

DÉFIS D'UNE ENQUÊTE DE GRANDE ENVERGURE SUR LES POSTES VACANTS : PASSER DE LA CONCEPTUALISATION À LA MISE EN ŒUVRE

Danielle LEBRASSEUR¹ (), Étienne RASSART (*), Susan DEMEDASH (*)*

() Statistique Canada, Direction de la méthodologie, Division des méthodes
d'enquêtes auprès des entreprises*

Résumé

Dans le but de combler des lacunes sur la demande de main d'œuvre au Canada, le développement d'une enquête d'envergure sur les postes vacants et les salaires (EPVS) a été lancé à la fin de l'été 2014 à Statistique Canada. L'EPVS comprendra deux volets : un volet trimestriel sur les postes vacants et un volet annuel sur les salaires. Le volet sur les postes vacants a été mis en œuvre en février 2015 et se répétera sur une base trimestrielle auprès d'un échantillon de 100 000 entreprises au Canada. La première diffusion est prévue en août 2015. Le volet annuel sur les salaires doit s'intégrer à l'échantillon de l'EPVS dès janvier 2016 et les premières estimations doivent être publiées d'ici la fin de la même année. Le but principal de l'enquête est de publier des estimations sur les postes vacants et les salaires selon les régions économiques et les professions.

Des enquêtes antérieures ont démontré que la collecte de données sur les postes vacants et les salaires auprès des entreprises constitue un défi de taille. De plus, une enquête d'une telle envergure impose un fardeau de réponse préoccupant autant pour les petites entreprises que pour les grandes. Tirant profit de certaines leçons apprises, cet article exposera le plan de sondage qui a été adopté pour la réalisation du volet trimestriel sur les postes vacants et abordera la stratégie envisagée pour l'intégration du volet annuel sur les salaires.

Abstract

In order to fill gaps in labour statistics data on job vacancies and wages, the Job Vacancy and Wage Survey (JVWS) was launched in July 2014. The JVWS has two components: a quarterly component on job vacancies and an annual component on wages. The survey started to collect data on job vacancies from a quarterly sample of 100,000 business locations in Canada in February 2015. Starting in January 2016, a wage component will be collected once a year, using the same sample. The main goal of this survey is to produce estimates of job vacancies and wages by economic region and occupation. This paper will cover the survey design for the job vacancy component and will also address the proposed strategy for the integration of the wage component.

Mots-clés :

Enquête trimestrielle, emplacements, Système de classification nationale des professions, sondage de Poisson, nombres aléatoires permanents

¹ danielle.lebrasseur@statcan.gc.ca

Introduction

Les ministres provinciaux du marché du travail de l'ensemble du Canada se réunissaient en juillet 2014 afin de prendre des mesures pour que les Canadiens puissent posséder les compétences dont ils ont besoin pour être compétitifs dans le contexte d'une économie mondialisée. Au cours de ce Forum, les ministres ont convenu qu'une meilleure information sur le marché du travail de meilleure qualité est nécessaire pour appuyer la prise de décision des chercheurs d'emploi, des étudiants, des employeurs et des responsables des politiques. Ces informations permettent notamment aux Canadiens de prendre des décisions éclairées concernant leur formation et leurs études, ainsi qu'à se tenir au courant des emplois offerts. Le gouvernement du Canada a donc accepté de financer deux nouvelles enquêtes, au coût de 14 millions de dollars par année, afin de recueillir des données fiables et comparables sur les salaires et les emplois disponibles et de bien comprendre les pressions actuelles et émergentes sur le marché du travail.

Le développement d'une enquête sur les postes vacants et les salaires (EPVS) de grande envergure a été lancé à la fin de l'été 2014 à Statistique Canada. L'EPVS est composée de deux volets : un volet trimestriel sur les postes vacants et un volet annuel sur les salaires. Le volet de l'enquête sur les postes vacants a été mis en œuvre en février 2015 sur un échantillon trimestriel de 100 000 entreprises au Canada dans le but de produire des estimations du nombre de postes vacants et des taux de postes vacants à des niveaux détaillés. Les niveaux visés de publication sont les régions économiques et les professions. Des premières estimations doivent être publiées pour le mois d'août 2015.

Un volet annuel, soit celui des salaires, doit s'intégrer à l'échantillon de l'enquête sur les postes vacants dans le but de recueillir des données sur les salaires payés des postes existants au sein des entreprises sélectionnées. L'intégration de ce volet est prévue en janvier 2016 et les premières estimations annuelles doivent être publiées d'ici la fin de la même année.

Des enquêtes antérieures ont démontré que la collecte de données sur les postes vacants et les salaires auprès des entreprises constitue un défi de taille. Avec une enquête d'une telle envergure, le fardeau de réponse est un aspect préoccupant autant pour les petites que pour les grandes entreprises. Cet article exposera en détail le plan de sondage qui a été adopté pour le volet sur les postes vacants et abordera brièvement la stratégie envisagée pour l'intégration du volet annuel sur les salaires.

1. Plan de sondage de l'EPVS – volet sur les postes vacants

1.1. Objectif principal

L'objectif principal du volet sur les postes vacants est de publier des estimations trimestrielles du nombre de postes vacants et du taux de postes vacants selon les régions économiques et les professions. Il importe de mentionner que le partenaire principal, le ministère de l'Emploi et Développement social Canada, s'intéresse plus à l'évolution des changements d'un trimestre à l'autre qu'au niveau d'un trimestre donné. La première diffusion des données de l'enquête est prévue en août 2015.

Un poste vacant est défini comme un « poste pour lequel l'entreprise recrute activement et qu'elle prévoit pourvoir au cours des 30 prochains jours ». Cette définition est celle qui a été utilisée par l'Enquête sur le milieu de travail (EMT) de 2011² et qui est actuellement utilisée par l'Enquête sur l'emploi, la rémunération et les heures (EERH). Cette dernière diffuse actuellement des estimations du nombre et du taux de postes vacants à l'échelle provinciale et secteur industriel.

Le taux de postes vacants est défini comme le ratio du nombre de postes vacants sur le total du nombre d'employés et des postes vacants. Il est actuellement très difficile à partir du volet sur les

² Demedash, S., Jang, L., Lothian, J., « Enquête sur le milieu de travail – postes vacants : des défis en abondance », *Recueil du Symposium 2013 de Statistique Canada: Produire des estimations fiables à partir de bases imparfaites*, Ottawa, 2014, pp. 232-238.

postes vacants de produire des estimations du taux de postes vacants selon la profession, car le nombre d'employés selon la profession n'est pas disponible. Cependant, il sera possible de produire un taux de postes vacants selon la région économique, toutes professions confondues. On espère se servir du volet salaire de l'EPVS pour fournir une estimation du nombre d'employés selon la profession, bien que ce ne soit pas l'objectif principal de ce volet qui sera discuté plus tard.

Le Canada est constitué de 76 régions économiques. Elles représentent des unités géographiques normalisées pour l'analyse de l'activité économique. Le détail visé des statistiques sur les professions correspond au code à quatre chiffres du Système de classification nationale des professions (CNP³). Cette classification contient 500 codes de professions organisés selon une structure hiérarchique. Voici un exemple de cette structure:

- CNP-1: 4 - Enseignement, droit et services sociaux, communautaires et gouvernementaux
- CNP-2: 40 - Personnel professionnel en services d'enseignement
- CNP-3: 401 - Professeurs/professeures d'université et assistants/assistantes d'enseignement au niveau postsecondaire
- CNP-4: 4011 - Professeurs/professeures et chargés/chargées de cours au niveau universitaire

1.2. Autres indicateurs potentiels

Lors de la planification de l'enquête, les indicateurs potentiels suivants ont été identifiés :

- Proportion de postes vacants pour une profession et/ou un secteur d'activité qui sont à pourvoir depuis 60 jours ou plus (durée intermédiaire pour les postes difficiles à pourvoir).
- Proportion de postes vacants pour une profession qui paye au moins X,YY \$/heure
- Taux de postes vacants selon le type de poste (permanent/temporaire)
- Taux de postes vacants selon la taille de l'entreprise (selon le nombre d'employés)
- Distribution des postes vacants selon les salaires horaires les plus faibles proposés
- Distribution des postes vacants selon les exigences en matière d'éducation

La référence au secteur d'activité correspond aux 20 secteurs industriels définis essentiellement selon les codes à 2 chiffres du Système de classification des industries de l'Amérique du Nord (SCIAN⁴).

2. Base de sondage et population

2.1. Base de sondage : Le Registre des entreprises de Statistique Canada

La Division du Registre des entreprises (DRE) joue un rôle clé pour ce qui est d'assurer la cohérence des programmes d'enquête économiques de Statistique Canada en fournissant une base centrale harmonisée qu'est le Registre des entreprises (RE). Ce dernier est une base qui englobe essentiellement toutes les entreprises en exploitation au Canada ainsi que les entreprises étrangères qui entretiennent des liens avec les entreprises canadiennes. Il fournit aux statisticiens chargés de la production et de l'analyse de statistiques économiques une base de très grande qualité du point de vue de la couverture et des éléments de données.

Le RE suit les concepts fondamentaux du Système de comptabilité nationale (SCN) et il permet d'entreposer, de consulter, d'actualiser et d'extraire des renseignements utilisés par les programmes économiques. Il fournit de l'information au sujet de la composition de la population des entreprises au Canada, à savoir leur structure organisationnelle, leur activité industrielle, leur taille et leur emplacement géographique. Sur demande, le RE crée une liste de personnes ressources et de codes d'identification des questionnaires d'enquête qui sont nécessaires pour les processus de collecte de

³ La Classification nationale des professions 2011 est une structure hiérarchique à quatre niveaux de groupes de professions, comportant des niveaux successifs de désagrégation. Ces différents niveaux sont : les grandes catégories professionnelles, les grands groupes, les groupes intermédiaires et les groupes de base.

⁴ Le Système de classification des industries de l'Amérique du Nord (SCIAN) est un système hiérarchique composé de secteurs (codes à deux chiffres), sous-secteurs (codes à trois chiffres), groupes (codes à quatre chiffres) et classes (codes à cinq et à six chiffres).

données. Il surveille le fardeau de réponse imposé à chaque entreprise canadienne et il fournit des réponses pertinentes pour gérer efficacement les problèmes liés aux réponses.

Sur le RE, une entreprise est considérée comme une structure intégrée composée d'unités d'exploitation juridiques et non juridiques. On utilise les unités d'exploitation de la structure intégrée pour les besoins de l'échantillonnage. Les entreprises sont structurées de façon hiérarchique selon les quatre éléments statistiques suivants : l'entreprise statistique (unités juridiques), la compagnie statistique (centres d'investissement), l'établissement statistique (centres de profit) et l'emplacement statistique qui correspond à la plus petite entité au sein de la structure intégrée (unités locales, en Europe). L'emplacement peut être un immeuble ou une usine.

Une entreprise est complexe quand elle a plus d'une unité d'exploitation. Sur le RE, une structure complexe est définie comme une entreprise qui affiche un indicateur d'emplacements multiples ou un indicateur de structure complexe, tous les emplacements associés à l'entreprise étant complexes. Dans le cas d'une entreprise simple, tous les éléments statistiques sont équivalents comme s'ils faisaient un. Il est donc important de bien comprendre la structure de l'entreprise au moment d'établir un plan d'enquête.

Le RE produit également des facteurs d'allocation permettant de désagréger des données disponibles à un niveau supérieur de la structure d'une entreprise complexe vers des niveaux inférieurs. Ces facteurs permettent par exemple aux enquêtes de produire des estimations pour de nouvelles ramifications. Par exemple, des données d'emploi disponibles au niveau de l'entreprise peuvent être réparties au niveau des établissements ou au niveau des emplacements. Ainsi, des estimations à des niveaux industriels ou géographiques plus fins peuvent être produites. La qualité de ces estimations dépend de la qualité des facteurs d'allocation qui repose sur la structure établie d'une entreprise.

Plus d'une centaine d'enquêtes auprès des entreprises utilisent le RE comme base de sondage et la majorité d'entre elles échantillonnent au niveau de l'établissement. Le choix du RE comme base de sondage va de soi et l'EPVS ne fait pas exception à cette pratique corporative. Pour l'EPVS, le choix de l'unité de la base de sondage aux fins de l'échantillonnage s'est arrêté sur l'emplacement. Ce choix repose sur les deux arguments suivants:

- L'expérience de la collecte des données de l'EMT a conclu que l'information sur les besoins de recrutement est plus souvent connue au niveau de l'emplacement, contrairement aux informations financières ou économiques, qui sont plus faciles à obtenir au niveau de l'établissement ;
- L'EPVS a pour but de publier des estimations à l'échelle des régions économiques, lesquelles correspondent à des régions infraprovinciales bien définies. Des données qui seraient obtenues au niveau de l'établissement devraient être subséquentement réparties parmi les régions économiques dans le cas des entreprises complexes. En l'absence de facteurs d'allocations fiables pour le nombre de postes vacants, l'emplacement demeure la meilleure option pour le choix de l'unité d'échantillonnage.

2.2. Informations-clés de la base pour la stratification

Outre le choix de l'emplacement comme l'unité optimale à la fois pour l'échantillonnage et la collecte des informations visées, d'autres renseignements importants de la base servent à les stratifier de façon à optimiser le plan d'échantillonnage en fonction des domaines d'estimations visés par l'enquête. Dans cette optique, trois types de renseignements disponibles sur le RE ont été identifiés pour l'EPVS, soit la région économique, le secteur industriel et la taille des emplacements.

La région économique est considérée la dimension la plus importante, après la profession, dans les objectifs d'estimation. Le secteur industriel (SCIAN-2) a été jugé utile pour cerner certaines professions qui seraient répandues seulement dans de petits secteurs industriels. Bien qu'un niveau plus détaillé aurait été encore plus discriminant, il a fallu se limiter à un nombre raisonnable de strates résultant du croisement avec la région économique. Finalement, on a retenu la taille des emplacements définie selon le nombre d'employés pour la relation qui pourrait lier le nombre de postes vacants au nombre d'employés. Bien qu'un certain nombre d'entreprises de petite et moyenne taille peuvent afficher un nombre non négligeable de postes vacants (entreprises saisonnières ou nouvelles), on présume que ce nombre pourrait être plus élevé au fur et à mesure que la taille de l'entreprise augmente.

Deux sources d'information sont disponibles sur le RE pour le nombre d'employés : l'emploi profilé ou l'emploi alloué selon une source administrative. Cette dernière réfère à des formulaires administratifs de retenues sur la paie (PD7) remplis par les entreprises pour l'Agence du revenu du Canada et transmis sur une base régulière à Statistique Canada. Le nombre d'employés sur le RE représente pour chaque source le deuxième maximum observé au cours des 12 derniers mois. Aux fins de l'EPVS, la taille d'un emplacement donné a été définie selon le nombre le plus récent des deux sources. Pour la grande majorité des emplacements, l'emploi alloué des formulaires PD7 est sélectionné.

2.3. Populations cible et enquêtée

La population cible de l'EPVS a été extraite du RE et couvre tous les emplacements qui ont rapporté avoir au moins un employé au cours des 17 derniers mois, de toutes les régions économiques du Canada et de tous les secteurs industriels, y compris l'agriculture, la pêche et la foresterie, sauf les entreprises du secteur public référant aux administrations provinciales et fédérale. On prévoit toutefois inclure ce secteur ultérieurement.

La population enquêtée est définie par l'ensemble des emplacements de la population cible ayant au moins deux employés. Cette décision provient d'une recommandation du Comité de gestion du fardeau de réponse de Statistique Canada qui veille à contrôler le fardeau de réponse auprès des petites entreprises. Cette règle est cohérente avec celle appliquée dans l'EERH au niveau de l'établissement. La population couverte et exclue du processus de collecte sera traitée à l'étape d'estimation par le calage. Pour le premier trimestre de l'enquête, la population couverte se chiffrait à 1 165 000 emplacements et la population enquêtée à 870 000 emplacements.

3. Plan d'échantillonnage

Depuis plusieurs années, on utilise à Statistique Canada un système généralisé d'échantillonnage servant à la sélection des échantillons destinés aux enquêtes. La dernière version de ce système, G-Sam, a été utilisée pour l'EPVS, à la fois pour la stratification, la répartition, la sélection et la rotation. Un plan d'échantillonnage de Bernoulli stratifié à une phase est utilisé pour l'enquête. Ce plan a été choisi pour sa simplicité et sa flexibilité.

3.1. Stratification

La population d'emplacements a été stratifiée selon l'industrie (SCIAN-2, ou secteur), la géographie (région économique) et la taille de l'emplacement définie selon le nombre d'employés. Pour la stratification par taille, deux ou trois catégories de taille sont utilisées, selon le nombre d'emplacements compris dans le croisement de la région économique et du secteur industriel. Les limites des strates ont été définies à partir de la méthode géométrique de Gunning-Horgan (2004)⁵ qui s'avère utile pour les populations asymétriques. Dans une première étape, les unités influentielles (soit les plus grosses, dont la borne de taille a été définie au 99^{ème} percentile) ont été isolées dans chaque croisement « région économique – secteur » et mises de côté dans des strates spéciales à tirage complet. On leur assignera donc un poids de 1. Pour les unités restantes, les bornes forment une suite géométrique ($ar^0, ar^1, ar^2, ar^3, \dots$) où le terme initial et la raison sont calculés à l'aide des valeurs minimale et maximale observées dans le croisement : a = valeur minimale et r = maximum / minimum. Avec quatre classes de taille, on aura donc typiquement trois strates à tirage partiel (petits, moyens, gros) et une strate à tirage complet (très gros).

Dans l'établissement de la stratégie de stratification, il a été établi que la population des emplacements serait restratifiée chaque trimestre selon les informations les plus à jour de la base de sondage. Il a été évalué qu'environ 3% des emplacements seraient réassignés à différentes strates d'un trimestre à l'autre. On a préféré une restratification trimestrielle à une restratification annuelle, car elle permet d'intégrer les changements de strates au fur et à mesure, ce qui représente moins de changements à la fois et donc n'affecterait pas d'une façon importante la comparaison des estimations

⁵ Gunning P., Horgan J.M., « Un nouvel algorithme pour la construction de bornes de stratification dans les populations asymétriques », *Techniques d'enquête*, vol. 30, n° 2, pp. 177-185, décembre 2004

d'un trimestre à l'autre. La restratification annuelle, qui amènerait plus de changements, pourrait causer un bris plus important qui compliquerait l'analyse des tendances de la série trimestrielle lors du passage d'une année à l'autre. Ajoutons que la restratification trimestrielle permet de produire des estimations chaque trimestre sur une population transversale à jour.

3.2. Répartition

En raison de la disparité de la taille des régions économiques au Canada, nous avons opté pour une répartition en puissance (*power allocation*). Ce type de répartition a pour but d'établir un bon équilibre de qualité entre les estimations des petites et des grandes régions économiques. La taille d'échantillon totale a ainsi été attribuée de façon approximativement proportionnelle à la racine carrée du nombre d'employés estimé dans chacune des strates. Cette stratégie assigne une plus grande taille d'échantillon dans les petites régions et favorise une plus grande qualité des estimations publiées pour elles, sans trop compromettre la qualité des estimations pour les plus grandes régions économiques, de même que les estimations provinciales et nationale.

3.3. Sélection

Pour sélectionner l'échantillon, on a utilisé l'échantillonnage de Poisson. Pour ce faire, un nombre aléatoire permanent (NAP) choisi uniformément dans l'intervalle [0, 1] a été attribué initialement à chaque unité de la population. Le processus de répartition produit, pour chaque strate, une fraction de sondage. Les unités d'une strate ont toutes cette même fraction de sondage comme probabilité de sélection. Une unité est choisie si son NAP est inférieur ou égal à cette probabilité de sélection. Cette méthode de sélection doit être modifiée lorsqu'on a aussi une rotation de l'échantillon (voir plus bas).

3.4. Rotation et mise à jour

Le partenaire principal de cette enquête s'intéresse plus aux estimations des tendances qu'à celle des niveaux. Pour répondre à cette demande, la sélection des 100 000 emplacements de l'enquête a été mise en oeuvre selon un plan de rotation de 1/8. Ainsi, une fois sélectionné, un emplacement demeurera dans l'échantillon pendant deux ans (8 trimestres), avant d'en sortir (sauf pour les emplacements à tirage complet). Ce plan de rotation permet de mettre en oeuvre un processus de collecte plus robuste, en limitant le nombre de contacts préliminaires requis pour chaque cycle d'enquête. Il contribue aussi à stabiliser les estimations et à augmenter la précision des tendances. Avant le début de chaque nouveau cycle, l'échantillon sera aussi mis à jour pour tenir compte des créations et des suppressions d'emplacements.

3.5. Échantillons de 2015 et mesures spéciales de démarrage

Pour la première année de mise en oeuvre de l'enquête, il a été décidé de ne pas faire de rotation de l'échantillon pour favoriser un meilleur rodage des opérations de collecte. Cependant, l'échantillon sera mis à jour pour refléter les changements dus à la restratification des unités de la population ainsi que l'entrée de nouveaux emplacements ou la suppression d'emplacements dans la population visée.

La taille de l'échantillon sélectionné pour le premier trimestre de 2015 est de 101 000 emplacements dont 18 500 dans des strates à tirage complet et 82 500 dans des strates à tirage partiel. Une initiative corporative mise en oeuvre en janvier 2015 sur la réduction du fardeau de réponse a également imposé d'exclure quelques petites entreprises de l'échantillon de l'enquête. À ces exclusions s'est rajoutée le retrait de presque 3000 emplacements afin de minimiser le chevauchement avec l'échantillon de l'EERH. Ces exclusions de l'échantillon seront considérées comme des cas de non-réponse à l'étape de traitement des données.

4. Collecte

La question de la période de référence pour la collecte des données sur les postes vacants nous a menés à un choix arbitraire, puisqu'il n'y a actuellement pas de règle acceptée à l'échelle internationale. Selon les pays, la collecte de statistiques trimestrielles sur les postes vacants peut se

faire un jour particulier du trimestre (p. ex., le quinzième jour du mois intermédiaire, ou le dernier jour ouvrable ou jour de travail du trimestre) ou selon une moyenne sur trois mois. Toutefois, au niveau conceptuel, nous a semblé préférable d'avoir une période de référence fixe dans le trimestre. Comme nous avons décidé d'aller vers une stratégie de collecte mensuelle, la période de référence a alors été établie au premier jour de chaque mois du trimestre.

4.1. Stratégie de collecte mensuelle

L'échantillon d'environ 100 000 emplacements, sélectionné chaque trimestre, est subdivisé d'une façon aléatoire en trois échantillons mensuels. Cela donne lieu à un échantillon plus équilibré au cours du trimestre de référence, du fait que l'on tient compte, dans un sens, de chaque mois du trimestre. Cette stratégie permet également aux bureaux régionaux d'organiser le travail de la collecte sur les douze mois de l'année, ce qui représente un avantage pour la planification des ressources requises. Même si la tâche est répartie chaque mois, elle n'en demeure pas moins un défi, étant donné la taille importante de l'échantillon mensuel d'environ 33 333 emplacements.

Afin de maximiser la collecte de l'échantillon trimestriel, le calendrier de collecte a été établi de manière à laisser chaque mois de collecte ouvert. Ceci signifie que la collecte du premier et deuxième échantillons mensuels d'un trimestre donné pourra se prolonger après la fin de leur mois respectif, tout en s'assurant que ces cas soient finalisés avant la fin de la période de collecte assignée au dernier mois de chaque trimestre. Par exemple, pour le deuxième trimestre de 2015, la collecte de l'échantillon d'avril s'effectuera du 1^{er} au 30 avril, mais restera ouverte jusqu'au 10 juillet 2015; cette date étant l'échéance de la collecte de l'échantillon de juin. Cette stratégie de collecte assure ainsi que les entreprises ont assez de temps pour répondre à l'enquête.

4.2. Contenu du questionnaire et fardeau de réponse

Le questionnaire de l'EPVS pour le volet sur les postes vacants a été établi selon un questionnaire développé aux États-Unis par le *National JVS Workgroup* (le « modèle du Minnesota », plus récemment utilisé au Kansas). Dans ce questionnaire, on demande à la personne-ressource de fournir les renseignements suivants à l'emplacement sélectionné :

- Confirmation de l'emplacement
- Nombre d'employés rémunérés
- Nombre de postes vacants
- Des détails sur les postes vacants par profession (par exemple, titre, description du poste, temporaire ou permanent, niveau d'éducation requis, salaire offert, l'effort requis pour pourvoir le poste, etc.)

L'obtention de détails sur les postes vacants par profession pourrait représenter un fardeau de réponse grandissant avec la taille de l'entreprise, et encore plus important pour les grandes entreprises complexes qui couvrent un grand spectre de professions. Étant donné que l'on s'attend à un chevauchement de 80% et plus des échantillons d'un trimestre à l'autre, il est planifié d'alléger ce fardeau en transférant des renseignements sur les titres des postes vacants fournis par le répondant à un cycle donné au cycle suivant. On prévoit ajouter cette fonctionnalité à l'outil de collecte.

4.3. Mode de collecte et suivi de réponse

Le seul mode de collecte offert est le questionnaire électronique (QE). Le premier jour ouvrable du mois, toutes les entreprises ayant une adresse courriel reçoivent une invitation à remplir le QE pour participer à l'enquête. Les invitations initiales ne sont pas accompagnées d'un questionnaire papier. S'il n'y a pas d'adresse courriel, on envoie une lettre avec un code d'accès au QE.

De l'aide sous forme de collecte téléphonique assistée par informatique (CATI) est fournie sur demande seulement, ou lors du suivi de la non-réponse effectué par des intervieweurs dans les bureaux régionaux. Lors de la conception du QE, il a été planifié d'inviter le répondant à contacter le personnel des bureaux régionaux si l'une des deux situations suivantes survient : le répondant veut rapporter des données combinées pour plus de cinq emplacements, ou le répondant a plus de 20 professions à rapporter au moment d'établir sa liste de postes vacants. Dans la première situation, le personnel du bureau régional cherche à clarifier le lien entre le répondant et l'emplacement

échantillonné pour lequel la collecte devait être effectuée. Dans la deuxième situation, on offre au répondant l'option d'utiliser un outil de déclaration personnalisé se présentant sous la forme d'un chiffrier ; ce qui devrait faciliter et accélérer son processus de réponse.

4.4. Opération de précontact préliminaire

Suite à l'expérience de l'EMT, un des nombreux défis pendant la période de collecte est d'obtenir les données de l'emplacement sélectionné. Pour ce faire, il est primordial d'identifier la bonne personne-ressource pour chaque emplacement sélectionné, surtout pour ceux associés à une structure d'entreprise complexe. Malgré qu'il s'agit d'une enquête économique d'envergure, les coordonnées de personnes-ressources pour certains emplacements échantillonnés, particulièrement dans le cas de grandes entreprises complexes, n'ont pas toujours été disponibles sur le RE. De plus, les coordonnées de la personne-ressource contenues dans le RE sont généralement celles d'un comptable ou d'un contrôleur, étant donné que la majorité des enquêtes économiques recueillent des renseignements financiers. Aux fins de la présente enquête, on croit que le directeur des ressources humaines ou le gestionnaire du recrutement sont les plus aptes à fournir l'information sur les postes vacants. Dans le but de combler ces lacunes, une opération de contact préliminaire pour l'EPVS a été effectuée en décembre 2014 en préparation pour le début des opérations de collecte en février 2015. Cette opération visait à déterminer non seulement le bon répondant pour les questions sur l'emploi et les postes vacants, mais également à obtenir les adresses électroniques.

Aux fins du contact préliminaire, un échantillon initial de 17 600 emplacements a été tiré de l'échantillon trimestriel de 100 000 emplacements à l'aide du RE de novembre 2014. Cet échantillon regroupait les grands emplacements complexes se classant au-dessus du 75^e percentile pour le nombre d'employés dans chacune des 76 régions économiques. Ont été ajoutés à cet échantillon des hôpitaux, des conseils scolaires et des municipalités, puisque bon nombre de ces entités sont considérées comme des unités à emplacement unique dans la base de données du RE. Ont également été inclus dans cet échantillon initial du contact préliminaire les emplacements sous gestion de portefeuille d'entreprises (GPE). Ces derniers sont des emplacements de grandes entreprises au Canada qui sont suivis par des gestionnaires de GPE à Statistique Canada.

En raison de contraintes budgétaires, l'échantillon du contact préliminaire a été réduit à quelque 12 000 emplacements. Les emplacements comptant le plus grand nombre d'employés ont été retenus dans l'échantillon définitif du contact préliminaire. Un seuil a été défini de manière à obtenir un échantillon de la taille visée, y compris tous les emplacements sous GPE.

Le contact préliminaire a été effectué pendant environ 15 jours ouvrables en décembre 2014. Le taux de réponse final s'est établi à 87 %. Les 76 régions économiques ont été couvertes au cours de cette période. Les questions posées lors du contact préliminaire a permis de vérifier que la personne inscrite au dossier était bien celle qui pouvait répondre aux questions sur l'emploi et les postes vacants. Les renseignements demandés au sujet du répondant comprenaient le nom, le titre, l'adresse, l'adresse électronique et le numéro de téléphone.

Bien qu'on espérait obtenir un répondant distinct pour chaque emplacement, dans certains cas, une seule personne tenait à agir comme contact pour tous les emplacements de l'établissement ou de l'entreprise ou bien, pouvait fournir des données uniquement pour un ensemble combiné d'emplacements, ce que l'on appelle habituellement une « déclaration combinée ». Dans ce cas, on a recueilli des renseignements concernant les emplacements couverts, afin de pouvoir attribuer la réponse à l'emplacement échantillonné le moment venu. Cet ajustement *a posteriori* des réponses à l'emplacement initialement choisi a bien été couvert par l'EMT (voir [1]) et la même approche sera utilisée pour l'EPVS.

4.4. Premier cycle de collecte

La collecte du premier trimestre de 2015 a débuté seulement en février dû à un calendrier d'activités trop serré pour permettre de débiter en janvier. Cependant, le plan d'échantillonnage a été appliqué tel que planifié, c.-à.-d. qu'un échantillon d'environ 100 000 a été sélectionné pour le trimestre et subdivisé parmi les mois de janvier, février et mars. L'échantillon de janvier n'a pas été envoyé à la collecte, et celle-ci a débuté en février avec l'échantillon alloué pour ce mois. Ensuite, l'échantillon de

mars a été envoyé à la collecte en mars, et ainsi de suite. Bien que la collecte de chaque mois est concentrée sur l'échantillon mensuel correspondant, la collecte reste ouverte pour tous les échantillons du premier trimestre jusqu'au 10 avril, selon le calendrier de collecte établi. Ce premier trimestre sera considéré comme un trimestre partiel puisqu'environ deux tiers seulement de l'échantillon aura été envoyé à la collecte. Bien que cela représente tout de même 66 000 emplacements, l'omission de certaines unités importantes en janvier – notamment un tiers des emplacements des strates à tirage complet – risque cependant de compromettre la représentativité des estimations dans des domaines particuliers pour le premier trimestre. Cet enjeu ne sera pas présent pour le deuxième trimestre et les suivants. Les répercussions sur la qualité des estimations qui seront diffusées en août 2015 seront évaluées le moment venu.

4.5. Suivi de non-réponse et priorisation

Une série de rappels par courriel et d'appels de suivi à des intervalles prédéterminés sont planifiés dans le calendrier de collecte pour les cas de non-réponse. En vue de planifier cette étape, la question de priorisation des cas pour le suivi de non-réponse a été soulevée. Un premier scénario envisagé était de prioriser seulement les emplacements de grande taille, soit ceux avec le plus grand nombre d'employés. Cependant, la stratégie de priorisation est étroitement liée à la méthode de traitement envisagée pour corriger pour la non-réponse totale. Pour l'EPVS, l'approche privilégiée pour le traitement de la non-réponse totale est la repondération. Pour que cette approche soit envisageable, il faut s'assurer que chaque emplacement non répondant, qu'il soit de petite ou de grande taille, ait une chance non nulle d'être sélectionné pour le suivi de non-réponse. Puisque les emplacements de grande taille peuvent être importants pour l'estimation des postes vacants dans une région économique donnée, il a été recommandé d'utiliser la méthode d'échantillonnage séquentielle de Poisson élaborée par Ohlsson⁶.

Avec l'échantillonnage séquentiel de Poisson, un rang a été attribué à chaque emplacement de l'échantillon pour le suivi de non-réponse mis en œuvre pour l'échantillon de février. Afin d'assurer que les plus grands emplacements de l'échantillon bénéficient d'un suivi, on a attribué un rang de 1 à tous les emplacements dont le nombre d'employés était supérieur à 500. Ces emplacements forment essentiellement une strate à tirage complet pour les besoins du suivi (cette strate est suffisamment petite pour que des ressources puissent toujours lui être affectées). Ensuite, un rang est attribué aux emplacements restants à partir du rang 2 en utilisant le classement par ordre selon l'échantillonnage séquentiel de Poisson.

Pour déterminer le rang des emplacements pour le suivi selon cette méthode, un nombre aléatoire u_i (entre 0 et 1) est associé à chaque emplacement i de l'échantillon. Puis, les valeurs $\xi_i = u_i/X_i$ sont calculées, X_i étant le nombre d'employés de l'emplacement. Les emplacements sont alors triés selon les valeurs croissantes des ξ_i , et un rang allant de 2 à la taille d'échantillon n est attribué à chaque unité selon l'ordre des ξ_i .

Les avantages de cette méthode en ce qui concerne le suivi de non-réponse de l'EPVS sont les suivants :

- La taille du sous-échantillon du suivi de non-réponse ne doit pas être spécifiée à l'avance, de sorte que l'on peut utiliser le classement par ordre pour attribuer un rang à chaque unité de l'échantillon avant même de commencer le suivi ;
- Comme l'échantillonnage séquentiel de Poisson est approximativement un échantillonnage de probabilité proportionnelle à la taille, il favorise les emplacements comptant le plus grand nombre d'employés et, par conséquent, les emplacements les plus susceptibles de contribuer de manière importante au nombre total de postes vacants.

5. Traitement des données et estimation

Les détails concernant les stratégies de traitement et d'estimation sont en cours de développement. En gros, il est envisagé de traiter la non-réponse partielle par l'imputation et la non-réponse totale par la repondération. Il est prévu d'utiliser les systèmes généralisés de Statistique Canada BANFF et G-Est, respectivement pour la vérification et l'imputation des données, et pour l'estimation.

⁶ Ohlsson E., « Sequential Poisson Sampling », *Journal of Official Statistics*, vol. 14, n° 2, pp. 149-162, 1998

6. Calage

Dans le contexte du plan d'échantillonnage utilisé pour l'EPVS, il est prévu de procéder à un calage des estimations. Le calage, qui consiste en un ajustement des poids initiaux, peut être effectué pour différentes raisons. D'abord, avec le calage, on peut s'assurer que les effectifs estimés à partir de l'échantillon correspondent aux effectifs de population. On peut aussi compenser la partie exclue de la population enquêtée dans l'estimation finale d'une caractéristique de la population. Dans le contexte de l'EPVS, le premier ajustement est requis dû au caractère aléatoire de la taille d'échantillon par strate, une caractéristique inhérente au choix du plan d'échantillonnage de Poisson. Le deuxième ajustement sera effectué pour compenser l'exclusion des emplacements avec un employé à l'estimation du nombre total de postes vacants. Il importe de mentionner que l'EERH procède à cet ajustement en utilisant des totaux de contrôle fondés sur des données d'emploi des formulaires PD7.

Le poids initial attribué à chaque emplacement de l'échantillon a été établi de manière à représenter les effectifs de la population enquêtée à l'intérieur d'une strate (région économique x SCIAN-2 x taille). L'ajustement des poids par calage pourra se faire à un niveau aussi détaillé que les strates ou à un niveau moins détaillé, selon la disponibilité et la qualité des totaux de contrôle. Il sera également important de considérer dans la stratégie de calage les domaines d'estimation visés par l'enquête.

En plus d'assurer une cohérence, le calage vise à améliorer les estimations, et donc requiert que les totaux de contrôle proviennent d'une source fiable. Pour l'EPVS, la stratégie de calage suivante est envisagée :

- 1) Calage sur les effectifs d'emplacements de la population enquêtée par strate, ou sinon, par région économique et taille. Les effectifs d'emplacements sont établis à partir du RE.
- 2) Calage sur le nombre d'employés par province x SCIAN-2 provenant des données des formulaires PD7, validées et allouées au niveau établissement, ce qui est équivalent aux totaux publiés par l'EERH. Notons que ce calage ne sera pas appliqué au secteur de l'agriculture, car il est exclu de la population visée de l'EERH.

7. Intégration du volet annuel sur les salaires

Le volet annuel sur les salaires doit s'intégrer à l'échantillon de l'enquête sur les postes vacants dans le but de recueillir des données sur les salaires payés des postes existants au sein des entreprises sélectionnées. L'intégration de ce volet est prévue en janvier 2016 et les premières estimations annuelles doivent être publiées d'ici la fin de la même année.

On espère se servir du volet salaire de l'EPVS pour produire des estimations annuelles sur les salaires payés par profession, selon les régions économiques. Pour ce faire, on utilisera le même échantillon que celui pour le volet sur les postes vacants. On conservera aussi la même optique, soit de répartir la mesure et la collecte sur l'ensemble des mois de l'année, au lieu d'un ou de plusieurs mois donnés. Pour ce faire, chaque emplacement de l'échantillon recevra un deuxième NAP 1, 2, 3 ou 4 qui déterminera dans quel trimestre de l'année le questionnaire sur les salaires lui sera envoyé. Ceci est illustré dans la figure ci-dessous.

			1er trimestre			2ème trimestre			3ème trimestre			4ème trimestre		
			1	2	3	4	5	6	7	8	9	10	11	12
Trimestre assigné dans l'année	1	Mois assigné dans le trimestre	1	S										
			2		S									
			3			S								
	2	Mois assigné dans le trimestre	1				S							
			2					S						
			3						S					
	3	Mois assigné dans le trimestre	1						S					
			2							S				
			3								S			
	4	Mois assigné dans le trimestre	1								S			
			2									S		
			3											S

Les principaux avantages de l'approche proposée sont les suivants. Premièrement, la collecte sera répartie de façon uniforme chaque mois de l'année, ce qui équilibrera le fardeau de réponse. Deuxièmement, seulement un quart des emplacements recevront les deux volets chaque trimestre. Ceci favorisera un meilleur taux de réponse trimestriel pour le volet sur les postes vacants. En effet, même si le volet sur le salaire causait une baisse de la réponse, cette baisse affecterait seulement un quart de l'échantillon du trimestre. Finalement, on aura une période de référence identique pour les emplacements qui répondront aux deux volets.

Bibliographie

- [1] Demedash, S., Jang, L., Lothian, J., « Enquête sur le milieu de travail – postes vacants : des défis en abondance », *Recueil du Symposium 2013 de Statistique Canada: Produire des estimations fiables à partir de bases imparfaites*, Ottawa, 2014, pp. 232-238.
- [2] Gunning P., Horgan J.M., « Un nouvel algorithme pour la construction de bornes de stratification dans les populations asymétriques », *Techniques d'enquête*, vol. 30, n° 2, pp. 177-185, décembre 2004.
- [3] Ohlsson E., « Sequential Poisson Sampling », *Journal of Official Statistics*, vol. 14, n° 2, pp. 149-162, 1998.
- [4] Statistique Canada, Classification nationale des professions (CNP) 2011, <http://www.statcan.gc.ca/subjects-sujets/standard-norme/noc-cnp/2011/index-indexe-fra.htm>
- [5] Statistique Canada, Le Système de classification des industries de l'Amérique du Nord (SCIAN) Canada 2012, <http://www.statcan.gc.ca/subjects-sujets/standard-norme/naics-scian/2012/index-indexe-fra.htm>