Nations Unies ECE/CES/GE.42/2011/3

Conseil économique et social

Distr. générale
5 juillet 2011

Original: français

Commission économique pour l'Europe

Conférence des statisticiens européens

Groupe d'experts des registres des entreprises

Douzième session
Paris, 14-15 septembre 2011
Point 5 de l'ordre du jour provisoire
Allègement de la charge de travail des répondants - le rôle des registres des entreprises

La construction d'un répertoire statistique, lié au répertoire interadministratif d'entreprises du Système national d'Identification du Répertoire des Entreprises et de leurs Etablissements : enregistrer les échantillons afin de mesurer la charge de réponse

Note établie par l'Institut national de la statistique et des études économiques

Résumé

Depuis longtemps la France dispose d'un répertoire administratif dont l'unité de base est l'unité légale. A vocation administrative, ce répertoire a été largement utilisé à de fins statistiques car dans le système statistique d'entreprises français la notion d'entreprise était jusqu'à présent assimilée à la notion d'unité légale.

Au milieu des années 2000, un vaste programme de refonte des statistiques structurelles d'entreprises, a eu pour ambition de refondre les statistiques d'entreprises, en particulier d'introduire une vraie définition de l'entreprise, qui ne serait plus l'unité légale. La construction d'un répertoire statistique enregistrant les entreprises avait alors été actée dès le lancement de ce programme.

Ce répertoire, en cours de construction, permet de mobiliser de nouvelles unités statistiques, en particulier les « entreprises », de connaitre leur composition en unités légales et d'assurer l'homogénéité des caractéristiques statistiques au sein du système statistique d'entreprises.

Il sera également à terme l'outil permettant de mesurer la charge de réponse des entreprises, demande récurrente (cf. « French methods and practices regarding the statistical burden ») à laquelle il n'est pas possible actuellement de répondre simplement.

Merci de recycler

I. Du Système national d'Identification du Répertoire des Entreprises et de leurs Etablissements, répertoire administratif…

1. Jusqu'en 2010, la France disposait d'un répertoire administratif exhaustif, décrit en détails dans "Improving the quality of Business Registers" (cf. références).

2. Toutes les unités légales et tous leurs établissements situés en France sont immatriculés dans un répertoire national d'identification appelé SIRENE[1]. Ce répertoire s'inscrit dans un cadre réglementaire dans lequel l'Institut national de la statistique et des études économiques (INSEE) a pour mission d'identifier les unités légales et leurs établissements, à partir de documents transmis par des centres de formalités des entreprises (CFE), centres créés pour prendre en compte l'ensemble des démarches administratives qui incombent aux unités légales. Depuis une loi du 11/02/1994 créant le numéro unique d'identification, il est également au cœur du système inter-administratif, cette identification s'imposant à toutes les administrations pour leurs relations avec les entreprises.

3. Le répertoire SIRENE enregistre l'état civil des unités légales, leur forme juridique, l'adresse de leur siège social, la date de leur création et, le cas échéant, leur cessation d'activité. Pour les établissements sont indiquées les adresses ainsi que les dates de création. Outre ces données d'identification, le répertoire contient, pour chacune des unités inscrites, l'effectif salarié et le code d'Activité Principale Exercée (APE) attribué par l'INSEE en référence à la Nomenclature d'Activités Française (NAF).

4. Cette richesse fait, qu'en dehors de l'utilisation réglementaire due au caractère inter-administratif du répertoire, SIRENE était aussi exploité dans la sphère statistique et dans la sphère commerciale : petit à petit il avait donc été enrichi de données « statistiques » non règlementaires (indicatrice de consistance économique, des variables concernant les effectifs salariés au 31/12 de chaque année, en équivalents temps plein, chiffre d'affaires annuel, etc.).

5. Ce répertoire présente toutefois deux inconvénients majeurs lorsqu'il est utilisé pour produire des statistiques. Il s'agit d'une part d'un répertoire limité aux unités légales (les sociétés et leurs établissements) : avec la mondialisation des groupes et le nombre croissant, l'assimilation « unité légale » / « entreprise »2 n'est plus pertinente dans certains cas. Ces cas ont malheureusement un poids économique important et l'assimilation de l'unité légale à l'entreprise peut amener, par le biais de double compte en termes de chiffre d'affaires par exemple, à donner une image fausse de l'économie.

6. D'autre part, dans un répertoire administratif, on ne peut pas « faire ce qu'on veut »: on ne peut ni ajouter, ni cesser des unités sans que les associés réglementaires ou les unités enregistrées elles-mêmes n'en soient informés, ce qui ne permet pas d'avoir la souplesse qu'on pourrait attendre d'un répertoire plus statistique.

[1] Créé par le décret modifié du n° 73-314 du 14 mars 1973.

[2] c'est-à-dire « plus petite combinaison d'unités légales qui constitue une unité organisationnelle de production de biens et de services jouissant d'une certaine autonomie de décision, notamment pour l'affectation de ses ressources courantes ».

II. ... au Système d'Identification au Répertoire des Unités Statistiques, répertoire statistique

A. Refonte des Statistiques Annuelles d'Entreprises, un programme structurant

7. Le début de l'année 2006 a vu la naissance à l'INSEE d'un projet très structurant pour les statistiques d'entreprises : le programme de Refonte des Statistiques Annuelles d'Entreprises (RESANE) dont les réalisations se sont échelonnées depuis 2009 et continuent aujourd'hui. Ses trois objectifs étaient :

- de diminuer la charge de réponse des entreprises en recourant le plus systématiquement possible aux sources administratives ;
- de faire des gains de productivité globaux et pour cela, procéder à une réingénierie complète des processus mis en œuvre jusqu'à présent pour produire des statistiques structurelles d'entreprises ;
- d'améliorer la qualité des statistiques produites en termes de délai (gains de productivité) et en termes de pertinence et donc, de se rapprocher de la définition exacte de l'entreprise (définie par le règlement européen 696/93) en prenant en compte la dimension groupe de sociétés3 dans l'élaboration des statistiques d'entreprises.

8. Ce dernier objectif « prise en compte de la dimension groupe dans la statistique d'entreprises » vise à suivre deux types d'unités supplémentaires que ne gérait pas jusqu'ici le répertoire administratif à savoir : l'entreprise et le groupe de sociétés.

9. Pour y parvenir, deux projets complémentaires ont démarré en parallèle[4] :

(a) le premier, appelé « Profilage », consiste à définir, à partir des groupes d'unités légales, des entreprises au sens du règlement européen 696/93. La France procèdera par cibles successives :

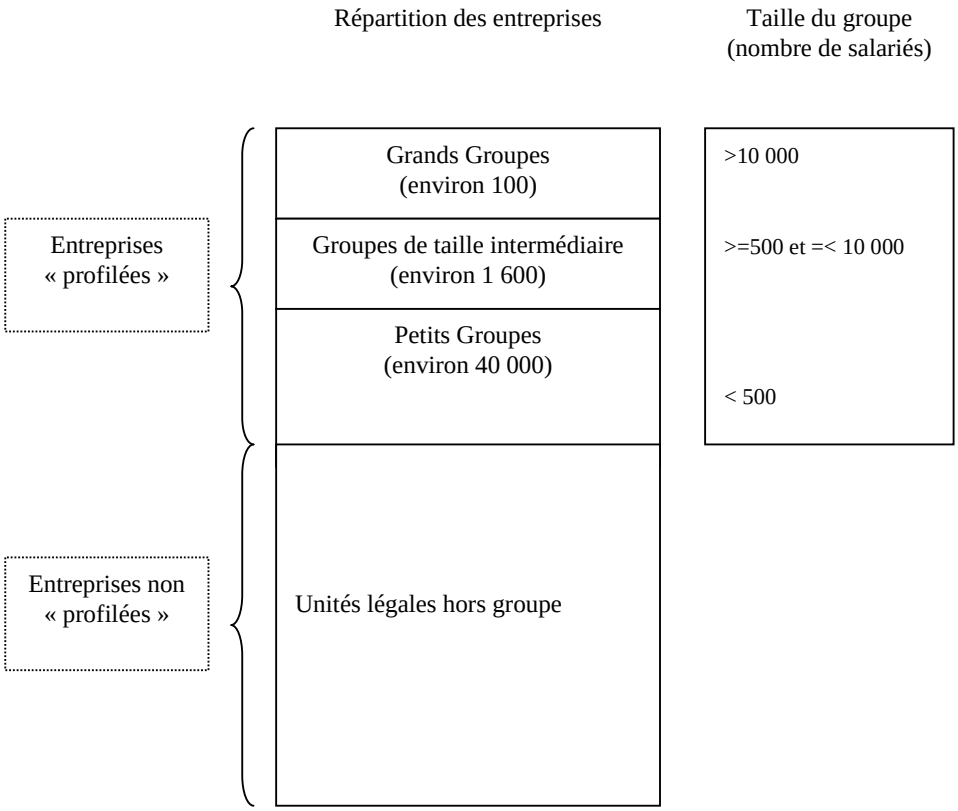
- Pour une centaine de groupes (22 % salariés, 26 % de la valeur ajoutée) parmi les plus grands et/ou les plus complexes et/ou les plus « demandeurs », les entreprises seront définies en accord avec le groupe à partir duquel elles sont construites : le profilage sera donc manuel, les entreprises définies lors d'un entretien annuel avec le groupe ;
- Pour près de 38,000 « petits » groupes (moins de 500 salariés ou possédant 2 filiales françaises au plus), l'entreprise sera assimilée au groupe dans sa totalité et ses caractéristiques (secteur d'activité par exemple) obtenues grâce à un algorithme ; ces groupes représentent 19 % salariés et 22 %de la valeur ajoutée ;
- Enfin environ 1,600 groupes de taille intermédiaire (17 % des salariés et de la valeur ajoutée) seront « profilés » grâce à une méthode intermédiaire, mi-manuelle mi-automatisée, moins fine donc que le profilage individuel des grands groupes. Cette méthode pourrait s'appuyer sur les résultats d'une enquête spécifique, qui viserait à déterminer des coefficients de consolidation au sein des groupes.

[3] Un groupe est un ensemble d'unités légales juridiquement indépendantes, mais liées entre elles par des participations au capital et parmi lesquelles l'une de ces sociétés exerce sur les autres un pouvoir de décision. En France, les groupes sont suivis, par enquête, depuis 1980.

[4] Ce sont d'ailleurs les mêmes équipes qui travaillent sur ces projets !

Toutes les entreprises ainsi définies seront appelées « entreprises profilées ». Les unités légales n'appartenant pas à un groupe resteront quant à elles assimilées à des entreprises.

Figure 1
Définition des entreprises à partir des groupes



(b) le second sous-projet (dont les développements ont commencé en 2009 mais qui avait été imaginé dès 2005 dans la réflexion ayant mené au programme RESANE) consiste à mettre en place un répertoire statistique avec pour unité principale « l'entreprise » dont les contours auront été définis par le projet décrit dans le point « a ».

B. Le répertoire du Système d'Identification au Répertoire des Unités Statistiques

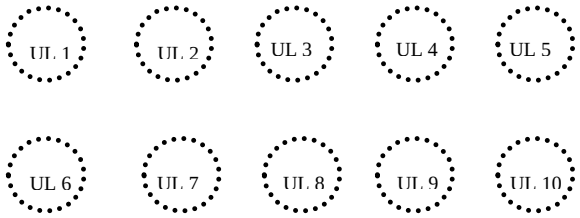
1. La prise en compte de nouvelles unités statistiques et leurs liens entre elles

10. Concrètement, comment est construit ce répertoire statistique? Le Système d'Identification au Répertoire des Unités Statistiques (SIRUS), enregistre les entreprises profilées décrites par l'application « Profilage », puis récupère, dans le répertoire administratif SIRENE, les unités légales qui n'appartiennent pas aux contours de ces entreprises profilées : ces dernières, qui sont donc indépendantes (non liées à une entreprise), sont donc des « entreprises ».

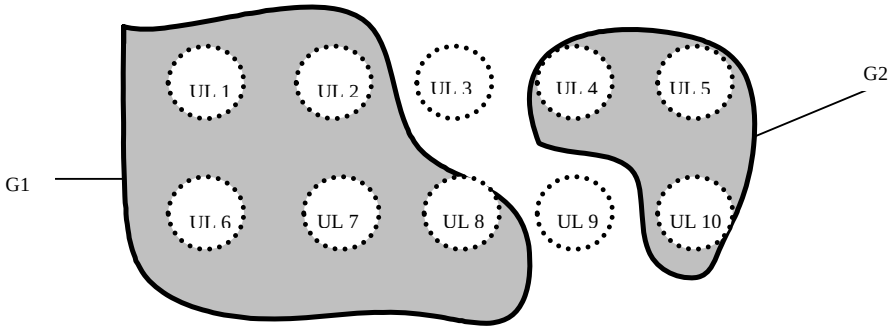
11. SIRUS ne se limitera pas aux unités « entreprises » puisqu'il enregistrera d'autres unités statistiques d'intérêt, d'où son nom « Système d'Identification au Répertoire des Unités Statistiques » : citons les groupes de sociétés (qui permettent de déterminer les

entreprises profilées), les unités locales des entreprises, les unités légales qui les composent, les réseaux[5] et tous les liens qui unissent ces différentes unités.

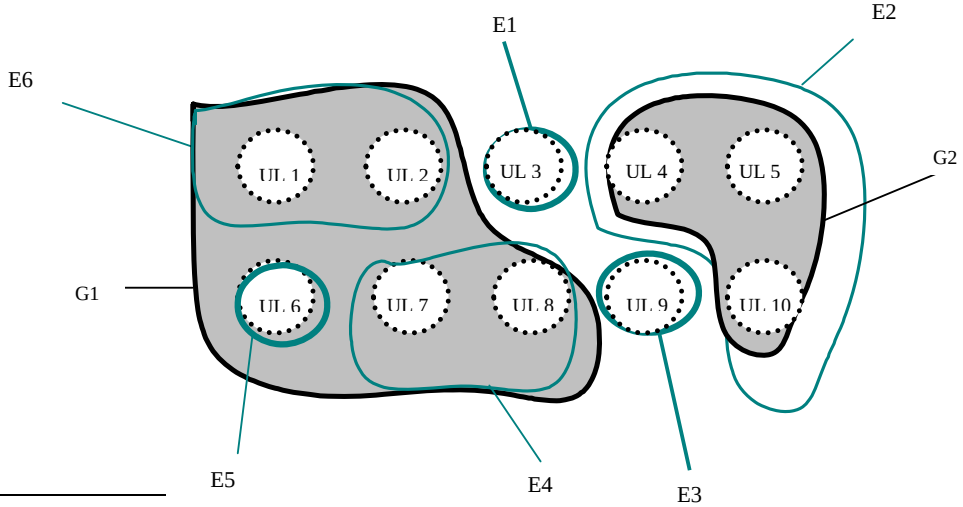
Prenons l'exemple d'une population de 10 unités légales (UL) de SIRENE :



Plusieurs d'entre elles sont détenues par des groupes de sociétés G1 (grand groupe) et G2 (petit groupe) :



Après avoir profilé les groupes, la population des 10 unités légales devient une population de 6 entreprises comme le décrit la figure ci-dessous (les 2 unités légales indépendantes, les 3 entreprises issues du profilage manuel du grand groupe et le petit groupe qui a été assimilé à une entreprise):



[5] ensemble de points de vente (unités légales ou établissements) liés à une tête de réseau, qui entretiennent des relations durables en établissant une communauté d'intérêts.

2. Un nouveau mode d'enregistrement des données

12. SIRUS enregistrera les entreprises (profilées ou non profilées) mais évidemment aussi certaines caractéristiques de ces entreprises comme la taille, le classement sectoriel, l'effectif salarié, le chiffre d'affaires, entre autres. Pour mener à bien cette mission, il centralisera les demandes de mises à jour émises par des applications ou enquêtes dites « partenaires » et acceptera ou refusera les demandes de mises à jour en fonction de règles de priorités entre les sources de mises à jour. Par exemple si, suite à une enquête, une unité apparait « inactive », l'information pourra être signalée à SIRUS qui cessera l'unité, sauf signalement contraire d'une autre enquête. Le répertoire jouera en quelque sorte un rôle « d'arbitre ».

13. Toute modification d'une valeur dans SIRUS donnera lieu à historisation, et SIRUS informera les applications « partenaires » des modifications concernant des unités de leur champ d'intérêt. Par ce circuit de mise à jour[6]/historisation/transmission des données, on assurera qu'une entreprise, à un instant donné et pour une période de référence donnée, se verra affecter à la même population de référence dans toutes les enquêtes et y aura les mêmes caractéristiques[7]. De plus, cette valeur unique et de référence pour certaines statistiques permettra une meilleure cohérence globale du système : par exemple, pour la diffusion des résultats, le classement sectoriel d'une entreprise (APE) sera identique quelle que soit la source à l'origine de la publication.

14. Toutes les caractéristiques de SIRUS seront datées (on connaitra donc la valeur qu'avait une variable à une date donnée) et tous les traitements seront datés (on connaitra donc la date à laquelle une variable a pris la valeur qu'elle a). Ces deux informations permettront de construire des bases de sondage de plusieurs natures.

15. Prenons le cas d'une unité, classée dans l'industrie en 2009, qui passe dans le secteur des transports en décembre 2010. Cette information n'est transmise à SIRUS qu'en juillet 2011. Le répertoire enregistrera donc :

Secteur	Date d'effet	Date de traitement de l'information dans SIRUS
Industrie	01.03.2009	01.03.2009
Transports	31.12.2010	15.07.2011

Si une enquête demande à SIRUS la création d'une base de sondage au 31.12.2010, alors :

* si la base est créée le 01.01.2011, l'entreprise sera classée « Industrie »

* si la base est créée le 01.08.2011, l'entreprise sera classée « Transports »

* si la base est créée le 01.08.2011, et que l'utilisateur demande à SIRUS l'information telle qu'elle était connue du répertoire en mars 2011, l'entreprise sera classée « Industrie ».

3. La prise en compte de nouvelles variables

16. SIRUS mettra également à disposition de tous des informations nouvelles et exclusivement accessibles par ce répertoire. Par exemple, en décembre 2008, un décret

6 cessation d'une unité, changement de secteur, embauche de salariés par exemple.

7 aujourd'hui, certaines applications utilisent des « photos » du répertoire SIRENE, produites à une date t, sans prendre en compte les modifications enregistrées par SIRENE après t, même si elles concernent un événement antérieur à t.

d'application[8] de la loi française de modernisation de l'économie a défini des catégories d'entreprises, en s'appuyant sur le concept d'entreprise tel qu'il est défini par le règlement européen sur les unités statistiques : microentreprises, petites et moyennes entreprises, entreprises de taille intermédiaire, grandes entreprises. Cette information relative à la catégorie est calculée dans le répertoire statistique.

17. Autre information nouvelle : afin de faciliter la communication avec les entreprises, SIRUS enregistrera, en particulier pour les entreprises profilées, un correspondant unique (de l'INSEE). Ce sera la personne que contactera en priorité l'entreprise en cas de problème transversal affectant plusieurs enquêtes, ou pour être orientée vers le bon correspondant spécifique à une enquête.

18. En résumé, SIRUS sera l'outil idéal pour :

- définir les populations de référence à employer dans les enquêtes statistiques ;
- constituer des référentiels ou des bases de sondage – décrites par les maîtrises d'ouvrage –, que l'unité d'intérêt soit l'entreprise, le groupe ou, au niveau le plus fin, l'unité locale ;
- décrire les unités statistiques à partir des unités qui la composent.

19. Il sera le seul instrument permettant de mettre en œuvre à la fois le règlement « unités statistiques » 696/93 et le règlement « répertoires » 177/2008, ces deux règlements définissant les unités statistiques requises par les différents règlements européens (SBS, STS, Labour Cost, etc.).

20. Brièvement, quelques éléments de calendrier : tous ces travaux de création de SIRUS, de mises à jour par des applications « partenaires » sont actuellement en cours de développement. La mise en production de SIRUS et la création des premières « populations de référence » aura lieu en septembre 2011. A compter de cette date commenceront les travaux sur l'enregistrement des enquêtes.

III. Enregistrer la charge d'enquêtes

21. Le répertoire statistique SIRUS présenté, revenons-en à son rôle dans la réduction de la charge qui pèse sur les enquêtés. Avant même de chercher à diminuer cette charge, il faut l'enregistrer pour la mesurer.

22. En France, une entreprise donnée peut être interrogée par plusieurs directions de l'INSEE (enquêtes sur des données de salaires, sur des données économiques), voire par des services statistiques ministériels spécialisés (un pour les problématiques d'emploi, un pour les sujets de transport ou construction par exemple). Or actuellement, en l'absence de centralisation des échantillons d'enquêtes réalisées par une direction de l'INSEE ou un service ministériel, il est impossible d'évaluer, à faible coût et rapidement, le nombre de questionnaires reçus par une unité légale ou une entreprise donnée, le nombre d'enquêtes pour lesquelles celle-ci a été sollicitée et auxquelles elle a *in fine* répondu. Quant à quantifier le temps passé à répondre à ces enquêtes, c'est (presque) mission impossible.

23. Les choses vont changer lorsque SIRUS sera en production, puisque SIRUS enregistrera les échantillons de toutes les enquêtes auprès des entreprises conduites par le système statistique public. Pour cela, il travaillera en lien avec :

[8] décret d'application de la loi n°2008-776 de modernisation de l'économie du 4 août 2008.

- l'Unité méthodologie statistique - entreprises (UMS-E) : une grande partie des tirages d'échantillon des enquêtes-entreprises réalisées par l'INSEE ou les services statistiques ministériels est réalisée par l'UMS-E, un service spécialisé de l'INSEE. En amont du tirage d'échantillon, SIRUS livrera à l'UMS-E, des bases de sondage, à partir des spécifications qui lui auront été transmises (champ, type d'unités enquêtées). En retour, après le tirage de l'échantillon, l'UMS-E livrera à SIRUS la base de sondage complétée d'informations de sondage (strate d'échantillonnage pour toutes les unités, poids des unités appartenant à l'échantillon). SIRUS saura ainsi si telle unité a été interrogée, ou si elle était dans la base de sondage, tout en disposant de données de stratification ;
- les autres services : SIRUS pourra en tant que de besoin livrer des bases de sondage et sera destinataire, pour enregistrement, de tous les échantillons des autres enquêtes menées par le système statistique d'entreprises.

24. On saura donc si une unité a été échantillonnée, quand, par quel service, pour quelles enquêtes et quelles années, si elle était dans telle ou telle base de sondage. On le saura pour tous les types d'unités, que ce soit une entreprise profilée, un groupe, une unité légale, une unité locale. On s'efforcera d'ailleurs de récupérer les échantillons des enquêtes passées afin de constituer un historique de cette charge d'enquête.

25. A l'avenir, pour s'assurer que SIRUS est exhaustif dans l'enregistrement des échantillons, le répertoire se tiendra informé des enquêtes menées. Pour cela, il se rapprochera du Conseil National de l'Information Statistique (CNIS), instance de dialogue entre les statisticiens et les représentants de la société civile (syndicats professionnels, syndicats de salariés, chercheurs, etc.) devant laquelle l'INSEE et les services statistiques ministériels présentent annuellement leur programme de travail. C'est le CNIS qui donne son accord pour la réalisation des enquêtes statistiques. Il dispose donc pour une année donnée de l'ensemble des enquêtes réalisé par la statistique publique française.

IV. Mesurer la charge de réponse

26. Le répertoire SIRUS envisage d'aller encore un peu plus loin dans sa démarche et d'enregistrer la charge de réponse aux enquêtes, c'est-à-dire le temps passé par l'entreprise répondante à remplir le questionnaire. Certaines enquêtes posent déjà systématiquement une question sur la charge de réponse (cf. papier CES/2009/35). De plus, une des mesures de simplification administrative présentées récemment par le gouvernement français à destination des entreprises stipule que figure, à partir de janvier 2012, une question sur le temps de réponse dans chaque enquête statistique. Cette information sera récupérée dans SIRUS ; à défaut de réponse de la part de l'entreprise à cette question, c'est le fait que l'unité a répondu ou non à l'enquête qui sera enregistrée.

27. Là encore, une coordination sera nécessaire avec tous les services enquêteurs, afin de récupérer des données post-enquêtes (liste des unités répondantes, temps de réponse), y compris lorsque les échantillons auront été établis sans faire appel à SIRUS. L'idéal serait cependant, à terme, de réduire au maximum cette dernière catégorie d'enquête afin d'optimiser au mieux la charge d'enquête auprès des entreprises.

28. La structure du répertoire permettra d'enregistrer la charge pour les unités légales comme pour les entreprises : la plus intéressante sera cependant la charge mesurée au niveau entreprise (si des enquêtes sont faites au niveau unité légale, on sommera ces données afin d'obtenir l'information au niveau entreprise).

29. Au final dans l'idéal, les deux informations suivantes seront mobilisables pour chaque entreprise de la base :

- l'entreprise a été interrogée X fois au cours de l'année écoulée (il sera en outre possible de lister les enquêtes auxquelles elle a participé et donc le nombre de questionnaires reçus) ;
- le temps consacré par l'entreprise pour répondre aux enquêtes statistiques est de XX h au cours de l'année écoulée.

V. Réduire la charge d'enquêtes

30. Si mesurer la charge statistique pesant sur les entreprises est intéressant, s'efforcer de la réduire l'est encore plus. Comment SIRUS pourra-t-il contribuer ?

A. Interroger au niveau de l'entreprise, qu'il s'agisse d'un groupe ou d'un sous-groupe, entraînera de facto une baisse de la charge pour les entreprises (profilées).

31. En effet, connaissant le nombre de questionnaires d'enquêtes envoyés à toutes les unités légales en général, et à celles appartenant à une entreprise profilée en particulier, on pourra par exemple mesurer les gains réalisables sur cette entreprise, à qui un seul questionnaire d'enquête aura pu être envoyé grâce au profilage. C'est d'ailleurs un élément dont disposent les personnes responsables du Profilage des grands groupes à l'heure actuelle pour convaincre les groupes de leur intérêt à participer au profilage.

32. Illustrons ce propos avec les retours obtenus sur les premiers groupes profilés : 30 groupes (plutôt les plus gros) regroupent 8 900 sociétés dans le monde dont 3 400 unités légales en France. Après concertation avec ces groupes, on devrait retenir 95 entreprises (de 1 à 7 entreprises selon les groupes, 3 en moyenne). Ainsi, lorsque ce sera possible, l'interrogation au niveau entreprise permettra de n'envoyer que 95 questionnaires pour mobiliser l'information qui serait obtenue aujourd'hui par l'envoi des 3 400 questionnaires.

33. Un problème subsiste : des tests actuels auprès de groupes montrent qu'il ne sera pas toujours possible d'interroger l'entreprise pour obtenir certaines informations.

34. En outre, les procédures légales de collecte (obligation légale de répondre) existent aujourd'hui au niveau des unités légales, mais il faut étudier comment le dispositif peut s'adapter à une interrogation des entreprises.

35. Ainsi, certaines enquêtes seront probablement toujours menées au niveau unité légale et dans ce cas (comme évoqué ci-dessus), la charge au niveau de l'entreprise sera obtenue en sommant la charge observée sur chaque unité légale du contour de l'entreprise.

B. Le Système d'Identification au Répertoire des Unités Statistiques permettra également de mettre en œuvre la coordination négative

36. Au début des réflexions sur SIRUS, il avait été imaginé que le répertoire statistique soit en mesure de constituer des bases de sondage à partir des spécifications de champ fournies par le concepteur de l'enquête mais également de tirer les échantillons d'enquêtes grâce à un outil dédié. Cette ambition a été abandonnée mais le fait que SIRUS contienne, outre un enregistrement des échantillons, des éléments sur la constitution des bases de sondage et sur la charge de réponse reste un atout.

37. Pour réduire la charge d'enquêtes, une méthode actuellement utilisée à l'INSEE pour les enquêtes-entreprise consiste à faire de la coordination négative d'échantillons et à favoriser, lors du tirage d'un échantillon, la sélection d'entreprises n'ayant pas déjà été

sélectionnées lors d'enquêtes récentes, tout en conservant le caractère sans biais des échantillons. Cette méthode, décrite dans l'encadré « la coordination d'échantillons deux à deux », est basée sur l'affectation à chaque unité d'un numéro aléatoire. Malheureusement, elle ne permet de disposer d'échantillons d'entreprises coordonnés négativement que deux à deux.

38. Or des travaux de recherche, menés à l'INSEE il y a quelques années ont conduit à proposer une méthode différente, actuellement en cours de test, qui devrait permettre de faire de la coordination négative de plusieurs enquêtes.

39. Elle repose sur l'utilisation de la charge d'enquête. Cette charge évolue bien sûr à chaque tirage d'échantillon, puisqu'elle dépend des enquêtes précédentes auxquelles a participé l'entreprise : chaque tirage consistera à sélectionner, dans chaque strate, les unités ayant les plus petits cumuls des charges.

40. Cette méthode a été décrite par Christian Hesse (cf . Généralisation des tirages à numéros aléatoires permanents, ou la méthode JALES+) et est actuellement expérimentée à l'INSEE, l'objectif étant la mise au point d'un logiciel. Dans ce cadre, le répertoire statistique SIRUS sera mobilisé pour fournir au logiciel, pour chaque unité, les informations sur la charge d'enquête puisque lui seul centralisera tous les échantillons.

41. Or, on a dit que SIRUS centraliserait également les charges de réponse : cette information pourrait être mobilisée dans la méthode en cours de développement car elle permettrait de différencier les enquêtes, de ne pas donner à toutes le même « poids ». Dit autrement, une entreprise qui répond à un questionnaire de dix pages qui nécessite en moyenne 3 h de remplissage aurait moins de chances d'être tirée dans les enquêtes futures qu'une unité qui a répondu à une enquête d'une demi-page dont le temps de réponse moyen est estimé à 10 minutes.

42. En conclusion, on peut donner quelques éléments de calendrier : le répertoire mis en place à l'automne 2011 sera progressivement enrichi par les échantillons des années passées, en particulier ceux de 2011, voire 2010. A l'été 2012, il devrait donc être possible de disposer de premiers éléments sur la charge d'enquêtes et de réponse.

43. Le répertoire statistique français SIRUS permettra donc, grâce à la prise en compte de la dimension « entreprise » et au calcul effectif de la charge statistique qui pèse sur ces unités, de mettre en œuvre de nouvelles méthodes pour la coordination d'échantillon, ce qui devrait réduire la charge d'enquêtes statistiques qui pèse sur les entreprises.

Encadré : la coordination d'échantillons deux à deux

44. Dans le cas de plans de sondage aléatoires simples stratifiés, l'algorithme de tirage d'échantillon généralement utilisé à l'INSEE consiste à attribuer à chaque unité de la base de sondage d'une première enquête un numéro aléatoire compris entre 0 et 1, puis à sélectionner les unités possédant les plus petits numéros aléatoires de chaque strate. A l'issue du tirage, une permutation de ces numéros aléatoires est opérée afin que les unités sélectionnées dans l'échantillon qui avaient le plus petit numéro aléatoire se retrouvent avec les numéros aléatoires les plus grands et vice versa (les unités non échantillonnées se retrouvent avec des numéros aléatoires plus petits), comme l'illustre le schéma ci-dessous.

45. Prenons l'exemple d'une enquête qui contiendrait deux strates (Industrie avec un taux de sondage à 50 %) et transport avec un taux de sondage à 10 %). Les unités sont classées dans leur strate par numéro aléatoire croissant. Les unités grisées sont celles qui sont sélectionnées.

A	**B**	**C**	**D**	**E**	F	G	H	I	J

Secteur de l'industrie (taux : 50 %)

K	L	M	N	O	P	Q	R	S	T

Secteur des transports (taux : 10 %)

A l'issue du tirage, la permutation des numéros aléatoires permet d'aboutir à la situation suivante :

F	G	H	I	J	**A**	**B**	**C**	**D**	**E**	

Secteur de l'industrie (taux : 50 %)

L	M	N	O	P	Q	R	S	T	**K**

Secteur des transports (taux : 10 %)

46. Pour le tirage de l'échantillon suivant, après ajout d'un numéro aléatoire pour les unités absentes de la première base de sondage (les unités U et V dans le schéma ci-dessous), on sélectionne les unités ayant les plus petits numéros aléatoires dans chaque nouvelle strate de la seconde base de sondage, comme le montre le schéma suivant.

47. Imaginons qu'à l'issue de la première enquête soit tirée une seconde enquête, dont le taux de sondage serait de 60 % et qui ne contiendrait qu'une strate. Le résultat serait le suivant

F	**L**	**G**	**M**	**N**	**H**	**I**	**U**	**O**	**J**	**A**	**P**	**B**	Q	C	R	V	S	D	T	E	K

Secteur des transports et de l'industrie (taux : 60 %)

Donc seuls A et B reçoivent deux questionnaires : on observe un léger recouvrement mais on minimise le nombre d'unités recevant deux questionnaires.

VI. Références :

"French methods and practices regarding the statistical burden", papier présenté par la France, conférence des DGINS 2008, Vilnius

Depoutot R, "Reengineering French structural business statistics: an overview", Conference Q10, Helsinki

INSEE [2009] "Redesigning French structural business statistics: how can the response burden on companies be lessened", UN ECE, paper CES/2009/35

INSEE [2007] "Improving the quality of Business Registers", UN ECE, paper ECE/CES/GE 42/2007/3

Hesse Chr. : "Généralisation des tirages à numéros aléatoires permanents, ou la méthode JALES+", INSEE, Direction des Statistiques d'Entreprises.

Laflamme, G., Rivière, P. (1997): "OCEAN, toward a general system for sample collection and coordination", Courrier des Statistiques, English version n°3.

Hecquet V., 2010 : "Le profilage et son impact sur la représentation de l'appareil productif, papier présenté au colloque de l'Association de comptabilité nationale.