

PROCÉDURES POUR TENIR COMPTE DES NOUVELLES UNITÉS DANS LES ENQUÊTES AUPRÈS DES ENTREPRISES¹

Carol S. King et Robert E. Struble²

RÉSUMÉ

L'Economic Directorate du United States Bureau of the Census réalise diverses enquêtes conçues pour produire des estimations pour plusieurs secteurs économiques, dont l'exploitation minière, la fabrication, le commerce de gros et de détail et les services. La plupart de ces programmes s'appuient sur des échantillons probabilistes pour représenter la population visée. Les populations réelles sont en évolution constante: de nouvelles entreprises apparaissent, d'autres disparaissent, des modifications organisationnelles ont lieu et les classifications subissent des modifications. Pour nous assurer que nos échantillons demeurent représentatifs de ces populations changeantes, nous avons élaboré et mis en œuvre des méthodes de mise à jour. Les méthodes suivies pour tenir compte des nouvelles unités et des unités entrantes sont un élément important de cette mise à jour. Le présent article décrit et compare les méthodes de mise à jour des échantillons d'entreprises tenant compte des nouvelles unités et des unités entrantes mises en œuvre à l'heure actuelle pour les divers secteurs.

MOTS CLÉS: Mise à jour des échantillons; nouvelles unités; unités entrantes.

1. INTRODUCTION

L'identification des entreprises qui entrent ou sont créées dans les secteurs du commerce de détail, du commerce de gros, des services et de la fabrication repose en grande partie sur des données administratives transmises par d'autres organismes du gouvernement des États-Unis, notamment l'Internal Revenue Service (IRS), la Social Security Administration (SSA) et le Bureau of Labor Statistics (BLS). Le respect du caractère confidentiel de ces renseignements et l'autorisation de les utiliser sont strictement réglementés (United States Code). Les données administratives que transmettent ces organismes comprennent la raison sociale et l'adresse de l'entreprise, le code de classification de la branche d'activité, la paye trimestrielle et annuelle, l'effectif, le montant annuel des ventes ou des recettes, et l'affiliation de l'entreprise. Ces données sont celles sur lesquelles se fondent la création et la tenue à jour du Registre des entreprises (RE) du Census Bureau. Le RE contient la liste de tous les établissements avec salariés recensés aux États-Unis.

Les nouvelles entreprises qui emploient des salariés doivent présenter une demande de numéro fédéral d'identification d'employeur (Federal Employer Identification Number / EIN) à l'IRS. Les entreprises avec salariés utilisent l'EIN attribué par l'IRS comme numéro d'identification du contribuable lorsqu'elles transmettent les données sur la paye et sur l'effectif. Conformément aux dispositions légales, l'IRS transmet régulièrement au Census Bureau certaines données extraites des déclarations de revenus en vue de la production de statistiques officielles. Dans le RE, l'EIN est utilisé comme identificateur principal des entreprises ne comptant qu'un seul établissement et comme identificateur secondaire des entreprises comptant plusieurs établissements. Cette information est utilisée pour repérer les entreprises nouvellement créées dans les secteurs du commerce de détail, du commerce de gros et des services, ainsi que les entreprises entrantes dans le secteur de la fabrication.

¹ Cet article présente les résultats de travaux de recherche et d'analyse entrepris par les employés du Census Bureau. Il a fait l'objet d'un examen de portée plus limitée que celui auquel sont soumis les publications officielles. Le présent rapport est diffusé en vue de tenir les parties intéressées au courant des travaux de recherche en cours et d'en favoriser la discussion.

² United States Census Bureau, Washington, D.C., États-Unis, 20233

Dans le présent article, l'expression «entreprises entrantes» s'entend des nouvelles entreprises créées dans le secteur étudié, ainsi que des entreprises existantes qui entrent dans le secteur à la suite d'une modification de la classification des branches d'activité. Les procédures décrites pour les secteurs du commerce de détail, du commerce de gros et des services (section2) ont trait exclusivement aux nouvelles entreprises, tandis que celles décrites pour le secteur de la fabrication (section3) couvrent à la fois les nouvelles entreprises et certains changements de classification.

2. PROCÉDURES POUR TENIR COMPTE DES NOUVELLES UNITÉS DANS LES SECTEURS DU COMMERCE DE DÉTAIL, DU COMMERCE DE GROS ET DES SERVICES

2.1 Échantillons de départ

Tous les cinq ans environ, le plus tôt possible après le recensement économique le plus récent, de nouveaux échantillons sont sélectionnés pour les enquêtes sur les secteurs du commerce de détail, du commerce de gros et des services. Les données recueillies dans le cadre de ces enquêtes sont utilisées pour estimer les totaux et les tendances au niveau national pour diverses classes d'entreprises, définies dans le Système de classification des industries de l'Amérique du Nord (SCIAN) de 1999. Les données recueillies sur les secteurs du commerce de détail et de gros incluent le chiffre des ventes, la valeur des stocks à la fin de l'exercice et la valeur des achats. L'enquête sur le secteur des services fournit des données sur les recettes, ainsi que des données propres à certaines branches d'activité définies dans le SCIAN.

Pour ces enquêtes, la base de sondage de départ regroupant les entreprises avec salariés comprend deux catégories d'unités d'échantillonnage, à savoir les sociétés et les numéros fédéraux d'identification d'employeur ou EIN. Les sociétés sont des entreprises comptant plusieurs établissements dont le chiffre de ventes annuel excède un seuil spécifique pour leur branche d'activité, tandis que les EIN sont des EIN d'entreprises comptant plusieurs établissements non échantillonnées en tant que sociétés ou des EIN d'entreprises comptant un seul établissement. Tous les EIN sélectionnés dans les échantillons de départ, y compris ceux associés à des sociétés confirmées, portent dans le RE un code dénotant leur situation d'échantillonnage: sélectionné pour l'enquête sur le commerce de détail, le commerce de gros ou les services; soumis à l'échantillonnage et non sélectionné ou hors du champ de l'enquête, à cause de la catégorie d'activité (p.ex., fabrication) ou de l'emplacement géographique (p.ex., Canada).

2.2 Dépistage des nouvelles unités

S'assurer que les échantillons d'enquête représentent, dans la mesure du possible, la population étudiée durant la période de référence pour laquelle sont calculées les estimations fait partie intégrante du maintien de la qualité des estimations produites d'après les données d'enquête. Par conséquent, il est nécessaire de mettre à jour les échantillons de départ dans les plus brefs délais lorsque de nouvelles entreprises ouvrent leurs portes après la sélection de l'échantillon. Pour les enquêtes mensuelles et annuelles sur le commerce de détail, le commerce de gros et les services, toutes les nouvelles entreprises ajoutées à l'échantillon de départ le sont sous forme d'unités d'échantillonnage (EIN). Nous appelons les EIN associés à ces nouvelles entreprises des nouveaux EIN, ou nouvelles unités.

Notre objectif est de repérer les nouvelles unités et de les représenter dans les échantillons aussitôt que possible après le démarrage des opérations. En vue d'atteindre cet objectif, trimestriellement, nous extrayons les nouvelles unités ajoutées au RE et nous les soumettons à l'échantillonnage. Rappelons que les unités dépistées sont les nouveaux EIN et qu'une nouvelle unité n'est pas nécessairement une nouvelle entreprise; il pourrait s'agir d'une entreprise existante qui, pour une certaine raison (p.ex., réorganisation), a acquis un nouvel EIN. Bien qu'un nouvel EIN puisse correspondre à une entreprise comptant plusieurs établissements, la plupart sont associés à des entreprises ne comptant qu'un seul établissement. Quoi qu'il en soit, habituellement, lorsque nous parlons de l'EIN, nous entendons l'unité commerciale particulière à laquelle il est associé. Cette signification est celle que nous attribuerons à l'expression «nouvelle unité» dans la suite de l'exposé.

Les critères d'extraction des nouvelles unités sont les suivants:

1. n'est pas considérée comme ayant été soumise à l'échantillonnage ou comme étant hors du champ de l'enquête;
2. est considérée comme faisant partie du champ de l'enquête sur le commerce de détail, le commerce de gros ou les services;
3. si elle n'est pas catégorisée, la nouvelle unité:
 - 1) compte plus de 3 employés ou
 - 2) a déclaré une paye non nulle pour au moins un des quatre derniers trimestres;
4. est une unité active dans le fichier maître des entreprises (Business Master File/BMF), autrement dit est tenue de produire la formule 941 de l'IRS, Employer's Quarterly Federal Tax Return. L'IRS envoie une formule 941 par la poste à toute entreprise qui a des employés et l'entreprise utilise l'EIN pour produire le rapport trimestriel;
5. possède une adresse postale valide.

2.3 Première phase du processus d'échantillonnage

Les nouvelles unités extraites du RE sont soumises à une première phase d'échantillonnage qui consiste, pour commencer, à repérer les nouvelles unités associées aux unités déjà échantillonnées et à les éliminer des étapes suivantes du processus d'échantillonnage. Puis, deux coefficients de pondération sont attribués à chaque nouvelle unité retenue. L'un est établi d'après le montant trimestriel de la paye de l'unité et l'autre, d'après son effectif. Le coefficient de pondération appliqué en bout de ligne est celui qui donne à la nouvelle unité la probabilité de sélection la plus forte. Dans la plupart des cas, il s'agit du coefficient de pondération fondé sur la paie, car on dispose plus souvent de données sur celle-ci que sur l'effectif. Nous procédons ensuite à un échantillonnage systématique avec probabilités égales dans les branches d'activité définies conformément au SCIAN, par strate de taille. Les nouvelles unités non codées selon le SCIAN sont échantillonnées dans une branche d'activité non classifiée, par strate de taille. De 125000 à 170000 nouvelles unités sont recensées chaque trimestre. Environ 10% de nouvelles unités sont sélectionnées lors de la première phase d'échantillonnage.

Nous envoyons par la poste, aux nouvelles unités échantillonnées, un questionnaire appelé SQ-CLASS, Business and Professional Classification Report. Les questionnaires sont mis à la poste à Jeffersonville, Indiana et renvoyés là-bas dûment remplis aux fins de leur traitement administratif et de la saisie des données. L'enquête auprès des nouvelles unités sélectionnées à la première phase d'échantillonnage permet de recueillir des données sur deux mois de ventes (revenus), la catégorie d'activité, l'organisation de l'entreprise, la situation d'imposition ou d'exemption d'impôt, les stocks des établissements de gros et le commerce électronique. La participation à l'enquête n'est pas obligatoire et le taux de réponse est de 70% environ.

2.4 Deuxième phase du processus d'échantillonnage

Les nouvelles unités auxquelles est envoyé par la poste un questionnaire SQ-CLASS sont soumises à la deuxième phase du processus d'échantillonnage. Cette phase vise en partie à mettre à jour les données administratives sur les nouvelles unités depuis l'envoi par la poste du questionnaire de première phase. Ceci nous permet d'utiliser les données les plus récentes pour déterminer la catégorie d'activité et la taille. Au même titre que certaines nouvelles unités sont exclues de la première phase d'échantillonnage, certaines sont exclues de la deuxième phase également. Ces nouvelles unités sont celles pour lesquelles est établie une association avec des unités figurant déjà dans les échantillons, celles qui étaient non classifiées au moment de l'envoi du questionnaire par la poste et pour lesquelles on a déterminé qu'elles étaient hors du champ de l'enquête d'après les réponses au questionnaire SQ-CLASS, celles associées aux bureaux centraux administratifs ou celles associées à des entités gouvernementales.

D'après les renseignements recueillis au moyen du questionnaire SQ-CLASS, nous attribuons à chaque nouvelle unité un nouveau code du SCIAN ou un code plus précis, et une mesure de taille (MT). Nous calculons la mesure de taille, dont nous nous servons pour déterminer le poids de sondage de deuxième

phase appliqué à la nouvelle unité, d'après le chiffre de ventes pour les deux mois (revenus), s'il a été déclaré. Pour les nouvelles unités qui ne fournissent pas de données mensuelles, nous calculons la mesure de taille d'après la paye ou l'effectif.

Les nouvelles unités qui ne sont pas encore en exploitation ou qui n'ont pas fourni suffisamment de renseignements pour être échantillonnées sont retenues dans un registre appelé Registre des nouvelles unités non échantillonnées. Si la mise à jour des données administratives qui a lieu au trimestre suivant produit suffisamment de renseignements, la nouvelle unité qui figure dans le registre des nouvelles unités non échantillonnées peut être soumise à l'échantillonnage au cours du trimestre suivant. Sinon, ce trimestre-là, l'entreprise associée à la nouvelle unité reçoit par la poste un questionnaire SQ-CLASS, c'est-à-dire six mois après le premier envoi par la poste destiné à cette nouvelle unité. L'entreprise doit encore être en exploitation et avoir déclaré une paye récemment. L'envoi par la poste à intervalle de six mois se poursuivra jusqu'à l'obtention d'une réponse ou la collecte de suffisamment de données administratives. Lors d'un trimestre donné, il existe environ 15 000 nouvelles unités dans le registre des nouvelles unités non échantillonnées, la plupart figurant dans ce registre parce que les données concernant la classification selon le SCIAN ne suffisent pas pour qu'on puisse les échantillonner.

L'échantillonnage de deuxième phase a lieu trois mois après l'échantillonnage de première phase. Environ 20% des unités sélectionnées lors de la première phase sont sélectionnées lors de l'échantillonnage de deuxième phase. Ce dernier est un échantillonnage systématique avec probabilité proportionnelle à la taille réalisé de trimestre en trimestre.

Les nouvelles unités sélectionnées lors de la deuxième phase sont ajoutées aux échantillons mensuels et annuels. Le délai pour l'ajout d'une nouvelle unité aux échantillons est de six à neuf mois, en moyenne. Sa longueur dépend de plusieurs facteurs, dont le temps écoulé entre le moment où la nouvelle unité est entrée en exploitation et la production de la formule SS-4 (Application for Employer Identification Number) de l'IRS et le temps nécessaire pour procéder aux envois par la poste pour les EIN sélectionnés lors de la première phase d'échantillonnage. Comme pour les EIN sélectionnés dans les échantillons de départ, chaque nouvelle unité reçoit un code d'état qui reflète les résultats des opérations d'échantillonnage.

2.5 Problèmes de qualité que pose le traitement des nouvelles unités

Il est généralement reconnu qu'il est important de réduire les effets des erreurs non dues à l'échantillonnage afin de produire des estimations de qualité. Le processus trimestriel de sélection de nouvelles unités vise à résoudre les problèmes de couverture qui affectent nos programmes économiques grâce à l'intégration de procédures permettant de repérer les nouvelles unités dans le champ de l'enquête, de procéder à un échantillonnage à deux phases pour déterminer correctement la catégorie d'activité et obtenir des renseignements exacts sur la taille, et de repérer les unités couvertes en double. Cependant, notre processus pose deux problèmes importants, l'un étant la rapidité d'exécution et l'autre, l'exactitude des codes précisant la catégorie d'activité de l'entreprise.

2.5.1 Rapidité d'exécution du processus de sélection des nouvelles unités

Idéalement, nous aimerions que toute nouvelle unité soit représentée dans nos échantillons aussitôt qu'elle entre en activité. Comme nous l'avons mentionné plus haut, le retard avec lequel les nouvelles unités sont intégrées dans les enquêtes est de six à neuf mois. Nous introduisons ces nouvelles unités dans les enquêtes mensuelles tous les trois mois. D'une façon réaliste, nous ne pourrions intégrer les nouvelles unités dans nos enquêtes plus rapidement que mensuellement. Cependant, plusieurs facteurs entrent en ligne de compte lorsque l'on essaye d'intégrer les nouvelles unités plus rapidement que trimestriellement. En voici la liste.

1. Réception des dossiers administratifs. Nous ne recevons pas toutes les données administratives utilisées pour repérer les nouvelles unités faisant partie du champ des enquêtes de façon uniforme durant l'année. Par conséquent, le nombre de nouvelles unités repérées trimestriellement est plus uniforme que le nombre qui le serait mensuellement. Ce problème d'uniformité aurait une incidence

sur le nombre de nouvelles unités intégrées mensuellement dans les échantillons et pourrait fort bien influencer nos estimations des variations mensuelles.

2. Traitement. À l'heure actuelle, nous donnons aux répondants 30 jours pour répondre au questionnaire SQ-CLASS. Nous pourrions réduire le délai de réponse à moins de 30 jours si 1) nous avons une bonne raison de le faire et que 2) nous expliquions cette bonne raison à l'Office of Management and Budget (OMB). L'OMB exige que les organismes décrivent dans la demande d'autorisation toute circonstance nécessitant une réponse dans un délai de moins de 30 jours après la réception d'un instrument de collecte de renseignements.
3. Effet de l'intégration des nouvelles unités dans les estimations mensuelles. Nous devrions mettre au point une méthode efficace d'intégration des nouvelles unités de sorte que celle-ci n'ait aucun effet artificiel sur les estimations de la variation mensuelle des ventes et des stocks. À l'heure actuelle, une méthode est en place pour le processus de traitement trimestriel.

2.5.2 Classification des entreprises

La classification des entreprises, si elle existe au moment du processus de repérage des nouvelles unités, peut provenir de diverses sources, notamment la SSA, l'IRS et le BLS.

- La SSA se fonde, pour coder les branches d'activité, sur la formule SS-4 de l'IRS dont elle reçoit un exemplaire pour pouvoir produire un code de classification d'entreprise. Cette formule contient des renseignements sur l'activité principale de l'entreprise, le nombre prévu d'employés, le genre d'organisation et la catégorie de clients.
- Trimestriellement, le Census Bureau transmet au BLS la liste des entreprises comptant un seul établissement inscrites dans le Registre des entreprises qui ne possèdent pas de code de classification, dont le code de classification est incomplet ou auquel on a attribué un code de classification au hasard. Le BLS apparie ces établissements avec les fichiers qui les concernent et transmet au Census Bureau tout code de classification d'entreprise qu'il possède.
- Un code d'activité principale de l'entreprise peut aussi être transmis par l'IRS dans les enregistrements du fichier maître des entreprises (BMF) que nous recevons.

Malgré ces sources de renseignements, environ 41% de nouvelles unités extraites lors d'un trimestre donné restent non classées. En outre, nombre des nouvelles unités classées ne possèdent qu'un code partiel, c'est-à-dire un code du SCIAN contenant moins que les six chiffres prévus. En outre, environ 20 % d'unités reprises sous un code donné du SCIAN au moment de l'envoi par la poste sont recodées et se voient octroyer un code du SCIAN entièrement différent (qui peut ou non figurer dans le champ des enquêtes sur le commerce de détail, le commerce de gros et les services). Environ 30% de codes concordent au niveau à cinq chiffres, qui est le nombre minimal de chiffres nécessaires pour l'échantillonnage.

2.6 Travaux futurs

Nous évaluons pour le moment la faisabilité d'un échantillonnage à une seule phase. Un tel processus nécessiterait la détermination exacte de la catégorie d'activité de l'entreprise, ainsi que l'estimation de la taille de l'entreprise d'après des données administratives uniquement. En outre, nous envisageons de procéder à l'intégration annuelle plutôt que trimestrielle des nouvelles unités dans nos échantillons mensuels et annuels.

Il a également été proposé d'étendre le processus de traitement des nouvelles unités de façon à:

- inclure les nouvelles unités considérées à l'heure actuelle comme étant hors du champ des enquêtes sur le commerce de détail, le commerce de gros et les services;

- augmenter la taille de l'échantillon de première phase pour certains secteurs/sous-secteurs afin de mettre l'accent sur des branches d'activité pour lesquelles on sait ou soupçonne qu'il existe des problèmes de classification (p.ex., commerce de gros par opposition à commerce de détail);
- augmenter le taux d'échantillonnage pour les petites unités;
- baisser le seuil de paye fixé pour déterminer si une nouvelle unité existe réellement.

Cette expansion présente les avantages qui suivent.

- Élimination de la nécessité de garder en réserve les EIN hors du champ des enquêtes. Durant la phase de dépistage du processus trimestriel de traitement des nouvelles unités, nous manquons certains EIN des secteurs du commerce de détail, du commerce de gros et des services qui ont été classifiés erronément comme étant hors du champ des enquêtes. Pour compenser, nous gardons dans les enquêtes les EIN hors champ qui avaient été considérés comme étant dans le champ lors de l'envoi par la poste de première phase, mais qui ont été codés plus tard comme des unités hors du champ des enquêtes d'après les renseignements figurant sur le questionnaire rempli par les nouvelles unités.
- Amélioration de la classification du genre d'activité des nouvelles unités.
- Obtention de renseignements sur la répartition des nouvelles unités conformément au SCIAN.
- Utilisation des codes du SCIAN pour évaluer la qualité des codes administratifs de branche d'activité.
- Collecte de renseignements auxiliaires, comme l'affiliation de l'entreprise et les lieux d'exploitation, qui amélioreraient la couverture des entreprises comptant plusieurs établissements par le Registre des entreprises.

3. UNITÉS ENTRANTES DANS LE SECTEUR DE LA FABRICATION

3.1 Caractéristiques de l'enquête sur le secteur de la fabrication

Le Census Bureau utilise diverses enquêtes pour étudier le secteur de la fabrication. La plupart de ces enquêtes ont pour unité d'échantillonnage et de collecte des données l'établissement individuel (plutôt que la société). Autrement dit, si une société est en exploitation à plusieurs emplacements (un établissement par emplacement), alors chaque emplacement est considéré comme distinct pour ce qui est de l'échantillonnage et de la collecte de données. Par conséquent, les méthodes visant à repérer les nouvelles unités dans le secteur de la fabrication doivent, elles aussi, être appliquées au niveau de l'établissement.

La fréquence des enquêtes sur le secteur de la fabrication est un autre facteur important de l'élaboration du processus de dépistage des nouvelles unités manufacturières. Le Census of Manufactures est l'enquête fondamentale de collecte des données sur le secteur de la fabrication. Ce recensement, qui est réalisé tous les cinq ans, sert de base de sondage pour presque toutes les autres enquêtes sur le secteur de la fabrication. La Manufacturing Energy Consumption Survey, réalisé tous les trois à quatre ans, s'appuie sur les données du recensement, mises à jour pour tenir compte des nouvelles unités, des unités disparues et d'autres changements, comme base de sondage et source de données sur la taille des entreprises. L'Annual Survey of Manufactures (ASM) est réalisé annuellement. Un nouvel échantillon est sélectionné tous les cinq ans d'après des renseignements provenant du recensement le plus récent. De nouveau, nous nous servons du recensement pour créer une base de sondage et obtenir des données sur la taille des unités. L'échantillon de l'ASM est mis à jour annuellement afin de refléter les nouvelles unités et d'autres changements décelés lors des opérations de mise à jour des échantillons. Plusieurs autres enquêtes sur le secteur de la fabrication (Survey of Plant Capacity, Pollution Abatement Costs and Expenditures Survey, et nombre d'enquêtes du programme des Current industrial Reports) sont également réalisées annuellement en se fondant sur des renseignements provenant du recensement des manufactures le plus récent. Les nouvelles unités sont également intégrées annuellement dans les échantillons de ces enquêtes.

Puisque la plupart des enquêtes sur le secteur de la fabrication se fondent sur des renseignements provenant du recensement quinquennal, mais sont réalisées à intervalles plus fréquents, leur échantillon doit être mis à jour régulièrement afin que les données restent représentatives de la population en évolution. En outre, puisque la majorité de ces enquêtes sont réalisées annuellement, les mises à jour requises doivent être établies annuellement afin de pouvoir mettre les panels à jour avant l'envoi par la poste des questionnaires

du cycle d'enquête suivant. Les procédures permettant de repérer les unités entrantes du secteur de la fabrication ont pour but de produire les listes d'unités manufacturières entrantes qu'il faut intégrer dans les bases de sondage des enquêtes annuelles sur le secteur de la fabrication.

3.2 Sources et catégories d'unités entrantes

Les sources de renseignements sur les unités entrantes du secteur de la fabrication sont multiples. Certaines unités entrantes sont repérées d'après les données des enquêtes existantes du Census Bureau, tandis que d'autres le sont grâce aux données administratives transmises par l'IRS, la SSA et le BLS. Puisqu'il constitue le dépôt des données administratives transmises par ces sources externes, ainsi que des données provenant des programmes économiques du Census Bureau, le Registre des entreprises (RE) devient la source directe de dépistage des unités entrantes du secteur de la fabrication. Le cycle de mise à jour du RE, qui est annuel, coïncide avec le cycle de traitement des données de la Company Organization Survey réalisée par le Bureau. Chaque année, au moment de la réinitialisation du RE pour tenir compte de la nouvelle année de traitement, la version de l'année précédente est sauvegardée dans un fichier d'archive. L'existence de versions pluriannuelles du RE facilite le dépistage des unités entrantes dans le secteur de la fabrication.

Comme nous le mentionnons dans l'introduction, les unités entrantes du secteur de la fabrication englobent les nouveaux établissements de fabrication (nouvelles unités) et les établissements existants qui entrent dans le secteur de la fabrication à cause de certaines modifications apportées aux classifications. Plus précisément, les trois catégories d'unités entrantes du secteur de la fabrication que nous examinerons dans le présent article sont les nouveaux établissements de fabrication des entreprises comptant plusieurs établissements, les entreprises de fabrication ne comptant qu'un seul établissement dont l'exploitation débute durant l'année de référence (nouvelles unités) et les entreprises de fabrication ne comptant qu'un seul établissement qui étaient en exploitation l'année précédente, mais qui n'avait pas été codées selon la classification des industries cette année-là. Des procédures distinctes sont appliquées pour repérer chacune de ces catégories d'unités entrantes. Les établissements dont la classification par branche d'activité est modifiée de sorte qu'ils passent d'un secteur non manufacturier (année précédente) au secteur de la fabrication (année courante) ne sont pas étudiés dans le présent article.

3.2.1 Nouvelles unités - Entreprises comptant plusieurs établissements

Dans le cas des entreprises comptant plusieurs établissements, la plupart des nouvelles unités sont repérées grâce à la Company Organization Survey (COS). Cette enquête, réalisée annuellement par le Census Bureau, est conçue pour obtenir une liste exacte et à jour des établissements qui forment les entreprises comptant plusieurs établissements sélectionnés. Le questionnaire de la COS est envoyé par la poste à toutes les grandes et moyennes entreprises de fabrication comptant plusieurs établissements (entreprises employant au moins 50 personnes) ainsi qu'à un échantillon d'entreprises de plus petite taille. Environ 90 % des établissements de fabrication faisant partie d'entreprises à plusieurs établissements sont recensés annuellement grâce à la COS. Les entreprises qui reçoivent le questionnaire de la COS ont pour instruction d'inscrire sur celui-ci tous les établissements qui n'y sont pas encore énumérés, de décrire l'activité principale de ces établissements et d'indiquer s'il s'agit d'établissements « nouveaux » ou acquis. Un code de la Classification type des industries est attribué aux « nouveaux » établissements d'après la description fournie de leurs activités. Tous les « nouveaux » établissements classés dans le secteur de la fabrication sont ajoutés à l'univers de la fabrication et reçoivent automatiquement par la poste un questionnaire de l'ASM. Grâce à cette méthode, les nouvelles unités correspondant à des entreprises comptant plusieurs établissements sont intégrées dans la base de sondage de l'ASM l'année où elles sont repérées. Environ 3 000 nouvelles unités correspondant à des entreprises comptant plusieurs établissements sont décelées annuellement. Le Bureau produit également la liste de ces nouvelles unités afin qu'elles puissent être intégrées à d'autres enquêtes sur le secteur de la fabrication.

3.2.2 Nouvelles unités - Entreprises comptant un seul établissement

Le Census Bureau repère annuellement les nouvelles entreprises comptant un seul établissement par comparaison du RE de l'année courante (contenant les mises à jour transmises par l'IRS et la SSA) et du RE

de l'année précédente. Toutes les entreprises de fabrication ne comptant qu'un seul établissement sont repérées dans le RE courant (d'après les codes de secteur et d'activité), puis le fichier contenant les entreprises ainsi repérées est apparié à un fichier des entreprises de fabrication ne comptant qu'un seul établissement figurant dans le RE de l'année précédente. Les enregistrements non appariés du fichier de l'année courante sont alors examinés afin d'éliminer les cas de nouvelles entreprises du secteur de la fabrication qui résultent d'une modification de la classification des branches d'activité. Les établissements retenus sont considérés comme des nouvelles entreprises de fabrication ne contenant qu'un seul établissement.

À ce stade, les nouvelles unités sont réparties en trois fichiers d'après la taille de l'établissement. Le Bureau utilise les données sur la paye transmises par l'IRS pour distinguer les nouvelles unités «très grandes » ou «très petites » des autres nouvelles unités. Celles considérées comme « très grandes » sont éliminées du processus de traitement des nouvelles unités, car, en général, il s'agit simplement de nouveaux EIN produits pour des entreprises existantes comptant plusieurs établissements plutôt que de vraies nouvelles entreprises comptant un seul établissement. Les fichiers transmis par l'IRS et la SSA ne font pas la distinction entre les nouveaux EIN attribués à des entreprises comptant plusieurs établissements et la création de nouvelles entreprises comptant un seul établissement. Le Census Bureau s'efforce de relier les nouveaux EIN aux entreprises à plusieurs établissements existantes, mais, souvent, ces couplages ne sont pas repérés immédiatement.

À l'autre extrémité de l'échelle de taille, les nouvelles unités «très petites » repérées d'après les seuils de paye qui varient selon la branche d'activité. Ordinairement, les petites entreprises ne comptant qu'un seul établissement sont exclues du volet postal des enquêtes sur le secteur de la fabrication. Par exemple, dans le cas de l'ASM, la partie du panel qui ne participe pas à l'enquête postale comprend 166 000 des 366 000 établissements de fabrication. Bien qu'il s'agisse d'un pourcentage élevé de la population d'établissements, la partie du panel qui ne participe pas à l'enquête par la poste représente seulement 2 % du total des livraisons du secteur de la fabrication. Les données sur cette partie du panel sont estimées d'après les renseignements transmis par l'IRS et la SSA. Ces estimations sont incluses dans les estimations publiées, calculées d'après l'ASM. Les nouvelles entreprises ne comptant qu'un seul établissement dont le chiffre de paye est inférieur au seuil établi (environ 4 000 par année) sont incluses dans le volet non postal des estimations fondées sur l'ASM. Pour d'autres enquêtes sur le secteur de la fabrication (Survey of Plant Capacity, Pollution Abatement Costs and Expenditures Survey et les enquêtes du programme des Current Industrial Reports), ces petites unités nouvelles ne sont pas ajoutées aux panels, puisque ces enquêtes couvrent uniquement les établissements de fabrication dont le chiffre de paye est supérieur au seuil établi pour la participation à l'enquête par la poste.

Les autres nouvelles entreprises ne comptant qu'un seul établissement (c'est-à-dire celles qui ne sont ni «très grandes» ni « très petites ») sont ajoutées annuellement aux bases de sondage de l'ASM et des autres enquêtes sur le secteur de la fabrication. Ces nouvelles unités sont au nombre d'environ 6 000. Dans le cas de l'ASM, le Census Bureau s'est efforcé, vers le milieu des années 1990, d'apporter une correction pour les nouvelles entreprises ne comptant qu'un seul établissement classifiées erronément comme étant des entreprises non manufacturières lors de leur enregistrement dans le RE. L'application de facteurs de rajustement des coefficients de pondération est la méthode qui a été choisie. Souvent, le code de classification selon la branche d'activité attribué au départ aux nouvelles entreprises ne comptant qu'un seul établissement ajoutées dans le RE ne représente pas l'activité réelle de ces entreprises. En général, ces codes incorrects ne sont corrigés qu'au moment où un questionnaire est envoyé aux établissements par la poste lors du recensement économique suivant. En ce qui concerne le secteur de la fabrication, un flux important de ces nouvelles unités reclassées a été observé durant l'année de recensement. La correction de la pondération des nouvelles entreprises ne comptant qu'un seul établissement est la stratégie qui a été adoptée pour le cycle de 1994-1998 de l'ASM afin de répartir l'effet de ces cas sur les années intercensitaires. Cette stratégie a toutefois été abandonnée pour le panel courant de l'ASM, à cause des effets incertains que les modifications récentes du système de codage pourraient avoir sur la qualité et sur la cohérence des codes de classification.

3.2.3 Entrées d'entreprises ne comptant qu'un seul établissement non classifié antérieurement

Au moment où la SSA attribue le code initial de branche d'activité, une proportion importante d'établissements ne peuvent être classés parce que la description de l'activité principale de l'entreprise est insuffisante ou trop vague. Comme nous l'avons mentionné plus haut, le Census Bureau obtient des renseignements supplémentaires pour la classification de ces établissements grâce à une entente conclue avec le BLS. À mesure qu'ils sont transmis par ce dernier, les codes de classification mis à jour sont intégrés dans le Registre des entreprises. Une opération distincte permet de repérer et d'intégrer les enregistrements antérieurement non classés auxquels est finalement attribué un code correspondant au secteur de la fabrication.

La première étape de cette opération consiste à repérer les établissements non classifiés dans la version du RE de l'année précédente. Le fichier de ces établissements est apparié à la version courante du RE afin d'établir leur situation de classification la plus récente. Les établissements qui, à ce stade, sont classés dans le secteur de la fabrication sont extraits du Registre en vue de leur inclusion dans les programmes visant le secteur de la fabrication. Ces unités entrantes non classifiées antérieurement sont au nombre d'environ 5 000 par année (y compris 4 000 dont le chiffre de paye est inférieur au seuil d'inclusion dans l'enquête postale). Le Bureau crée ensuite un fichier comprenant les établissements de fabrication dont la taille est supérieure au seuil d'exclusion de l'enquête postale. Ce fichier est utilisé pour ajouter les unités entrantes antérieurement non classifiées aux échantillons des enquêtes sur le secteur de la fabrication, sauf l'ASM. Pour cette dernière, le Bureau suit une procédure différente.

Dans le cas de l'ASM, le fichier complet des enregistrements entièrement non classifiés auxquels a été attribué un code correspondant au secteur de la fabrication (y compris les enregistrements d'établissements dont la taille est inférieure au seuil d'exclusion de l'enquête postale) est étendu de sorte qu'il soit représentatif des enregistrements non classifiés figurant dans le RE de l'année courante qui seront, en dernière analyse, classés dans le secteur de la fabrication. Pour calculer ce facteur d'extension (ajustement de la pondération), nous tenons à jour un registre du nombre d'établissements non classifiés antérieurement qui sont maintenant classés dans le secteur de la fabrication, ainsi que du nombre d'établissements classés maintenant dans un autre secteur que celui de la fabrication. Nous déterminons aussi le nombre d'établissements enregistrés dans le RE courant qui ne sont pas classifiés. D'après ces dénombrements, nous calculons la proportion relative d'établissements classés dans le secteur de la fabrication et appliquons cette proportion au nombre total d'établissements non classifiés qui figurent dans le RE courant. Nous obtenons ainsi une estimation du nombre prévu d'établissements couramment non classifiés qui finiront par être classés dans le secteur de la fabrication. Ces cas sont représentés dans les estimations de l'année courante de l'ASM grâce à l'application de facteurs de correction de la pondération aux unités antérieurement non classifiées qui entrent dans le secteur de la fabrication.'

3.3 Problèmes de qualité que posent les unités qui entrent dans le secteur de la fabrication

Comme, dans le cas du secteur de la fabrication, le traitement des unités entrantes est annuel, le délai d'intégration n'est pas un problème aussi important que pour les secteurs du commerce de détail, du commerce de gros et des services. Dans le cas de la fabrication, les principaux problèmes de qualité des données sont liés à la classification erronée des nouvelles entreprises ne comptant qu'un seul établissement et au volume fortement variable d'enregistrements ne contenant pas de code de branche d'activité dans le RE. Comme nous l'avons mentionné plus haut, une proportion importante de nouvelles unités se voient attribuer un code initial de branche d'activité qui ne reflète pas l'activité réelle de l'établissement. Si les codes ne sont pas mis à jour avant le recensement économique suivant, l'univers de la fabrication présente des lacunes durant les années inter censitaires. Pour le cycle de 1994 à 1998 de l'ASM, nous avons résolu le problème en estimant le nombre d'unités non classifiées (d'après le recensement antérieur) et nous les avons réparties sur les années inter censitaires. Cependant, à cause de la conversion au SCIAN et des modifications subséquentes apportées aux systèmes de classification, nous avons interrompu cette procédure. Lorsque les changements de classification auront été mis en œuvre, nous aimerions réaliser une étude de la qualité et la fiabilité des nouvelles méthodes de classification.

Le nombre d'enregistrements non classifiés qui figurent dans le RE varie de moins de 100 000 à plus d'un quart de million. Comme la classification selon la branche d'activité des nouveaux EIN dépend fortement des renseignements transmis par d'autres organismes gouvernementaux, toute interruption de ce flux d'information risque d'avoir un effet important sur le nombre d'enregistrements non classifiés. Bien que la proportion d'enregistrements non classifiés auxquels on attribue, en bout de ligne, un code du secteur de la fabrication soit relativement faible (ordinairement de 4 % à 5 %), elle introduit néanmoins chaque année un certain degré d'incertitude en ce qui concerne l'univers de la fabrication.

4. CONCLUSION

Les programmes des enquêtes sur le commerce de détail, le commerce de gros et les services et des enquêtes sur le secteur de la fabrication ont mis en place des procédures en vue de mettre les échantillons de départ à jour en tenant compte des nouvelles entreprises. Le moment de l'introduction des nouvelles unités et des unités entrantes dans chaque échantillon dépend directement de la fréquence des enquêtes. Cependant, nous étudierons l'effet de l'ajout annuel plutôt que trimestriel des nouvelles unités dans les échantillons dans le cas des enquêtes sur le commerce de détail, le commerce de gros et les services.

À l'heure actuelle, le processus trimestriel est partiellement utile au programme d'enquêtes sur le secteur de la fabrication, grâce aux codes du SCIAN attribués lors de l'enquête postale de la première phase d'échantillonnage. Certaines nouvelles unités non classifiées au départ, ainsi que certaines unités classées au départ dans le secteur du commerce de détail, du commerce de gros ou des services sont classées dans le secteur de la fabrication durant ce processus. Le processus de codage du genre d'activité des nouvelles unités rendra un meilleur service à l'Economic Directorate si la proposition visant à étendre le processus trimestriel de traitement des nouvelles unités est mise en œuvre.

BIBLIOGRAPHIE

- BSR-2K Action Memo 2N1, "First and Second Stage Birth Sampling," février 2001, document interne.
- Konschnik, C. (1988), "Coverage Error in Establishment Surveys," article présenté à l'Annual Meeting of the American Statistical Association, Nouvelle Orléans, LA..
- Konschnik, C., Monsour, N., Detlefsen, R. (1985), "Constructing and Maintaining Frames and Samples For Business Surveys," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp.113-122.