

REVUE DES STRATÉGIES DE COLLECTE DE L'ENQUÊTE UNIFIÉE AUPRÈS DES ENTREPRISES DE STATISTIQUE CANADA

Rose-Carline Evra et Sylvie DeBlois¹

RÉSUMÉ

L'Enquête unifiée auprès des entreprises (EUE) est une enquête annuelle majeure de Statistique Canada qui est en cours depuis 1998. La collecte de données pour cette enquête pose certains défis dus, entre autres, aux contraintes de temps et de coûts. Dans le but de pouvoir déterminer à quelles entreprises un suivi prioritaire devrait être accordé dans le cas de non-réponse, une fonction de score fut introduite en 2002. Celle-ci est présentement revue, afin de la rendre plus efficace. Cet article portera sur la réévaluation de cette fonction et, en particulier sur l'amélioration du traitement des entreprises qui demandent un délai (rendez-vous) pour finaliser la collecte des données de leur questionnaire.

MOTS CLÉS : Collecte; entreprises; fonction de score; rendez-vous.

ABSTRACT

The Unified Enterprise Survey (UES) is a major annual survey at Statistics Canada that has been conducted since 1998. Data collection for this survey creates some challenges due, in part, to time and cost constraints. In order to determine the enterprises for which a high priority follow-up should be assigned in cases of non-response, a score function was introduced in 2002. This function is currently being reviewed in order to make it more efficient. This paper will discuss the re-evaluation of that function and, in particular, the improvement of the treatment of enterprises which ask for a delay (appointment) in order to finalize the data collection for their questionnaire.

KEY WORDS: Appointment, Collection, Enterprises, Score function.

1. INTRODUCTION

1.1 Collecte de l'Enquête unifiée auprès des entreprises

L'Enquête unifiée auprès des entreprises (EUE) est une enquête annuelle majeure de Statistique Canada qui est en cours depuis 1998. Elle est le principal véhicule de production des estimations annuelles pour plusieurs secteurs d'activité économique : statistique du commerce, de l'industrie de service, du secteur du transport, de l'agriculture et du secteur manufacturier (Gauthier, 2005). La collecte de données auprès des entreprises représente un défi intéressant, en effet, nous devons faire face à des contraintes de temps et de coûts, ce qui a un impact direct sur le taux de réponse. La période de collecte pour cette enquête s'étend du mois de février jusqu'au début du mois d'octobre. L'ensemble de l'information sur la collecte est conservé dans un système central nommé BLAISE conçu par les Pays-Bas et adapté par Statistique Canada pour ses besoins. L'information qui fut utilisée pour effectuer les différentes analyses présentées dans cet article a été extraite des fichiers de transactions qui sont créés par l'application Blaise.

Après la sélection de l'échantillon, un fichier de questionnaires est créé afin de déterminer quelles entreprises recevront un questionnaire. Les questionnaires sont envoyés par la poste puis, quelque temps après l'envoi, un processus de suivi auprès des non-répondants est entamé dans le but d'obtenir le plus de questionnaires complétés possible. Dans le cas qui nous intéresse, le suivi consistera en un appel téléphonique auprès du répondant, cependant d'autres méthodes de suivi

¹ Rose-Carline Evra, Statistique Canada, 100 Tunney's Pasture Driveway Ottawa, K1A 0T6, Canada, rose.evra@statcan.ca
Sylvie DeBlois, Statistique Canada, 150 100 Tunney's Pasture Driveway Ottawa, K1A 0T6, Canada, sylvie.deblois@statcan.ca

sont utilisées lors de la collecte (télécopie, renvoi de questionnaire, etc.). Pour ce qui est du rendez-vous, il peut être vu comme une indication du désir du répondant de discuter de son questionnaire ou de fournir l'information manquante nécessaire pour compléter son questionnaire à un moment ultérieur. Lorsqu'un rendez-vous est indiqué, le moment où celui-ci aura lieu peut avoir été déterminé par le répondant (pour lui permettre, entre autres, de trouver l'ensemble de l'information dont il a besoin ou d'être contacté à un moment plus opportun) ou par l'interviewer (pour lui rappeler de poursuivre le suivi auprès du répondant à un autre moment).

Dans cet article se trouvent les résultats d'une étude menée sur l'impact de la prise de rendez-vous lors de la collecte sur le taux de réponse de l'EUE. Avant de vous présenter ces résultats, un bref aperçu de la fonction de score de l'EUE vous sera donné.

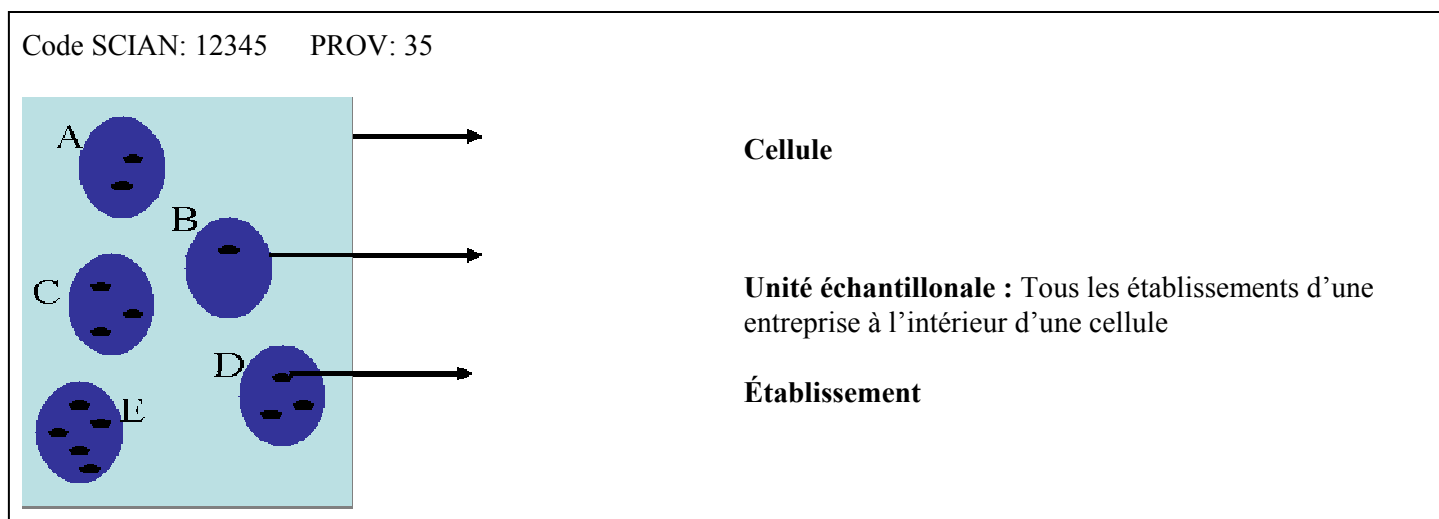
1.2 Fonction de score

En 2002, une fonction de score fut introduite avec l'objectif d'avoir un plan de suivi efficace auprès des entreprises (Pursey, 2003). Cette fonction accorde des priorités de suivi à la plupart des unités se trouvant dans l'EUE (certaines unités incluses dans l'EUE font partie d'un autre processus de collecte). Les entreprises ayant une priorité de suivi de 0 auront un suivi moindre que celles ayant une priorité de 1. Actuellement, un nombre élevé d'entreprises ayant une priorité de 1 demande un délai (rendez-vous) pour finaliser la collecte des données de leur questionnaire.

Examinons la fonction de score de plus près. Cette fonction a pour objectifs principaux l'obtention d'un taux de couverture pondéré élevé à la collecte et la réduction des coûts de collecte tout en maintenant la qualité des données. Cette fonction est dynamique, car elle met à jour la priorité attribuée aux questionnaires de façon régulière durant la collecte en tenant compte des codes de réponse attribués à ceux-ci.

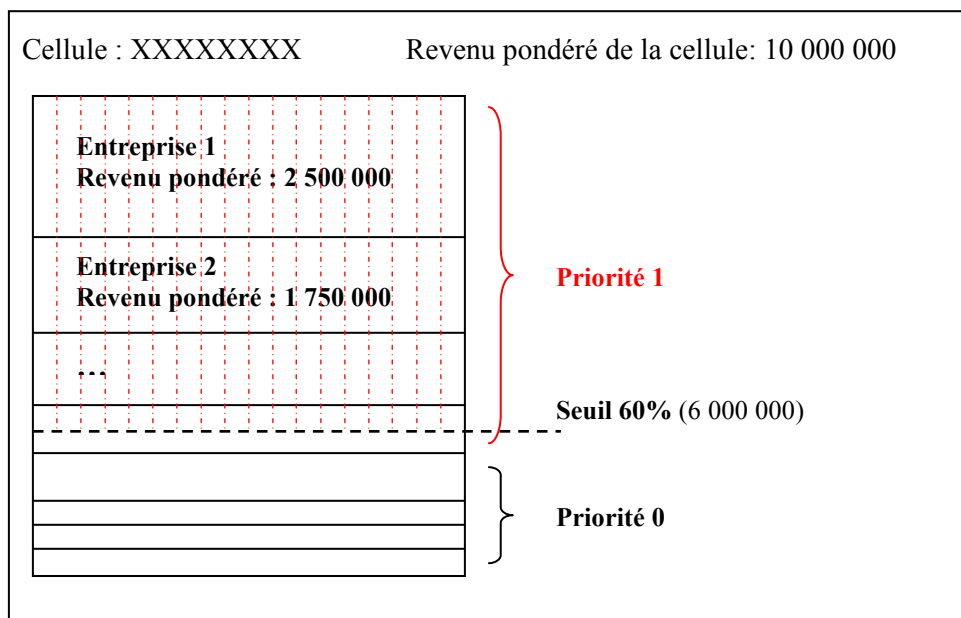
La priorité accordée à un questionnaire est déterminée à l'aide d'un score calculé pour chacun de ceux-ci au niveau de chaque cellule. Une cellule est une combinaison de code SCIAN² * PROVINCE pour chacune des enquêtes de commerce et d'agriculture et une combinaison de code SCIAN * PROVINCE * STRATE pour chacune des enquêtes de services et de transport (voir figure 1). Le score est le revenu pondéré du questionnaire en pourcentage du revenu total de la cellule. Comme mentionné précédemment, la fonction de score est dynamique, car les priorités des questionnaires sont mises à jour régulièrement durant la collecte. Après avoir trié les scores en ordre décroissant, on fait la somme des scores, un après l'autre, jusqu'à l'obtention d'un seuil de couverture prédéterminé. Tous les questionnaires sous le seuil auront une priorité de 1 alors que le reste des questionnaires recevront une priorité de 0 (voir figure 2).

Figure 1 : Cellule



² SCIAN: Système de classification des industries de l'Amérique du Nord

Figure 2 : Attribution des priorités initiales



Ces priorités sont mises à jour régulièrement durant la collecte selon les codes de réponse des questionnaires et ce qui reste à couvrir pour atteindre le seuil fixé. Les questionnaires reçus (complétés ou non) contribuent à l'atteinte du seuil et font donc diminuer la partie du seuil qu'il reste à atteindre. Les non-répondants ne contribuent pas à l'atteinte du seuil, alors que les unités hors champ sont retirées complètement de la cellule. Le tableau 1 résume les différents statuts des questionnaires.

Tableau 1 : Statuts possibles des questionnaires et contribution au taux de couverture	
Réponse (reçu ou complet)	Questionnaire contribuant à l'atteinte du seuil
Non-réponse	Questionnaire ne contribuant pas à l'atteinte du seuil
Hors champ	Questionnaire enlevé de la cellule (donc, réduction du revenu total de la cellule)
En traitement	Questionnaire toujours en processus de collecte (comprend les rendez-vous)

2. IMPACT D'UN RENDEZ-VOUS

2.1 Rendez-vous

Lors de la collecte de l'Enquête unifiée auprès des entreprises (EUE), un répondant peut souhaiter qu'on le rappelle à une date ultérieure. Dans ce cas, le code de réponse 90 (rendez-vous, appel téléphonique) est attribué au questionnaire de cette entreprise. Dans le but de déterminer l'impact de cette partie du processus de suivi, nous avons procédé à l'analyse de données de l'année de référence 2003 (AR2003).

Pour l'AR2003, le code de réponse 90 a été attribué à 9 766 questionnaires sur un total de 23 498, soit 41,6 % de ceux-ci. Suite à la prise du rendez-vous,³ 41,7 % des questionnaires furent remplis (partiellement ou complètement), 52,9 % ne furent pas remplis et dans 5,5 % des cas, le questionnaire avait été envoyé à une entreprise qui s'est révélée être hors champ lors du contact. On peut aussi noter que 37 % des répondants avaient pris un rendez-vous lors du suivi de la collecte, chez les non-répondants ce taux s'élève à 53,1 % et il est de 19,3 % pour les unités déclarées hors champ après un contact. (Voir le tableau 2, de la page suivante pour plus de détails)

³ Rendez-vous = code de réponse 90, et ce dans tout l'article.

Tableau 2 : Taux de réponse avec et sans rendez-vous			
	Avec rendez-vous	Sans rendez-vous	Total
Répondants %	4070 (37,0 %) 41,7 %	6939 (63,0 %) 50,5 %	11009 46,90 %
Non-répondants %	5162 (53,1 %) 52,9 %	4560 (47,9 %) 33,2 %	9722 41,40 %
Hors champ %	534 (19,3 %) 5,5 %	2233 (80,7 %) 16,3 %	2767 11,80 %
Total	9766 (41,6 %)	13732 (58,4 %)	23498

Suite à des analyses effectuées à l'aide de régressions logistiques, on constate que plus le rendez-vous est pris tôt, plus les chances que le questionnaire soit rempli augmentent. De plus, les questionnaires auxquels le code de réponse 90 n'a pas été attribué (sans rendez-vous) ont un taux de réponse supérieur aux autres. On constate aussi que plus le revenu est élevé, plus les chances que le questionnaire soit classé dans la catégorie « avec rendez-vous » sont élevées. Pour ce qui est de la durée pendant laquelle le questionnaire demeure avec le code de réponse 90, plus elle est longue, plus les chances que le questionnaire soit répondu diminuent.

Les rendez-vous offrent un choix intéressant aux répondants et nous permettent d'obtenir des questionnaires complétés qui autrement ne l'auraient pas été. Toutefois, il semblerait que pour augmenter les probabilités de réponse, il faudrait réduire l'utilisation des rendez-vous. Cependant, si on doit en prendre un, il devrait être pris tôt. De plus, le suivi devrait avoir lieu peu de temps après que le rendez-vous ait été pris. Sinon, les chances que ce questionnaire soit complété diminuent, et ce, même si le rendez-vous a été pris au début de la période de collecte.

2.2 Fonction de score et rendez-vous

Les questionnaires pour lesquels un rendez-vous est pris ont tous une priorité de 1 (suivi intensif). Sachant qu'un pourcentage plutôt élevé de questionnaires avec un code de réponse 90 ne sera pas complété (non-réponse), nous étudions présentement la possibilité d'augmenter le nombre d'unités avec priorité de 1 afin de compenser pour la perte possible de ces unités avec rendez-vous. Ceci pourra être fait en évaluant le pourcentage de la cellule que représentent les unités avec rendez-vous et en attribuant une priorité de 1 aux prochaines unités de la liste de scores triés qui auparavant avait une priorité de 0. Sachant que les unités en rendez-vous sont "en traitement" et peuvent l'être pour une longue période, un temps précieux peut être perdu pendant la période de collecte si ces unités deviennent non répondantes. En ajoutant des unités avec la priorité de 1, le temps d'attente pourra servir à réduire la partie du seuil qu'il reste à atteindre en allant chercher ces nouvelles unités. Le tout dans le but d'avoir un meilleur taux de couverture pondéré.

3. CONCLUSION

3.1 Conclusion

En ce qui concerne les rendez-vous, plus de recherches devraient être effectuées sur le sujet afin d'en améliorer leur utilisation et d'avoir une meilleure idée de l'effort réel mis sur ces unités lors de la collecte. Cependant, suite à cette étude, il apparaît que l'usage des rendez-vous lors de la collecte devrait être reconsidéré.

La collecte est l'une des premières étapes d'une enquête et demeure une étape primordiale. Le processus de collecte doit être revu et amélioré régulièrement, les données recueillies étant cruciales à la production d'estimations de bonne qualité. Afin d'améliorer le processus de collecte de l'EUE, nous prévoyons poursuivre nos travaux de recherche sur plusieurs aspects de la collecte.

RÉFÉRENCES

- Gauthier, S. (2005). « *De 7 à 26 enquêtes – Complexité grandissante de la gestion de l'échantillonnage pour l'Enquête unifiée auprès des entreprises* ». Compte rendu du Colloque francophone sur les sondages 2005, Québec.
- Pursey, S. (2003). "Use of the Score Function to Optimize Data Collection Resources in the Unified Enterprise Survey". Compte rendu de l'Assemblée annuelle de la SSC, juin 2003, Halifax.