

LES EXTENSIONS RÉGIONALES DE L'ENQUÊTE INNOVATION : COMMENT JUGER DE LEUR PERTINENCE STATISTIQUE ?

Corinne AUTANT-BERNARD ()*

Michelle MONGO ()*

*Benoît BUISSON (**)*

*Lionel DOISNEAU (***)*

() Université de Saint-Étienne, GATE-LSE*

*(**) Insee, Pôle Ingénierie Statistique Entreprises*

*(***) Insee, PSAR Études Économiques Régionales*

Introduction

Dans le domaine des entreprises, de multiples enquêtes thématiques - souvent définies au niveau européen - fournissent potentiellement d'importants éléments de diagnostic et d'évaluation au niveau national sur des sujets variés. **Les besoins d'informations économiques** aux niveaux régional et local se renforcent, dans un contexte de mise en place de nombreuses politiques publiques. Il est rarement facile de répondre à ces besoins de manière simple et pertinente. Les questions tout comme le champ d'observation - avec les notions d'entreprises, d'unités légales et d'établissements - sont complexes et le système d'information n'est pas construit pour fournir des réponses à des niveaux géographiques infranationaux. Face à cette situation, deux « solutions » peuvent apparaître en terme d'enquêtes pour la statistique publique. D'une part les enquêtes régionales propres qui couvrent de manière pertinente la problématique des acteurs locaux, sans toutefois bénéficier de possibilités de comparaisons spatiales. Nous pouvons citer ici l'enquête sur la filière automobile menée dans la région Nord-Pas-de-Calais ou encore l'enquête sous-traitance aéronautique réalisée dans les régions Aquitaine et Midi-Pyrénées. D'autre part **les extensions régionales d'enquêtes nationales** : le questionnement est exactement le même qu'au niveau national mais l'échantillon est renforcé dans la région considérée pour obtenir des résultats robustes à ce niveau géographique. Ces extensions ont l'avantage de pouvoir permettre des comparaisons géographiques (région / France notamment) mais le questionnement n'est de fait pas toujours adapté à des problématiques locales particulières. Comme exemple nous pouvons citer ici les extensions aux enquêtes SINE (Système d'Informations sur les Nouvelles Entreprises) ou aux enquêtes innovation, objet de cet article. Ces deux « solutions » nécessitent des moyens, dans un contexte délicat de ce point de vue, et bien sûr sollicitent directement les entreprises. La tentation est donc grande de voir s'il est toujours justifié d'y recourir. L'objectif de cet article est d'apporter des éléments de réponse à cette question en partant d'une collaboration Insee - Université de Saint-Étienne sur la recherche des déterminants de l'innovation. L'objectif de cette collaboration est notamment de voir s'il existe des effets « régionaux » dans les déterminants de l'innovation. Cet article analyse en quoi les résultats de cette étude de nature économétrique, associés aux calculs d'estimateurs petits domaines, peuvent apporter des éléments de réponse à **la justification statistique (ex post) du recours aux extensions régionales** dans le domaine de l'innovation.

1. Présentation de l'enquête innovation et des extensions régionales

1.1. L'enquête innovation CIS 2008

L'enquête communautaire sur l'innovation est menée dans l'ensemble des pays de l'Union européenne et prend appui sur des définitions harmonisées au niveau international, manuel d'Oslo de l'OCDE notamment. Elle permet de décrire le processus d'innovation, d'en mesurer le poids économique, d'évaluer ses effets et d'apprécier ses mécanismes. Cette enquête, **CIS 2008**¹, a été réalisée entre octobre 2009 et février 2010. Elle porte sur la période 2006-2008 et couvre le champ des sociétés actives, marchandes et exploitantes de 10 salariés ou plus implantées en France (métropole et DOM). Les unités interrogées dans cette enquête sont **les unités légales**, et correspondent donc à la définition « juridique » de l'entreprise. Lorsqu'elles appartiennent à un groupe, elles ont répondu pour elles-mêmes et non pour l'ensemble du groupe. Les secteurs interrogés correspondent aux divisions 05 à 81 (hors division 75) de la nomenclature NAF rév. 2. Le questionnaire de l'enquête CIS 2008 a été envoyé par voie postale ou électronique à un échantillon d'environ **25 000** unités légales. L'échantillon national est exhaustif pour les unités de 250 salariés ou plus, et par sondage en deçà. A noter qu'il intègre toutes les unités de 10 salariés ou plus du secteur des industries agroalimentaires (divisions 10 à 12, nommées par la suite IAA) pour trois régions : Aquitaine, Bourgogne et Midi Pyrénées. La couverture est également exhaustive pour les entreprises de 50 salariés ou plus d'une partie du transport (groupes 49.1 et 49.2, divisions 50 et 51), sans spécificité régionale.

1.2. Pourquoi des extensions régionales ?

Les besoins d'informations chiffrées sur l'innovation en région sont très importants. Lors de la mise en place des stratégies régionales d'innovation (SRI) en 2009, il est apparu aux décideurs publics locaux (conseils régionaux, préfetures) qu'ils disposaient de peu d'indicateurs chiffrés spécifiques sur le domaine de l'innovation. Ce manque d'indicateurs était également souligné dans des rapports nationaux sur le sujet comme le rapport Prager. La demande d'informations était donc importante dans le contexte de mise en place de politiques publiques à un niveau régional sur le sujet.

Une précédente étude d'un groupe de travail Insee sur l'enquête CIS 2006² montrait clairement que **les résultats nationaux de cette enquête ne pouvaient pas directement être déclinés au niveau régional.** En effet la non-représentativité de l'échantillon au niveau régional entraînait une grande imprécision des résultats à ce niveau. Pour la plupart des régions, le taux d'innovation régional tiré de l'enquête nationale souffrait d'une imprécision de plus ou moins 10 points. Seules les grandes régions (Île-de-France, Rhône-Alpes) enregistraient une précision acceptable.

¹ Pour en savoir plus sur [l'enquête CIS 2008](#), il est possible de consulter le site internet de l'INSEE.

² Pour les lecteurs internes INSEE, cette [étude](#) est accessible sur le site intranet du pôle ingénierie statistique entreprises.

En lien avec cette demande d'informations et l'impossibilité d'exploiter directement de manière fiable l'enquête nationale, il a été décidé de réaliser des extensions régionales. Le questionnaire est identique à celui de l'enquête nationale mais des unités sont sélectionnées en plus pour atteindre une représentativité régionale.

1.3. Le contour des extensions régionales et les fichiers d'études

Sept régions ont donc décidé de réaliser des extensions régionales. Les extensions régionales couvrent **les unités légales de 10 à 249 salariés, mono-régionales ou quasi-mono-régionales**, non enquêtées dans l'échantillon national. Les unités légales multirégionales ne sont pas dans le champ des extensions, du fait que nous ne savons pas localiser dans ce cas les comportements d'innovation par région. L'ensemble des extensions régionales regroupe **5 000 unités**. Selon les régions, elles couvrent des secteurs d'activité différents.

Le tableau 1 récapitule les secteurs d'activités (en grisé) et les régions pour lesquelles il est possible de disposer d'informations au niveau régional, du fait des extensions régionales.

Tableau 1 : Couverture des extensions régionales

	<i>Industrie</i>	<i>Commerce de gros</i>	<i>Transport</i>	<i>Services technologiques</i>	<i>Services intellectuels</i>
Divisions	10 à 33	46	49, 52	58 à 63, 72	69-71, 73
<i>Nord-Pas-de-Calais</i>					
<i>Franche-Comté</i>					
<i>Haute Normandie</i>					
<i>Pays de la Loire*</i>					
<i>PACA</i>					
<i>Île-de-France</i>					
<i>Corse</i>					

Pour les Pays de la Loire, l'extension régionale dans le secteur des services intellectuels concerne uniquement les activités d'architecture et d'ingénierie (poste 71 de la nomenclature)

L'échantillon national est utilisé pour produire des résultats France entière. **Il n'intègre pas les entreprises enquêtées dans les extensions régionales.** Les extensions régionales ne portent que sur les entreprises de moins de 250 salariés mono-régionales ou quasi-mono-régionales. Les résultats régionaux sont produits avec les résultats des entreprises des extensions, ainsi qu'avec les entreprises de l'échantillon national, dans le champ des extensions, pour lesquelles c'est leur région de principale d'implantation.

Concrètement, nous disposons donc en fin d'enquête de deux types de fichiers :

- **Un fichier national**, composé de l'échantillon national, dans lequel chaque entreprise a un poids fixé ;
- **Sept fichiers régionaux**, chacun étant composé de l'échantillon « extension » de la région, ainsi que des entreprises de l'échantillon national avec pour région d'implantation la région considérée, sur le champ des extensions.

Afin de tirer profit de toute l'information disponible, **un fichier unique** de toutes les entreprises enquêtées par CIS 2008 a été élaboré. Il s'agit de la concaténation du fichier de l'enquête nationale, complété des entreprises des fichiers régionaux qui ne sont pas présentes dans le fichier national (les entreprises des extensions donc). Les variables étant identiques, cette agrégation ne pose pas de problème. L'identification de l'appartenance d'une unité au fichier national ou à celui des extensions est possible avec la variable « champ ». Une nouvelle variable a été construite : « poids_unique », qui donne le poids de chaque unité. Ce poids est différent de celui du fichier national ou des fichiers régionaux. C'est le poids qui aurait été utilisé si les extensions régionales avaient été prises en compte pour réaliser les estimateurs nationaux.

L'étude présentée étant axée sur l'existence d'un effet régional dans les déterminants de l'innovation, son champ est restreint aux **entreprises mono ou quasi monorégionales**. La prise en compte des entreprises plurirégionales, qui ont une activité significative dans plusieurs régions, sans que l'on sache dans quelle(s) région(s) elles ont innové aurait été une source de biais dans l'analyse de l'effet régional. Seules les entreprises mono-régionales et quasi-mono-régionales ont donc été sélectionnées. Cette information provient d'un algorithme à partir du répertoire Clap (connaissance locale de l'appareil productif). Dans un second temps, le champ a été restreint **aux entreprises de moins de 250 salariés**. Cette restriction permet de caler complètement l'étude sur le champ des extensions régionales. Le tableau 2 donne le nombre d'entreprises répondantes suivant les différents champs.

Tableau 2 : *Éléments de volumétrie de l'enquête nationale et des extensions régionales*

	Nombre d'entreprises	Entreprises régionales	Entreprises régionales <250 salariés
Fichier national	20 114	19 830	16 255
Extensions	4 146	4 146	4 146
Nord-Pas-de-Calais	774	774	774
Franche Comté	791	791	791
Haute-Normandie	624	624	624
Pays de la Loire	668	668	668
PACA	676	676	676
Île-de-France	421	421	421
Corse	192	192	192
Total	24 260	23 976	20 401

Champ : CIS 2008, entreprises répondantes

2. La modélisation du taux d'entreprises innovantes : une justification des extensions régionales ?

2.1. Brève revue de la littérature et choix des variables explicatives

Les informations disponibles dans l'enquête innovation ont permis **le développement des analyses empiriques des déterminants de l'innovation** (Mairesse et Mohnen 2010). Cette littérature porte pour l'essentiel sur l'industrie et ne fournit donc pas d'indications concernant les déterminants dans les secteurs des services. En outre, les approches antérieures ne prennent pas en **compte la dimension spatiale de l'innovation**. Or, les travaux réalisés en économie spatiale montrent qu'il existe des déterminants régionaux de l'innovation. Jusqu'à présent, la prise en compte de la dimension spatiale dans les enquêtes innovation n'était pas possible en raison de la non-représentativité de l'enquête à l'échelle régionale. La réalisation d'extensions régionales à l'enquête CIS 2008 dans sept régions change la donne dans ce domaine.

Prenant pour point de départ les nombreuses analyses théoriques, parfois contradictoires, sur les déterminants de l'innovation, un ensemble de travaux empiriques s'est développé depuis le début des années 1990. Ils étudient de manière large ou au contraire très ciblée l'impact sur l'innovation ou l'investissement en R&D de facteurs spécifiques à l'entreprise ou à son domaine d'activité, tels que la taille des entreprises, la structure de marché ou encore le rôle d'inputs spécifiques, en particulier l'accès au financement. Ces travaux ont notamment pris appui sur le développement des enquêtes innovation, qui fournissent, comme le détaillent Mairesse et Mohnen (2008), trois principaux types d'information : des informations sur les inputs d'innovation, des informations sur la production d'innovation et des indications sur les modalités de l'innovation.

Un ensemble conséquent de travaux s'est développé depuis les années 1990, exploitant les enquêtes innovation pour rendre compte des facteurs déterminant la propension des entreprises à innover. Il ne s'agit pas ici de détailler chacun de ces travaux, fort nombreux (Hollenstein, H. (1996), Blundell, R., Griffith, R. and van Reenen, J. (1999), Mairesse, J., Mohnen, P. (2002 and 2005) Raymond, W., Mohnen, P., Palm, F. and Schim van der Loeff, S. (2007), etc.) et détaillés dans un récent article par Mairesse et Mohnen (2010). L'objectif est ici de donner une vue d'ensemble des principaux résultats. Le plus récurrent est **l'effet de la taille et du secteur d'activité** sur la probabilité d'innover. A noter cependant que l'effet taille ne se retrouve pas lorsque l'on s'intéresse non pas au fait d'innover ou non mais à l'intensité de l'innovation. Au-delà de ces facteurs traditionnels de l'innovation, quelques études mettent l'accent sur **l'impact des coopérations** (Belderbos, R., Carree, M. and Lokshin, B. (2006), Leiponen, A. (2005)) ou des **aides financières publiques à l'innovation** (Arundel, A., Bordoy., C., Mohnen, P. and Smith, K. (2008), Mohnen, P. and Röller, L.-H. (2005)). Dans cette perspective, très peu de travaux s'intéressent aux secteurs des **services**. Une littérature récente reconnaît le rôle incontestable des activités de services dans la dynamique d'innovation des territoires. Ainsi, de façon similaire à l'industrie, l'effet sectoriel influence la probabilité d'innover au sein des services. D'autres facteurs alimentent la dynamique d'innovation au sein des services. Parmi lesquels,

on retrouve le rôle positif sur l'innovation : d'une **main-d'œuvre qualifiée** (Evangelista et Savona, 2003), du recours aux sources d'information interne et externe à l'entreprise (Pires, Sarkar and Carvalho, 2008) et de la **proximité géographique** entre les clients et les prestataires de services (Muller, Zenker, 2001). Enfin, au regard de la littérature sur l'innovation, il y a une reconnaissance de **l'étendue de marché** comme facteur déterminant de la capacité d'innovation des entreprises. Autrement dit, il existe une relation positive entre concurrence et innovation (Encaoua et Ulph, 2000). Ce déterminant très peu abordé dans le cas du secteur des services interroge sur son rôle d'incitation à l'innovation au sein de ce secteur.

La question de **l'impact régional** n'a jusqu'à présent pas été appréhendée à l'aide des enquêtes innovation (à l'exception du cas du Royaume-Uni). Les déterminants régionaux sont donc analysés à l'aide d'autres approches, reposant sur les données de **brevets** et de **dépenses en R&D**. Ces travaux qui appartiennent au domaine de la géographie de l'innovation se sont développés depuis la fin des années 90 (Jaffe (1989) Acs , Audretsch and Feldman (1991) Anselin Varga and Acs (1997) Audretsch and Feldman (1996) Bottazzi and Peri (2003) Autant-Bernard and LeSage (2010)). Les résultats attestent de l'existence de disparités spatiales persistantes dans les capacités à innover. Dans la majorité des cas, les zones urbaines ayant des structures industrielles diversifiées sont les plus innovantes. Ces travaux se concentrent toutefois sur les secteurs industriels. On ne dispose donc pas, hormis quelques études de cas, d'évaluation systématique des différences entre industrie et service concernant l'impact régional.

En conclusion, si l'enquête innovation a favorisé le développement des analyses empiriques des déterminants de l'innovation, **la littérature existante** porte pour l'essentiel sur l'industrie, et ne rend pas compte de **la dimension spatiale de l'innovation**. L'étude, issue d'une collaboration INSEE Université de Saint-Étienne, vise donc à tenter de combler ces déficits, en identifiant, grâce à la « régionalisation » de l'enquête CIS, les éventuelles singularités régionales dans ces processus d'innovation.

2.2. Un modèle national avec introduction de la région

En lien avec cette analyse de la littérature et la disponibilité des données, **les déterminants potentiels de l'innovation** retenus sont les suivants :

- le secteur d'appartenance de l'entreprise soit au niveau A38 de la nomenclature naf rev2, soit en utilisant la nomenclature de l'OCDE selon l'intensité technologique des secteurs ;
- l'appartenance ou non à un groupe ;
- la taille de l'entreprise mesurée à l'aide du nombre de salariés ;
- la composition de la main d'œuvre avec notamment la part d'ingénieurs et de techniciens dans l'entreprise ;
- le chiffre d'affaires ;
- le taux d'exportation ; le taux d'investissement ; la rentabilité économique.

Nous avons mené une analyse descriptive du lien entre le taux d'innovation au sens large et chaque variable explicative. Pour l'analyse, c'est bien la variable « **innovation au sens large** » qui sera privilégiée. Une unité légale sera déclarée innovante au sens large si elle est innovante dans au moins un des aspects qui suivent : innovation de produits, innovation de procédés, innovation de marketing, innovation d'organisation. Cette analyse nous a permis d'envisager deux aspects importants : le choix entre variables continues ou le passage en classes, les éventuels problèmes d'endogénéité. Nous ne présentons pas les résultats complets de cette analyse, ils seront décrits dans un document de travail INSEE. Nous mentionnerons uniquement les conclusions importantes pour la suite :

- il existe bien des **problèmes d'endogénéité** relativement aux variables économiques et financières. Pour cette raison, pour les variables chiffre d'affaires, taux d'exportation, taux d'investissement et taux de rentabilité, nous prendrons comme année de référence l'année 2005 (un an avant le début de la période d'observation de l'enquête).
- Pour la plupart des variables il apparaît plus pertinent de raisonner en **classes**. Seul le chiffre d'affaires (son logarithme de fait) sera conservé en variable quantitative.
- En ce qui concerne le **secteur d'activité**, nous avons privilégié le secteur en 38 postes pour l'industrie et en 88 postes pour les services, en effectuant des regroupements pour éviter les classes de trop petite taille.

Une analyse descriptive montre une forte variabilité au niveau régional des taux d'innovation. Toutefois cette variabilité est source d'interrogation pour les régions sans extension, pour lesquelles le taux d'innovation peut s'avérer être un indicateur peu robuste. Ce degré de fiabilité des résultats peut s'appréhender, comme pour toute enquête par sondage, par la notion **d'intervalle de confiance**. Ainsi en Haute-Normandie dans l'industrie (extension) le taux d'innovation est estimé à 46,3% et appartient à l'intervalle de confiance à 95% compris entre 44,8% et 47,7%. Si la Haute-Normandie n'avait pas réalisé d'extension dans l'industrie, le taux d'innovation aurait été estimé à 48,8% avec l'enquête nationale, l'intervalle de confiance à 95% étant compris entre 41,3% et 56,3%. Pour les « petites régions » sans extension l'amplitude de l'intervalle de confiance peut dépasser 15 points. Il importe également de déterminer si ces différences spatiales peuvent s'expliquer par **des disparités dans les caractéristiques structurelles des firmes locales**. Il est probable qu'au moins une partie des disparités spatiales n'est que la résultante d'une inégale distribution des entreprises ayant un fort potentiel innovant (en raison par exemple de dynamiques de spécialisation industrielles spécifiques, ou en raison d'un processus d'auto-sélection des firmes innovantes dans des zones aux caractéristiques particulières). Dans ce cas, il serait peut-être possible d'éviter la mise en place d'un dispositif d'extensions régionales, en exploitant par exemple les données régionales issues d'autres enquêtes, afin de prendre en compte les caractéristiques structurelles régionales à l'origine de la singularité des comportements locaux d'innovation. Mais il est aussi possible, c'est en tout cas l'idée défendue par les théoriciens des systèmes régionaux d'innovation, que **des processus d'innovation spécifiques se produisent dans certaines régions**, au-delà de ces effets structurels. La présence d'externalités de connaissance, l'accès à des infrastructures spécifiques, de recherche, de communication etc, sont autant de facteurs localisés susceptibles d'affecter la propension des firmes à

innover. Déterminer s'il existe de telles spécificités et en identifier les déterminants est essentiel pour comprendre quelles sont les particularités régionales, autrement dit les forces et les faiblesses de l'innovation des entreprises en région.

Sur cette base, un modèle logit est mis en œuvre dans lequel la probabilité d'innover (au sens large) est expliquée par les facteurs mentionnés précédemment. **Notre objectif est ici de voir si la région de principale implantation d'une unité légale apparaît comme un des déterminants de l'innovation.** L'analyse décrite ci-dessous est réalisée sur le fichier d'ensemble - enquête nationale et extensions régionales - afin d'avoir le maximum d'individus et de tirer profit de l'existence des extensions régionales. Nous raisonnons ici sur **le champ de l'industrie, des services technologiques et intellectuels.** C'est en effet dans ces secteurs que la majorité des extensions se concentre. Une fois enlevées les quelques unités légales pour lesquelles nous ne disposons pas de données 2005, l'analyse porte sur 11182 individus. Dans un premier temps nous n'avons pas introduit la région de principale implantation comme déterminant potentiel afin de sélectionner les déterminants « économiques » de l'innovation. Nous avons testé l'introduction de la région de principale implantation dans un second temps, pour juger de la significativité de son influence. Le modèle est estimé en utilisant une procédure stepwise, afin de sélectionner le modèle ayant le meilleur pouvoir explicatif. Les régressions sont effectuées sur **les données non pondérées.** Nous avons également effectué des régressions sur données pondérées, les résultats sont quasi-identiques.

La variable de localisation régionale des entreprises apparaît globalement significative lorsqu'elle est introduite dans les modèles. Même si elle améliore le modèle des déterminants de l'innovation, cette variable de localisation a un **pouvoir explicatif moins important** que les variables structurelles (tranche d'effectif salarié, comportement à l'exportation...). En termes d'interprétation, il faut probablement rester prudent à ce stade. Il serait un peu rapide d'affirmer que la localisation géographique des entreprises, mesurée au niveau régional, est un déterminant de l'innovation. Néanmoins nous pouvons dire que prendre en compte la localisation géographique des firmes améliore les modèles de recherche de déterminants de l'innovation. La localisation géographique permet donc de compléter l'analyse des déterminants, au-delà des variables structurelles déjà citées. Plusieurs régions ressortent à ce stade comme ayant des comportements particuliers en regard de l'innovation. Nous retrouvons ici deux régions avec extension (Pays de la Loire et Corse), une région sans extension mais de taille conséquente (Rhône-Alpes) et des petites régions sans extension pour lesquelles les résultats sont probablement plus fragiles (Basse-Normandie et Bretagne par exemple). Il n'est pas dans l'objectif de cet article d'analyser dans le détail les déterminants de l'innovation, d'expliquer les causalités mises en évidence. Le tableau 3 page suivante précise toutefois les déterminants retenus.

De ce fait, la recherche d'un premier modèle explicatif des comportements d'innovation montre la significativité de la variable de localisation régionale, la région de principale implantation. Ce premier résultat va dans le sens de la **justification des extensions régionales de l'enquête innovation.** En

plus de nombreux déterminants de nature économique, financière ou relatifs à la composition de la main d'œuvre, la localisation explique une partie des comportements d'innovation.

Tableau 3 : les déterminants de l'innovation au sens large

<i>Déterminant</i>	<i>Coef. estimé</i>	<i>Déterminant</i>	<i>Coef. estimé</i>
Constante	- 0,20 *	Guadeloupe	- 0,81
Effectif < 20 salariés	- 0,14 ***	Martinique	+ 0,22
Effectif 20-50 salariés	ref	Guyane	+ 0,40
Effectif > 50 salariés	+ 0,55 ***	La Réunion	+ 0,27
IAA	ref	Île-de-France	ref
Textile, habillement	- 0,10	Aquitaine	+ 0,35 ***
Bois, papier	- 0,10	Auvergne	+ 0,46 ***
Chimie, pharmacie	+ 0,75 ***	Basse-Normandie	+ 0,43 **
Caoutchouc, plastique	+ 0,27 ***	Bourgogne	+ 0,11
Métallurgie	- 0,02	Bretagne	+ 0,32 **
Informatique, électro.	+ 0,46 ***	Corse	+ 0,45 **
Equip électriques	+ 0,14	Centre	+ 0,04
Machines nca	+ 0,37 ***	Champagne-Ardenne	+ 0,15
Matériels de transports	+ 0,17	Franche-Comté	+ 0,13
Réparation, installation	- 0,18 **	Haute-Normandie	- 0,01
Edition	+ 0,53 ***	Languedoc-Roussillon	+ 0,41 **
Films, télé, vidéo	- 0,05	Limousin	+ 0,16
Programmation,diffusion	- 0,65 **	Lorraine	+ 0,14
Télécommunications	+ 0,04	Midi-Pyrénées	+ 0,16
Conseil et informatique	+ 0,48 ***	Nord-Pas-de-Calais	+ 0,00
Services d'information	+ 0,10	PACA	+ 0,00
Juridiques, comptables	- 0,14	Pays de la Loire	+ 0,21 **
Conseil de gestion	- 0,08	Picardie	+ 0,15
Ingénierie, architec.	+ 0,04	Poitou-Charentes	+ 0,03
Recherche et développ.	+ 0,89 ***	Rhône-Alpes	+ 0,23 ***
Publicité	+ 0,36 **		
Pas d'ingénieurs	- 0,31 ***		
Part ingénieur faible	ref		
Part ingénieur moyenne	+ 0,31 ***		
Part ingénieur forte	+ 0,65 ***		
Pas d'exports	- 0,21 ***		
Taux d'export faible	ref		
Taux d'export moyen	+ 0,14 **		
Taux d'esport élevé	+ 0,24 ***		
Taux d'invest très faible	- 0,10 *		
Taux d'invest faible	ref		
Taux d'invest moyen	+ 0,17 ***		
Taux d'invest élevé	+ 0,34 ***		

Lecture : niveau de significativité par rapport à la modalité de référence : * significatif au niveau 10% ; significatif au niveau 5% ; significatif au niveau 1 %.

2.3. Une autre voie : des modèles régionaux ?

Le second type d'analyse que permettent les extensions régionales est la réalisation des **estimations séparément pour chacune des régions couvertes par les extensions**. L'estimation de la propension des firmes à innover région par région, permet ainsi de mieux cerner l'origine des effets régionaux. Les tailles d'échantillon incitent à la prudence en regard des conclusions mises en évidence. Pour cette raison nous avons exclu la région Corse des traitements ultérieurs. Toutefois cette analyse peut fournir des renseignements complémentaires à celle menée précédemment. Le tableau 4 résume les déterminants, région par région, de la propension à innover au sens large.

Tableau 4 : les déterminants de l'innovation au sens large, région par région

<i>Région</i>	<i>Nombre d'obs.</i>	<i>Déterminants de l'innovation</i>
Île-de-France	1920	Part en ingénieurs et techniciens Taux d'exportation Secteur Taux d'investissement Taux de rentabilité économique Tranche d'effectifs
Haute-Normandie	784	Tranche d'effectifs Taux d'exportation
Nord-Pas-de-Calais	1029	Tranche d'effectifs Secteur Chiffre d'affaires
Franche-Comté	938	Tranche d'effectifs Taux d'exportation Taux d'investissement
Pays de la Loire	1095	Part en ingénieurs et techniciens Taux d'investissement Tranche d'effectifs
PACA	1007	Part en ingénieurs et techniciens Taux d'exportation Secteur Tranche d'effectifs

Ces résultats attestent donc de **profils régionaux différents en matière d'innovation**. Les déterminants ne sont pas les mêmes d'une région à l'autre. Par exemple la part en ingénieurs et techniciens joue un rôle prédominant dans trois régions (Île-de-France, Pays de la Loire et PACA) alors que cela n'est pas un déterminant significatif dans les trois autres (Haute-Normandie, Nord-Pas-de-Calais et Poitou-Charentes). Le taux d'exportation semble être un déterminant de premier plan dans quatre régions alors que son influence n'est pas mise en avant en Nord-Pas-de-Calais et en Pays de la Loire. Ces résultats sont à interpréter plus dans le détail région par région. Toutefois, dans le cadre de cette étude, elle nous apporte **un indice supplémentaire du**

bien-fondé des extensions régionales de l'enquête innovation, du fait de déterminants différenciés du comportement d'innovation selon les régions.

2.4. Estimateurs régionaux à partir de l'approche modèle

La recherche de modèles économétriques des déterminants de l'innovation peut permettre d'aller plus loin dans l'analyse du bien-fondé des extensions régionales, pour les régions concernées. Pour ces régions nous connaissons de manière fiable les taux d'innovation au sens large, pour l'industrie et/ou les services selon la couverture des extensions (**estimateur cible**). Nous pouvons voir dans quelle mesure il n'aurait pas été possible de fournir une estimation fiable de ces taux d'innovation à partir des données uniquement de l'enquête nationale. Pour cela nous recherchons un modèle des déterminants de l'innovation, proche de celui mis en évidence dans la partie 2.2, mais cette fois-ci uniquement sur **les unités légales de l'enquête nationale**, sur le champ de l'industrie, des services technologiques et intellectuels. De ce fait le modèle porte sur 7375 unités légales et non plus 11182. Nous recherchons un modèle ici sans introduire la région de principale implantation comme variable potentiellement explicative. Une fois ce modèle déterminé, nous pouvons en déduire une probabilité d'innover pour chaque unité légale des régions à extension, unité légale présente dans l'échantillon national ou dans les extensions régionales. En utilisant les poids à bon escient nous pouvons de fait calculer un estimateur des taux d'innovation, à partir du modèle national (estimateur modèle national dans le tableau). Nous pouvons également raisonner de manière légèrement différente. Pour les unités légales de l'enquête nationale, nous connaissons leur comportement d'innovation d'après l'enquête. Nous conservons cette information, nous nous servons de la probabilité d'innover uniquement pour les unités issues des extensions régionales (estimateur par prédiction). Le tableau 5 compare l'estimateur cible des taux d'innovation par secteur et région avec ceux obtenus par **l'approche « modèle national »**. Pour information nous faisons figurer également dans ce tableau l'estimateur issu des modèles régionaux mis en évidence dans la partie précédente.

Tableau 5 : comparaison des estimateurs cibles avec les estimateurs issus du modèle national et des modèles régionaux.

Région	Secteur	Est. cible	Est modèle national	Est. prédiction	Est modèle régional
Île-de-France	Services	57,9 %	63,3 %	61,7 %	58,6 %
Haute-Normandie	Industrie	46,3 %	50,6 %	50,6 %	46,6 %
Nord-Pas-de-Calais	Industrie	46,5 %	50,5 %	48,6 %	47,4 %
Nord-Pas-de-Calais	Services	41,6 %	47,9 %	46,0 %	40,1 %
Franche-Comté	Industrie	51,6 %	52,7 %	54,1 %	52,3 %
Franche-Comté	Services	45,3 %	43,2 %	44,1 %	40,8 %
Pays de la Loire	Industrie	53,4 %	51,5 %	50,9 %	52,7 %
Pays de la Loire	Services	57,9 %	57,5 %	58,5 %	59,4 %
PACA	Industrie	45,8 %	50,2 %	49,4 %	45,5 %
PACA	Services	61,0 %	62,5 %	63,6 %	61,0 %

L'écart moyen entre l'estimateur cible et l'estimateur issu du modèle national s'élève à 3,1 points. La « qualité » de l'estimateur issu du modèle national se révèle très variable d'une région et d'un secteur à l'autre. Cette adéquation apparaît très bonne pour les services dans les Pays de la Loire, bonne dans la région Franche-Comté en général, dégradé pour l'Île-de-France et le Nord-Pas-de-Calais. **Le modèle national n'est donc pas applicable dans toutes les situations mais la construction d'un estimateur basé sur ce modèle peut se révéler un proxy acceptable dans plusieurs cas.** L'estimateur par prédiction peut améliorer l'adéquation avec l'estimateur cible, notamment pour les régions pour lesquelles l'estimateur issu du modèle national était peu adapté, Nord-Pas-de-Calais par exemple. Nous reviendrons plus en détail dans la partie suivante sur la notion de « qualité du modèle national », en abordant la notion d'estimateurs sur des petits domaines. Pour information, la formalisation des estimateurs décrits dans les tableaux 5, 6 et 7 est explicitée dans un encadré à la fin de la partie 3.2.

3. Les estimateurs de type petits domaines

L'objectif de l'étude menée conjointement entre l'Insee et l'Université de Saint-Étienne est donc de mettre en évidence **les déterminants des processus d'innovation**, notamment en analysant les éventuels effets de la localisation géographique des firmes. Cette étude peut indirectement, comme cela est montré dans les paragraphes précédents, fournir des informations précieuses sur la pertinence d'avoir réalisé des extensions régionales à l'enquête innovation. Cette étude se poursuit actuellement pour aller plus loin dans l'explication des déterminants liés à la localisation géographique, en introduisant d'autres variables de nature régionale (les dépenses de recherche et développement, l'intensité des collaborations au sein d'une région par exemple). L'objectif initial de cette collaboration n'était donc pas de construire des **estimateurs du type petits domaines** les plus pertinents possibles. Toutefois la mise en évidence de modèles des déterminants de l'innovation apporte des informations précieuses dans le cadre de la construction de ce type d'estimateurs. En effet, la mise en évidence des déterminants les plus significatifs, dans le modèle national et pour les modèles régionaux, permet de définir avec plus de précision **les variables de calage** à utiliser en priorité pour élaborer les estimateurs petits domaines. Ces aspects sont développés dans cette troisième partie, en distinguant les estimateurs directs - élaborés à partir uniquement des unités légales de la région considérée - des estimateurs indirects qui prennent en compte les unités légales sur une zone qui contient la région considérée.

3.1. Les estimateurs directs

En se servant uniquement des unités légales enquêtées par l'enquête nationale, le moyen le plus « simple » pour estimer les taux d'innovation (au sens large) par région et grands secteurs consiste à prendre en compte seulement les unités légales de la région considérée, tout en utilisant leur poids issu de ce même fichier national. Il est bon de rappeler ici que la région de principale implantation géographique n'était pas prise en compte directement³ dans les strates de tirage pour l'échantillonnage de l'enquête CIS 2008. **L'estimateur direct des taux d'innovation** que nous obtenons souffre bien sûr d'une grande imprécision, imprécision d'autant plus forte que la taille de l'échantillon dans le domaine (la région de principale implantation) est réduite. Outre le fait que le nombre d'individus sur lesquels se base cet estimateur peut être réduit, nous pouvons avoir un problème lié à l'utilisation des **poids** issus du fichier national. Ces poids, issus d'un processus de correction de la non-réponse totale par repondération et d'un calage⁴ sur le fichier national, sont davantage pertinents au niveau national que régional. La mise en évidence de modèles, à un niveau national et pour les régions à extension, permet de mettre en évidence les déterminants les plus significatifs des comportements d'innovation. Il pourrait donc s'avérer pertinent de pratiquer **un calage** sur les strates croisées issues de ces déterminants, qui ne mettent pas toujours en évidence en premier lieu la tranche d'effectif et le secteur d'activité économique. Ainsi nous contrôlerons, avec les poids issus de ce calage, la « représentativité » des unités utilisées par rapport aux déterminants les plus significatifs de l'innovation. Cela peut s'avérer important, notamment pour les régions de taille réduite, pour lesquels la structure des unités enquêtées et répondantes peut se révéler différente de la structure de la base de sondage relativement aux principaux déterminants de l'innovation. Par contre si nous retenons tous les déterminants mis en évidence dans les modèles, nous obtiendrons des strates de calage avec des tailles beaucoup trop réduite. Il est donc nécessaire de faire un choix ici, un arbitrage entre taille des strates de calage et richesse des déterminants utilisés. Nous allons pratiquer d'une part un calage en utilisant les deux déterminants les plus significatifs du modèle national (part d'ingénieurs et de techniciens, tranche d'effectifs) et d'autre part des calages par région où nous avons calé sur les deux déterminants les plus significatifs explicités dans le tableau 4⁵. Les résultats sont présentés dans le tableau 6 qui compare l'estimateur cible, à l'estimateur direct, à l'estimateur direct calé sur les déterminants nationaux et enfin à l'estimateur direct calé sur les déterminants régionaux.

³ Les critères de stratification étaient la tranche de taille et le secteur économique d'appartenance des unités légales. Dans chaque strate les unités étaient triées par région de principale implantation. Le tirage était effectué de manière systématique dans chaque strate, ce qui permettait de contrôler le nombre d'unités tirées dans chaque strate par région de principale implantation.

⁴ Calage uniquement sur le nombre d'unités légales par strate de diffusion, croisement d'une tranche d'effectif avec un secteur économique donné

⁵ Sauf pour l'Île-de-France où nous avons privilégié le secteur au taux d'exportation

Tableau 6 : comparaison des estimateurs directs avec l'estimateur cible

Région	Secteur	Est. cible	Est direct	Est direct calé modèle national	Est direct calé modèle régional
Île-de-France	Services	57,9 %	52,2 %	59,7 %	58,1 %
Haute-Normandie	Industrie	46,3 %	47,0 %	49,1 %	44,5 %
Nord-Pas-de-Calais	Industrie	46,5 %	45,9 %	46,3 %	44,6 %
Nord-Pas-de-Calais	Services	41,6 %	42,2 %	45,7 %	47,7 %
Franche-Comté	Industrie	51,6 %	57,2 %	65,1 %	63,7 %
Franche-Comté	Services	45,3 %	43,7 %	45,5 %	45,5 %
Pays de la Loire	Industrie	53,4 %	52,7 %	53,3 %	51,5 %
Pays de la Loire	Services	57,9 %	54,1 %	65,7 %	63,4 %
PACA	Industrie	45,8 %	45,2 %	45,9 %	45,6 %
PACA	Services	61,0 %	53,9 %	69,4 %	71,2 %

L'écart moyen avec l'estimateur cible est plus faible pour l'estimateur direct (2,7 points) que pour l'estimateur direct calé sur le modèle national (3,9 points) et l'estimateur direct calé sur le modèle régional (4,0 points). Ce résultat peut surprendre, tout comme la bonne adéquation de l'estimateur direct dans le cas de régions de taille réduite, comme pour la Haute-Normandie dans l'industrie ou la Franche-Comté dans les services. Lorsque nous pratiquons un calage, l'adéquation peut donc se révéler moins bonne comme en Haute-Normandie ou pour les services en Nord-Pas-de-Calais. Cela peut probablement traduire une forte instabilité de l'estimateur direct, qu'il faudrait approfondir en calculant **les variances** de ces estimateurs. Dans ce cas il est probable que l'estimateur direct se base sur trop peu d'observations, ce qui le rend très instable. Cela est probablement le cas, par exemple, pour la Franche-Comté dans l'industrie. Par contre pour **les régions de taille intermédiaire**, associées à des secteurs qui comprennent un grand nombre d'unités légales, l'estimateur direct calé sur le modèle national peut se révéler approprié comme pour les services en Île-de-France ou pour l'industrie en Nord-Pas-de-Calais, Pays de la Loire ou PACA. Pour l'estimateur direct calé sur le modèle régional, il ne faut pas oublier que les extensions régionales ont été nécessaires pour déterminer le modèle régional. Toutefois, comme le verrons dans la partie 4, l'enquête CIS étant programmée tous les deux ans, un modèle régional déterminé sur les extensions une année donnée, pourrait être utilisé deux ans plus tard pour cette même région, qui ne ferait plus d'extension cette fois-ci.

3.2. Les estimateurs indirects

Pour parer à la forte imprécision des estimateurs directs sur les petits domaines, il est possible d'estimer des taux d'innovation par région en se basant non plus sur les unités légales de la région mais sur celles de l'ensemble du territoire national. Il sera nécessaire de supposer par contre ici que **les comportements sont identiques**, à un niveau de stratification donné, entre les unités nationales

et les unités régionales. Ceci est une hypothèse forte, notamment en regard des conclusions de la partie 2. Ce type d'estimateur peut toutefois se révéler judicieux pour **les « petites » régions pour lesquels le modèle national s'applique correctement**. Outre le fait de prendre les unités de l'enquête nationale, il va falloir savoir quel poids utiliser pour calculer ce type d'estimateur. Appliquer les poids de l'enquête nationale, reviendrait de fait à calculer des estimateurs nationaux et donc à supposer que les estimateurs régionaux sont identiques aux estimateurs nationaux. Nous pouvons évidemment proposer mieux en utilisant le résultat de la modélisation économétrique. Ainsi il est possible de **caler relativement à la structure régionale**, pour la région considérée, relativement aux variables les plus déterminantes du modèle explicatif. Dans le tableau page suivante, nous avons utilisé les deux déterminants les plus « significatifs » au niveau national (part d'ingénieurs et de techniciens, tranche d'effectifs) en distinguant l'industrie des services.

Comme le montre le tableau suivant, l'écart moyen entre l'estimateur cible et l'estimateur synthétique s'élève à 2,9 points. L'adéquation de l'estimateur synthétique de type post-stratifié se révèle particulièrement pertinente pour la région Franche-Comté. Nous sommes typiquement dans le cas d'une « petite » zone géographique pour laquelle l'estimateur direct peut se révéler fortement imprécis et pour laquelle la pertinence du modèle national s'avère correcte. Le cas de la région Nord-Pas-de-Calais peut également être riche d'enseignements. L'estimateur synthétique se montre peu pertinent, notamment dans les services. Nous sommes ici dans un cas où le modèle national ne s'appliquait pas bien. Le calage sur les déterminants nationaux apporte donc, ce qui somme toute paraît logique, peu d'intérêt. Lorsque nous sommes dans ce cas nous avons un double problème avec ce type d'estimateur synthétique. D'une part nous utilisons les unités légales nationales, dont le comportement en terme d'innovation demeure potentiellement différent des unités régionales. D'autre part nous calons sur les variables pertinentes du modèle national, qui ne sont pas forcément adaptées aux régions dont l'adéquation n'est pas optimale avec le modèle national.

Tableau 7 : comparaison des estimateurs cibles avec l'estimateur synthétique post-stratifié

Région	Secteur	Est. cible	Est synthétique post stratifié
Île-de-France	Services	57,9 %	57,5 %
Haute-Normandie	Industrie	46,3 %	50,5 %
Nord-Pas-de-Calais	Industrie	46,5 %	50,4 %
Nord-Pas-de-Calais	Services	41,6 %	48,6 %
Franche-Comté	Industrie	51,6 %	51,2 %
Franche-Comté	Services	45,3 %	45,6 %
Pays de la Loire	Industrie	53,4 %	51,8 %
Pays de la Loire	Services	57,9 %	55,8 %
PACA	Industrie	45,8 %	49,3 %
PACA	Services	61,0 %	55,8 %

Dans cette troisième partie, nous avons beaucoup utilisé la notion **d'adéquation avec le modèle national**. Pour cela nous nous sommes basés sur les résultats issus du tableau 5. Pour calculer l'estimateur issu du modèle national, nous avons sommé la probabilité d'innover de chaque unité légale, en appliquant le modèle national. Ce choix se révèle plus pertinent, notamment pour les petites régions avec peu d'unités dans l'échantillon national, que d'avoir un raisonnement dichotomique du type : lorsque la probabilité estimée est supérieure à 0,5 l'unité innove et elle n'innove pas dans le cas contraire. Un autre moyen pour estimer l'adéquation au modèle national est de calculer, par post strate issue du modèle national, **une distance entre le comportement des entreprises nationales et régionales**. Il s'agit de calculer, par post strate, un taux d'innovation issu de l'enquête nationale et de le comparer au taux d'innovation régional, calculé en tirant parti des extensions régionales. En pondérant par le poids des post-strate au niveau de chaque région, nous arrivons à une distance moyenne pondérée, qui peut être comparée d'une région à l'autre. Il apparaît que les résultats sont proches selon les régions. L'écart le plus faible est enregistré pour la Franche-Comté dans les services alors que nous enregistrons l'écart le plus fort en Nord-Pas-de-Calais dans les services. Nous pouvons directement relier cela aux résultats du tableau 3 relativement à l'écart entre estimateur cible et estimateur synthétique post-stratifié.

Encadré : formalisation des estimateurs décrits dans les tableaux 5, 6 et 7

Notations utilisées :

$Y_i=1$ si l'unité légale innove, 0 dans le cas contraire

S_R : échantillon régional (pour la région R)

S_N : échantillon national

S_{ER} : échantillon des extensions (pour la région R)

Nous avons donc : $S_R = S_{ER} + S_{N \cap R}$

p_i : poids de l'unité légale i utilisé pour la production d'estimateurs au niveau régional

q_i : poids de l'unité légale i utilisé pour la production d'estimateurs au niveau national

N_H^R : nombre d'unités légales dans la région R et dans la post strate H.

$$\text{Estimateur cible} = \sum_{i \in S_R} p_i Y_i$$

$$\text{Estimateur issu du modèle national (tableau 5)} = \sum_{i \in S_R} p_i Y_i^{estim / mod .national}$$

$$\text{Estimateur par prédiction (tableau 5)} = \sum_{i \in S_{N \cap R}} p_i Y_i + \sum_{i \in S_{ER}} p_i Y_i^{estim / mod .national}$$

$$\text{Estimateur issu du modèle régional (tableau 5)} = \sum_{i \in S_R} p_i Y_i^{estim / mod .régional}$$

$$\text{Estimateur direct (tableau 6)} = \sum_{i \in S_{N \cap R}} q_i Y_i$$

$$\text{Estimateur direct calé sur le modèle national (tableau 6)} = \sum_{i \in S_{N \cap R}} Y_i p_i^{calnat}$$

$$\text{Avec pour chaque post-strate } H_1 \text{ du modèle national : } \sum_{i \in S_{N \cap R}} p_i^{calnat} = N_{H_1}^R$$

$$\text{Estimateur direct calé sur le modèle régional (tableau 6)} = \sum_{i \in S_{N \cap R}} Y_i p_i^{calreg}$$

$$\text{Avec pour chaque post-strate } H_2 \text{ du modèle régional : } \sum_{i \in S_{N \cap R}} p_i^{calreg} = N_{H_2}^R$$

$$\text{Estimateur synthétique post strate (tableau 7)} = \sum_{i \in S_N} Y_i p_i^{post_strate}$$

$$\text{Avec pour chaque post strate } H_1 : \sum_{i \in S_N} p_i^{post_strate} = N_{H_1}^R$$

3.3. Un effet régional qui justifie les extensions ?

Suite à la partie 2, il apparaissait l'existence d'un déterminant régional dans les comportements d'innovation, ce qui peut justifier ex-post le recours à des extensions régionales dans le domaine de l'innovation. Dans cette partie 3, nous avons proposé la construction d'estimateurs petits domaines, que nous pouvions comparer à un estimateur cible robuste en lien avec l'existence d'extensions régionales. Même s'il serait nécessaire d'aller plus loin en estimant les variances de ces estimateurs petits domaines notamment, nous avons pu tirer **plusieurs enseignements**. Dans le cas où le modèle national s'applique correctement à la région, l'estimateur synthétique post strate se révèle performant. C'est notamment le cas pour la Franche-Comté. Dans le cas où l'application du modèle national demeure plus risquée, il s'avère judicieux dans certains cas d'utiliser l'estimateur direct calé sur les variables du modèle qui ressortent comme déterminants principaux de l'innovation. Ce résultat s'applique aux régions de taille moyenne ou grande pour des secteurs bien représentés : l'Île-de-France dans les services, les régions PACA, Pays de la Loire ou Nord-Pas-de-Calais dans l'industrie. Ce type d'estimateur donne également de bons résultats dans l'industrie en Haute-Normandie, c'est-à-dire pour une région plus petite que celles précédemment citées. Il reste toutefois des cas où une analyse supplémentaire s'avère nécessaire pour construire des estimateurs plus adaptés. Cela est notamment le cas dans les services pour les régions PACA et Nord-Pas-de-Calais. Tous ces résultats pourront être mis à profit pour limiter les extensions régionales, en supposant la stabilité temporelle des modèles, ce projet étant détaillé dans la partie 4.2.

4. Un projet : calcul d'estimateurs petits domaines pour limiter le recours aux extensions régionales

4.1. Présentation de l'enquête innovation CIS 2010 et des extensions régionales associées

L'enquête nationale innovation CIS 2008 et les extensions régionales associées se sont terminées en février 2010. Les publications régionales (cf. bibliographie) ont été publiées en novembre 2010. Ces publications ont été l'objet de reprises de presse et ont suscité des débats avec les partenaires locaux de l'Insee en région. Devant le succès de cette opération et dans l'objectif d'évaluer **les stratégies régionales d'innovation** dans de nombreuses régions, la demande s'est révélée forte pour organiser des extensions régionales à **l'enquête innovation CIS 2010**. Il a été décidé en février 2011 que dix régions, sept régions métropolitaines⁶ et trois départements d'outre-mer⁷ pourraient bénéficier d'extensions régionales. Cela représente une volumétrie de 8000 extensions régionales contre 5000 pour la précédente édition, pour un échantillon national de 21000 unités légales. La collecte de l'information se terminera fin janvier 2012, les taux de réponse apparaissant dès maintenant très satisfaisants. Les fichiers redressés seront disponibles fin mai 2012, pour des publications nationales

⁶ Île-de-France, Bretagne, Rhône-Alpes, Poitou-Charentes, Franche-Comté, Languedoc-Roussillon et Haute-Normandie

⁷ Guadeloupe, Martinique, La Réunion

et régionales à échéance de novembre 2012. Pour cette édition, il est fort probable que les estimateurs nationaux prennent en compte les éléments collectés dans le cadre des extensions régionales, ce qui améliorera leur précision. Comme déjà signalé et montré, l'intérêt des extensions régionales dans le domaine de l'innovation se comprend aisément. Toutefois il intervient dans une période délicate à double titre. D'une part la tension sur les moyens à l'Insee fait qu'il n'est pas toujours évident de dédier une équipe de collecte à la gestion de ces extensions régionales. D'autre part pour les unités légales de moins de 250 salariés, cela reste synonyme d'un accroissement de la charge liée aux enquêtes. Pour ces deux raisons, après avoir assuré les travaux post-collecte de l'enquête innovation CIS 2010, soit à partir de juillet 2012, le pôle ingénierie statistique entreprises souhaite mener une étude pour voir comment **le recours à des estimateurs petits domaines pourrait limiter les extensions régionales.**

4.2. Des extensions tous les 4 ans avec des estimateurs petits domaines comme point intermédiaire ?

L'idée générale de ce projet d'étude est de permettre, pour une région donnée et sur un champ donné, de réaliser des extensions régionales tous les 4 ans et d'obtenir des **estimateurs intermédiaires** sur les principaux agrégats de l'enquête. Pour les régions qui ont réalisé des extensions à l'enquête innovation CIS 2008, l'idée générale est donc de construire des estimateurs issus de l'enquête CIS 2010 de type petits domaines au moins en considérant les estimateurs qui suivent :

- estimateur direct calé sur les déterminants de l'innovation mis en évidence à partir des modèles régionaux de CIS 2008 ;
- estimateur indirect de type post-stratifié en calant sur les déterminants du modèle national qui lui peut-être recalculé de manière robuste sur CIS 2010.

En lien avec les conclusions mises en évidence dans la partie 3, nous pourrions appliquer ces estimateurs de manière différenciée selon les régions et secteurs concernés. Un des points importants sera **la stabilité des modèles régionaux** entre CIS 2008 et CIS 2010. Nous supposons de fait cette stabilité mais nous aurons néanmoins des éléments pour la tester. Les régions Haute-Normandie et Franche-Comté⁸ ont réalisé des extensions aux enquêtes CIS 2008 et CIS 2010 sur le même champ. Pour ces deux régions, nous pourrions tester sans problème la stabilité des modèles. Pour les régions de taille plus conséquente (Île-de-France notamment) il faudra voir s'il n'est pas possible d'obtenir des informations sur la stabilité des modèles en lien avec la taille des unités échantillonnées dans l'enquête nationale. Cette étude se réalisera au deuxième semestre 2012.

⁸ L'Île-de-France a réalisé des extensions sur les enquêtes CIS 2008 et CIS 2010 mais sur le champ des services technologiques pour CIS 2008 et pour l'industrie en CIS 2010.

Bibliographie

Sur l'innovation et ses déterminants

- [1] Anselin, L., Varga, A. and Acs, Z. (1997), "Local Geographic Spillovers Between University Research and High Technology Innovations ", *Journal of Urban Economics*, no. 42, p. 422-448.
- [2] Arundel, A., C. Bordoy., P. Mohnen and K. Smith (2008), *Innovation Surveys and Policy: Lessons from the CIS*, in *Innovation Policy in Europe: Measurement and Strategy* (C. Nauwelaers and R. Wintjes, eds.), Edward Elgar, 3-28.
- [3] Audretsch, D. B. and Feldman, M. P. (1996) R&D Spillovers and the Geography of Innovation and Production, *The American Economic Review*, vol. 86, no. 3, jun, p. 630-640.
- [4] Autant-Bernard C. and LeSage J. (2010) Quantifying knowledge spillovers using spatial econometric tools, *Journal of Regional Science*, no. doi: 10.1111/j.1467-9787.2010.00705.x.
- [5] Bottazzi, L. and Peri, G. (2003) Innovation and spillovers in regions: Evidence from European patent data, *European Economic Review*, 47-4, pp. 687-710.
- [6] Belderbos R., Carree M. and Lokshin B. (2004) Cooperative R&D and firm performance, *Research Policy* 33, pp. 1477–1492
- [7] Blundell, R., R. Griffith and J. van Reenen (1995), Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms, *Review of Economic Studies*, 66, 529-554.
- [8] Encaoua et Ulph (2000) Catching-up or leapfrogging ? The effects of competition on innovation and growth. *Les cahiers de l'innovation Paris I*, 00017.
- [9] Evangelista R. & Savona M. (2003) Innovation, employment and skills in services. Firm and sectoral evidence, *Structural Change and Economic Dynamics* 14, pp.449-474.
- [10] Hollenstein, H. (1996) A Composite Indicator of a Firm's Innovativeness. An Empirical Analysis Based on Survey Data for Swiss Manufacturing, *Research Policy*, 25, 633-645.
- [11] Jaffe, A.B., (1989) Real Effects of Academic Research, *The American Economic Review*, vol. 79, no. 5, December, p. 957-970.
- [12] Leiponen, A. (2005) Skills and Innovation, *International Journal of Industrial Organization*, 23 (5-6), 303-323.
- [13] Mohnen, P. and L.-H. Röller (2005) Complementarities in Innovation Policy, *European Economic Review*, 49(6), 1431-1450.
- [14] Mairesse, J. and Mohnen, P. (2002) Accounting for Innovation and Measuring Innovativeness: An Illustrative Framework and an Application, *The American Economic Review*, 92, 226-230.
- [15] Mairesse, J. and Mohnen, P. (2005) The Importance of R&D for Innovation: A Reassessment Using French Survey Data, *The Journal of Technology Transfer*, special issue in memory of Edwin Mansfield, 30, 183-197.
- [16] Mairesse, J., and Mohnen, P. (2008) Innovation surveys and innovation policy. mimeo CREST and UNU-MERIT.
- [17] Mairesse, J., and Mohnen, P. (2010) Using innovation surveys for econometric analysis, mimeo UNU-MERIT.

- [18] Muller E, Zenker A., (2001) Business services as actors of knowledge transformation: the role of KIBS in regional and national innovation systems, *Research Policy* 30(9):1501–16.
- [19] Pires, C., Sarkar, S. and Carvalho, L. (2008) Innovation in services: how different from manufacturing?, *The Service Industries Journal*, 28(10), pp. 1339-1356.
- [20] Raymond, W., P. Mohnen, F. Palm, and S. Schim van der Loeff (2010) Persistence of Innovation in Dutch Manufacturing: Is it Spurious?, *The Review of Economics and Statistics*, 92(3): 495–504

Sur l'innovation en région

- [1] Autant-Bernard C., Massard N., Charlot S., Les déterminants géographiques de l'innovation, diffusion spatiale des connaissances et choix de localisation, *Innovation et compétitivité des régions Complément F*, Conseil d'analyse économique rapport 77, La Documentation française, 2008.
- [2] Barret C., Les PME de Provence-Alpes-Côte d'Azur en déficit d'innovation, Insee PACA, Sud Insee l'essentiel numéro 145
- [3] Blazévic B., L'innovation dans les PMI Hauts-normandes, une dynamique d'innovation relativement faible dans les petites structures industrielles, cahier d'aval numéro 89
- [4] Bruley F., L'innovation bénéficie principalement à l'environnement, Insee Franche-Comté, l'Essentiel numéro 124
- [5] Gallard E., La Corse surtout tournée vers l'innovation non technologique, Insee Corse, Quant'île numéro 13
- [6] Girard P., La machine à innover des PME franciliennes : un rendement perfectible au regard du potentiel élevé, Insee Île-de-France, Île-de-France à la page numéro 345
- [7] Godet M., Durance P., Mousli M, 2010, Création et innovation dans les territoires, Conseil d'analyse économique rapport n°4/2010.
- [8] Madies T., Prager J.C., Innovation et compétitivité des régions, Conseil d'analyse économique rapport 77, La Documentation française, 2008.
- [9] Prager J.C., Méthode de diagnostic du système d'innovation dans les régions françaises, Ministère de l'économie, de l'industrie et de l'emploi, 2008.
- [10] Seguin S., L'innovation dans les Pays de la Loire : l'industrie ne manque pas d'atouts, Insee Pays de la Loire, Etudes n°89
- [11] Vaillant E., Innovation en Nord-Pas-de-Calais : un potentiel à optimiser, Insee Nord-Pas-de-Calais, pages de profils numéro 79

Sur les estimateurs petits domaines

- [1] Albouy V, Estimation sur petits domaines : le cas de l'enquête handicaps, incapacités, dépendance, rapport de stage ENSAE été 2000
- [2] Ardilly P., Les techniques de sondages, éditions technip 2006
- [3] Ardilly P., Panorama des méthodes d'estimation sur petits domaines, Méthodologie statistique, Insee Document de travail n°602
- [4] Ardilly P., Estimation sur petits domaines, photocopié formation CEPE
- [5] Attal-Toubert K., Sautory O., Estimation de données régionales à l'aide de techniques d'analyse multidimensionnelle, Méthodologie statistique, Document de travail n°9807
- [6] Caron N., Estimation sur petits domaines, note Insee N°051/E210 mai 2002
- [7] Couet C., Estimations locales dans le cadre de l'enquête HID, Insee document de travail DSDS F0207 novembre 2002
- [8] Destandau S., Estimation sur petits domaines, JMS 2006, Insee Méthodes n°69-70-71