

ÉVOLUTION RÉCENTE DE LA COLLECTE ÉLECTRONIQUE DES DONNÉES À STATISTICS NEW ZEALAND

Nancy McBeth, Stuart Pitts et Steven Johnston¹

RÉSUMÉ

Ces quelques dernières années, Statistics New Zealand (SNZ) a vu s'accroître la quantité de données communiquées par courrier électronique aux fins des enquêtes-entreprises, mais jusqu'à présent, cet organisme n'a pas mis en place les mécanismes d'enquête-entreprises pouvant lui permettre de soutenir la collecte électronique de manière à répondre aussi bien à ses propres besoins qu'à ceux des fournisseurs de données. C'est pourquoi SNZ s'est efforcé, pendant la dernière année, d'étudier la meilleure façon d'aborder les problèmes et les possibilités d'une collecte électronique. Le présent document décrit sommairement ce qu'il se propose de faire pour passer d'une collecte par courrier électronique à un système sécuritaire de transmission électronique et, à plus long terme, à un système de collecte par Internet. Il expose en outre une étude de cas des données actuellement communiquées par courrier électronique dans le cadre de l'enquête mensuelle sur le commerce de détail. Cette étude illustre certains des avantages d'une collecte électronique, mais aussi un certain nombre de problèmes de qualité de données et de coûts pour l'organisme. Il fait ressortir la nécessité de considérer la méthodologie de collecte de données dans la perspective plus générale du cycle entier d'une enquête.

MOTS CLÉS : Courrier électronique; Commerce de détail; Qualité des données

1. INTRODUCTION

À l'instar de bien d'autres organismes statistiques nationaux, Statistics New Zealand (SNZ) s'est rapidement attaqué aux conséquences de la multiplication des données fournies par voie électronique. Il y a encore quelques années, l'organisme ne se souciait pas outre mesure de ce qui n'était encore qu'une déclaration électronique occasionnelle ou spéciale, mais avec un recours croissant au courrier électronique, la collecte électronique de données a gagné en importance sans que SNZ dispose des mécanismes d'enquête-entreprises pouvant lui permettre de soutenir cette collecte de manière à répondre au mieux tant à ses propres besoins qu'à ceux des fournisseurs de données. Voilà pourquoi il a depuis un an mis beaucoup de temps à étudier comment le mieux aborder les problèmes et les possibilités d'une telle collecte.

2. CADRE DE COLLECTE DE DONNÉES AUPRÈS DES ENTREPRISES

L'économie néo-zélandaise est modeste à l'aune planétaire avec une population totale d'un peu moins de 3,8 millions de personnes au recensement de mars 2001. La Nouvelle-Zélande comptait alors quelque 1,8 million d'emplois répartis dans près de 300 000 entreprises. Les grands changements économiques des décennies 1980 et 1990 ont transformé la nature de l'économie de ce pays qui, de sa vaste dépendance de toujours à l'égard de l'agriculture et de ses produits, s'est de plus en plus mise à l'heure de la diversification. Un aspect primordial de cette évolution a été l'ouverture de son économie non seulement

¹ Nancy McBeth et Stuart Pitts, Survey Design and Development Division, Steven Johnston, Survey Methods Division, Statistics New Zealand, C.P. 2922, Wellington, Nouvelle-Zélande.

dans ses débouchés, mais aussi dans sa structure de propriété. En février 2000, environ 2,3 % des entreprises néo-zélandaises étaient en propriété étrangère dans une proportion d'au moins 25 %. Comme dans bien d'autres économies, la propriété extérieure n'est pas d'une répartition égale, comme en témoignent les secteurs des finances et des assurances où plus de 12 % des entreprises sont au moins du quart en propriété étrangère.

Ce régime de mondialisation a pour conséquence que, de plus en plus, les sièges des grandes sociétés du pays se situent à l'étranger. Comme il est très probable que la réponse aux enquêtes du SNZ soit faite par le bureau chef, cela a voulu dire que SNZ a dû découvrir une façon plus systématique de faire enquête auprès des entreprises ayant leur siège dans d'autres pays.

Un important facteur dans la production d'indicateurs économiques sûrs et en temps opportun est le besoin de trouver un juste milieu entre les demandes de données et les charges de réponse. Comme d'autres organismes, SNZ a subi des pressions politiques en vue d'un allègement du fardeau de réponse. On a de plus en plus insisté en pareil cas sur l'importance d'un recours à la technologie comme moyen de réduction des charges de réponse. Dans une récente étude des coûts d'observation de la réglementation publique pour les entreprises, on a recommandé que SNZ presse le pas dans ses travaux visant à ménager une transmission sécuritaire des données et cherche au plus tôt à améliorer la circulation de l'information et les consultations par la technologie Web.

Il faut donc voir les éléments d'évolution qu'évoque le reste de cet exposé dans le double contexte des pressions grandissantes en vue d'un allègement du fardeau de réponse et de la nécessité de mener des activités de collecte dans une économie mondialisée.

3. ÉVOLUTION À SNZ

3.1 Il y a d'abord eu le courrier électronique

SNZ fait presque toutes ses enquêtes-entreprises sur papier, mais cela ne devait pas empêcher les enquêtés de renvoyer leurs données sous diverses formes électroniques et, dans certains cas, c'est un mouvement qui a été activement favorisé. La télécopie a probablement été la forme la plus courante de renvoi électronique et le courriel est de plus en plus utilisé en gros de la même manière comme instrument normal à la disposition des entreprises. Depuis 1997, SNZ a fait figurer une adresse générale de courrier électronique (surveys@stats.govt.nz) sur une foule de ses questionnaires d'enquête-entreprises, et on ne s'étonnera pas qu'un certain nombre d'enquêtés se soient mis à fournir leurs données par courrier électronique à cette adresse. L'enquête la plus importante pour le nombre de tels renvois a été l'enquête mensuelle sur le commerce de détail (ECD), qui est décrite dans l'étude de cas qui suit.

3.2 Politique de collecte électronique de données

Une fois que SNZ a commencé à considérer à la possibilité d'implanter un système Internet de collecte de données, il a jugé en conséquence qu'il lui fallait se doter d'une politique de recours permanent à Internet comme support de collecte. Voici comment on peut très succinctement résumer cette politique :

Statistics New Zealand n'acceptera par Internet que les données électroniques qui ont été à la fois encodées et authentifiées selon les normes recevables de l'industrie.

La politique en question vise principalement à faciliter l'adoption d'un mode Internet de collecte de données de manière tant à respecter les obligations qu'impose à l'organisme la loi néo-zélandaise sur la statistique qu'à sauvegarder les principes d'intégrité et de confiance qui font de ce même organisme un service statistique national efficace.

3.3 Plans conçus pour un système de collecte Web

Comme première mesure prise en application de sa politique, SNZ a entrepris de mettre en place un système sûr de transmission électronique de données provisoirement appelé « Secure Deposit Box ». Avec ce « coffret de sûreté », on se trouve à aborder la notion d'une collecte Internet à un niveau tout à fait fondamental. On se ménage les bases sur lesquelles reposera toute évolution future. Les enquêtés acquerront la possibilité de communiquer leurs données sous toute forme (texte, fichier joint, etc.) plutôt que de recevoir au départ des questionnaires sur support électronique. Cela vient compliquer toute tentative de pleine intégration aux systèmes de traitement de données d'enquête, mais donne de la souplesse aux enquêtés. Comme il sera impossible d'utiliser le coffret de sûreté sans autorisation préalable, une des clés du système sera l'obligation faite à l'ensemble des utilisateurs de conclure un accord officiel de collecte électronique avec SNZ, lequel se chargera des autres aspects de la gestion du système (gestion des mots de passe, etc.). Le résultat net est que SNZ exercera un contrôle plus étroit sur l'exploitation du système, ce qui en facilitera, devons-nous espérer, le développement et le perfectionnement.

Un système de collecte Web parvenu à maturité sera sans doute issu des améliorations apportées au coffret de sûreté. Comme premier perfectionnement, il y aura sans doute l'intégration de modèles de données d'enquête par voie électronique qui « contrôlent » la structure de présentation des données renvoyées. Une fois cette structure uniformisée, il devient plutôt aisé de faire le lien avec les systèmes de traitement de données d'enquête, les données électroniques étant traitées comme les données entrées par imagerie, par lecture optique ou sur clavier. Une des grandes questions que devra résoudre SNZ lorsqu'il passera à ce modèle de collecte plus complexe sera celle de l'« effet modal » possible de la diversité de modes de collecte. Comme nous l'évoquerons plus loin, on a commencé à travailler à cette question en ce qui concerne les données actuellement renvoyées à l'organisme par courrier électronique. Avec l'implantation de systèmes de collecte plus perfectionnés et plus uniformes, il devrait être plus facile de constater tout effet modal et de trouver une solution.

4. ÉTUDE DE CAS : RENVOIS PAR COURRIER ÉLECTRONIQUE AUX FINS DE L'ENQUÊTE SUR LE COMMERCE DE DÉTAIL

4.1 Contexte

Dans l'enquête sur le commerce de détail (ECD), il y a collecte mensuelle du seul chiffre d'affaires (valeur des ventes) de chaque établissement (unité géographique ou GEO) visé. La valeur mensuelle totale des ventes au détail s'établit actuellement à environ 3,6 milliards de dollars néo-zélandais (elle approche des 4,5 milliards en décembre). Avec cette valeur des ventes, on veut aussi savoir si le chiffre comprend la taxe sur les produits et services (TPS) et quelle est la période effectivement visée par la déclaration. Tous les trois mois, on demande également la valeur des stocks de fermeture.

La sélection aux fins de l'ECD s'opère au niveau des entreprises, chacune comprenant une ou plusieurs GEO. Tous les établissements GEO de vente au détail d'une entreprise sélectionnée sont visés par l'enquête. À l'heure actuelle, l'échantillon est formé de 4 000 entreprises et d'environ 7 900 unités géographiques, dont la répartition n'est pas homogène entre les entreprises échantillonnées. Quelque 80 % des entreprises visées ont une seule GEO. En revanche, dans une des grandes entreprises à établissements multiples, on dénombre plus de 200 GEO. Dans l'ECD, SNZ obtient invariablement un taux de réponse qui est de 88 % à 90 % des GEO de l'échantillon.

Comme nous l'avons dit, l'adresse de courrier électronique surveys@stats.govt.nz a été de plus en plus utilisée par les répondants. C'est d'avril à mai 1999 qu'on a entrepris de demander un renvoi par courrier électronique aux fins de cette enquête. C'est aussi à ce moment-là qu'on a uniformisé une structure de réponse que la plupart des enquêtés semblent avoir continué à employer. Les derniers mois de la même année, SNZ a réitéré sa demande de renvoi par courrier électronique, tentant ainsi de réduire les délais

d'enquête. Dans les opérations de collecte, on a l'habitude de poster un questionnaire papier, ce qui veut dire que le courrier électronique n'est pas là au départ, sauf dans un petit nombre de cas, comme invitation à une réponse par voie électronique. Les mêmes enquêtés répondent généralement à chaque cycle de l'enquête par courrier électronique. À l'occasion, un répondant s'ajoutera à ce groupe, mais il s'agit le plus souvent d'utilisateurs réguliers de ce mode électronique de collecte.

On accepte les déclarations par courrier électronique en texte ou par fichier joint. La figure 1 dégage la tendance de ces renvois dans le cadre de l'ECD pour les mois compris entre juin 1999 et juillet 2001.

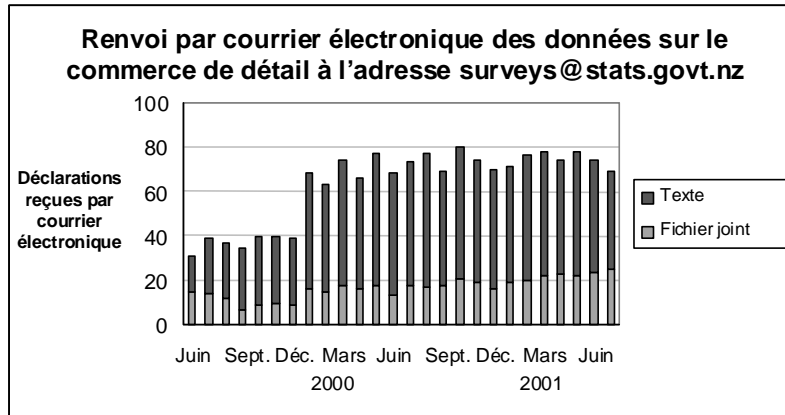


Figure 1

Les observations qui suivent sont tirées d'une analyse de quatre ensembles de déclarations ECD par courrier électronique à intervalles bimestriels depuis décembre 2000. Dans cette étude de cas, l'analyse porte sur l'entreprise plutôt que sur l'établissement GEO, car c'est la première qui est le répondant aux fins de cette enquête.

On a reçu par courrier électronique les réponses de 70 entreprises en moyenne (c'est là environ 2 % des entreprises de l'échantillon). Le nombre mensuel de communications électroniques est relativement modeste, mais il faut noter que les entreprises comptant un grand nombre de GEO paraissent plus enclines à répondre par courrier électronique. La figure 2 indique qu'un peu plus de 60 % des entreprises visées ayant plus de 50 GEO ont répondu par ce moyen, contre 1 % seulement des entreprises à GEO unique.

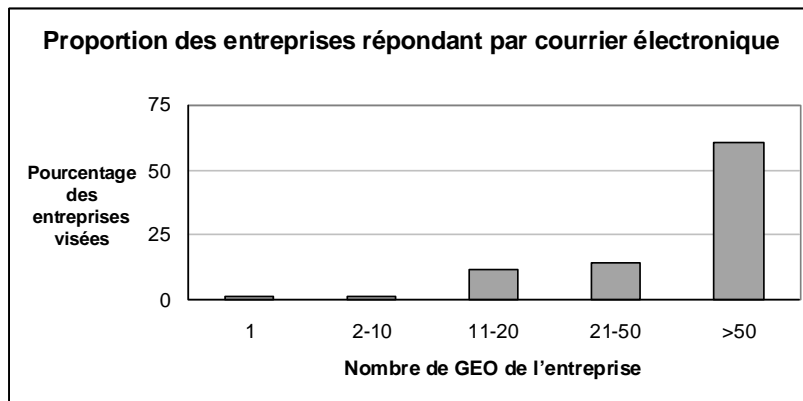


Figure 2

Non seulement les entreprises à établissements multiples ont plus de chances de répondre par courrier électronique, mais si elles le font, ce sera plus probablement par fichier joint. Plus de 80 % des entreprises à GEO multiples ont produit leurs déclarations par courrier électronique sous forme de fichiers joints (sur tableur Excel le plus souvent) comparativement à 7 % seulement des entreprises à GEO unique. Cela a sans doute à voir avec la façon dont les grandes entreprises extraient des données d'enquête de leurs systèmes financiers. Il est plus facile à un gros organisme de mettre en place un système automatisé d'extraction de réponses multiples en bloc et de les reporter sur tableur ou autre fichier. Les entreprises à GEO multiples qui produisaient leurs déclarations en texte étaient fort peu nombreuses. Les plus grandes comptaient seulement 5 GEO.

Le nombre effectif de renvois par courrier électronique ne paraît pas élevé, mais la valeur en dollars des déclarations en question représente environ 9 % de la valeur totale des ventes diffusée par l'enquête. Si on regarde la ventilation des ventes totales selon les types de détaillants, on constate une ample variation de l'importance relative des déclarations par courrier électronique. Comme l'indiquent les éléments d'analyse de la figure 3, la contribution des déclarations électroniques à la valeur des ventes est de plus de 60 % dans le cas des grands magasins et nulle dans les secteurs des spiritueux, des chaussures et des meubles et revêtements de sol.

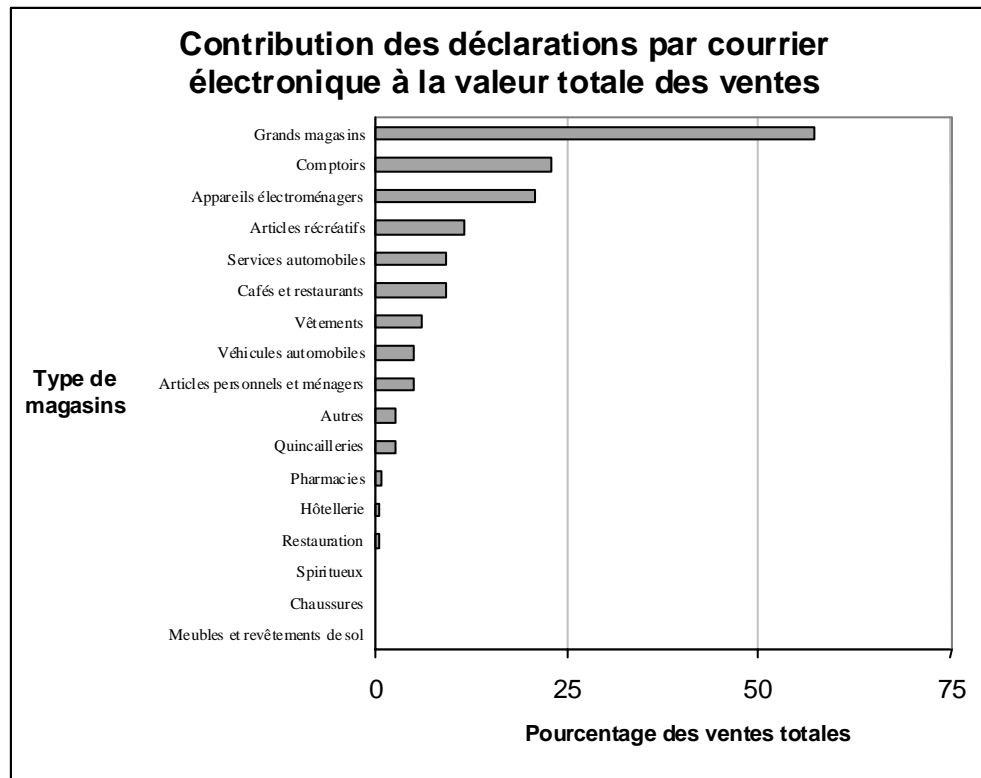


Figure 3

*La contribution dégagée est une moyenne pour les quatre mois analysés.

4.2 Rapidité de réponse

Pour ce qui est des délais de réponse, il est intéressant d'observer les tendances des réponses par courrier électronique pendant la période de collecte. On a fréquemment supposé que le courrier électronique constitue un important outil de réponse au moment critique de la fin du cycle de collecte, mais c'est une hypothèse qui ne semble pas confirmée par les résultats obtenus. C'est ce qu'on peut voir en prenant la moyenne du taux de réponse cumulatif sur trois mois (décembre étant exclu à cause des fêtes de fin

d'année) et en comparant le taux de réponse par courrier électronique au taux global de réponse comme à la figure 4.

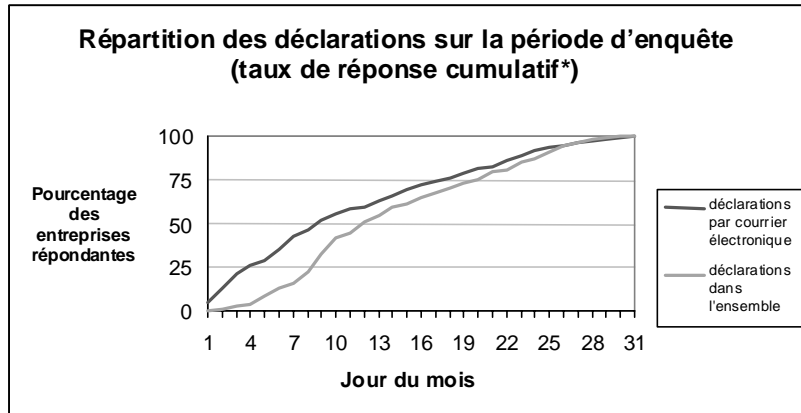


Figure 4

*Comme on ignore le nombre de non-répondants par courrier électronique, le taux de réponse est fondé sur une proportion de toutes les déclarations reçues.

On doit se rappeler deux choses à propos de ce graphique. D'abord, les questionnaires sont postés vers la fin d'un mois quelconque avant même que les enquêtés n'aient terminé la période visée par la déclaration. Ainsi, le renvoi le plus hâtif ne se fera qu'au terme du dernier jour de commerce du mois visé. En second lieu, l'opération générale de codage se fait au traitement des questionnaires plutôt qu'à leur réception, mais les dates de réception électronique utilisées sont les dates effectives d'arrivée. Le plus souvent, les questionnaires seront traités le jour de leur réception pour que les données ne soient pas faussées outre mesure.

Cela étant dit, il est toujours possible de distinguer une nette différence sur le plan de la rapidité de réponse. Ainsi, on parvient à un taux de renvoi de 50 % le 9^e jour pour les déclarations par courrier électronique et le 12^e jour pour l'ensemble des déclarations. La figure 5 décrit la rapidité de réponse d'une autre manière qui dégage encore mieux la « culmination » des déclarations par courrier électronique plus tôt dans le cycle de réponse, ce qui s'explique peut-être par la durée relative de renvoi des déclarations (qui est de quelques minutes au plus pour le courrier électronique et de quelques jours pour la poste). Comme nous l'avons signalé, le renvoi par courrier électronique ne paraît pas nettement gagner en importance en fin de cycle de déclaration.

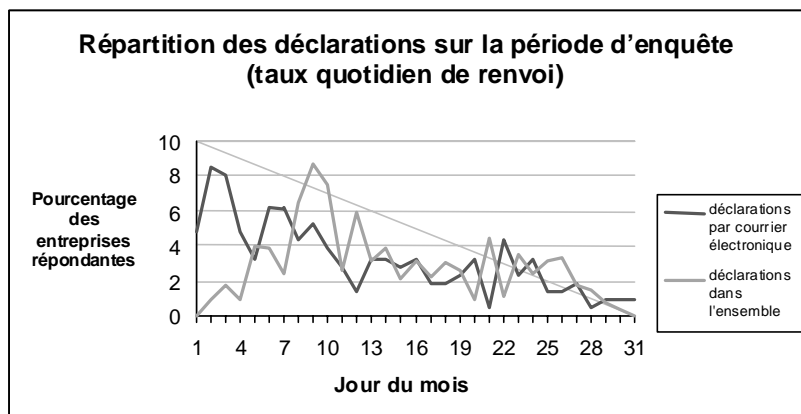


Figure 5

4.3 Qualité des données

Pour ce qui est de la qualité des déclarations par voie électronique, il a été impossible de bien comparer les données brutes reçues par courrier électronique aux données correspondantes acheminées par la poste, seules les données en communication électronique ayant été examinées en détail. Cela ne doit cependant pas nous empêcher de reconnaître quatre grands sujets d'inquiétude :

- absence d'indications de TPS;
- absence de chiffres sur les stocks;
- absence de données relatives à un certain nombre de GEO;
- absence de codes de répondant.

Il y a d'abord la question de l'inclusion d'indications de TPS. Comme pour toutes les données de cette nature, il importe de voir si les données reçues d'un répondant comprennent la TPS ou non. Comme l'indique la figure 6, il n'y avait aucune indication de TPS dans 24 % environ des déclarations par courrier électronique, tendance qui ne variait pas pour les quatre mois étudiés. Un premier examen des questionnaires renvoyés sur papier révèle qu'environ 4 % des répondants n'avaient pas précisé si la TPS figurait ou non dans la valeur déclarée des ventes. Il semble y avoir là un « effet modal ». On doit étudier la question plus à fond, mais il est possible que cette omission s'explique par la séparation du questionnaire d'avec la réponse.

Il convient de noter que, en cas d'absence d'indication de TPS, la prise en compte de cette taxe peut habituellement être supposée par les réponses antérieures de l'unité déclarante, les entreprises n'ayant pas tendance à modifier leur façon de déclarer cette taxe. On s'inquiétera peut-être moins dans ce cas des indications manquantes de TPS, mais le problème se posera toujours si les répondants changent leur façon de déclarer (en cas de changement de responsable de la déclaration au sein de l'entreprise, par exemple).

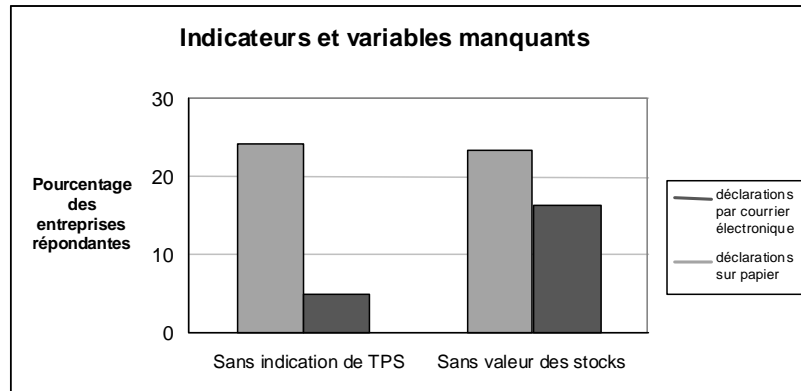


Figure 6

L'analyse de la question sur les stocks qui est posée tous les trois mois est moins concluante. L'examen des déclarations par courrier électronique ne comportait que deux mois de déclaration des stocks, à savoir décembre 2000 et juin 2001. La figure 6 indique que 23 % des entreprises ayant répondu par courrier électronique ont déclaré leurs ventes, mais non leurs stocks pour au moins une GEO (on notera qu'une valeur nulle déclarée de stocks ou toute autre indication fournie d'absence de stocks compte comme déclaration effective de cet élément). Pour les entreprises déclarant sur papier les mêmes mois, le pourcentage correspondant est d'environ 16 %. Il est vrai que les résultats relatifs à deux mois seulement peuvent ne pas être représentatifs, mais on est porté à penser que les répondants par courrier électronique risquent plus de ne pas déclarer de valeur de stocks. En poussant l'analyse, on pourra peut-être mieux discerner si un effet modal intervient ici (dans une analyse, par exemple, sur le plan de l'importance relative des valeurs manquantes de déclaration des stocks).

À en croire des données fragmentaires émanant des préposés au traitement des données d'enquête, il semblerait que les mois de déclaration des stocks posent un problème et que, assez souvent, les enquêtés ont de la difficulté à déclarer les valeurs de leurs stocks. Il ressort également de l'examen des questionnaires renvoyés tant par la poste que par courrier électronique que les valeurs communiquées sont dans bien des cas des estimations. Il s'agit en effet de valeurs arrondies (au millier le plus proche, par exemple), alors que les valeurs des ventes paraissent plus précises.

Un troisième problème de qualité se pose dans le cas de la non-réponse partielle là où des entreprises à GEO multiples fournissent des données pour une partie seulement de leurs établissements. La figure 7 indique la proportion de ces entreprises déclarantes qui ne renseignent pas sur au moins une de leurs GEO. Le plus remarquable ici, c'est la variation considérable de notre mesure de qualité des données des répondants par courrier électronique pour les quatre mois visés. Le pourcentage varie en effet de 28 % à 5 %, ce qui est peut-être en partie imputable au petit nombre d'entreprises à GEO multiples qui produisent leurs déclarations par courrier électronique. Il y a par ailleurs des indications nettes selon lesquelles les répondants par courrier électronique risqueraient davantage d'oublier quelques une de leurs GEO. Cela pourrait tenir à la séparation de la réponse d'avec le questionnaire, les établissements à l'égard desquels une entreprise est priée de communiquer des données étant imprimés d'avance sur le questionnaire.

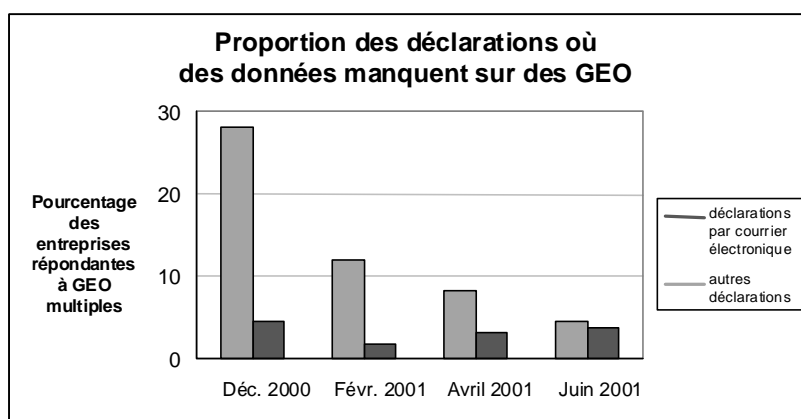


Figure 7

Comme dernier problème de qualité, il y a l'absence de codes de répondant dans certaines déclarations reçues par courrier électronique. Dans un nombre appréciable de ces déclarations, il n'y a aucune indication de l'unité statistique sur laquelle portent les données. Dans presque 14 % des déclarations électroniques, on ne trouvait aucun code de répondant et, dans 6 %, on relevait un code d'entreprise, mais aucun code d'établissement GEO. Dans le cas des entreprises à établissement unique, on résout habituellement le problème en appariant les noms et les unités de la base de sondage de SNZ, mais la tâche risque de se révéler plus laborieuse dans le cas des structures plus complexes. Parfois, les codes de répondant fournis (au niveau des entreprises ou des GEO) avaient été mal inscrits. Là encore, le problème tient sans doute à la séparation du questionnaire d'avec la réponse, car tous les questionnaires papier envoyés par SNZ (y compris ceux qui servent de rappel) portent des étiquettes ou des mentions imprimées de code de répondant.

4.4 Intégration des systèmes

Jusqu'ici, les systèmes d'enquête de SNZ ont généralement été centrés sur la réception et la saisie de déclarations sur papier. Il en va de même de l'enquête sur le commerce de détail. Il s'est avéré difficile d'intégrer les données communiquées par courrier électronique aux systèmes en place. Ainsi, toutes les déclarations en texte par courrier électronique (par opposition aux déclarations par fichier joint) doivent être examinées par le personnel de l'organisme et les données doivent en être transférées manuellement au

système de traitement. Cette opération manuelle est un peu facilitée par l'uniformité de la structure de présentation des déclarations mêmes par courrier électronique, puisque la plupart des répondants semblent avoir reçu un modèle à appliquer et s'y être tenus pendant un certain nombre de cycles d'enquête.

Les déclarations par fichier joint exigent encore plus une intervention manuelle que les déclarations en texte. Les récentes tentatives d'automatisation des opérations manuelles – par téléchargement automatique des données présentées par virgules de séparation dans le système de traitement – ont révélé des variations de présentation. Tous les fichiers joints ne sont pas des tableurs Excel, mais même pour ces tableurs, presque tous les fichiers annexés diffèrent de disposition et de forme de présentation. Voici quelques-unes de ces différences :

- positions dans les feuilles de calcul;
- espacement des rangées;
- inclusion de codes d'établissement;
- chiffres présentés en milliers;
- chiffres présentés en valeurs négatives.

Pour pouvoir se servir du système de téléchargement automatique, on doit retravailler chaque fichier joint à des fins d'uniformisation des présentations. Parfois, le travail est considérable. Dans un cas extrême, on a constaté que des valeurs avaient été déclarées pour près de 600 « lieux » ne correspondant pas à des unités statistiques SNZ.

4.5 Leçons à tirer de l'étude de cas

L'étude de cas fait voir d'intéressantes expériences d'utilisation du courrier électronique comme moyen de collecte de données. Les éléments d'analyse que nous avons présentés n'indiquent pas que cette utilisation crée des biais systématiques dans les données, mais l'étude de cas dégage certains points à étudier plus avant. Pour rester à l'écoute des besoins des enquêtés en matière de déclaration, SNZ a été prêt à offrir un éventail de supports de collecte. L'étude de cas illustre la nécessité de déterminer si aucun des problèmes de qualité de l'information relevés dans le cas du courrier électronique se remarque aussi dans le cas des autres supports. On a entrepris de soumettre à une analyse semblable les déclarations par télécopieur et téléphone, ce qui aidera SNZ et les utilisateurs de données à cerner l'incidence des modes de collecte sur les données.

Ce complément d'analyse sera également utile au moment d'aborder les problèmes de qualité de l'information que semblent poser les moyens de communication de données, c'est-à-dire les problèmes d'absence d'indications de TPS, de données relatives à des GEO et de codes de répondant. Il importera de bien constater si ces difficultés ont plus à voir avec la taille des déclarants qu'avec les modes de communication de données. Des données fragmentaires obtenues avant la réception de déclarations par courrier électronique semblent indiquer qu'un certain nombre de répondants se dotaient de tableurs et faisaient leurs déclarations SNZ par impression ou télécopie de feuilles de travail. Si les problèmes observés sont liés à la séparation du questionnaire d'avec la réponse, on ne saurait considérer qu'ils ont uniquement à voir avec les moyens de transmission des données.

L'ECD est sur le point de vivre d'importants changements, car un nouveau système de traitement de données d'enquête sera mis en place et le gros du travail de traitement et de production se transportera dans un autre centre (mais la réception à l'interne de tous les types de données sera toujours coordonnée au centre actuel). Comme nous l'avons signalé, SNZ est aussi en voie de créer un système de collecte Web qui remplacera le système de courrier électronique. Les changements apportés à l'ECD en général et l'implantation d'un nouveau système de collecte en particulier sont l'occasion de s'attaquer à certains des problèmes précités. Un net resserrement des liens avec les déclarants par voie électronique est un des résultats attendus du passage à un système de collecte Web. En établissant et en entretenant de telles relations, il devrait être possible d'exercer un meilleur contrôle sur la nature et la présentation des données actuellement communiquées. Un autre résultat sera l'obligation faite aux répondants de mentionner au moins un code d'identification de haut niveau dans toute déclaration. Cela n'aidera pas nécessairement à

résoudre les problèmes de déclarations multiples, mais on aura du moins la certitude que toute déclaration électronique émane d'une source identifiable.

Une autre façon d'améliorer la qualité des données serait de créer des modèles ou des spécifications en Excel à l'intention des grandes entreprises à établissements multiples avec des questionnaires papier en complément ou au lieu de tels questionnaires. Si la forme de présentation des données communiquées peut être la même pour tous les répondants, on aura à travailler bien moins fort pour préparer les déclarations sur tableur à un téléchargement automatique dans les systèmes de traitement des données d'enquête. En mettant des modèles individualisés à leur disposition, on pourra aussi aider les enquêtés à relier les données qu'ils communiquent à des codes d'identification et on leur rappellera constamment qu'ils doivent fournir des indications de TPS.

À l'autre extrémité de l'échelle où les répondants font généralement leurs déclarations en texte, on pourrait intégrer un type de formule électronique au système de collecte Web, ce qui aiderait grandement à garantir que toutes les données d'enquête exigées figurent dans les déclarations présentées.

Il n'y a pas que dans l'ECD que l'on incite les enquêtés à produire leurs données sous la forme qui leur convient le mieux. Dans un organisme comme SNZ où les responsabilités de gestion des enquêtés ont été centralisées, on gère cette incitation en demeurant souple devant les besoins des enquêtés. L'étude de cas fait toutefois ressortir l'importance de veiller à ce que les responsables de la collecte de données reçoivent régulièrement et systématiquement de la rétroaction au sujet de l'incidence des divers modes de collecte sur la qualité des données finales.

5. CONCLUSIONS

L'étude de cas des expériences d'utilisation du courrier électronique pour la collecte des données de l'enquête mensuelle sur le commerce de détail met en relief un certain nombre d'avantages et de problèmes de qualité de l'information qui peuvent être liés à une collecte électronique.

Au nombre des avantages, on compte une communication rapide des données et un meilleur étalement des renvois sur la période de réponse. L'étude de cas dégage néanmoins aussi un certain nombre de grands problèmes de qualité de données qui ne se posent pas dans un cadre de déclaration sur papier, qu'il s'agisse de l'absence de codes d'identification ou de celle de variables explicatives clés. On peut imputer de tels problèmes de qualité à la séparation des questions et des réponses dans les données finalement communiquées, mais on peut aussi nettement constater la nécessité de pousser l'étude du comportement des répondants.

L'étude de cas fait également ressortir le besoin de considérer l'incidence des changements de méthodologie de collecte dans la grande perspective du cycle d'enquête. L'expérience de l'ECD nous enseigne que le recours au courrier électronique à des fins de collecte n'a pas apporté, sur le plan des coûts, les économies attendues. Les coûts de cette enquête demeurent excessifs tant pour la solution des problèmes de qualité que nous avons évoqués que pour la transformation des données de courrier électronique sous des formes conformes aux cadres de traitement en place.

En Nouvelle-Zélande, le cadre de collecte des données des enquêtes-entreprises est bien plus difficile. Il y a des pressions qui s'exercent en vue de garantir que le fardeau de réponse imposé aux entreprises, et surtout aux entreprises petites et moyennes, sera le plus léger possible. On voit de plus en plus la technologie comme une partie de la solution, mais ainsi que nous le rappelle l'étude de cas, les solutions adoptées au stade de la collecte doivent s'intégrer à tout le cycle d'enquête pour que soit comblé le besoin d'obtention de données de qualité en temps utile et en toute rentabilité (rapport efficacité-coût).