245 km S	5 & 50 & 245 km
45 %	15 %	30 %	0 %	0 %
50km	50km S	50 & 245 km S		
36 %	6 %	0 %		
245km	245km S			
17 %	17 %			

Note de lecture : 19% des secteurs manufacturiers sont concentrés à 5km; 1% des secteurs manufacturiers sont concentrés seulement à la distance de 5km.

Source : Insee - Calculs des auteurs.

Pourquoi les services sont-ils plus concentrés que les secteurs manufacturiers ?

Les secteurs de services sont moins contraints dans leur choix de localisation par l'accès à des ressources naturelles que peuvent l'être certains secteurs manufacturiers. La concentration dans ces secteurs est donc la résultante de forces d'agglomération et de dispersion. La question est alors de comprendre en quoi ces forces peuvent être différentes dans les services et le manufacturier.

Comme évoquée en introduction, la première force de dispersion concerne les coûts à l'échange et l'accès à la demande. Il est possible que l'introduction des nouvelles technologies de l'information et de la communication ait réduit les coûts de transfert de manière plus importante dans certains services que dans le manufacturier. Les échanges générés par les secteurs manufacturiers nécessitent le transport d'un bien physique, alors que nombre de produits du secteur des services sont immatériels et donc peu coûteux à transporter. Ainsi, dans le secteur des études de marché et sondage (un des secteurs de service les plus concentrés), on peut supposer qu'un grand nombre des échanges entre les clients et le fournisseur peut se faire via des moyens de communication modernes et peu coûteux.

Une seconde force de dispersion concerne les effets de concurrence. Dès lors que les coûts aux échanges se réduisent pour un secteur, le marché pour ses produits s'étend, la concurrence locale s'efface, devient nationale voire internationale. Dans ce cas, les effets de concurrence locale qui peuvent être un frein à l'agglomération sont relâchés.

Concernant les forces d'agglomération, certains services reposent sur des inputs spécialisés, notamment une main-d'œuvre aux compétences très particulières. Ces services peuvent alors être

incités à se regrouper spatialement pour partager un ensemble plus large de travailleurs ayant ces compétences. Les secteurs de services, notamment les services financiers et une partie des services aux entreprises, emploient une main-d'œuvre particulièrement qualifiée : en effet, pour ces secteurs la part des travailleurs qualifiés est plus du double de la part moyenne sur l'ensemble des secteurs. Almazan, de Motta et Titman (2007) proposent un modèle où les besoins en capital humain d'une industrie déterminent ses choix de localisation. Ils montrent que, dans un cluster, le coût de formation de la main-d'œuvre et le coût d'ajustement suite à un choc sont plus faibles mais que la concurrence sur le marché du travail y est plus âpre.

Enfin, il est envisageable que les spillovers de connaissance jouent un rôle plus important dans les services que dans les secteurs manufacturiers traditionnels. Duranton et Puga (2001) montrent comment les secteurs où l'échange d'idées est primordiale se localisent préférentiellement dans le coeur des villes, alors que les industries plus matures, ne profitant plus des spillovers de connaissance, se relocalisent en périphérie de ces villes, où la pression foncière est moins forte.

Conclusions intermédiaires

Trois messages importants ressortent de cette partie :

- le classement des secteurs est très sensiblement modifié lorsque l'on passe d'un indicateur discret à un indicateur en continu.
- les secteurs de services dévient plus souvent de la distribution aléatoire que les secteurs manufacturiers, majoritairement dans le sens d'une plus grande concentration.
- parmi les secteurs concentrés, les secteurs de services le sont généralement à plus courte distance que les secteurs manufacturiers, ce qui peut s'expliquer par des gains plus élevés à la concentration et/ou des moindres coûts de transfert.

L'organisation spatiale des secteurs de services n'apparaît pas dans le débat de politique publique, les pôles de compétitivité relevant avant tout de la politique industrielle. L'analyse précédente montre que du point de vue politique, la question des services est certainement à aborder à une échelle géographique plus fine, par exemple celle de la ville.

5 Dynamique de concentration entre 1996 et 2005

Houdebine (1999) analyse l'impact des créations et destructions d'établissements sur la dynamique de concentration spatiale des activités. Il montre que deux facteurs interdépendants déterminent cet impact : (i) la structure de localisation des créations et des destructions d'une part (les créations sont-elles plus ou moins concentrées que le reste du secteur ?) et (ii) la structure de localisation de ces créations (ou destructions) relativement au stock d'établissements pérennes (les établissements créés le sont-ils dans des zones déjà spécialisées dans le secteur d'activité correspondant ?).

Pour illustrer cette idée, reprenons l'exemple proposé par Houdebine (1999). La région A possède un établissement et la région B, deux établissements, tous les trois dans le même secteur d'activité. On suppose qu'un des trois établissements disparaît. S'il s'agit de l'établissement de la région A, la concentration géographique du secteur augmentera. Au contraire, s'il s'agit de l'un des établissements de la région B, la concentration géographique diminuera. La différence réside bien dans l'interaction entre les localisations des établissements pérennes et des établissements détruits : faible dans le premier cas, plus élevée dans le second.

Houdebine (1999) montre ainsi une croissance de la concentration géographique des activités au niveau des zones d'emploi entre 1990 et 1992. Celle-ci est due au choix des établissements nouvellement créés de se localiser dans les zones d'emploi déjà dynamiques dans leur secteur, alors que la distribution spatiale des établissements qui disparaissent est moins corrélée à celle des établissements pérennes.

Dans ce qui suit, nous nous proposons d'étudier cette question sur une période plus longue 1996-2005, en retenant une approche légèrement différente de celle de Houdebine (1999). En effet, pour étudier la dynamique de concentration d'un secteur, l'identité des établissements n'a que peu d'importance. Nous ne nous intéressons donc pas dans la suite à l'entrée et à la sortie d'établissements mais plutôt à l'entrée et à la sortie de sites actifs. Étant contraints à localiser les établissements au centre de chaque commune, nous étudions donc l'évolution sur trois dates 1996, 2000 et 2005 du nombre de sites actifs par secteur et commune. Autrement dit, si dans un même secteur et une même commune, un établissement en remplace un autre, ceci n'aura aucun impact sur l'évolution de la concentration du secteur.

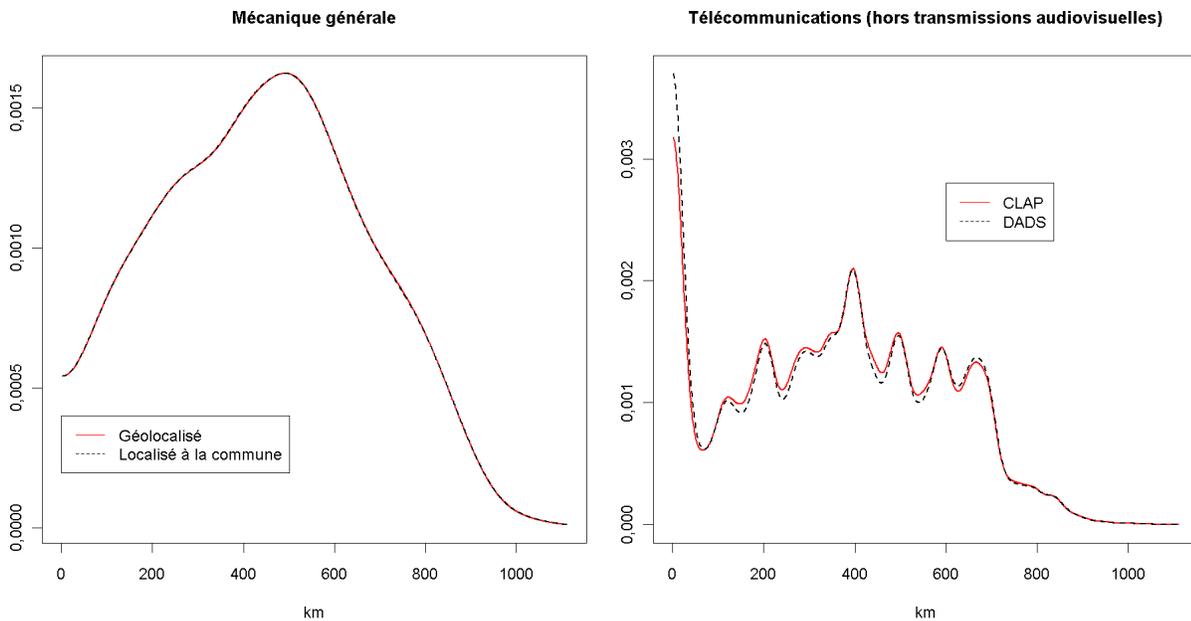
5.1 Comparaison avec les différentes données

La localisation fine de l'emploi à partir de CLAP, ainsi que la géolocalisation, ne sont pas disponibles sur longue période. Pour étudier la dynamique des phénomènes de concentration, nous utilisons donc les DADS (Déclaration Annuelle de Données sociales) sur la période 1996-2005. Cette source administrative à l'avantage d'être exhaustive pour l'ensemble des établissements ayant au moins un salarié. Ces données présentent cependant deux inconvénients par rapport à celles utilisées dans la partie précédente (Clap 2005). D'une part, les établissements sont localisés au centre de la commune. D'autre part, le dégroupement n'est pas effectué, c'est-à-dire qu'une entreprise multi-établissements peut légalement déclarer l'ensemble de ses salariés dans un seul de ses établissements. Ceci fait disparaître de la base des établissements auxquels aucun salarié n'est associé. Les entreprises ne sont pas contraintes de choisir le même établissement

pour déclarer leurs salariés d’une année sur l’autre, le problème du dégroupement peut donc être particulièrement délicat lors de l’étude de la dynamique.

Afin d’avoir une idée de l’impact de ces approximations sur les résultats trouvés, nous calculons les distributions $\hat{K}(d)$ pour trois jeux de données de la même année : CLAP 2005 géolocalisé, CLAP 2005 localisé à la commune et les DADS 2005 localisées à la commune. Nous effectuons deux comparaisons : CLAP géolocalisé avec CLAP localisé à la commune, puis CLAP localisé à la commune avec les DADS. La première comparaison nous permet d’obtenir l’effet de la moindre précision de la localisation, et la seconde l’effet du dégroupage. Pour les deux comparaisons, nous cherchons le secteur pour lequel la différence relative²⁷ est la plus grande. Les résultats sont présentés dans la figure 9.

FIG. 9 – Effet de la localisation à la commune et du dégroupement



Source : Insee - Calculs des auteurs.

La perte de précision résultant de la localisation à la commune ne modifie pas visiblement la distribution des distances entre les paires d’établissements (figure 9, graphique de gauche). A contrario, la distribution semble plus sensible au dégroupage. Cependant même si les courbes ne sont pas parfaitement confondues, l’allure générale reste préservée. Notons que pour la plupart des secteurs le dégroupage ne modifie pratiquement pas la distribution, nous représentons ici les secteurs pour lequel la différence entre les sources est la plus grande. Deux secteurs paraissent particulièrement problématiques : les « Télécommunications (hors transmissions audiovisuelles) » et les « Postes nationales »²⁸. Ces deux secteurs sont éliminés dans l’analyse

²⁷Somme des carrés des différences entre les deux courbes à chaque kilomètre.

²⁸Rathelot et Sillard (2007) montrent également que ces deux secteurs sont problématiques dans leur étude de l’impact des Zones Franches Urbaines (ZFU) sur l’emploi salarié et la création d’établissements.

ultérieure. Pour 23% des secteurs, il n’y a aucune différence entre la distribution calculée à partir de CLAP et celle calculée à partir des DADS (figure 9).

5.2 Évolution de la concentration entre 1996 et 2005

Nous calculons les indicateurs de concentration ($\Gamma(d)$ et Γ) et de dispersion ($\Psi(d)$ et Ψ) pour trois années : 1996, 2000 et 2005²⁹. Nous calculons tout d’abord les indices sectoriels (Γ et Ψ) afin de voir si le classement des secteurs les plus concentrés (respectivement les plus dispersés) varie au cours du temps. Nous ajoutons à cette comparaison temporelle une comparaison pour 2005 entre les résultats issus des données de CLAP et ceux issus des DADS afin de mesurer l’effet de la source utilisée.

Les corrélations de rang entre les différentes années (tableaux 10 et 11) sont relativement élevées. Néanmoins, il semble y avoir quelques changements dans le classement des secteurs les plus concentrés sur 10 ans, puisque la corrélation de rang entre 1996 et 2005 est plus faible, s’élevant à 0,84. Pour les indicateurs de dispersion (tableau 11), ces changements semblent plus marqués puisque la corrélation tombe à 0,61 entre 1996 et 2005. En revanche, notons que la source utilisée importe assez peu, les corrélations entre les classements calculés à partir de CLAP 2005 et les DADS 2005 s’élèvent à 0,98 pour les indicateurs de concentration et 0,90 pour la dispersion.

TAB. 10 – Corrélations de rang de l’indice de concentration sectoriel (Γ)

	(1)	(2)	(3)	(4)
DADS 1996 (1)	1,00	-	-	-
DADS 2000 (2)	0,92	1,00	-	-
DADS 2005 (3)	0,84	0,90	1,00	-
CLAP 2005 (4)	-	-	0,98	1,00

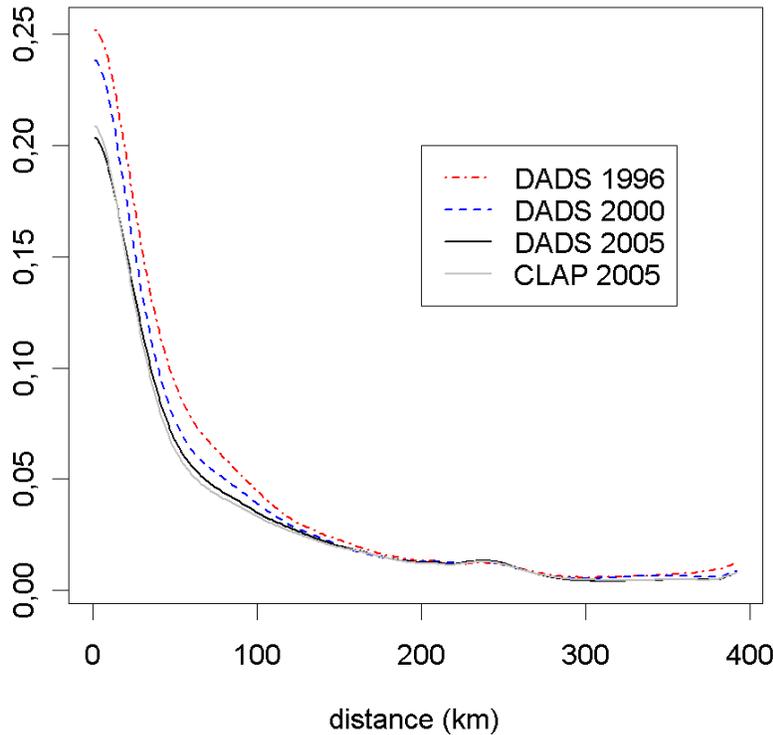
TAB. 11 – Corrélations de rang de l’indice de dispersion sectoriel (Ψ)

	(1)	(2)	(3)	(4)
DADS 1996 (1)	1,00	-	-	-
DADS 2000 (2)	0,69	1,00	-	-
DADS 2005 (3)	0,61	0,78	1,00	-
CLAP 2005 (4)	-	-	0,90	1,00

²⁹Pour des raisons de temps de calcul, nous avons pour l’instant utilisé les intervalles de confiance de 2005. Rappelons que pour un secteur donné, l’intervalle de confiance dépend de la distribution des distances bilatérales de l’ensemble des établissements (tous secteurs confondus) et du nombre d’établissements dans le secteur. La distribution des distances bilatérales de l’ensemble des établissements étant relativement similaire sur les trois années, seule la variation de taille du secteur peut modifier l’intervalle de confiance sur les trois années. Nous avons donc associé les secteurs de 1996, 2000 et 2005 (données DADS) à des secteurs de 2005 (données CLAP) de même taille. Cette méthode n’est pas parfaite mais permet d’avoir une première idée des résultats pour la dynamique sans être obligé de faire tous les calculs d’intervalles de confiance. Les résultats présentés mériteraient cependant d’être confirmés avec un calcul exact des intervalles de confiance.

Nous nous intéressons ensuite à l'évolution des indicateurs par distance : $\Gamma(d)$ et $\Psi(d)$. Pour avoir une image de la variation globale de la concentration, nous sommes ces indicateurs sur l'ensemble des secteurs. Nous constatons alors une légère diminution, notamment aux distances les plus courtes, de l'indicateur de concentration ; ceci signifie que les secteurs concentrés le sont de moins en moins sur la période (figure 10). Remarquons tout de même que l'allure globale de la courbe reste inchangée sur la période, les phénomènes de concentration sont toujours dominants à courte distance. Pour les indicateurs de dispersion aucune tendance n'apparaît.

FIG. 10 – Évolution de $\Gamma(d)$



Source : Insee - Calculs des auteurs.

Ces premiers résultats nous amènent plutôt à conclure que, si l'image prédominante est celle d'une assez grande immobilité des activités économiques sur la période, les secteurs les plus concentrés enregistrent tout de même une légère tendance à la déconcentration sur la période 1996-2005. Toutefois, à concentration relativement stable, il est possible qu'il y ait des phénomènes de compensation entre les entrées et sorties de sites actifs. Autrement dit, si les évolutions nettes sont assez faibles, il est possible que les entrées et sorties de sites actifs soient davantage marquées spatialement. Pour reprendre l'analyse d'Houdebine (1999), si ces entrées et sorties ne modifient que légèrement la concentration globale du secteur, elles peuvent par contre en modifier plus largement la « géographie » si la localisation de ces entrées et sorties n'est que peu corrélée avec celle des sites pérennes. Pour tester ce dernier point, nous utilisons le principe développé par Durantou et Overman (2008).

5.3 Évolution de la localisation de l'activité sectorielle

L'idée de Duranton et Overman (2008) est de tester l'influence de la distribution des distances bilatérales impliquant des établissements entrants (ou sortants) sur la distribution de l'ensemble des distances bilatérales. Les distances bilatérales impliquant des établissements entrants (respectivement sortants) sont de deux types : celles entre établissements entrants uniquement (respectivement sortants uniquement) et les distances bilatérales entre un entrant (respectivement sortant) et un établissement pérenne. Duranton et Overman (2008) étudient séparément ces deux ensembles de distances. Pour étudier le premier, ils calculent le K-estimateur pour les entrants uniquement ; pour le second, ils calculent le K-estimateur de co-localisation³⁰ entre les entrants (sortants) et les pérennes. Dans les deux cas, le contrefactuel est défini par l'ensemble des sites occupés par les établissements d'un secteur³¹.

Pour se prémunir des difficultés inhérentes au suivi longitudinal des firmes, nous mesurons les entrants et les sortants à partir des variations du nombre d'établissements par secteur-commune. Ainsi, si pour un secteur et une commune donnés le nombre d'établissements augmente entre 1996 et 2000, nous considérons qu'il y a autant d'entrants que d'établissements supplémentaires dans cette commune et aucun sortant. Autrement dit, si au sein d'une commune et d'un secteur, un établissement est remplacé par un autre, cela n'affectera pas nos résultats. Par conséquent, les entrants sont donc des sites actifs qui apparaissent et les sortants des sites actifs qui disparaissent. Nous mesurons les entrants sur la période 1996-2000 et les sortants sur 2000-2005. Ainsi pour les deux catégories, la distribution de référence est la distribution des sites actifs en 2000.

Nous optons pour une méthode légèrement différente de celle proposée par Duranton et Overman (2008) puisque nous regroupons les deux ensembles de distances, entre entrants uniquement et entre entrants et pérennes. Plus précisément, nous comparons la distribution des distances bilatérales impliquant au moins un entrant (sortant) et la distribution de l'ensemble des distances. Le but est de mesurer l'impact des entrants (sortants) sur la distribution des distances. Or cet impact ne peut être déterminé qu'en prenant en compte toutes les distances inter-établissements induites par un entrant (sortant). Ainsi, on calcule la distribution suivante :

$$\hat{K}_e(d) = \frac{1}{(n_e(n_e - 1)/2 + n_e(n - n_e))h} \sum_{i=1}^{n_e} \sum_{j=i+1}^n f\left(\frac{d - d_{i,j}}{h}\right)$$

avec n_e le nombre d'entrants (ou sortants), les autres variables étant définies comme précédemment. Pour savoir si les entrants (ou les sortants) modifient significativement la distribution des distances bilatérales, nous redistribuons les entrants (ou des sortants) aléatoirement sur l'ensemble des sites actifs du secteur pour calculer l'intervalle de confiance³². Dans cette partie, nous nous limitons aux secteurs qui ont plus de 100 établissements les trois années (afin d'éviter

³⁰L'indicateur de co-localisation entre un groupe A composé de n_A établissements et un groupe B de n_B établissements se calcule à partir de la distribution suivante : $\hat{K}(d) = \frac{1}{n_A n_B h} \sum_{i \in A} \sum_{j \in B} f\left(\frac{d - d_{i,j}}{h}\right)$.

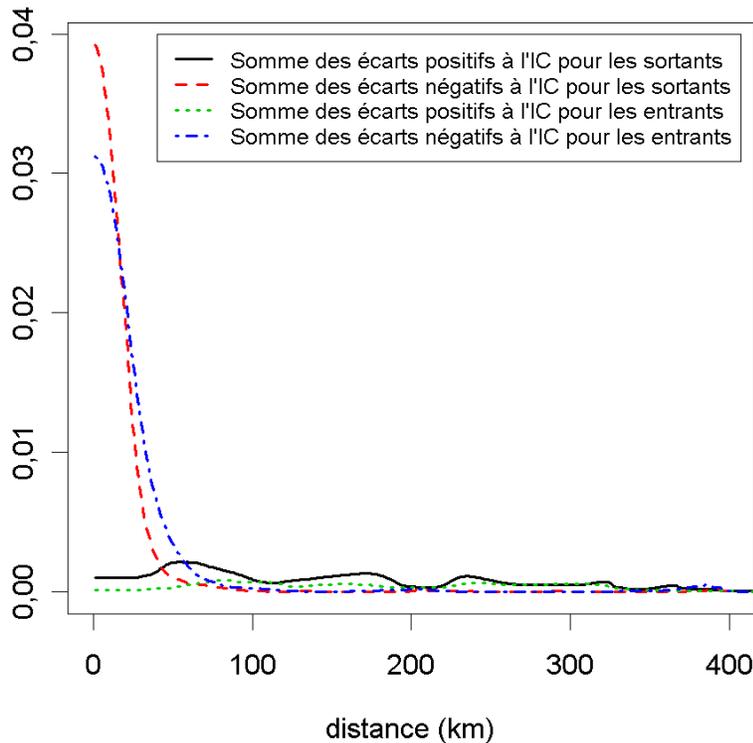
³¹Il y a ici un parallèle évident avec l'analyse de Houdebine (1999). Le K-estimateur entre entrants rend compte de la structure de localisation des créations (les créations sont-elles plus ou moins concentrées que le reste du secteur ?). L'indicateur de co-localisation rend compte de la structure de localisation de ces créations relativement aux stocks d'établissements pérennes (les établissements créés le sont-ils dans des zones déjà spécialisées dans le secteur d'activité correspondant ?).

³²La distance médiane pour calculer l'intervalle de confiance global est la distance médiane entre les établissements du secteur en 2000, elle est donc sectorielle.

les variations erratiques) et dont les entrants et les sortants représentent au moins 10 % des établissements en 2000. Nous effectuons également 1000 simulations pour chaque secteur. Au final, nous travaillons sur 252 secteurs.

Pour les secteurs dont la distribution des entrants (ou des sortants) dévie de l'intervalle de confiance, cette déviation est généralement à la baisse, c'est-à-dire que les distributions induites par les entrants ou les sortants sont sous l'intervalle de confiance. L'écart entre la distribution observée et l'intervalle de confiance est généralement plus grand à courte distance, entre 0 et 50 km (figure 11). Les entrants (resp. les sortants) tendent donc à déconcentrer (resp. concentrer) les secteurs, notamment à courte distance. La distribution des entrants (respectivement des sortants) est significativement inférieure à la distribution de l'ensemble des établissements entre 0 et 50km pour 33% (respectivement 20%) des secteurs. Ces secteurs sont le plus souvent des secteurs concentrés à moins de 50 km.

FIG. 11 – Somme sur les secteurs des écarts à l'intervalle de confiance



Source : Insee - Calculs des auteurs.

Ceci nous conduit à plusieurs conclusions. Il n'y a ni apparition, ni disparition de pôles d'activité spécialisés sur les 10 années étudiées. En effet, de tels phénomènes auraient pour conséquence de faire apparaître des pics de concentration pour les entrants ou les sortants à courte distance. Les entrants ne renforcent pas la concentration du secteur, autrement dit ces entrants ne se localisent pas dans les zones déjà spécialisées. On peut supposer que ceci est le signe soit d'un phénomène de congestion des pôles de spécialisation existants, soit de l'absence de bénéfices

privés à la localisation dans ces pôles (même si des bénéfices sociaux peuvent exister). Les sortants disparaissent davantage en dehors des zones spécialisées. Ceci peut laisser penser que le regroupement de sites en grappe les protège contre les chocs externes.

Pour illustrer cette dernière conclusion, nous régressons, secteur par secteur, la probabilité de sortie d'un site actif sur le nombre de sites actifs du même secteur à moins de 5 ou 10 km. Cette dernière variable a le plus souvent un effet significativement négatif (voir tableau 12), en particulier lorsque le secteur est significativement concentré à moins de 10 km. Ceci montre qu'il existe une corrélation négative entre le fait d'appartenir à un pôle et la probabilité de disparaître.

TAB. 12 – Effet du nombre d'établissements à moins de x km sur la probabilité de sortir

	Ensemble des secteurs			Secteurs concentrés à moins de 10km		
	Signif. négatif	Non signif.	Signif. positif	Signif. négatif	Non signif.	Signif. positif
5km	37%	57%	6%	62%	37%	1%
10km	28%	64%	8%	54%	43%	3%
15km	25%	64%	11%	51%	45%	4%

Note de lecture : Pour 37% des secteurs le nombre de sites actifs à moins de 5km a un effet significativement négatif sur la probabilité de disparition d'un site actif.

Source : Insee - Calculs des auteurs.

Au final, nous pouvons retenir de cette partie une assez grande stabilité dans la localisation des activités économiques en France sur la décennie passée. Cependant, une légère tendance à la déconcentration des secteurs les plus concentrés semble à l'œuvre. Ceci est la résultante de deux effets inverses : les entrants se localisent moins souvent dans les pôles spécialisés existants, les sortants disparaissent davantage lorsqu'ils sont loin de ces pôles. Les premiers étant plus nombreux que les seconds, la résultante est bien une légère déconcentration. Même si cette analyse n'a pas la force d'une preuve économétrique, elle semble montrer que les points de concentration n'attirent pas les entrants, mais peuvent mieux protéger contre les chocs externes. Cette dernière idée mériterait d'être analysée plus avant.

Conclusion et perspectives

La politique des *pôles de compétitivité*, mise en place en juillet 2005, repose sur l'idée que la concentration géographique d'un secteur d'activité, au sein de *clusters*, est source de performance et de croissance pour les acteurs qui les composent (voir Duranton, Martin, Mayer et Mayneris (2008) pour une analyse critique des arguments économiques sous-jacents aux politiques de clusters). L'intervention publique en matière d'organisation spatiale des secteurs d'activité est justifiée dès lors que l'on pense ou, encore mieux, que l'on prouve que les forces de marché n'atteignent pas l'optimum social. Avant toute intervention, il est donc nécessaire de se questionner sur l'étendue *naturelle* du phénomène de concentration géographique.

L'objectif de cette étude a donc été de quantifier la concentration sectorielle dans l'industrie manufacturière et les services en France, à l'aide d'un indicateur en continu développé par Duranton et Overman (2005). L'idée de cet indicateur est de comparer la distribution des distances bilatérales entre établissements d'un secteur à celle d'un secteur fictif possédant le même nombre d'établissements distribués de manière aléatoire sur l'ensemble des sites actifs. Cette méthodologie permet à la mesure de la concentration de s'abstraire du problème de la sensibilité au découpage géographique inhérent aux autres indicateurs utilisés jusque là (indices d'Ellison-Glaeser ou de Maurel-Sédillot). Nous utilisons donc la richesse des données géolocalisées disponibles à l'Insee pour construire cet indicateur dans le cas français.

Nous avons organisé notre étude autour de trois axes. Un premier a consisté à comparer la concentration des secteurs manufacturiers en France et au Royaume-Uni. Il apparaît que les quatre conclusions mises en exergue par Duranton et Overman (2005) restent valables pour la France : (i) le nombre de secteurs apparaissant concentrés avec un indicateur en continu est notablement plus faible qu'avec un indicateur discret de type Ellison-Glaeser ou Maurel-Sédillot, (ii) la concentration a lieu à relativement faible distance, (iii) il existe une très grande hétérogénéité dans les valeurs des indicateurs de concentration, (iv) au sein d'un même sur-secteur (SIC à 2 chiffres), les formes de concentration des secteurs en SIC à 4 chiffres sont relativement homogènes.

Néanmoins, un certain nombre de secteurs en France sont concentrés à moyenne distance 60-150 km, ce qui n'apparaît pas au Royaume-Uni. Nous associons cette différence à la fois aux différences de géographie entre les deux pays et à une tendance à la périurbanisation des industries manufacturières. Une analyse secteur par secteur montre ensuite que seuls quatre secteurs appartiennent à la fois aux dix secteurs les plus concentrés en France et au Royaume-Uni. Les secteurs les plus concentrés en France sont soit liés à des particularismes régionaux (cidrerie en Normandie), soit à un héritage historique (tissage de soie dans la région lyonnaise). Ce dernier point n'exclut pas pour autant qu'il existe des bénéfices à la concentration, une forme de rendements d'échelle dans ces secteurs.

Le deuxième axe a consisté à comparer les phénomènes de concentration dans l'industrie manufacturière et dans les services aux entreprises. Ces derniers semblent dévier plus souvent de la distribution aléatoire, surtout en faveur d'une plus forte concentration. De plus, cette concentration a généralement lieu à plus faible distance que pour les secteurs manufacturiers. Ceci est

conforme à l'idée que ces secteurs se localisent principalement dans le cœur des villes et, lorsqu'ils ne sont pas contraints par la localisation de leur demande, dans un petit nombre d'entre elles.

Le dernier axe s'intéresse à la dynamique de ces formes de concentration sur la période 1996-2005. Ne disposant pas de données géolocalisées sur l'ensemble de la période, nous montrons que le biais introduit par une analyse à la commune est négligeable. Si l'image globale reste celle d'une assez grande immobilité des secteurs d'activité, il est quand même possible de mettre en évidence une légère déconcentration des secteurs sur la période. Il apparaît ensuite que les sites actifs qui se créent et qui sont détruits sont en règle générale plus dispersés que l'ensemble des sites actifs du secteur. Les entrées étant plus nombreuses que les sorties, il est logique que ces flux bruts conduisent à une légère décroissance de la concentration sectorielle. Une simple statistique descriptive montre enfin que la probabilité de sortie d'un site actif est d'autant plus grande que celui-ci est éloigné du reste des sites actifs du secteur, suggérant que la résistance aux chocs est plus grande dans les zones de plus grande concentration du secteur.

Ce dernier exercice n'a pas force de preuve. Il nous semble qu'une extension intéressante de cette étude serait d'utiliser cette méthodologie pour caractériser les forces et les faiblesses face aux chocs externes des sites actifs en fonction de leur position relative au reste du secteur. En effet, si Duranton, Martin, Mayer et Mayneris (2008) émettent des doutes quant à la capacité des pôles de compétitivité à réellement être un moteur de performance et de croissance, ils évoquent l'idée que ceux-ci aident leurs acteurs à mutualiser les risques liés par exemple à la conjoncture internationale. Dans cette veine, Guido et Schivardi (2007) montrent comment les interactions sociales, au sein des *clusters*, réduisent les coûts à l'ajustement suite à des chocs externes. Cette idée peut faire l'objet d'un prolongement pertinent de notre analyse.

Références

- ALMAZAN, A., A. DE MOTTA et S. TITMAN (2007) : « Firm Location and the Creation and Utilization of Human Capital », *The Review of Economic Studies*, 74(4), 1305–1327.
- BRIANT, A., P.-P. COMBES et M. LAFOURCADE (2007) : « Do the Size and Shape of Spatial Units Jeopardize Economic Geography Estimations? », mimeo.
- COMBES, P.-P. et M. LAFOURCADE (2005) : « Transport costs: measures, determinants, and regional policy implications for France », *Journal of Economic Geography*, 5(3), 319–349.
- COMBES, P.-P., T. MAYER et J.-F. THISSE (2006) : *Economie Géographique, L'intégration des régions et des nations*. Economica.
- DEVEREUX, M. P., R. GRIFFITH et H. SIMPSON (2004) : « The Geographic Distribution of Production Activity in the UK », *Regional Science and Urban Economics*, 34(5), 533–564.
- DURANTON, G., P. MARTIN, T. MAYER et F. MAYNERIS (2008) : *Les pôles de compétitivité: Que peut-on en attendre ?* Centre pour le Recherche Économique et ses Applications - Éditions de l'École Normale Supérieure (CEPREMAP/ENS).
- DURANTON, G. et H. G. OVERMAN (2005) : « Testing for Localization Using Micro-Geographic Data », *Review of Economic Studies*, 72(4), 1077–1106.
- (2008) : « Exploring The Detailed Location Patterns of U.K. Manufacturing Industries Using Microgeographic Data », *Journal of Regional Science*, 48(1), 213–243.
- DURANTON, G. et D. PUGA (2001) : « Nursery Cities: Urban Diversity, Process Innovation and the Life-cycle of Product », *American Economic Review*, 91(5), 1454–1477.
- (2004) : « Microfoundations of Urban Agglomeration Economies », in *Handbook of Urban and Regional Economics*, vol. 4. Elsevier Science Publisher.
- (2005) : « From Sectoral to Functional Urban Specialisation », *Journal of Urban Economics*, 57(2), 343–370.
- ELLISON, G. et E. L. GLAESER (1997) : « Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach », *Journal of Political Economy*, 105(5), 889–927.
- GILLI, F. (2006) : « La région parisienne entre 1975 et 1999 : une mutation géographique et économique », *Economie et Statistiques*, 387, 3–32.
- GUIDO, L. et F. SCHIVARDI (2007) : « Spillovers in Industrial Districts », *Economic Journal*, 117, 68–93.
- HOLMES, T. J. et J. J. STEVENS (2002) : « Geographic Concentration and Establishment Scale », *The Review of Economics and Statistics*, 84(4), 682–690.
- HOUBEINE, M. (1999) : « Concentration géographique des activités et spécialisation des départements français », *Économie et statistiques*, 326–327.
- KOLKO, J. (2000) : « Can I Get Some Service Here? Information Technology, Service Industries and the Future of Cities », mimeo.
- (2007) : « Agglomeration and Co-Agglomeration of Services Industries », mimeo.
- LACOUR, C., A. DELAMARRE et M. THOIN (2008) : *40 ans d'aménagement du territoire*. DATAR-La Documentation Française.
- LAFOURCADE, M. et G. MION (2007) : « Concentration, Agglomeration and the Size of Plants », *Regional Science and Urban Economics*, 37(1), 46–68.

- MARCON, E. et F. PUECH (2007) : « Measures of the Geographic Concentration of Industries: Improving Distance-Based Methods », mimeo.
- MAUREL, F. et B. SÉDILLOT (1999) : « A Measure of Geographic Concentration in French Manufacturing Industries », *Regional Science and Urban Economics*, 29, 575–604.
- RATHELOT, R. et P. SILLARD (2007) : « Zones Franches Urbaines : Quels effets sur l'emploi salarié et les créations d'établissements? », Document de travail INSEE/DESE.
- SILVERMAN, B. (1986) : *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.

A Indices de concentration discrets

Cette annexe vise à présenter succinctement les différents indices de concentration utilisés dans la littérature (Ellison et Glaeser (1997), Maurel et Sédillot (1999), Devereux, Griffith et Simpson (2004)). Pour plus de détails, le lecteur pourra se reporter au chapitre 10 de Combes, Mayer et Thisse (2006).

Les notations suivantes seront utilisées pour tous les indices.

S	désigne le nombre de secteurs
R	désigne le nombre de régions
x_r^s	est l'emploi total du secteur s dans la région r
x^s	est l'emploi total du secteur s
x_r	est l'emploi total de la région r
x	est l'emploi total
$\lambda_r^s = \frac{x_r^s}{x^s}$	est la part du secteur s dans la région r
$\lambda_r = \frac{x_r}{x}$	est la part de la région r dans l'emploi total

Les indices absolus

Ces indices mesurent la concentration de l'emploi d'un secteur par rapport à une situation où l'emploi serait également réparti dans l'ensemble des R régions. Nous utiliserons ci-après deux indices absolus : l'indice d'Herfindahl et l'indice de Gini. L'indice d'Herfindahl est défini par :

$$H^s = \sum_{r=1}^R (\lambda_r^s)^2$$

Il vaut $\frac{1}{R}$ si l'emploi est uniformément réparti et 1 si l'emploi est concentré dans une seule région.

Pour calculer l'indice de Gini, il faut classer les régions de la plus petite (en termes d'emploi dans le secteur s) à la plus grande. On définit alors $\lambda_{r(n)}^s = \sum_{r=1}^n \lambda_r^s$ la somme cumulée des parts d'emploi du secteur s dans les n régions les plus petites régions. L'indice de Gini est égal au double de l'aire entre la courbe des $\lambda_{r(n)}^s$ (courbe de Lorentz) et la bissectrice (qui représente la répartition uniforme de l'emploi). Formellement, l'indice absolu de Gini pour un secteur s s'écrit :

$$G^s = 1 - \sum_{n=1}^R \frac{1}{R} \left[\lambda_{r(n)}^s + \lambda_{r(n+1)}^s \right]$$

Les indices relatifs

Une première amélioration de ces deux indices est de ne plus considérer que la situation de référence est la répartition uniforme de l'emploi mais que c'est la répartition de l'emploi total.

Autrement dit, on cherche à savoir si un secteur est plus ou moins concentré que la moyenne. Pour cela, on compare λ_r^s et λ_r . L'indice d'Herfindahl relatif se définit par :

$$H_{rel}^s = \frac{1}{R} \sum_{r=1}^R \lambda_r \left(\frac{\lambda_r^s}{\lambda_r} \right)^2$$

L'indice de Gini relatif se définit par :

$$G_{rel}^s = 1 - \sum_{n=1}^R \left[\lambda_{r(n)}^s + \lambda_{r(n+1)}^s \right] \left[\lambda_{r(n)} - \lambda_{r(n+1)} \right]$$

Les indices tenant compte de la concentration industrielle

Ellison et Glaeser (1997) remarquent que, d'après les indices ci-dessus, si un secteur est très concentré industriellement alors il sera également concentré spatialement, ce qui n'est pas forcément souhaitable. En effet, la concentration industrielle et la concentration spatiale (appelée agglomération dans la littérature) n'ont ni les mêmes causes ni les mêmes conséquences. Il convient donc que les indices d'agglomération puissent permettre de discerner ces deux formes de concentration. Ellison et Glaeser (1997) introduisent l'indice suivant d'agglomération, qui corrige de la concentration industrielle.

$$EG^s = \frac{G_{EG}^s - H^s}{1 - H^s}$$

où H^s est un indice d'Herfindahl absolu au niveau établissement et :

$$G_{EG}^s = \frac{\sum_{r=1}^R (\lambda_r^s - \lambda_r)^2}{1 - \sum_{r=1}^R \lambda_r^2}$$

Maurel et Sédillot (1999) proposent une variante de cet indice :

$$MS^s = \frac{G_{MS}^s - H^s}{1 - H^s}$$

où H^s est un indice d'Herfindahl absolu et :

$$G_{MS}^s = \frac{\sum_{r=1}^R ((\lambda_r^s)^2 - \lambda_r^2)}{1 - \sum_{r=1}^R \lambda_r^2}$$

Les deux derniers indices, contrairement aux précédents, sont dérivés de modèles probabilistes de choix de localisation des firmes. Ces modèles permettent de dériver les variances des indicateurs G_{EG}^s et G_{MS}^s .

B Critères de qualité d'un bon indicateur de concentration

Critère 1 : La mesure de la concentration spatiale doit être effectuée par rapport à une référence clairement établie.

Les premiers indices de concentration utilisés (Herfindahl ou Gini absolu) prennent la distribution uniforme comme référence. En d'autres termes, un secteur est peu concentré si ses établissements sont uniformément répartis sur le territoire. Cette référence apparaît rapidement peu pertinente. La répartition uniforme n'est pas plausible : certaines zones au relief accidenté ne peuvent accueillir d'établissements. Les indices relatifs (Gini relatif, par exemple) améliorent ce point puisqu'ils prennent comme référence la répartition de l'activité globale.

Critère 2 : La mesure de la concentration spatiale doit être comparable entre secteurs.

Ellison et Glaeser (1997) construisent leur indicateur en considérant explicitement ce problème. Contrairement à l'indice de Gini, leur indicateur corrige la concentration géographique des effets liés aux différences dans le nombre d'établissements entre secteurs et de la répartition de l'emploi entre ces établissements, c'est à dire de la *concentration industrielle*. Reprenons à titre d'illustration l'exemple développé par Ellison et Glaeser (1997) : aux États-Unis, dans le secteur de fabrication des aspirateurs, 75% de l'emploi appartient à quatre établissements seulement. Un indice de Gini relatif conclura à la concentration géographique de ce secteur, même si les établissements se localisent de manière aléatoire dans l'espace. Un tel indice échoue donc à rendre compte de l'existence de forces d'agglomération externes aux établissements.

Critère 3 : La mesure de la concentration spatiale doit être comparable entre zones géographiques et insensible à un changement de définition des unités géographiques.

Les indicateurs de Gini, Ellison-Glaeser ou Maurel-Sédillot reposent tous sur un découpage géographique arbitraire. Ils sont donc sensibles au choix de ce découpage. Tout d'abord, si la taille ou la forme des entités géographiques change, alors la valeur de l'indicateur change, bien que la concentration du secteur n'ait pas été modifiée. Ce problème est connu dans la littérature anglo-saxonne sous le terme *Modifiable Areal Unit Problem* ou *Problème des Unités Spatiales Modifiables* (Briant, Combes et Lafourcade (2007)).

Un bon indicateur doit également tenir compte de la position relative des zones géographiques les unes par rapport aux autres. En effet, un secteur réparti sur deux zones géographiques contigües peut avoir la même valeur de l'indicateur d'Ellison et Glaeser ou de Maurel et Sédillot qu'un secteur de même taille réparti identiquement entre deux zones aux deux extrémités de la France, pour peu que ces zones soient de même taille. Pourtant, la simple observation visuelle nous permettrait de conclure à la concentration du premier par rapport au second. En d'autres termes, un bon indicateur doit être sensible à la permutation des zones géographiques les unes par rapport aux autres. Ce critère n'est rempli ni par l'indice d'Ellison et Glaeser, ni par l'indice de Maurel et Sédillot.

Critère 4 : La mesure de la concentration spatiale doit permettre de déterminer si des différences significatives par rapport à la référence ou entre deux situations (zones, périodes ou secteurs) existent.

Ellison et Glaeser (1997) et Maurel et Sédillot (1999) dérivent la variance de leurs estimateurs respectifs. Les premiers montrent que 446 (des 459) secteurs qu'ils étudient présentent un indice de concentration supérieur à ce que prédirait une distribution aléatoire. Parmi ces 446, l'écart entre la valeur de l'indice et le contrefactuel (distribution aléatoire) est supérieur à deux fois son écart-type dans 369 secteurs³³. De même, Maurel et Sédillot (1999) dérivent la variance de leur estimateur et concluent dans 270 secteurs (sur 273) à une valeur de l'indicateur significativement différente de zéro³⁴. Parmi ces secteurs, l'indicateur est significativement positif dans 211 secteurs (soit 77%).

Critère 5 : la mesure de la concentration spatiale doit être insensible à un changement de définition des secteurs. À notre connaissance, aucun indicateur de concentration ne remplit cette condition.

³³Ceci ne constitue pas un test formel de significativité. En effet, Ellison et Glaeser (1997) sont capables de dériver les deux moments (moyenne et variance) de leur indicateur mais n'en connaissent pas la loi asymptotique. Ils ne peuvent donc pas formellement conclure sur le seuil de significativité.

³⁴Maurel et Sédillot (1999) dérivent la variance de leur estimateur mais en aucun cas la distribution asymptotique. Aussi lorsque les auteurs parlent d'intervalles de confiance à 95% (p582), il est difficile de connaître la distribution de référence considérée.

C Détails sur les données utilisées

La construction de l'indice de Duranton et Overman (2005) nécessite de mobiliser de nombreuses sources de données.

Le projet Clap (Connaissance Locale de l'Appareil Productif) constitue une synthèse de sources administratives dont l'objectif est de localiser l'ensemble de l'activité économique au travers des effectifs et des rémunérations. Clap rapproche des données du répertoire Sirene, des DADS (Déclarations administratives de données sociales), des Urssaf, des fichiers de salariés de la fonction publique et de Suse (Système Unifié de Statistiques d'Entreprises). Clap comprend l'ensemble des établissements employeurs ou non-employeurs, dès lors qu'ils ont eu au moins un jour d'activité dans l'année d'exercice. Les ménages employeurs domestiques et les non-salariés ne font pas partie du champ de Clap. Notre base initiale est constituée de 4 370 346 établissements en France continentale (hors Corse) actifs en 2005. Suivant Duranton et Overman (2005), nous ne retenons dans cette étude que les établissements actifs au 31 décembre 2005 et déclarant au moins un salarié. Nous supprimons: 1) les établissements ayant cessé leur activité au cours de l'année (la géolocalisation des établissements n'étant disponible qu'au 1er janvier de chaque année, ces derniers ne sont pas présents dans le répertoire SIREN géolocalisé au 1er janvier 2006); 2) les mono-établissements ne déclarant aucun salarié; 3) les établissements sans salarié d'entreprises pluri-établissements. Deux possibilités : soit l'établissement n'enregistre effectivement aucun salarié, soit l'entreprise regroupe ses déclarations de salariés dans un petit nombre de ces établissements effectivement actifs.

Ce projet permet de traiter les questions de localisation fine de l'emploi et de localiser l'emploi à la commune. En effet, les DADS ne permettent pas toujours de décrire avec précision la situation locale de l'activité productive. Il arrive souvent que les salariés travaillant dans un même établissement ne soient pas effectivement déclarés dans un seul et même établissement. L'objectif de Clap est de localiser l'emploi à la commune. Un certain nombre d'entreprises possèdent plusieurs établissements dans une même commune. Il arrive également que les déclarations administratives, sur lesquelles sont fondées les données de Clap, soient regroupées sur un seul établissement de la commune (cas des banques). La conséquence est que la donnée au niveau du croisement de la commune et de l'activité est satisfaisante mais, au niveau infra-communal, la localisation de l'emploi n'est plus aussi précise. De plus, l'établissement « regroupeur » peut atteindre et dépasser les seuils définissant les grands établissements.

Pour les établissements non géolocalisés (i.e. pour lesquels les coordonnées géographiques ne sont pas disponibles), nous avons eu recours à une technique de géocodage qui consiste à retrouver les coordonnées géographiques d'un établissement à partir d'une adresse postale. Cette association n'est évidemment pas parfaite : certains établissements ne peuvent être localisés qu'à un numéro proche de la rue ou bien à la rue. Dans ce dernier cas, l'erreur commise est donc de quelques centaines de mètres. Au final, pour les secteurs retenus, cette technique permet de porter à 89% la part des établissements de l'échantillon qui sont effectivement géolocalisés.

G 9001	J. FAYOLLE et M. FLEURBAEY Accumulation, profitabilité et endettement des entreprises		Macro-economic import functions with imperfect competition - An application to the E.C. Trade	G 9311	J. BOURDIEU - B. COLIN-SEDILLOT Les décisions de financement des entreprises françaises : une évaluation empirique des théories de la structure optimale du capital		analyse économique des politiques française et allemande
G 9002	H. ROUSSE Détection et effets de la multicolinéarité dans les modèles linéaires ordinaires - Un prolongement de la réflexion de BELSLEY, KUH et WELSCH	G 9203	I. STAPIC Les échanges internationaux de services de la France dans le cadre des négociations multilatérales du GATT Juin 1992 (1ère version) Novembre 1992 (version finale)	G 9312	L. BLOCH - B. CŒURÉ Q de Tobin marginal et transmission des chocs financiers	G 9412	J. BOURDIEU - B. CŒURÉ - B. COLIN-SEDILLOT Investissement, incertitude et irréversibilité Quelques développements récents de la théorie de l'investissement
G 9003	P. RALLE et J. TOUJAS-BERNATE Indexation des salaires : la rupture de 1983	G 9204	P. SEVESTRE L'économétrie sur données individuelles-temporelles. Une note introductive	G 9313	Equipes Amadeus (INSEE), Banque de France, Méric (DP) Présentation des propriétés des principaux modèles macroéconomiques du Service Public	G 9413	B. DORMONT - M. PAUCHET L'évaluation de l'élasticité emploi-salaire dépendelle des structures de qualification ?
G 9004	D. GUELLEC et P. RALLE Compétitivité, croissance et innovation de produit	G 9205	H. ERKEL-ROUSSE Le commerce extérieur et l'environnement international dans le modèle AMADEUS (réestimation 1992)	G 9314	B. CREPON - E. DUGUET Research & Development, competition and innovation	G 9414	I. KABLA Le Choix de breveter une invention
G 9005	P. RALLE et J. TOUJAS-BERNATE Les conséquences de la désindexation. Analyse dans une maquette prix-salaires	G 9206	N. GREENAN et D. GUELLEC Coordination within the firm and endogenous growth	G 9315	B. DORMONT Quelle est l'influence du coût du travail sur l'emploi ?	G 9501	J. BOURDIEU - B. CŒURÉ - B. SEDILLOT Irreversible Investment and Uncertainty : When is there a Value of Waiting ?
G 9101	Equipe AMADEUS Le modèle AMADEUS - Première partie - Présentation générale	G 9207	A. MAGNIER et J. TOUJAS-BERNATE Technology and trade : empirical evidences for the major five industrialized countries	G 9316	D. BLANCHET - C. BROUSSE Deux études sur l'âge de la retraite	G 9502	L. BLOCH - B. CŒURÉ Imperfections du marché du crédit, investissement des entreprises et cycle économique
G 9102	J.L. BRILLET Le modèle AMADEUS - Deuxième partie - Propriétés variantielles	G 9208	B. CREPON, E. DUGUET, D. ENCAOUA et P. MOHNEN Cooperative, non cooperative R & D and optimal patent life	G 9317	D. BLANCHET Répartition du travail dans une population hétérogène : deux notes	G 9503	D. GOUX - E. MAURIN Les transformations de la demande de travail par qualification en France Une étude sur la période 1970-1993
G 9103	D. GUELLEC et P. RALLE Endogenous growth and product innovation	G 9209	B. CREPON et E. DUGUET Research and development, competition and innovation : an application of pseudo maximum likelihood methods to Poisson models with heterogeneity	G 9318	D. EYSSARTIER - N. PONTY AMADEUS - an annual macro-economic model for the medium and long term	G 9504	N. GREENAN Technologie, changement organisationnel, qualifications et emploi : une étude empirique sur l'industrie manufacturière
G 9104	H. ROUSSE Le modèle AMADEUS - Troisième partie - Le commerce extérieur et l'environnement international	G 9301	J. TOUJAS-BERNATE Commerce international et concurrence imparfaite : développements récents et implications pour la politique commerciale	G 9319	G. CETTE - Ph. CUNÉO - D. EYSSARTIER - J. GAUTIÉ Les effets sur l'emploi d'un abaissement du coût du travail des jeunes	G 9505	D. GOUX - E. MAURIN Persistence des hiérarchies sectorielles de salaires: un réexamen sur données françaises
G 9105	H. ROUSSE Effets de demande et d'offre dans les résultats du commerce extérieur manufacturé de la France au cours des deux dernières décennies	G 9302	Ch. CASES Durées de chômage et comportements d'offre de travail : une revue de la littérature	G 9401	D. BLANCHET Les structures par âge importent-elles ?	G 9505	D. GOUX - E. MAURIN Bis Persistence of inter-industry wages differentials: a reexamination on matched worker-firm panel data
G 9106	B. CREPON Innovation, taille et concentration : causalités et dynamiques	G 9303	H. ERKEL-ROUSSE Union économique et monétaire : le débat économique	G 9402	J. GAUTIÉ Le chômage des jeunes en France : problème de formation ou phénomène de file d'attente ? Quelques éléments du débat	G 9506	S. JACOBZONE Les liens entre RMI et chômage, une mise en perspective <i>NON PARU - article sorti dans Economie et Prévision n° 122 (1996) - pages 95 à 113</i>
G 9107	B. AMABLE et D. GUELLEC Un panorama des théories de la croissance endogène	G 9304	N. GREENAN - D. GUELLEC / G. BROUSSAUDIER - L. MIOTTI Innovation organisationnelle, dynamisme technologique et performances des entreprises	G 9403	P. QUIRION Les déchets en France : éléments statistiques et économiques	G 9507	G. CETTE - S. MAHFOUZ Le partage primaire du revenu Constat descriptif sur longue période
G 9108	M. GLAUDE et M. MOUTARDIER Une évaluation du coût direct de l'enfant de 1979 à 1989	G 9305	P. JAILLARD Le traité de Maastricht : présentation juridique et historique	G 9404	D. LADIRAY - M. GRUN-REHOMME Lissage par moyennes mobiles - Le problème des extrémités de série	G 9601	Banque de France - CEPREMAP - Direction de la Prévision - Erasme - INSEE - OFCE Structures et propriétés de cinq modèles macro-économiques français
G 9109	P. RALLE et alii France - Allemagne : performances économiques comparées	G 9306	J.L. BRILLET Micro-DMS : présentation et propriétés	G 9405	V. MAILLARD Théorie et pratique de la correction des effets de jours ouvrables	G 9602	Rapport d'activité de la DESE de l'année 1995
G 9110	J.L. BRILLET Micro-DMS NON PARU	G 9307	J.L. BRILLET Micro-DMS - variantes : les tableaux	G 9406	F. ROSENWALD La décision d'investir	G 9603	J. BOURDIEU - A. DRAZNIKS L'octroi de crédit aux PME : une analyse à partir d'informations bancaires
G 9111	A. MAGNIER Effets accélérateur et multiplicateur en France depuis 1970 : quelques résultats empiriques	G 9308	S. JACOBZONE Les grands réseaux publics français dans une perspective européenne	G 9407	S. JACOBZONE Les apports de l'économie industrielle pour définir la stratégie économique de l'hôpital public	G 9604	A. TOPIOL-BENSAÏD Les implantations japonaises en France
G 9112	B. CREPON et G. DUREAU Investissement en recherche-développement : analyse de causalités dans un modèle d'accélérateur généralisé	G 9309	L. BLOCH - B. CŒURE Profitabilité de l'investissement productif et transmission des chocs financiers	G 9408	L. BLOCH, J. BOURDIEU, B. COLIN-SEDILLOT, G. LONGUEVILLE Du défaut de paiement au dépôt de bilan : les banquiers face aux PME en difficulté	G 9605	P. GENIER - S. JACOBZONE Comportements de prévention, consommation d'alcool et tabagie : peut-on parler d'une gestion globale du capital santé ? <i>Une modélisation microéconométrique empirique</i>
G 9113	J.L. BRILLET, H. ERKEL-ROUSSE, J. TOUJAS-BERNATE "France-Allemagne Couplées" - Deux économies vues par une maquette macro-économétrique	G 9310	J. BOURDIEU - B. COLIN-SEDILLOT Les théories sur la structure optimal du capital : quelques points de repère	G 9409	D. EYSSARTIER, P. MAIRE Impacts macro-économiques de mesures d'aide au logement - quelques éléments d'évaluation	G 9606	C. DOZ - F. LENGART Factor analysis and unobserved component models: an application to the study of French business surveys
G 9201	W.J. ADAMS, B. CREPON, D. ENCAOUA Choix technologiques et stratégies de dissuasion d'entrée	G 9311	J. BOURDIEU - B. COLIN-SEDILLOT Les théories sur la structure optimal du capital : quelques points de repère	G 9410	F. ROSENWALD Suivi conjoncturel de l'investissement	G 9607	N. GREENAN - D. GUELLEC La théorie coopérative de la firme
G 9202	J. OLIVEIRA-MARTINS, J. TOUJAS-BERNATE			G 9411	C. DEFEUILLEY - Ph. QUIRION Les déchets d'emballages ménagers : une		

G 9608	N. GREENAN - D. GUELLEC Technological innovation and employment reallocation
G 9609	Ph. COUR - F. RUPPRECHT L'intégration asymétrique au sein du continent américain : un essai de modélisation
G 9610	S. DUCHENE - G. FORGEOT - A. JACQUOT Analyse des évolutions récentes de la productivité apparente du travail
G 9611	X. BONNET - S. MAHFOUZ The influence of different specifications of wages-prices spirals on the measure of the NAIRU : the case of France
G 9612	PH. COUR - E. DUBOIS, S. MAHFOUZ, J. PISANI-FERRY The cost of fiscal retrenchment revisited: how strong is the evidence ?
G 9613	A. JACQUOT Les flexions des taux d'activité sont-elles seulement conjoncturelles ?
G 9614	ZHANG Yingxiang - SONG Xueqing Lexique macroéconomique Français-Chinois
G 9701	J.L. SCHNEIDER La taxe professionnelle : éléments de cadrage économique
G 9702	J.L. SCHNEIDER Transition et stabilité politique d'un système redistributif
G 9703	D. GOUX - E. MAURIN Train or Pay: Does it Reduce Inequalities to Encourage Firms to Train their Workers?
G 9704	P. GENIER Deux contributions sur dépendance et équité
G 9705	E. DUGUET - N. IUNG R & D Investment, Patent Life and Patent Value An Econometric Analysis at the Firm Level
G 9706	M. HOUEBINE - A. TOPIOL-BENSAÏD Les entreprises internationales en France : une analyse à partir de données individuelles
G 9707	M. HOUEBINE Polarisation des activités et spécialisation des départements en France
G 9708	E. DUGUET - N. GREENAN Le biais technologique : une analyse sur données individuelles
G 9709	J.L. BRILLET Analyzing a small French ECM Model
G 9710	J.L. BRILLET Formalizing the transition process : scenarios for capital accumulation
G 9711	G. FORGEOT - J. GAUTÉ Insertion professionnelle des jeunes et processus de déclassement
G 9712	E. DUBOIS High Real Interest Rates: the Consequence of a Saving Investment Disequilibrium or of an insufficient Credibility of Monetary Authorities?
G 9713	Bilan des activités de la Direction des Etudes et Synthèses Economiques - 1996

G 9714	F. LEQUILLER Does the French Consumer Price Index Overstate Inflation?
G 9715	X. BONNET Peut-on mettre en évidence les rigidités à la baisse des salaires nominaux ? Une étude sur quelques grands pays de l'OCDE
G 9716	N. IUNG - F. RUPPRECHT Productivité de la recherche et rendements d'échelle dans le secteur pharmaceutique français
G 9717	E. DUGUET - I. KABLA Appropriation strategy and the motivations to use the patent system in France - An econometric analysis at the firm level
G 9718	L.P. PELÉ - P. RALLE Âge de la retraite : les aspects incitatifs du régime général
G 9719	ZHANG Yingxiang - SONG Xueqing Lexique macroéconomique français-chinois, chinois-français
G 9720	M. HOUEBINE - J.L. SCHNEIDER Mesurer l'influence de la fiscalité sur la localisation des entreprises
G 9721	A. MOURougANE Crédibilité, indépendance et politique monétaire Une revue de la littérature
G 9722	P. AUGERAUD - L. BRIOT Les données comptables d'entreprises Le système intermédiaire d'entreprises Passage des données individuelles aux données sectorielles
G 9723	P. AUGERAUD - J.E. CHAPRON Using Business Accounts for Compiling National Accounts: the French Experience
G 9724	P. AUGERAUD Les comptes d'entreprise par activités - Le passage aux comptes - De la comptabilité d'entreprise à la comptabilité nationale - A <i>paraître</i>
G 9801	H. MICHAUDON - C. PRIGENT Présentation du modèle AMADEUS
G 9802	J. ACCARDO Une étude de comptabilité générationnelle pour la France en 1996
G 9803	X. BONNET - S. DUCHÊNE Apports et limites de la modélisation « Real Business Cycles »
G 9804	C. BARLET - C. DUGUET - D. ENCAOUA - J. PRADEL The Commercial Success of Innovations An econometric analysis at the firm level in French manufacturing
G 9805	P. CAHUC - Ch. GIANELLA - D. GOUX - A. ZILBERBERG Equalizing Wage Differences and Bargaining Power - Evidence from a Panel of French Firms
G 9806	J. ACCARDO - M. JLASSI La productivité globale des facteurs entre 1975 et 1996

G 9807	Bilan des activités de la Direction des Etudes et Synthèses Economiques - 1997
G 9808	A. MOURougANE Can a Conservative Governor Conduct an Accomodative Monetary Policy ?
G 9809	X. BONNET - E. DUBOIS - L. FAUVET Asymétrie des inflations relatives et menus costs : tests sur l'inflation française
G 9810	E. DUGUET - N. IUNG Sales and Advertising with Spillovers at the firm level: Estimation of a Dynamic Structural Model on Panel Data
G 9811	J.P. BERTHIER Congestion urbaine : un modèle de trafic de pointe à courbe débit-vitesse et demande élastique
G 9812	C. PRIGENT La part des salaires dans la valeur ajoutée : une approche macroéconomique
G 9813	A.Th. AERTS L'évolution de la part des salaires dans la valeur ajoutée en France reflète-t-elle les évolutions individuelles sur la période 1979-1994 ?
G 9814	B. SALANIÉ Guide pratique des séries non-stationnaires
G 9901	S. DUCHÊNE - A. JACQUOT Une croissance plus riche en emplois depuis le début de la décennie ? Une analyse en comparaison internationale
G 9902	Ch. COLIN Modélisation des carrières dans Destinie
G 9903	Ch. COLIN Evolution de la dispersion des salaires : un essai de prospective par microsimulation
G 9904	B. CREPON - N. IUNG Innovation, emploi et performances
G 9905	B. CREPON - Ch. GIANELLA Wages inequalities in France 1969-1992 An application of quantile regression techniques
G 9906	C. BONNET - R. MAHIEU Microsimulation techniques applied to inter-generational transfers - Pensions in a dynamic framework: the case of France
G 9907	F. ROSENWALD L'impact des contraintes financières dans la décision d'investissement
G 9908	Bilan des activités de la DESE - 1998
G 9909	J.P. ZOYEM Contrat d'insertion et sortie du RMI Evaluation des effets d'une politique sociale
G 9910	Ch. COLIN - FI. LEGROS - R. MAHIEU Bilans contributifs comparés des régimes de retraite du secteur privé et de la fonction publique
G 9911	G. LAROQUE - B. SALANIÉ Une décomposition du non-emploi en France
G 9912	B. SALANIÉ Une maquette analytique de long terme du marché du travail
G 9912	Ch. GIANELLA
Bis	Une estimation de l'élasticité de l'emploi peu qualifié à son coût
G 9913	Division « Redistribution et Politiques Sociales » Le modèle de microsimulation dynamique DESTINIE
G 9914	E. DUGUET Macro-commandes SAS pour l'économétrie des panels et des variables qualitatives
G 9915	R. DUHAUTOIS Evolution des flux d'emplois en France entre 1990 et 1996 : une étude empirique à partir du fichier des bénéficiaires réels normaux (BRN)
G 9916	J.Y. FOURNIER Extraction du cycle des affaires : la méthode de Baxter et King
G 9917	B. CRÉPON - R. DESPLATZ - J. MAIRESSE Estimating price cost margins, scale economies and workers' bargaining power at the firm level
G 9918	Ch. GIANELLA - Ph. LAGARDE Productivity of hours in the aggregate production function: an evaluation on a panel of French firms from the manufacturing sector
G 9919	S. AUDRIC - P. GIVORD - C. PROST Evolution de l'emploi et des coûts par qualification entre 1982 et 1996
G 2000/01	R. MAHIEU Les déterminants des dépenses de santé : une approche macroéconomique
G 2000/02	C. ALLARD-PRIGENT - H. GUILMEAU - A. QUINET The real exchange rate as the relative price of nontradables in terms of tradables: theoretical investigation and empirical study on French data
G 2000/03	J.-Y. FOURNIER L'approximation du filtre passe-bande proposée par Christiano et Fitzgerald
G 2000/04	Bilan des activités de la DESE - 1999
G 2000/05	B. CREPON - F. ROSENWALD Investissement et contraintes de financement : le poids du cycle Une estimation sur données françaises
G 2000/06	A. FLIPO Les comportements matrimoniaux de fait
G 2000/07	R. MAHIEU - B. SÉDILLOT Microsimulations of the retirement decision: a supply side approach
G 2000/08	C. AUDENIS - C. PROST Déficit conjoncturel : une prise en compte des conjonctures passées
G 2000/09	R. MAHIEU - B. SÉDILLOT Equivalent patrimonial de la rente et souscription de retraite complémentaire
G 2000/10	R. DUHAUTOIS Ralentissement de l'investissement : petites ou grandes entreprises ? industrie ou tertiaire ?
G 2000/11	G. LAROQUE - B. SALANIÉ Temps partiel féminin et incitations financières à l'emploi
G2000/12	Ch. GIANELLA Local unemployment and wages

G2000/13	B. CREPON - Th. HECKEL - Informatisation en France : une évaluation à partir de données individuelles - Computerization in France: an evaluation based on individual company data
G2001/01	F. LEQUILLER - La nouvelle économie et la mesure de la croissance du PIB - The new economy and the measurement of GDP growth
G2001/02	S. AUDRIC La reprise de la croissance de l'emploi profite-t-elle aussi aux non-diplômés ?
G2001/03	I. BRAUN-LEMAIRE Evolution et répartition du surplus de productivité
G2001/04	A. BEAUDU - Th. HECKEL Le canal du crédit fonctionne-t-il en Europe ? Une étude de l'hétérogénéité des comportements d'investissement à partir de données de bilan agrégées
G2001/05	C. AUDENIS - P. BISCOURP - N. FOURCADE - O. LOISEL Testing the augmented Solow growth model : An empirical reassessment using panel data
G2001/06	R. MAHIEU - B. SÉDILLOT Départ à la retraite, irréversibilité et incertitude
G2001/07	Bilan des activités de la DESE - 2000
G2001/08	J. Ph. GAUDEMET Les dispositifs d'acquisition à titre facultatif d'annuités viagères de retraite
G2001/09	B. CRÉPON - Ch. GIANELLA Fiscalité, coût d'usage du capital et demande de facteurs : une analyse sur données individuelles
G2001/10	B. CRÉPON - R. DESPLATZ Evaluation des effets des dispositifs d'allègements de charges sociales sur les bas salaires
G2001/11	J.-Y. FOURNIER Comparaison des salaires des secteurs public et privé
G2001/12	J.-P. BERTHIER - C. JAULENT R. CONVENEVOLE - S. PISANI Une méthodologie de comparaison entre consommations intermédiaires de source fiscale et de comptabilité nationale
G2001/13	P. BISCOURP - Ch. GIANELLA Substitution and complementarity between capital, skilled and less skilled workers: an analysis at the firm level in the French manufacturing industry
G2001/14	I. ROBERT-BOBEE Modelling demographic behaviours in the French microsimulation model Destinie: An analysis of future change in completed fertility
G2001/15	J.-P. ZOYEM Diagnostic sur la pauvreté et calendrier de revenus : le cas du "Panel européen des ménages"
G2001/16	J.-Y. FOURNIER - P. GIVORD La réduction des taux d'activité aux âges extrêmes, une spécificité française ?

G2001/17	C. AUDENIS - P. BISCOURP - N. RIEDINGER Existe-t-il une asymétrie dans la transmission du prix du brut aux prix des carburants ?
G2002/01	F. MAGNIEN - J.-L. TAVERNIER - D. THESMAR Les statistiques internationales de PIB par habitant en standard de pouvoir d'achat : une analyse des résultats
G2002/02	Bilan des activités de la DESE - 2001
G2002/03	B. SÉDILLOT - E. WALRAET La cessation d'activité au sein des couples : y a-t-il interdépendance des choix ?
G2002/04	G. BRILHAULT - Rétropolation des séries de FBCF et calcul du capital fixe en SEC-95 dans les comptes nationaux français - Retropolation of the investment series (GFCF) and estimation of fixed capital stocks on the ESA-95 basis for the French balance sheets
G2002/05	P. BISCOURP - B. CRÉPON - T. HECKEL - N. RIEDINGER How do firms respond to cheaper computers? Microeconomic evidence for France based on a production function approach
G2002/06	C. AUDENIS - J. DERUYON - N. FOURCADE L'impact des nouvelles technologies de l'information et de la communication sur l'économie française - un bouclage macro-économique
G2002/07	J. BARDAJI - B. SÉDILLOT - E. WALRAET Évaluation de trois réformes du Régime Général d'assurance vieillesse à l'aide du modèle de microsimulation DESTINIE
G2002/08	J.-P. BERTHIER Réflexions sur les différentes notions de volume dans les comptes nationaux : comptes aux prix d'une année fixe ou aux prix de l'année précédente, séries chaînées
G2002/09	F. HILD Les soldes d'opinion résumant-ils au mieux les réponses des entreprises aux enquêtes de conjoncture ?
G2002/10	I. ROBERT-BOBÉE Les comportements démographiques dans le modèle de microsimulation Destinie - Une comparaison des estimations issues des enquêtes Jeunes et Carrières 1997 et Histoire Familiale 1999
G2002/11	J.-P. ZOYEM La dynamique des bas revenus : une analyse des entrées-sorties de pauvreté
G2002/12	F. HILD Prévisions d'inflation pour la France
G2002/13	M. LECLAIR Réduction du temps de travail et tensions sur les facteurs de production
G2002/14	E. WALRAET - A. VINCENT - Analyse de la redistribution intragénérationnelle dans le système de retraite des salariés du privé - Une approche par microsimulation - Intragenerational distributional analysis in the french private sector pension scheme - A microsimulation approach

G2002/15	P. CHONE - D. LE BLANC - I. ROBERT-BOBEE Offre de travail féminine et garde des jeunes enfants
G2002/16	F. MAUREL - S. GREGOIR Les indices de compétitivité des pays : interprétation et limites
G2003/01	N. RIEDINGER - E. HAUVY Le coût de dépollution atmosphérique pour les entreprises françaises : Une estimation à partir de données individuelles
G2003/02	P. BISCOURP et F. KRAMARZ Création d'emplois, destruction d'emplois et internationalisation des entreprises industrielles françaises : une analyse sur la période 1986-1992
G2003/03	Bilan des activités de la DESE - 2002
G2003/04	P.-O. BEFFY - J. DERUYON - N. FOURCADE - S. GREGOIR - N. LAÏB - B. MONFORT Évolutions démographiques et croissance : une projection macro-économique à l'horizon 2020
G2003/05	P. AUBERT La situation des salariés de plus de cinquante ans dans le secteur privé
G2003/06	P. AUBERT - B. CRÉPON Age, salaire et productivité La productivité des salariés décline-t-elle en fin de carrière ?
G2003/07	H. BARON - P.O. BEFFY - N. FOURCADE - R. MAHIEU Le ralentissement de la productivité du travail au cours des années 1990
G2003/08	P.-O. BEFFY - B. MONFORT Patrimoine des ménages, dynamique d'allocation et comportement de consommation
G2003/09	P. BISCOURP - N. FOURCADE Peut-on mettre en évidence l'existence de rigidités à la baisse des salaires à partir de données individuelles ? Le cas de la France à la fin des années 90
G2003/10	M. LECLAIR - P. PETIT Présence syndicale dans les firmes : quel impact sur les inégalités salariales entre les hommes et les femmes ?
G2003/11	P.-O. BEFFY - X. BONNET - M. DARRACQ-PARIES - B. MONFORT MZE: a small macro-model for the euro area
G2004/01	P. AUBERT - M. LECLAIR La compétitivité exprimée dans les enquêtes trimestrielles sur la situation et les perspectives dans l'industrie
G2004/02	M. DUÉE - C. REBILLARD La dépendance des personnes âgées : une projection à long terme
G2004/03	S. RASPILLER - N. RIEDINGER Régulation environnementale et choix de localisation des groupes français
G2004/04	A. NABOULET - S. RASPILLER Les déterminants de la décision d'investir : une approche par les perceptions subjectives des firmes

G2004/05	N. RAGACHE La déclaration des enfants par les couples non mariés est-elle fiscalement optimale ?
G2004/06	M. DUÉE L'impact du chômage des parents sur le devenir scolaire des enfants
G2004/07	P. AUBERT - E. CAROLI - M. ROGER New Technologies, Workplace Organisation and the Age Structure of the Workforce: Firm-Level Evidence
G2004/08	E. DUGUET - C. LELARGE Les brevets accroissent-ils les incitations privées à innover ? Un examen microéconométrique
G2004/09	S. RASPILLER - P. SILLARD Affiliating versus Subcontracting: the Case of Multinationals
G2004/10	J. BOISSINOT - C. L'ANGEVIN - B. MONFORT Public Debt Sustainability: Some Results on the French Case
G2004/11	S. ANANIAN - P. AUBERT Travailleurs âgés, nouvelles technologies et changements organisationnels : un réexamen à partir de l'enquête « REPONSE »
G2004/12	X. BONNET - H. PONCET Structures de revenus et propensions différentes à consommer - Vers une équation de consommation des ménages plus robuste en prévision pour la France
G2004/13	C. PICART Évaluer la rentabilité des sociétés non financières
G2004/14	J. BARDAJI - B. SÉDILLOT - E. WALRAET Les retraites du secteur public : projections à l'horizon 2040 à l'aide du modèle de microsimulation DESTINIE
G2005/01	S. BUFFETEAU - P. GODEFROY Conditions de départ en retraite selon l'âge de fin d'études : analyse prospective pour les générations 1945 à 1974
G2005/02	C. AFSA - S. BUFFETEAU L'évolution de l'activité féminine en France : une approche par pseudo-panel
G2005/03	P. AUBERT - P. SILLARD Délocalisations et réductions d'effectifs dans l'industrie française
G2005/04	M. LECLAIR - S. ROUX Mesure et utilisation des emplois instables dans les entreprises
G2005/05	C. L'ANGEVIN - S. SERRAVALLE Performances à l'exportation de la France et de l'Allemagne - Une analyse par secteur et destination géographique
G2005/06	Bilan des activités de la Direction des Études et Synthèses Économiques - 2004
G2005/07	S. RASPILLER La concurrence fiscale : principaux enseignements de l'analyse économique
G2005/08	C. L'ANGEVIN - N. LAÏB Éducation et croissance en France et dans un panel de 21 pays de l'OCDE
G2005/09	N. FERRARI Prévoir l'investissement des entreprises

	Un indicateur des révisions dans l'enquête de conjoncture sur les investissements dans l'industrie.
G2005/10	P.-O. BEFFY - C. L'ANGEVIN Chômage et boucle prix-salaires : apport d'un modèle « qualifiés/peu qualifiés »
G2005/11	B. HEITZ A two-states Markov-switching model of inflation in France and the USA: credible target VS inflation spiral
G2005/12	O. BIAU - H. ERKEL-ROUSSE - N. FERRARI Réponses individuelles aux enquêtes de conjoncture et prévision macroéconomiques : Exemple de la prévision de la production manufacturière
G2005/13	P. AUBERT - D. BLANCHET - D. BLAU The labour market after age 50: some elements of a Franco-American comparison
G2005/14	D. BLANCHET - T. DEBRAND - P. DOURGNON - P. POLLET L'enquête SHARE : présentation et premiers résultats de l'édition française
G2005/15	M. DUÉE La modélisation des comportements démographiques dans le modèle de microsimulation DESTINIE
G2005/16	H. RAOUI - S. ROUX Étude de simulation sur la participation versée aux salariés par les entreprises
G2006/01	C. BONNET - S. BUFFETEAU - P. GODEFROY Disparités de retraite de droit direct entre hommes et femmes : quelles évolutions ?
G2006/02	C. PICART Les gazelles en France
G2006/03	P. AUBERT - B. CRÉPON - P. ZAMORA Le rendement apparent de la formation continue dans les entreprises : effets sur la productivité et les salaires
G2006/04	J.-F. OUVRARD - R. RATHELOT Demographic change and unemployment: what do macroeconomic models predict?
G2006/05	D. BLANCHET - J.-F. OUVRARD Indicateurs d'engagements implicites des systèmes de retraite : chiffrages, propriétés analytiques et réactions à des chocs démographiques types
G2006/06	G. BIAU - O. BIAU - L. ROUVIERE Nonparametric Forecasting of the Manufacturing Output Growth with Firm-level Survey Data
G2006/07	C. AFSA - P. GIVORD Le rôle des conditions de travail dans les absences pour maladie
G2006/08	P. SILLARD - C. L'ANGEVIN - S. SERRAVALLE Performances comparées à l'exportation de la France et de ses principaux partenaires Une analyse structurelle sur 12 ans
G2006/09	X. BOUTIN - S. QUANTIN Une méthodologie d'évaluation comptable du coût du capital des entreprises françaises : 1984-2002

G2006/10	C. AFSA L'estimation d'un coût implicite de la pénibilité du travail chez les travailleurs âgés
G2006/11	C. LELARGE Les entreprises (industrielles) françaises sont-elles à la frontière technologique ?
G2006/12	O. BIAU - N. FERRARI Théorie de l'opinion Faut-il pondérer les réponses individuelles ?
G2006/13	A. KOUBI - S. ROUX Une réinterprétation de la relation entre productivité et inégalités salariales dans les entreprises
G2006/14	R. RATHELOT - P. SILLARD The impact of local taxes on plants location decision
G2006/15	L. GONZALEZ - C. PICART Diversification, recentrage et poids des activités de support dans les groupes (1993-2000)
G2007/01	D. SRAER Allègements de cotisations patronales et dynamique salariale
G2007/02	V. ALBOUY - L. LEQUIN Les rendements non monétaires de l'éducation : le cas de la santé
G2007/03	D. BLANCHET - T. DEBRAND Aspiration à la retraite, santé et satisfaction au travail : une comparaison européenne
G2007/04	M. BARLET - L. CRUSSON Quel impact des variations du prix du pétrole sur la croissance française ?
G2007/05	C. PICART Flux d'emploi et de main-d'œuvre en France : un réexamen
G2007/06	V. ALBOUY - C. TAVAN Massification et démocratisation de l'enseignement supérieur en France
G2007/07	T. LE BARBANÇON The Changing response to oil price shocks in France : a DSGE type approach
G2007/08	T. CHANEY - D. SRAER - D. THESMAR Collateral Value and Corporate Investment Evidence from the French Real Estate Market
G2007/09	J. BOISSINOT Consumption over the Life Cycle: Facts for France
G2007/10	C. AFSA Interpréter les variables de satisfaction : l'exemple de la durée du travail
G2007/11	R. RATHELOT - P. SILLARD Zones Franches Urbaines : quels effets sur l'emploi salarié et les créations d'établissements ?
G2007/12	V. ALBOUY - B. CRÉPON Aléa moral en santé : une évaluation dans le cadre du modèle causal de Rubin
G2008/01	C. PICART Les PME françaises : rentables mais peu dynamiques

G2008/02	P. BISCOURP - X. BOUTIN - T. VERGÉ The Effects of Retail Regulations on Prices Evidence from the Loi Galland
G2008/03	Y. BARBESOL - A. BRIANT Économies d'agglomération et productivité des entreprises : estimation sur données individuelles françaises
G2008/04	D. BLANCHET - F. LE GALLO Les projections démographiques : principaux mécanismes et retour sur l'expérience française
G2008/05	D. BLANCHET - F. TOUTLEMONDE Évolutions démographiques et déformation du cycle de vie active : quelles relations ?
G2008/06	M. BARLET - D. BLANCHET - L. CRUSSON Internationalisation et flux d'emplois : que dit une approche comptable ?
G2008/07	C. LELARGE - D. SRAER - D. THESMAR Entrepreneurship and Credit Constraints - Evidence from a French Loan Guarantee Program
G2008/08	X. BOUTIN - L. JANIN Are Prices Really Affected by Mergers?
G2008/09	M. BARLET - A. BRIANT - L. CRUSSON Concentration géographique dans l'industrie manufacturière et dans les services en France : une approche par un indicateur en continu