

## DÉVELOPPEMENT DE L'ANNUAL BUSINESS INQUIRY DU ROYAUME-UNI

Gareth Jones<sup>1</sup>

### RÉSUMÉ

Le présent article décrit le projet de l'Annual Business Inquiry (ABI) entrepris par l'Office for National Statistics (ONS) en vue d'intégrer les principales enquêtes annuelles auprès des entreprises des divers secteurs économiques et de fusionner les enquêtes sur l'emploi et les données financières. L'ABI permet aussi de produire une gamme variée d'analyses de niveau infranational, un autre objectif du projet. Les aspects méthodologiques abordés dans l'article incluent le plan d'échantillonnage, l'estimation et le traitement des valeurs aberrantes, la ventilation des données fournies par les unités déclarantes entre les unités locales (emplacements individuels), ainsi que les méthodes d'estimation infranationale et régionale. La méthode d'estimation infranationale s'appuie sur des modèles logistiques et log-linéaires.

MOTS CLÉS : Intégration d'enquêtes; cohérence; analyse infranationale; modèles logistiques et log-linéaires.

### 1. INTRODUCTION

Le présent article décrit le développement de l'Annual Business Inquiry (ABI) du Royaume-Uni, c'est-à-dire une nouvelle enquête intégrée sur l'emploi et les données comptables réalisée auprès des entreprises et d'autres établissements de la plupart des secteurs de l'économie. L'ABI remplace les systèmes d'enquêtes annuelles suivants :

- Annual Employment Survey (AES)
- Annual Censuses of Production and Construction (ACOP/C), qui incluent la Purchases Inquiry (PI);
- les six enquête annuelle Distribution and Services inquiries (DSI), c'est-à-dire :
  - Annual Wholesale Inquiry;
  - Annual Retail Inquiry;
  - Annual Motor Trades Inquiry;
  - Annual Catering Inquiry;
  - Annual Property Inquiry;
  - Annual Service Trades Inquiry.

Outre la restructuration et l'intégration des enquêtes, l'ABI a permis d'apporter d'importantes améliorations méthodologiques et de les appliquer uniformément aux diverses composantes. L'ABI fournit sur les entreprises des données annuelles plus cohérentes, qui couvrent toute une gamme de variables pour l'économie dans son ensemble. Diverses données de sortie de l'ABI ont servi à améliorer la qualité des comptes nationaux. En outre, le projet a permis de produire de meilleures statistiques du genre « par habitant », comme la production par habitant.

### 2. CONTEXTE

Le projet de l'ABI a eu pour toile de fond deux événements importants dans le domaine de la statistique des entreprises, à savoir :

---

<sup>1</sup> Dr D Gareth Jones, UK Office for National Statistics, Room 1.202 Gov't Bldgs, Cardiff Road, Newport, NP10 8XG, UK. Courriel: gareth.jones@ons.gov.uk

- la mise en œuvre, couronnée de succès, de l'Inter-departmental Business Register (IDBR) en 1994 et 1995, qui a servi de fondement à une enquête plus intégrée couvrant « l'ensemble de l'économie »;
- l'introduction par Eurostat (le bureau de la statistique de la Communauté européenne (CE)) d'un nouveau Règlement relatif aux statistiques structurelles sur les entreprises (Règlement SSE) qui étend de plusieurs façons les exigences en matière de données énoncées dans les directives antérieures de la CE. Plus précisément, le Règlement SSE couvre pour la première fois les secteurs économiques des services. Ses dispositions sont entrées en vigueur intégralement pour l'année de référence 1999.

L'ABI remplace les systèmes énumérés plus haut dans l'introduction. L'Annual Employment Survey visait à recueillir des renseignements sur l'effectif de l'entreprise (avec une ventilation hommes/femmes, temps plein/temps partiel) et sur la classe industrielle conformément à la Classification type des industries, CTI (92). Ces données étaient recueillies séparément pour chaque emplacement appartenant à une entreprise sélectionnée dans l'échantillon d'enquête. L'AES (qui a débuté en 1995) fournissait des données utilisées principalement pour produire les estimations annuelles des emplois salariés et remplaçait, elle-même, les recensements antérieurs de l'emploi qui étaient réalisés moins fréquemment. Les données provenant de l'AES fournissaient aussi des renseignements importants pour la mise à jour du Registre des entreprises, particulièrement en ce qui a trait à la structure interne des entreprises, c'est-à-dire les « unités locales » ou emplacements individuels où les entreprises poursuivent leurs activités. Ces renseignements sont importants pour l'analyse régionale, ou à d'autres niveaux de détail géographique, des données sur les entreprises.

Les Annual Censuses of Production and Construction et les Annual Distribution and Services Inquiries étaient des enquêtes par sondage conçues pour recueillir annuellement des données comptables auprès des entreprises, habituellement à un niveau plus détaillé que celui des comptes d'entreprise. L'Annual Census of Production permettait aussi de recueillir certaines données sur l'emploi, quoique de façon légèrement différente de l'AES. Les données comptables tirées de ces enquêtes étaient utilisées à diverses fins. L'une des plus importantes était l'estimation de la valeur ajoutée brute utilisée dans les comptes nationaux pour estimer le produit intérieur brut (PIB). La Purchases Inquiry, réalisée auprès d'un sous-échantillon de l'échantillon de l'ACOP, fournissait des renseignements plus détaillés sur la ventilation des achats de biens et de services par les entreprises. Chaque année, les enquêtes couvraient le cinquième environ des branches d'activité du secteur de la fabrication avec un renouvellement sur cinq ans.

### 3. OBJECTIFS

Les objectifs du projet de l'ABI rentrent dans quatre grandes catégories :

- uniformiser et normaliser les enquêtes et en améliorer la méthodologie;
- satisfaire les exigences internationales;
- produire de nouvelles analyses, particulièrement pour les secteurs des services;
- évaluer la qualité des données et améliorer la tenue à jour du Registre des entreprises.

Le fait de recueillir simultanément les données sur l'emploi et les données financières au moyen d'une même enquête augmente la validité des mesures de la productivité, puisque la valeur ajoutée et l'emploi sont estimés de façon cohérente d'après les données fournies par les mêmes unités déclarantes appartenant aux mêmes branches d'activité. Dans le cas des systèmes d'enquête antérieurs, le manque de comparabilité des données sur l'emploi et sur la valeur ajoutée posait un problème sérieux. L'intégration des diverses branches d'activité en une seule enquête permet d'éviter les problèmes d'omission et de double compte qui étaient possibles par le passé. L'ABI permet aussi d'appliquer la même méthode d'enquête aux divers secteurs économiques, y compris normaliser les définitions. En outre, conformément aux exigences du Système de comptabilité nationale et d'Eurostat, l'ABI couvre l'ensemble du Royaume-Uni, alors que l'AES ne couvrait pas l'Irlande du Nord.

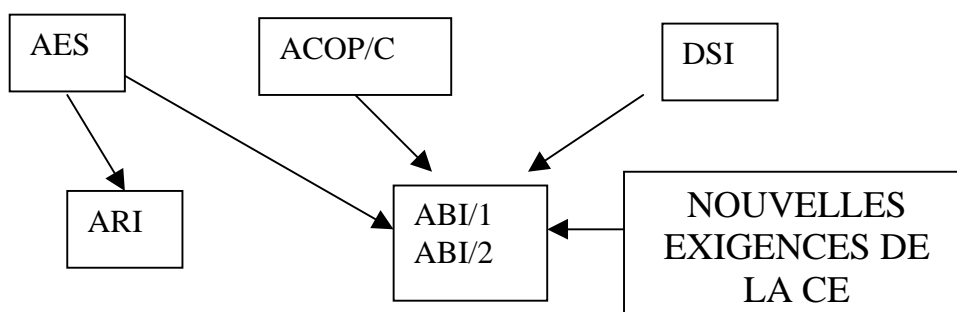
Le Règlement SSE de la Communauté européenne ajoute de nouvelles exigences qui ont été satisfaites en grande partie grâce à l'ajout de questions à l'ABI de 1999. Le règlement exige en outre que l'on utilise des

définitions normalisées afin de permettre les comparaisons entre pays. Il définit aussi un ensemble d'analyses types qui ont été adoptées pour l'ABI. Ces analyses incluent des ventilations selon la région et la taille, ainsi que la ventilation plus classique selon la branche d'activité. Nombre des exigences de la CE, ainsi que d'autres analyses que l'ABI permet de réaliser couvrent tous les secteurs économiques, y compris les services. En particulier, il est maintenant possible de procéder à l'analyse régionale de la valeur ajoutée pour les branches d'activité du secteur des services.

Comme nous l'expliquerons plus bas, une deuxième enquête, appelée Annual Register Inquiry (ARI), a été élaborée parallèlement à l'ABI. L'ARI a pour objectif d'évaluer annuellement divers aspects de la qualité du Registre et d'améliorer le processus de mise à jour de ce dernier et, par conséquent, sa qualité. Elle est conçue pour répondre aux exigences concernant la mise à jour des registres énoncées dans le règlement relatif aux registres des entreprises de la Communauté européenne et fournit des données sur la taille (déterminée d'après l'effectif), la classe industrielle et l'emplacement de l'entreprise complète ainsi que les emplacements individuels où elle poursuit ses activités (unités locales).

#### 4. STRUCTURE DU SYSTÈME

Le diagramme qui suit montre les liens entre les systèmes antérieurs et celui de l'ABI et de l'ARI.



Aujourd'hui, les données sur l'emploi sont recueillies en même temps que les données comptables au moyen de l'ABI, et les enquêtes sectorielles ACOP/C et DSI sont fusionnées. Le questionnaire de l'ABI comprend deux parties, l'une ayant trait aux données sur l'emploi (ABI/1) et l'autre, aux données comptables (ABI/2). Les données sur l'emploi sont recueillies pour une date de référence fixée en décembre et les deux parties du questionnaire sont envoyées aux entreprises à la fin de l'année. Le questionnaire est divisé en deux parties principalement parce que les entreprises sont capables de fournir les données sur l'emploi beaucoup plus tôt que les données comptables. La clôture de la collecte des données au moyen du questionnaire ABI/1 a lieu à la fin de mars de l'année suivante, et au moyen du questionnaire ABI/2, environ six mois plus tard, afin de permettre aux entreprises de fournir les renseignements extraits de leurs propres comptes annuels. Le retour rapide de la partie ABI/1 du questionnaire permet de publier les estimations de l'emploi conformément au même calendrier que celui qui était établi pour l'AES. Le fait que les données sur l'emploi et les données comptables sont souvent fournies par des services différents d'une organisation est une autre raison pour laquelle le questionnaire est divisé en deux parties, car celles-ci peuvent être adressées à des personnes différentes.

En 1998, l'échantillon auquel a été envoyé le questionnaire ABI/1 comptait environ 78 500 entreprises. Le questionnaire ABI/2 est rempli par un sous-échantillon de l'échantillon ABI/1 dans lequel certaines branches d'activité ne sont pas représentées et dont la taille est d'environ 75 000 entreprises. L'ABI/1 couvre les divisions 2 à 93 inclusivement de la CTI (92) – c'est-à-dire toutes les divisions sauf celles de l'agriculture, des ménages privés avec employés et des organismes extra-territoriaux. En 1998, l'ABI/2 couvrait les branches d'activité susmentionnées, sauf les divisions 2 (foresterie), 5 (pêche), 65-67 (services financiers), 75 (administrations publiques) et le secteur des services publics d'éducation (division 80) et de santé et de services sociaux (division 85). Les médecins, les dentistes et les organismes de bienfaisance classés dans la division 85 ont également été omis. En principe, la couverture des branches d'activité par

l'ABI/2 sera étendue dans l'avenir (par exemple les divisions 2 (foresterie) et 5 (pêche) de la CTI ont été couvertes par l'ABI/2 en 2000). L'échantillon de la Purchases Inquiry a été retenu provisoirement comme sous-échantillon de l'ABI/2, les questions supplémentaires sur les achats étant ajoutées au questionnaire de l'ABI/2 pour les secteurs pertinents. Nous sommes en train de procéder à l'expansion graduelle de la Purchases Inquiry. À l'heure actuelle, il est prévu d'étendre progressivement la couverture de l'ABI/2 d'année en année jusqu'à ce que toutes les branches d'activité qui doivent être couvertes le soient à compter de 2001. Cependant, nous pourrions décider de réaliser la Purchases Inquiry en tant qu'enquête distincte pour éviter de surcharger le questionnaire de l'ABI/2.

Comme nous l'avons mentionné plus haut, le lancement de l'ABI a été accompagné de celui d'une autre enquête, l'Annual Register Inquiry (ARI), dont les opérations ont débuté en juillet 1999. L'ARI comprend les activités antérieures de vérification du Registre, remplace l'AES en ce qui concerne la mise à jour des données sur la structure des unités locales des entreprises, et est conçue en fonction de cette dernière activité plutôt que pour la mesure directe de l'emploi. Elle couvre les grandes entreprises chaque année et les moyennes entreprises (de 20 à 99 employés) tous les quatre ans. Pour les entreprises dont l'effectif est inférieur au seuil de 20, la vérification n'est pas systématique étant donné leur très grand nombre. Cependant, une partie de l'échantillon de l'ARI est affectée à l'évaluation de la qualité des diverses caractéristiques du Registre (particulièrement la classification des industries). Une autre partie de l'échantillon est réservée au ciblage ponctuel de domaines du Registre que l'on sait poser des problèmes de qualité. Par conséquent, l'ARI est un outil fort souple de mise à jour de l'IDBR d'évaluation et d'amélioration de sa qualité. Chaque année, cette enquête couvre environ 400 000 unités locales appartenant à 68 000 entreprises.

## 5. CALENDRIER DU PROJET

Le calendrier qui suit décrit l'introduction par étape du système de l'ABI. Les dates indiquées ici sont les années de référence auxquelles se rapportent les données et non les dates de communication des résultats.

- 1996 AES, ACOP/C, DSI
- 1997 AES, ACOP/C modifiée et DSI pour une exécution en parallèle
- 1998 AES, ABI
- 1999 ARI, ABI (conforme à toutes les exigences de la CE)

Les premiers changements ont eu lieu pour les enquêtes concernant 1997. L'ACOP/C et la DSI ont été modifiées de façon à recueillir des données sur l'emploi selon des critères comparables à ceux de l'AES pour 1997. Dans le cas de l'ACOP/C, il a fallu étendre le questionnaire existant sur l'emploi et modifier la définition pour passer d'un chiffre annuel moyen à un chiffre pour une date de référence donnée. Les mêmes questions ont été introduites dans le questionnaire de la DSI qui, au préalable, ne contenait aucune question sur l'emploi. Ces changements avaient pour but de répondre à la demande des clients qui souhaitaient que l'on recueille parallèlement des données selon la méthodologie de l'ABI et selon celle des systèmes antérieurs. Le logiciel de calcul et d'analyse des résultats de l'ABI a été rédigé en vue d'être utilisé pour la première fois avec les données recueillies pour 1997, bien qu'elles l'aient été au moyen des systèmes d'enquête antérieurs. L'exécution en parallèle visait à la fois à évaluer la nouvelle méthodologie et à estimer toute rupture introduite dans la série chronologique à cause des changements méthodologiques.

Les autres changements apportés pour 1997 incluaient, dans le cas de la DSI, le passage de l'utilisation du chiffre d'affaires à celle de l'effectif pour la stratification par taille prévue dans le plan de sondage, afin de réaliser la DSI et l'ACOP/C dans les mêmes conditions et de combler certaines lacunes mineures du champ d'observation des branches d'activité de l'ACOP/C et de la DSI. Certains travaux ont visé à vérifier la qualité de l'IDBR pour les branches d'activité en question, en prévision de cette extension de la couverture.

En outre, le logiciel de collecte des données utilisé pour l'ACOP/C a été modifié de façon à ce qu'il concorde avec celui de la DSI, qui utilisant déjà un logiciel standard de collecte des données.

Le lancement de l'ABI pour l'année de référence 1998 a achevé l'intégration des systèmes de l'ACOP/C et de la DSI. L'exercice a nécessité un examen complet des plans d'échantillonnage, des types de questionnaires, des questions et des définitions dans la perspective d'un système unique plutôt qu'un ensemble d'enquêtes sectorielles distinctes. Pour 1998, l'AES a été réalisé parallèlement à l'ABI afin d'obtenir une exécution parallèle plus complète pour les questions sur l'emploi. La transition au nouveau système de l'ABI et de l'ARI a été achevée pour l'année de référence 1999, au moyen du lancement de l'ARI en juillet 1999 et de l'ajout de questions supplémentaires aux questionnaires de l'ABI de 1999 en vue de répondre à toutes les exigences du Règlement SSE de la Communauté européenne. Toutefois, comme nous l'avons décrit plus haut, l'extension de la Purchases Inquiry ne sera pas achevée avant 2001.

L'adoption officielle des estimations du nombre d'emplois salariés calculées d'après l'ABI/1 plutôt que d'après l'AES s'est compliquée du fait que les estimations calculées pour 1998 d'après l'ABI/1 excédaient d'environ un million celles calculées d'après l'AES. Le programme approfondi de recherche qu'il a fallu mettre sur pied pour déterminer les raisons de cet écart a mené à la conclusion que les estimations fondées sur l'ABI sont les plus fiables. Par conséquent, la concordance entre les estimations de l'emploi au Royaume-Uni calculées d'après les enquêtes auprès des entreprises et l'Enquête sur la population active est désormais nettement meilleure.

## **6. QUESTIONNAIRES TYPES, QUESTIONS ET DÉFINITIONS**

Durant les préparatifs du lancement de l'ABI à la fin de 1998, nous avons examiné minutieusement tous les questionnaires types utilisés, les questions qu'ils contenaient et les définitions de ces questions. Nous nous sommes efforcés de réduire au minimum le nombre de questionnaires types tout en retenant le contenu considéré comme pertinent pour les entreprises auxquelles les questionnaires s'adressaient. Les questionnaires types sont personnalisés selon le secteur et le sous-secteur d'activité et, en 1998, il existait trois questionnaires types de base pour l'ABI/1 et 21 pour l'ABI/2. En outre, pour chaque questionnaire types de base, il existe habituellement un questionnaire abrégé équivalent destiné à recueillir des renseignements sur les totaux principaux, mais non les données ventilées qui figurent sur les questionnaires types de base. Les questionnaires abrégés sont utilisés pour réduire le fardeau de réponse des entreprises. Dans le cas de l'ABI/2, une certaine proportion d'entreprises faisant partie de l'échantillon reçoivent le questionnaire abrégé. Cette proportion augmente à mesure que la taille des entreprises diminue. Les entreprises qui reçoivent le questionnaire abrégé sont sous-échantillonnées aléatoirement à partir de l'échantillon total de chaque strate. Dans le cas de l'ABI/1, la démarche est différente, le questionnaire abrégé étant utilisé pour les entreprises qui ont également été échantillonnées pour l'ARI ou pour l'enquête sur l'emploi à court terme réalisée au quatrième trimestre, ainsi que pour toutes les entreprises d'Irlande du Nord. Des méthodes d'imputation sont appliquées pour extrapoler l'information recueillie au moyen des questionnaires abrégés au niveau de détail requis pour l'analyse.

Les principales questions posées dans le cadre de l'ABI/1 ont trait au nombre d'employés ventilé selon le sexe (hommes/femmes) et le genre d'emploi (temps plein/temps partiel) (comme dans l'AES), ainsi que le nombre de propriétaires/associés actifs et le nombre d'autres employés non rémunérés (par exemple, membre de la famille), ainsi que l'effectif total qui inclut toutes les catégories susmentionnées de travailleurs. De surcroît, les définitions concordent aux définitions des employés utilisées pour l'AES.

En général, le questionnaire de l'ABI/2 contient un plus grand nombre de questions que celui de l'ABI/1 et la gamme de questions varie davantage selon la branche d'activité. Cependant, il comprend un ensemble de questions de base qui couvrent le chiffre d'affaires (c'est-à-dire les ventes de biens et de services), les coûts salariaux, les achats de biens et de services, les impôts et les subventions, les stocks et l'investissement en biens d'équipement et qui figurent sur presque tous les questionnaire types. En outre, les questionnaires types de base contiennent habituellement des questions sur des ventilations plus détaillées de ces agrégats. Les définitions de ces quantités ont été examinées et normalisées. En général, les définitions utilisées sont

maintenant conformes aux exigences du Règlement SSE qui, elles-mêmes, concordent habituellement avec celles du Système européen des comptes (SEC).

En outre, le questionnaire de l'ABI/2 contient plusieurs « questions filtres » utilisées pour repérer les entreprises dont les activités rentrent dans des catégories particulières, afin d'améliorer la base de sondage en prévision d'autres enquêtes plus approfondies. Des questions filtres sur les activités de recherche et de développement et sur le commerce international dans le secteur des services, ainsi que des questions sur le commerce électronique, ont été incluses dans les questionnaires de l'ABI de 2000.

## 7. ANALYSES DISPONIBLES

Le tableau qui suit résume la gamme d'analyses des données de l'ABI disponibles sous réserve de problème d'exactitude et de confidentialité.

Recueillies	Dérivées	Analysées par
Emploi	Valeur ajoutée brute (VAB)	Classe de la CTI
Chiffre d'affaires	Production totale	Zone géographique
Coûts salariaux	Excédent d'exploitation	Taille (effectif ou chiffre d'affaires)
Achats	VAB /chiffre d'affaires	Situation juridique
Impôt/subventions	Productivité du travail	Pays de propriété
Stocks	Charge salariale unitaire	Période (c.-à-d. séries chronologiques)
Investissement	Ratios sur les stocks	
	Ratios sur l'investissement	

Des ventilations sont également disponibles pour la plupart des variables principales pour lesquelles les données sont recueillies. Par exemple les coûts salariaux sont subdivisés en salaires et traitements, cotisations aux régimes de pension, cotisations aux régimes de sécurité sociale et indemnités de départ. Les ventilations produites pour le chiffre d'affaires varient selon la branche d'activité et, dans le cas du commerce de détail, des données détaillées sont calculées selon le produit. En outre, la Purchases Inquiry fournit des renseignements fort détaillés sur les achats pour les branches d'activité faisant partie du champ de l'enquête.

Plusieurs variables importantes peuvent être dérivées des données recueillies, notamment la valeur ajoutée brute et la production totale, toute deux aux prix de base, l'excédent d'exploitation, la productivité du travail et la charge salariale unitaire. Tant pour les variables recueillies que pour les variables dérivées, les données peuvent être analysées en fonction de plusieurs autres variables, choisies principalement dans l'IDBR. En plus des analyses selon la classe de la CTI et la taille, les entreprises peuvent être classées et analysées selon le pays de propriété et la situation juridique (c.-à-d. sociétés, sociétés de personnes, entreprises individuelles, sociétés publiques, organismes sans but lucratif, administrations centrales ou locales). Les données peuvent aussi être analysées selon des variables géographiques, conformément au système de classification hiérarchique de la Nomenclature des unités territoriales statistiques (NUTS) de la Communauté européenne. La NUTS compte cinq niveaux hiérarchiques qui correspondent à peu près à l'organisation des administrations locales :

- NUTS1 Région
- NUTS2 Groupe de comtés ou d'autorités unitaires
- NUTS3 Comté ou autorité unitaire
- NUTS4 District ou autorité unitaire
- NUTS5 Quartier

Il est, en outre, possible de comparer diverses années de référence et de produire des séries chronologiques, au besoin.

## 8. MÉTHODOLOGIE DE L'ABI

### 8.1 Plan de sondage

L'échantillon de l'ABI est sélectionné à partir de la population d'unités déclarantes du Registre des entreprises (IDBR). Une unité déclarante (UD) est habituellement une entreprise unique complète, mais, dans certains cas, des dispositions ont été prises avec les entreprises afin de recueillir des données soit pour plusieurs entreprises regroupées, soit pour des parties d'entreprise de nature hétérogène. Chaque UD qui figure dans l>IDBR est en outre constituée d'une ou de plusieurs unités locales (c.-à-d. emplacements individuels) où l'entreprise est en exploitation. Les unités locales sont particulièrement importantes pour toute forme d'analyse géographique.

Pour 1998, la taille de l'échantillon de l'ABI/1 était d'environ 78 500 unités. Celle de l'ABI/2 était légèrement plus faible, soit environ 75 000 unités, parce que certaines branches d'activité ne sont pas couvertes par cette enquête. La sélection de l'échantillon se fait pour l'ABI/1, puis l'échantillon de l'ABI/2 est obtenu automatiquement par élimination des branches d'activité appropriées. Les unités sont sélectionnées selon un plan d'échantillonnage aléatoire stratifié à trois dimensions de stratification. Les strates sont définies en fonction de :

- six catégories d'effectif (1 à 9 employés, 12 à 19, 20 à 49, 50 à 99, 100 à 249, plus de 250);
- la région (Angleterre et pays de Gales combinés/Écosse/Irlande du Nord);
- la branche d'activité selon la CTI.

Pour l'Angleterre et le pays de Gales, la stratification selon la branche d'activité se fait au niveau à quatre chiffres de la CTI. Pour l'Irlande du Nord, elle se fait au niveau à deux chiffres et pour l'Écosse, à un niveau hybride de 2/3/4 chiffres. Toutes les variables de stratification sont sélectionnées dans l>IDBR. Des dispositions particulières ont été prises avec les bureaux de l'Écosse et de l'Irlande du Nord en vue d'allouer une part proportionnellement plus grande de l'échantillon à ces deux régions (c.-à-d. 9 000 unités et 3 000 unités, respectivement).

Sous réserve des contraintes de taille susmentionnées, l'échantillon a été réparti entre les strates selon la méthode de répartition optimale de Neyman qui consiste à minimiser la variance prévue du chiffre d'affaires total sur l'ensemble des strates. On obtient de la sorte, pour la strate qui correspond aux entreprises les plus grandes, un dénombrement complet pour toutes les branches d'activité. Dans la plupart des cas, le dénombrement est également complet pour certaines strates correspondant à des catégories de taille plus petite. En outre, les cas pour lesquels le chiffre d'affaires consigné dans le Registre est élevé tandis que l'effectif est faible (définis comme un chiffre d'affaires égal ou supérieur à 50 millions de £ et un effectif de moins de 10 employés) font aussi l'objet d'un dénombrement complet. Cette démarche produit une stratification limitée selon la variable de chiffre d'affaires et réduit considérablement la variance prévue des estimations calculées d'après l'échantillon. En 1999, cette méthode a été étendue afin de procéder à un dénombrement complet des entreprises dont le chiffre d'affaires enregistré était élevé, indépendamment de l'effectif enregistré. Un examen complet de la répartition de l'échantillon est prévu maintenant que l'on dispose de données de l'ABI pour deux années complètes.

Pour les entreprises tombant sous le seuil de dénombrement complet, l'échantillon est renouvelé comme suit. Les entreprises comptant moins de 10 employés sont remplacés entièrement chaque année. Celles dont l'effectif varie de 10 employés au seuil de dénombrement complet sont échantillonnées avec un taux de renouvellement de 50 % (c.-à-d. renouvellement de la moitié des unités chaque année). Ce plan de renouvellement vise à répartir le fardeau de réponse entre les entreprises tout en produisant des échantillons présentant un degré raisonnable d'appariement entre années consécutives pour augmenter l'exactitude des estimations de la variation annuelle.

### 8.2 Analyses au niveau national

Les données provenant de l'échantillon sont étendues à l'ensemble de la population de l>IDBR par la méthode de l'estimateur combiné du quotient, les strates correspondant à diverses catégories de taille étant

ainsi regroupées si l'échantillon est trop petit autrement. Cette estimateur s'appuie sur des variables auxiliaires du Registre afin d'améliorer l'exactitude de l'estimation.

La variable auxiliaire utilisée pour étendre à l'ensemble de la population observée les données sur l'emploi de l'ABI/1 et celles sur les coûts salariaux de l'ABI/2 est la variable d'emploi de l'IDBR. Pour toutes les autres données de l'ABI/2, la variable auxiliaire utilisée est le chiffre d'affaires. La variable auxiliaire a été choisie après examen de la taille des erreurs prévues d'échantillonnage pour diverses variables auxiliaires.

Une méthode de traitement des observations atypiques (valeur aberrantes) a été mise au point pour améliorer la robustesse de l'estimateur susmentionné. Cette méthode consiste à créer des strates a posteriori supplémentaires dans lesquelles les unités aberrantes sont classées comme si elles avaient été dénombrées entièrement au stade de l'échantillonnage (elles sont également supprimées de leur strate « naturelle »). Les critères utilisés pour considérer une unité comme étant aberrante ont été établis empiriquement de façon à obtenir des estimations robustes. Dans le cas de l'ABI/1, une unité est considérée comme aberrante si le rapport de son effectif déclaré à son effectif enregistré dans la base de sondage est supérieur à 20. Dans le cas de l'ABI/2, l'unité est considérée comme aberrante si elle est jugée aberrante pour l'ABI/1 ou que le rapport de son chiffre d'affaires déclaré à son chiffre d'affaires enregistré dans la base de sondage est supérieur à 50. Toute unité considérée comme aberrante est traitée en tant que telle pour toutes les variables visées par l'enquête.

L'importance de l'erreur d'échantillonnage à laquelle il faut s'attendre pour l'effectif de l'échantillon/plan de sondage et l'estimateur connexe est la suivante. Le coefficient de variation de l'estimation de l'effectif total pour toutes les branches d'activité est d'environ 0,4 % et le coefficient de variation de la valeur ajoutée totale pour l'ensemble des branches d'activité est d'environ 1,5 %. Autrement dit, pour l'emploi total, l'intervalle de confiance à 95 % et de  $\pm 0,8$  % et, pour la valeur ajoutée totale, il est de  $\pm 3$  %.

### **8.3 Répartition des unités locales**

Comme les données de l'ABI sont recueillies au niveau des unités déclarantes (habituellement des entreprises complètes), la première étape de toute analyse au niveau infranational consiste à répartir les données par les unités déclarantes entre les unités locales (UL) que contient l'unité déclarante. La répartition est exécutée, pour chaque variable étudiée X, par modélisation du rapport  $R = X$  divisé par l'effectif enregistré dans la base de sondage. Le modèle comprend deux éléments. Premièrement, on suppose que la probabilité d'une valeur non nulle est donnée par un modèle logistique linéaire ayant pour variables indépendantes la catégorie d'effectif, le code de la CTI et le code géographique qui figurent dans le Registre des entreprises. Deuxièmement, étant donné une valeur non nulle, on suppose que R est représenté par un modèle log-linéaire ayant le même ensemble de variables indépendantes que le modèle logistique; on se sert des paramètres estimés pour ces modèles pour calculer, pour chaque unité locale, les proportions (dont la somme est égale à l'unité pour l'ensemble des unités locales comprises dans une unité déclarante) qui déterminent la répartition du total de l'unité déclarante entre les unités locales. Les modèles sont ajustés aux données recueillies pour l'unité déclarante, puis appliqués aux unités locales afin de procéder à la répartition (les données sur les variables indépendantes figurent dans l'IDBR pour les unités déclarantes ainsi que pour les unités locales). Afin que cette conversion de l'unité déclarante à l'unité locale soit aussi fiable que possible, il n'est pas tenu compte des unités déclarantes les plus grandes lors de l'ajustement du modèle (en fait, seules les unités déclarantes comptant moins de 100 employés et moins de 3 unités locales sont utilisées pour ajuster le modèle).

Lors de la modélisation, des décisions doivent également être prises quant au degré approprié d'agrégation des catégories d'effectif, des codes de la CTI et des codes géographiques en vue de leur utilisation comme variables indépendantes. Il a été décidé de travailler au niveau d'agrégation le plus élevé qui permette de ne pas sacrifier une part importante de la « qualité de l'ajustement » du modèle. Nous avons obtenu ainsi les groupements qui suivent, qui sont utilisés pour toutes les variables d'enquête :

- catégorie d'effectif : 1 à 2 employés, 3 à 4, 5 à 9, 10 à 19, 20 à 49, 50 à 99, 100 à 249, plus de 250;
- branches d'activité : code à 2 chiffres de la CTI pour le modèle logistique et code à 3 chiffres de la CTI pour le modèle log-linéaire;



- Région : niveau NUTS2

Pour les questions sur l'emploi de l'ABI/1 (c.-à-d. nombre total d'employés et ventilation hommes/femmes, plein temps/temps partiel), d'autres sources de données peuvent être utilisées pour répartir la valeur obtenue pour l'UD entre les UL, à savoir l'ARI pour la Grande-Bretagne et le recensement biennal de l'emploi pour l'Irlande du Nord. Les renseignements sur la répartition des employés des unités locales provenant de ces sources pourraient être meilleurs que ceux obtenus d'après le procédé de modélisation susmentionné et sont donc utilisés de préférence aux données du modèle, s'ils existent.

## 8.4 Analyses au niveau infranational

Les analyses au niveau infranational des données de l'ABI sont réalisées par application d'une combinaison de méthodes classiques et synthétiques d'estimation. Au niveau le plus élevé d'agrégation, on utilise les méthodes classiques d'estimation, tandis qu'aux niveaux plus faibles, les méthodes classiques sont complétées par une estimation synthétique au moyen des données de l>IDBR. Ces méthodes sont décrites plus bas.

Pour commencer, on procède à une stratification a posteriori de toutes les unités locales comprises dans l'échantillon et dans la population d'UD. Les variables de stratification a posteriori sont la région de l'UL, la catégorie d'effectif de l'UD mère et le code de la CTI de l'UD mère. La catégorie d'effectif de l'UD mère est la même que celle utilisée pour la stratification dans le cas des analyses au niveau national. Sinon, les niveaux d'agrégation sont la région NUTS1 de l'UL et le code CTI à deux chiffres de l'UD mère, sauf dans certains cas où une agrégation diverse des strates a posteriori (en ce qui concerne tant la région que le code CTI) est nécessaire pour obtenir une taille d'échantillon appropriée. L'estimateur utilisé est l'estimateur par quotient ordinaire (non combiné), avec l'effectif enregistré de l'UL comme variable auxiliaire.

L'utilisation d'une autre stratification a posteriori pour les analyses infranationales pose un problème de concordance entre les résultats nationaux et infranationaux, dans la mesure où le total des résultats régionaux pour le Royaume-Uni dans son ensemble pourrait ne pas être égal à la valeur obtenue pour le Royaume-Uni d'après le système d'estimation nationale. Pour surmonter ce problème, on a produit pour chaque variable d'enquête des facteurs d'échelle qui sont appliqués aux coefficients de pondération calculés d'après le système d'estimation infranationale et fixés de telle sorte que les conditions d'additivité énoncées plus haut soient respectées. Enfin, il convient de souligner que les totaux et d'autres formes de variables dérivées sont calculés après extension des données d'échantillon à l'ensemble de la population au lieu d'être étendus individuellement à l'ensemble de la population dans le système d'estimation infranationale, car cette façon de procéder donnerait lieu à une autre forme de non-additivité ou d'incohérence.

En principe, la méthode décrite plus haut fournira des estimations pour toute sous-population requise (ou domaine, selon la terminologie utilisée plus bas), y compris les quartiers relevant des autorités locales. Cependant, en pratique, les erreurs d'échantillonnage sont fort importantes pour les très petits domaines. En outre, certains domaines ne contiennent aucune unité d'échantillonnage, si bien que l'estimation résultante est nulle. Bien qu'une estimation nulle ne soit pas nécessairement en contradiction avec les erreurs d'échantillonnage prévues, elle pose des problèmes de présentation, puisque l'on pourrait savoir d'avance que l'estimation correcte pour le domaine est supérieure à zéro.

Il n'est, par conséquent, pas réaliste d'utiliser la méthode d'estimation susmentionnée pour les très petits domaines et il a fallu définir un ensemble de « domaines minimaux » auxquels elle est applicable. Ces domaines minimaux sont bidimensionnels en ce sens qu'ils sont définis en fonction d'un niveau de désagrégation donné de la CTI et un niveau de désagrégation géographique donné, qui peuvent l'un et l'autre être différents pour un domaine minimal particulier. On a tenu compte, pour définir les domaines minimaux, de l'erreur d'échantillonnage associée à ces domaines et de la taille d'échantillon en fonction du nombre d'unités locales qu'ils contiennent.

Au dessous du niveau des domaines minimaux, on recourt à une forme d'estimation synthétique pour obtenir les estimations voulues. Essentiellement, si un domaine correspond à un sous-ensemble d'un

domaine minimal, les estimations le concernant sont produites par calcul proportionnel d'après l'effectif total enregistré dans le Registre pour le domaine étudié et l'effectif total enregistré pour le domaine minimal. Les unités échantillonnées sont éliminées avant le calcul proportionnel et rajoutées après. Tout domaine qui chevauche deux domaines minimaux ou plus est divisé en domaines plus petits qui ne chevauchent pas les limites des domaines minimaux. Les estimations sont alors calculées séparément pour ces plus petits domaines, puis totalisées. Le recours à l'estimation synthétique pour les très petits domaines signifie qu'on ne dispose pas d'estimation de l'erreur. L'erreur d'échantillonnage n'augmente pas au-delà du niveau du domaine minimal, mais d'autres formes d'erreurs, non quantifiables, pourraient être introduites.

## **9. TRAVAUX FUTURS**

Les quatre domaines où il est prévu d'entreprendre des travaux dans l'avenir sont les suivants :

- perfectionnement des questionnaires types et de la conception des questionnaires en vue de réduire le coût d'information, particulièrement pour les petites entreprises;
- étude visant à déterminer si l'exactitude des estimations de l'emploi pourrait être améliorée grâce à l'utilisation de données administratives sur la paye comme données auxiliaires lors de l'estimation;
- recherche d'une meilleure méthode de traitement des valeurs aberrantes, particulièrement l'utilisation d'une méthode à test bilatéral plutôt qu'unilatéral
- séparation éventuelle de la Purchases Inquiry et de l'ABI/2 afin d'éviter le surchargement du questionnaire.