11-617F
no.88-14
c.2

Statistique
Canada

STATISTICS CANADA STATISTIQUE CANADA
NOV 10 1994
LIBRARY
BIBLIOTHÈQUE

Methodology Branch

Business Survey Methods Division

Direction de la méthodologie

Division des méthodes d'enquêtes-entreprises

Canada

WORKING PAPER NO. METH-88-014F

METHODOLOGY BRANCH

CAHIER DE TRAVAIL NO. METH-88-014F

DIRECTION DE LA MÉTHODOLOGIE

LE FARDEAU DE RÉPONSE: SON CONTRÔLE ET SA RÉPARTITION

par

C. Poirier, R. Lussier et D. Leroux

Mai 1988

LE FARDEAU DE RÉPONSE: SON CONTRÔLE ET SA RÉPARTITION

par

C. Poirier, R. Lussier et D. Leroux

(Mai 1988)

PROJET DE REMANIEMENT DES ENQUÊTES-ENTREPRISES

LE FARDEAU DE RÉPONSE: SON CONTRÔLE ET SA RÉPARTITION

CLAUDE POIRIER, ROBERT LUSSIER ET DIDIER LEROUX

DIVISION DES MÉTHODES D'ENQUÊTES-ENTREPRISES

DÉCEMBRE 1987

(RÉVISÉ EN MAI 1988)

RÉSUMÉ

Le fardeau de réponse est généralement causé par une sélection répétée des unités dans une même enquête ou par la sélection multiple d'une unité par plusieurs enquêtes. S'il devient trop important, il peut pousser le répondant à fournir des données erronnées ou incomplètes, ou encore à ne pas répondre.

Ce document se veut un répertoire de méthodes visant à réduire les conséquences néfastes du fardeau de réponse. Il fournit premièrement des idées afin de déterminer et respecter les capacités de répondre de chaque répondant. Par la suite, des méthodes opérant au niveau de la sélection de l'échantillon y sont décrites. On y trouve alors l'échantillonnage synchronisé et différentes rotations échantillonnales. Finalement, des méthodes opérant au niveau de la collecte de données sont présentées. On y décrit entre autres, l'utilisation de questionnaires communs à plusieurs enquêtes.

ABSTRACT

Response burden is generally caused by repeated selection of units within a survey or by multiple selection of units by a number of surveys. If it becomes too important, it can lead a respondent to give erroneous or incomplete data, or even no response at all.

This document contains a list of methods aimed at reducing negative effects of response burden. First, it gives some ideas in order to determine and to respect respondents' abilities to answer. Then, methods working at the sample selection level are described. These include synchronized sampling and sample rotations. Finally, methods working at the data collection level are presented. Amongst other methods, the use of common questionnaires for a given set of surveys is considered.

TABLE DES MATIÈRES

1 Introduction 1
1.1 Objectifs 1
1.2 Présentation 1

2 Quelques notions sur le fardeau de réponse 3
2.1 Introduction 3
2.2 Définition 3
2.3 Effets 4
2.4 Evaluation 4

3 Contrôle du fardeau de réponse 7
3.1 Introduction 7
3.2 Respect des limites 7

4 Prévention du chevauchement entre les échantillons de différentes enquêtes 10
4.1 Introduction 10
4.2 Echantillonnage synchronisé 10

5 Répartition du fardeau de réponse à l'intérieur d'une enquête 13
5.1 Introduction 13
5.2 Répartition à partir d'un échantillonnage aléatoire simple 14
5.3 Répartition à partir d'un échantillonnage colloqué avec tailles fixées 17
5.4 Répartition à partir d'un échantillonnage de Poisson 20
5.5 Comparaison des méthodes de rotation 24

6 Diminution du fardeau de réponse 26
6.1 Introduction 26
6.2 Stratégie des questionnaires unifiés 26
6.3 Mise à la poste commune 28

7 Conclusion 29

Bibliographie 31

1 INTRODUCTION

1.1 OBJECTIFS

Lorsque des échantillons sont tirés indépendamment pour plusieurs sondages, un problème majeur se pose. Il s'agit de la sélection multiple de certains répondants qui se voient alors assigner une tâche plus importante. Ainsi, dans le cadre du projet d'assurance de la qualité à Statistique Canada, on réserve une place à l'étude du fardeau de réponse dans le but de diminuer le plus possible les conséquences néfastes qu'il peut entraîner.

La présente étude, ayant comme objectif de faire un compte rendu des principales méthodes de réduction du fardeau, se veut une suite logique au rapport "Etude empirique sur le fardeau du répondant" de Didier Leroux et Robert Lussier [7]. On y présente dans un premier temps, comment contrôler le fardeau de réponse de chaque entreprise afin de ne pas leur accorder plus de travail que ce qu'elles peuvent absorber. Bien entendu, il est difficile de déterminer ce que chaque répondant peut absorber mais, pour les besoins de ce rapport, nous supposerons que ces capacités sont connues.

On introduira ensuite différents moyens visant à réduire le fardeau de réponse entre les enquêtes de même qu'à l'intérieur des enquêtes tout en respectant le contrôle déjà établi. Prenant d'abord idée sur les méthodes existantes, nous transformerons ces dernières au besoin, pour qu'elles puissent s'adapter à nos différentes exigences.

1.2 PRÉSENTATION

La section 2 se voudra une base au présent rapport. On y définira le fardeau de réponse et on y décrira les effets importants qu'il peut avoir, autant sur le répondant que sur Statistique Canada.

La section 3 nous décrira deux principes visant à contrôler le fardeau de chacun des répondants. Ils seront tous deux basés sur les capacités des répondants en tant qu'accumulation de fardeau.

Enfin, la section 4 expliquera comment il est possible de coordonner plusieurs enquêtes afin de diminuer le fardeau

de réponse entre les enquêtes. On y introduira ainsi l'échantillonnage synchronisé.

La section 5 nous donnera trois façons de répartir le fardeau à l'intérieur d'une même enquête. Ces trois méthodes, adaptées respectivement aux échantillonnages aléatoire simple, colloqué et de Poisson, seront basées sur des principes de rotation.

La dernière section ne fournira que quelques idées sur des méthodes non-mathématiques visant à réduire le travail du répondant. Ainsi, on verra qu'il est possible de réduire le fardeau de réponse en utilisant un seul questionnaire, que l'on dira unifié pour plusieurs enquêtes, ou simplement en envoyant les questionnaires à chaque répondant dans une seule enveloppe.

Tout au long de ce rapport on fera une distinction entre "enquête" et "sondage". L'enquête se définit comme l'étude d'une certaine population par l'entremise de sous-études régulières. Chacune de ses sous-études constitue un sondage. Ainsi donc, chaque enquête exige généralement la tenue de sondages périodiques. Par exemple, SEPH, MRTS et MRCTS, dont font mention Leroux et Lussier [7] dans leur étude, sont trois enquêtes qui demandent à ce qu'un sondage soit effectué tous les mois.

2 QUELQUES NOTIONS SUR LE FARDEAU DE RÉPONSE

2.1 INTRODUCTION

Cette section présente comment peut se définir le fardeau de réponse, quels en sont les effets sur les répondants et donne une suggestion de quelques critères sur lesquels son évaluation pourrait être basée. On y établit ainsi les notions qui seront utilisées dans les sections suivantes.

2.2 DÉFINITION

Lorsqu'il est demandé à un répondant de compléter un questionnaire, celui-ci doit utiliser certaines ressources en plus de faire face à une dépense d'énergie occasionnée par les recherches, les efforts voués à la compréhension des questions, les demandes auprès des services comptables, qu'ils soient internes ou externes, etc. Ceci introduit donc une notion de fardeau de réponse auquel toutes les entreprises échantillonnées sont confrontées.

Ce fardeau, comme le suggère Parlor [8], peut se diviser en deux grandes parties: la partie brute et la partie bénifique. La partie brute est constituée des facteurs qui peuvent décourager le répondant à compléter un questionnaire, qu'il s'agisse d'une dépense économique, du temps consacré à remplir le questionnaire, d'un manque de motivation, d'une haine envers le gouvernement ou d'une non-confiance en Statistique Canada pour ce qui est de la confidentialité des informations, etc. La partie bénifique, elle, comprend tous les avantages dont le répondant peut bénéficier. Il pourrait par exemple, apprendre à mieux connaître son entreprise en répondant à des questions qu'il ne s'était jamais posées auparavant. Il peut aussi profiter des résultats finaux de l'enquête pour orienter avantageusement ses transactions, ajuster le prix de ses biens et services, modifier sa production, etc.

Malgré l'existence de ces deux composantes, on ne se préocupe que de la partie brute dans cette étude. Elle constituera donc le seul aspect considéré dans les développements qui suivent.

2.3 EFFETS

Lorsque le répondant perçoit le fardeau de réponse comme étant trop élevé, il se voit attribuer une tâche plus grande que celle à laquelle il voudrait bien faire face. Ceci peut provoquer chez lui un découragement et/ou une baisse de motivation qui peuvent le pousser à fournir des informations incorrectes, incomplètes ou en retard et peut-être même une combinaison des trois. Pour Statistique Canada, il devient alors primordial de minimiser ces effets car ils pourraient fausser ses estimés en plus de retarder la sortie de ses résultats.

Un contrôle du fardeau de réponse profiterait donc à l'enquêteur comme à l'enquêté.

2.4 ÉVALUATION

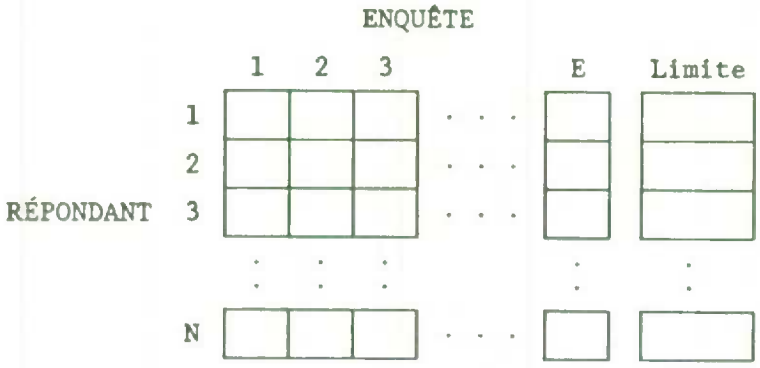
Une façon de contrôler le fardeau imposé aux répondants, débuterait par l'attribution d'un poids évaluant leur perception du fardeau de réponse pour chacune des enquêtes. Mais de telles évaluations sont loins d'être faciles puisque certaines composantes du fardeau de réponse comme le manque de motivation ou la non-confiance envers le gouvernement sont très difficiles à mesurer.

On évaluerait aussi pour chaque répondant, une borne maximale du fardeau qu'il pourrait accumuler durant une période de temps préétablie P. Ces valeurs constitueraient ainsi des limites à ne pas dépasser. Elles pourraient, par exemple, être accordées de façon proportionnelle aux revenus des entreprises, mais encore là, ces évaluations ne sont pas évidentes.

Toutefois, pour les besoins de la prochaine section, portant sur le contrôle du fardeau de réponse, on considérera que ces valeurs sont connues.

Ainsi, supposons que l'on ait à travailler avec E enquêtes et que l'union des E populations correspondantes forme une superpopulation de taille N. Les sections qui suivent supposeront alors que les E·N perceptions du fardeau et les N limites de la matrice suivante (introduite initialement par Parlor [8]), sont connues.

Figure 1: Matrice du fardeau de réponse



Par exemple, si on n'avait seulement que les enquêtes SEPH et MRTS (E=2), on aurait, selon Leroux et Lussier [7], un recouvrement des deux populations comme suit:

La matrice de la figure 1 serait alors de taille 2·N, où N = 724,020 (539,154 + 135,867 + 48,999). Dans un tel cas, plusieurs entrées de la matrice seront nulles puisque 588,153 répondants ne sont compris que dans une population (en principe on devrait n'avoir que 859,887 entrées non-nulles).

Remarquons que toutes ces évaluations pourraient être faites en regroupant les répondants. Nous aurions alors moins de travail à faire et nous pourrions travailler avec une matrice réduite. Ainsi, on rassemblerait les N individus de la superpopulation pour former G groupes homogènes selon un critère prédéterminé. Cette partition pourrait être basée sur des classes de revenu, des classes d'emploi, les provinces ou bien l'appartenance au IP ou au NIP. A la limite, elle pourrait même considérer un seul groupe. La figure 2 illustre une telle matrice réduite, dans le contexte où seules les enquêtes SEPH et MRTS sont présentes.

Figure 2: Matrice réduite du fardeau de réponse pour SEPH et MRTS

		SEPH	MRTS	Limite
GROUPE	1			
	2			
	.	.	.	.
	.	.	.	.
	G			

Mais pour être plus général, les dévelopements qui suivent feront référence à la matrice non-réduite de la figure 1.

3 CONTRÔLE DU FARDEAU DE RÉPONSE

3.1 INTRODUCTION

Le contrôle du fardeau de réponse se définit comme étant une façon d'éviter que des répondants n'aient à subir plus qu'un fardeau de réponse prédéterminé sur une période de temps donnée. Dans un premier temps, quelques critères seront proposés afin d'évaluer les limites de fardeau dans le contexte matriciel de la section précédente. Par la suite, deux méthodes seront introduites dans le but de respecter ces limites.

3.2 RESPECT DES LIMITES

3.2.1 Introduction

Les limites de fardeau, associées à chaque répondant, ont été définies, dans la section 2, sur une certaine période de temps P. Ainsi, lorsqu'on veut savoir si le fardeau d'un répondant tel qu'il a été défini par la matrice de la figure 1, respecte les limites préétablies, il faudrait le cumuler sur un intervalle de temps de longueur P. De cette façon, le fardeau imposé avant cette période ne serait pas considéré.

3.2.2 Principe d'éligibilité

Peu importe la méthode d'échantillonnage, il est possible de respecter les fardeaux limites préétablis. Une des façons d'y arriver consiste à considérer comme étant non-éligibles à l'échantillonnage, les unités ayant déjà atteint leur fardeau limite (bien entendu, cette limite ayant déjà été définie pour un intervalle de temps, les unités non-éligibles à l'échantillonnage ne le seront que pour une période bien déterminée). On n'échantillonnera alors que parmi les unités ayant un fardeau accumulé encore inférieur à leur limite, en supposant qu'il y en ait assez.

Malheureusement, une telle méthode pourrait biaiser les résultats, et l'importance de ce biais ne dépendrait essentiellement que du choix des limites. Par exemple, si les limites sont proportionnelles aux revenus des entreprises, les petites entreprises auront tendances à devenir non-éligibles plus souvent que les grosses, ce qui augmenterait sensiblement les estimés de production ou de ventes. Donc, pour que ce principe de non-éligibilité

n'engendre aucun biais, il ne devrait y avoir aucun lien de dépendance entre les estimés produits et les critères de base servant à l'évaluation des limites à l'intérieur de chacune des strates. Un autre désavantage consiste en l'augmentation des probabilités de sélection pour les unités encore éligibles. Finalement, un tel principe est difficilement applicable dans de petites strates puisqu'elles comprennent trop peu d'unités pour en rendre non-éligibles.

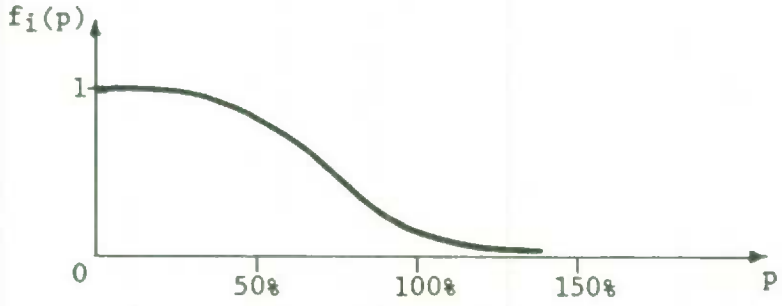
3.2.3 Réduction des probabilités d'échantillonnage

Une autre façon de respecter les limites, a été donnée par Sunter [10]. Il s'agit d'appliquer un facteur de réduction aux probabilités de sélection des unités lorsque leur fardeau accumulé (sur une période de temps P) s'approche de leur maximum. Par exemple, si pour l'entreprise *i* on designe par p la proportion du fardeau limite utilisée i.e.

p = fardeau accumulé / fardeau limite,

on aurait un facteur de réduction $f_i(p)$, fonction de p. La figure 3 donne un exemple d'une telle fonction.

Figure 3: Représentation d'un facteur de réduction



Soulignons ici que le choix de ces fonctions est purement subjectif. Nous aurions aussi très bien pu prendre une seule et même fonction f(p) pour tous les répondants. Tout au moins, il est naturel de choisir un $f_i(p)$ près de 1 lorsque p est petit, et près de 0 lorsque p est grand pour que la probabilité d'inclusion soit gardée intacte au départ et qu'elle diminue lorsque le fardeau de l'unité s'approche ou dépasse le fardeau limite fixé.

Toutefois, un tel principe apporte les mêmes problèmes que le principe d'éligibilité. Il entraine un accroissement

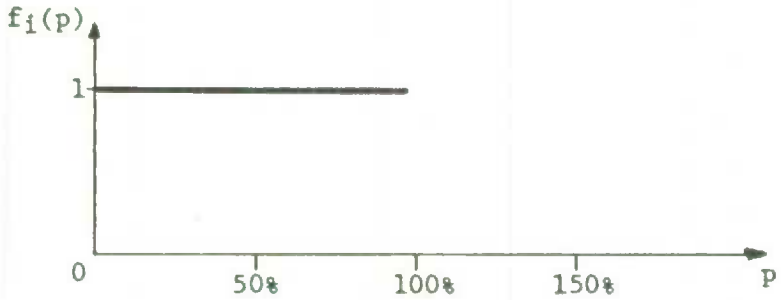
des probabilités de sélection pour les unités encore éligibles (si on veut garder des tailles échantillonales constantes) et il peut créer un biais dans les estimés, lorsque les limites sont mals définies.

Il est facile de remarquer que le principe d'éligibilité, défini en 3.2.2, n'est qu'un cas particulier de la présente technique. En effet, il peut être représenté par

$$f_i(p) = f(p) = \begin{cases} 1 & \text{si } 0 \leq p < 1 \\ 0 & \text{sinon.} \end{cases}$$

La figure 4 illustre cette fonction.

Figure 4: Facteur de réduction pour l'éligibilité à l'échantillonnage



Un avantage statistique dans l'utilisation des facteurs de réduction est la possibilité d'éliminer l'exclusion systématique. Ceci se fait avec des fonctions $f_i(p)$ décroissantes mais toujours supérieures à 0.

En comparant ces deux principes on peut s'appercevoir que le critère d'éligibilité est plus facile à mettre en oeuvre parce qu'il ne laisse pas de place à la subjectivité comme c'est le cas dans le choix des facteurs de réduction. Nous verrons en plus dans les prochaines sections que, contrairement au principe d'éligibilité, l'utilisation des facteurs de réduction est souvent très difficile à concevoir.

4 PRÉVENTION DU CHEVAUCHEMENT ENTRE LES ÉCHANTILLONS DE DIFFÉRENTES ENQUÊTES

4.1 INTRODUCTION

Lorsqu'on veut réduire le fardeau de réponse des entreprises, on peut le faire entre les enquêtes ou à l'intérieur des enquêtes. Ainsi, réduire le fardeau inter-enquêtes consiste essentiellement à diminuer les possibilités de sélection d'un répondant par plusieurs d'entres elles tandis que la réduction du fardeau intra-enquête se fait surtout à l'aide de rotations échantillonnales.

Cette section a pour but de définir une méthode visant à réduire le fardeau inter-enquêtes (le fardeau intra-enquêtes sera traité dans la section suivante). Pour ce faire, elle décrit le fonctionnement de l'échantillonnage synchronisé.

4.2 ÉCHANTILLONNAGE SYNCHRONISÉ

L'échantillonnage synchronisé a pour but de prévenir la sélection d'une unité d'échantillonnage par différentes enquêtes. Il synchronise donc les échantillons afin de minimiser leur chevauchement.

Soit une population de taille N, séparée en H strates de $N_1, N_2, \ldots, N_H$ unités respectivement. Pour procéder à plusieurs échantillonnages simultanés dans la strate h, Hinde et Young [6] proposent d'associer à chacune des N_h unités un nombre aléatoire compris entre 0 et 1, que l'on reporte ensuite sur l'intervalle [0,1). On arrive ainsi à ordonner aléatoirement ces unités, comme le montre la figure 5. Si par la suite, une nouvelle unité venait qu'à se joindre aux autres, on n'aurait qu'à lui associer un nombre aléatoire pour ainsi l'ordonner par rapport au reste de la population.

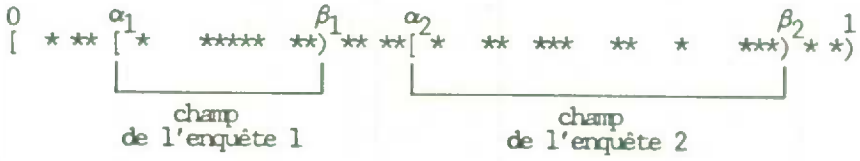
Figure 5: Représentation des N_h unités d'une certaine strate sur l'intervalle [0,1)

```
0                                                                  1
[ * **  *   ***** ** ** ** *  ** ***  **  *   *** * *)
```

Hinde et Young proposent ensuite de déterminer, pour chacune des enquêtes, un champ $[\alpha,\beta)$ inclus dans $[0,1)$, qui représenterait les unités susceptibles d'être échantillonnées. L'échantillon de chaque enquête sera ainsi tiré aléatoirement dans le champ $[\alpha,\beta)$ correspondant. Ces champs devront être assez grands pour permettre des rotations échantillonnales convenables, en plus de présenter le moins de chevauchements possibles afin d'éviter la possibilité qu'une unité soit sélectionnée par plusieurs enquêtes.

La figure 6 montre comment deux champs peuvent être déterminés pour deux enquêtes.

Figure 6: Représentation de deux champs pour deux enquêtes dans une certaine strate h



Un obstacle auquel on devra faire face si on utilise une telle méthode est la possibilité d'avoir des stratifications différentes pour chacune des enquêtes. Dans un cas pareil, il suffit de ne pas faire toute la stratification avant la détermination des champs $[\alpha,\beta)$ et ainsi s'en garder une partie pour après.

Une première façon d'y arriver serait de trouver le plus haut niveau de stratification, commune à toutes les enquêtes (par exemple: SIC et provinces) afin de déterminer les champs $[\alpha,\beta)$. Il suffirait ensuite de terminer la stratification à l'intérieur de ces champs pour satisfaire chacune des enquêtes.

Une deuxième façon, un peu plus simple, considérerait d'abord $H = 1$ (i.e. la population entière) pour déterminer les champs et n'incorporerait la stratification que par après, à l'intérieur de ceux-ci.

Ces deux posibilités augmenteraient la complexité de l'échantillonnage mais pourraient accommoder toutes les enquêtes.

Quoi qu'il en soit, il reste quand même que la mise sur pied d'un processus permettant de déterminer ces champs, pourrait nous causer beaucoup de problèmes. En effet, si les sondages totalisent une taille échantillonnale supérieure à la taille de la strate, il est impossible d'éliminer les intersections entre chaque champ. De plus, les morts et naissances peuvent faire changer le nombre d'unités dans chaque champ pour ainsi les débalancer. De même, étant donné que la longueur d'un champ devrait être liée à l'importance de l'enquête correspondante, sa détermination devient quelque chose de bien délicat à faire.

Il existe évidemment plusieurs façons d'échantillonner à l'intérieur de ces champs. La section qui vient en proposera trois et introduira pour chacune, une rotation permettant de répartir le fardeau à l'intérieur d'une enquête (ou d'un champ).

5 RÉPARTITION DU FARDEAU DE RÉPONSE À L'INTÉRIEUR D'UNE ENQUÊTE

5.1 INTRODUCTION

Faisons ressortir, pour commencer, la différence entre contrôle et répartition. Comme on l'a déjà expliqué, le contrôle vise à limiter les accumulations de fardeau pour chaque répondant. La répartition, elle, a pour but de ne favoriser aucun répondant en ce qui a trait aux accumulations de fardeau. Elle doit donc changer l'échantillon d'un sondage à l'autre, afin d'accroître le nombre d'unités enquêtées et ainsi uniformiser le fardeau de chacune d'elles.

Il est vrai que le choix d'un même échantillon pour plusieurs sondages consécutifs nous fournit de meilleurs estimés des tendances et minimise les coûts d'échantillonnage et de formation du répondant. Par contre, il diminue l'apport de l'échantillon à la mise à jour de la base de sondage et il ne concentre le fardeau que sur quelques répondants. D'un autre côté, le choix d'un échantillon totalement différent à chacun des sondages ne semble pas meilleur puisqu'il ne fait qu'inverser les qualités et les défauts de l'autre. Il s'agirait donc de faire un juste choix pour que, de sondage en sondage, une partie de l'échantillon précédant soit gardée et que l'autre partie soit renouvelée. On dira alors qu'on effectue des rotations R, où $0 \leq R \leq 1$ représente la proportion de l'échantillon remplacée d'un sondage à l'autre.

Dans ce qui suit, nous traiterons trois méthodes d'échantilonnage que l'on essayera d'intégrer à la matrice de fardeau de la section 2. Ces méthodes pourraient être utilisées à l'intérieur des champs $[\alpha,\beta)$ de l'échantillonnage synchronisé et auraient comme but de répartir le fardeau à travers la population. Cette répartition sera dite uniforme si tous les répondants accumulent leur fardeau à une vitesse à peu près égale. On aura une répartition équitable si leur accumulation est proportionnelle à leur limite (voir figure 1). Dans le contexte du fardeau de réponse, il semble bien que la répartition équitable soit préférable à la répartition uniforme.

On expliquera ensuite pour chaque méthode d'échantillonnage comment effectuer des rotations, et on fera ressortir les avantages et désavantages de chacune. On terminera la section par un résumé visuel des caractéristiques de ces trois méthodes.

5.2 RÉPARTITION À PARTIR D'UN ÉCHANTILLONNAGE ALÉATOIRE SIMPLE

5.2.1 Description de l'échantillonnage

Supposons qu'on veuille tirer un échantillon de taille n_h dans une strate de N_h unités réparties aléatoirement le long de $[\alpha,\beta)$, comme à la figure 6. Il suffit alors de choisir aléatoirement un point D dans $[\alpha,\beta)$, et de déterminer un point F à la droite de D pour que l'intervalle de sélection [D,F) comprenne exactement n_h unités. Les points D et F définiraient alors le début et la fin de l'échantillon respectivement. La figure 7 permet de visualiser cet échantillonnage.

Figure 7: Représentation d'un échantillon de taille n_h=6

Si notre enquête exige la tenue de sondages périodiques, le même échantillon sera choisi à chacun des sondages, à moins que des naissances ou des morts ne viennent changer le nombre d'unités incluses dans [D,F). Il est donc utile de se donner une rotation bien définie qui permettrait de répartir le fardeau entre les N_h unités.

5.2.2 Répartition du fardeau par rotation

Il y a plusieurs principes de rotation pouvant s'appliquer à un échantillon [D,F) comme celui qu'on vient de décrire. En effet, Hidiroglou & Srinath [5] ont mis sur pied un système, basé sur des panels, qui permet entre autre l'obtention de pourcentages de rotation constants et qui traite facilement les morts et naissances. Notons que cette méthode des panels sera celle utilisée par les enquêtes-entreprises de Statistique Canada. De plus, Tambay [11] transformait cette méthode pour que le pourcentage de rotation soit déterminé à partir du temps maximum qu'un répondant peut passer dans l'échantillon et du temps minimum qu'il peut ne pas y être, en ayant des unités préabablement ordonnées de façon équidistante sur [0,1). Par contre, sa méthode traitait difficilement les morts et naissances.

Etant donné que ces deux méthodes ont déjà été traitées à Statistiques Canada et qu'elles sont un peu complexes, nous avons choisi d'en décrire une plus simple. Il s'agit de celle proposée par Hinde et Young [6], dont le principe de base reste quand même semblable aux deux autres.

Supposons qu'à chaque sondage, on veuille renouveler une partie de l'échantillon pour que ce dernier soit complètement renouvelé au moins à tous les k sondages. Ceci sous-entend nécéssairement que le nombre d'unités contenues dans notre champ est au moins deux fois supérieur à notre taille échantillonnale. On aurait alors une rotation moyenne de $R_h = 1/k$ à chacun des sondages.

Ainsi donc, si nous voulons que l'échantillon soit totalement renouvelé à tous les k sondages, il suffit de procéder comme suit. Commençons par déterminer K le plus grand entier inférieur ou égal à k et divisons l'intervalle de sélection [D,F) en K blocs à l'aide de K-1 séparations. Nous pouvons alors effectuer une rotation à chacun des sondages tout simplement en décalant cet intervalle d'un bloc vers la droite. Ceci s'obtient en déplaçant D vers la droite jusqu'à la première séparation et F, encore vers la droite, d'autant d'unités que D. Lorsque l'intervalle [D,F) arrivera au bout de $[\alpha,\beta)$, il suffira de continuer la rotation à partir de α comme si le tout formait une boucle fermée. Si des naissances ou des morts venaient changer la taille échantillonnale n_h, on n'aurait qu'à inclure plus ou moins d'unités que prévu dans le nouveau bloc de droite. La figure 8 montre à peu près toutes les possibilités qui peuvent survenir.

figure 8: Représentation de la rotation avec K=3 et n_h=6

```
                α                D                F                                                β
sondage 1       [  * *[  **|*  *|* *)   * * *   *  *  *   * * *     *  ** *)
                                 └────┴──────séparations
                α            D                F                                                β
sondage 2       [  * *   **[*  *|* *|   * *)*   *  *  *   * * *     *  ** *)
                α                D                F                                                β
sondage 3       [  * *   ** *  *[* *|   * *|*   * )*   *   * * *     *  ** *)
                α                                D                F                                β
sondage 4       [  * *   ** *  * * *[   * *|.   * |*   *   *)* *     *  ** *)
                                        mort└──┘
                α                                D                F                                β
sondage 5       [  * *   ** *  * * *   * *[   . |.   **   *|* *     *)  ** *)
                                        morts└──┘└┘   └──naissance
                α                                D                F                                β
sondage 6       [  * *   ** *  * * *   * *     [   **   *|* *     *| )** *)
                α                                                D                F                β
sondage 7       [  * *   ** *  * * *   * *             **   *[***     *| |**)*)
                                                                 └──naissance
```

5.2.3 Caractéristiques de la rotation

L'utilisation d'une méthode comme celle que l'on vient de décrire comporte certains avantages et désavantages. Premièrement, le développement d'un système informatique pour mettre en oeuvre une telle méthode serait assez simple, ce qui constitue un atout certain. De plus, toutes les unités du champ $[\alpha,\beta)$ auront été sélectionnées après un certain nombre de sondages. Alors si la réunion des champs correspond à la population entière, on pourra mieux tirer profit de l'échantillonnage pour mettre à jour la base de sondage.

D'un autre côté, le fait que deux unités voisines sur l'intervalle $[\alpha,\beta)$ soient presque toujours sélectionnées ensembles, nous porte à croire que cette méthode reproduit des échantillons cycliques très semblables. Cette similitude sera d'autant plus forte si des cycles rapides sont obtenus à l'aide d'un petit K et d'une fraction échantillonnale n_h/N_h élevée. Aussi, le pourcentage de rotation est loin d'être constant. Pour s'en convaincre, il suffit de remarquer que

les échantillons des sondages 4, 5, 6 et 7 de la figure 8 sont obtenus avec des rotations respectives de 50%, 67%, 0% et 50%. Il devient alors difficile de donner des estimés de tendance.

Un autre côté négatif réside dans le fait que, en sélectionnant chaque unité K fois de suite, le fardeau sera réparti uniformément entre les unités. On considère cette caractéristique comme étant désavantageuse puisque ce qu'on souhaiterait obtenir, c'est plus une répartition équitable du fardeau par rapport aux limites de chacun. Par exemple, si l'unité A possède une limite deux fois plus élevée que l'unité B, il serait normal que son accumulation de fardeau se face deux fois plus rapidement que celle de B. Les deux unités s'approcheraient ainsi de leur limite à des vitesses relativement égales. Mais tel n'est pas le cas dans cette rotation. Finalement, un surplus de naissances peut ne pas toujours être épongé par le bloc de droite. On en aurait eu un exemple si, au sondage 6 de la figure 8, on avait eu une naissance.

5.2.4 Respect des limites

Si on veut respecter les limites à l'aide du principe d'éligibilité défini en 3.2.2, il faudrait cumuler le fardeau des répondants sur un intervalle de temps P comme il a été expliqué à la section 3. Les répondants devant être rendus non-éligibles à l'échantillonnage, dû à des accumulations de fardeau supérieures à leur limite, pourraient alors être considérés comme des morts. Après un certain temps, ils pourront revenir éligibles grâce à des naissances.

Pour ce qui est du principe de réduction des probabilités proposé en 3.2.3, remarquons qu'il est difficile à appliquer pour une telle rotation puisque les répondants sont sélectionnés selon leur position dans une suite d'unités ordonnées, et non pas selon une expérience aléatoire à chacun des sondages.

5.3 RÉPARTITION À PARTIR D'UN ÉCHANTILLONNAGE COLLOQUÉ AVEC TAILLES FIXÉES

5.3.1 Description de l'échantillonnage

Considérons encore une fois une population de taille N séparée en H strates de N_1, N_2,..., N_H unités respectivement. L'idée principale de l'échantillonnage colloqué est

d'ordonner les unités dans chaque strate d'une façon prédéterminée afin d'en tirer les n_h premières pour les inclure dans l'échantillon, d'où le mot colloqué.

Dans le contexte actuel, cette mise en ordre sera faite de façon croissante par rapport aux accumulations de fardeau. En choisissant les n_h premières unités, nous sélectionnerons ainsi les répondants ayant les plus bas fardeaux accumulés.

De Ree [3] propose une méthode qui consiste premièrement à assigner à chaque unité d'échantillonnage (dans notre champ $[\alpha,\beta)$) un nombre aléatoire compris entre 0 et 1. Cette valeur représente alors un fardeau initial fictif dont le but est de briser les égalités possibles entre deux répondants. En ordonnant la population de façon croissante selon ces nombres aléatoires, nous pouvons tirer un premier échantillon en choisissant les n_h premières unités de chaque strate. En ajoutant ensuite pour chaque répondant, le fardeau de réponse de ce premier sondage et en mettant encore en ordre croissant les unités selon leur fardeau accumulé, on arrive à l'échantillonnage colloqué avec tailles fixées en sélectionnant les n_h premières unités de chaque strate. En répétant cette opération pour chaque sondage, les échantillons choisis seront composés des répondants les moins éprouvés de chaque strate.

Remarquons que cette méthode peut se faire en cumulant le fardeau pour différentes enquêtes, mais ceci exigerait qu'elles aient une stratification commune. Par exemple, si SEPH et MRTS avaient une stratification commune, les unités appartenant aux deux populations pourraient se voir accroître leur fardeau par les deux enquêtes.

Toutefois, un échantillonnage semblable ne répond pas nécéssairement à nos attentes quant au choix du pourcentage de rotation, même s'il fournit une excellente répartition du fardeau. En effet, il ne laisse aucune liberté dans ce choix. Voyons comment il est possible d'obtenir un pourcentage de rotation bien précis.

5.3.2 Répartition du fardeau par rotation

Supposons que l'on veuille obtenir une rotation R_h dans la strate h (ceci signifie que $R_h \cdot 100\%$ des unités sélectionnées dans cette strate seront remplacées à chaque sondage). Supposons aussi que la taille N_h de cette strate soit assez grande pour permettre une telle rotation, i.e. $R_h \cdot n_h \leq N_h - n_h$. En ordonnant toujours les unités selon leur

fardeau accumulé, de Ree [3] propose de garder les $(1-R_h)\cdot n_h$ premières unités de l'échantillon précédant et de remplacer les $R_h\cdot n_h$ autres par les unités non-échantillonnées les moins imposées au point de vue du fardeau.

5.3.3 Caractéristiques de la rotation

La programmation de cette méthode de rotation est relativement simple, même si elle doit contenir une procédure de tri. De plus, après un certain nombre de sondages, toutes les unités de notre population auront été sélectionnées. Ainsi on pourra utiliser l'échantillonnage pour mettre à jour la base de sondage. La rotation permet aussi d'obtenir facilement le pourcentage de rotation désiré lorsque N est suffisament grand. Par contre, cette rotation cherche trop à répartir le fardeau uniformément à travers la population, plutôt que d'en faire une répartition équitable par rapport aux limites de chacun. Toutefois nous verrons plus loin qu'il est possible d'en arriver à une distribution équitable, avec une variante très simple.

Il est bon de mentionner qu'une telle rotation répartira bien dans le temps le fardeau d'un répondant si R_h est grand, tandis que s'il est petit, son fardeau sera plutôt concentré sur plusieurs sondages consécutifs. Il revient à l'utilisateur de choisir les R_h qui lui semblent les plus avantageux.

5.3.4 Respect des limites

Comme pour l'échantillonnage synchronisé, il serait facile de cumuler le fardeau imposé à chaque unité, tel que défini par la matrice de la figure 1. On appliquerait alors le principe d'éligibilité simplement en sélectionnant le nombre d'unités voulues parmi celles de la population encore éligibles à l'échantillonnage. Par contre, avec la rotation proposée, les unités d'une même strate auraient tendance à avoir une croissance assez semblable de leur fardeau cumulé. Alors, lorsqu'une unité deviendrait non-éligible, une quantité d'autres le deviendraient aussi, si les limites sont définies par groupe (voir figure 2). On verrait donc des blocs d'unités devenir non-éligibles à l'échantillonnage ce qui pourrait biaiser nos estimés par manque de représentants de certains groupes. Pour ce qui est du principe de réduction des probabilités (3.2.3), il serait encore une fois difficile à appliquer pour les mêmes raisons que l'on a vues dans la rotation décrite précédemment.

5.3.5 Variante de la rotation

Il serait possible de changer quelque peu la rotation afin qu'elle entraine une distribution équitable du fardeau. Ceci réglerait donc le problème soulevé en 4.3.3, en faisant croitre les accumulations de fardeau à des vitesses moyennes relativement égales. Une façon d'y arriver serait de transformer les entrées de la matrice de fardeau, avant de les cumuler. On choisirait ainsi une transformation qui favoriserait les unités ayant de petites limites. Par exemple, diviser le fardeau de réponse par les limites fixées serait un bon moyen d'y arriver. Les limites seraient ainsi toutes ramenées à une valeur standard de 1. Aussi, en échantillonnant les unités selon ce nouveau fardeau accumulé, la rotation aurait pour effet de répartir le fardeau d'une façon équitable selon les limites de chacun.

Cette variante aurait les mêmes caractéristiques que la rotation originale sauf qu'elle n'aurait plus le désavantage de répartir uniformément le fardeau entre les répondants.

5.4 RÉPARTITION À PARTIR D'UN ÉCHANTILLONNAGE DE POISSON

5.4.1 Description de l'échantillonnage

Cet échantillonnage considère une population de N unités de laquelle on voudrait tirer un échantillon de taille espérée $\bar{n}$. Nous disons ici taille espérée puisque le nombre d'unités contenues dans l'échantillon est une variable aléatoire.

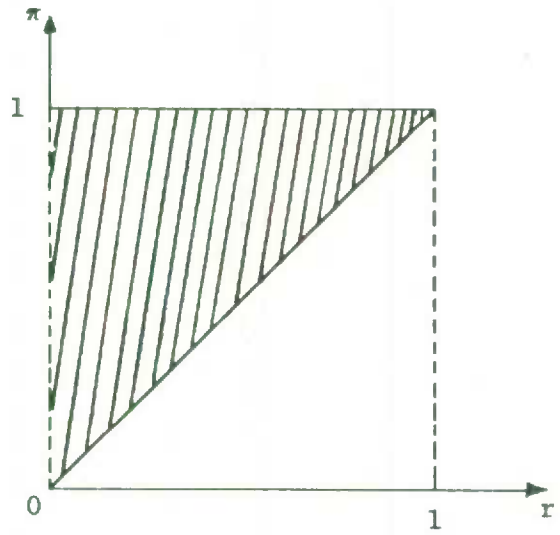
L'échantillonnage de Poisson consiste à selectionner chaque unité selon une probabilité π_i qui lui est propre. Bien que ces probabilités sont souvent définies selon l'importance des unités dans notre enquête, on s'approcherait déjà de notre objectif en choisissant des π_i proportionnels aux limites puisque les questionnaires seraient ainsi répartis équitablement à travers la population. Quoi qu'il en soit, ces π_i seront choisis de sorte que

$$\sum_{i=1}^{N} \pi_i = \bar{n}$$

Une façon simple de déterminer si l'unité i fera partie de l'échantillon est de lui accorder un nombre aléatoire r_i

de loi U(0,1), et de la sélectionner si $r_i \leq \pi_i$. Précisons ici que ces r_i, définis de façon aléatoire, resteront fixes pour tous les sondages à venir. De façon géométrique, ceci correspond à sélectionner l'unité *i* si le couple (r_i, π_i) est dans la partie hachurée de la figure suivante.

Figure 9: Région de sélection pour l'échantillonnage de Poisson



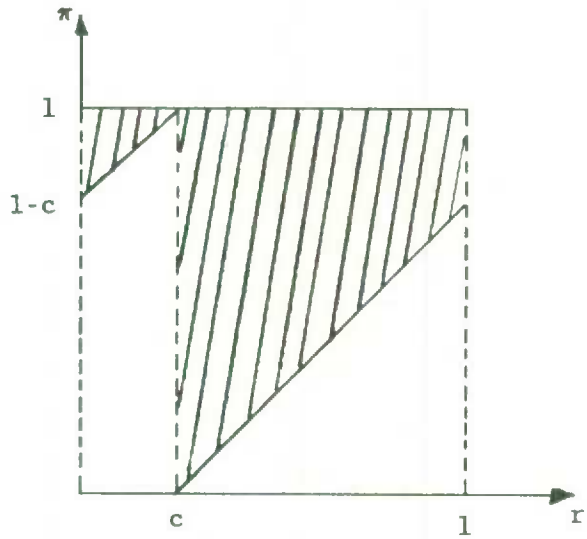
Sunter [10] et Brewer, Early & Joyce [2] donnent en appendice de leur article, les principaux estimateurs de total utilisés avec cet échantillonnage, ainsi que leur variance.

5.4.2 Répartition du fardeau par rotation

Le critère de rotation est basé ici aussi sur la proportion de renouvellement R de l'échantillon à chacun des sondages, comme c'était le cas pour l'échantillonnage colloqué. Il suffit de transformer la région de sélection pour que certains répondants du sondage précédant soient remplacés par de nouveaux répondants.

La transformation qui nous semble la plus naturelle, qui représente le mieux l'idée de rotation, qui fournit la meilleur répartition du fardeau et qui simplifie les calculs le plus possible a été décrite par Sunter [9] et est illustrée à la figure 10. Elle consiste à décaler la région de sélection entière, d'une longueur *c* le long de l'axe *r*.

Figure 10: Région de sélection déplacée de c



Ainsi donc l'unité *i* sera dans le deuxième échantillon si (r_i, π_i) est dans la nouvelle partie hachurée. De façon mathématique, ceci consiste à choisir l'unité *i* dans le deuxième échantillon si

$$r_i \leq \pi_i - (1-c) \quad \text{ou bien si} \quad c < r_i \leq \pi_i + c.$$

Il est possible de trouver ce qu'on devrait obtenir comme rotation échantillonnale, lorsqu'on fait un tel déplacement de la région de sélection. Sunter [9] nous indique comment cette rotation espérée $\bar{R}$ peut être exprimée en fonction de *c*. Voici l'équation qu'il établit si $c < 1/2$:

$$(1) \qquad \bar{R} = [\, \bar{n} - \sum_{\{i \mid c<\pi_i \leq 1-c\}} (\pi_i - c) - \sum_{\{i \mid 1-c<\pi_i\}} (2\pi_i - 1) \,] \,/\, \bar{n}$$

Il est à noter ici que la valeur de *c* peut être changée d'un sondage à l'autre pour ainsi permettre l'obtention de différentes rotations, selon l'exigence de chaque sondage. Mais une difficulté survient justement lorsqu'on veut obtenir une rotation R précise et qu'ainsi, on voudrait déterminer le *c* qui nous en permette l'obtention. En effet, il faudrait pouvoir exprimer *c* en fonction de $\bar{R}$. Nous devrions donc isoler *c* dans l'équation (1). Malheureusement, ceci ne peut se faire de façon explicite.

5.4.3 Caractéristiques de la rotation

Une rotation comme celle-ci semble présenter moins d'avantages que celle de l'échantillonnage colloqué.

Pour commencer, l'échantillonnage de Poisson nous donne la chance d'accorder des probabilités de sélection différentes pour chaque unités d'échantillonnage. Ceci nous permet de sélectionner certains répondants plus souvent que d'autres en plus de rendre possible toute restratification, en changeant simplement les π_i. De plus, la rotation nous fournit une règle de sélection assez simple tout en gardant des expressions peu complexes pour les estimateurs et leur variance.

D'un autre côté, le plus grand désavantage consiste en l'obtention d'une taille échantillonnale aléatoire, ce qui peut être particulièrement risqué dans les strates contenant peu d'unités. Par contre, on verra un peu plus loin que cet inconvénient peut être partiellement évité par un choix plus judicieux des r_i. Un autre point négatif vient de la difficulté qu'on peut éprouver lorsque vient le temps d'évaluer le décalage c afin d'obtenir une rotation précise. De même, comme dans l'échantillonnage synchronisé, une similitude cyclique des échantillons peut se produire ici aussi. Elle s'observera par un positionnement presque identique du triangle de sélection, d'un cycle à l'autre.

Aussi les morts, naissances, fusions et divisions d'entreprises sont difficile à traiter. Si on veut garder fixe notre taille échantillonnale espérée, il faudrait presque refaire l'assignation de tous les π_i. On pourrait ajouter à ça, le fait qu'il puisse s'écouler beaucoup de temps avant que toutes les unités soient sélectionnées au moins une fois. Par contre, ceci peut nous sembler moins pire si on remarque que ce sont les unités ayant de petits π_i qui manqueront ainsi à l'appel. En effet, ces unités n'étant parcourues que par la "pointe du triangle de sélection", elles ne seront sélectionnées que très rarement.

5.4.4 Respect des limites

Contrairement aux échantillonnages colloqué et synchronisé, le principe de réduction des probabilités s'adapte très bien à l'échantillonnage de Poisson. Il s'agit toujours de cumuler le fardeau des répondants (sur une période déjà définie) et de diminuer les π_i selon le ou les facteurs de

réduction correspondants. Bien entendu, cette diminution des probabilités occasionnera aussi une diminution de la taille échantillonnale.

Le principe d'éligibilité étant un cas particulier de la réduction des probabilités, il s'appliquerait d'une façon aussi simple.

5.4.5 Obtention de tailles échantillonnales plus stables

Comme l'ont expliqué Brewer, Early & Hanif [1], il est possible, dans certains cas, de rendre les tailles échantillonnales plus stables en choisisant les r_i d'une façon plus méthodique. La méthode qu'ils décrivent fournit des r_i' liés entre eux. Elle consiste à ordonner aléatoirement les unités de la population et de leur accorder des nombres aléatoires r_i' selon leur rang par rapport aux autres. Ainsi, pour la $i^{ème}$ unité on suggère de prendre

$$r_i' = \frac{r + i - 1}{N}$$

où r est un nombre aléatoire compris entre 0 et 1. Cette procédure permet de choisir des r_i' distant de 1/N l'un de l'autre et ainsi de les distribuer uniformément le long de [0,1). En procédant alors à l'échantillonnage de Poisson, on obtiendra encore des tailles aléatoires, mais elles seront généralement plus stables qu'avec des r_i indépendants.

Quant aux estimateurs de total, ils seront identiques à ceux utilisés dans l'échantillonnage classique de Poisson. Seules les expressions pour la variance seront un peu plus compliquées, mais quand même manipulables.

5.5 COMPARAISON DES MÉTHODES DE ROTATION

Nous présentons ici un tableau récapitulatif des différentes caractéristiques des quatres méthodes de rotation associées à leur échantillonnage respectif. Ce tableau donnera dans un premier temps les avantages de chacune puis ensuite les désavantages.

Pour éviter les répétitions, on ne considérera pas comme un désavantage le fait de ne pas avoir tel ou tel avantage. De même, ce ne sera pas un avantage de ne pas avoir certains

désavantages. Les caractéristiques sont ainsi inscrites dans l'une ou l'autre de ces deux catégories selon l'importance qu'elles y apportent.

Rapellons que la distribution uniforme du fardeau dont il est question dans ce tableau fait référence au cas où les accumulations de fardeau sont semblables pour tous les répondants, tant qu'ils n'atteindront pas leur limite. La distribution équitable, par contre, entraîne des accumulations plus rapides à certains répondants qu'à d'autres, tout dépendant du fardeau limite de chacun.

Tableau 1: Caractéristiques des quatres méthodes de rotation associées à leur échantillonnage respectif

	Rotation associée à l'échantillonnage:			
	ALÉA. SIMPLE	COLLOQUÉ	COLLOQUÉ*	De POISSON
Avantages				
Distribution équitable du fardeau			oui	si les π_i sont proportionnels aux limites
Probabilités de sélection différentes pour chaque unité				possible
programmation facile	oui	oui	oui	oui
Rotation désirée facile à obtenir		oui	oui	
RESPECT DES LIMITES Principe d'éligibilité	possible	possible	possible	possible
Réduction des probabilités				possible
Sélection de toutes les unités après un certain nombre de sondages	oui	oui	oui	exepté celles qui ont de petits π_i
Désavantages				
Distribution uniforme du fardeau	oui	oui		
tailles échantillonnales aléatoires				oui
Similitude cyclique des échantillons	oui			oui

* Avec la variante de la rotation expliquée en 5.3.5

6 DIMINUTION DU FARDEAU DE RÉPONSE

6.1 INTRODUCTION

La présente section donne d'autres alternatives afin de traiter le fardeau de réponse entre les enquêtes. Ainsi, on verra que le contrôle et la répartition du fardeau de réponse ne sont pas les seuls moyens disponibles afin de minimiser les effets néfastes de ce dernier. En effet, l'utilisation de données administratives au lieu de la tenue d'une enquête est très efficace au point de vu du fardeau. Par contre, ceci nécéssite que ces données soient disponibles et qu'elles correspondent à ce qu'on voudrait mesurer.

Des critères de précision moins sévères ou l'utilisation d'une méthode d'estimation plus précise constituerait une autre solution. Elle nous permettrait de réduire les tailles échantillonnales, au profit du répondant.

Une notion d'alternance pourrait aussi être envisagée lorsqu'une unité est sélectionnée par plus d'une enquête. Une même unité fournirait les données du mois M à l'enquête X et celles du mois M+1 à l'enquête Y, suivant un principe d'alternance pour les mois subséquents. Cette notion pourrait être utilisée lorsque la croissance économique est jugée relativement stable dans une industrie.

De plus, il est possible de réagir dès la conception des questionnaires. Il s'agirait alors de créer un questionnaire commun pour plusieurs enquêtes. C'est ce qui est expliqué ici, dans la stratégie des questionnaires unifiés ou par une mise à la poste commune des différents questionnaires.

6.2 STRATÉGIE DES QUESTIONNAIRES UNIFIÉS

Cette stratégie est utile lors de la sélection d'échantillons pour plusieurs enquêtes. Elle consiste à n'envoyer qu'un seul questionnaire aux unités échantillonnées par plus d'une enquête.

En 1982, la Division du travail à Statistiques Canada concevait un questionnaire pilote [4], dont la composition différait d'un répondant à l'autre. Ces questionnaires étaient formés de blocs de questions, adaptés à chacun des répondants. Les commentaires des répondants furent très positifs. Malheureusement, un manque de ressources les empêcha de pousser plus loin son utilisation. Aussi, des responsables du système de développement des enquêtes

généralisées [12] en sont actuellement à perfectionner une idée semblable. Ils travaillent effectivement à la création d'un système pouvant construire des questionnaires par blocs de questions, et ce de façon automatique tout en offrant la possibilité de changer les questions en cours de route. Leurs blocs sont choisis en regard avec le status IP/NIP des répondants, leur code de classification industrielle, etc.

En s'inspirant de cette idée, on peut créer, à partir de blocs de questions, des questionnaires fait sur mesure pour chacun des répondants. Ainsi, chaque bloc correspondrait à une enquête donnée. Le questionnaire qu'on enverrait à un répondant serait alors construit afin de satisfaire les enquêtes pour lesquelles il a été échantillonné.

Supposons par exemple qu'une entreprise *i* est sélectionnée par SEPH, MRTS et MRCTS. On n'aurait qu'à faire sortir les trois blocs de questions correspondants pour les mettre sur un même questionnaire qu'on lui ferait parvenir. On ajouterait à ça, un bloc supplémentaire servant à son identification.

Ces blocs devraient nécéssairement être conçus pour qu'aucune question ne puisse figurer à plus d'un endroit sur le questionnaire. On aura finalement un questionnaire unifié pour SEPH, MRTS et MRCTS qui serait moins long que les trois questionnaires initiaux mis ensembles. Il diminuerait ainsi le fardeau de réponse des unités appartenant aux trois échantillons. Voyons comment l'utilisation de ce questionnaire pourrait être plus avantageuse que des questionnaires indépendants.

Etant donné qu'un questionnaire unifié diminue beaucoup le travail des répondants, le fardeau total imposé par les E enquêtes disons, serait moindre avec l'utilisation d'un tel questionnaire. Cette différence sera d'autant plus élevée si les chevauchements des échantillons sont importants. Ainsi, ne serait-il pas plus avantageux de maximiser ces chevauchements pour maximiser par la suite, le temps non-enquêté par chacune des enquêtes? Une telle pratique minimiserait à long terme le fardeau de réponse imposé à l'ensemble des répondants. Le nombre d'unités enquêtées serait par le fait même minimisé, procurant ainsi des économies certaines pour Statistique Canada. De plus, une telle statégie est réalisable même si toutes les enquêtes ont leur propre stratification. Pour ce qui est de la collaboration des enquêtes, le projet de remaniement des enquêtes-entreprises nous offre l'opportunité de les rendre possibles.

Par contre, on peut imaginer une difficulté qui pourrait survenir dans une telle évantualité. La maximisation des chevauchements est un aspect qui se complique rapidement lorsque le nombre E d'enquêtes est grand. Aussi, une telle éventualité suppose que les enquêtes envoient leurs questionnaires en même temps, ce qui n'est pas toujours le cas. Dans une telle situation, il faudrait que quelques enquêtes acceptent un certain délai dans leur collecte de données. De plus, le fait que les enquêtes n'aient pas toutes la même unité répondante (même si elles ont une même unité statistique) complique un peu la tâche du répondant. Par exemple, les données financières sont l'affaire des comptables tandis que les données de production ne sont fournies que par les contremaîtres. Mais remarquons que ceci n'est sûrement le cas que des grosses entreprises.

6.3 MISE À LA POSTE COMMUNE

Une simplification à envisager consisterait simplement à envoyer, aux unités répondantes faisant partie de plusieurs enquêtes, les différents questionnaires dans une même enveloppe plutôt que d'utiliser un questionnaire unifié. Cette formule nous épargnerait ainsi la conception des questionnaires unifiés, et elle diminuerait quand même la perception du fardeau par les répondants. Elle permettrait en plus aux répondants de retourner leurs questionnaires dans une même enveloppe. Mais encore faut-il que les enquêtes envoient leurs questionnaires en même temps.

7 CONCLUSION

Avec les chiffres fournis par Leroux et Lussier [7] concernant le chevauchement des échantillons de SEPH et MRTS (ils affirment que 10% des entreprises échantillonnées par MRTS le sont aussi par SEPH), on s'appercoit que notre préocupation envers le fardeau de réponse n'est pas un inutile. Aussi, la faisabilité de mise sur pied d'un système visant à réduire le fardeau de réponse nous donne raison sur ce point. Remarquons de plus que cette préocupation donnerait à Statistique Canada, une raison de plus pour obliger les entreprises de répondre à leur questionnaires.

Que ce soit pour garder la motivation du répondant ou bien pour minimiser le biais des estimés de Statistique Canada, un contrôle du fardeau serait évidemment très profitable.

Ce contrôle peut se faire en déterminant les capacités de chaque répondant. Mais encore faut-il pouvoir déterminer le fardeau occasionné par chaque sondage à chacun des répondants.

Des deux principes permettant de contrôler le fardeau de réponse, il semble bien que l'éligibilité à l'échantillonnage soit le plus facilement réalisable et le plus intuitif. Ainsi, le principe de réduction des probabilités, étant trop subjectif et souvent difficile à appliquer, ne satisferait sûrement que bien peu d'utilisateurs.

Mais quelque soit le principe qu'on choisit, il est bon de voir à ce que le fardeau soit bien réparti à travers la population. Ainsi, le fardeau entre les enquêtes peut être réparti par l'échantillonnage synchronisé tandis qu'on peut avoir recours aux rotations échantillonnales pour diminuer les effets du fardeau à l'intérieur d'une même enquête. De telles rotations nous permettent de changer graduellement nos échantillons tout en gardant de bons estimés des tendances.

Si on en vient à calculer le fardeau de façon différente pour chaque répondant (avec une matrice comme celle de la section 2) il semble bien que la rotation transformée de l'échantillonnage colloqué soit celle qui donne la meilleure répartition du fardeau. En effet, elle est la seule qui fournit une distribution équitable du fardeau par rapport aux limites de chaque répondant. Elle facilite aussi l'obtention du renouvellement désiré et n'entraîne aucune similitude cyclique des échantillons. Par contre, si le fardeau n'est calculé qu'en terme de questionnaires répondus alors Srinath

et Hidiroglou nous recommendent d'utiliser une méthode de panels, semblable à celle décrite en 5.2.

Pour en revenir à la diminution du fardeau entre les enquêtes, notons que l'utilisation de questionnaires unifiés offre l'opportunité d'aller directement à la source du problème pour le régler. Ces questionnaires minimiseraient ainsi le fardeau des répondants sélectionnés par plusieurs enquêtes. Ils nous permettraient alors de maximiser le chevauchement des échantillons pour ensuite maximiser le temps non-enquêté pour chacune des entreprises. De plus, cette évantualité est possible, peu importe la stratifi-cation des enquêtes, ce qui lui donne un atout certain sur l'échantillonnage synchronisé, même si elle apporte des problèmes quand les unités répondantes ne sont pas les mêmes pour toutes les enquêtes.

BIBLIOGRAPHIE

[1] Brewer,K.R.W., Early,L.J. et Hanif,M. (1984). "Poisson, Modified Poisson and Collocated Sampling." *Journal of Statist. Planning and Inf., 10, 15-30.*

[2] Brewer,K.R.W., Early,L.J. et Joyce,S.F. (1984). "Selecting Several Samples from a Single Population." *Austral. Journal of Statist., 14 (3), 231-239.*

[3] De Ree,S.J.M. (1983). "A System of Co-ordonated Sampling to Spread Response Burden of Entreprises." *Document de travail du 'Netherlands Central Bureau of Statistics', première ébauche, 18 mai 1983.*

[4] Division du travail (1982). "Survey of Employee Compensation." *Questionnaire pilote utilisé par la division du travail de Statistique Canada, en 1982.*

[5] Hidiroglou,M.A. et Srinath,K.P. (1987). "Sample Rotation." *Document de travail de la division des méthodes d'enquêtes-entreprises, Statistique Canada, avril 1987.*

[6] Hinde,R. et Young,D. (1983). "The Synchronised Sampling Selection Procedure." *Extrait d'un document d'Australie, 1er septembre 1983.*

[7] Leroux,D. et Lussier,R. (1987). "Etude empirique sur le fardeau du répondant." *Document de travail de la division des méthodes d'enquêtes-entreprises, Statistique Canada, juillet 1987.*

[8] Parlor,M. (1987). "A Model of Respondent Relations within the context of the Business Survey Redesign Project." *Document de travail de la division de la base centrale, Statistique Canada, 26 juin 1987.*

[9] Sunter,A.B. (1977). "An Integrated System for Surveys of Small Businesses."

[10] Sunter,A.B. (1977). "Response Burden, Sample Rotation, and Classification Renewal in Economic Surveys." *Internat. Statist. Review, 45, 209-222.*

[11] Tambay,J.-L. (1987). "General Sample Selection Procedures." *Document de travail de la division des méthodes d'enquêtes-entreprises, Statistique Canada, juillet 1987.*

[12] Système de développement des enquêtes généralisées (1987). "Survey of Manufactures." *Questionnaire pilote en cours de conception, décembre 1987.*

CA 005

STATISTICS CANADA LIBRARY
BIBLIOTHÈQUE STATISTIQUE CANADA
1010176842