CATALOGUE 71-526 OCCASIONAL c.2

CATALOGUE 71-526 HORS SÉRIE

Methodology of the Canadian labour force survey

1976

Méthodologie de l'enquête sur la population active du Canada

1976

Labour Canada Travail
RECEIVED – REÇU
MAR 22 1983
LIBRARY – BIBLIOTHÈQUE

LABOUR CANADA LIBRARY SERVICES
SERVICES DE LA BIBLIOTHÈQUE DE TRAVAIL CANADA
LENT TO / PRÊTÉ À Jane Ward
DATE BORROWED / DATE D'EMPRUNT FEB 10 87
DATE DUE / DATE DE RETOUR MAR 11 87
Lab/Trav 816 (2/86)

Statistics Canada Statistique Canada

STATISTICS CANADA — STATISTIQUE CANADA

Household Surveys Development Division — Division de l'élaboration d'enquêtes auprès des ménages

METHODOLOGY OF THE CANADIAN LABOUR FORCE SURVEY

MÉTHODOLOGIE DE L'ENQUÊTE SUR LA POPULATION ACTIVE DU CANADA

1976

Published by Authority of
The Minister of Industry, Trade and Commerce

Publication autorisée par
le ministre de l'Industrie et du Commerce

October - 1977 - Octobre
6-3000-501

Price—Prix: $2.10

Statistics Canada should be credited when republishing all or any part of this document
Reproduction autorisée sous réserve d'indication de la source: Statistique Canada

Ottawa

PREFACE

The preparation of this report was primarily undertaken by R. Platek, Director, Household Surveys Development Division and M.P. Singh, Assistant Director of the same Division.

The report presents a general but comprehensive description of the Canadian Labour Force Survey (LFS) Design as developed and implemented during 1972-1975. The LFS sample is redesigned after each decennial census and the last sample redesign took place during 1963-1966. While the basic framework of the new design, namely that of a stratified multi-stage sample design, has been kept the same as in the old design, several refinements have been introduced at various stages such as stratification, selection, allocation, estimation, etc. Extensive use of the 1971 Census data has been made and practically all the operational stages in sample selection and rotation have been computerized, a detailed description of which is available in the Sample Design System manual.

The present program of sample redesign was treated as an integral part of the overall Labour Force Survey revision program which embraced a thorough review of concepts, definitions, questionnaire design, data capture, processing sample expansion, etc. The entire Labour Force Survey revision program benefited from extensive consultation with a variety of users. Of necessity, one of the main features of the new design is the increased degree of flexibility for incorporating changes in the design parameters at any given level.

The determination of the broad technical objectives of redesign and general guidance in the development and implementation of the survey design was provided by I.P. Fellegi, Assistant Chief Statistician. Technical direction of the development and application of statistical methods was given by R. Platek, Director, Household Surveys Development Division.

The detailed technical development, planning and organization of the design work was the responsibility of M.P. Singh (Chairman, Sample Redesign Committee). Several officers from various divisions provided valuable assistance in technical and operational aspects of the program, among whom special mention should be made of; P.F. Timmons, G.B. Gray, B. Petrie, J. Coombs, I. Macredie, R. Baxter, R.C. Muirhead, D. Drew, F. Mayda, J. Derkach, S. Ross, H. O'Haver, E. McLeod, J. Tourigny, D. Beckstead, E. Maguire and the late Mrs. G.K. Giroux.

PETER G. KIRKHAM,

Chief Statistician of Canada.

PRÉFACE

Le directeur de la Division de l'élaboration d'enquêtes auprès des ménages, M.R. Platek, et le directeur adjoint de la même division, M. M.P. Singh, sont à l'origine de ce rapport.

Il donne une description globale, mais complète du plan de sondage de l'enquête sur la population active du Canada (E.P.A.), qui a été élaboré et appliqué entre 1972 et 1975. L'échantillon de l'E.P.A. est remanié après chaque recensement décennal et le dernier remaniement s'est fait de 1963-1966. La conception de l'enquête est demeurée fondamentalement la même, c'est-à-dire qu'elle se fonde sur un sondage stratifié à plusieurs degrés; cependant certaines étapes, dont la stratification, le sélection, la répartition et l'estimation, ont été améliorées. On a puisé abondamment dans les données du recensement de 1971 et pratiquement toutes les phases opérationnelles de la sélection et du renouvellement de l'échantillon ont été automatisées. Le manuel traitant du système du plan de sondage donne une description détaillée de cette automatisation.

Le programme actuel de remaniement de l'échantillon fait partie intégrante de l'ensemble du programme de l'enquête sur la population active qui a pour but d'étudier à fond les concepts, les définitions, la conception du questionnaire, la saisie et le traitement des données, l'élargissement de l'échantillon, etc. Ce programme de révision s'est inspiré de nombreuses consultations avec divers utilisateurs. Par conséquent, le nouveau plan de sondage se caractérise notamment par une flexibilité accrue permettant de procéder, à toute étape, à des changements dans les paramètres du plan de sondage.

Le statisticien en chef adjoint, M. I.P. Fellegi, a figé les grands objectifs techniques du remainiement et a donné l'orientation générale à l'élaboration et à l'application du plan de sondage. Le directeur de la Division de l'élaboration d'enquêtes auprès des ménages, M. R. Platek, a assuré la direction technique de l'élaboration et de l'application des méthodes statistiques.

M. M.P. Singh (président du comité chargé du remaniement de l'échantillon) s'est chargé des détails de la mise au point technique, de la planification et de l'organisation du remaniement. Plusieurs fonctionnaires de diverses divisions ont fourni une aide précieuse quant aux aspects techniques et opérationnels du programme; nous citerons, entre autres, Mad. S. Ross, MM. P.F. Timmons, G.B. Gray, B. Petrie, J. Coombs, I. Macredie, R. Baster, R.C. Muirhead, D. Drew, F. Mayda, J. Derkach, H. O'Haver, E. McLeod, J. Tourigny, D. Beckstead, E. Maguire et feu Mme G.K. Giroux.

Le statisticien en chef du Canada,

PETER G. KIRKHAM.

TABLE OF CONTENTS

Chapter	Page
1. Introduction	11
2. LFS Design up to 1975	15
2.1 General Description of the Design	15
2.2 Consideration for Redesign	16
2.2.1 Changes in Economic Region (ER) Boundaries	16
2.2.2 Annexation to SRU	16
2.2.3 Extent of Restratification	17
2.2.4 Sampling Procedure	17
3. Main Objectives of the LFS Revision Project	19
3.1 Objectives of LFS Revision	19
3.2 Main Phases of the Project	20
4. Stratification	23
4.1 Introduction	23
4.2 Classification of SR Areas	24
4.2.1 Lower Limit	24
4.2.2 Formation of Subunits	24
4.3 NSR Design Constraints on Stratification and PSU Delineation	25
4.3.1 Choice of Stratification Variables	25
4.3.2 Determination of Stratum and PSU Sizes	26
4.3.3 Contiguity Constraints	28
4.4 Methods of Stratification	28
4.5 Primary Sampling Units	30
4.5.1 Basic Features	30
4.5.2 Delineation of PSUs	30
5. Self-representing Units	33
5.1 General Description	33
5.2 Delineation of Clusters	33
5.3 Choice of the Selection Method	34
5.4 Selection of Clusters	36
5.5 Selection of Households	37
5.6 Sample Expansion	37
6. Non-self-representing Units	39
6.1 Introduction	39
6.2 Selection of Primary Sampling Units	39
6.2.1 Choice of Selection Procedure	39
6.2.2 Selection of PSUs	39

TABLE DES MATIÈRES

Chapitre	Page
1. Introduction	11
2. Conception de l'E.P.A. jusqu'en 1975	15
2.1 Description générale du plan de sondage de l'enquête	15
2.2 Raisons motivant le remaniement de l'enquête	16
2.2.1 Modifications des limites des régions économiques (R.É.)	16
2.2.2 Annexion à l'U.A.R.	16
2.2.3 Importance de la restratification	17
2.2.4 Méthode d'échantillonnage	17
3. Principaux objectifs du projet de révision de l'E.P.A.	19
3.1 Objectifs de la révision de l'E.P.A.	19
3.2 Principales étapes du projet	20
4. Stratification	23
4.1 Introduction	23
4.2 Classification des U.A.R.	24
4.2.1 Limite inférieure	24
4.2.2 Formation des sous-unités	24
4.3 Contraintes imposées par le plan de sondage des U.N.A.R. pour la stratification et la délimitation des U.P.É.	25
4.3.1 Choix des variables de stratification	25
4.3.2 Établissement de la taille de la strate et de l'U.P.É.	26
4.3.3 Contraintes imposées par la contiguité	28
4.4 Méthodes de stratification	28
4.5 Unités d'échantillonnage primaires	30
4.5.1 Caractéristiques fondamentales	30
4.5.2 Délimitation des U.P.É.	30
5. Unités autoreprésentatives	33
5.1 Description globale	33
5.2 Délimitation des grappes	33
5.3 Choix de la méthode de sélection	34
5.4 Sélection des grappes	36
5.5 Sélection des ménages	37
5.6 Élargissement de l'échantillon	37
6. Unités non autoreprésentatives	39
6.1 Introduction	39
6.2 Sélection des unités primaires d'échantillonnage	39
6.2.1 Choix de la méthode de sélection	39
6.2.2 Sélection des U.P.É.	39

TABLE OF CONTENTS – Continued

Chapter	Page
6.3 Sub-sampling in Urban Areas	40
6.3.1 Determination of Lower Limit for Urban Centers	40
6.3.2 Cluster Delineation	41
6.3.3 Selection of Clusters	41
6.4 Sub-sampling in Rural Areas	43
6.4.1 Selection of Rural Groups	43
6.4.2 Delineation and Selection of Clusters in Rural Groups	44
6.5 Selection of Households	44
6.6 Sample Expansion	44
7. Special Selections	47
7.1 Introduction	47
7.2 Apartment Frame	47
7.2.1 Definition and Coverage	47
7.2.2 Stratification and Selection Procedures	47
7.3 Special Dwelling Units	48
7.4 Special Area Frame	49
8. Sample Rotation	51
8.I Introduction	51
8.2 Method of Rotation	51
8.3 Operational Steps Involved in Sample Rotation	52
8.4 Assigning Rotation Numbers	54
9. Numerical Identification System	57
10. Data Collection, Conversion and Processing	61
10.1 Introduction	61
10.2 Data Collection	61
10.3 Data Entry Procedures in the Regional Office	62
10.4 Data Processing in the Head Office	63
11. Estimation	65
11.1 Introduction	65
11.2 Basic Weighting Factor	66
11.3 Rural-urban Factor	66
11.4 Treatment of Non-response	67
11.5 Cluster Sub-weight	68
11.6 Age-sex Factor	69
11.7 Final Weights	69
11.8 Algebraic Description of the Estimation	69
11.9 Example of Weighting of a Record	72

TABLE DES MATIÈRES – suite

Chapitre	Page
6.3 Sous-échantillonnage dans les régions urbaines	40
6.3.1 Détermination de la limite inférieure dans les centres urbains	40
6.3.2 Délimitation des grappes	41
6.3.3 Sélection des grappes	41
6.4 Sous-échantillonnage dans les régions rurales	43
6.4.1 Sélection des groupes ruraux	43
6.4.2 Délimitation et choix des grappes dans les groupes ruraux	44
6.5 Sélection des ménages	44
6.6 Élargissement de l'échantillon	44
7. Sélections spéciales	47
7.1 Introduction	47
7.2 Base d'immeubles d'appartements	47
7.2.1 Définition et champ d'observation	47
7.2.2 Méthodes de stratification et de sélection	47
7.3 Unités spéciales de logement	48
7.4 Base de secteurs spéciaux	49
8. Renouvellement de l'échantillon	51
8.1 Introduction	51
8.2 Méthode de renouvellement	51
8.3 Processus de renouvellement de l'échantillon	52
8.4 Attribution des numéros de renouvellement	54
9. Système de désignation numérique	57
10. Collecte, conversion et traitement des données	61
10.1 Introduction	61
10.2 Collecte des données	61
10.3 Entrée des données en ordinateur au bureau régional	62
10.4 Traitement des données au bureau central	63
11. Estimation	65
11.1 Introduction	65
11.2 Facteur de pondération de base	66
11.3 Facteur rural-urbain	66
11.4 Traitement de la non-réponse	67
11.5 Facteur de sous-pondération de grappe	68
11.6 Facteur âge-sexe	69
11.7 Poids final	69
11.8 Description algébrique de l'estimation	69
11.9 Exemple de pondération d'un dossier	72

TABLE OF CONTENTS – Continued

Chapter	Page
12. Sampling Error	75
12.1 Introduction	75
12.2 Variance Estimation	76
12.3 An Example of the Variance Calculations	77
12.4 Presentation of Errors	79
13. Assessment and Control of Non-sampling Errors	81
13.1 Introduction	81
13.2 Observation Program	83
13.3 Re-interview Program	83
13.4 Evaluation and Developmental Projects	84
14. Sample Allocation	87
14.1 Allocation of the Original Sample	87
14.2 Allocation of the Expanded Sample	88
Appendix	
Newfoundland	
0.1 Economic Region Map	93
0.2 Composition of Non-self-representing Units	94
0.3 1971 SRU List, by ER and Population	95
Prince Edward Island	
1.1 Economic Region Map	96
1.2 Composition of Non-self-representing Units	97
1.3 1971 SRU List, by ER and Population	97
Nova Scotia	
2.1 Economic Region Map	98
2.2 Composition of Non-self-representing Units	99
2.3 1971 SRU List, by ER and Population	100
New Brunswick	
3.1 Economic Region Map	101
3.2 Composition of Non-self-representing Units	102
3.3 1971 SRU List, by ER and Population	103

TABLE DES MATIÈRES – suite

Chapitre	Page
12. Erreur d'échantillonnage	75
12.1 Introduction	75
12.2 Estimation de la variance	76
12.3 Exemple du calcul de la variance	77
12.4 Présentation des erreurs	79
13. Évaluation et correction des erreurs non dues à l'échantillonnage	81
13.1 Introduction	81
13.2 Programme d'observation	83
13.3 Programme de réinterview	83
13.4 Projets d'évaluation et d'amélioration	84
14. Répartition de l'échantillon	87
14.1 Répartition de l'échantillon initial	87
14.2 Répartition de l'échantillon élargi	88
Annexe	
Terre-Neuve	
0.1 Carte des régions économiques	93
0.2 Composition des unités non autoreprésentatives	94
0.3 Liste des U.A.R. de 1971, par R.É. et selon la population	95
Île-du-Prince-Édouard	
1.1 Carte des régions économiques	96
1.2 Composition des unités non autoreprésentatives	97
1.3 Liste des U.A.R. de 1971, par R.É. et selon la population	97
Nouvelle-Écosse	
2.1 Carte des régions économiques	98
2.2 Composition des unités non autoreprésentatives	99
2.3 Liste des U.A.R. de 1971, par R.É. et selon la population	100
Nouveau-Brunswick	
3.1 Carte des régions économiques	101
3.2 Composition des unités non autoreprésentatives	102
3.3 Liste des U.A.R. de 1971, par R.É. et selon la population	103

TABLE OF CONTENTS – Continued

Appendix — Page

Québec

4.1 Economic Region Map 104
4.2 Composition of Non-self-representing Units 105
4.3 1971 SRU List, by ER and Population . . 106

Ontario

5.1 Economic Region Map 107
5.2 Composition of Non-self-representing Units 108
5.3 1971 SRU List, by ER and Population . . 109

Manitoba

6.1 Economic Region Map 110
6.2 Composition of Non-self-representing Units 111
6.3 1971 SRU List, by ER and Population . . 112

Saskatchewan

7.1 Economic Region Map 113
7.2 Composition of Non-self-representing Units 114
7.3 1971 SRU List, by ER and Population . . 115

Alberta

8.1 Economic Region Map 116
8.2 Composition of Non-self-representing Units 117
8.3 1971 SRU List, by ER and Population . . 118

British Columbia

9.1 Economic Region Map 119
9.2 Composition of Non-self-representing Units 120
9.3 1971 SRU List, by ER and Population . . 121

TABLE DES MATIÈRES – suite

Annexe — Page

Québec

4.1 Carte des régions économiques 104
4.2 Composition des unités non autoreprésentatives 105
4.3 Liste des U.A.R. de 1971, par R.É. et selon la population 106

Ontario

5.1 Carte des régions économiques 107
5.2 Composition des unités non autoreprésentatives 108
5.3 Liste des U.A.R. de 1971, par R.É. et selon la population 109

Manitoba

6.1 Carte des régions économiques 110
6.2 Composition des unités non autoreprésentatives 111
6.3 Liste des U.A.R. de 1971, par R.É. et selon la population 112

Saskatchewan

7.1 Carte des régions économiques 113
7.2 Composition des unités non autoreprésentatives 114
7.3 Liste des U.A.R. de 1971, par R.É. et selon la population 115

Alberta

8.1 Carte des régions économiques 116
8.2 Composition des unités non autoreprésentatives 117
8.3 Liste des U.A.R. de 1971, par R.É. et selon la population 118

Colombie-Britannique

9.1 Carte des régions économiques 119
9.2 Composition des unités non autoreprésentatives 120
9.3 Liste des U.A.R. de 1971, par R.É. et selon la population 121

TABLE OF CONTENTS – Concluded

Appendix		Page
10.1	Map for Regina SRU	122
10.2	NSR – PSU Map	123
10.3	NSR – Group Map	124
10.4	SRU – Cluster Map	125
10.5	NSRU – (rural) Cluster Map	126
11.1	Frequency Distribution of Clusters by their Size in SRUs (Built-up)	127
11.2	Frequency Distribution of Clusters by their Size in SRUs (Fringe)	128
12.1	Population Summary by Province	129
13.1	Household Record Docket	130
13.2	Labour Force Survey Questionnaire	132
13.3	Definitions and Explanations	134
14.1	Diagramatic Representation of LFS Frame	137
References		138

TABLE DES MATIÈRES – fin

Annexe		Page
10.1	Carte de l'U.A.R. de Régina	122
10.2	Carte U.N.A.R. – U.P.E.	123
10.3	Carte U.N.A.R. – Groupe	124
10.4	Carte U.A.R. – Grappe	125
10.5	Carte U.N.A.R. – Grappe (Régina rurale)	126
11.1	Distribution de la fréquence des grappes selon leur taille dans les U.A.R. (zones bâties)	127
11.2	Distribution de la fréquence des grappes selon leur taille dans les U.A.R. (banlieue)	128
12.1	Aperçu de la population par province	129
13.1	Dossier du ménage	130
13.2	Questionnaire de l'enquête sur la population active	132
13.3	Définitions et explications	134
14.1	Diagramme du cadre de l'E.P.A.	137
Bibliographie		138

CHAPTER – 1 – CHAPITRE

INTRODUCTION

The Canadian Labour Force Survey is the largest continuing household survey conducted by Statistics Canada. The Survey, established in 1945 was originally designed to provide quarterly estimates, and since 1952 has been conducted monthly. Prior to the present revision, a major redesign of the sample took place in 1963 and a brief description of the design, used from 1964 to 1975, is given in Chapter 2.

The survey covers the civilian, non-institutional population of 14 years of age and over from the ten provinces in Canada. It excludes the Yukon and the Northwest Territories and population living on Indian Reserves and Crown lands, inmates of institutions and members of armed forces. The population in excluded areas accounts for about 2% of the total population of Canada.

The information provided by the Labour Force Survey enables Statistics Canada to publish a wide range of estimates on the size, composition and characteristics of the Labour Force. These monthly Labour Force statistics serve as important indicators of economic conditions, and are used extensively by federal and provincial governments, business, labour, and research institutions for understanding and interpretation of trends in the Canadian economy.

The responsibility for the operation of the LFS is divided amongst the Labour Force Survey Division, Field Division and Household Surveys Development Division, with substantial participation by the Systems and Data Processing Branch. The Labour Force Survey Division has the overall responsibility for the concepts, definitions, analysis, publication and user liaison, as well as for editing, processing and timeliness of different stages of operations.

Field Division is responsible for all aspects of the field work inclusive of sample maintenance, and the recruiting and training of interviewers required for the survey. It has eight Regional Offices (ROs) and about 700 interviewers (for the current sample) who are engaged in collecting information from approximately 35,000 households (about 80,000 respondents) during a specified week (the survey week) of each month for information pertaining to the previous week (the reference week). By early 1977, the sample size is expected to be approximately 55,000 households, in-

L'enquête sur la population active du Canada est la plus vaste enquête permanente sur les ménages menée par Statistique Canada. À sa création en 1945, elle avait pour but de fournir des estimations trimestrielles, mais depuis 1952, elle est menée tous les mois. Un important remaniement de l'échantillon a eu lieu en 1963, précédant la réforme actuelle. Le chapitre 2 contient une brève description du plan de sondage utilisé entre 1964 et 1975.

L'enquête couvre la population civile hors institutions âgée de 14 ans et plus, des dix provinces du Canada. En sont exclus les habitants du Yukon et des Territoires du Nord-Ouest, la population vivant dans les réserves indiennes et sur les terres de la Couronne, les détenus des institutions et les membres des forces armées. La population des régions exclues représente environ 2 % de la population totale du Canada.

Les renseignements obtenus au cours de l'enquête sur la population active permettent à Statistique Canada de publier une grande variété d'estimations sur la taille, la composition et les caractéristiques de la population active. Ces statistiques mensuelles constituent d'importants indicateurs des conditions économiques et sont largement utilisées par le gouvernement fédéral et les gouvernements provinciaux, par les entreprises commerciales, par les organismes spécialisés en main-d'oeuvre ou en recherches, pour mieux saisir et interpréter les tendances de l'économie canadienne.

Les Divisions de l'enquête sur la population active, des opérations régionales et de l'élaboration d'enquêtes auprès des ménages se partagent la responsabilité de l'enquête sur la population active et bénéficient de la collaboration soutenue de la Direction des systèmes informatiques et du traitement des données. La Division de l'enquête sur la population active est chargée de l'élaboration des concepts, des définitions, de l'analyse, de la publication, de la liaison avec les utilisateurs, de la vérification et du traitement informatique des données ainsi que de voir à ce que les délais prescrits pour les diverses étapes des travaux soient respectés.

La Division des opérations régionales s'occupe de tous les aspects du travail sur le terrain, y compris la mise à jour de l'échantillon, le recrutement et la formation des interviewers requis pour l'enquête. Elle a huit bureau régionaux (B.R.) et près de 700 interviewers (pour l'échantillon actuel) qui interrogent chaque mois 35,000 ménages environ, soit approximativement 80,000 répondants, au cours d'une semaine donnée (la semaine d'enquête) pour obtenir des renseignements sur la semaine précédente (la semaine de référence). On prévoit qu'au début de 1977, la taille de l'échantillon sera de 55,000 ménages et que l'on

volving about 1,200 interviewers. The interviews are carried out primarily by telephone in the larger cities and by personal interviews in other areas. Intensive programs of observation and re-interviewing are employed as regular features of the survey. The questionnaires filled in by the interviewers are sent to the ROs where data are captured and transmitted to the Head Office for detailed editing and processing. The entire operation of data conversion and transmission has been automated (through the use of mini-computers in each of the ROs) as part of the general revision program designed to minimize the time-lag and increase the efficiency of the data handling system.

The Household Surveys Development Division (HSDD) has overall responsibility for the methodology of all household surveys. Particularly as this relates to the LFS, HSDD is responsible for directing and carrying out research and development of the LFS methodology. This involves carrying out all design related work; that is, activities such as development and evaluation of methodology, development of evaluation programs and controls at different stages of the operation of the survey, maintenance of the quality of the design over time, and preparation of sampling material for field operations.

In this report it is intended to provide a comprehensive description of the sampling methodology developed and applied in revising the LFS during 1972-1975. Historically, a redesign of the LFS sample has usually been undertaken after each decennial census. In the meantime, the survey relies on regular updating programs to take account of population growth. Each Census presents a unique opportunity to examine all phases of the existing design in light of the changes in the structure of the labour force as determined from the latest Census, the more recent developments reported in survey sampling and the results of the evaluation studies conducted for the LFS itself. It also provides the opportunity to introduce other changes which on their own would not be feasible.

A brief description of the old design and considerations leading to a complete redesign are given in Chapter 2. The objectives of the overall revision program, and the approach taken to achieve the objective, are summarized in Chapter 3. Chapters 4-9 are devoted to a detailed description of the redesign methodology. This includes stratification in Chapter 4, allocation and selection procedure for self-representing units and non-self-representing units in Chapters 5 and 6, apartment sample and special area sample in Chapter 7, the rotation scheme in Chapter 8 and the numerical identification system in Chapter 9. A brief description of data collection and processing is given in Chapter 10. Chapters 11 and 12 respectively provide description

comptera 1,200 interviewers. Dans les grandes villes, les interviews se font surtout par téléphone, alors qu'ailleurs, on procède par interviews sur place. L'enquête est caractérisée par des programmes intensifs d'observation et de réinterview. Les interviewers envoient les questionnaires qu'ils ont remplis aux B.R. où les données sont saisies et transmises au bureau principal; elles y subissent une vérification approfondie avant leur traitement informatique. Le processus de conversion et de transmission des données a été entièrement automatisé (grâce à l'utilisation de mini-ordinateurs dans tous les B.R.) dans le cadre du programme général de révision de l'enquête visant à réduire les délais et à améliorer l'efficacité du système de traitement des données.

La Division de l'élaboration d'enquêtes auprès des ménages est chargée d'établir la méthodologie pour toutes les enquêtes auprès des ménages, ainsi que de diriger et d'effectuer la recherche qui s'y rattache. Cette responsabilité s'applique en particulier à l'E.P.A. A cette fin, elle doit s'acquitter de toutes les tâches liées à la conception de l'enquête, telles qu'élaborer et évaluer la méthodologie mettre au point des programmes d'évaluation et des contrôles pour les diverses étapes de l'enquête assurer, en tout temps, le maintien de la qualité du plan de sondage et préparer le matériel nécessaire à l'échantillonnage sur le terrain.

Le présent rapport a pour but de présenter une description complète des méthodes d'échantillonnage qui ont été élaborées et appliquées au cours de la refonte de l'E.P.A. qui a eu lieu de 1972-1975. Ordinairement, après chaque recensement décennal, l'échantillon de l'enquête sur la population active est remanié. Entre temps, des programmes de mise à jour permettent de tenir compte de l'accroissement de la population. Chaque recensement offre l'occasion unique d'étudier toutes les facettes du plan de sondage utilisé, à la lumière des changements intervenus dans la structure de la population active et mis en évidence par le dernier recensement, des tendances nouvelles dans les méthodes d'échantillonnage et des résultats des études d'évaluation entreprises pour l'E.P.A. Il permet également l'introduction d'autres changements qui ne seraient pas réalisables autrement.

Le chapitre 2 contient une brève description de l'ancien plan de sondage et donne les raisons qui en ont motivé le remaniement complet. Le chapitre 3 résume les objectifs de l'ensemble du programme de refonte de l'enquête et l'approche adoptée pour réaliser ces objectifs. Les chapitres 4-9 sont consacrés à la description détaillée des méthodes adoptées lors de la refonte de l'enquête. Le chapitre 4 porte sur la stratification, les chapitres 5 et 6 décrivent le processus de répartition et de sélection dans le cas des unités autoreprésentatives et celui des unités non autoreprésentatives, le chapitre 7 traite des échantillons d'immeubles d'appartements et de secteurs spéciaux, le chapitre 8 étudie le mode de renouvellement de l'échantillon, et le chapitre 9, le système de désignation numérique.

of the estimation and variance estimation procedures as used in the survey. Procedures employed for assessment and control of non-sampling errors are discussed in Chapter 13. A brief discussion on the allocation of the sample to various provinces is presented in the last chapter. Tables providing general information on the composition of self-representing units, non-self-representing units, economic regions, etc. for each province are included in the appendices.

Le chapitre 10 contient une brève description de la méthode de collecte et de traitement des données. Les chapitres 11 et 12 décrivent respectivement les méthodes d'estimation et d'estimation de la variance utilisées dans l'enquête. Les méthodes employées pour l'évaluation et le contrôle des erreurs non dues à l'échantillonnage sont exposées dans le chapitre 13. Le dernier chapitre renferme une brève analyse de la répartition de l'échantillon dans les diverses provinces. L'appendice contient des tableaux qui donnent des renseignements généraux sur la composition des unités autoreprésentatives, des unités non autoreprésentatives, des régions économiques, etc., pour chaque province.

CHAPTER – 2 – CHAPITRE

LFS DESIGN UP TO 1975

2.1 General Description of the Design

The LFS sample which was designed during 1963 - 1966 continued to be the source for obtaining estimates for labour force characteristics up to the end of 1975. The sample design was based on information from the Census of 1961. Basically the LFS consisted of two main parts: Self-representing Units (SRUs) and Non-self-representing Units (NSRUs). The SRU's were cities whose population exceeded 15,000 persons or whose unique characteristics demanded their establishment as SRUs. Within all SRUs the sample was selected independently so that each of them is represented in the survey by a sample of its own population and hence the name "self-representing". In addition, the larger SRUs were divided into subunits (strata) of sizes ranging from 15,000 - 30,000 persons. Each subunit was subdivided into segments (mostly city blocks with well defined boundaries) which formed the first stage units for selection. Six (or multiples of six) segments were selected systematically with probability proportional to size (PPS), the size being the number of households within a segment. A sample of households was then selected systematically by applying the segment sub-sampling ratio within the selected segment.

The NSRUs were the areas outside the SRUs, containing rural areas and small urban centers. Each province in Canada is divided into Economic Regions (ER) whose boundaries remain, more or less, stable over time. Within each ER the NSRU area was stratified according to some relevant characteristics. The sample of dwellings was selected within each stratum in four stages. From each stratum two Primary Sampling Units (PSUs) were selected without replacement with probabilities proportional to their size (population) following Fellegi's [2] procedure. Each selected PSU was divided into rural segments and urban segments (usually census enumeration areas (EAs), and the required number of segments were selected systematically with PPS, independently from rural and urban parts of the selected PSU. The third stage of sampling was clusters which were selected with PPS. Finally, dwellings were chosen systematically. The design was made self weighting for both SRUs and NSRUs at the province level.

In addition to the above area frames, the design had recognized and incorporated separate frames for

CONCEPTION DE L'E.P.A. JUSQU'EN 1975

2.1 Description générale du plan de sondage de l'enquête

Le plan de sondage de l'E.P.A., conçu au cours des années 1963 - 1966, a servi à obtenir des estimations des caractéristiques de la population active jusqu'en 1975. Il a été établi à partir des renseignements recueillis au cours du recensement de 1961. L'E.P.A. se composait essentiellement de deux grandes parties: les unités autoreprésentatives (U.A.R.) et les unités non autoreprésentatives (U.N.A.R.). Les U.A.R. représentaient les cités qui avaient une population de plus de 15,000 habitants ou qui, en raison de leurs caractéristiques particulières, devaient être considérées comme des U.A.R. Dans toutes les U.A.R., l'échantillon était choisi de façon indépendante afin que chaque unité puisse être représentée par un échantillon de sa propre population. C'est la raison pour laquelle on les a appelé unités "autoreprésentatives". En outre, les grandes U.A.R. étaient divisées en sous-unités (les strates) dont la taille variait de 15,000 - 30,000 habitants. Chaque sous-unité était subdivisée en segments (pour la plupart des îlots urbains dont les limites étaient bien déterminées) qui constituaient les premières unités de sélection. Six segments (ou des multiples de six) étaient choisis systématiquement avec probabilité proportionnelle à la taille (P.P.T.), la taille étant le nombre de ménages composant le segment. Ensuite, on sélectionnait systématiquement un échantillon de ménages en utilisant la fraction de sous-sondage du segment dans le segment choisi.

Les U.N.A.R. étaient les régions à l'extérieur des U.A.R., qui se composaient de régions rurales et de petits centres urbains. Toutes les provinces canadiennes sont divisées en régions économiques (R.É.) dont les limites sont plus ou moins permanentes. Dans chaque R.É., les U.N.A.R. étaient stratifiées en fonction de certaines caractéristiques. Le choix de l'échantillon de logements dans chaque strate se faisait en quatre étapes. D'abord, on tirait dans chaque strate deux unités primaires d'échantillonnage (U.P.É.) sans remise, avec probabilité proportionnelle à la taille (en l'occurrence, la population), conformément à la méthode de Fellegi [2]. On divisait ensuite chaque U.P.É. choisie en segments ruraux et en segments urbains – en général, des secteurs de dénombrement du recensement (S.D.), puis on tirait systématiquement le nombre de segment requis, avec probabilité proportionnelle à la taille (P.P.T.), séparément pour les parties rurales et urbaine de l'U.P.É. choisie. Le troisième degré d'échantillonnage consistait à sélectionner des grappes avec probabilité proportionnelle à la taille (P.P.T.). Enfin, on choisissait les logements par tirage systématique. Au niveau des provinces, le plan de sondage était à auto-pondération aussi bien pour les U.A.R. que pour les U.N.A.R.

Outre les bases des régions mentionnées précédemment, le plan de sondage comprenait des bases distinctes

apartment buildings and special areas. In 12 larger SRUs a separate frame for apartment buildings was set up and a PPS systematic sample of apartments was selected from each of 12 cities. With regards to regular updating of the sample, the apartment frame was kept open ended in order to allow new apartments to be included in the frame. In cases of excessive growth within subunits, special selection procedures, referred to as up-dating procedures, were applied.

Special areas were first divided into four strata and an independent sample was then selected from each stratum.

A detailed description of the methodology is given in [16].

2.2 Consideration for Redesign

Due to differences in the design as well as the extent of updating used in different parts (viz. SRU, NSRU, apartment and special area) of the overall LFS frame, it was decided at an early stage of the development of the redesign program that each part would be reviewed individually and the decision on the extent of redesign in one part would not directly influence the decision in another part. In this context the following factors were examined and as a result complete redesign was undertaken.

2.2.1 Changes in Economic Region (ER) Boundaries

In the old design, the Department of Defence Production's (DDP) ERs of 1961 were used as primary strata. A review of DDP, ERs was carried out with the result that in November 1971, the Regional Statistical Policy and Co-ordinating Committee in consultation with the provinces recommended a new set of regions. The principle recommendation of the committee was that these 1971 ERs be "utilized uniformly by the bureau divisions as an additional level of spatial reference for the purpose of data tabulation". Following further discussions between Statistics Canada and data users, particularly those in provincial government departments, the set of ER boundaries shown in the appendices was adopted for the design of the revised LFS sample frame.

2.2.2 Annexation to SRU

The boundaries of many SRUs (i.e., as defined in terms of 1971 versus 1961 EA descriptions of cities with qualifying populations) had changed due to growth or due to some other administrative considerations. As a result of the changes, parts of NSR strata, or in some cases entire strata, were included in a neighbouring SRU. Consequently, the existing NSR stratification in such ERs would have been adversely affected, which

pour les immeubles d'appartements et les secteurs spéciaux. Une base distincte a été établie pour les immeubles d'appartements dans 12 grandes U.A.R. et un échantillon systématique d'appartements a été choisi avec P.P.T. dans les 12 cités. En vue de la mise à jour régulière de l'échantillon, on avait gardé la base d'immeubles d'appartements ouverte pour pouvoir y introduire de nouveaux immeubles d'appartements. Lorsque les sous-unités connaissaient une trop forte croissance, on utilisait des méthodes de sélection spéciales qu'on appelle des méthodes de mise à jour.

Les secteurs spéciaux avaient d'abord été divisés en quatre strates, puis on tirait un échantillon de façon indépendante dans chaque strate.

Vous trouverez une description détaillée de la méthode utilisée dans [16].

2.2 Raisons motivant le remaniement de l'enquête

À cause des divergences dans le plan de l'enquête et de l'ampleur de la mise à jour effectuée dans les diverses parties de la base de l'E.P.A. (c'est-à-dire les U.A.R., les U.N.A.R., les immeubles d'appartements et les secteurs spéciaux), il a été décidé, au début du programme de refonte de l'enquête, de réviser séparément chaque partie et de ne pas assujettir l'importance du remaniement dans une partie à la décision prise pour une autre partie. Compte tenu de ces faits, on a étudié les facteurs ci-après et entrepris une refonte intégrale de l'enquête.

2.2.1 Modification des limites des régions économiques (R.É.)

L'ancien plan de sondage de l'enquête utilisait les R.É. de 1961, établies par le ministère de la Production de défense, comme strates primaires. On a entrepris la révision de ces R.É. et, en novembre 1971, le Comité d'orientation et de coordination des statistiques régionales a recommandé, de concert avec les provinces, la formation de nouvelles régions. Il a recommandé principalement que les diverses divisions du bureau utilisent uniformément les R.É. créées en 1971 comme référence spatiale additionnelle pour classer les données sous forme de tableaux. À la suite de discussions ultérieures entre Statistique Canada et les utilisateurs de données, en particulier ceux des ministères provinciaux, la délimitation des R.É. indiquée dans les annexes a été adoptée en vue de la conception de la nouvelle base de sondage de l'E.P.A.

2.2.2 Annexion à l'U.A.R.

Les limites de plusieurs U.A.R. (créées selon la description des S.D. de 1971, comparativement à celle de 1961, pour les cités dont la population justifie cette dénomination) ont été modifiées en raison de l'accroissement de la population ou de certaines autres considérations d'ordre administratif. À la suite de ces modifications, certaines parties des strates des U.N.A.R., ou dans certains cas des strates entières, étaient incluses dans l'U.A.R.

could have resulted in poor efficiency of the LFS estimates. The number of SRUs affected due to annexation and the boundary changes in ERs was one factor which determined the extent of restratification necessary.

2.2.3 Extent of Restratification

It was decided that all the ERs which were affected by either boundary changes or annexation required restratification. Examination showed 41 out of 67 ERs fell into this category. Of the remaining 26 ERs, six ERs consisted of a single stratum, and six ERs (distributed in different provinces) did not have any NSR area. Thus, in 1971 there were only 14 out of a total of 67 ERs which would not require restratification on account of boundary changes or annexation.

In addition, other operational and theoretical considerations had to be taken into account when deciding upon the need for restratification of an ER. Some of these considerations were boundary changes in EAs (used in formation of primary and secondary sampling units in NSR areas), others were deterioration in the Index of Stratification, and formation of new SRUs.

Joint consideration of the above factors led to the complete redesign of the NSR areas. The main reason for redesigning SR areas was to update the size measures and incorporate a selection procedure that would simplify up-dating of the sample.

2.2.4 Sampling Procedure

The selection methods employed in the old design were examined according to the following criteria: (a) simplicity of updating procedure, (b) flexibility in the procedure to increase the sample size at any stage of sampling, (c) theoretical efficiency of the estimates and variance estimates. Criterion (a) is quite important for any continuous survey where uneven population (size) growth occurs frequently, and this was a major consideration for the choice of the sampling procedure in the SRUs (Platek and Singh [11]). Criterion (b) was necessitated by the possibility of a substantial increase in the size of the LFS, and thus had a bearing on the selection procedure employed in the NSR areas in the old design. Detailed discussion of these is given in the subsequent chapters.

voisine, ce qui aurait nui à la stratification d'alors des U.N.A.R. dans ces R.É. et réduit sensiblement l'efficacité des estimations de l'E.P.A. Le nombre d'U.A.R. touchées par l'annexion et par les modifications apportées à leur frontières est l'un des facteurs qui ont déterminé l'importance de la restratification nécessaire.

2.2.3 Importance de la restratification

On a jugé nécessaire de restratifier toutes les R.É. touchées par des changements de frontières ou par l'annexion. Selon une étude, 41 des 67 R.É. entraient dans cette catégorie. Sur les 26 R.É. qui restaient, six étaient composées d'une seule strate et six autres (réparties dans différentes provinces) ne comptaient pas d'U.N.A.R. C'est la raison pour laquelle en 1971 seules 14 R.É. sur un total de 67 n'ont pas été touchées par la restratification à cause de changements de frontières ou d'annexion.

En outre, il a fallu tenir compte d'autres aspects d'ordre pratique ou théorique, au moment de décider de la nécessité de la restratification d'une R.É. Il s'agissait, dans certains cas, de modifications apportées aux limites des S.D. (servant à former les unités primaires et secondaires d'échantillonnage dans les U.N.A.R.), et, dans d'autres, de la détérioration de l'indice de stratification et de la formation de nouvelles U.A.R.

L'étude de tous ces facteurs a entraîné le remaniement complet des U.N.A.R. Le but principal du remaniement des U.A.R. était de mettre à jour les mesures de la taille et d'introduire une méthode de sélection permettant de simplifier la mise à jour de l'échantillon.

2.2.4 Méthode d'échantillonnage

Les méthodes de sélection utilisées dans l'ancien plan de sondage ont été étudiées en fonction des critères suivants: a) simplicité de la mise à jour, b) flexibilité de la méthode utilisée permettant d'élargir l'échantillon à n'importe quel degré de l'échantillonnage, c) efficacité théorique des estimations et des estimations de la variance. Le critère a) est très important pour toute enquête permanente dans laquelle il faut tenir compte de fréquents accroissements irréguliers de la population et constitue un facteur décisif du choix de la méthode d'échantillonnage dans les U.A.R. (Platek et Singh [11]). Le critère b) est devenu nécessaire pour faire face à l'éventualité d'une forte augmentation de l'effectif de l'E.P.A. et il a influencé le choix de la méthode de sélection utilisée pour les U.N.A.R. dans l'ancien plan de sondage. Les chapitres suivants donnent une analyse détaillée de ces critères.

CHAPTER – 3 – CHAPITRE

MAIN OBJECTIVES OF THE LFS REVISION PROJECT

3.1 Objectives of LFS Revision

In a continuing large scale, multi-purpose survey such as the Canadian LFS, the major objectives and priorities are reviewed from time to time and the survey is revised accordingly. Some aspects of the survey may be revised due to changes in data requirements, others are due to changes in methodology and procedure.

Since the middle of 1972, Statistics Canada has been engaged in the development of a major revision and extension of the LFS in order to meet the new data requirements which in recent years have emerged and continue to emerge at a rapide rate.

The requirements for new information can be summarized in the following general categories:

(i) Demand for new types of information to provide a more comprehensive picture of the current labour market situation than is revealed by the simple distinction between employment, unemployment, and non-labour activities.

(ii) Demand for new information on the dynamic aspects of the labour force to supplement the monthly "snapshots" the new survey provides, much more information is required on the gross movements between various labour force categories each month, along with somewhat longer-term longitudinal data covering the six month period over which respondents are surveyed.

(iii) Demand for a larger sample to produce more reliable estimates for provincial, sub-provincial and regional levels, and provide finer levels of cross-classification (e.g., unemployed classified according to more detailed demographic breakdowns) than are now available.

Meeting the new data requirements outlined above was the major, but by no means the only objective of the overall revision of the LFS. Included among the other targets were the following:

(i) Up-dating the sample by partial or complete redesign.

(ii) Introduction of new methodology and procedures at the level of stratification, sample allocation and formation of sampling units. Examination of new sampling methods with respect to their efficiency, cost, variance estimation and operational convenience.

(iii) Improving data collection techniques, quality control and evaluation procedures, etc., in order to provide more and better quality data without sacrificing timeliness.

PRINCIPAUX OBJECTIFS DU PROJET DE RÉVISION DE L'E.P.A.

3.1 Objectifs de la révision de l'E.P.A.

De temps à autre, il faut procéder à la révision des principaux objectifs et des priorités d'une enquête permanente à grande échelle et à buts multiples comme l'E.P.A. du Canada et la réviser en conséquence. Certains aspects de l'enquête peuvent être réexaminés en raison de changements intervenus dans les besoins en données, d'autres, en fonction des nouvelles méthodes et pratiques adoptées.

Depuis le milieu de 1972, Statistique Canada a entrepris une vaste révision et un élargissement de l'E.P.A. afin de satisfaire les nouveaux besoins en données qui se sont manifestés au cours des dernières années et qui augmentent rapidement.

On peut grouper les demandes en nouveaux renseignements selon les catégories suivantes:

(i) Nouveaux types de renseignements donnant une description plus précise de la situation actuelle du marché du travail que ne le fait la simple distinction entre l'emploi, le chômage et l'inactivité.

(ii) Nouveaux renseignements sur les aspects dynamiques de la population active, qui compléteraient les communiqués mensuels issus de la nouvelle enquête; on demande davantage de renseignements sur les mouvements bruts mensuels entre les diverses catégories de la population active, et des données longitudinales qui couvriraient la période de six mois pendant laquelle les répondants sont soumis à l'enquête.

(iii) Plus grand échantillon pour produire des estimations provinciales, infraprovinciales et régionales plus fiables, et de meilleures classifications recoupées que celles dont on dispose actuellement (par exemple, il faudrait que les chômeurs soient classés selon des répartitions démographiques plus détaillées).

Bien que la révision générale de l'E.P.A. ait eu pour objectif principal la satisfaction de nouveaux besoins en données, d'autres raisons l'ont également motivée, parmi lesquelles on peut citer:

(i) La mise à jour de l'échantillon à l'aide d'une refonte partielle ou totale.

(ii) L'introduction de nouvelles méthodes et procédés de stratification, de répartition de l'échantillon et de formation d'unités d'échantillonnage; l'étude des nouvelles méthodes d'échantillonnage au point de vue efficacité, coût, estimation de la variance et praticabilité.

(iii) L'amélioration des techniques relatives à la collecte des données, au contrôle de la qualité et aux méthodes d'évaluation, etc., afin de produire davantage de données de meilleure qualité sans sacrifier leur actualité.

More elaborate discussion of these objectives and different phases of the revision program are given in a paper entitled "Project Review" [9]. Only the sample redesign is discussed in detail in this report, although in the remainder of this chapter various aspects of the revision are briefly discussed in the context of how they affect the redesign methodology.

In order to produce new types of data, to portray the sensitive picture and the complexities of the labour market situation, both conceptual changes and changes in the survey technique were considered. After considerable consultation and discussion with users, it was decided not to make any major changes in fundamental concepts, since none of the possible changes would meet the desired objectives of all users, and since the present concepts were in close agreement with International Labour Organization standards. Rather, it was decided to incorporate basic changes in the survey technique to achieve the objectives. These changes required a shift from a general to a highly structured approach in the questionnaire design, from an indirect style of questioning to a direct style of asking a series of more specific and factual questions. The revised questionnaire hence is similar in nature to that used in the U.S. *Current Population Survey* and in a number of comparable surveys in other industrialized countries. The many reasons leading to various decisions on questionnaire content and design have been described in detail in papers [8], [9], [10] and [14].

Le document intitulé "Étude du projet" [9] renferme une analyse plus détaillée de ces objectifs et des diverses phases du programme de révision. Seul le remaniement de l'échantillon est traité en détail dans le présent rapport, même si la fin de ce chapitre analyse brièvement les divers aspects de la révision de l'enquête et son influence sur les méthodes du nouveau plan.

On a étudié la possibilité d'apporter des changements tant aux concepts qu'aux techniques d'enquête pour produire de nouveaux types de données dans le but de donner une description fidèle de la situation et des aspects complexes du marché du travail. Après de nombreuses consultations et discussions avec les utilisateurs, nous avons décidé de ne procéder à aucun changement important dans les concepts fondamentaux, car aucun des changements possibles ne satisferait les objectifs de tous les utilisateurs, et que les concepts actuels respectent parfaitement les normes de l'Organisation internationale du travail. Nous avons plutôt décidé d'introduire des changements fondamentaux dans la technique de l'enquête afin de réaliser ces objectifs. Ces changements exigent, pour la conception du questionnaire, l'abandon d'une approche globale pour une méthodologie très structurée et le passage d'un style d'interrogation indirecte à l'interrogation directe au moyen d'un ensemble de questions plus précises et plus concrètes. Le questionnaire révisé ressemble donc à celui utilisé aux É.-U. pour la *Current Population Survey* et à celui de nombreuses enquêtes semblables menées dans d'autres pays industrialisés. Les nombreuses raisons motivant les diverses décisions prises sur la conception et sur le contenu du questionnaire sont énoncées en détail dans les documents [8], [9], [10] et [14].

3.2 Main Phases of the Project

The overall approach taken in the LFS revision projects was to develop a new survey, incorporating the desired changes, to replace the former survey in January of 1976. Main phases of the program [9] were: (i) development of the revised questionnaire and survey processing system, (ii) sample redesign, (iii) phase-in of the new sample, (iv) parallel run of the new and old surveys, (v) sample increase. The developments in (i) and (ii) took place concurrently and the phase-in was spread over a period of eight months, from May to December 1974. The sample listing system and data capture and transmission system were completely tested during this phase-in period. Since the survey technique in general for the new sample was different from the ongoing LFS, an entirely separate set of interviewers, (about 700 interviewers throughout the eight ROs) were hired, trained and inducted during this period.

3.2 Principales étapes du projet

Les projets de révision de l'E.P.A. visaient à élaborer une nouvelle enquête tenant compte des changements désirés, qui remplacerait l'enquête précédente à compter de janvier 1976. Les principales étapes de ce programme [9] sont les suivantes: (i) élaboration du nouveau questionnaire et du système de traitement de données de l'enquête, (ii) remaniement de l'échantillon, (iii) introduction du nouvel échantillon, (iv) déroulement simultané de la nouvelle enquête et de l'ancienne, (v) élargissement de l'échantillon. Les étapes (i) et (ii) se sont déroulées en même temps et l'introduction du nouvel échantillon s'est étalée sur huit mois soit de mai à décembre 1974. Le système de listage de l'échantillon et celui de saisie et de transmission des données ont été entièrement mis à l'épreuve au cours de cette période d'introduction de l'échantillon. Comme dans l'ensemble, la technique d'enquête pour le nouvel échantillon est différente de celle de l'E.P.A. en cours, on a engagé, formé et nommé de nouveaux interviewers au cours de cette période (environ 700 interviewers pour les huit B.R.).

These first three phases gave rise to the Revised LFS and prepared for the PARALLEL RUN, to be followed by the last phase of sample expansion. The span of the parallel run was the full calendar year of 1975. During this period both the old LFS and the Revised LFS were conducted as two independent survey operations to obtain data enabling direct comparison of survey results for purposes of historical revision and seasonal adjustment. The sample expansion was planned for 1976 - 1977 for the purpose of providing more detailed data for smaller regions, and called for an increase in sample size from approximately 35,000 - 55,000 households per month. Allocation of this increase to various provinces was directed mainly to the objective of reducing the differences in the level of reliability of unemployment estimates for the individual provinces. Discussions on various allocation strategies are presented in papers by Platek, Singh and Gray [13], and Singh and Lawes [17].

Ces trois premières étapes ont permis la création de la N.E.P.A., après quoi on a pu procéder au DÉROULEMENT SIMULTANÉ DES DEUX ENQUÊTES et, par la suite, à la dernière étape, soit l'élargissement de l'échantillon. Les deux enquêtes ont été menées simultanément pendant toute l'année civile 1975, et indépendamment l'une de l'autre, pour obtenir des données permettant la comparaison directe des résultats afin d'établir des liens entre les deux enquêtes et de procéder à des rajustements saisonniers. L'élargissement de l'échantillon, prévu pour 1976 - 1977 permettra d'obtenir des données plus détaillées sur les petites régions. L'échantillon passera de 35,000 ménages approximativement à 55,000 ménages par mois. La répartition de cette augmentation entre les diverses provinces avait pour objectif principal de réduire les différences de fiabilité des estimations relatives au taux de chômage de chaque province. Les rapports de Platek, Singh et Gray [13] et de Singh et Lawes [17] contiennent des analyses de divers modes de répartition.

CHAPTER – 4 – CHAPITRE

STRATIFICATION

4.1 Introduction

Stratification in an area frame is basically a process of classifying (usually compact) area units into certain collections called strata. Though the main advantage of stratified sampling is the possible increase in efficiency per unit cost in estimating the population characteristics, stratification also introduces considerable flexibility in the sense that, depending upon the information available, sampling and estimation procedures may differ from stratum to stratum. Further, in a continuous survey like the LFS, stratification provides an added flexibility of updating or redesigning the sample of a specified stratum or groups of strata, without affecting the design in the remaining strata.

Each of the ten provinces in Canada is divided into a number of economic regions (ERs). An ER has areas of similar economic structure formed on the basis of recent information and are stable over a period of time. These ERs are treated as primary strata and further stratification is carried out within the self-representing unit (SRU) and non-self-representing unit (NSRU) parts independently in each ER. Two main reasons for stratifying ERs separately were:

(i) use of ERs as primary strata is more manageable than the province as a whole, and

(ii) it would facilitate obtaining sub-provincial estimates in future.

The extent of stratification carried out in the two parts of the LFS, viz., self-representing units (SRUs) and non-self-representing units (NSRUs) varies greatly. Optimum stratification is carried out within NSRU portions of the ERs (on the basis of major LF characteristics), while within SRU areas subunits are delineated on the basis of geography, dwelling counts and potential for growth.

After presenting a brief discussion on the formation of subunits (sub-strata) in the SR areas in section 2, the remaining sections (3-5) of this chapter are devoted to a general description of stratification and Primary Sampling Unit (PSU) formation in NSR areas of the survey. Section 3 deals with the constraints that are imposed on the stratification and PSU formation in the context of the total survey design such as the choice of stratification variables and the size of strata and PSUs. Contiguity and other operational constraints

STRATIFICATION

4.1 Introduction

La stratification de la base d'un sondage aréolaire consiste essentiellement à regrouper les unités (généralement compactes) dans des ensembles appelés strates. L'efficacité accrue, par unité de coût, de l'estimation des caractéristiques de la population constitue le principal avantage de la stratification bien que cette dernière donne également une flexibilité considérable en ce sens que, selon les renseignements disponibles, les méthodes d'échantillonnage et d'estimation peuvent différer d'une strate à l'autre. En outre, pour une enquête permanente comme l'E.P.A., la stratification permet de mettre à jour ou de remanier l'échantillon d'une strate donnée ou de plusieurs groupes de strates, sans nuire à la conception adoptée pour les autres strates.

Chacune des dix provinces canadiennes est divisée en un certain nombre de régions économiques (R.É.). Une R.É. est composée de secteurs ayant des structures économiques semblables, selon des renseignements récents, et demeurant stables pendant un certain temps. Ces R.É. sont considérées comme des strates primaires et on procède à une nouvelle stratification à l'intérieur des régions autoreprésentatives (A.R.) et des régions non autoreprésentatives (N.A.P.), indépendanmment dans chaque R.É. Les deux principales raisons pour lesquelles les R.É. sont stratifiées séparément sont les suivantes:

(i) il est plus pratique d'utiliser les R.É. comme strates primaires que d'utiliser toute la province, et

(ii) cela permettrait ultérieurement d'obtenir plus facilement des estimations infraprovinciales.

L'ampleur de la stratification faite dans les deux parties de l'E.P.A., c'est-à-dire les unités autoreprésentatives (U.A.R.) et les unités non autoreprésentatives (U.N.A.R.), varie considérablement. Dans la partie N.A.R. des R.É., on procède à une stratification optimale par (rapport aux caractéristiques principales de la population active), tandis que les régions A.R. sont subdivisées en sous-unités en fonction de leur situation géographique, du nombre de logements et de leurs possibilités d'accroissement.

Le paragraphe 2 du présent chapitre présente une brève analyse de la formation des sous-unités (les sous-strates) dans les U.A.R. les autres paragraphes (3-5) donnent une description globale de la stratification et de la formation des unités primaires d'échantillonnage (U.P.É.) dans les régions N.A.R. Le paragraphe 3 indique les contraintes que la conception globale de l'enquête impose pour la stratification et la formation des U.P.É., notamment quant au choix des variables de stratification et de la taille des strates et des U.P.É. Ce paragraphe traite

are also discussed in this section. The two methods of stratification, viz. manual method, pattern analysis and automated system are briefly described in section 4. Lastly, formation of PSUs in NSR areas is discussed in section 5. A step by step account of various operations involved may be found in the "Sampling Manual".

4.2 Classification of SR Areas

4.2.1 Lower Limit

For operational and theoretical reasons, the lower limit (size of the smallest SRU) in the new design is made to vary from province to province. The guideline for classifying a city as SRU was that the city should yield an expected take of 20 households, the minimum acceptable assignment size for an interviewer. The lower limits were thus obtained as a product of inverse of sampling ratio in the province x 20 x 4, where the average household size is assumed to be four. For example, this limit was 24,000 (300 x 20 x 4) persons, in Ontario and Quebec and 14,500 (180 x 20 x 4) persons in Manitoba, Saskatchewan and Alberta. Some cities with population less than the lower limit determined for the province were also classified as SRUs in cases where they possess unique labour force characteristics. Further, in some cases smaller urban centers were combined to form an SRU due to factors such as their unique characteristics, the distance between the urban centers, or the total urban and rural population in the ER.

4.2.2 Formation of Subunits

Larger SRU, are subdivided into subunits (strata). The subunit sizes vary, from province to province, ranging from 1,000 to 12,000 dwellings. These subunits ensure geographical representativeness and provide a manageable sized unit which may be redesigned separately when the need arises.

By definition, Census Tracts are small permanent units established in cities with a minimum population of 50,000. A Census tract has a population between 2,500 and 8,000, and it is an area that is as homogeneous as possible in terms of economic status and social living conditions, and has easily recognizable boundaries. In the SRU with census tracts, subunits were comprised of contiguous census tracts, while smaller SRUs were divided into subunits on the basis of field counts by city blocks and block faces. These field counts were obtained for both areas, and formed the basis for subsequent delineation of sampling units.

également du problème de la contiguïté et d'autres contraintes d'ordre pratique. Le paragraphe 4 décrit les deux méthodes de stratification, soit la méthode non automatisée, ou analyse de modèle, et le système automatisé. Pour finir, le paragraphe 5 analyse la formation des U.P.É. dans les régions N.A.R. Le "Manuel d'échantillonnage" donne un compte rendu détaillé des diverses opérations effectuées.

4.2 Classification des U.A.R.

4.2.1 Limite inférieure

Pour des raisons opérationnelles et théoriques, la limite inférieure (taille de la plus petite U.A.R.) établie dans le nouveau plan de sondage varie d'une province à l'autre. Pour qu'une cité soit classée comme U.A.R., elle doit avoir un rendement prévu de 20 ménages, ce qui constitue le nombre minimal de ménages à confier à un interviewer. Les limites inférieures ont donc été obtenues en multipliant la fraction de sondage inverse de la province par 20 et par 4, en fonction d'un ménage moyen composé de quatre personnes. Par exemple, cette limite est de 24,000 personnes (300 x 20 x 4), en Ontario et au Québec, et de 14,500 personnes (180 x 20 x 4) au Manitoba, en Saskatchewan et en Alberta. Certaines cités ayant une population plus petite que la limite inférieure fixée pour la province ont également été classées comme des U.A.R. parce qu'elles présentent des caractéristiques de la population active uniques en leur genre. En outre, certains petits centres urbains ont été groupés en une seule U.A.R. notamment à cause de leurs caractéristiques particulières, de la distance qui les sépare, ou de la population urbaine et rurale totale de la R.É.

4.2.2 Formation des sous-unités

Les grandes U.A.R. sont subdivisées en sous-unités (les strates). La taille des sous-unités varie entre 1,000 et 12,000 logements d'une province à l'autre. Ces sous-unités assurent la représentativité géographique et ont l'avantage d'être maniables et de pouvoir, au besoin, être remaniées séparément.

Par définition, les secteurs de recensement sont de petites unités permanentes établies dans des cités ayant une population d'au moins 50,000 habitants. Le secteur de recensement a une population variant de 2,500 à 8,000 habitants dont le statut économique et les conditions de vie sociale sont aussi homogènes que possible, et a des limites facilement reconnaissables. Dans les U.A.R. divisées en secteurs de recensement, les sous-unités se composent de secteurs de recensement contigus, et dans les petites U.A.R., d'un certain nombre d'îlots urbains et de côtés d'îlots selon les comptes sur le terrain. Ces dénombrements sont faits pour les deux types D'U.A.R. et constituent la base de la délimitation ultérieure des unités d'échantillonnage.

The subunits of each SRU are further classified as built-up, fringe, and combinations of built-up and fringe, depending on potential for future population growth. This classification is important for two reasons: firstly it identifies the subunits which require closer attention for updating in the future, secondly, it helps to minimize the variation within subunits, as households in core areas of the larger cities are likely to have different labour force characteristics than those in the fringe areas.

Further, in 17 larger cities, a separate frame of apartments has been set up for apartment buildings having at least five storeys and 30 or more units. Details of stratification of this frame are given in Chapter 7.

4.3 NSR Design Constraints on Stratification and PSU Delineation

4.3.1 Choice of Stratification Variables

It is important that the variables (characteristics) upon which the stratification is based meet certain requirements if the stratification is to result in the increase of design efficiency. These requirements for the RLFS were as follows:

(i) The stratification variables should be related to the variables under study. In this case the stratification is intended specifically for the LFS, therefore, the stratification variables should be related to the characteristics of the labour force,

(ii) The characteristics should be stable over time in order to retain the efficiency of stratification for a longer period of time.

(iii) The number of persons having the characteristic should vary from area to area within the ER making meaningful the concept of similar and dissimilar areas with respect to the characteristic.

(iv) The number of persons having the characteristic should account for a sizeable proportion of the ER population.

Following these guidelines, the proportions of the labour force employed by industry were considered as possible stratification variables for an ER, fulfilling requirements (i) and (ii). The labour force was partitioned into seven categories: agriculture, forestry or fishing, mining, manufacturing, construction, transportation, and services, hence-forth referred to as the labour force characteristics.

Census Enumeration Areas (EAs) belonging to NSR areas of an ER were identified (i.e., EAs belonging to SRUs, Indian Reserves and special areas were excluded). Data at the EA level with sub-totals at appropriate levels within each ER was retrieved from Census

Les sous-unités de toutes les U.A.R. sont par la suite classées comme zone bâtie, banlieue et mélange de zone bâtie et de banlieue, selon les possibilités d'accroissement futur de leur population. Cette classification est importante pour deux raisons: premièrement, elle indique les sous-unités qu'il faudra surveiller de plus près en vue d'une mise à jour ultérieure, deuxièmement, elle aide à amoindrir la variation à l'intérieur des sous-unités, puisque les ménages situés au coeur des grandes cités sont susceptibles de présenter des caractéristiques de la population active différentes de celles des banlieues.

En outre, dans 17 grandes cités, on a établi une base d'immeubles d'appartements distincte pour les immeubles d'appartements d'au moins cinq étages et 30 unités. Le chapitre 7 donne des détails sur la stratification de cette base.

4.3 Contraintes imposées par le plan de sondage des U.N.A.R. pour la stratification et la délimitation des U.P.É.

4.3.1 Choix des variables de stratification

Les variables (caractéristiques) en fonction desquelles on procède à la stratification doivent nécessairement répondre à certaines exigences si l'on veut que la stratification augmente l'efficacité du plan de sondage de l'enquête. Dans le cas de la N.E.P.A. ces exigences sont les suivantes:

(i) Les variables de stratification doivent être apparentées aux variables qui font l'objet de l'étude. Dans le cas présent, la stratification est particulièrement destinée à l'E.P.A., par conséquent, les variables de stratification doivent être apparentées aux caractéristiques de la population active.

(ii) Ces caractéristiques doivent être stables pour que l'efficacité de la stratification se maintienne plus longtemps.

(iii) Le nombre de personnes qui présentent une certaine caractéristique doit varier d'une partie à l'autre de la R.É., pour que le concept de régions semblables et différentes soit justifié quant à cette caractéristique.

(iv) Le nombre de personnes qui présentent une certaine caractéristique doit représenter une proportion importante de la population de la R.É.

Suivant ces principes directeurs, on a considéré les fractions de la population active employée par activité économique comme des variables de stratification possibles pour une R.É. puisqu'elles satisfont les exigences (i) et (ii). La population active a été divisée en sept catégories: agriculture, forêts ou pêche, industries minières, industries manufacturières, bâtiment et travaux publics, transports et services, que l'on considère comme des caractéristiques de la population active.

On a fait un relevé des secteurs de dénombrement du recensement (S.D.) faisant partie des U.N.A.R. d'une R.É. (ainsi les S.D. des U.A.R., des réserves indiennes et des secteurs spéciaux ont été exclus). Des données relatives aux S.D. et certains de leurs sous-totaux dans chaque

files by use of STATPAK, a data retrieval system. From the seven labour force characteristics for each ER, the three characteristics best fulfilling conditions (iii) and (iv) were used as stratification variables. As a measure of this, the importance factor (IF) was defined for the ℓth characteristic as the variance in the proportion employed in characteristic ℓ at the EA level weighted by the EA population multiplied by the proportion in the ER having characteristic ℓ. The weighted variance expression is given by:

$$\sigma^2_\ell = \sum_{j=1}^{n} \frac{N_{.j}}{N_{..}} (P_{\ell j} - P_{\ell .})^2$$

where

ℓ = characteristic

j = EA

$N_{\ell j}$ = population having characteristic ℓ in EA j

$N_{..}$ = population of ER

$P_{\ell j} = N_{\ell j}/N_{.j}$ = proportion of labour force in EA j having characteristic ℓ

$P_{\ell .} = N_{\ell .}/N_{..}$ = proportion having characteristic ℓ in ER

Then the importance factor is given by:

$$IF_\ell = P_{\ell .}\, \sigma^2_\ell$$

A list of the important characteristics for each of of the ERs is provided in appendice.

It should be noted that in the Prairie Provinces due to the high proportion engaged in agriculture, stratification variables used were related to different types of agricultural activities.

4.3.2 Determination of Stratum and PSU Sizes

Two design considerations determined guidelines for PSU and stratum sizes from province to province. First, studies and consultation with field representatives indicated that an interviewer's assignment should be about 50 households. The second consideration was the anticipated increase in the size of the LFS. In order to retain the option of accomplishing such an increase by selecting more PSUs, it was desirable that there be a sufficient number of PSUs to draw from in each stratum. Adherence to these criteria leads to determination of PSU and stratum sizes by province in the following manner.

R.É. ont été obtenues des dossiers du recensement à l'aide d'un système de retrait des données, le STATPAK. Sur les sept caractéristiques de la population active, on a utilisé comme variables de stratification pour chaque R.É. les trois caractéristiques qui remplissaient le mieux les conditions (iii) et (iv). Pour déterminer cela, on a défini le facteur d'importance (F.I.) pour la ℓe caractéristique comme étant la variance dans la proportion employée dans la caractéristique ℓ pour le S.D., pondérée par la population du S.D., multipliée par la proportion de la R.É. présentant la caractéristique ℓ. La variance pondérée s'obtient par le calcul suivant:

$$\sigma^2_\ell = \sum_{j=1}^{n} \frac{N_{.j}}{N_{..}} (P_{\ell j} - P_{\ell .})^2$$

où

ℓ = caractéristique

j = S.D.

$N_{\ell j}$ = population ayant la caractéristique ℓ dans le S.D.j.

$N_{..}$ = population de la R.É.

$P_{\ell j} = N_{\ell j}/N_{.j}$ = proportion de la population active dans le S.D.j. ayant la caractéristique ℓ

$P_{\ell .} = N_{\ell .}/N_{..}$ = proportion de la population active ayant la caractéristique ℓ dans la R.É.

Le facteur d'importance s'obtient par le calcul suivant:

$$F.I._\ell = P_{\ell .}\, \sigma^2_\ell$$

L'annexe donne une liste des caractéristiques importantes de toutes les R.É.

Il est à noter qu'en raison de la forte proportion de la population travaillant dans l'agriculture dans les provinces des Prairies, on a utilisé des variables de stratification ayant trait aux divers types d'activité agricole.

4.3.2 Établissement de la taille de la strate et de l'U.P.É.

La taille des U.P.É. et des strates de chaque province a été déterminée en fonction de ce que devrait être la tâche d'un interviewer, selon les études et les consultations avec des préposés sur le terrain, soit 50 ménages environ, et en fonction de l'élargissement prévu de la taille de l'E.P.A. Il est souhaitable que chaque strate contienne un nombre suffisant d'U.P.É. pour permettre l'élargissement de l'enquête par le choix d'un plus grand nombre d'U.P.É. Ces critères permettent de déterminer la taille des U.P.É. et des strates de chaque province de la manière décrite ci-après.

For province p, if we assume we have n_p^* PSUs per stratum, all of approximately equal sizes, with two PSUs selected, then the inverse sub-sampling ratio within the selected PSU is given by:

$$R_{pi} = \Pi_i R_p$$

where R_p = inverse provincial sampling ratio and Π_i = probability of inclusion of ith PSU ($\Pi_i \doteq 2/n_p^* = \Pi$). Taking into consideration field requirements, the average PSU size, N_p^*, should yield a sample of about 50 households, which is equivalent to one interviewer's assignment size (E) then we have:

$$N_p^*/R_{pi} = m\,E$$

i.e.:

$$N_p^* = m\,E\,\Pi\,R_p$$

where m ñ number of persons per household (≈4).

It was decided that the average number of PSUs, n_p^* per stratum should be about 1.5 except in the Maritimes where $n_p^* = 12$, to offset the effect there of a low provincial inverse sampling ratio. The exact number of PSUs in any stratum was allowed to vary provided the size of each PSU was within the allowable limits as described below. The number of PSUs by stratum are given in the appendix for each ER of a province.

The average stratum size within a province was obtained as:

$$N_p' = n_p^* N_p^*.$$

For each ER, the number of strata to be formed in the NSR area was based only on the NSRU population of the ER. While N_p^* and N_p' above correspond to average PSU and stratum sizes respectively, the actual sizes were allowed to vary within a specified range.

Table 4.1 gives ranges and average values for the sizes of PSUs and strata as they differ across Canada.

Pour la province p, si nous considérons que nous avons n_p^* U.P.É. par strate, de tailles sensiblement égales, et que deux U.P.É. seront choisies, la fraction de sous-sondage inverse dans l'U.P.É. choisie s'obtient par le calcul suivant:

$$R_{pi} = \Pi_i R_p$$

où R_p = la fraction de sondage inverse provinciale et Π_i = la probabilité que soit incluse la i^e U.P.É. ($\Pi_i \doteq 2/n_p^* = \Pi$). Si l'on tient compte des exigences du travail sur le terrain, l'U.P.É. de taille moyenne N_p^* doit donner un échantillon d'environ 50 ménages, ce qui équivaut à la tâche d'interviewer (E). Nous avons alors:

$$N_p^*/R_{pi} = m\ E$$

soit:

$$N_p^* = m\ E\ \Pi\ R_p$$

où m = le nombre de personnes composant le ménage (≈4).

On a décidé que le nombre moyen d'U.P.É. n_p^* par strate devrait être de 15, sauf dans les Maritimes où $n^*p = 12$, pour contrebalancer l'effet de la faible fraction de sondage inverse provinciale. Le nombre exact d'U.P.É. composant chaque strate peut varier, pourvu que la taille de chaque U.P.É. respecte les limites décrites ci-dessous. L'annexe donne le nombre d'U.P.É. par strate pour chaque R.É. d'une province.

La taille moyenne de la strate d'une province s'obtient de la façon suivante:

$$N_p' = n_p^* N_p^*.$$

Pour chaque R.É. le nombre de strates à former dans la partie NAR ne dépend que de la population des U.N.A.R. de la R.É. Même si N_p^* et N_p' calculés précédemment correspondent à la taille moyenne de l'U.P.É. et de la strate, leur taille réelle peut varier selon un intervalle de variation précis.

Le tableau 4.1 indique les divers intervalles de variation et les moyennes établies pour la taille des U.P.É. et des strates, car ils diffèrent d'une province canadienne à l'autre.

TABLE 4.1. PSU and Stratum Size

TABLEAU 4.1. Taille de l'U.P.É. et de la strate

Region – Région	R_p province ISR – F.S.I. province	R_p PSU ISR – E.S.I. de l'U.P.É.	N_p^* PSU size – Taille de l'U.P.É.	PSU range – Intervalle de variation de l'U.P.É.	N_p' Stratum size – Taille de la strate	Stratum range – Intervalle de variation de la strate
Maritimes	62.5	2/12 x 62.5	2,000	1,700 - 2,300	25,000	20,000 - 30,000
Ontario - Québec	200	2/15 x 200	5,300	4,600 - 6,100	80,000	64,000 - 96,000
Prairies	125	2/15 x 125	3,300	2,900 - 3,800	50,000	40,000 - 60,000
British Columbia – Colombie-Britannique	143	2/15 x 143	3,800	3,300 - 4,300	56,000	45,000 - 68,000

4.3.3 Contiguity Constraints

The optimal grouping of Census EAs from a variance reduction point of view need not necessarily yield contiguous strata, and when they are forced to form contiguous strata, some loss of efficiency could result. On the other hand discontiguous strata could result in higher interviewer's travel costs. As a compromise, partially contiguous stratification was thought to be a practical alternative where contiguity constraints were to be relaxed to the extent that any collection of EAs sufficient to form at least one PSU could remain on its own as a part of a stratum. Some experiments were conducted to compare the efficiency of the partially contiguous strata and the contiguous strata, however, no definite trend could be established for predicting whether a significant gain in efficiency would result from partially contiguous strata without actually forming the two types of strata individually for each ER and comparing the efficiencies. Therefore, on the assumption that contiguous strata are more likely to retain homogeneity over a period of time than if they were discontiguous and that user's requirements for small area estimates, wherever applicable, may be better served through contiguous strata, this type of stratification was adopted in the LFS.

4.3.3 Contraintes imposées par la contiguité

Le groupement optimal des secteurs de S.D. en vue de réduire la variance ne doit pas nécessairement donner des strates contiguës, et si l'on tente de former des strates contiguës, on peut nuire à l'efficacité du groupement. D'autre part, si les strates ne sont pas contiguës, les frais de voyage de l'interviewer peuvent être plus élevés. Un compromis aurait été la stratification partiellement contiguë, où tout groupement de S.D. suffisant pour former au moins une U.P.É. peut faire partie d'une strate. Pour comparer l'efficacité des deux modes de regroupement, on a tenté quelques expériences, mais on n'a pas pu définir de tendance précise et prévoir si la formation de strates partiellement contiguës s'avèrerait plus efficace que celle de strates contiguës, sans réellement former ces deux types de strates séparément pour chaque R.É. et comparer leur efficacité respective. Par conséquent, en supposant que les strates contiguës sont plus susceptibles de rester homogènes pendant un certain temps que les strates qui ne le sont pas et qu'elles permettent mieux d'établir les estimations relatives aux petites régions dont ont besoin les utilisateurs, on a adopté ce genre de stratification dans l'E.P.A.

4.4 Methods of Stratification

Basically two methods of stratification were examined: pattern analysis (manual) and an automatic system. In the manual method, the unit chosen for stratification was the "combined municipality" which is the geographical area enclosed within a rural municipality and as such often contains within its boundaries urban municipalities which are geographically smaller. The combined municipality to which an EA belongs was obtainable from the Census EA descriptions so that it was possible to retrieve tabulations at the combined municipal level. The number of combined municipalities in an ER varies from about 10 - 60.

4.4 Méthodes de stratification

On a essentiellement étudié deux méthodes de stratification: l'analyse de modèle (manuelle) et le système automatisé. Dans la méthode non automatisée, l'unité choisie pour la stratification est la "municipalité groupée", qui représente la région géographique comprise dans une municipalité rurale et qui, par conséquent, renferme souvent des municipalités urbaines géographiquement plus petites. Les descriptions des S.D. du recensement indiquent les municipalités groupées auxquelles les S.D. appartiennent et ont permis d'obtenir des données au niveau des municipalités groupées. Les R.É. peuvent contenir de 10 - 60 municipalités groupées.

The printouts of Census Data contained $N_{\ell j}$ the population having characteristic ℓ in ith EA, $P_{\ell j}$ the proportion having characteristic ℓ in ith EA, and $P_{\ell j}/P_{\ell}$ the proportion having characteristic ℓ in ith EA divided by the proportion having characteristic ℓ in the ER. In addition similar counts and proportions were obtained at the combined municipality level. The method of importance factors (IF_{ℓ} in 4.3.2) was used for determination of the three most important characteristics for each ER.

The manual method of stratification involved a procedure known as "pattern analysis". The $P_{\ell i}/P_{\ell}$ ratios for combined municipalities (i = municipality) were plotted for the three most important characteristics. This ratio compared the proportion of the labour force engaged in a characteristic in a combined municipality to the proportion in the ER as a whole. Any departure from a value of one for this ratio indicated the relative departure from similarity between the combined municipality and the ER. These three ratios were plotted graphically and joined by straight lines. The angles formed between the two resulting straight lines could be classified into nine different patterns. Rules which depended on the number of strata to be formed were adopted to group together certain subsets of the patterns. Combined municipalities were thus assigned to one of these subsets on the basis of their patterns, and this formed the initial approximation to the final strata. Adjustments were usually necessary in order to satisfy the size and contiguity constraints on strata. Further, it was also necessary that each stratum should have a reasonable proportion of urban population to facilitate PSU formation.

The final strata were composed then of a contiguous collection of combined municipalities such that each stratum had a population within predetermined limits. These strata were demarcated on large ER maps which were subsequently used for PSU delineation.

It should be noted that in Newfoundland, due to its unique demographic characteristics, it was decided to adopt a rural-urban stratification. In each ER, one or two urban strata composed of all NSR centers having a population of at least 500 were formed, and in the remainder of the NSR areas contiguous rural strata were formed by the method of pattern analysis.

An automated system of stratification was also developed and an evaluation of the two methods showed that they were roughly equally efficient. Manual or automated systems were used for stratification in individual ERs depending on the operational suitability.

Des relevés mécanographiques des données du recensement, on obtient $N_{\ell j}$, la population qui présente la caractéristique ℓ dans l'iᵉ S.D., $P_{\ell j}$, la proportion de la population qui présente la caractéristique ℓ dans l'iᵉ S.D., et le rapport $P_{\ell j}/P_{\ell}$, la proportion de la population qui présente la caractéristique ℓ dans l'iᵉ S.D. divisée par la proportion de la population qui présente la caractéristique ℓ dans la R.É. En outre, on a obtenu des calculs et des rapports similaires pour les municipalités groupées. Pour déterminer les trois caractéristiques les plus importantes de chaque R.É., on a utilisé la méthode du facteur d'importance (F.I.ℓ), décrite au paragraphe 4.3.2.

La méthode non automatisée utilise "l'analyse de modèle". Le rapport $P_{\ell i}/P_{\ell}$ pour les municipalités groupées (i = municipalité) a été calculé pour les trois caractéristiques les plus importantes. Ce rapport compare la proportion de la population active présentant une caractéristique donnée dans une municipalité groupée, à la même proportion dans toute la R.É. Tout écart de la valeur de un pour ce rapport indique une dissemblance relative entre la municipalité groupée et la R.É. Ces trois rapport ont été portés sur un diagramme et reliés par des lignes droites. On peut représenter par neuf tendances différentes les angles obtenus par la rencontre de ces deux lignes droites. On a adopté des règles, établies en fonction du nombre de strates à former, pour réunir certains sous-ensembles de ces tendances. Ainsi, les municipalités groupées ont été classées dans l'un de ces sous-ensembles, suivant les tendances qui y ont été décelées, et ceci constitue l'esquisse de la strate finale. En général, il a fallu procéder à certains rajustements pour satisfaire les exigences relatives à la taille et à la contiguité des strates. En outre, il fallait également que chaque strate ait une proportion raisonnable de population urbaine pour faciliter la formation d'U.P.É.

Par conséquent, les strates finales sont composées d'une série de municipalités groupées contiguës et la population de chaque strate doit respecter certaines limites déterminées à l'avance. Ces strates sont tracées sur de grandes cartes des R.É. que l'on a par la suite utilisées pour délimiter les U.P.É.

Il est à noter qu'en raison des caractéristiques démographiques uniques de Terre-Neuve, on a décidé d'y adopter une stratification rurale et urbaine. Dans chaque R.É., on a formé une ou deux strates urbaines composées de toutes les U.N.A.R. ayant une population d'au moins 500 habitants et, dans les autres régions N.A.R., on a formé des strates rurales contiguës selon la méthode de l'analyse de modèle.

On a également utilisé un système de stratification automatisé et l'évaluation des deux méthodes a prouvé qu'elles sont d'une efficacité comparable. La stratification des R.É. s'est faite à l'aide des système automatisé ou non automatisé selon que l'un ou l'autre s'avérait le plus pratique.

4.5 Primary Sampling Units

4.5.1 Basic Features

Having obtained the average PSU sizes as given in section 4.3.2 for individual provinces, geographically contiguous and compact PSUs were delineated within each stratum. Each PSU represented, to the extent possible, the stratum with respect to the important LF characteristics. The following criteria were used in their formation:

(i) For each of the three important LF characteristics, the percentage of the labour force engaged in a given characteristic in each PSU should not deviate from the stratum percentage by more than 5%.

(ii) The population of the PSU should lie within the size constraints, i.e., between .86 N_p^* and 1.14 N_p^*, where N_p^* is the average provincial PSU size.

(iii) The proportion of urban and rural population in each PSU should not deviate from the corresponding stratum proportion by more than 5%. Depending on the size of the urban centers, it was often necessary to allocate (split) them amongst more than one PSU in order to maintain the stratum rural-urban ratio. Such splitting was not done geographically but on a fractional basis. Later, at the time of selection, if a PSU sharing a proportion T_{ij} of an urban center (j) was selected, the entire urban center would be sampled with an inverse sampling ratio of T_{ij} R_i, where R_i is the PSU inverse sampling ratio. The splitting of urban centers was avoided except when necessary due to size or uniqueness of its LF characteristics. Further, when the urban part of a PSU was made up of more than one urban center, it was important to ensure that these centers were close to one another and with sufficient communication facilities between them. It may be noted that all places that had been designated as urban by Census were not necessarily considered as urban centers for the purpose of LFS design (see section 6.3.1). A cut-off point in the range of 200-500 was established for different provinces and only those census designated urban places were considered as urban centers which had a population higher than the cut-off point. In a few instances however where the urban population in the stratum was very small and the characteristics of the urban places were sufficiently unique, they were considered as urban even though their size was below the cut-off point.

4.5.2 Delineation of PSUs

Since each of the PSUs to be formed may vary in size within the range ($.86N_p^*$, $1.14N_p^*$), then there was some flexibility in the number of PSUs, n_{ph}, that can be formed in stratum h, viz.:

4.5 Unités d'échantillonnage primaire

4.5.1 Caractéristiques fondamentales

Une fois la taille moyenne de l'U.P.É. obtenue par province, comme l'indique le paragraphe 4.3.2, on délimite des U.P.É. contiguës et compactes dans chaque strate. Chaque U.P.É. présente, dans la mesure du possible, les caractéristiques dominantes de la strate. Les critères suivants ont servi à les former:

(i) Pour les trois caractéristiques dominantes, l'écart entre le pourcentage de la population active présentant une caractéristique donnée dans chaque U.P.É. et dans la strate à laquelle elle appartient doit être inférieur à 5 %.

(ii) La population de l'U.P.É. doit respecter les limites fixées pour la taille de l'U.P.É., c'est-à-dire qu'elle doit varier entre .86 N_p^* et 1.14 N_p^*, N_p^* étant la taille moyenne de l'U.P.É. pour la province.

(iii) L'écart entre la proportion de population urbaine et rurale de chaque U.P.É. et celle de la strate à laquelle elle appartient doit être inférieur à 5 %. Les plus grands centres urbains ont dû souvent être répartis (divisés) entre plus d'une U.P.É., afin de répondre à cette exigence. Il ne s'agit pas d'une division géographique, mais d'un fractionnement. Ensuite, si au moment de la sélection on a choisi une U.P.É. comprenant une fraction T_{ij} d'un centre urbain, l'ensemble du centre urbain est échantillonné selon la fraction de sondage inverse de T_{ij} R_i, R_i représentant la fraction de sondage inverse de l'U.P.É. On a évité de diviser les centres urbains, sauf lorsque leur taille ou les caractéristiques de leur population active le justifiaient. En outre, si la partie urbaine de l'U.P.É. était composée de plus d'un centre urbain, il a fallu s'assurer que ces centres soient proches les uns des autres et que des moyens de communication suffisants les relient. Il est à noter que le plan de sondage de l'E.P.A. n'a pas nécessairement considéré comme des centres urbains toutes les régions désignées urbaines aux fins du recensement (voir le paragraphe 6.3.1). On a fixé pour chaque province un point limite, dans l'intervalle de 200-500 habitants, et seules les régions désignées urbaines aux fins du recensement, dont la population était supérieure à cette limite, ont été considérées comme des centres urbains. Cependant, dans quelques cas où la population urbaine de la strate était très petite et où les caractéristiques des régions urbaines étaient suffisamment particulières, on les a considérées comme des centres urbains, même si leur taille était inférieure aux limites fixées.

4.5.2 Délimitation des U.P.É.

Étant donné que la taille des U.P.É. que l'on devait former pouvait varier entre $.86N_p^*$ et 1.14 N_p^*, le nombre d'U.P.É., n_{ph}, que l'on pouvait former dans la strate h était quelque peu flexible, soit:

$$(N_{ph}/1.14\ N_p^*) \leqslant n_{ph} \leqslant (N_{ph}/.86\ N_p^*)$$

Where N_{ph} = population of stratum h.

An example: N_{ph} = 80,000, N_p^* = 5,300, then n_{ph} will lie between 13 and 17.

For a given value of n_{ph}, in order to satisfy (ii) and (iii) (section 4.5.1) simultaneously an amount of urban given by $(1 \pm .10)\ N_{phi}^{(u)}$ (where $N_{ph}^{(u)}$ = urban population of stratum) would be assigned to each PSU. All urban centers whose population did not lie within this range were split amongst more than one PSU. Thus n_{ph} was chosen to minimize the splitting necessary to allocate an acceptable amount of urban to each PSU.

Having decided on n_{ph} and on the urban parts, an appropriate amount of rural population had to be assigned to each urban part in order to balance the proportions between the two types of area. While the rural area allocated to an urban should be as near to the urban parts as possible, in some cases the distribution of urban throughout a stratum was such that there could still be considerable distance between the rural and urban parts of a PSU. The rural part of a PSU consisted then of a collection of contiguous EAs whose population, $N_{phi}^{(r)}$ was in the range

$$.95\ N_{phi}^{(u)}\ S_{ph} \leqslant N_{phi}^{(r)} \leqslant 1.05\ N_{phi}^{(u)}\ S_{ph}$$

where S_{ph} is the ratio of rural and urban populations in stratum h, province p.

Due to the rural-urban stratification adopted in Newfoundland, formation of urban PSUs is analogous to splitting of the urban into n_{ph} parts having population N_{ph} while formation of the rural PSUs amounts to grouping together rural EAs on the basis of population and contiguity. In both cases adjustments to take into account characteristic differences were necessary.

$$(N_{ph}/1.14\ N_p^*) \leqslant n_{ph} \leqslant (N_{ph}/.86\ N_p^*)$$

où N_{ph} = représente la population de la strate h.

Par exemple: si N_{ph} = 80,000 et N_p^* = 5,300, n_{ph} se situera entre 13 et 17.

Pour une valeur donnée de n_{ph}, il faut, afin de remplir les conditions (ii) et (iii) (voir le paragraphe 4.5.1) simultanément, attribuer à chaque U.P.É. une portion de population urbaine qui est dans l'intervalle $(1 \pm .10)$ $N_{phi}^{(u)}$ (où $N_{ph}^{(u)}$ = la population urbaine de la strate). Tous les centres urbains dont la population ne se situait pas dans cet intervalle ont été divisés entre plus d'une U.P.É. Par conséquent, n_{ph} a été choisi pour amoindrir le fractionnement nécessaire pour répartir une quantité acceptable de la population urbaine entre chaque U.P.É.

Une fois n_{ph} et les zones urbaines déterminés une quantité appropriée de la population rurale doit être affectée à chaque zone urbaine pour équilibrer les rapports entre ces deux types de régions. Même si la région rurale affectée à une zone urbaine doit être aussi proche que possible de cette dernière, dans certains cas la répartition de la population urbaine d'une strate était telle qu'il pouvait encore y avoir une distance considérable entre les zones rurales et urbaines d'une U.P.É. La partie rurale d'une U.P.É. se composait alors d'un ensemble de S.D. contigus dont la population $N_{phi}^{(r)}$ avait un intervalle de variation donné par

$$.95\ N_{phi}^{(u)}\ S_{ph} \leqslant N_{phi}^{(r)} \leqslant 1.05\ N_{phi}^{(u)}\ S_{ph}$$

où S_{ph} représentait le quotient des populations rurales et urbaines de la strate h, de la province p;

En raison de la stratification rurale et urbaine adoptée à Terre-Neuve, la formation des U.P.É. urbaines est analogue au fractionnement de la population urbaine en n_{ph} parties ayant une population N_{ph}/n_{ph}, tandis que la formation des U.P.É. rurales consiste en un regroupement des S.D. ruraux en fonction de la population et de la contiguité. Dans les deux cas, il a fallu faire des rajustements pour tenir compte des caractéristiques différentes.

CHAPTER – 5 – CHAPITRE

SELF-REPRESENTING UNITS

5.1 General Description

A stratified two-stage sampling design has been adopted in SRUs. The strata in SRUs (see section 4.2) are referred to as subunits (or also as PSUs for administrative purposes), which are further sub-divided into clusters (normally a city block). The cluster is the first stage of selection within subunits, the second and final stage of selection being the dwelling.

The subunits are classified into one of the following three types on the basis of their potential for future growth: type 1, built-up; type 2, fringe; type 3, combination (built-up and fringe).

In each subunit, six (or in some cases 12) "groups" of clusters are formed at random, and each group being either a built-up group (i.e., random collection of only built-up clusters) or an analogously fringe group. In the case of type 3 subunits, in order to facilitate variance estimation it was ensured that both built-up and fringe groups were in equal proportion. Further, a special sample of apartments was set up in 17 larger cities for apartment buildings having at least five storeys and 30 or more units, thus removing this problem area from the "regular" sample.

The provincial minimum and maximum sizes for subunits of types 1 and 2 were determined as the product of the inverse sampling ratio for the province, the number of groups to be formed and respectively the minimum and maximum allowable density factors for the corresponding group type. The range for the density factor (expected take) was established as 4 to 6 in built-up groups and 2 to 3.5 in fringe groups where it was lower to account for higher travel cost and the possibility of high growth in these areas. For any combination of fringe and built-up groups the approximate sizes of the fringe and built-up components of the subunit were separately determined as explained above for types 1 and 2 subunits. The final determination of the number of fringe and built-up groups was made by the computer program for cluster selection on the basis of which combination of groups gave the most suitable density factors. As an example, a map showing subunits for the city of Regina is given in Appendix 10.1.

5.2 Delineation of Clusters

The delineation of clusters was in accordance with the following criteria:

UNITÉS AUTOREPRÉSENTATIVES

5.1 Description globale

Pour les U.A.R., on a adopté un plan d'échantillonnage stratifié à deux degrés. Les strates des U.A.R. (voir le paragraphe 4.2) sont appelées des sous-unités (ou encore des U.P.É. pour les besoins de l'administration), et elles sont subdivisées en grappes (qui représentent normalement des îlots urbains). La grappe constitue le premier degré de sélection dans les sous-unités, le deuxième et dernier degré étant la sélection des logements.

Les sous-unités sont classées, selon leurs possibilités de croissance future, dans l'un des trois types suivants: type 1, zone bâtie, type 2, banlieue, type 3, mixte (zone bâtie et banlieue).

Dans chaque sous-unité, on forme au hasard six (ou dans certains cas 12) "groupes" de grappes, chaque groupe étant composé soit de zones bâties (c'est-à-dire d'un groupement aléatoire de grappes de zones bâties, soit d'un groupe de grappes de banlieues. Dans le cas des sous-unités du type 3, on a fait en sorte que les groupes composés de zones bâties et de banlieues soient en proportion égale pour faciliter l'estimation de la variance. En outre, dans 17 grandes villes, on a fait un échantillon d'appartements spécial, composé d'immeubles d'appartements d'au moins cinq étages et 30 unités, et on a ainsi retiré cette région critique de l'échantillon "ordinaire".

On détermine les tailles minimales et maximales des sous-unités des types 1 et 2 pour la province en multipliant la fraction de sondage provinciale inverse, le nombre de groupes à former et les facteurs de densité minimaux et maximaux acceptables pour le type de groupe correspondant. L'intervalle de variation du facteur de densité (rendement prévu) se situe entre 4 et 6 pour les groupes de zones bâties et entre 2 et 3.5 pour les groupes de banlieues, où il est plus faible vu les coûts de déplacement plus élevés et la possibilité d'un fort accroissement de la population de ces régions. Pour toutes les sous-unités mixtes, (banlieue et zone bâtie), on a calculé la taille approximative de la partie banlieue et celle de la partie zone bâtie composant la sous-unité, de la même façon que dans le cas des sous-unités des types 1 et 2 précédemment décrites. Pour le calcul final du nombre de groupes de banlieues et de zones bâties, on a utilisé un programme d'ordinateur qui sélectionne les grappes. Vous trouverez à l'annexe 10.1, comme exemple, une carte montrant les sous-unités de la cité de Régina.

5.2 Délimitation des grappes

La délimitation des grappes se fit selon les critères suivants:

(i) The number of dwellings per cluster should be between specified lower and upper limits. These limits were obtained keeping in mind both operational and theoretical considerations. The minimum number of dwellings required for each cluster was three times the density factor, and thus it was 18 dwellings for the built-up clusters and about ten dwellings for the fringe clusters. The maximum number of dwellings per cluster was set at 100 for Ontario and Quebec, and 60 for other provinces in order to minimize the problem of listing maintenance which becomes cumbersome for very large clusters.

If a city block had less than the minimum number of dwellings, it was combined with one or more neighbouring blocks to form a cluster. If the number of dwelling units exceeded the maximum size specified for the cluster, then a breakdown of the counts was obtained using well identifiable physical features (e.g., parks, church, block faces, etc.). In this way the block was divided into two or more clusters. Only in exceptional cases, where no such features could be established, was the cluster allowed to have a larger number of dwellings.

(ii) The whole area of the subunit had to be accounted for. Areas containing no dwellings and apparently no space for residential construction (e.g., parks, industrial plants, etc.) were joined with other contiguous areas to form clusters.

The clusters formed in this way were outlined on the subunit map and were serially numbered. A cluster for Calgary is on page 125. The completed "cluster record forms" showing the number of dwellings in each cluster provided the frame from which the sample was selected.

5.3 Choice of the Selection Method

From each subunit, six (in some cases 12) clusters were selected using the Random Group method of selection [15]. The following considerations led to the choice of the random group method.

(a) **Simplicity in updating**: Since changes in the size measures are inherent in any continuous survey, the adaptability of the selection method to a simple procedure of revision of the selection probabilities was considered an important criterion for the choice of a selection method in the initial design. For many complex selection methods involving two or more selections, the revision of initial selection probabilities might require complex treatments due to their diverse joint inclusion probabilities.

(i) Le nombre de logements compris dans chaque grappe doit être conforme aux limites minimales et maximales fixées. Ces limites tiennent compte des nécessités opérationnelles et théoriques. Le nombre minimal de logements requis pour chaque grappe est égal au triple du facteur de densité et, par conséquent, il est de 18 logements pour les grappes composées de zones bâties et d'environ dix logements pour les grappes composées de banlieues. Le nombre maximal de logements compris dans chaque grappe a été fixé à 100 pour l'Ontario et le Québec, et à 60 pour les autres provinces, afin de minimiser les problèmes soulevés par la tenue des listes, tâche très fastidieuse lorsque les grappes sont très grosses.

Lorsqu'un îlot urbain a un nombre de logements inférieur au nombre de logements requis, il est groupé avec un ou plusieurs ilots voisins pour former une grappe. Lorsque le nombre d'unités de logements est plus grand que la taille maximale prévue pour la grappe, on divise l'îlot en utilisant des éléments facilement reconnaissables (par ex., les parcs, les églises, les côtés d'îlots, etc.). Ainsi l'îlot peu être divisé en une ou plusieurs grappes. On admet que la grappe soit composée d'un nombre de logements supérieur au maximum prévu, uniquement dans quelques cas exceptionnels où cette division est irréalisable.

(ii) Il faut tenir compte de toute la région composant la sous-unité. Les régions qui ne contiennent pas de logements et qui, apparemment, n'ont aucun espace prévu pour la construction de logements (par exemple les parcs, les parcs industriels etc.) ont été réunies à d'autres régions contiguës pour former des grappes.

Les grappes constituées de cette manière sont indiquées sur les cartes des sous-unités et un numéro de série leur est attribué. Vous trouverez un schéma d'une grappe de Calgary à la page 125. Les régistres des grappes indiquant le nombre de logements contenu dans chaque grappe servent de base dans laquelle l'échantillon est choisi.

5.3 Choix de la méthode de sélection

Dans chaque sous-unité, six grappes (parfois 12) sont choisies à l'aide de la méthode des groupes aléatoires [15]. Le choix de cette méthode a été dicté par les considérations suivantes:

a) **Simplicité de la mise à jour**: Étant donné que dans toute enquête permanente la taille des unités de sélection change, la facilité d'adaptation de la méthode de sélection à un procédé simple de révision des probabilités de sélection constitue un critère important, qui influe sur le choix de la méthode de sélection à utiliser dans le plan de sondage initial. Pour beaucoup de méthodes de sélection complexes, la révision des probabilités de sélection initiales peut nécessiter des traitements complexes en raison des probabilités composées d'inclusion.

The random group method, like the method of selecting one cluster per subunit, could be easily adapted to simple revision of selection probabilities. A detailed discussion on this is given in a paper by Platek and Singh [11].

(b) **Number of sample clusters**: There are six rotation groups in the survey and each selected cluster is assigned a rotation number (see Chapter 8). It is necessary to select six (or multiple of six) clusters from each subunit in order that every subunit is represented in each rotation group. Thus, the rotation pattern (discussed later) imposes the restriction on the choice of any selection method. The method selected should permit selection of six (or multiple of six) clusters from each subunit. It may be noted that most of the unequal probability schemes suggested in the literature consider the selection of two units (clusters) only, and therefore these were not considered suitable for application in the SRUs of the LFS. One might make the subunits smaller so that two clusters or only one cluster would be selected from each of the smaller subunits. In that case, however, it would not be possible to have representation for each rotation group from every such subunit. This was one of the main reasons for not adopting the method of one cluster per subunit, although this method does possess the similar advantage in updating as the random group method.

(c) **Sample expansion**: As mentioned in section 3.2, the flexibility of the initial selection method for increasing the sample size to a desired level was considered as another important criterion for the choice. The random group method was found to possess this criterion, as described separately in section 5.6.

With regard to the efficiency of the estimates and the variance estimates, the random group method compares well with any other unequal probability without replacement scheme and is always more efficient than with replacement scheme when the random groups contain an equal number of clusters, which was the case in the application of this method in the SRUs. Two special features of this method exploited in the present design are that (i) after the formation of the random groups, sampling within group is done independently, and (ii) each random group by itself is a sub-sample of clusters from the subunit.

The method involves dividing the N clusters of a subunit into K groups at random, each with N/K clusters, and selecting one cluster with probability proportional to its "size" from each group. The number of groups (K) is taken as six (or multiple of six in some cases) since there are six rotation numbers, to ensure that

La méthode des groupes aléatoires, comme la méthode qui consiste à choisir une grappe dans chaque sous-unité, peut facilement être adaptée à la simple révision des probabilités de sélection. Un document signé par Platek et Singh en donne une analyse détaillée [11].

b) **Nombre de grappes d'échantillon**: Il y a six groupes de renouvellement dans l'enquête et chaque grappe choisie reçoit un **numéro de renouvellement** (voir le chapitre 8). Il faut choisir six grappes (ou un multiple de six) dans chaque sous-unité pour que toutes les sous-unités soient représentées dans chaque groupe de renouvellement. Le mode de renouvellement (analysé ultérieurement) impose donc certaines restrictions à toute méthode de sélection choisie. La méthode choisie doit permettre de sélectionner six grappes (ou un multiple de six) dans chaque sous-unité. Il est à noter que la plupart des méthodes connues de sélection avec probabilités inégales tiennent compte de la sélection de deux unités (grappes) uniquement et, par conséquent, elles n'ont pas été jugées applicables aux U.A.R. de l'E.P.A. On peut former des sous-unités plus petites pour permettre la sélection de deux grappes ou d'une seule grappe dans chacune de ces petites sous-unités. Cependant, dans pareil cas, il ne serait pas possible de représenter chaque groupe de renouvellement dans chaque sous-unité. C'est l'une des principales raisons pour lesquelles on n'a pas adopté la méthode qui consiste à choisir une grappe dans chaque sous-unité, même si cette méthode présente les mêmes avantages que la méthode des groupes aléatoires pour la mise à jour.

c) **Élargissement de l'échantillon**: Comme il a été mentionné au paragraphe 3.2, la flexibilité de la méthode de sélection utilisée pour élargir l'échantillon jusqu'au niveau désiré constitue un autre critère important pour le choix de la méthode. La méthode des groupes aléatoires possède cette caractéristique, qui est décrite au paragraphe 5.6.

En ce qui concerne l'efficacité des estimations et des estimations de la variance, la méthode des groupes aléatoires se compare bien à toute autre méthode de tirage sans remise avec probabilités inégales et est toujours plus efficace que les méthode avec remise, lorsque les groupes aléatoires contiennent un nombre égal de grappes, ce qui est le cas lorsqu'on applique cette méthode aux U.A.R. Cette méthode, employée dans le plan de sondage actuel, présente deux caractéristiques particulières: (i) après la formation des groupes aléatoires, l'échantillonnage à l'intérieur des groupes se fait de façon autonome, et (ii) chaque groupe aléatoire est en lui-même un sous-échantillon des grappes formées dans la sous-unité.

Cette méthode consiste à diviser les N grappes d'une sous-unité entre K groupes au hasard, chacun ayant N/K grappes, et à choisir dans chaque groupe une grappe avec probabilité proportionnelle à sa "taille". Étant donné qu'il y a six numéros de renouvellement, on considère que le nombre de groupes (K) est de six (ou, dans certains cas,

each rotation is represented the same number of times in each subunit. This was easy to accomplish in built-up and fringe subunits. In case of combination (built-up and fringe) subunits, the groups in fringe and built-up (say K_1 and K_2 groups respectively) were formed independently and the restriction on the choice of K_1 and K_2 was that they should be even numbers and that the sum $K_1 + K_2 = 6$ (or 12). After having established the number of groups, the selection procedure, as described below, is applied separately to the built-up and fringe clusters within a subunit.

5.4 Selection of Clusters

The input list of N_i clusters of the same type from ith subunit is randomized and K_i groups are formed by dividing the randomized list (henceforth, the order is assumed to be that of the randomized list), for instance if:

$$N_i^* = [N_i/K_i]$$

and $$m_i = N_i - N_i^* K$$

then m_i groups had $N_i^* + 1$ clusters and the remaining groups had N_i^* clusters.

The expected number of households is then computed for each group as $d_{ij} = H_{ij}/R$, where H_{ij} is the number of households in group j of ith subunit and R is the provincial inverse sampling ratio.

The cluster inverse sampling ratio R_{ijk} was then obtained for all clusters as $R_{ijk} = [H_{ijk}/d_{ij} + .5]$ where H_{ijk} is the number of households in cluster k. The ratios were rounded, if necessary, so that $\sum_k R_{ijk} = R$ to ensure unbiasedness. These R_{ijk}'s were treated as the size of the cluster and one cluster is selected from each group with probability proportional to its size.

Each group is assigned a rotation number following a procedure explained in section 8.4, with the objective of distributing the sample evenly with regards to expected take, so that the sample among the rotation is as balanced as possible.

The selection of clusters is an automated procedure [1], for which the input consists of the number of built-up and fringe clusters, the number of dwellings in each cluster, and in addition, information required for assigning rotation numbers. The output consists of printouts showing the sampling frame and indicating which clusters are selected. Also, the program creates rotation schedules which are used in rotation of the sample (Chapter 8).

un multiple de six) pour que tous les renouvellements puissent être représentés le même nombre de fois dans chaque sous-unité. Cela s'est fait aisément dans les sous-unités composées uniquement de zones bâties ou de banlieues. Dans le cas des sous-unités mixtes (composées de zones bâties et de banlieues), les groupes compris dans la banlieue et dans la zone bâtie (soit K_1 et K_2 respectivement) ont été formés indépendamment et la restriction imposée au choix de K_1 et K_2 était qu'ils soient des nombres pairs et que la somme $K_1 + K_2 = 6$ (ou 12). Une fois le nombre de groupes établi, la méthode de sélection décrite ci-après fut appliquée séparément aux grappes composées de zones bâties et de banlieues dans une sous-unité.

5.4 Sélection des grappes

On dresse la liste aléatoire des N_i grappes de même type provenant de l'i^e sous-unité et on forme K_i groupes en divisant cette liste (désormais, on considère que l'ordre est celui de la liste aléatoire); par exemple, si:

$$N_i^* = [N_i/K_i]$$

et que $$m_i = N_i - N_i^* K$$

alors m_i groupes ont $N_i^* + 1$ grappes et les autres groupes ont N_i^* grappes.

On calcule alors le nombre de ménages prévu dans chaque groupe comme étant $d_{ij} = H_{ij}/R$, où H_{ij} représente le nombre de logements compris dans le groupe J de l'i^e sous-unité et R la fraction de sondage inverse provinciale.

Ensuite, on obtient la fraction de sondage inverse de la grappe en calculant $R_{ijk} = [H_{ijk}/d_{ij} + .5]$, où H_{ijk} représente le nombre de ménages compris dans la grappe k. Au besoin, les rapport sont arrondis de sorte que $\sum_k R_{ijk} = R$, évitant ainsi les biais possibles. On considère ces R_{ijk} comme représentant la taille de la grappe et on choisit une grappe dans chaque groupe avec probabilité proportionnelle à sa taille.

On donne un numéro de renouvellement à chaque groupe en suivant la méthode expliquée au paragraphe 8.4, avec objectif de répartir l'échantillon également quant au rendement prévu.

La sélection des grappes est automatisée[1], on donne le nombre de grappes formées de zones bâties et de banlieues, le nombre de logements compris dans chaque grappe et, en outre, les renseignements requis pour l'attribution des numéros de renouvellement; on obtient des relevés indiquant la base d'échantillonnage et les grappes qui ont été choisies. Ce programme génère également les calendriers de renouvellement utilisés pour le renouvellement de l'échantillon (voir chapitre 8).

5.5 Selection of Households

A random start r_k^* is determined for each selected cluster by choosing a random integer so that $1 \leqslant r_k^* \leqslant R_{ijk}$. For each cluster a Form 01 (page 125) is prepared which includes a diagram showing cluster identifications, cluster inverse sub-sampling ratio and the random start. These forms are sent to the Regional Office (RO) where all the dwellings are listed on FO2 along with their identification particulars. This list then becomes the input to the Sample Listing System, and a sample of dwellings is selected systematically, applying the random start and the inverse sub-sampling ratio for the selected cluster. Each of the selected dwellings is visited by the LFS interviewer during a specified week and all eligible persons in the household occupying the dwellings are interviewed.

5.6 Sample Expansion

The primary objective for increasing the sample size from approximately 35,000 households selected initially to 55,000 households is to enable the provision of more detailed and more reliable estimates, especially for provincial data. Having determined the allocation of the additional sample (20,000 households) for various provinces to meet the above objective, the allocation between SR and NSR areas of a province was carried out in an optimum manner, see Chapter 14. As a result, the increase in sample from SRUs varies greatly from province to province. For example, there is no increase in SRU sample for Ontario and Quebec; while for Saskatchewan and Prince Edward Island the sample is increased to three times the initial sample.

The random group method used for selection of the initial sample from SRUs allowed selection of one cluster from each random group. While increasing the sample, it was necessary, because of field requirements and cost implications, to retain the initially selected clusters; furthermore it was necessary to devise a selection mechanism which would allow for the selection of the desired number of additional clusters and at the same time would retain basic feature of the random group method as outlined earlier in section 5.3.

The selection of the additional sample was thus achieved as follows: each existing random group was sub-divided into the desired number of sub-groups (i.e., 2 or 3 sub-groups depending on whether the sample in a particular subunit is to be doubled or tripled). As much as possible the sub-groups formed were of equal sizes so that the difference in the expected take among the sub-groups was minimized. The cluster being already surveyed as part of the initial sample was retained to represent the sub-group to which it belonged and one cluster was selected independently from each of the

5.5 Sélection des ménages

On détermine une origine aléatoire r_k^* pour chaque grappe sélectionnée en choisissant un nombre entier aléatoire tel que $1 \leqslant r_k^* \leqslant R_{ijk}$. Pour chaque grappe, on remplit une formule 01 (page 125) qui comprend un schéma décrivant la grappe, la fraction de sous-sondage inverse de la grappe et l'origine aléatoire. Ces formules sont envoyées au bureau régional (B.R.) où tous les logements sont inscrits sur la formule F02 avec leurs particularités. On introduit ensuite cette liste dans le système de listage de l'échantillon, puis on sélectionne systématiquement un échantillon de logements avec l'origine aléatoire et la fraction de sous-sondage inverse de la grappe choisie. Au cours d'une semaine donnée, l'interviewer de l'E.P.A. visite tous les logements sélectionnés et interroge toutes les personnes admissibles du ménage occupant ces logements.

5.6 Élargissement de l'échantillon

L'objectif principal de l'élargissement de l'échantillon, qui passe d'environ 35,000 ménages initialement sélectionnés à 55,000 ménages, est l'obtention d'estimations plus détaillées et plus fiables, en particulier pour les données provinciales. Après avoir déterminé la répartition de l'échantillon additionnel (20,000 ménages) entre les diverses provinces, en vue de réaliser l'objectif précédemment mentionné, on a procédé à une répartition optimale entre les U.A.R. et les U.N.A.R. de chaque province, voir le chapitre 14. Par conséquent, l'accroissement de l'échantillon provenant des U.A.R. varie considérablement d'une province à l'autre. Par exemple, l'échantillon des U.A.R. n'a pas augmenté en Ontario et au Québec, tandis qu'en Saskatchewan et à l'Île-du-Prince-Édouard, la taille de l'échantillon initial a triplé.

La méthode des groupes aléatoires utilisée pour la sélection a augmenté l'échantillon intial tiré des U.A.R. permet la sélection d'une grappe dans chaque groupe aléatoire. Lorsqu'on a augmenté l'échantillon, il a fallu en raison des exigences des opérations sur le terrain et des coûts, conserver les grappes initialement choisies; en outre, il était nécessaire que le mécanisme de sélection permit de choisir le nombre voulu de grappes additionnelles tout en conservant les caractéristiques essentielles de la méthode des groupes aléatoires, mentionnées au paragraphe 5.3 précédent.

Par conséquent, la sélection de l'échantillon additionnel se fit de la manière suivante: chaque groupe aléatoire fut subdivisé en un nombre désiré de sous-groupes (c'est-à-dire 2 ou 3 sous-groupes, selon que l'échantillon d'une sous-unité donnée doit être doublé ou triplé). Autant que possible on a formé des sous-groupes de taille égale afin de minimiser la différence de rendement prévu entre les sous-groupes. La grappe qui faisait partie de l'échantillon initial, et qui a donc déjà été echantillonnée fut retenue pour reprérenter le sous-groupe auquel elle appartenait; on choisit une grappe de façon indépendente

remaining sub-groups with probabilities proportional to the sizes based only on the clusters belonging to these sub-groups.

dans chacun dans des sous-groupes restants, avec probabilité proportionnelle à leur taille, uniquement en fonction des grappes qui font partie de ces sous-groupes.

CHAPTER – 6 – CHAPITRE

NON-SELF-REPRESENTING UNITS

6.1 Introduction

NSRUs are the areas lying outside the SR areas and are comprised of rural areas and small urban centers. In the NSR areas, the sample of households was selected in each ER following a stratified multi-stage sampling design. The first stage unit is called the PSU. The size of strata and PSUs differs from province to province, but the main principles of their formation were the same, as described in Chapter 4. In each urban part (center) of selected PSUs a two-stage sample was drawn while in the rural parts of the sampled PSUs, due to large distances involved in travelling, the households were selected in three stages.

Details of the sub-sampling in the urban and rural parts of the sampled PSUs are described in sections 6.3 and 6.4. In section 6.2 the selection procedure for PSUs is discussed.

6.2 Selection of Primary Sampling Units

6.2.1 Choice of Selection Procedure

Given the possibility of sample expansion, as mentioned in Chapter 3, the main criterion for the choice of a selection procedure at the design stage was that the adopted procedure should be flexible enough to allow expansion of the sample. Fellegi's [2] method used in the earlier design, although quite efficient, becomes very complicated when the number of PSUs to be selected is large (say, larger than three). This method, therefore, limits the flexibility of the design particularly with respect to increasing the sample in terms of number of PSUs. Theoretical and empirical studies conducted showed that Randomized PPS (Probability Proportional to Size) systematic sampling method [6] meets the necessary requirements. In other words, this method allows for the expansion of the sample from two originally selected PSUs to any number of additional PSUs, while maintaining the selection of the two originally sampled PSUs. The method of sample expansion is briefly described in section 6.6.

6.2.2 Selection of PSUs

After having obtained the average size of PSUs within a stratum (which is the same for all strata within a province), PSUs were delineated following the criteria given in section 4.5. Every effort was made to delineate PSUs in such a way that they represent the stratum with respect to important LF characteristics and with respect to the proportion of urban and rural popula-

UNITÉS NON AUTOREPRÉSENTATIVES

6.1 Introduction

Les U.N.A.R. sont des secteurs situés à l'extérieur des secteurs A.R. et sont formées de régions rurales et de petits centres urbains. Dans les secteurs N.A.R., on a formé des échantillons de ménages dans chaque R.É. suivant un plan de sondage à plusieurs degrés. L'unité obtenue à la première étape s'appelle U.P.É. La taille des strates et des U.P.É. varie d'une province à l'autre, mais les principes généraux qui déterminent leur formation, et qui sont décrits au chapitre 4, demeurent les mêmes. Dans chaque partie urbaine (centre) des U.P.É. choisies, on a procédé a un sous-sondage à deux degrés tandis que dans les régions rurales des U.P.É. échantillonnés, en raison des distances nécessitant de longs déplacements, les ménages ont été échantillonnés en trois étapes.

Les paragraphes 6.3 et 6.4 contiennent une description détaillée du sous-échantillonnage des parties rurales et urbaines des U.P.É., tandis que le paragraphe 6.2 présente le processus de sélection des U.P.É.

6.2 Sélection des unités primaires d'échantillonnage

6.2.1 Choix de la méthode de sélection

Étant donné la possibilité de devoir élargir l'échantillon, comme il est expliqué au chapitre 3, le critère de base pour le choix d'une méthode de sélection, à l'étape de la conception de l'échantillon, a été la flexibilité de cette méthode, permettant l'élargissement de l'échantillon. La méthode Fellegi [2], utilisée jusqu'alors avec efficacité, se révèle d'emploi compliqué lorsqu'il s'agit de choisir un nombre élevé d'U.P.É. (plus de trois U.P.É., par exemple). En conséquence, cette méthode limite le flexibilité du plan de sondage, en particulier en ce qui a trait à l'élargissement de l'échantillon par l'augmentation du nombre d'U.P.É. Des études empiriques et théoriques ont démontré que la méthode d'échantillonnage systématique avec classement aléatoire et probabilité proportionnelle à la taille (E.S.C.A.P.P.T.) [6] satisfait aux exigences. En d'autres termes, cette méthode permet d'élargir l'échantillon à partir des deux U.P.É. originellement choisies en leur ajoutant le nombre voulu d'U.P.É. Le paragraphe 6.6 décrit brièvement la méthode d'élargissement de l'échantillon.

6.2.2 Sélection des U.P.É.

Après avoir déterminé la taille moyenne des U.P.É. dans une strate (qui est la même pour toutes les strates d'une même province), on a délimité les U.P.É. conformément aux critères exposés au paragraphe 4.5. On s'est efforcé de délimiter les U.P.É. de façon à ce qu'elles représentent la strate au point de vue des caractéristiques importantes de la population active et des proportions de la

tion. The number of PSUs within a stratum is generally (with minor exception) between 10 and 19. (The actual number by stratum and ER is given in the appendix). From each stratum two PSUs were originally selected using randomized systematic PPS sampling. The following steps are involved in the selection of PSUs:

(a) within each stratum, the list of PSUs was randomized,

(b) the inverse sampling ratio (R_i) for PSU i (i = 1, 2, ..., N) was obtained as:

$$R_i = [(n\ H_i/H)\ R + .5] \qquad (6.1)$$

where n(=2) is the number of PSUs to be selected, H_i is the population for PSU_i, $H = \sum_1^n H_i$ and R is the provincial inverse sampling ratio.

Further, to ensure unbiasedness the individual R_i's were re-rounded off, if necessary, so that the sum of all PSUs within the stratum was equal to n times the provincial inverse sampling ratio.

(c) Using R_i's as the size measures, R as the interval and by choosing random integer r ($1 \leqslant r \leqslant R$), a PPS systematic sample of two PSUs was selected from each stratum.

The sampled PSUs were sub-stratified into urban and rural parts, each part comprising exclusively urban or rural groups. In the rural part of a sampled PSU, the subsampling was carried out in three stages: group, cluster and household; while in the urban part, groups (urban centers) are selected with certainty resulting in a two stage sample. The subsampling ratios for each part are obtained by using the PSU sampling ratio. The details of the delineation of rural and urban groups and the subsampling procedure adopted, are described in the following sections.

6.3 Sub-sampling in Urban Areas

6.3.1 Determination of Lower Limit for Urban Centers

The 1971 Census of Canada identified certain places as "urban" but for the purpose of the LFS the definition of "urban centers" is somewhat different. It is intended to meet certain design constraints in the field and in variance estimation procedures. Following the rules described below the lower population limits were established for the identification of urban centers. These limits were mainly dictated by the field requirement that an interviewer has at least one day's assignment (about six sample households) in an urban center.

population rurale et urbaine. Le nombre d'U.P.É. par strate se situe généralement, à quelques exceptions près, entre 10 et 19 (le nombre exact par strate et par R.É., est donné en annexe). On a tout d'abord sélectionné deux U.P.É. par strate à l'aide de la méthode d'échantillonnage systématique avec classement aléatoire et probabilité proportionnelle à la taille. La sélection des U.P.É. se fit comme suit:

a) pour chaque strate, on a établi une liste aléatoire des U.P.É..

b) la fraction de sondage inverse (R_i) pour l'U.P.É. i (i = 1, 2, ..., N) s'obtenait par la formule suivante:

$$R_i = [(n\ H_i/H)\ R + .5] \qquad (6.1)$$

où n(=2) représente le nombre d'U.P.É. à sélectionner, H_i la population de l'UPE_i, $H = \sum_1^n H_i$ et R est la fraction de sondage inverse provinciale.

De plus, pour éviter tout biais, chaque R_i a été arrondi à nouveau chaque fois que cela était nécessaire, de sorte que la somme relative à toutes les U.P.É. de chaques strate soit égale à n fois la fraction de sondage provinciale inverse.

c) En prenant les R_i comme mesure de la taille, R comme intervalle et en choisissant r comme nombre entier aléatoire ($1 \leqslant r \leqslant R$), on a choisi dans chaque strate un échantillon systématique de deux U.P.É.

Chaque U.P.É. a été sous-stratifiée en une partie urbaine et une partie rurale, chaque partie étant formée exclusivement de groupes urbains ou ruraux. Dans la partie rurale de l'U.P.É. échantillonnée, le sous-échantillonnage s'est effectué en trois étapes: groupe, grappe et ménage, tandis que dans la partie urbaine, les groupes (centres urbains) ont tous été choisis, ce qui a donné un échantillon à deux degrés. La fraction de sous-échantillonnage pour chaque partie s'obtient à l'aide de la fraction de sondage de l'U.P.É. La délimitation des groupes ruraux et urbains et la méthode adoptée pour le sous-échantillonnage sont décrits en détail dans les paragraphes qui suivent.

6.3 Sous-échantillonnage dans les régions urbaines

6.3.1 Détermination de la limite inférieure dans les centres urbains

Le recensement du Canada de 1971 avait désigné "urbaines" certaines agglomérations, mais pour les besoins de l'E.P.A., l'expression "centres urbains" a une définition quelque peu différente. Cette définition a pour but de répondre à certaines contraintes d'ordre pratique au point de vue conception d'enquête et méthodes d'estimation de la variance. D'après les règles décrites ci-après, les limites minimales de population furent arrêtées aux fins d'identification des centres urbains. Ces limites ont surtout été établies en fonction des exigences des opérations sur le

This implied that the minimum number of households (H_j) in an urban center should be 6 R_j, where R_j is the inverse sampling ratio for PSU_j, as defined in (6.1).

The lower limit in terms of persons was obtained by considering four persons as the average household size. Thus the minimum size (population) of urban places to qualify as "urban centers" for the LFS was obtained as 24 R_j. For this purpose, it was sufficient to obtain R_j as the product of R and $(2/\bar{N})$, where $\bar{N}$ is the average number of PSU per stratum in a given province. The limits vary from province to province depending on R and N, for example, this limit in the Maritimes was 250 persons, 400 persons in British Columbia and 650 persons in Ontario and Quebec.

In general, these limits were followed in classifying towns and villages as urban centers, however in some instances, smaller places were also considered as "urban centers" particularly when the characteristics of the place were similar to other large centers.

The urban part of a PSU is composed of urban centers totally belonging to that PSU and/or fractional parts of relatively larger urban centers. That is, some larger urban centers are shared amongst two or more PSUs belonging to the same stratum. Each time an urban center appears in a selected PSU it is called a group for that PSU. Since each urban center is selected with certainty, in order to reduce travel costs, where possible the number of urban centers per PSU was limited to two.

6.3.2 Cluster Delineation

As mentioned above, urban centers (groups) were selected with certainty. Thus for all urban centers associated with the selected PSUs, maps were prepared and forwarded to the appropriate RO where a physical count of all the dwellings within the group was obtained. These counts were entered on the maps in as detailed a manner as possible and in such a way as to respect all permanent physical features such as roads, rivers, railroad tracks, etc.

When these maps were returned, each urban center was subdivided into a number of clusters, a cluster being a well-defined area with boundaries recognizable both on the maps and in the field. Clusters contained a minimum of two to three dwellings and where possible, a maximum equal to three times the inverse sampling ratio to be applied to the cluster list.

6.3.3 Selection of Clusters

Clusters were selected independently from each urban center belonging to the sampled PSUs following randomized PPS systematic sampling. If we denote the inverse of the proportion of the jth urban center be-

terrain, la tâche de l'interviewer par centre urbain devant équivaloir à au moins une journée d'enquête (environ six ménages échantillonnés), ce qui suppose que le nombre minimal de ménages (H_j) par centre urbain devrait être de 6 R_j, où R_j est la fraction de sondage inverse de l'U.P.É.$_j$, comme il est défini au paragraphe 6.1.

On a estimé à environ quatre personnes la taille moyenne des ménages, ce qui élève à 24 R_j la taille minimale (au point de vue population) des agglomérations urbaines pour se qualifier comme "centres urbains" aux fins de l'E.P.A. À cet effet, il suffisait d'obtenir R_j comme le produit de R et $(2/\bar{N})$, où $\bar{N}$ est le nombre moyen d'U.P.É. par strate dans une province donnée. Ces limites varient d'une province à l'autre suivant la valeur de R et de $\bar{N}$. Par exemple, la limite était de 250 personnes pour les Maritimes, 400 personnes pour la Colombie-Britannique et 650 personnes pour le Québec et l'Ontario.

De façon générale, on a respecté ces limites pour la classification des villes et villages en centres urbains, sauf dans certains cas où de petites agglomérations ont été considérées "centres urbains" du fait qu'elles possédaient certaines caractéristiques des grands centres.

La partie urbaine d'une U.P.É. est composée uniquement de centres urbains appartenant à cette U.P.É. et(ou) de portions de centres urbains relativement plus importants. Ainsi, certains grands centres urbains sont partagés entre deux ou plusieurs U.P.É. appartenant à la même strate. Chaque centre urbain d'une U.P.É. choisie est appelé groupe de cette U.P.É. Étant donné que chaque centre urbain est sélectionné avec certitude, on a limité leur nombre à deux par U.P.É. en vue de réduire, dans la mesure du possible, les frais de déplacement du personnel.

6.3.2 Délimitation des grappes

Comme il a déjà été mentionné plus haut, la sélection des centres urbains (groupes) s'est faite avec certitude. On a alors pour tous les centres urbains relatifs aux U.P.É. choisies, préparé les cartes nécessaires qui ont ensuite été envoyées aux B.R. pour qu'ils dénombrent les logements par groupe. Ils ont reporté les chiffres obtenus sur les cartes avec tous les détails possibles, en respectant les éléments cartographiques permanents tels que: routes, rivières, voies de chemin de fer, etc.

Une fois ces cartes remplies, on a divisé chaque centre urbain en grappes (une grappe étant un secteur bien défini dont les limites sont reconnaissables à la fois sur les cartes et sur le terrain). Chaque grappe contient un minimum de deux ou trois logements et, lorsque c'est possible, un maximum égal à trois fois la fraction de sondage inverse à appliquer à la liste des grappes.

6.3.3 Sélection des grappes

Les grappes ont été sélectionnées indépendamment dans chaque centre urbain appartenant aux U.P.É. échantillonnées, suivant la méthode d'échantillonnage systématique avec classement aléatoire et P.P.T. Si nous posons

longing to PSU i as T_{ij}, (T_{ij} = 1 if the urban center belongs only to PSU i) then the inverse sampling ratio within urban center i is given by $T_{ij}R_i$ where R_i is the PSU inverse sampling ratio. This means the entire urban center will be clustered and a sample of clusters will be drawn from the whole urban center with an inverse sampling ratio of $T_{ij}R_i$. The following steps are involved in the selection of clusters within the group:

(a) Within each selected urban center, the list of clusters was randomized.

(b) The desired expected take per cluster (or density factor) was usually taken as 3. But in the groups where there was a high frequency of clusters with only 2 households, it was taken as 2 or 2.5. The number of clusters (n_{ij}) to be selected from the urban center was then determined so that the density factor:

$$d_{ij} = \frac{H_{ij}}{n_{ij}\, T_{ij}R_i} \qquad (6.2)$$

was as close as possible to 3 (or 2 or 2.5).

(c) All clusters for which the number of households (H_{ijk}) was larger than or equal to the ratio $\sum_k H_{ijk}/n_{ij}$ were selected with certainty and they were assigned an inverse sampling ratio equal to the inverse sampling ratio of the group (i.e., $R_{ijk} = T_{ij}R_i$).

(d) The inverse sampling ratios (R_{ijk}) for the remaining clusters were obtained as:

$$R_{ijk} = [\frac{H_{ijk}}{d_{ij}} + .5] \qquad (6.3)$$

where d_{ij} was similarly defined for the remaining set of clusters only. These individual R_{ijk} were re-rounded if necessary to ensure unbiasedness.

(e) Using R_{ijk} as the cluster size, a systematic sample of clusters was selected in the usual manner.

Shared urban centers – In the event that the second selected PSU (= i_2) also shared the same urban center with inverse of the proportion denoted by T_{i_2j} then a second sample of clusters was drawn from the entire urban center in such a manner that samples were as far apart as possible. Since the two inverse sampling ratios $T_{i_1j}R_i$ and $T_{i_2j}R_{i_2}$ were different, independent application of the above procedure of cluster selection to the same list of clusters (i.e., in the same urban center) would result in a different number of clusters in the sample and also overlap of the samples. For practical reasons this was not desirable. Therefore, in order to

que T_{ij} est l'inverse de la proportion du j^e centre urbain appartenant à l'U.P.É.$_i$ (T_{ij} = 1 lorsque le centre urbain n'appartient qu'à l'U.P.É.$_i$), la fraction de sondage inverse de ce centre urbain sera donc $T_{ij}R_i$, où R_i est la fraction de sondage inverse de l'U.P.É. Dans ce cas, le centre urbain sera entièrement divisé en grappes et un échantillon de grappes sera prélevé du centre urbain complet dont la fraction de sondage inverse sera $T_{ij}R_i$. La sélection des grappes à l'intérieur d'un groupe se fait en procédant comme suit:

a) Pour chaque centre urbain choisi, la liste des grappes a été établie au hasard.

b) On a estimé en général à 3 le rendement prévu et souhaitable par grappe (ou facteur de densité), mais dans les groupes présentant une haute fréquence de grappes à deux ménages, ce rendement a été fixé à 2 ou 2.5. Le nombre de grappes à choisir (n_{ij}) dans le centre urbain a été déterminé de telle façon que le facteur de densité:

$$d_{ij} = \frac{H_{ij}}{n_{ij}\, T_{ij}R_i} \qquad (6.2)$$

se rapprochait le plus possible de 3 (ou de 2 ou de 2.5).

c) Toutes les grappes ayant un nombre de ménages (H_{ijk}) plus grand ou égal à la fraction $\sum_k H_{ijk}/n_{ij}$ ont été choisies avec certitude et la fraction de sondage inverse qui leur a été assignée était égale à la fraction de sondage inverse du groupe (soit: $R_{ijk} = T_{ij}R_i$).

d) Les fractions de sondage inverses (R_{ijk}) des autres grappes ont été calculées à l'aide de la formule suivante:

$$R_{ijk} = [\frac{H_{ijk}}{d_{ij}} + .5] \qquad (6.3)$$

où d_{ij} a été défini comme d_{ij} mais que pour les grappes restantes. Chaque R_{ijk} a été arrondi à nouveau, au besoin, pour éviter tout biais.

e) À l'aide de R_{ijk} comme taille de la grappe, on a choisi de la façon habituelle un échantillon systématique de grappes.

Centres urbains partagés – Dans le cas où la seconde U.P.É. choisie (= i_2) partageait également le même centre urbain avec l'inverse de la proportion donnée par T_{i_2j}, on a alors prélevé un second échantillon de grappes dans tout le centre urbain de telle façon que les deux échantillons fussent le plus éloignés possible l'un de l'autre. Étant donné que les deux fractions de sondage inverses $T_{i_1j}R_{i_1}$ et $T_{i_2j}R_{i_2}$ étaient différentes, la sélection distincte de grappes selon la méthode précitée à partir de la même liste de grappes (c'est-à-dire dans le même centre urbain) pourrait résulter un nombre différent de grappes dans l'échantillon ainsi qu'en un chevauchement des échantillons. Pour

ensure $n_{i1j} = n_{i2j} = n.j$, i.e., the same number of clusters were selected from the shared urban centers for the two PSUs, $n.j$ was determined to best satisfy the constraints on the density factors, considering both PSUs together.

Further, since the cluster inverse sampling ratios depended on $T_{i1j}R_{i1}$ and $T_{i2j}R_{i2}$, there were two different sets of cluster sizes. The random start for the first PSU was chosen in the usual manner, that is, the random integer r_{i1j} for the first PSU was $1 \leqslant r_{i1j} \leqslant T_{i1j}R_{i1}$, while the random integer for the second PSU r_{i2j} was spaced at a maximum distance from r_{i1j}.

With the two samples spaced a maximum distance apart on the randomized list, and the constraints on allowable cluster sizes, the same cluster will not be selected for both PSUs.

Remark – The new subsampling procedures as distinct from the old design are: selection of all urban centers with certainty which reduced within PSU component of the variation, randomization of cluster list to enable unbiased variance estimation, and more complex treatment of shared urban centers to avoid repetition of the same clusters and households in the sample and thus reducing the respondent burden.

6.4 Sub-sampling in Rural Areas

6.4.1 Selection of Rural Groups

The rural part of a PSU was composed of a collection of contiguous EAs. Individual EAs for the most part made up groups, except in cases where the population of an EA was either too small or too large. EAs too small were combined with adjacent EAs to form a group in such a way that the combined population fell within the allowable range. The population of larger EAs was equally divided to form two or three groups, each of which shared the common area as in the case of urban centers shared between PSUs.

The selection of rural groups was according to randomized PPS systematic method. The first step was to randomly order the rural groups for the selected PSU i. A density factor was determined by the relation

$$d_i = H_i^*/(n\,R_i) \qquad (6.4)$$

where n representing the number of groups to be selected, was the smallest integer $\geqslant 2$ for which $d_i \leqslant 20$, H_i^* was the total rural households in the PSU and R_i was the PSU inverse sampling ratio in (6.1).

des raisons d'ordre pratique, ce résultat n'était pas souhaitable. En conséquence, afin d'assurer que $n_{i1j} = n_{i2}$, c'est-à-dire que le même nombre de grappes soit sélectionné dans les centres urbains partagés pour les deux U.P.É., $n.j$ a été déterminé de façon à répondre au mieux aux exigences imposées aux facteurs de densité, en tenant compte des deux U.P.É. à la fois.

De plus, étant donné que les fractions de sondage inverses des grappes étaient dépendantes de $T_{i1j}R_{i1}$ et de $T_{i2j}R_{i2}$, on a obtenu deux séries de taille de grappes. L'origine choisie au hasard pour la première U.P.É. a été déterminée par la méthode habituelle, c'est-à-dire le nombre entier aléatoire r_{i1j} pour la première U.P.É. étant $1 \leqslant r_{i1j} \leqslant T_{i1j}R_{i1}$, tandis que le nombre entier aléatoire pour la seconde U.P.É., r_{i2j}, a été situé à une distance maximale de r_{i1j}.

Du fait de la distance maximale séparant les deux échantillons sur la liste aléatoire et des exigences imposées à la taille des grappes, il sera impossible de choisir une même grappe pour les deux U.P.É.

Remarque – Les méthodes de sous-échantillonnage qui diffèrent de celles de l'ancien plan de sondage sont les suivantes: sélection automatique de tous les centres urbains, ce qui réduit l'élément de variation à l'intérieur de l'U.P.É., établissement d'une liste aléatoire de grappes; ce qui permet une estimation sans biais de la variance, et traitement plus minuteux des centres urbains partagés afin d'éviter la répétition des mêmes grappes et des mêmes ménages dans l'échantillon et de réduire la charge imposée aux répondants.

6.4 Sous-échantillonnage dans les régions rurales

6.4.1 Sélection des groupes ruraux

La partie rurale d'une U.P.É. se composait de plusieurs S.D. contigus. Chaque S.D. formait en général un groupe, sauf dans le cas où sa population était soit insuffisante, soit trop importante. Un S.D. trop petit était réuni à un S.D. adjacent pour former un groupe de sorte que les deux populations réunies respectent les limites permises. Dans le même ordre d'idée, la population d'un S.D. trop important était divisée en deux ou trois groupes partageant chacun un secteur commun, comme dans le cas de centres urbains partagés entre plus d'une U.P.É.

Le choix des groupes ruraux s'est effectué selon la méthode d'échantillonnage systématique avec classement aléatoire et P.P.T. En premier lieu, on a classé au hasard les groupes ruraux pour l'U.P.É. i choisie. Le facteur de densité a été déterminé par la formule

$$d_i = H_i^*/(n\,R_i) \qquad (6.4)$$

où n, représentant le nombre de groupes à choisir, était le plus petit nombre entier $\geqslant 2$ pour lequel $d_i \leqslant 20$, H_i^* représentait le nombre total de ménages ruraux dans l'U.P.É. et R_i était la fraction de sondage inverse de l'U.P.É. dans la formule (6.1).

The group sizes or inverse sampling ratios, R_{ij} were calculcated, as:

$$R_{ij} = [H^*_{ij}/d_i + .5]. \qquad (6.5)$$

where H^*_{ij} was the number of households in the jth group. If a group was selected with certainty, $R_{ij} = R_i$. R_{ij}'s were re-rounded if necessary so that $\sum_j R_{ij} = n\ R_i$ and groups (if any) sharing equal fractions of the same collection of EAs had equal inverse sampling ratios.

6.4.2 Delineation and Selection of Clusters in Rural Groups

Maps of selected groups were prepared and forwarded to the ROs for field counts. A detailed description of the procedure adopted for field counts, its transcription on maps and subsequent delineation of clusters with well defined boundaries recognizable in the field, is given in the Sampling Manual. A cluster map for NSRU rural area is Appendix 10.5.

The method of cluster selection in rural groups was the same as discussed earlier for the urban groups. The list of clusters was first randomized, n_{ij}, R_{ijk} and d_{ij} were computed and then the sample was selected using PPS systematic method. Implementation of the procedure was however much simpler since for the majority of groups, T_{ij} was unity; that is, there were only a few EAs which were shared between two selected groups. These shared EAs were treated in the same way as the shared urban centers.

6.5 Selection of Households

The procedure followed for the selection of dwellings in NSRU clusters was the same as those for SRU clusters as described in 5.5. For each sample cluster, a Form 01 was prepared and sent to the RO for listing. The lists of dwellings in selected clusters were stored on computer and the random starts were imput from files created by the sample selection programs [1]. Dwellings were selected systematically by applying a random start and sampling interval equal to the cluster inverse sampling ratio for the selected cluster.

6.6 Sample Expansion

The additional sample was allocated to various provinces and then to SR and NSR areas within a province following certain criteria (see Chapter 14.) In order

La taille des groupes, ou fractions de sondage inverses, R_{ij} a été calculée à l'aide de la formule:

$$R_{ij} = [H^*_{ij}/d_i + .5]. \qquad (6.5)$$

où H^*_{ij} représentait le nombre du ménages dans le j^e groupe. Lorsqu'un groupe était choisi automatiquement, R_{ij} était égal à R_i. Au besoin, les R_{ij} ont été arrondis à nouveau pour que $\sum_j R_{ij} = n\ R_i$ et que les groupes (éventuels) partageant les mêmes fractions du même groupe de S.D. aient des fractions de sondage inverses égales.

6.4.2 Délimitation et choix des grappes dans les groupes ruraux

Les cartes des groupes choisis ont été envoyées aux B.R. pour fins d'inscription des chiffres relevés sur le terrain. On trouvera dans le Manuel d'échantillonnage une description détaillée de la méthode de dénombrement sur le terrain, de la transcription sur les cartes et de la délimitation des grappes avec des limites définies repérables sur le terrain. L'annexe 10.5 présente une carte de grappes dans une région rurale d'U.N.A.R.

La méthode de sélection de grappes dans les groupes ruraux était la même que celle utilisée pour les groupes urbains et expliquée plus haut. La liste des grappes a d'abord été établie au hasard, n_{ij}, R_{ijk} et d_{ij} ont été calculés, puis l'échantillon a été choisi à l'aide de la méthode d'échantillonnage systématique avec classement aléatoire et P.P.T. La mise en application de la méthode s'est cependant effectuée de façon beaucoup plus simple étant donné que pour la majorité des groupes, T_{ij} était l'unité il n'y avait que quelques S.D. qui étaient partagés entre deux groupes choisis. On a procédé de la même façon pour ces S.D. partagés que pour les centres urbains partagés.

6.5 Sélection des ménages

La méthode adoptée pour la sélection des ménages dans les grappes d'U.N.A.R. était la même que celle utilisée pour les grappes d'U.A.R., décrite au paragraphe 5.5. Une formule 01 a été établie pour chaque grappe de l'échantillon, puis adressée au B.R. pour qu'il établisse la liste des logements contenus dans la grappe. Ces listes ont été entrées en ordinateur et les origines choisies au hasard ont été introduites à partir des dossiers créés par les programmes de sélection de l'échantillon [1]. La sélection des logements s'est faite systématiquement à l'aide d'une origine choisie au hasard et d'un intervalle d'échantillonnage égal à la fraction de sondage inverse de la grappe choisie.

6.6 Élargissement de l'échantillon

On a assigné l'échantillon supplémentaire à plusieurs provinces, puis aux secteurs A.R. et N.A.R. situés dans une province, selon certains critères

to facilitate the field work and to ensure better geographic spread, the sample expansion in the NSR area was achieved by increasing the number of PUSs in the sample. Depending on the allocation of the additional sample to a province, one to four additional PSUs were selected from each stratum. For example, one additional PSU was selected in Ontario, Quebec, Nova Scotia and New Brunswick while four additional PSUs were selected from each stratum in Prince Edward Island.

The randomized systematic PPS sampling method originally used for selection of two PSUs from each stratum allowed selection of any number of additional PSUs, while retaining the original selections provided $n_{p_i} \geq 1$, n being the total number of PSUs and p_i is the selection probability. For example, expansion of sample from two PSUs to three or four PSUs was achieved by applying to the randomized list of PSUs a random start r + 1/2 or r - 1/2 depending on the original r used for sampling two PSUs was $\leq 1/2$ or $> 1/2$. This procedure was roughly equivalent to selecting four PSUs in the original sample itself; although in most cases the joint inclusion probabilities would differ slightly in the two situations. Following the same principle, any number of additional PSUs may be selected at any time. In the case that the additional PSUs were odd in number then an even number of PSUs, given by the old number plus one, was first selected systematically after which one was dropped at random.

Sampling within additional PSUs was undertaken in the same manner as described in earlier sections in relation to the originally sampled PSUs. That is, the desired number of urban and rural groups were selected from thes PSUs and sent to the RO for obtaining field counts, after which clusters were delineated and selected. Although the algorithm for selection of clusters from shared urban centers became more complicated, because in many cases the same urban center was shared among three or more PSUs, the same procedure of selection was used. That is, the sample of clusters for each PSU was selected from the entire urban center, by spacing different random starts (r_{i1j}, r_{i2j}, r_{i3j}, r_{i4j} ...) at the maximum possible distance from each other.

(voir chapitre 14). Afin de faciliter la tâche sur le terrain et pour assurer une meilleure couverture géographique, l'élargissement de l'échantillon dans les secteurs N.A.R. s'est effectué par l'augmentation du nombre d'U.P.É. dans l'échantillon. Selon l'importance de l'échantillon supplémentaire accordé à une province, on a choisi de une à quatre U.P.É. supplémentaires dans chaque strate. Par exemple, on a choisi une U.P.É. supplémentaire en Ontario, au Québec, en Nouvelle-Écosse et au Nouveau-Brunswick, et quatre par strate dans l'Île-du-Prince-Édouard.

La méthode d'échantillonnage systématique avec classement aléatoire et P.P.T. qui était utilisée à l'origine pour la sélection de deux U.P.É. par strate permettait de choisir un nombre quelconque d'U.P.É. supplémentaires, tout en conservant les sélections originales, à la condition que $n_{p_i} \geq 1$, n étant le nombre total d'U.P.É. et p_i la probabilité de sélection. Par exemple, l'élargissement d'un échantillon de deux U.P.É. à trois ou quatre U.P.É., s'effectuait en appliquant à la liste des U.P.É. établie au hasard, une origine choisie au hasard r + 1/2 ou r - 1/2 suivant que le r original utilisé pour l'échantillannage de deux U.P.É. était $\leq 1/2$ ou $> 1/2$. Cette méthode équivalait approximativement à choisir quatre U.P.É. dans l'échantillon original, bien que dans la majorité des cas, les probabilités composées de leur inclusion auraient sensiblement différé dans les deux cas. Suivant ce même principe, on peut choisir n'importe quel nombre d'U.P.É. supplémentaires en tout temps. Lorsqu'il y avait un nombre impair d'U.P.É. supplémentaires, un nombre pair d'U.P.É. était sélectionné systématiquement, en ajoutant un au chiffre impair, après quoi on éliminait une U.P.É. au hasard.

On a procédé à l'échantillonnage à l'intérieur des U.P.É. supplémentaires suivant la méthode décrite dans les paragraphes précédents relatifs à l'échantillonnage des U.P.É. premières: le nombre recherché de groupes urbains et ruraux ont été choisis dans ces U.P.É., puis adressés aux B.R. pour qu'ils soient dénombrés, après quoi on a délimité et choisi les grappes. Ce choix s'est fait suivant la même méthode bien que l'algorithme pour la sélection de grappes dans des centres urbains partagés devienne plus compliqué du fait qu'en de nombreux cas, plus de deux U.P.É. portageaient le même centre urbain. D'après cette méthode, l'échantillon de grappes pour chaque U.P.É. a été choisi dans centre urbain tout entier, en espaçant au maximum différentes origines choixies au hasard (r_{i1j}, r_{i2j}, r_{i3j}, r_{i4j}, ...).

CHAPTER – 7 – CHAPITRE

SPECIAL SELECTIONS

7.1 Introduction

In addition to the two regular area frames; viz., SR areas and NSR areas, the LFS is comprised of two other frames namely, (i) apartment frame and (ii) special area frame. Both these frames were in existence in the earlier design as well, although the coverage, and sampling methodology used for these frames were different.

7.2 Apartment Frame

7.2.1 Definition and Coverage

A separate frame of apartment buildings was created in 17 of the larger SRUs, namely Montréal, Toronto, Ottawa, Hamilton, London, Winnipeg, Calgary, Edmonton, Vancouver, Halifax, Québec City, Oshawa, Burlington, Windsor, Kitchener, Saskatoon and Victoria. To belong to the apartment frame, an apartment building must have at least 30 units and five floors of living quarters.

The purpose of setting up this frame was to provide another stage of stratification in larger cities, so as to ensure the representation of apartment dwellers, as they may differ in many characteristics from persons living in other types of dwellings. Also, by keeping the present and future apartments in a separate frame, it minimizes the chances of unusual growth in a particular cluster. Thus, it avoids increase in the sample size in the cluster due to construction of new apartment buildings and therefore greatly minimizes the need for the revision of the cluster.

In addition, there are other advantages to having a separate apartment frame. The selection of samples for other surveys and studies are simplified and also the updating of the sample can be carried out independently in each frame.

The apartment frame is kept open-ended and new apartment buildings, which satisfy the required criteria are added to the existing frame. The frame is kept up-to-date through a system of reporting in collaboration with the offices of the Central Mortgage and Housing Corporation and RO activity.

7.2.2 Stratification and Selection Procedures

Following the results of some theoretical and operational studies, the apartment frame has been stratified wherever possible by geographical location

SELECTIONS SPÉCIALES

7.1 Introduction

Outre les deux bases de sondage ordinaires, secteurs A.R. et N.A.R., l'E.P.A. utilise deux autres bases: (i) la base d'immeubles d'appartements et (ii) la base des secteurs spéciaux. Ces deux bases existaient déjà dans le précédent plan de sondage bien que leur couverture et la méthode d'échantillonnage employée dans leur cas étaient différentes.

7.2 Base d'immeubles d'appartements

7.2.1 Définition et champ d'observation

On a créé une base distincte d'immeubles d'appartements dans 17 des plus importantes U.A.R., soit à Montréal, Toronto, Ottawa, Hamilton, London, Winnipeg, Calgary, Edmonton, Vancouver, Halifax, Québec, Oshawa, Burlington, Windsor, Kitchener, Saskatoon et Victoria. Un immeuble d'appartements doit contenir au minimum 30 unités de logement et cinq étages pour faire partie d'une base d'immeubles d'appartements.

Ce type de base a été établi pour que l'on profite d'une stratification supplémentaire dans les grands centres urbains, garantissant une juste représentation des occupants d'appartements, lesquels pouvant présenter plusieurs caractéristiques qui diffèrent de celles des personnes résidant dans d'autres types de logement. De plus, en classant les immeubles d'appartements actuels et futurs dans une base distincte, on peut réduire les risques de croissance anormale dans une grappe en particulier. On évite ainsi la possibilité d'agrandissement de la taille de l'échantillon dans la grappe attribuable à la construction de nouveaux immeubles d'appartements, réduisant ainsi de beaucoup la nécessité de réviser la grappe.

Il existe d'autres avantages à disposer d'une base d'immeubles d'appartements distincte. La sélection d'échantillons destinés à d'autres enquêtes ou études s'en trouve simplifiée et la mise à jour de l'échantillon peut être entreprise séparément dans chaque base.

La base d'immeubles d'appartements demeure ouverte et tout nouvel immeuble d'appartements qui répond aux critères peut lui être adjoint. Sa mise à jour se fait par un système de déclaration de données auquel collaborent les bureaux de la Société centrale d'hypothèques et de logement et les B.R.

7.2.2 Méthodes de stratification et de sélection

Pour se conformer aux résultats d'études théoriques et pratiques, chaque fois que cela était possible, on a stratifié la base d'immeubles d'appartements par situation

and by size, the size being the number of suites. Consequently, the two largest SRUs, namely Toronto and Montréal, were stratified by both the location and the size of the apartment buildings, while seven medium sized SRUs (viz., Ottawa, Hamilon, London, Winnipeg, Calgary, Edmonton and Vancouver) were stratified by size only. In the remaining eight SRUs, since there were not as many qualifying apartment buildings, stratification was not considered necessary. The following table gives the number of sub-strata formed in various SRUs.

géographique et par taille, la taille étant le nombre d'appartements. Ainsi, la stratification des deux plus importantes U.A.R., Toronto et Montréal, s'est faite selon la situation géographique et la taille des immeubles d'appartements, tandis que celle des sept U.A.R. de taille moyenne (Ottawa, Hamilton, London, Winnipeg, Calgary, Edmonton et Vancouver) s'est effectuée selon la taille seulement. Dans les huit autres U.A.R. qui ne possédaient pas un assez grand nombre d'immeubles d'appartements répondant aux critères, la stratification n'a pas été jugée nécessaire. Le tableau ci-dessous présente le nombre de sous-strates formées dans les différents types d'U.A.R.

TABLE 7.1. Stratification of Apartment Sample Frame

TABLEAU 7.1. Stratification d'une base d'échantillon d'immeubles d'appartements

SRUs — U.A.R.	Serial number of geographical strata — Numéro de série de la strate géographique	Number of strata by size — Nombre de strates selon la taille
Toronto	1	3
	2	3
	3	3
	4	2
Montréal	1	2
	2	2
7 medium size SRUs – U.A.R. de taille moyenne	–	2 each – chacune
8 smaller SRUs – U.A.R. de petite taille	–	1

Within each ultimate stratum the apartment buildings (clusters) were selected systematically with probability proportional to size by applying a prespecified density factor. Households, as the second stage of sampling, were selected systematically following the procedure described in 5.5.

Les immeubles d'appartements (grappes) ont été sélectionnés dans chaque strate par la méthode systématique avec probabilité proportionnelle a la taille et en utilisant un facteur de densité prédéterminé. La deuxième étape de l'échantillonnage a consisté à sélectionner systématiquement les ménages, suivant la méthode décrite au paragraphe 5.5.

7.3 Special Dwelling Units

Smaller collective type dwellings or institutional type dwellings such as homes for the aged, convents, nursing homes, hostels, etc., were considered as "special dwelling units". Lists of special dwelling units were compiled at the time of field counts prior to clustering. Each time a new NSR group rotates into the sample, its special dwelling units are identified. The principal conceptual difficulty with the special dwelling units was in relating them to ordinary dwellings. For sampling purposes, in some cases an entire special dwelling unit may have been considered as one dwelling, while in other instances it may have been considered as comprising a number of dwelling units. Each dwelling unit

7.3 Unités spéciales de logements

Les logements de type collectif moins importants ou les institutions, tels que les foyers pour personnes âgées, les couvents, les maisons de repos, les refuges, etc. ont été classés "unités spéciales de logement". Avant de déterminer les grappes, on a établi au cours du dénombrement sur le terrain des listes d'unités spéciales de logement. Chaque fois qu'un nouveau groupe N.A.R. est introduit par renouvellement dans l'échantillon, on détermine les unités spéciales de logement qu'il contient. La principale difficulté qu'ont présentées les unités spéciales de logement, au niveau conceptuel, a été de les situer par rapport aux logements ordinaires. Pour les besoins de l'échantillonnage toute une unité spéciale de logement peut être, dans

belonging to the special dwelling, was assigned a probability of selection equal to that of other dwelling units in the cluster. This treatment of special dwellings ensured against instability in the sample yields and in the variance, which otherwise might have resulted when a special dwelling unit was rotated into the sample.

7.4 Special Area Frame

Institutions such as hospitals, schools, hotels, military establishments, remote areas, etc. were designated as special areas. For sampling purposes, the special area frame was divided into the following four group:

1. Military establishments.

2. Hospitals.

3. Other institutions.

4. Remote areas.

It may be noted that only the civilian population living on military establishments are included in the survey and that in the case of institutions, inmates of the institutions are not included in the survey.

Two primary reasons for separating special areas from the regular (SR or NSR) areas were:

(i) The special areas possessed characteristics differing from those of the population at large and therefore, if they were included in the regular areas, the sampling variance would increase. Further, if not segregated, they would introduce instability in the sample which would be particularly pronounced at the time these areas rotated in or out of the sample.

(ii) There have been instances when due to inaccessibility or other reasons, the method and/or the frequency of data collection has been modified. For example in some remote areas, interviews are conducted only once every six months with the data being copied during intervening months. Also in remote areas of Alberta, the Royal Canadian Mounted Police have, on occasion, done the interviewing. For administrative reasons then, it was convenient to have what represented the majority of the problem areas restricted to one part of the frame where they could be given closer attention.

certains cas, considérée comme un seul logement, alors que dans d'autres circonstances, on la considérera comme plusieurs unités de logement. On attribue à chaque unité de logement appartenant à un logement spécial une probabilité de sélection égale à celle des autres unités de logement dans la grappe. Cette façon de procéder permet d'éviter des fluctuations du rendement de l'échantillon et de la variance, qui pourraient se produire au moment où une unité spéciale de logement serait intégrée à l'échantillon par renouvellement.

7.4 Base de secteurs spéciaux

Certaines institutions telles que les hôpitaux, les écoles, les hôtels, les établissements militaires, les régions éloignées, etc. ont été désignées secteurs spéciaux. Aux fins d'échantillonnage, on a divisé la base de secteurs spéciaux en quatre groupes:

1. Les établissements militaires.

2. Les hôpitaux.

3. Les autres institutions.

4. Les régions éloignées.

Il faut noter que dans les établissements militaires, seule la population civile qui y réside est visée par l'enquête sur la population active et que les détenus dans les établissements pénitenciers ne font pas partie de l'enquête.

Les deux principales raisons pour lesquelles on a séparé les secteurs spéciaux des secteurs réguliers (A.R. ou N.A.R.) sont les suivantes:

(i) Les secteurs spéciaux possèdent des caractéristiques différentes de celles de la population en général; leur inclusion dans les secteurs réguliers augmenterait donc la variance d'échantillonnage. De plus, s'ils n'avaient pas été groupés à part, ils auraient constitué un facteur d'instabilité à l'échantillon, qui se serait manifesté particulièrement au moment du renouvellement.

(ii) Dans certains cas particuliers, pour des raisons d'inaccessibilité et (ou) pour d'autres motifs, il a fallu modifier la méthode ou la fréquence de collecte des données. Par exemple, dans certaines régions éloignées, les interviews n'ont lieu que tous les six mois et, dans l'intervalle, les données sont reproduites. La Gendarmerie royale du Canada s'est parfois chargée des interviews dans les régions éloignées de l'Alberta. Pour des raisons administratives, il fut jugé plus pratique de concentrer la plupart des secteurs susceptibles de poser des problèmes dans une section donnée de la base pour qu'on puisse leur accorder une attention plus poussée.

With a few exceptions each of the above four groups formed a stratum in each province except Prince Edward Island where there were no special areas.

The special areas were sample in three stages. The first stage units corresponded to EAs and were selected systematically from an ordered list with PPS, the eligible LFS population being the size measure. Subsequent stages of sampling were clusters and households. In some cases, however, clustering did not take place, for instance in hospitals and low population EAs.

À quelques exceptions près, chacun des quatre groupes précités a constitué une strate dans chaque province, sauf à l'Île-du-Prince-Édouard qui ne possède pas de secteurs spéciaux.

L'échantillonnage des secteurs spéciaux s'est fait en trois étapes. Les unités d'échantillonnage de la première étape correspondaient aux S.D.; elles ont été sélectionnées systématiquement à partir de listes ordonnées, avec P.P.T., la population visée par l'E.P.A. étant la mesure de la taille. Aux étapes suivantes, l'échantillon a été divisé en grappes et en ménages. Cependant, dans certains cas comme les hôpitaux et les S.D. à faible population, il n'y a pas eu de constitution de grappes.

CHAPTER – 8 – CHAPITRE

SAMPLE ROTATION

8.1 Introduction

Rotation of the LFS sample involves complete or partial replacement of sampling units carried out in such a way that the sample for each month is a probability sample. Rotation of sampling units occurs at each stage of the multi-stage design adopted for the LFS. The ultimate unit of selection, the household, is replaced every six months, while successively higher staged units remain in the sample for longer periods of time. The determination of six months as the period for rotation of households is a trade off between the cost of rotation and the increase in non-response that might occur if respondents were asked to remain in the survey for a longer period of time. In order to ensure uniform interviewer workloads and to minimize the effect of any bias due to the number of months a household has been in the survey, a rotation scheme was adopted whereby one sixth of the households rotate each month. This was achieved by associating with each PSU-GR-CLUS a rotation number between one and six which determines the months in which rotation of households takes place (rotation number one implies rotation of households in months January and July; two in February and August; etc.). Described in the following sections are the method by which rotation of units takes place, the operational steps involved in sample rotation, and the methods by which rotation numbers are assigned.

8.2 Method of Rotation

For the sampling methods employed in the design (randomized PPS systematic in the NSR areas, random group method in the SRUs, and PPS systematic in apartment and special areas) sample rotation was relatively straight forward. The ordering of units, given the initial selection, determined the sequence of rotation. The timing of the rotations was generally determined as follows. When a unit was selected at any stage of sampling prior to the household, a number, r, between one and the inverse sampling ratio, (i.s.r.) of the units was determined as a random start within the selected unit. In the case of clusters, a systematic sample of dwellings was drawn, using this random start and the cluster inverse sampling ratio as a step interval. Every six months the random start of the cluster is increased by one, and this is done in the month corresponding to the rotation number of the cluster. This results in a different sample of dwellings or rotation of households. Where the random start exceeds the cluster i.s.r., rotation of clusters takes place, that is the next cluster on the randomized

RENOUVELLEMENT DE L'ÉCHANTILLON

8.1 Introduction

Le renouvellement de l'échantillon de l'E.P.A. comporte le remplacement partiel ou total des unités d'échantillonnage, effectué de sorte à avoir chaque mois un échantillon probabiliste. Le renouvellement des unités d'échantillonnage se produit à chaque étape du plan de sondage à plusieurs degrés adopté pour l'E.P.A. La dernière unité de sélection, le ménage, est remplacée tous les six mois, tandis que les unités de sélection des étapes antérieures demeurent plus longtemps dans l'échantillon. La fixation à six mois pour la période de renouvellement des ménages est un compromis entre le coût du renouvellement et l'accroissement du taux de non-réponse qui peut se produire lorsque l'enquête vise les mêmes répondants pendant une période prolongée. Afin d'assurer une charge de travail égale aux interviewers et de réduire le biais de réponse qui pourrait se produire en raison du nombre de mois pendant lesquels un ménage est soumis à l'enquête, on a adopté un plan de renouvellement par lequel à chaque mois un sixième des ménages de l'échantillon est remplacé. À cette fin, chaque U.P.É.-G.R.-G.R.A. a reçu un numéro de renouvellement situé entre un et six, qui détermine le mois dans lequel le renouvellement du ménage aura lieu (le numéro de renouvellement un signifie que les ménages portant ce numéro seront renouvelés en janvier et juillet; deux, en février et en août; etc.). Les paragraphes suivants décrivent la méthode de renouvellement des unités d'échantillonnage, le processus de renouvellement de l'échantillon et les méthodes d'attribution des numéros de renouvellement.

8.2 Méthode de renouvellement

Dans le cas des méthodes d'échantillonnage utilisées dans le plan de sondage (méthode d'échantillonnage systématique avec classement aléatoire et P.P.T. dans les secteurs N.A.R., méthode des groupes aléatoires dans les U.A.R. et P.P.T. dans les secteurs spéciaux et les secteurs d'immeubles d'appartements), le renouvellement de l'échantillon a été relativement simple. Le classement des unités à partir de la sélection originale a déterminé la séquence de renouvellement. Le moment où le renouvellement s'effectuerait a été établi de la façon suivante. Lorsqu'une unité est sélectionnée à une étape d'échantillonnage précédant celle du ménage, un nombre r situé entre un et la fraction de sondage inverse, (F.S.I.) de l'unité détermine l'orgine choisie au hasard dans l'unité sélectionnée. Dans le cas des grappes, on tire un échantillon systématique de logements à l'aide de cette origine choisie au hasard, en utilisant la fraction de sondage inverse de la grappe comme intervalle de sélection. Tous les six mois, l'origine choisie au hasard de la grappe est augmenté d'une unité, au cours du mois correspondant au numéro de renouvellement de

list of clusters replaces the one already in the sample. The new cluster is given a random start of one, which is incremented in a similar fashion every six months.

For the SRUs as there was only one stage of selection prior to the household, each group remains in the sample for the life of the design, so that the only units which rotate are clusters and households.

In NSRU areas, in addition to the clusters, rural groups and PSUs also rotate. As urban centers were sampled with certainty within selected PSUs, they remain in the sample for the life of a PSU. Rotation for these units proceeds in a manner similar to cluster rotation. Every six months the random start for the unit is increased by one until it exceeds the i.s.r. of the unit. When this happens, then rotation takes place into the next unit on the randomized list, with a start of one, irrespective of whether all the units at lower stages have been enumerated or not. On an average, NSR rural groups stay in the sample for three to four years.

Due to the high cost of PSU rotation in NSR areas, a minimum period of roughly five years has been adopted before rotation takes place. This is achieved in such a manner that probabilities of selection are not altered. The minimum period for retention of PSUs is staggered throughout the strata of each province, to ensure against any "bumps" in statistics due to rotating many units at the same time.

8.3 Operational Steps Involved in Sample Rotation

Rotation of the sample has been automated to the extent of identifying at each survey, the units which are rotating in and out of the sample. These rotating units require manual processing in Sample control, Household Surveys Development Division, in terms of identifying the geographical area represented, and formation of lower staged sampling units where applicable.

Each of the cluster selection computer programs create **rotation records** which describe the pattern of rotation following each initial cluster selection. In addition, for parts of the frame selected manually, such as apartments and special areas, the frame was captured on computer and similarly, rotation records were created.

la grappe, ce qui produit un échantillon différent de logements ou un renouvellement des ménages. Lorsque l'origine choisie au hasard est supérieure à la F.S.I. de la grappe, il y a renouvellement des grappes, c'est-à-dire que la grappe suivante sur la liste aléatoire de grappes remplace celle de l'échantillon. On attribue à cette nouvelle grappe une origine choisie au hasard de un, qui se trouve ensuite augmentée d'une unité tous les six mois de la façon décrite plus haut.

Dans le cas des U.A.R., comme il n'y a qu'une étape de sélection avant celle des ménages, chaque groupe demeure dans l'échantillon pendant toute la durée de l'enquête, de sorte que seuls les grappes et les ménages sont renouvelés.

Dans les secteurs N.A.R., en plus des grappes, les groupes ruraux et les U.P.É. sont renouvelés. Les centres urbains des U.P.É. choisies étant échantillonnés automatiquement, ils demeurent dans l'échantillon aussi longtemps que l'U.P.É. Le renouvellement de ces unités s'effectue de la même façon que celui des grappes. Tous les six mois, l'origine choisie au hasard de chaque unité est augmentée de un jusqu'à ce qu'elle dépasse la F.S.I. de cette unité. Lorsque cette limite est atteinte, le renouvellement porte alors sur l'unité suivante de la liste aléatoire, avec une origine de un, que toutes les unités des étapes subséquentes soient dénombrées ou qu'elles ne le soient pas. Les groupes ruraux N.A.R. demeurent en moyenne de trois à quatre ans dans l'échantillon.

En raison du coût élevé du renouvellement des U.P.É. dans les secteurs N.A.R., on a fixé à environ cinq ans au minimum l'intervalle de renouvellement lequel s'effectue de telle façon que cela n'affecte pas les probabilités de sélection. La période minimale de conservation des U.P.É. dans l'échantillon varie parmi les strates de chaque province pour éviter les résultats en "dents de scie" que provoquerait le renouvellement d'un trop grand nombre d'unités en une seule fois.

8.3 Processus de renouvellement de l'échantillon

À chaque enquête, le renouvellement de l'échantillon se fait automatiquement en ce qui a trait à l'établissement des unités qui seront intégrées à l'échantillon ou en seront exclues. Ces unités sont traitées manuellement par la Section du contrôle et du développement de l'échantillon, qui doit déterminer la région géographique représentée et, éventuellement, former des unités d'échantillonnage pour les étapes subséquentes.

Chaque programme informatique de sélection de grappes crée des **fiches de renouvellement** qui décrivent le processus de renouvellement après la sélection de la grappe initiale. En outre, pour les éléments de la base sélectionnés manuellement, tels que les secteurs spéciaux et les secteurs d'immeubles d'appartements, la base a été mise sur support informatique et des fiches de renouvellement ont aussi été créées.

These rotation records describe a sequence of advancing starts and the subsequent pattern of rotation corresponding to each selection within a group. Each record lasts the life of the group or up to a maximum of 40 random starts. The collection of these records taken together, constitutes the master rotation file which is used to automatically rotate the sample. The master rotation file is changed as rotation patterns are updated (as in the open-ended apartment frame), and as old records become exhausted and are replaced by new ones.

The way in which rotation of households is accomplished is then as follows. Seven months in advance of a particular survey date, a file is created which contains design information on all clusters newly rotating into the sample for that survey date. This file is used to initiate the data base for listing purposes, and to identify for Sample Control the clusters for which cluster diagrams F01's are required. The F01's are sent to the ROs to be received co-incident with the Sample Status Report which is generated 22-23 weeks before the date of introduction and identifies all the new clusters requiring listing. The proeess of listing requires that a separate list line be completed on the Cluster List (F02) for each dwelling in the cluster.

The lists are entered on the mini-computers in the ROs and transmitted to the central data base. Prior to the survey data a file (which is created from the rotation files by Sample Design System) is fed into the data base; this file specifies the random start and inverse sampling ratio to be applied to each cluster in the sample. Thereafter, selection of households is accomplished by a systematic sample of the lists, and households dockets are printed for each selected household. Households remain in the sample for six months (or less if they are introduced in a month other than that associated with their rotation number as when phasing in new parts of the sample) at which time another set of random starts are reserved on the data base, thereby resulting in selection of new households.

In NSR areas, a rotation record lasts only as long as a group remains in the sample. Thus, as records become exhausted, it can be determined which groups are exhausted, and by referring to the group frame, replacement groups are determined. In the groups identified as rotating in, Sample Control determines their geographical location, and if it is a previously unclustered area, prepares maps to be sent to the field for detailed counts prior to clustering. After the counts are returned, the area is clustered, cluster record forms are completed, and clusters are selected by the cluster selection program which creates rotation records to be merged with the master rotation file. If the area was previously clustered, a copy of the original cluster record form is sent

Ces fiches décrivent une série d'origines choisies au hasard et la méthode de renouvellement correspondant à chaque sélection dans un groupe donné. Chaque fiche est valable pour la durée d'un groupe ou applicable à un maximum de 40 origines choisies au hasard. Toutes ces fiches réunies forment le fichier principal de renouvellement qui sert à renouveler automatiquement l'échantillon. Ce fichier principal est modifié chaque fois que le mode de renouvellement est mis à jour (cas de la base ouverte d'immeubles d'appartements) et au fur et à mesure que les fiches deviennent périmées et sont remplacées.

Le renouvellement des ménages se fait de la façon suivante. Sept mois avant la date officielle de l'enquête, on constitue un dossier contenant tous les renseignements sur la constitution des nouvelles grappes qui feront partie de l'échantillon à la date de l'enquête. Ce dossier sert à créer une base de données aux fins d'établissement de listes et à indiquer à la Section du contrôle et du développement de l'échantillon les grappes pour lesquelles il faut établir des schémas F01. Ces schémas sont adressés aux B.R. qui devront les recevoir en même temps que le Rapport sur l'état de l'échantillon, lui-même établi 22-23 semaines avant la date officielle de l'enquête et qui indique toutes les nouvelles grappes à porter sur la Liste de grappes (F02), qu'on dresse en inscrivant un logement de la grappe par ligne.

Les listes sont introduites dans les mini-ordinateurs des B.R. qui les transmettent à la base centrale des données. Avant l'introduction des données d'enquête dans la base des données, on a introduit un relevé, créé à partir des fiches de renouvellement par le Système de conception de l'échantillon; ce relevé indique l'origine choisie au hasard et la fraction de sondage inverse à appliquer à chaque grappe de l'échantillon. Ensuite, la sélection des ménages s'effectue par échantillonnage systématique des listes et un dossier est imprimé pour chaque ménage choisi. Les ménages font partie de l'échantillon pendant six mois (ou moins de six mois s'ils ont été pris en compte comme nouveaux éléments de l'échantillon au cours d'un mois autre que celui qui correspond à leur numéro de renouvellement), à la suite de quoi une nouvelle série d'origines choisies au hasard est réservée dans la base de données, ce qui permet la sélection de nouveaux ménages.

Dans les secteurs N.A.R., la fiche de renouvellement dure aussi longtemps que le groupe demeure dans l'échantillon. On peut ainsi déterminer quels groupes sont épuisés d'après les fiches périmées et quels seront les groupes de remplacement en consultant la base de groupes. La Section du contrôle et du développement de l'échantillon détermine la situation géographique des groupes devant faire partie de l'échantillon et s'il s'agit d'un secteur qui n'a pas été constitué en grappes, elle adresse les cartes de ce secteur au bureau régional concerné qui effectuera sur place le dénombrement. Sur réception des documents, la Section divise le secteur en grappes, les formules relatives aux grappes sont remplies et les grappes sont choisies par le programme de sélection des grappes qui, alors, crée des

along with control information necessary to ensure that when the data are processed by the cluster selection program, the same random ordering of units will result. Ordinarily, at least 52 weeks before the date of introduction, the process of group rotation begins.

With rotation of PSUs, the subsequent stages of selection, namely groups and clusters, will have to be carried out. Approximately the same length of time is required for replacement of PSUs, as formation and selection of groups is not very time consuming.

8.4 Assigning Rotation Numbers

In assigning rotation numbers, the objective is to evenly distribute the sample in terms of expected take and other design criteria (such as rural-urban proportion in NSRUs and fringe to built-up proportion in SRUs), achieving this distribution simultaneously for the sample as a whole, and at the same time for the smallest possible geographical sub-sets of it. Adherence to these objectives implies the following: (i) the workload of interviewers is stable, as roughly equal numbers of units are rotating each month; (ii) the effect of any bias due to the number of months a household has been in the sample is minimized regardless of the area for which estimates are being sought; (iii) the sample is effectively divided into six equally representative parts, which may be used when sub-samples from the RLFS frame are desired.

To achieve the aforementioned objectives in NSRUs, selected clusters were assigned rotation numbers in such a manner as to balance the rural-urban and total expected take within each stratum. The assignment of rotation numbers was accomplished independently within each stratum, but a factor was built in to randomly distribute that expected take which remained after equal amounts were assigned to each rotation number. Due to the randomness of the procedure, the unbalanced expected takes at the stratum level should tend to cancel out at higher levels. To further enhance the balancing of the expected take, within the rural and urban parts of a PSU, groups were ranked according to group density factors and were then processed from the largest to the smallest.

On the average, about 30 clusters were selected per stratum. After assigning rotation numbers the difference in expected take at the stratum level over the six rotations would be, at the most, approximately three households.

Applying the objective of dividing the sample into six equal representative parts of the SR universe, rotation numbers were assigned in such a manner that:

fiches de renouvellement qui seront ensuite jointes au fichier principal de renouvellement. Lorsque le secteur a déjà été constitué en grappes, une copie du relevé original des grappes est envoyée, accompagnée des données de contrôle nécessaires afin que lorsque les données seront traitées par le programme de sélection des grappes, on obtienne le même classement au hasard des unités. En général, le renouvellement des groupes débute au moins 52 semaines avant la date officielle de l'enquête.

À la suite du renouvellement des U.P.É., il faut appliquer les méthodes de sélection aux étapes suivantes, soit les groupes et les grappes. Il faut environ le même temps pour le remplacement des U.P.É., car la sélection et la formation des groupes exige peu de temps.

8.4 Attribution des numéros de renouvellement

Lors de l'attribution des numéros de renouvellement, il faut voir à répartir uniformément l'échantillon des points de vue rendement prévu et autres exigences du plan de sondage (telles que la proportion secteur urbain-secteur rural dans les U.N.A.R. et banlieue-zone bâtie dans les U.A.R.), tant dans l'ensemble de l'échantillon que pour la plus petite division géographique possible. Un tel objectif implique que: (i) la charge de travail des interviewers est stable étant donné, que chaque mois, un nombre à peu près égal d'unités est renouvelé; (ii) l'effet de toute déviation due au nombre de mois pendant lesquels un ménage demeure dan l'échantillon se trouve réduit quel que soit le secteur visé par l'évaluation; (iii) l'échantillon est effectivement divisé en six parties également représentatives dont on peut se servir lorsqu'on désire obtenir des sous-échantillons de la base de la N.E.P.A.

Afin d'atteindre les objectifs susmentionnés dans les U.N.A.R., on a attribué des numéros de renouvellement aux grappes choisies de façon à équilibrer le contenu urbain-rural et le rendement prévu dans chaque strate, L'attribution des numéros de renouvellement se fait indépendamment dans chaque strate, mais on introduit un facteur pour distribuer au hasard le rendement prévu restant après attribution de nombres égaux à chaque numéro de renouvellement. En raison du caractère aléatoire de la méthode utilisée, les rendements prévus non compensés au niveau de la strate devraient s'annuler aux paliers supérieurs. Pour assurer la compensation du rendement prévu dans les parties urbaines et rurales d'une U.P.É., les groupes sont classés suivant leur facteur de densité et traités par ordre de grandeur décroissant.

En moyenne, environ 30 grappes sont tirées au hasard dans chaque strate. Après attribution des numéros de renouvellement, la différence dans le rendement prévu au niveau de la strate, à la suite des six renouvellements, serait au plus d'environ trois ménages.

Dans le but de diviser l'échantillon en six parties représentatives égales dans l'univers A.R., les numéros de renouvellement sont attribués de façon à obtenir les résultats suivants:

(i) The expected take in each rotation is equal to 1/6 of the total expected take.

(ii) In every subunit, each rotation number was assigned the same number of times. Thus, each SRU would be represented in any sub-sample composed of sampling units with the same rotation number.

(iii) For a collection of subunits, the proportion of fringe to built-up groups assigned to each rotation should be equal. The collections of subunits taken were areas which themselves might be of particular interest at some future time, such as larger SRUs or groups of geographically nearby smaller SRUs. This measure in combination with (i) and (ii) ensured that over a collection of subunits, the sample among the rotation was as balanced as possible.

In some special cases, assignment of rotation numbers was done manually in accordance with the aforementioned objectives. Specifically, this applied to the apartment frame, the special area frame and unclustered NSR areas.

In the apartment frame, the first selected cluster (apartment) was randomly assigned a rotation number, and each new selection was assigned the next rotation number in a circular fashion from one to six.

A similar approach was used in the special area frame and in unclustered NSRU groups. In the case of special areas and unclustered NSRU, groups form the penultimate sampling units, however prior to sampling in these areas, it may have been decided to cluster them. If such was the case, the clusters were selected by the NSR cluster selection program. However all clusters were assigned the rotation number of the group.

(i) Pour chaque renouvellement, le rendement prévu équivaut à 1/6 du rendement prévu total.

(ii) Dans chaque sous-unité, le numéro de renouvellement est attribué le même nombre de fois; ainsi, chaque U.A.R. sera représentée dans tous les sous-échantillons formés d'unités d'échantillonnage ayant le même numéro de renouvellement.

(iii) Dans un ensemble de sous-unités, la proportion de groupes de banlieue et de groupes de zones bâties, lors de chaque renouvellement, devrait être égale. Les groupes de sous-unités en question sont des secteurs qui peuvent présenter un certain intérêt par la suite, tels que de plus grandes U.A.R. ou des groupes de plus petites U.A.R. avoisinantes. Cette disposition appliquée conjointement avec celles énoncées en (i) et (ii) garantit dans un ensemble de sous-unités la meilleure compensation possible de l'échantillon soumis aux renouvellements.

Dans certains cas particuliers, l'attribution des numéros de renouvellement se fait manuellement et en conformité avec les objectifs précités. Ce mode d'attribution s'applique en particulier à la base d'immeubles d'appartements, à la base de secteurs spéciaux et aux secteurs N.A.R. non constitués en grappes.

Dans le case de la base d'immeubles d'appartements, la première grappe choisie (appartement) reçoit un numéro de renouvellement choisi au hasard entre un et six, la seconde grappe choisie reçoit le suivant, et ainsi de suite.

On utilise une méthode identique pour la base de secteurs spéciaux et pour les groupes N.A.R. non constitués en grappes. Dans ces deux cas, les groupes forment les avant-dernières unités d'échantillonnage; cependant, avant de les échantillonner, on peut décider de les constituer en grappes, et dans ce cas, les grappes sont choisies par le programme de sélection des grappes N.A.R. Toutefois, toutes les grappes reçoivent le numéro de renouvellement du groupe.

CHAPTER – 9 – CHAPITRE

NUMBERICAL IDENTIFICATION SYSTEM

The numbering system identifies each record according to the primary sampling unit or subunit, group, cluster, rotation number, and household to which it corresponds. The numbering system consists of 15 digits divided into six fields as follows:

Digits	Field
1 - 5	PSU code
6 - 7	Group number
8 - 10	Cluster number
11	Rotation number
12 - 14	Listing number
15	Multiple households

The structure of each of these fields is presented in tabular form in Table 9.1, followed by a verbal description.

SYSTÈME DE DÉSIGNATION NUMÉRIQUE

Le système de numérotation sert à identifier chaque fiche par rapport à l'unité primaire d'échantillonnage ou la sous-unité, au groupe, à la grappe, au numéro de renouvellement et au ménage auquel elle correspond. Ce système se compose de 15 chiffres répartis en six zones comme suit:

Chiffres	Zones
1 - 5	Code de l'U.P.É.
6 - 7	Numéro de groupe
8 - 10	Numéro de grappe
11	Numéro de renouvellement
12 - 14	Numéro de liste
15	Ménages multiples

Le tableau 9.1 présente la répartition numérique par zone, suivie d'une description de chaque code numérique.

TABLE 9.1

TABLEAU 9.1

Term digit – Terme chiffre / Type of area – Type de secteur	PSU – U.P.É.				Group – Groupe	Cluster – Grappe	Rot. No. – Nº de renouv.	List – Liste	Mult. – Ménages mult.
	1	2	3	4 - 5	6 - 7	8 9 10	11	12 13 14	15
NSRU – U.N.A.R.	Province 0 - 9	ER R.É. 0 - 9	0	Stratum PSU – Strate U.P.É. 01 - 19 21 - 39 41 - 59 etc.	61 - 69 urban – urbain 71 - 89 rural	1 - 999 (000 if unclustered) – (000 si non constitué en grappes)	1 - 6	1 - 999 hhlds per clus. – ménages par grappe	0 - 9 No. of Mult. Hhlds. – Nbre de ménages mult.
SRU – U.A.R.	Province	ER – R.É.	1 - 8	Sub-units – Sous-unités 1 - 99	Random group – Groupe aléatoire 01 - 18 regular – normal 21 - 56 apt. – app.	1 - 499 built up – zone bâtie 501-999 fringe – banlieue	1 - 6	1 - 999 hhlds per clus. – ménages par grappe	0 - 9 No. of Mult. Hhlds. – Nbre de ménages mult.
Special – Spécial	Prov. ince	0	9	Stratum – Strate 00 - 29 30 - 59 60 - 89 90 - 99	90 - 99	1 - 999 (000 if unclustered) – (000 si non constitué en grappes)	1 - 6	1 - 999 hhlds per clus. – ménages par grappe	0 - 9 No of Mult. Hhlds. – Nbre de ménages mult.

Digit 1 – Digit 1 in all cases gives provincial code (0: Newfoundland, 1: Prince Edward Island, 2: Nova Scotia, ... 9: British Columbia). If the Northwest Territories and Yukon are brought into the sample, then column 1 will be an alpha-numeric code with the 11th province or territory being A, etc.

Digit 2 – Digit 2 indicates the ER in NSRUs and SRUs and takes on values 1 - 9 and 0 as at most there are 10 ERs per province. ERs are not identified for special areas and 0 is put under this column.

Digit 3 – In digit 3, 0 indicates NSRU, 1 - 8 indicates the SRU number within the ER (the largest SRU being given number 1, etc.), and 9 indicates special areas.

Digits 4 - 5 NSRU – Digits 4 - 5 indicate PSU and stratum number in the NSRUs. The numbering system allows for up to five strata each containing up to 19 PSUs. The numbers 01 - 19 are reserved for PSUs in stratum one, 21 - 39 for PSUs in stratum two, etc.

Digits 4 - 5 SRU – In the SRUs, digits 4 - 5 indicate subunit number allowing up to a maximum of 99 subunits per SRU, numbered consecutively starting from 01. If the SRU contains the apartment frame, then those subunits in the apartment frame constitute the last subunits, although the number of apartment subunits varies.

There are two exceptions to be noted in regards to the PSU number field in SRU. Both occur in ER 46 which contains Montréal. Montréal contains over 100 subunits so that Montréal was divided into the north shore with city number one and the south shore with city number two. Also in ER 46 there are eight SRUs, but since Montréal is counted as two, the two smallest SRUs (Joliette and Granby) have been treated as one by assigning eight to both in digit 3, however the subunits (one in each case) are numbered 46801 – Joliette and 46811 – Granby to distinguish that they are separate SRUs.

Digits 6 - 7 NSRU – In the NSRU, urban groups (i.e., urban centers) are numbered starting from 61 up to the number of urban centers in the PSU (maximum number of urban centers are three in the redesign). For NSRU rural in most cases a group corresponds to a rural EA, while in some cases due to size restrictions a group may comprise several EAs or an EA may comprise up to three groups. Numbering of rural groups is from 71 - 89.

Digits 6 - 7 SRU – These digits identify the group. In the regular SRU the groups correspond to a random collection of clusters of a particular type

1er chiffre – Le 1er chiffre indique dans tous les cas le code provincial (0: Terre-Neuve, 1: Île-du-Prince-Édouard, 2: Nouvelle-Écosse, ... 9: Colombie-Britannique). Si les Territoires du Nord-Ouest et le Yukon sont intégrés à l'échantillon, on les inscrira dans la colonne 1 à l'aide d'un code alphanumérique, soit: A pour la 11e inscription, etc.

2e chiffre – Le 2e chiffre indique la R.É. dans les U.N.A.R. et U.A.R. et peut se situer entre 1 - 9 ou 0, le nombre maximal de R.É. par province étant de 10. La R.É. n'est pas indiquée pour les secteurs spéciaux et l'on inscrit un 0 dans cette colonne.

3e chiffre – Pour le 3e chiffre, 0 indique l'U.N.A.R., 1 - 8 le numéro d'U.A.R. dans la R.E. (la plus grande U.A.R. recevant le numéro 1, etc.), et 9 un secteur spécial.

4e - 5e chiffres U.N.A.R. – Les 4e - 5e chiffres indiquent le numéro de strate et d'U.P.É. dans les U.N.A.R. Le système de numérotation permet un maximum de cinq strates pouvant contenir chacune jusqu'à 19 U.P.É. Les nombres 01 - 19 sont réservés aux U.P.É. de la strate un, 21 - 39 aux U.P.É. de la strate deux, etc.

4e - 5e chiffres U.A.R. – Dans les U.A.R., les 4e - 5e chiffres indiquent le numéro de sous-unité avec un maximum de 99 sous-unités par U.A.R., numérotées à partir de 01. Si l'U.A.R. contient la base d'immeubles d'appartements, les sous-unités de cette base constituent les dernières sous-unités, bien que le nombre de sous-unités d'immeubles d'appartements puisse varier.

En ce qui concerne la zone du numéro d'U.P.É. dans l'U.A.R., il faut noter deux exceptions ayant trait à la R.É. 46 qui contient Montréal. Montréal, contenant plus de 100 sous-unités, a été divisée en deux: la rive nord, portant le numéro de cité un et la rive sud, avec le numéro de cité deux. De plus, la R.É. 46 contient huit U.A.R., mais étant donné que Montréal est divisé en deux, les deux plus petites U.A.R. (Joliette et Granby) ont été réunies en une seule par l'attribution à chacune du numéro huit comme 3e chiffre; toutefois, les sous-unités (une dans les deux cas) portent les numéros 46801 (Joliette) et 46811 (Granby) pour bien indiquer qu'elles sont des U.A.R. distinctes.

6e - 7e chiffres U.N.A.R. – Dans les U.N.A.R. urbaines, le numérotage des groupes (ou centres urbains) débute à 61 jusqu'au nombre de centres urbains dans l'U.P.É. (le nombre maximal de centres urbains est fixé à trois dans la refonte de l'enquête). Dans les U.N.A.R. rurales, un groupe correspond en général à un S.D. rural, tandis que dans certains cas, pour des raisons de taille, un groupe peut se composer de plusieurs S.D. ou un S.D. peut contenir trois groupes. Le numérotage des groupes ruraux se fait de 71 - 89.

6e - 7e chiffres U.A.R. – Ces chiffres indiquent le groupe. Dans les U.A.R. normales, les groupes correspondent à un ensemble de grappes formées au hasard, appartenant à

(built-up or fringe) within a subunit. Up to 18 random groups numbered 01 - 18 are permitted (most of the subunits have six random groups while at present the maximum is 12 groups). Each of these groups have different density factors and if up-dating takes place due to growth, then these random groups will individually be revised.

In all those provinces where sample size was increased in the SR areas, the desired number of sub-groups were formed within the groups at the time of sample expansion. These sub-groups are not given separate identification numbers; although they can be identified through the cluster identification whenever necessary.

In the SRU apartment frame, group numbers run from 21 - 56. Here a group corresponds to numerically consecutive apartments on the list from which one is selected, so that there are at maximum 36 selections per apartment PSU. For cities in which there are more than 36 apartment selections, or in which there is stratification in the apartment frame there is more than one apartment PSU.

Digits 4, 5, 6, 7 – Special areas – In special areas as noted in section 7.4, stratification by type of special area takes place, and within each stratum a systematic sample is selected from a list composed of EAs of each type throughout the entire province. The numbering is then as follows: the 7-digit EA numbering system is adapted to the system by considering the first five digits as the "PSU" and the remaining two digits as a "Group". "Group" numbers, then, range from 90 - 99 within each "PSU", and "PSU" numbers range as follows:

A0900 - A0929 – Military establishments
A0930 - A0959 – Hospitals
A0960 - A0989 – Other institutions
A0900 - A0999 – Remote areas
where A = Province number (0 - 9).

According to the numbering system described above, a "special area PSU" represents a collection of ten EAs, however, this terminology is purely artificial in the context of sample design and estimation procedures.

Digits 8, 9, 10 – Within sampled groups, clusters are generally delineated from the stage of selection immediately prior to the household. In parts of the special areas and certain NSRU areas in Newfoundland however, no clustering takes place within selected groups so that dummy cluster numbers of 000 are used. Except in unclustered areas, rotation numbers have been assigned to the clusters. Three digits will enable

un type particulier (zone bâtie ou banlieue) et situé dans une sous-unité. On peut trouver un maximum de 18 groupes aléatoires numérotés de 01 - 18 (la plupart des sous-unités contiennent six groupes aléatoires tandis que le maximum actuel est de 12 groupes). Chaque groupe a un facteur de densité distinct et advenant l'augmentation de l'échantillon, la mise à jour de chaque groupe se fera séparément.

Dans toutes les provinces où la taille de l'échantillon a été augmentée pour les secteurs A.R., le nombre voulu de sous-groupes a été formé à l'intérieur des groupes au moment de l'agrandissement de l'échantillon. Ces sous-groupes ne reçoivent pas de numéro particulier, mais on peut les retracer par leur numéro de grappe.

Dans la base d'immeubles d'appartements A.R., les numéros de groupe se situent entre 21 - 56. Chaque groupe correspond à des appartements dont les numéros se suivent sur la liste; en l'un d'entre eux est sélectionné de sorte qu'il y a au maximum 36 sélections par U.P.É. d'immeubles d'appartements. Les villes pour lesquelles il y a plus de 36 sélections d'immeubles d'appartements ou qui contiennent une base d'immeubles d'appartements stratifiée ont plus d'une U.P.É. d'immeubles d'appartements.

4e, 5e, 6e, 7e chiffres – Secteurs spéciaux – Comme le signalait le paragraphe 7.4, les secteurs spéciaux sont soumis à la stratification par type de secteur spécial, puis on tire dans chaque strate un échantillon systématique à partir d'une liste composée de S.D. de tous ces types de secteurs spéciaux relevés dans toute la province. Le numérotage se fait comme suit: on se sert du système de numérotation à 7 chiffres des S.D. en prenant les cinq premiers chiffres pour représenter l'"U.P.É." et les deux chiffres suivants pour représenter le "groupe". Dans chaque "U.P.É.," les numéros de "groupe" se situent entre 90 - 99, et les numéros d'"U.P.É." sont les suivants:

A0900 - A0929 – Établissements militaires
A0930 - A0959 – Hôpitaux
A0960 - A0989 – Autres institutions
A0990 - A0999 – Régions éloignées
A représente ici le numéro de la province (de 0 - 9).

D'après la description ci-dessus du système de numérotation, une "U.P.É. de secteurs spéciaux" représente un ensemble de dix S.D.; cependant, ces termes sont purement théoriques dans le contexte du plan de sondage et des méthodes d'évaluation.

8e, 9e, 10e chiffres – Les grappes sont généralement délimitées à l'intérieur des groupes échantillonnés, au moment de la sélection et immédiatement avant la sélection des ménages. Toutefois, à Terre-Neuve, dans certaines parties de secteurs spéciaux et certains secteurs N.A.R., on ne constitue pas de grappes dans les groupes choisis et on utilise les numéros de grappes fictives 000. Sauf dans les secteurs non constitués en grappes, on a

listing for 999 clusters per group in NSRU and 999 clusters per subunit in SRUs. The cluster numbers are the same as those used on maps in Sample Control. In SRU subunits, cluster numbers 001 - 499 are fringe, while in NSRU groups cluster numbers start at 001 and continue up to 999.

Digit 11 – Rotation number represents a separate field independent of fields which identify units on a map. It takes on values 1 - 6 and identifies in what month households in a PSU-GR-CLUS rotate (rotation one, households rotate in January, July, rotation two in February, August etc.). Rotation numbers are assigned (at the cluster or group level) so as to divide the sample into six representative parts.

Digits 12, 13, 14 – Listing line number. When a cluster is listed, the lister assigns a three digit number to each dwelling in the cluster. Further, each time new dwellings are discovered, the lister gives the next available listing line number to the dwelling.

Digit 15 – This field indicates the presence of a multiple household. It is blank until a multiple dwelling is detected (multiple dwellings are found only at the time of interviewing in already selected "dwellings"). In the case of a multiple dwelling, a code of 0 is given to dwellings listed originally, the other dwellings are listed using the same listing line number and multiple codes 1, 2, 3 ... as necessary.

attribué des numéros de renouvellement aux grappes. On peut inscrire 999 grappes par groupe dans les U.N.A.R. et par sous-unité dans les U.A.R. Ce sont les mêmes numéros de grappes que ceux notés sur les cartes de la Section du contrôle et du développement de l'échantillon. Dans les sous-unités A.R., les numéros de grappes 001 - 499 sont attribués à la banlieue, tandis que dans les groupes N.A.R., les numéros de grappes s'étalent de 001 - 999.

11e chiffre – Le numéro de renouvellement représente une zone distincte, indépendante de celles qui servent à désigner les unités sur les cartes. Sa valeur varie de 1 - 6; elle détermine le mois où sont renouvelés les ménages d'une U.P.É.-G.R.-G.R.A.P. (les ménages avec le numéro de renouvellement un sont renouvelés en janvier et en juillet, le numéro deux signifie février et août, etc.). Dans chaque grappe ou chaque groupe, les numéros de renouvellement sont attribués de façon à diviser l'échantillon en six parties représentatives.

12e, 13e, 14e chiffres – Numéro de ligne de la liste. La liste d'une grappe est établie en attribuant un numéro de trois chiffres à chaque logement de la grappe. Par la suite, chaque fois qu'on trouve un nouveau logement, on lui attribue le premier numéro de ligne disponible sur la liste.

15e chiffre – Cette zone indique la présence d'un logement multiple. Elle n'est remplie que lorsqu'on a trouvé un logement multiple dans l'U.P.É. (les logements multiples ne sont repérés qu'au moment de l'interview dans les "logements" déjà sélectionnés). Dans le cas d'un logement multiple, on attribue le chiffre 0 au logement original et les autres logements reçoivent le même numéro de ligne de liste suivis des codes numériques 1, 2, 3 ... suivant le cas.

CHAPTER – 10 – CHAPITRE

DATA COLLECTION, CONVERSION AND PROCESSING

10.1 Introduction

This chapter describes briefly the major operational steps from the selection of dwellings to the tabulation of the monthly Labour Force estimates. These steps include the conducting of interviews, followed by data entry into the Regional Office (RO) minicomputers and transmittal to head office. Finally it touches upon the head office processing system including editing, imputation, weighting and tabulation of the survey data.

10.2 Data Collection

Earlier chapters described the sampling processes leading up to the selection of dwellings. This section discusses briefly the procedures followed by interviewers while collecting data. A more detailed step by step description may be found in the "Interviewer's Manual". Procedures for newly selected dwellings are mentioned first. These dwellings are those selected by computer and which have just rotated into the sample, or those which are selected manually by the interviewer because of newly found additions to the list. For those selected by computer, a household record docket or Form 03 (F03), along with three labour force survey questionnaires (F05's), are pre-printed in the RO. These contain a sample identification code (Primary Sampling Unit (PSU) group, cluster, rotation number, listing number and a preplanned assignment code) which enables the F03's and F05's to be printed by interviewer assignment.

Upon identifying selected dwellings, the first contact by interviewer with persons in selected dwellings is always a personal visit. The interviewers are allowed to accept response about all eligible members of the household from any responsible member (14 years of age and over) of that household.

Data on household composition is completed on the F03, and then a labour force questionnaire (F05) is completed by the interviewer for each eligible member of the household. Since only three F05's are pre-printed in the first month, the interviewer carries an extra supply of blank F05's which are used as necessary. The questionnaire is designed so as to utilize direct questionning techniques in order to reduce non-sampling errors.

COLLECTE, CONVERSION ET TRAITEMENT DES DONNÉES

10.1 Introduction

Ce chapitre décrit brièvement les principales étapes de l'E.P.A., de la sélection des logements à la compilation des données mensuelles de la population active. Ces étapes comprennent la conduite des interviews, suivie de l'introduction des données dans les mini-ordinateurs des bureaux régionaux (B.R.) et la transmission de ces données au bureau central. Pour terminer, elles englobent le système de traitement au bureau central, incluant le contrôle, l'imputation, la pondération et l'exploitation des données d'enquête.

10.2 Collecte des données

Les précédents chapitres décrivaient le processus d'échantillonnage qui conduit à la sélection des logements. Ce paragraphe décrit brièvement les méthodes qu'utilisent les interviewers pour la collecte des données. On trouvera une description plus détaillée dans le "Manuel de l'interviewer". Voici d'abord la description des méthodes de collecte utilisées pour les logements nouvellement choisis. Il s'agit des logements sélectionnés par ordinateur et qui viennent d'être intégrés à l'échantillon par renouvellement ou des logements sélectionnés manuellement par l'interviewer par suite de récentes additions à la liste. Dans le cas de logements sélectionnés par ordinateurs, le bureau régional imprime un dossier du ménage ou formule 03 (F03) ainsi que trois questionnaires d'enquête sur la population active (F05). Ces formules portent un code de désignation de l'échantillon (unité primaire d'échantillonage (U.P.É.) groupe, grappe, numéro de renouvellement, numéro de liste et code de tâche) qui permet d'imprimer les formules F03 et F05 par tâche d'interviewer.

Une fois les logements choisis, l'interviewer rend une visite personnelle aux ménages demeurant dans ces logements. L'interviewer peut accepter les réponses de tout membre responsable (de 14 ans et plus) de ces ménages pour tous les membres admissibles.

Les données sur la compostion des ménages sont inscrites sur la formule F03 et l'interviewer remplit un questionnaire sur la population active (F05) pour chaque membre admissible du ménage. Étant donné que le premier mois, on n'imprime que trois F05 par ménage, l'interviewer emporte quelques formules supplémentaires en cas de besoin. Le questionnaire est conçu de façon à permettre l'utilisation de questions directes afin de réduire les erreurs qui n'ont pas trait à l'échantillonnage.

If no eligible person is available or agreeable to respond to the interviewer's questions, the interviewer enters a code on the F03 which classifies the reasons for non-interview as one of the various categories of non-response or vacancy. In the event that no one is at home on initial contact, the interviewer calls back later in the week in an attempt to elicit a response.

In successive months in which households are in the sample, information such as demographic data and membership status of persons in the household is turned around and pre-printed on F03's. In addition an F05 is pre-printed for each member, containing information such as the name of employer, industry, occupation and the type of work. This makes it unnecessary for the interviewer to collect this information again unless threre has been a change. In some of the larger self-representing unit (SRUs) where permission was granted in the first month, interviews in subsequent months are carried out by telephone rather than through personal visit.

At the end of each working day the interviewer mails the forms completed on that day to the RO.

10.3 Data Entry Procedure in the Regional Office

Upon receipt in the RO of labour force documents from the interviewers, various checks are carried out which relate to the numbers and types of forms submitted. The F03 and F05 forms are then separated and batches of each type of form are created in preparation for data entry. These batches are then passed on to the data entry operators who enter the F03 and F05 data into the RO mini-computer using visual display units. For households in the survey for the 2nd to 6th months, much of the information on the F03 need not be data entered, unless there has been a change in this information from the previous month. To monitor and control the quality of data entry, verification of the data entry operation is undertaken on a 100% or a sample basis. This varies according to criteria concerning the type of form and the past performance of the data entry operator. Each evening all of the verified information which has been entered into the mini-computer is transmitted to the head office. These procedures are described in greater detail in the "LFS Clerical Operations Procedures Manual" and "Computer Operations Manual".

Lorsque dans un ménage, aucune personne n'est admissible ou ne désire répondre aux questions de l'interviewer, celui-ci inscrit un code sur la formule F03 qui établit la raison de l'absence d'interview parmi les diverses catégories de non-réponse ou de vacance du logement. Si, lors de la première visite de l'interviewer, il n'y a personne au domicile, l'interviewer se présente plus tard au cours de la semaine pour tenter d'obtenir une réponse.

Durant les mois pendant lesquels un ménage fait partie de l'échantillon, certaines données telles que les données démographiques et le statut des membres du ménage, sont reproduites et imprimées sur la formule F03. De plus, on imprime une F05 pour chaque membre du ménage, contenant des renseignements tels que le nom de l'employeur, l'activité économique, la profession et le genre de travail. Ceci évite à l'interviewer d'avoir à recueillir à nouveau ce type de renseignements, à moins qu'entre temps un changement ne soit intervenu. Dans certaines grandes unités autoreprésentatives (U.A.R.) lorsqu'un accord est intervenu en ce sens avec l'interviewé dès le premier mois, les interviews se font par téléphone plutôt que par visites à domicile.

L'interviewer fait parvenir le jour même, au B.R., toutes les formules remplies durant chaque jour de travail.

10.3 Entrée des données en ordinateur au bureau régional

Dès la réception des documents sur la population active envoyés par les interviewers, le B.R. effectue diverses vérifications portant sur le nombre et le type de formules recueillies. Les formules F03 et F05 sont séparées et réunies en lots en vue de l'entrée des données en ordinateur. Ces lots sont ensuite confiés aux préposés à l'entrée des données qui, à l'aide de terminaux à écran de visualisation, entrent les données contenues dans les formules F03 et F05 dans les mini-ordinateurs du B.R. La plupart des renseignements contenus sur la formule F03 n'ont pas besoin d'être introduits à nouveau en ordinateur pour les ménages faisant partie de l'enquête, entre les deuxième et sixième mois, à moins que des changements ne soient intervenus depuis le mois précédent. Le contrôle qualitatif de l'entrée des données s'effectue par une vérification totale de l'opération ou sur une base de sondage. La portée de la vérification varie suivant les critères applicables au type de formule et la qualité reconnue du travail du préposé. À la fin de la journée, toutes les données entrées dans les mini-ordinateurs qui ont été vérifiées, sont transmises au bureau central. Toutes les étapes de cette opération sont décrites plus en détail dans le "Manuel des opérations de bureau de l'E.P.A." et dans le "Manuel des opérations informatiques".

10.4 Data Processing in the Head Office

The source of the data for head office processing is the structure and response file. The structure and response file is used to receive the data which is transmitted on a daily basis from the ROs and to transmit to the ROs the information required for the preprinting of F03's and F05's. Each record on the structure and response file contains F03 data and F05 data for every member of the household and data on these records are only read and not changed (that is, no processing is done on the file).

Each record on the structure and response file is examined, and according to certain rules, records containing data which require processing are selected, reformatted and copied onto the consolidated file. The consolidated file represents one respondent with the F03 data simply repeated on each consolidated record for all members covered by the F03. Certain errors (example a household member with a coded age of ten and a completed F05) are detected at this stage and the records in error are printed out for correction by the editor. After all data have been received from the ROs for households having no members with F05 data, a consolidated record is created from the F03 data on the structure and response file to ensure that the household is properly represented in the subsequent processing and weighting procedures. The consolidated file is the basis for all future processing. Each record on the consolidated file is the basis for all future processing. Each record on the consolidated file must successfully pass through a series of four process steps before it is classified as clean data.

Process step one examines each record on the consolidated file and prints out turn around documents (TAD) for records having information in the notes sections of the F05. The interviewer uses the notes sections if a code "Other" is used for any responses or if they encounter an unusual or difficult situation not covered by conventional means. The editors examine these notes to determine if an alternative code is appropriate or if changes in the record content are necessary as a result of the note.

Once a record passes process step one it is submitted to process step two for coding of industry, occupation and the class of worker. In process step two each consolidated record is examined and if any new information is entered in the current month for questions related to industry, occupation, or class of worker, the record is written out and routed to proper locations for industry, occupation and class of worker coding.

10.4 Traitement des données au bureau central

La source de données pour le traitement au bureau central est le fichier des éléments de base. Ce fichier sert à recevoir les données transmises quotidiennent par les B.R. et à transmettre à ces B.R. les données nécessaires pour imprimer les formules F03 et F05. Chaque fiche de ce fichier contient des données des formules F03 et F05 relatives à chaque membre du ménage, lesquelles ne peuvent être que lues par ordinateur et non modifiées (le fichier ne subit aucun traitement informatique).

Chaque fiche du fichier des éléments de base est examinée et, lorsqu'elle contient des données à traiter, est sélectionnée, rééditée et mise en mémoire dans le fichier principal, conformément aux méthodes d'explication. Chaque fiche du fichier principal représente un répondant et les données de la formule F03 sont simplement répétées sur la fiche de chacun des membres du ménage inscrits sur la F03. Certaines erreurs (par exemple, un membre d'un ménages dont l'âge codé correspond à dix ans et pour lequel on a rempli une formule F05) sont relevées à ce stade du traitement informatique et les fiches erronées sont imprimées aux fins de correction. Une fois que toutes les données ont été reçues des B.R., on crée, à partir du fichier des éléments de base, un fichier principal pour les ménages pour lesquels aucune formule F05 n'a été remplie afin que ces ménages soient adéquatement représentés au cours des opérations de traitement et de pondération. Ce fichier principal est la base sur laquelle porteront par la suite toutes les opérations du traitement informatique. Chaque fiche du fichier principal doit passer avec succès une série de quatre étapes de traitement pour que ses données soient classées exactes.

À la première étape, chaque fiche du fichier principal est examinée et des documents-navette sont imprimés pour les fiches ayant des données relevées dans les sections "Remarques" de la formule F05. L'interviewer remplit ces sections lorsqu'il utilise le code "Autre" pour une réponse donnée ou pour un cas complexe qu'il ne peut noter de façon conventionnelle. Le préposé à la vérification étudie ces inscriptions pour déterminer si un autre code peut leur être attribué ou s'il faut modifier le contenu de la fiche.

La fiche passe ensuite la seconde étape du traitement qui consiste à coder la branche d'activité, la profession et le statut professionnel. À cette étape, chaque fiche est examinée et lorsque, au cours du mois, on y a inscrit de nouvelles données relatives à la branche d'activité, la profession ou le statut professionnel, la fiche est imprimée aux fins de codage de ces trois types de données.

In process step three, records are examined for logic or validity errors in F05 data on the consolidated records and for those failing the edits TAD are printed for action by the editors. In the case of missing data on F05 information for an individual (e.g., responses to necessary questions left blank) attempts are made to copy from the individual's response for last month and if this is not possible data is substituted from a similar individual's F05 data. For certain types of non-response (no one at home, temporarily absent, or refusal to respond) where the household actually responded for the previous month, responses are imputed for the current month by copying the previous month's data after applying updating transformations where necessary. Once all individual records for a household pass process step three, the household is submitted to the next process step for demographic edits.

Process step four detects consolidated records having notes in the F03 section and prints the records for the editors' consideration. Logic or validity errors in individual demographic data are identified and the records are printed for correction. Process step four also detects any logic errors in the family relationships of all records.

There is a continual interaction between the head office processing system and manual procedures; that is errors are identified by the computer system which prints TADs for records requiring alterations. The editors then indicate the corrective actions on the right hand side of the TADs which are read by an optical character reader after which the corrections are made to the records on the consolidated file for subsequent runs of the editing process.

The data processing commences as soon as records are received in head office, usually on Tuesday of the survey week. Due to the deadline for publishing LF data, only two weeks, after the survey week, are alloted for editing and imputation. At this time, households failing edit specifications, if any, are deleted from the consolidated file and records are created for such households so that they are properly accounted for in the weighting system. The resultant consolidated file of clean records forms the input to the weighting system (as described in Chapter 11). Finally the weighted file is available for tabulation purposes. Weighting, tabulation and analysis of results takes place within a week prior to press release of the monthly Labour Force estimates, which is usually on the third Tuesday following survey week. A more detailed description of the head office processing of LFS data can be found in the "LFS Edit Manual".

La troisième étape consiste à examiner les fiches du point de vue logique ou vraisemblance des données tirées de la formule F05; dans le cas des fiches rejetées au contrôle, des documents-navette sont imprimés indiquant les corrections à apporter. Lorsque la formule F05 ne contient pas certaines données pour une personne (cas de réponses nécessaires laissées vides), on essaie de se procurer une copie de la réponse fournie par la personne le mois précédent et, lorsque cette méthode n'est pas applicable, on se sert de données d'une formule F05 établie pour un cas analogue. Pour certains types de non-réponse (personne à la maison, temporairement absent ou refus de répondre), lorsque le ménage a réellement répondu le mois précédent, on impute pour le mois en cours les données fournies le mois précédent, mises à jour si nécessaire. Une fois que toutes les fiches relatives à une ménage ont passé la troisième étape, on les soumet au contrôle démographique.

La quatrième étape du traitement informatique sert à relever les fiches contenant les remarques notées sur la formule F03 et à les imprimer pour les faire vérifier. Les erreurs de logique ou de vraisemblance faites dans les données démographiques individuelles sont repérées et les fiches sont imprimées aux fins de correction. La quatrième étape sert également à détecter toute erreur de logique dans les liens familiaux dans toutes les fiches.

Les deux modes de traitement, l'informatique et le manuel, sont interdépendants en ce sens que l'ordinateur détecte les erreurs et imprime des documents-navette pour les articles qui nécessitent des corrections, puis le préposé à la vérification indique à la droite du document-navette les corrections à apporter, qui sont lues par un lecteur optique de caractères, après quoi l'ordinateur corrige les fiches du fichier principal pour utilisation dans les vérifications subséquentes.

Le traitement des données se fait dès la réception des documents au bureau central, en général le mardi de la semaine d'enquête. En raison de la date limite imposée pour la publication des données de l'E.P.A., la vérification et l'imputation des données doivent être effectuées dans les deux semaines qui suivent celle de l'enquête. À cette date, les ménages qui ne remplissent pas les critères sont retirés du fichier principal et des fiches sont créés pour ces ménages afin qu'ils soient pris en compte dans le système de pondération. Le fichier principal ainsi constitué de fiches exactes contient les données à fournir au système de pondération (décrit au chapitre 11). Une fois soumis à la pondération, le fichier peut être mis en tableau. La pondération, la mise en tableau et l'analyse des résultats sont effectuées dans la semaine précédant la publication des relevés mensuels sur la population active, c'est-à-dire en général le troisième mardi qui suit la semaine d'enquête. On trouvera dans le "Manuel de vérification de l'E.P.A." une description plus détaillée du traitement des données de l'E.P.A. au bureau central.

ESTIMATION

11.1 Introduction

The principle behind the estimation procedure in a probability sample such as the LFS is that each person in the sample "represents" besides himself or herself, several other persons not in the sample. For example, in a simple random sample of 2%, each person in the sample represents 50 persons in the population. This could be achieved by producing 50 duplicates of each record in the sample and then proceeding to compile any aggregates or cross-classifications which would now refer to the entire population and would represent the estimates for the corresponding quantities in the population as obtained from the 2% sample.

In a probability sample, the sample design itself determines weights which may be used to produce unbiased estimates. Each record may be weighted by the inverse of the probability of selecting the person to whom the record refers (in the example of the 2% random sample this probability would be 0.02 for each person and so the records could be weighted by $1/0.02 = 50$). This may be called the **simple estimate.**

For the LFS the file created for tabulation purposes contains one record per selected person in the sample. Each record contains all labour force and demographic characteristics concerning selected individuals. Instead of physically duplicating the sample records, an overall weighting factor is entered on each record. The weighting factor refers to the number of times a particular record should be duplicated. For example, if the number of persons employed in manufacturing is to be estimated, this is done by sorting the records referring to those persons in the sample employed in manufacturing and summing the weights entered on these records.

Frequently we come across situations where objective information on certain relevant characteristics for the same universe is available from sources other than the survey itself. There are several estimation methods which utilize such auxiliary information in order to increase the reliability of the estimate. Ratio estimation is one of the most prevalent techniques of utilizing relevant information external to the survey. The main principle of **ratio estimation** may be summarized as follows: suppose that simple estimates of aggregates are produced for certain classification of the population (i.e., for age-sex groups or for the

ESTIMATION

11.1 Introduction

Le principe à la base de l'estimation dans un échantillon probabiliste tel que celui de l'E.P.A. est que chaque personne dans l'échantillon "représente", outre lui-même ou elle-même, plusieurs autres personnes qui ne font pas partie de l'échantillon. Par exemple, dans un échantillon aléatoire simple de 2 %, chaque personne dans l'échantillon représente 50 personnes de la population totale. On pourrait obtenir le même résultat en reproduisant à 50 exemplaires chaque dossier de l'échantillon et en procédant à la compilation de tous les groupements ou classifications recoupées qui porteraient alors sur la population entière et qui représenteraient les estimations des quantités correspondantes dans la population obtenues par l'échantillon de 2 %.

Dans un échantillon probabiliste, le plan de sondage détermine lui-même le coefficient de pondération à utiliser pour produire des estimations sans biais. Chaque dossier peut être pondéré par l'inverse de la probabilité de sélection de la personne à laquelle il se réfère (dans l'exemple de l'échantillon aléatoire à 2 %, cette probabilité serait de 0.02 pour chaque personne de sorte que le coefficient de pondération serait de $1/0.02 = 50$). On peut appeler cette méthode une **estimation simple.**

Dans l'E.P.A., le fichier créé aux fins de tabulation des données contient un dossier par personne sélectionnée faisant partie de l'échantillon. Chaque dossier possède toutes les caractéristiques démographiques et se rapportant à la population active des personnes sélectionnées. Au lieu de reproduire à plusieurs exemplaires les dossiers de l'échantillon, on inscrit sur chaque dossier un facteur global de pondération. Ce facteur de pondération désigne le nombre de fois qu'un dossier devrait être reproduit. Par exemple, pour estimer le nombre de personnes travaillant dans l'industrie manufacturière, on retire les dossiers relatifs aux personnes échantillonnées qui y travaillent et on fait la somme des coefficients de pondération attribués à ces dossiers.

Il est fréquent d'obtenir par des sources autres que l'enquête sur la population active, des données objectives sur certaines caractéristiques intéressantes. Il existe diverses méthodes d'estimation qui ont recours à ces données auxiliaires pour renforcer la fiabilité des estimations. L'estimation par quotient est l'une des méthodes les plus répandues pour l'emploi de données appropriées provenant de sources autres que l'enquête conventionnelle. On peut résumer le principe général de l'**estimation par quotient** à ceci: supposons que des estimations simples d'agrégats sont produites pour certaines classifications de la population (par exemple, pour des groupes selon l'âge et le sexe ou pour la

population in rural and urban areas, etc.) utilizing the simple estimating procedure described above. Assume also that reliable estimates or actual counts are available by aggregates from sources outside the survey for the same classifications of the population. One may then compare the estimates derived from the survey with those obtained from outside sources. The estimates from the outside sources are divided by the sample estimates for each classification and the weights of the records in each classification are adjusted by multiplying the weigts by this factor. After the adjustment of the weights the estimated aggregates will now agree with the estimate from the independent source for each classification. Ratio estimation is quite simple as compared to other methods of using external information and at the same time results in increased efficiency. The choice of external information is however, very crucial to the procedure as it leads to higher efficiency only if such information is highly correlated with the characteristics of interest in the survey.

In the LFS, the **final weight** attached to each record is derived in five steps and is the product of five factors. These are known as basic weight, rural-urban factor, balancing factor for non-response, cluster subweight and weighting to projected population (ratio estimate).

11.2 Basic Weighting Factor

As described in the introduction, the sample design itself determines a set of basic weights to be applied to each record referring to persons in the sample. This is called the basic weighting factor. The sample design is such that within the same province and same type of area (NSRU, SRU or special area), the basic weights are identical (except where specified) for each record (person) in the sample and are equal to the inverse of the sampling ratio. If data on all sampled households are available then the simple estimate is derived by applying the basic weights to each record in the sample.

11.3 Rural-urban Factor

As mentioned in the previous chapters, each primary sampling unit in the NSRU is composed of rural and urban areas, and the proportion of population belonging to the two areas differs from province to province and also from stratum to stratum within each province. Information concerning the total population in rural and urban areas is available from the 1971 Census for each PSU (whether it is in the sample or not) as well as for each province. Using the selected

population en zones rurales ou urbaines, etc.) suivant la méthode d'estimation simple décrite plus haut; supposons également que l'on dispose d'estimations fiables ou de chiffres absolus par agrégats pour ces mêmes classifications de la population et fournis par des sources autres que l'enquête; on peut alors comparer les estimations obtenues par voie d'enquête avec celles fournies par des sources extérieures. Les estimations de ces sources extérieures sont divisées par les estimations de l'échantillon pour chaque classification et on corrige les poids des dossiers de chaque classification en les multipliant par le facteur résultant de cette division. Une fois la correction des poids effectuée, les agrégats estimés correspondront avec l'estimation de chaque classification obtenue par une source indépendante. L'estimation par quotient est une méthode très simple comparativement aux autres méthodes faisant usage de données extérieures en plus d'être efficace. Le choix des données extérieures est cependant très important pour cette procédure car elle est plus efficace seulement si ces données extérieures, sont fortement corrélées avec les caractéristiques d'intérêt de l'enquête.

Dans l'E.P.A., **le poids final** de chaque dossier s'obtient au bout de cinq étapes et est le produit de cinq facteurs. Ce sont: la pondération de base, le facteur rural-urbain, le facteur de compensation pour la non-réponse, le facteur de sous-pondération de grappe et la pondération en accord avec l'extrapolation de la population (estimation par quotient).

11.2 Facteur de pondération de base

Comme l'introduction le précisait, le plan de sondage lui-même détermine un ensemble de poids de base à appliquer à chaque dossier relatif aux personnes faisant partie de l'échantillon, qu'on appelle le facteur de pondération de base. Le plan de sondage est conçu de telle façon que dans une même province et dans un même type de secteur (N.A.R., A.R. ou secteur spécial), ces facteurs de pondération sont identiques (sauf avis contraire) pour chaque dossier (personne) de l'échantillon et sont égaux à la fraction de sondage inverse. Lorsqu'on dispose de données sur tous les ménages échantillonnés, l'estimation simple s'obtient en appliquant les facteurs de pondération de base à chaque dossier de l'échantillon.

11.3 Facteur rural-urbain

Comme il a déjà été mentionné dans les précédents chapitres, chaque unité primaire d'échantillonnage dans une U.N.A.R. est formée de secteurs urbains et de secteurs ruraux, et la proportion de la population appartenant à ces deux secteurs diffère d'une province à l'autre, et dans chaque province, d'une strate à l'autre. On peut obtenir du recensement de 1971 des données sur la population en secteur urbain ou rural pour chaque U.P.É. (qu'elle fasse partie de l'échantillon ou qu'elle en soit

PSUs only, and dividing their 1971 rural or urban population by the known probability of selection, a "simple estimate" of the 1971 rural or urban population is obtained for each province. Comparison, by province, with the actual 1971 rural or urban census counts indicates whether the selected PSUs over or under represent the respective areas. The ratio of the actual rural-urban counts divided by the corresponding estimates, provides a measure of the representativeness of the sample. These two factors are termed as a rural and an urban factor for each province and are used in the form of ratio estimates. These two factors are computed for each province at the time of the selection of the PSUs and are entered on each sample record according to the appropriate area of that province. Changes in these factors are incorporated at the time of PSU rotations.

11.4 Treatment of Non-response

Some non-response is virtually certain to occur in any survey of human population whether it is because there is not one at home during the enumeration or for some other reason. In the LFS each month, the sample design completely specifies the households that are to be interviewed during interview week. Each interviewer is assigned a set of households and is given firm instructions to make every effort to interview these households. If in spite of all the attempts by the interviewer, certain households remain non-respondent, then the interviewer is asked to provide a reason for a non-response for each of these households.

In certain types of non-response such as "no one at home", "refusal to answer questions", or a "temporarily absent household", if the previous month's responses are available, then records are copied with suitable transformations being applied to certain fields and the response status is changed to that of the previous month. For estimation purposes these households are treated in the same way as any other responding household. These records are then flagged so that records will not be copied for more than one consecutive month.

To compensate for other types of non-response, such as "no call made" due to weather conditions, "no interviewer available", newly rotated households which are non-respondents, or households which are non-respondents for a second consecutive month, the "interviewed" households, are compensated by increasing their weight by a **balancing factor**. Balancing is carried out within each **balancing unit**.

exclue) ainsi que pour chaque province. En se servant seulement des U.P.É. sélectionnées et en divisant leur population rurale ou urbaine de 1971 par la probabilité de sélection connue, on peut obtenir pour chaque province une "estimation simple" de la population rurale ou urbaine de 1971. En comparant par province les résultats obtenus avec les données sur la population en secteur rural et urbain provenant du recensement de 1971, on peut déterminer si les U.P.É. sur-représentent ou sous-représentent ces secteurs respectifs. Le rapport des données du recensement concernant la population en secteur rural et urbain par les estimations correspondantes indique le degré de représentativité de l'échantillon. Ces deux facteurs sont désignés facteur rural et facteur urbain pour chaque province et servent aux estimations par quotient. Ces facteurs sont calculés pour chaque province au moment de la sélection des U.P.É. et sont inscrits sur chaque dossier de l'échantillon selon le secteur de cette province. Les modifications nécessaires sont ensuite apportées au moment du renouvellement des U.P.É.

11.4 Traitement de la non-réponse

Dans toute enquête menée auprès de la population, il est pratiquement certain qu'il y aura des cas de non-réponse, soit en raison de l'absence des répondants pendant la semaine de l'enquête, ou pour toute autre raison. Dans l'E.P.A., chaque mois le plan d'échantillonnage spécifie exactement quels sont les ménages à interviewer au cours de la semaine d'interview. Un groupe de ménages est attribué à chaque interviewer qui est fortement incité à prendre toutes les mesures nécessaires pour rencontrer ces ménages. Si, en dépit de tous les efforts de l'interviewer, certains ménages demeurent des non-répondants, il doit alors fournir les raisons de chaque cas de non-réponse rencontré.

Pour certains types de non-réponse, tels que "personne à la maison", "refus de répondre aux questions" ou "ménage temporairement absent", pour lesquels on dispose de données recueillies au cours de l'interview du mois précédent. Les dossiers du mois en cours sont créés à partir de ces données avec les modifications qui s'imposent et le code-réponse devient celui du mois précédent. Pour les besoins de l'estimation, ces ménages sont traités de la même façon que tout autre ménage ayant accepté de répondre. Ces dossiers sont alors marqués pour qu'ils ne soient pas à nouveau copiés le mois suivant.

Pour compenser les autres types de non-réponse, tels que "aucune visite" en raison des intempéries, "aucun interviewer disponible", les ménages qui viennent d'être introduits dans l'échantillon par renouvellement et qui sont non-répondants ou ceux qui sont des non-répondants pour le deuxième mois consécutif, on augmente à l'aide d'un **facteur de compensation** le poids des ménages "interviewés". Cette compensation est effectuée dans chaque **unité de compensation**.

In the NSR areas, each sampled PSU is divided into two balancing units (a-urban and b-rural parts) and in the SRUs each subunit is a balancing unit. For each balancing unit the number of households, which should have been interviewed is divided by the number actually interview or input for, on the basis of last month's records, and this ratio (the balancing factor) is then entered on each sample record in that balancing unit. The estimates which are weighted up to the balancing unit by this ratio are based on the assumption that the households that have been interviewed represent the characteristics of the households that should have been interviewed. However, if this assumption is not true the estimates will be biased and the bias will increase with a higher rate of non-response. The exact magnitude of bias introduced by the adjustment for non-response is impossible to calculate. Consequently, rather than depending entirely on the adjustments for non-response, every effort is made to reduce it int the field.

Dans les secteurs A.R., chaque U.P.É. choisie est divisée en deux unités de compensation (a - partie urbaine, b - partie rurale) et dans les U.A.R., chaque sous-unité forme une unité de compensation. Dans chaque unité de compensation, le nombre de ménages qui auraient dû être interviewés est divisé par le nombre de ménages effectivement interviewés ou pris en compte d'après les dossiers du mois précédent, et le quotient (le facteur de compensation) est porté sur chaque dossier d'échantillon de cette unité de compensation. Les estimations qui sont pondérées jusqu'à l'unité de compensation par ce quotient sont calculées en supposant que les ménages qui ont été interviewés représentent les caractéristiques des ménages qui auraient dû être interviewés. Toutefois, advenant que cette supposition soit erronée, les estimations seront biaisées et cette déviation augmentera avec le taux de non-réponse. Il est impossible d'évaluer l'ampleur de la déviation provoquée par la correction de la non-réponse. Par suite, plutôt que de dépendre entièrement des ajustements pour non-réponse, on tente par tous les moyens possibles de réduire le taux de non-réponse sur le terrain.

11.5 Cluster Sub-weight

Each interviewer is assigned a specific set of households to enumerate during the interview week of each month. In the NSRU, each PSU is designed to yield an expected take suitable to make up an interviewer assignment, while in the SRUs assignments are formed from contiguous subunits taking into account the expected sample take at the design stage.

Further, each cluster has been designed to yield a sample take of two to three or four to six households respectively in NSR or SRU areas. The actual take is fairly robust against departure from these figures when growth is moderate; indeed each 100% increase in the number of households listed in a cluster versus design count results in an increase of only two to six households. Thus, substantial growth can be withstood in an isolated cluster before the additional take presents a field problem. If growth takes place in more than one cluster in an assignment, then the cumulative effect of smaller increases may also create a problem. In clusters where such substantial growth has taken place, subsampling may be resorted to as a means of avoiding disruptions in field operations. Rather than enumerate all of the households which should be selected, the inverse sampling ratio of the cluster is modified, say to k times its original value, which results in only 1 out of every k originally selected households being selected. The records for these households are then weighted by an additional factor equal to k, as each of these records represent k times as many records as was expected by design.

11.5 Facteur de sous-pondération de grappe

Chaque interviewer reçoit chaque mois un groupe spécifique de ménages à énumérer au cours de la semaine d'enquête. Dans les U.N.A.R., chaque U.P.É. est conçue de façon à produire un rendement déterminé suffisant pour constituer une tâche d'interviewer, tandis que les tâches dans les U.A.R. sont formées de sous-unités contiguës tout en tenant compte du rendement prévu de l'échantillon au moment de la planification de l'enquête.

De plus, chaque grappe est délimitée de facon à produire un échantillon de deux à trois ménages dans les secteurs N.A.R. et de quatre à six ménages dans les secteurs A.R. Le rendement réel s'écarte peu de ces chiffres lorsque l'augmentation de la grappe est modérée. En effet, chaque augmentation de 100 % dans le nombre de ménages d'une grappe, comparativement au nombre fixé par le plan de sondage, ne se traduit que par une augmentation de deux à six ménages, de sorte qu'une augmentation substantielle peut être tolérée dans une grappe donnée sans que le supplément de rendement ne cause un problème d'exécution. Lorsque cette augmentation touche plusieurs grappes faisant partie d'une même tâche, l'effet cumulé d'augmentations minimes peut également causer un problème d'ordre pratique. Dans les cas de grappes subissant une augmentation substantielle, le sous-échantillonnage peut être utilisé pour éviter des interruptions dans les opérations d'enquête. Au lieu de dénombrer tous les ménages qui devraient être sélectionnés, on modifie la fraction de sondage inverse de la grappe, soit k fois sa valeur initiale, ce qui entraîne la sélection d'un seul ménage sur tous les k ménages originellement choisis. Les dossiers créés pour ces ménages sont ensuite pondérés par un autre facteur égal à k, étant donné que ces dossiers représentent k fois le nombre de dossiers prévus par le plan de sondage.

In addition to their use, in sub-sampling of growth areas, cluster sub-weights have been used in the context of sample expansion in SRUs. As pointed out in section 5.6, at the time of sample increase in SRUs, each random group was sub-divided into the required number of sub-groups of approximately equal size. The cluster already being surveyed was retained to represent the sub-group to which it belonged and one cluster was selected from each additional sub-group(s). The new inverse sampling ratios (i.s.r.'s) were determined for all clusters. For the cluster which was already being surveyed, to avoid the overlap of systematic sample that would occur otherwise, one solution adopted was to continue to sample using the old i.s.r. and to introduce a cluster sub-weight at the estimation stage. This cluster sub-weight was determined as the ratio of old i.s.r. to the new i.s.r. for this cluster. This feature introduced in the weighting system adds to the flexibility of the overall design.

11.6 Age-sex Factor

By applying the previously described four weighting factors, a valid estimate could be derived for any aggregates for which information was obtained during the enumeration. In particular, estimates of the total number of persons are produced in each of the ten provinces in each of 20 age-sex groups. Independent estimates are available monthly for the totals in these 200 province-age-sex classes, by projecting forward the 1971 Census counts. In each class the independent estimate is divided by the sample estimate and this ratio is called the age-sex factor (ratio estimate). The age-sex factor is entered on all the records belonging to the appropriate class. The ten age groups are 14, 15 - 16, 17 - 19, 20 - 24, 25 - 34, 35 - 44, 45 - 54, 55 - 64, 65 - 69 and 70 or over.

11.7 Final Weights

The final weight of each record is the product of the five factors described above. In the final tabulations the estimated aggregate of each classification is obtained by summing the final weights of those records which indicate the presence of the characteristic. For example, to obtain the estimated aggregate of unemployed, the final weights of those records that indicate "unemployment" are summed.

11.8 Algebraic Description of the Estimation

The estimation procedure described above can be stated algebraically, and the following subscripts will be needed for that purpose:

Outre leur utilisation dans les sous-échantillons de secteurs agrandis, les sous-poids de grappe sont également utilisés dans les cas d'élargissement de l'échantillon dans les U.A.R. Comme le signalait le paragraphe 5.6, au moment de l'agrandissement de l'échantillon dans les U.A.R., chaque groupe aléatoire est divisé en un nombre donné de sous-groupes de taille approximativement égale. La grappe déjà sélectionnée pour l'enquête est retenue pour représenter le sous-groupe auquel elle appartient et on sélectionne une grappe dans chaque (ou dans le) sous-groupe supplémentaire. On détermine la nouvelle fraction de sondage inverse (F.S.I.) pour chaque grappe. Dans le cas de la grappe qui a déjà été sélectionnée pour l'enquête, afin d'éviter le chevauchement de l'échantillon systématique, on poursuit l'échantillonnage à l'aide de l'ancienne F.S.I. et on n'emploie le facteur de sous-pondération de grappe qu'au stade de l'estimation. Ce facteur de sous-pondération de grappe équivaut au quotient de l'ancienne F.S.I. et de la nouvelle. Ce nouvel élément du système de pondération contribue à la flexibilité du plan de sondage tout entier.

11.6 Facteur âge-sexe

À l'aide des quatre facteurs de pondération précités, on pourrait obtenir une estimation valide pour tous les agrégats pour lesquels des données ont été recueillies au cours du dénombrement. On calcule, par exemple, des estimations du nombre total de personnes pour chacun des 20 groupes d'âge selon le sexe dans chacune des dix provinces. Chaque mois, on produit des estimations indépendantes pour ces 200 classes, par province, selon l'âge et le sexe, en extrapolant les chiffres du recensement de 1971. Pour chaque classe, cette estimation indépendante est divisée par l'estimation de l'échantillon et le quotient obtenu s'appelle le facteur âge-sexe (estimation par quotient). Le facteur âge-sexe est inscrit dans tous les dossiers de la classe appropriée. Les dix groupes d'âge sont les suivants: 14, 15 - 16, 17 - 19, 20 - 24, 25 - 34, 35 - 44, 45 - 54, 55 - 64, 65 - 69 et 70 et plus.

11.7 Poids final

Le poids final de chaque dossier est le produit des cinq facteurs décrits ci-dessus. Dans les calculs finals, l'agrégat estimé de chaque classification s'obtient en faisant la somme des poids finals des dossiers qui contiennent une caractéristique donnée. Par exemple, pour obtenir l'agrégat estimé des chômeurs, on fait la somme des poids finals des dossiers qui possèdent la caractéristique "chômage".

11.8 Description algébrique de l'estimation

On peut représenter algébriquement la méthode d'estimation décrite ci-dessus à l'aide des symboles suivants:

p – for province.
h – for stratum in the non-self-representing areas and subunit in self-representing areas.
i – for the selected primary sampling units in stratum h (it is redundant in self-representing areas).
j – for type of area (j=1 self-representing, j=2 non-self-representing urban, j=3 non-self-representing rural).
c – cluster, group identification.
a – for age-sex groups.

A quantity referring to province p, stratum (or subunit) h, primary sampling unit i (in non-self-representing areas), type of area j cluster-group c and age-sex group a will be referred to briefly as the quantity in (p, h, i, j, c, a).

The following notations are used:

P_{phijca} – number of persons in the sample in (p, h, i, j, c, a).
X_{phijca} – number of persons in the sample having a particular labour force characteristic in (p, h, i, j, c, a).
$P_{(71)\ phi\ j}$ – 1971 Census count of persons in province p, stratum h, primary sampling unit i, type of area j.
$P_{(71)\ phi.}$ – 1971 Census count of persons in province p, stratum h, primary sampling unit i.
$P_{(71)\ ph..}$ – 1971 Census count of persons in province p, stratum h.
$P_{(71)\ p..j}$ – 1971 Census count of persons in province p, type of area j.
P_{pa} – total population as projected forward from the 1971 Census count to the current month in province p, age-sex-group a.

First the **rural-urban factor** F_{pj} is calculated for province p, type of area j. For the self-representing areas, $F_{pj} = 1$, where j signifies a self-representing area (j=1). In the non-self-representing areas, the probability π_{phi}, of selecting primary sampling unit i in stratum h, province p is proportional to its 1971 Census population and is actually equal to:

$$\pi_{phi} = n_{ph} \frac{P_{(71)\ phi.}}{P_{(71)\ ph..}} \qquad (11.1)$$

were n_{ph} is the number of PSUs selected per stratum. Note that $n_{ph} = 2$ in the original sample, however after the sample expansion, one to four additional PSUs were sampled from a stratum.

p – pour province.
h – pour une strate dans les secteurs non autoreprésentatifs et pour une sous-unité dans les secteurs autoreprésentatifs.
i – pour les unités primaires d'échantillonnage choisies dans la strate h (ferait double emploi dans les secteurs autoreprésentatifs).
j – pour le type de secteur (j=1 autoreprésentatif, j=2 urbain non autoreprésentatif, j=3 rural non autoreprésentatif).
c – désignation du groupe ou de la grappe.
a – pour groupes selon l'âge et le sexe.

Un nombre relatif à la province p, la strate (ou la sous-unité) h, l'unité primaire d'échantillonnage i (dans les secteurs non autoreprésentatifs), le type de secteur j, le groupe-grappe c et le groupe selon l'âge et le sexe a, sera décrit comme étant une quantité contenue dans (p, h, i, j, c, a).

On utilise les notations symboliques suivantes:

P_{phijca} – nombre de personnes dans l'échantillon dans (p, h, i, j, c, a).
X_{phijca} – nombre de personnes dans l'échantillon ayant une caractéristique particulière de la population active dans (p, h, i, j, c, a).
$P_{(71)\ phi\ j}$ – nombre de personnes d'après le recensement de 1971, dans la province p, la strate h, l'unité primaire d'échantillonnage i et le type de secteur j.
$P_{(71)\ phi.}$ – nombre de personnes d'après le recensement de 1971, dans la province p, la strate h et l'unité primaire d'échantillonnage i.
$P_{(71)\ ph..}$ – nombre de personnes d'après le recensement de 1971, dans la province p, la strate h.
$P_{(71)\ p..j}$ – nombre de personnes d'après le recensement de 1971, dans la province p, le type de secteur j.
P_{pa} – population totale pour le mois en cours, extrapolée des chiffres du recensement de 1971, dans la province p, dans le groupe a selon l'âge et le sexe.

On calcule en premier lieu le **facteur rural-urbain** F_{pj} pour la province p, type de secteur j. Dans les secteurs autoreprésentatifs, $F_{pj} = 1$, j signifiant un secteur autoreprésentatif (j=1). Dans les secteurs non autoreprésentatifs, la probabilité π_{phi} de sélectionner l'unité primaire d'échantillonnage i dans la strate h de la province p, est proportionnelle à sa population dénombrée par le recensement de 1971 et égale à:

$$\pi_{phi} = n_{ph} \frac{P_{(71)\ phi.}}{P_{(71)\ ph..}} \qquad (11.1)$$

où n_{ph} est le nombre d'U.P.É. choisies par strate. On remarquera que $n_{ph} = 2$ dans l'échantillon initial; toutefois, après l'agrandissement de l'échantillon, on a sélectionné de une à quatre U.P.É. supplémentaires dans une state.

The estimate derived from the sampled PSUs for the 1971 population in province p and type of area j is therefore equal to:

$$\hat{P}_{(71)\,p..j} = \sum_h \sum_i \frac{P_{(71)\,phi\,j}}{\pi\,phi} \qquad (11.2)$$

The rural-urban factor is therefore equal to:

$$F_{pj} = \frac{P_{(71)\,phi\,j}}{\hat{P}_{(71)\,p..j}} \qquad (11.3)$$

were j signifies non-self-representing rural or urban areas (j=2 or j=3).

The **cluster sub-weight** $C_{phigc} = 1$ in all cases except in the few instances where sub-sampling due to growth takes place, or where cluster sub-weights were applied at increase.

Next the **balancing factor** (B_{phij}) is calculated:

$B_{phij} = \frac{\text{number of households expected}}{\text{number of households enumerated}}$ in the sample in province p, stratum (or subunit) h, primary sampling unit i (in non-self representing areas), type of area j.

The expected number of households refers to the number of dwellings selected for a given survey excluding ineligible dwellings such as vacant dwellings, dwellings under construction etc. Enumerated households are these for which response is available either as a result of a successful interview or through the imputation procedure as discussed in section 10.4.

The **basic weight**, is the inverse of the sampling ratio since the sample is self-weighting. This weight is denoted by:

W_{phj}

It may be noted that W_{phj} differs from province to province and within a province it varies between SR and NSR areas. Within a type of area, W_{phj} is generally the same for all strata; however, after sample expansion, some SRUs (for example Victoria) have a different value of W_{phj} than the rest of the SRUs of the same province.

The age-sex factor is computed for each of 200 province-age-sex classifications as the ratio of the projected count and the sample estimate of total population for that classification. The sample estimate for a

L'estimation faite d'après les U.P.E. sélectionnées pour la population de 1971 dans la province p et le type de secteur j est donc égale à:

$$\hat{P}_{(71)\,p..j} = \sum_h \sum_i \frac{P_{(71)\,phi\,j}}{\pi\,phi} \qquad (11.2)$$

Le facteur rural-urbain est donc égal à:

$$F_{pj} = \frac{P_{(71)\,p..j}}{\hat{P}_{(71)\,p..j}} \qquad (11.3)$$

où j signifie des secteurs non autoreprésentatifs ruraux ou urbains (j=2 ou j=3).

Le **facteur de sous-pondération de grappe** $C_{phigc} = 1$ dans tous les cas sauf lorsqu'il y a sous-échantillonnage dû à l'accroissement de la population ou lorsque les facteurs de sous-pondération de grappe sont appliqués lors de l'agrandissement de l'échantillon.

Le facteur de compensation (B_{phij}) se calcule de la façon suivante:

$B_{phij} = \frac{\text{nombre prévu de ménages}}{\text{nombre de ménages dénombrés}}$ dans l'échantillon, dans la province p, la strate (ou la sous-unité) h, l'unité primaire d'échantillonnage i (dans les secteurs non autoreprésentatifs), le type de secteur j.

Le nombre prévu de ménages se rapporte au nombre de ménages sélectionnés pour une enquête donnée, excluant les logements non admissibles tels que les logements vacants, les logements en construction, etc. Les ménages dénombrés sont ceux pour lesquels une réponse a été fournie, soit par interview ou par la méthode d'imputation décrite au paragraphe 10.4.

Le **poids de base** est l'inverse de la fraction de sondate étant donné que l'échantillon est autopondéré. Ce poids est représenté par la formule suivante:

W_{phj}

Il faut noter que W_{phj} varie d'une province à l'autre et que dans une même province, il varie entre secteurs A.R. et secteurs N.A.R. Dans un type de secteur donné, W_{phj} est généralement le même pour toutes les strates; cependant, après l'agrandissement de l'échantillon, dans certaines U.A.R. (telles que Victoria), la valeur de W_{phj} est différente de celle des autres U.A.R. de la même province.

Le facteur âge-sexe est calculé pour chacune des 200 classifications âge-sexe-province comme étant le rapport du chiffre prévu et de l'estimation de la population totale pour cette classification provenant de l'échantillon.

province-age-sex classification, based on the previously developed five weighting factors is equal to:

$$\hat{P}_{pa} = \sum_{h}\sum_{i}\sum_{j}\sum_{c} F_{pj}\, B_{phij}\, W_{phj}\, C_{phic}\, P_{phijca} \quad (11.4)$$

Therefore the **age-sex factor** is equal to:

$$A_{pa} = \frac{P_{pa}}{\hat{P}_{pa}} \quad (11.5)$$

Now the **final weight** is computed as the product of the five weighting factors:

$$W_{phijca} = F_{pj}\, B_{phij}\, W_{phj}\, C_{phic}\, A_{pa} \quad (11.6)$$

and the provincial estimate ($\hat{X}_p$) for a labour force characteristic is then obtained as:

$$\hat{X}_p = \sum_{h}\sum_{i}\sum_{j}\sum_{c}\sum_{a} W_{phijca}\, X_{phijca} \quad (11.7)$$

11.9 Example of Weighting of a Record

The weighting system is an integral part of the head office processing system. The inputs to the weighting system consist of an edited unweighted file, and a file containing cluster sub-weights, rural-urban factor and census projections by age-sex categories. Provincial inverse sampling ratios are contained in tabular form within the actual weighting program itself. The unweighted file is first processed to calculate balancing factors for non-response. At this stage, a file of balancing factors is created and a printout of the file is generated. Balancing units with unusually high non-response are collapsed together and the balancing factor file is updated according to the revised balancing factors. The unweighted file is then processed by the weighting program to produce simple estimates for each age-sex category at the province level. Using these simple estimates and the census projections, age-sex factors are calculated. The now partially weighted file is processed again to calculate the final weight which is the product of the age-sex factor and the simple weight.

For illustration, the weighting of a hypothetical record will be considered. Consider a record corresponding to a male, age 26 who was a respondent in PSU-GR-CLUS 81007-71-006 in the survey month in question. Other information necessary to weight this record will be assumed at the step at which it is required.

L'estimation d'une classification âge-sexe-province, basée sur les cinq facteurs de pondération définis plus haut, est égale à:

$$\hat{P}_{pa} = \sum_{h}\sum_{i}\sum_{j}\sum_{c} F_{pj}\, B_{phij}\, W_{phj}\, C_{phic}\, P_{phijca} \quad (11.4)$$

Ainsi, le **facteur âge-sexe** est égal à:

$$A_{pa} = \frac{P_{pa}}{\hat{P}_{pa}} \quad (11.5)$$

Puis, le **poids final** se calcule comme étant le produit des cinq facteurs de pondération:

$$W_{phijca} = F_{pj}\, B_{phij}\, W_{phj}\, C_{phic}\, A_{pa} \quad (11.6)$$

et l'estimation par province ($\hat{X}_p$) d'une caractéristique de la population active s'obtient par la formule suivante:

$$\hat{X}_p = \sum_{h}\sum_{i}\sum_{j}\sum_{c}\sum_{a} W_{phijca}\, X_{phijca} \quad (11.7)$$

11.9 Exemple de pondération d'un dossier

Le système de pondération fait partie intégrante du système de traitement au bureau central. Les données introduites dans ce système consistent en un fichier non pondéré et contrôlé et en un fichier contenant des sous-pondérations de grappe, un facteur rural-urbain et des projections de chiffres du recensement relatifs à des catégories selon l'âge et le sexe. Les fractions de sondage inverses provinciales sont incluses dans le programme de pondération sous forme de tableau. Le fichier non pondéré est traité en premier lieu afin de calculer les facteurs de compensation pour la non-réponse. À ce stade du traitement, un fichier de facteurs de compensation est créé et ce fichier est imprimé. Les unités de compensation ayant un taux exceptionnellement élevé de non-réponse sont combinées et le fichier de facteurs de compensation est modifié d'après les facteurs de compensation rectifiés. Le fichier non pondéré est alors traité par le programme de pondération pour produire des estimations simples pour chaque catégorie selon l'âge et le sexe, au niveau provincial. On calcule les facteurs âge-sexe à l'aide de ces estimations simples et des projections de chiffres de recensement. Le fichier ainsi partiellement pondéré est traité à nouveau pour calculer le poids final qui est le produit du facteur âge-sexe par le poids simple.

Pour illustrer ce processus de pondération, prenons pour exemple la pondération d'un dossier hypothétique. Soit un dossier correspondant à une personne de sexe masculin, âgée de 26 ans, et qui a été un répondant dans l'U.P.É.-G.R.-G.R.A. 81007-71-006 au cours d'un mois d'enquête donné. Les autres données nécessaires à la pondération de ce dossier seront fournies au fur et à mesure qu'elles entreront dans le calcul.

Calculation of balancing factor – The record under consideration is in the NSRU rural of stratum 81000 (the first stratum in Economic Regional (ER) 1 of Alberta). The rural portions of each PSU selected in the stratum therefore form balancing units, i.e., in this case the rural portion of PSU 81007. To obtain the balancing factor it is necessary to divide the expected households (the number of households selected minus the vacant type dwellings) by the number actually interviewed or imputed on the basis of last months records. Assume also that cluster sub-sampling takes place in cluster 81007-71-006, say that in this case the cluster inverse sampling ratio was changed from one to seven. The cluster **sub-weight** then is seven. Further assume that in the cluster, expected take is eight households (if the i.s.r. had not been modified, the expected take would have been approximately 7 x 8 = 56 households). Of the eight households, assume six were respondents, one was not at home this month but had responded last month, and the last household was non-respondent for the second consecutive month. The response status of the first month non-respondent household for which there was a previous month's response would have been changed to a response prior to weighting, so that for the purposes of weighting the respondent households numbered seven, with one non-respondent household. In the remainder of the balancing unit (rural of PSU 81007) assume that the expected households were 44, the number of enumerated households equals 35, and further that the only cluster in the balancing unit with a cluster sub-weight is cluster 81007-71-006. The expected households as of design in the balancing unit is then given by 44 + (cluster sub-weight x expected take) = 44 + (7 x 8) = 100, and the enumerated households is given by 35 + (cluster sub-weight x enumerated households) = 35 + (7 x 7) = 84 (since households in sub-sampled clusters are weighted to reflect expected sample as of design). The balancing unit weight B_{phij} = 100/84 = 1.19048. If B_{phij} had been greater than three, then collapsing of units would take place and the rural portion for the whole stratum would be considered as a balancing unit for non-response, instead of each PSU being considered separately, and the balancing factors would be recalculated on this basis.

Calculation of simple weight – The PSU-GR identifies the record as NSRU rural for Alberta which establishes the following weights: the basic weight W_{phj} = 125, and the rural-urban factor for the province F_{pj} = .997114 (reflecting the fact that in the selected PSUs in Alberta at this time the rural population was slightly over represented). The balancing factor B_{phij} = 1.19048 was calculated previously, and a cluster sub-weight of 7.0 is input and applied to the records in this cluster. Thus the simple weight for the cluster is given by:

Calcul du facteur de compensation – Le dossier que nous allons étudier se trouve dans l'U.N.A.R. rurale de la strate 81000 (la première strate de la région économique (R.É.) 1 de l'Alberta). Les parties rurales de chaque U.P.É. choisie dans la strate forment ainsi des unités compensées, c'est-à-dire dans ce cas la partie rurale de l'U.P.É. 81007. Pour obtenir le facteur de compensation, il faut diviser le nombre de ménages prévus (le nombre de ménages choisis moins le nombre de logements vacants) par le nombre de ménages réellement interviewés ou imputés d'après les chiffres des derniers mois. Supposons également que le sous-échantillonnage de la grappe vise la grappe 81007-71-006; dans ce cas, la fraction de sondage inverse a été changée de un à sept. Le **sous-poids** de la grappe est alors de sept. De plus, imaginons que dans la grappe, le rendement prévu est de huit ménages (si la F.S.I. n'avait pas été modifiée, le rendement prévu aurait été d'environ 7 x 8 = 56 ménages). De ces huit ménages, suppons que six ont effectivement répondu, un était absent mais avait répondu le mois précédent et le dernier a été un non-répondant pour le deuxième mois consécutif. Le statut de non-réponse du ménage n'ayant pas répondu pour la première fois et pour lequel une réponse avait été obtenue le mois précédent aurait été changé en "réponse donnée" avant la pondération de sorte qu'aux fins de pondération, les ménages ayant répondu se seraient chiffrés à sept plus un ménage non-répondant. Dans la partie restante de l'unité compensée (partie rurale de l'U.P.É. 81007), supposons que le nombre de ménages prévus s'élevait à 44, le nombre de ménages dénombrés à 35 et que, de plus, la seule grappe de l'unité compensée avec une sous-pondération de grappe est la grappe 81007-71-006. Le nombre de ménages prévus d'après le plan de sondage dans l'unité compensée se calcule comme suit: 44 + (sous-poids de grappe x rendement prévu) = 44 + (7 x 8) = 100, et le nombre de ménages dénombrés: 35 + (sous-poids de la grappe x ménages dénombrés) = 35 + (7 x 7) = 84 (étant donné que le nombre de ménages dans les grappes sous-échantillonnées est pondéré pour refléter l'échantillon prévu d'après le plan de sondage). Le poids de l'unité compensée B_{phij} est égal à: 100/84 = 1.19048. Si B_{phij} avait été supérieur à trois, il y aurait eu combinaison d'unités et on aurait considéré la portion rurale de la strate comme une unité compensée pour la non-réponse au lieu de considérer chaque U.P.É. séparément, et les facteurs de compensation seraient recalculés sur cette base.

Calcul du poids simple – L'U.P.É.-G.R. désigne le dossier comme étant une U.N.A.R. rurale de l'Alberta, ce qui donne les poids suivants: le poids de base W_{phj} = 125 et le facteur rural-urbain pour la province F_{pj} = .997114 (reflétant le fait que dans les U.P.É. choisies de l'Alberta, la population rurale était légèrement sur-représentée). Le facteur de compensation B_{phij} = 1.19048 a été calculé précédemment et un sous-poids de grappe de 7.0 est imputé et appliqué aux dossiers de cette grappe. Le poids simple de la grappe se calcule comme suit:

$S_{phij} = W_{phj} F_{pj} B_{phij} C_{phic}$
$= 125 \times .997114 \times 1.19048 \times 7$
$= 148.38053 \times 7$
$= 1038.6637$

Calculation of the age-sex factor and final weight – Each record falls into one of 200 age-sex categories, and as the simple weights are determined within a province, the weights are accumulated in each category. The ratio of the census age-sex projection for a province divided by the sum of the simple weights for the age-sex category gives the age-sex factor. In the example the record falls into the male 25 - 34 category, say for which the sum of the simple weights for Alberta is 111,919.4726. For this age-sex category the census projected population for the month in question is say 121,745.9822. The age-sex factor then is given by:

$A_{pa} = 121,745.9822/111,919.4726 = 1.0878.$

Once the age-sex factors are calculated the file is processed again to calculate the final weight given by:

$W_{phijca} = S_{phijc} A_{pa}$

which in this hypothetical case is given by:

$W_{phijca} = 1038.6637 \times 1.0878$
$= 1129.8586$

$S_{phij} = W_{phj} F_{pj} B_{phij} C_{phic}$
$= 125 \times .997114 \times 1.19048 \times 7$
$= 148.38053 \times 7$
$= 1038.6637$

Calcul du facteur âge-sexe et du poids final – Chaque dossier se situe dans l'une des 200 catégories selon l'âge et le sexe, et les poids simples étant déterminés pour chaque province, les poids sont cumulés pour chaque catégorie. Le rapport de la projection par âge-sexe du recensement pour une province, divisé par la somme des poids simples pour une catégorie selon l'âge et le sexe, donne le facteur âge-sexe. Dans cet exemple, le dossier appartient à la catégorie "sexe masculin de 25 - 34 ans", catégorie pour laquelle en Alberta la somme des poids simples est disons de 111,919.4762. Pour cette même catégorie, la population de recensement projetée pour le mois en question est disons de 121,745.9822. Le facteur âge-sexe est donc:

$A_{pa} = 121,745.9822/111,919.4726 = 1.0878.$

Une fois les facteurs âge-sexe déterminés, le fichier est traité à nouveau pour calculer le poids final à l'aide de la formule suivante:

$W_{phijca} = S_{phijc} A_{pa}$

qui, dans ce cas hypothéthique, est donné par la formule:

$W_{phijca} = 1038.6637 \times 1.0878$
$= 1129.8586$

CHAPTER – 12 – CHAPITRE

SAMPLING ERROR

12.1 Introduction

To appreciate the need for sample surveys, it is necessary to understand clearly the role of sampling and other errors in such surveys. In sample surveys, since inference is made about the entire **population** covered by the survey on the basis of data obtained from only a part (**sample**) of the population, the results are likely to be different from the "true" population values. The true population values in this context refer to the values that would have been obtained when the entire population was enumerated under the same general survey conditions. The error arising due to drawing inferences about the population on the basis of information from the sample is termed **sampling error**.

The sampling error, in addition to the size of the sample depends on factors such as variability in the population, sampling design, and method of estimation. For example, for a given size of the sample, the sampling error will depend on the stratification procedure employed, allocation of the sample, choice of sampling units, and method of selection employed at different stages of sampling in a multi-stage design. Further, even for the same sampling design, the estimation procedure plays an important role and we can make different types of calculation to arrive at the most efficient (i.e., with least sampling error) estimation procedure. It should also be recognized that even if the sample design, the size of the sample, and the estimation procedure were the same, different characteristics (on which data have been collected from the sample) would possess different sampling errors. This would be due to the fact that they have different degrees of variability in the population. For instance, the sampling error for "unemployed" estimates would be different from the sampling error for, say, "not in labour force" estimates, both estimates being based on the same sample.

The efficiency of a design is usually measured by the so called **mean square error** of one or more characteristics. The mean square error is defined as the average of the square of deviations of the estimate from the "true value", taken over all possible samples. For a probability sample, such as the LFS, we can obtain the frequency distribution of the estimates using the theory of probability. The average of the estimates over all possible samples is termed the **expected value** of the estimate.

ERREUR D'ÉCHANTILLONNAGE

12.1 Introduction

Pour pouvoir juger de la nécessité d'effectuer des enquêtes par sondage, il faut d'abord bien connaître les conséquences des erreurs d'échantillonnage et des erreurs imputables à d'autres facteurs dans ce genre d'enquête. Étant donné que, dans les enquêtes par sondage, on tire des conclusions pour l'ensemble de la **population** visée à partir de données recueillies auprès d'une partie (**échantillon**) de la population, il est tout à fait vraisemblable que les résultats de l'enquête diffèrent des valeurs réelles de la population. Ces valeurs réelles sont précisément les valeurs que l'on obtiendrait si l'ensemble de la population était dénombré selon le même plan d'enquête. L'erreur commise lors des déductions faites au sujet de l'ensemble d'une population à partir des renseignements obtenus d'un échantillon s'appelle **l'erreur d'échantillonnage**.

L'erreur d'échantillonnage dépend non seulement de la taille de l'échantillon mais également de facteurs tels que la variabilité de la population, le plan de sondage et la méthode d'estimation. Par exemple, pour un échantillon d'une certaine taille, l'erreur d'échantillonnage peut varier selon la méthode de stratification utilisée, la répartition de l'échantillon, le choix des unités d'échantillonnage et la méthode de sélection utilisée à diverses étapes dans le cas d'un plan de sondage à plusieurs degrés. En outre, pour un même plan de sondage, la méthode d'estimation joue un rôle important et il faut parfois effectuer une variété de calculs avant d'obtenir la méthode d'estimation la plus efficace, c'est-à-dire ayant la moindre erreur d'échantillonnage. Il faut également reconnaître que même si le plan de sondage, la taille de l'échantillon et la méthode d'estimation étaient toujours les mêmes, les diverses caractéristiques de la population au sujet desquelles des données sont recueillies dans un échantillon entraîneraient des erreurs d'échantillonnage différentes. Cela serait dû au fait qu'elles possèdent divers degrés de variabilité dans la population. Par exemple, l'erreur d'échantillonnage des estimations concernant les chômeurs peut être tout à fait différente de l'erreur d'échantillonnage des estimations relatives aux inactifs, même si les deux types d'estimations sont fondées sur le même échantillon.

On évalue habituellement l'efficacité d'un plan par ce que l'on appelle **l'erreur quadratique moyenne** d'une ou de plusieurs caractéristiques. L'erreur quadratique moyenne est définie comme la moyenne du carré de l'écart entre l'estimation et la valeur réelle pour l'ensemble des échantillons possibles. Dans le cas d'un échantillon probabiliste, par exemple l'E.P.A., nous obtenons la distribution de fréquence des estimations en appliquant la théorie des probabilités. La moyenne des estimations pour l'ensemble des échantillons possibles s'appelle la **valeur prévue** de l'estimation.

The difference between this expected value and the true value is termed as **bias** of the estimate. **Variance** of the sample estimate is thus defined as the average, over all possible samples, of the squared difference of the estimate from its expected value. For situations where the expected value is the same as the true value, the estimate is termed as **unbiased estimate** and its mean square error becomes the same as the **variance**. It is worth mentioning here that certain biased estimation procedures (e.g., the ratio estimation) yield smaller mean square errors as compared to the unbiased estimates. The LFS estimates are almost unbiased, that is, bias due to the ratio estimation used in the LFS is negligible due to its large sample size. The ratio estimation procedure used in the LFS is more efficient than the simple estimate (as discussed in the previous chapter) for most of the characteristics for which estimates are produced regularly. The square root of the variance is called the **standard deviation** (or standard error) of the estimate.

The most important feature of probability sampling is that the variance of the estimate (and hence the standard deviation) may be estimated from the sample itself.

The variance of the final estimates produced from the LFS may be estimated using the sample data. A description of the procedure used is given in the following section. The sources, effects and control of non-sampling errors will be discussed in the following chapters.

12.2 Variance Estimation

On the basis of sample data, estimates of the sampling variability of LFS estimates can be calculated. In the following discussion, the estimation of the variability of a characteristic total estimate at the province level is presented.

The methodology adopted for estimating variances of LFS estimated characteristic totals was initially developed by Keyfitz [7] in 1957, and was subsequently modified to the requirements of the LFS.

For purposes of variance estimation in NSRU areas, the n_h PSUs selected from within stratum h are assumed to have been selected independently. In SRU areas each subunit h is split into two components – one component consisting of all clusters having an assigned even rotation number and the second component consisting of all clusters having an assigned odd rotation number and denote these components by $i = 1, 2$. The so formed half samples or pseudo-PSUs from within the subunit are assumed to have been independently

L'écart entre la valeur prévue et la valeur réelle représente le biais de l'estimation. On définit la **variance** de l'estimation provenant de l'échantillon comme étant la moyenne, pour l'ensemble des échantillons possibles, du carré de l'écart entre l'estimation et la valeur prévue. Dans les cas où la valeur prévue est égale à la valeur réelle, on obtient alors une **estimation sans biais** et son erreur quadratique moyenne correspond à la **variance**. Il convient de souligner que certaines méthodes d'estimation biaisée, par exemple l'estimation par quotient, donnent des erreurs quadratiques moyennes moins élevées que celles que produit l'estimation non biaisée. Les estimations de l'E.P.A. sont presque dénuées de tout biais, c'est-à-dire que le biais provenant de l'estimation par quotient appliquée dans cette enquête est négligeable à cause de la taille considérable de l'échantillon. La méthode de l'estimation par quotient appliquée dans l'E.P.A. est plus efficace que l'estimation simple (voir le chapitre précédent), et ce pour la majorité des caractéristiques pour lesquelles on produit des estimations régulièrement. La racine carrée de la variance fournit **l'écart-type** de l'estimation.

L'échantillon probabiliste a cet avantage que la variance de l'estimation (ainsi que l'écart-type) peut être calculée à partir de l'échantillon même.

La variance des estimations finales découlant de l'E.P.A. peut être établie à partir des données de l'échantillon. La section suivante donne une description de la méthode utilisée pour calculer cette variance. Les causes, les conséquences et la rectification des erreurs non imputables à l'échantillonnage sont expliquées aux chapitres suivants.

12.2 Estimation de la variance

Il est possible de calculer les estimations de la variabilité d'échantillonnage des estimations de l'E.P.A. à partir des données de l'échantillon. Dans la présente section, on étudiera l'estimation de la variabilité de l'estimation globale d'une caractéristique à l'échelon provincial.

La méthode utilisée pour estimer la variance des totaux estimés pour telle caractéristique de l'E.P.A. est la méthode élaborée par Keyfitz [7] en 1957 et adaptée aux exigences de l'E.P.A.

Dans le but d'estimer la variance dans les régions U.N.A.R. les n_h U.P.É. choisies à l'intérieur de la strate h sont supposées avoir été sélectionnées indépendamment. Dans les régions autoreprésentatives, chaque sous-unité h est divisée en deux composantes (la première comprenant toutes les grappes qui portent un numéro de renouvellement pair et la seconde comprenant toutes les grappes dont le numéro de renouvellement est impair) qui sont désignées par $i = 1, 2$. Ces demi-échantillons ou pseudo-U.P.É. formés à partir d'une sous-unité sont censés avoir

selected. Within special areas, for purposes of variance estimation, pseudo-strata are formed as either a design strata or as collapsed design strata and each pseudo-stratum is divided into from two to four components (pseudo-PSUs). Again components are assumed to have been independently selected from within the stratum. Thus, for the entire sample, h represents a stratum or subunit and i represents a PSU or pseudo-PSU selected from within a stratum or subunit.

Let $\hat{X}_A$ and $\hat{P}_A$ be the estimated number of persons having the labour force characteristic X and total population P, respectively, in area A. $\hat{X}_A$ and $\hat{P}_A$ are simply the sum of the final weights on all records referring to sampled persons in area A possessing the characteristic X and total population P respectively. From within each stratum or subunit h, $\hat{X}_{h1}, \ldots, \hat{X}_{hn_h}$ are assumed to be n_h independent estimates of $(\frac{1}{n_h})$th of the stratum total X_h, i.e., of $\frac{X_h}{n_h}$. For purposes of variance estimation a D-value is defined for each PSU. For estimating the sampling variability of $\hat{X}_p$, the estimator of the total number of persons in province p possessing the characteristic X, the D-values are defined as:

$$D_{phi}(\hat{X}_p) = \hat{X}_{phi} - \sum_{a=1}^{20} \frac{\hat{X}_{pa}}{\hat{P}_{pa}} \hat{P}_{phia} \qquad (12.1)$$

for each selected PSU in the province. The second term in the above expression accounts for the ratio estimation procedure which is applied in each age-sex post stratum. The estimate of the sampling variability of $\hat{X}_p$ is then expressed as:

$$\hat{Var}(\hat{X}_p) = \sum_{h \epsilon p} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} [D_{phi}(\hat{X}_p) - \frac{1}{n_h} \sum_{i=1}^{n_h} D_{phi}(\hat{X}_p)]^2 \qquad (12.2)$$

where $\sum_{h \epsilon p}$ denotes summation over all strata and subunits in province p and $\sum_{i=1}^{n_h}$ denotes summation over the n_h selected PSUs (or pseudo-PSUs) in stratum (or subunit) h.

The variance estimate for an estimated characteristic total at the Canada level, is simply the sum of corresponding provincial variance estimates since the sampling and estimation procedures are carried out independently within each province.

12.3 An Example of the Variance Calculations

We shall present a hypothetical case to illustrate the computation. Consider a province p, which is composed of two strata h = 1, 2 with two PSUs selected

été choisis indépendamment. Dans les régions spéciales, des pseudo-strates sont constituées, pour les besoins du calcul de l'estimation de la variance, soit de strates du plan, ou de combinaisons de strates, et chaque pseudo-strate est divisée en deux à quatre composantes (pseudo-U.P.É.). Encore une fois, les composantes sont supposées avoir été choisies indépendamment à l'intérieur de chaque strate. Ainsi, pour l'échantillon au complet, h représente une strate ou une sous-unité et i, une U.P.É. ou pseudo-U.P.É. choisie à l'intérieur de la strate ou de la sous-unité.

Soit $\hat{X}_A$, le nombre estimatif de personnes qui possèdent la caractéristique X de la population active, et $\hat{P}_A$, la population globale P dans la région A. $\hat{X}_A$ et $\hat{P}_A$ sont la somme des coefficients de pondération finals de tous les dossiers des personnes échantillonnées dans la région A, qui possèdent la caractéristique X et la population globale P respectivement. Dans chaque strate ou sous-unité h, $\hat{X}_{h1}, \ldots, \hat{X}_{hn_h}$ sont censés représenter n_h estimations indépendantes de $(\frac{1}{n_h})$e de la strate globale X_h, c'est-à-dire de $\frac{X_h}{n_h}$. Pour les besoins de l'estimation de la variance, on attribue une valeur D à chaque U.P.É. Pour estimer la variabilité d'échantillonnage de $\hat{X}_p$ qui est l'estimateur du nombre de personnes dans une province p possédant la caractéristique X, les valeurs D représentent:

$$D_{phi}(\hat{X}_p) = \hat{X}_{phi} - \sum_{a=1}^{20} \frac{\hat{X}_{pa}}{\hat{P}_{pa}} \hat{P}_{phia} \qquad (12.1)$$

pour chaque U.P.É. choisie dans cette province. La deuxième partie de l'équation ci-dessus correspond à la méthode de l'estimation par quotient qui est appliquée à chaque strate âge-sexe établie a posteriori. L'estimation de la variabilité de l'échantillonnage de $\hat{X}_p$ prend donc la forme suivante:

$$\hat{Var}(\hat{X}_p) = \sum_{h \epsilon p} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} [D_{phi}(\hat{X}_p) - \frac{1}{n_h} \sum_{i=1}^{n_h} D_{phi}(\hat{X}_p)]^2 \qquad (12.2)$$

où $\sum_{h \epsilon p}$ représente la somme de toutes les strates et sous-unités dans la province p et $\sum_{i=1}^{n_h}$ représente la somme des n_h U.P.É. choisies (ou pseudo-U.P.É.) dans la strate (ou sous-unité) h.

L'estimation de la variance du total estimé d'une caractéristique à l'échelle nationale consiste à faire la somme des estimations de la variance correspondantes à l'échelon provincial, puisque les estimations sont calculées indépendamment dans chaque province.

12.3 Exemple du calcul de la variance

Nous allons étudier un cas hypothétique pour expliquer ce genre de calcul. Prenons une province p qui comprend deux strates h = 1, 2; on choisit deux U.P.É. dans la

from stratum 1 ($n_1 = 2$) and three PSUs selected from stratum 2 ($n_2 = 3$). Suppose that post-stratification is carried out for two age-sex classifications (a = 1, 2) and a ratio estimation procedure is applied independently within these post strata. In accordance with the weighting of individual records as previously described, estimated totals for a particular area and characteristic are obtained by summing the final weights on records for all individuals falling within the area and possessing the characteristic. Assume that the following table presents the final weighted estimates of population and characteristic totals by age-sex classification for each PSU weighted up to one half of the stratum totals in stratum 1 and one third of the stratum totals in stratum 2.

strate 1 ($n_1 = 2$) et trois U.P.E. dans la strate 2 ($n_2 = 3$). Supposons que l'on effectue une stratification a posteriori pour les deux classements âge-sexe (a = 1, 2) et qu'on applique la méthode de l'estimation par quotient à chacune des strates établies a posteriori. Selon la pondération des dossiers individuels expliquée précédemment, on évalue les totaux d'une région particulière et d'une caractéristique en faisant la somme des coefficients de pondération finals indiqués dans les dossiers pour toutes les personnes demeurant dans cette région et possédant cette caractéristique. Supposons que le tableau ci-dessous présente les estimations finales pondérées des totaux de la population et de la caractéristique, suivant le classement âge-sexe, pour chaque U.P.E. pondérée à une demie des totaux de la strate 1 et à un tiers des totaux de la strate 2.

PSU identification Désignation de l'U.P.É.			Age-sex gp 1 (a=1) Âge-sexe, gr. 1		Age-sex gp 2 (a=2) Âge-sexe, gr. 2		Total	
Prov.	Str.	PSU U.P.É.	Pop.	Charact. Caract.	Pop.	Charact. Caract.	Pop.	Charact. Caract.
Prov.	Str.	PSU U.P.É.	$\hat{P}_{phi1}$	$\hat{X}_{phi1}$	$\hat{P}_{phi2}$	$\hat{X}_{phi2}$	$\hat{P}_{phi}$	$\hat{X}_{phi}$
P	1	1	500	100	400	110	900	210
P	1	2	625	150	510	140	1135	290
P	2	1	410	90	375	75	785	165
P	2	2	400	110	410	100	810	210
P	2	3	425	100	450	150	875	250

P(entire province) P(province entière) $\hat{P}_{p1} = 2360$ $\hat{X}_{p1} = 550$ $\hat{P}_{p2} = 2145$ $\hat{X}_{p2} = 575$ $\hat{P}_p = 4505$ $\hat{X}_p = 1125$

The estimate of the characteristic total for province P is:

$$\hat{X}_p = \sum_{h=1}^{2} \sum_{i=1}^{n_h} \hat{X}_{phi} = 210 + 290 + 165 + 210 + 250 = 1125$$

The estimator of the variance of $\hat{X}_p$ denoted by $\hat{Var}(\hat{X}_p)$ can be expressed as:

$$\hat{Var}(\hat{X}_p) = \sum_{h \in p} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} [\hat{X}_{phi} - \sum_{a=1}^{2} \frac{\hat{X}_{pa}}{\hat{P}_{pa}} \hat{P}_{phia} - \frac{1}{n_h} \sum_{i=1}^{n_h} (\hat{X}_{phi} - \sum_{a} \frac{\hat{X}_{pa}}{\hat{P}_{pa}} \hat{P}_{phia})]^2$$

using the figures given in the table above $\hat{Var}(\hat{X}_p)$ is 3337.109.

Pour estimer le total d'une caractéristique d'une province P, il faut utiliser la formule suivante:

$$\hat{X}_p = \sum_{h=1}^{2} \sum_{i=1}^{n_h} \hat{X}_{phi} = 210 + 290 + 165 + 210 + 250 = 1125$$

L'estimateur de la variance de $\hat{X}_p$ représenté par $\hat{Var}(\hat{X}_p)$ peut prendre la forme suivante:

$$\hat{Var}(\hat{X}_p) = \sum_{h \in p} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} [\hat{X}_{phi} - \sum_{a=1}^{2} \frac{\hat{X}_{pa}}{\hat{P}_{pa}} \hat{P}_{phia} - \frac{1}{n_h} \sum_{i=1}^{n_h} (\hat{X}_{phi} - \sum_{a} \frac{\hat{X}_{pa}}{\hat{P}_{pa}} \hat{P}_{phia})]^2$$

à partir des chiffres figurant dans le tableau ci-dessus, $\hat{Var}(\hat{X}_p)$ est 3337.109.

The percent standard deviation of $\hat{X}_p$ (or coefficient of variation as it is usually called) is then given by $100\ [\hat{Var}\ (\hat{X}_p)]^{1/2}/\hat{X}_p = 5.135\%$.

12.4 Presentation of Errors

The press release and the detailed bulletin (Catalogue 71-001) are the two main monthly publications from the LFS. The sampling errors in these publications are presented in the form of lettered symbols for the majority of characteristics with respect to monthly estimates or rates. The lettered symbols refer to ranges of percentage coefficient of variation that are applicable to the characteristics on the basis of average coefficient of variation during the preceding calendar year. These symbols are as follows:

A(0.0 - 0.5%), B(0.6 - 1%), C(1.1 - 2.5%), D(2.6 - 5%), E(5.1 - 10%), F(10.1 - 16.5%), G(16.6 - 25%), H(25.1 - 33.3%), J(33.4 +)

and only data with categories A to F are generally released without any restrictions.

In addition to these symbols providing the magnitude of coefficient of variation, an explanatory note on the "Reliability of Data" is provided describing sampling error, non-sampling error and other factors on the limitations of the data.

L'écart-type en pourcentage de $\hat{X}_p$ (communément appelé le coefficient de variation) se calcule de la façon suivante: $100\ [\hat{Var}\ (\hat{X}_p)]^{1/2}/\hat{X}_p = 5.135\%$.

12.4 Présentation des erreurs

Le communiqué de presse et le bulletin détaillé (nº 71-001 au catalogue) sont les deux principales publications mensuelles tirées de l'E.P.A. Les erreurs d'échantillonnage sont présentées dans ces publications sous forme d'indicateurs alphabétiques pour la majorité des caractéristiques par rapport aux estimations ou proportions mensuelles. Ces indicateurs correspondent aux intervalles des coefficients de variation en pourcentage, qui s'appliquent aux caractéristiques en se basant sur le coefficient de variation moyen obtenu durant l'année civile précédente. Ces indicateurs sont les suivants:

A(0.0 - 0.5 %), B(0.6 - 1 %), C(1.1 - 2.5 %), D(2.6 - 5 %), E(5.1 - 10 %), F(10.1 - 16.5 %), G(16.6 - 25 %), H(25.1 - 33.3 %), J (33.4 +)

seules les données dont les symboles sont dans les catégories A à F sont publiées sans aucune mise en garde.

Outre ces symboles qui révèlent l'importance du coefficient de variation, il y a une note explicative concernant la "Fiabilité des données" qui explique l'erreur d'échantillonnage, l'erreur non imputable à l'échantillonnage et les autres facteurs qui entraînent des mises en garde contre les données.

ASSESSMENT AND CONTROL OF NON-SAMPLING ERRORS

13.1 Introduction

While discussing estimation and variance estimation in Chapters 11 and 12 we assumed that the "true" value of each unit in the population (or sample) can be obtained and processed without any error. However, this is not usually the case in practice. Apart from sampling error, discussed in Chapter 12, there are other types of errors in the data collected from surveys. Such errors are termed **non-sampling errors** and are present in sample surveys as well as in censuses. Thus in the case of censuses, non-sampling errors, and in the case of sample surveys, both sampling and non-sampling errors require control at the level at which their presence does not vitiate the use and interpretation of final results. The following sections describe briefly the programs used in LFS for the assessment and control of different components of non-sampling errors.

The non-sampling error (like the sampling error) may be divided into two main components, namely, **non-sampling variance** and **non-sampling bias**. Thus the **total error** consists of sampling variance and bias, non-sampling variance and bias and some interactions between the sources of non-sampling errors and the sampling units. The non-sampling variance has a tendency to be in either direction and thus the net effect of this component may be negligible over a large number of observations or for large areas; but for small areas or rare or sensitive characteristics it can be large, The non-sampling bias component, on the other hand, has a tendency to occur in one direction more than the other. For example, it may be the result of training or attitude of interviewer or it may be a fault in questionnaire design or the method of imputations for non-response in that it may tend to invite error in one direction more than the other.

Contributions to non-sampling variance or bias may arise through one or more of the following main sources of error, namely, coverage error, response error, non-response error and/or processing error. These errors are briefly discussed below.

ÉVALUATION ET CORRECTION DES ERREURS NON DUES À L'ÉCHANTILLONNAGE

13.1 Introduction

Lors de l'étude de l'estimation et de l'estimation de la variance aux chapitres 11 et 12, nous avons supposé que la valeur "réelle" de chaque unité dans la population (ou l'échantillon) peut être établie et exploitée sans erreur. Cependant, il en va tout autrement dans la pratique. Outre l'erreur d'échantillonnage expliquée au chapitre 12, on trouve également d'autres types d'erreur dans les données recueillies dans les enquêtes. Ces erreurs sont des erreurs non dues à l'échantillonnage que l'on trouve aussi bien dans les enquêtes que dans les recensements. Ainsi, les erreurs qui n'ont pas trait à l'échantillonnage commises dans les recensements et les erreurs d'échantillonnage et non imputables à l'échantillonnage trouvées dans les enquêtes doivent être corrigées avant qu'elles faussent l'interprétation et empêchent l'utilisation des résultats finals. Les sections suivantes décrivent brièvement les méthodes utilisées dans l'E.P.A. pour évaluer et contrôler les diverses composantes des erreurs non attribuables à l'échantillonnage.

L'erreur non due à l'échantillonnage, tout comme l'erreur d'échantillonnage, peut se décomposer en deux éléments principaux, à savoir la variance **non imputable à l'échantillonnage** et le **biais non imputable à l'échantillonnage**. Par conséquent, **l'erreur totale** comprend la variance et le biais d'échantillonnage, la variance et le biais non-attribuables à l'échantillon et un certain nombre d'interactions entre les causes des erreurs non-dues à l'échantillonnage et les unités de sondage. La variance non imputable à l'échantillonnage a tendance à se produire dans un sens comme dans l'autre; ainsi, les conséquences nettes de cette composante peuvent être négligeables dans le cas d'un grand nombre d'observations ou dans de grandes régions. Par contre, dans les petites régions ou lorsque les caractéristiques sont rares ou de nature à éveiller la susceptibilité, cette composante peut avoir de graves répercussions sur l'erreur totale. D'un autre côté, le biais d'observation a tendance à se produire plus souvent dans une direction que dans l'autre. Par exemple, il peut être le résultat de la formation ou de l'attitude de l'interviewer; il peut être causé par une mauvaise conception du questionnaire ou par la méthode de calcul utilisée dans les cas de non-réponse, qui favorise des erreurs dans un sens plutôt que dans l'autre.

La variance ou le biais non dus à l'échantillonnage peuvent provenir d'une ou de plusieurs des principales sources d'erreurs suivantes notamment l'erreur de couverture, l'erreur de réponse, l'erreur de non-réponse et l'erreur de dépouillement. Ces erreurs sont expliquées brièvement ci-dessous.

Coverage error – This error results from inadequate coverage of the population of interest. Coverage error may occur at several stages of the survey for example, while preparing the frame, defining the sampling units and assigning the selection probabilities for sampling purposes or at the data collection and processing stages.

Response error – This error may be due to questionnaire design, the characteristics of the question, attitude of the respondent, the interviewing approach or other general conditions under which the survey is conducted. These errors may occur when data is requested, provided, received and recorded. Questions may be misinterpreted by the respondent, he or she may not know, may not remember or may purposely want to distort the correct answer. Different interviewers may have different tendencies in explaining questions or interpreting response, etc.

Non-response error – Although the intention is to obtain response from every individual included in the survey, this is never fully achieved. Some respondents may refuse to answer questions, others may not be at home at the time of call, etc. Data for the non-responding units have to be imputed using the data from responding units or some earlier data on the non-responding units, if available. The extent of bias due to imputation is usually unknown and is very much dependent on the characteristic differences between the responding group and the non-responding group of individuals in the survey. This error increases with increase in the non-response rate, attempts are therefore made to obtain as high a response rate as possible for any survey.

Processing error – These are the errors that may occur at various stages of processing such as coding, data entry, verification, editing, weighting, tabulation, etc.

In the LFS, all attempts are made to minimize the non-sampling errors: units are defined in a most precise manner and the most up-to-date listing used; the questionnaires used in the survey are carefully designed to minimize different interpretations, interviewers are instructed to ask the questions as printed on the questionnaires; detailed acceptance testing is carried out for the different stages of editing and processing; efforts are made to reduce the non-response rate as well as respondent burden, etc. The following sections briefly describe two regular programs of the LFS which aim at controlling and assessing the non-sampling errors in the survey data.

Erreur de couverture – Cette erreur provient d'une mauvaise couverture de la population visée. Ce genre d'erreur peut survenir à n'importe quelle étape de l'enquête, par exemple au moment de la conception du cadre de l'enquête, lors de la définition des unités de sondage et du calcul des probabilités de sélection en vue de l'échantillonnage ou au moment de la collecte et du dépouillement des données.

Erreur de réponse – Ce type d'erreur peut être dû à la nature du questionnaire, aux caractéristiques de la question posée, à l'attitude de l'enquêté, à la méthode d'entrevue utilisée ou à d'autres conditions générales sous lesquelles l'enquête est effectuée. Ces erreurs peuvent se produire lorsque les données sont demandées, fournies, reçues et enregistrées. Il peut arriver que l'enquêté comprenne mal la question, qu'il ne connaisse pas la réponse on ne s'en souvienne pas, ou encore, qu'il veuille délibérément donner une réponse inexacte. Il y a également le fait que les interviewers peuvent expliquer les questions ou interpréter les réponses de différentes façons, etc.

Erreur de non-réponse – Bien que le but premier d'une enquête soit d'obtenir de chaque personne interviewée une réponse aux questions, cet objectif n'est pas toujours atteint. Il peut arriver que certains enquêtés refusent de répondre aux questions, que d'autres soient absents de leur domicile au moment de l'enquête, etc. Les données concernant les unités de non-réponse doivent être calculées à partir des données recueillies auprès des autres unités ou de données antérieures, si elles sont disponibles, établies pour les unités de non-réponse. L'importance du biais dû à l'imputation n'est habituellement pas connue et dépend beaucoup des différences de caractéristique entre le groupe de personnes qui ont répondu et celui des personnes qui n'ont pas répondu à l'enquête, par exemple. Ce type d'erreur augmente proportionnellement à l'augmentation du taux de non-réponse; on tente donc d'obtenir dans une enquête le taux de réponse le plus élevé possible.

Erreur de dépouillement – Il s'agit des erreurs qui peuvent survenir à n'importe quelle étape du dépouillement, par exemple au codage, à l'entrée des données, à la vérification, au contrôle, à la pondération, à la mise en tableaux, etc.

Dans l'E.P.A., nous tentons par tous les moyens de réduire au minimum les erreurs non attribuables à l'échantillonnage. Ainsi, les unités sont définies le plus exactement possible et l'on utilise uniquement les listes les plus récentes; les questionnaires sont soigneusement conçus pour minimiser la possibilité d'obtenir différentes interprétations des questions; les interviewers doivent poser les questions exactement comme elles figurent sur le questionnaire; des tests d'acceptabilité très poussés sont effectués aux diverses étapes du contrôle et du dépouillement; on tente de réduire le taux de non-réponse tout en n'augmentant pas le fardeau de l'enquête, etc. Les sections qui suivent donnent un bref aperçu des deux programmes de l'E.P.A. qui visent à vérifier et à évaluer les erreurs non dues à l'échantillonnage dans les données de l'enquête.

13.2 Observation Program

Control of errors is of particular concern because errors arising from various sources, especially those arising in the field collection of data, may be of much larger magnitude than commonly recognized. There are several programs for evaluating and controlling the work of interviewers in the LFS. A detailed description of the type of errors that may be introduced through the interviewers, the underlying causes of such errors, and programs for measurement and control of such errors in the context of the Canadian LFS is given in [12].

Uniformity in the interviewing standards can be achieved through a careful selection of the interviewer, a thorough training program as well as a regular program of control and retraining. Interviewers are hired on the basis of a written examination and a personal interview conducted by the supervisory personnel. After they are hired they are subjected to a thorough training program. During their first days of conducting interviews they are accompanied by supervisory personnel.

Observation is an integral part of the improvement process in the interviewer's performance. Observation occurs when a supervisory staff member accompanies an interviewer in the field for the purpose of evaluating and improving his or her performance. It provides an excellent opportunity for the observer to assess the interviewer's attitude towards his/her job, understanding of the concepts and procedures, and his/her interviewing technique in general. The scheduling of observations and selection of those to be observed is the responsibility of the supervisor. All interviewers are to be observed every six months in the course of systematic observations. In addition, those interviewers whose performance was below standard and those scheduled for first and second month induction are "observed" by the supervisory staff.

An **observation report** is completed by the observer for each day of observation. Depending upon the recommendation, an interviewer may be asked to go through a retraining program if found necessary. Detailed instructions about the scheduling of observation, the role of observer, the handling of problem areas etc. are given in the "Observation Instruction Manual".

13.3 Re-interview Program

One recent trend in the survey taking has been the increased attention to measuring the response error in the data collected in sample surveys. The new re-interview program in the LFS has accordingly been designed [18] to measure various components of the response error.

13.2 Programme d'observation

La vérification des erreurs est particulièrement importante parce que les erreurs provenant de diverses sources surtout les erreurs commises pendant la collecte des données sur le terrain, peuvent avoir des conséquences beaucoup plus graves qu'on ne le croit. Il existe plusieurs programmes d'évaluation et de vérification du rendement des interviewers affectés à l'E.P.A. On trouve en [12] des explications sur les divers types d'erreurs commises par les interviewers, sur les causes principales de ces erreurs et sur les programmes d'évaluation et de vérification des erreurs élaborés pour l'E.P.A.

Il est possible d'uniformiser les méthodes d'entrevue en sélectionnant avec soin les interviewers, en leur faisant suivre un cours de formation et en mettant sur pied des programmes de vérification et des cours complémentaires de formation. Les interviewers sont embauchés en fonction des résultats d'un examen écrit et d'une entrevue menée par le personnel cadre. Une fois embauchées, ces personnes doivent suivre un cours complet de formation. Pour leurs premières entrevues, les interviewers sont accompagnés d'un suiveillant.

L'observation fait partie intégrante du programme de perfectionnement de l'interviewer. L'observation s'effectue lorsqu'un membre du personnel cadre accompagne l'interviewer sur le terrain dans le but d'évaluer et d'améliorer son travail. Cela donne à l'observateur une excellente occasion d'évaluer le composrtement de l'interviewer vis-à-vis de ses tâches, de déterminer s'il a bien assimilé les concepts et compris les méthodes de travail et, de façon générale, d'évaluer la technique d'entrevue adoptée par l'interviewer. Les périodes d'observation et le choix des interviewers visés incombent au superviseur. On soumet tous les interviewers à cette observation au moins une fois tous les six mois. En outre, les interviewers dont le rendement ne satisfait pas aux normes et ceux qui en sont à leur premier et deuxième mois de formation sont "observés" par le personnel d'encadrement.

L'observateur remplit un **rapport d'observation** tous les jours de la période d'observation. Suivant les recommandations faites, un interviewer peut être appelé à suivre un cours complémentaire de formation, si cela est jugé nécessaire. Des directives détaillées concernant l'établissement du calendrier des périodes d'observation, le rôle de l'observateur, la solution de problèmes, etc. sont fournies dans le "Guide d'instruction pour l'observation".

13.3 Programme de réinterview

Il a été remarqué au cours de récentes enquêtes que l'on accorde une attention de plus en plus marquée à l'évaluation de l'erreur de réponse dans les données recueillies dans les sondages. Le nouveau programme de réinterview de l'E.P.A. a donc pour objet [18] d'évaluer les divers éléments de l'erreur de réponse.

The re-interview program rests upon the following design. Each month, the re-interview sample is randomly split into two parts. For both parts of the re-interview sample the re-interviewer, who is a senior interviewer responsible for the quality of the work performed by a group of interviewers, carries out the interview. In the first part, the re-interviews are conducted as independently as possible form the original interviews in order to achieve a repetition of the original survey under similar conditions. In the second part, having completed all interviews within a household, the re-interviewer compares the answers with those appearing on the copies of the original schedules, and tactfully discusses differences with the respondent. He will note the correct responses and code a reason to explain the discrepancies on the reconciliation form.

The accumulation of data from both parts of the sample will enable us to study general aspects of response errors. Thus, simple response variance may be estimated from the first part of the sample, under the assumption of a "repetition of a survey", and response bias may be estimated from the second part of the sample if it is assumed that the re-interviewer provides the "true" answers under the survey conditions. Also, it will be possible to conduct a general study of the variation of interviewers' performance, by post-stratifying according to interviewers' experience, duration of training, working area, season of the year and so on. Furthermore, the field control aspect is not neglected. The re-interviewers will have the responsibility of determining specific instances where the information originally collected by the interviewer is incomplete, wrong or non-existent with respect to the coverage of households, and persons within.

The periodical analysis of the data produced by the program should permit detection of sources of errors and should lead to a continuous improvement of the overall methodology, since it will have a feedback with respect to various components of the survey design such as questionnaire design, interview technique, interviewer training and so on.

On the other hand, the correlated response variance cannot be estimated from a simple re-interview program. It requires a special design based on interpenetrating samples. At this stage, it is planned to conduct a study periodically, in order to provide estimates of "correlated response variance". The variety and extent of such programs are, quite often, determined by the budgetary considerations.

13.4 Evaluation and Developmental Projects

The need for regular programs of evaluation and control for large scale continuous surveys like the LFS can hardly be over emphasized. The LFS incorporates

Ce programme a été élaboré selon le plan suivant: chaque mois, l'échantillon de réinterview est divisé au hasard en deux parties. Pour chacune de ces parties, la personne chargée des réinterviews, qui est un interviewer d'expérience responsable de la qualité du travail exécuté par un groupe d'interviewers, effectue les interviews. Dans la première partie, les réinterviews sont menées de la façon la plus indépendante possible des interviews originales afin d'effectuer une répétition de l'enquête originale dans les mêmes conditions. Dans la deuxième partie, une fois toutes les interviews effectuées à l'intérieur d'un ménage, la personne chargée des réinterviews compare les réponses obtenues avec celles qui figurent sur les dossiers précédents et discute des différences avec l'enquêté, puis elle inscrit les bonnes réponses et code les raisons qui justifient les divergences, sur la formule de comparaison.

L'accumulation de données provenant des deux parties de l'échantillon nous permet d'étudier les aspects généraux des erreurs de réponse. Ainsi, on peut évaluer la variance de réponse simple à partir de la première partie de l'échantillon, suivant l'hypothèse qu'il s'agit d'une "répétition de l'enquête", et le biais de la réponse à partir de la deuxième partie de l'échantillon, si l'on tient pour acquis que le préposé aux réinterviews a indiqué les "vraies" réponses obtenues dans les mêmes conditions d'enquête. En plus, ce programme permet d'effectuer une étude générale du rendement des interviewers en faisant une stratification a posteriori, selon l'expérience de l'interviewer, la durée de la formation, la zone de travail, la saison de l'année et ainsi de suite. En outre, il ne faut pas négliger l'aspect de vérification sur le terrain. Les personnes chargées des réinterviews doivent découvrir les cas particuliers où les renseignements recueillis au début par l'interviewer sont incomplets, inexacts ou faux concernant les ménages et leurs membres.

L'analyse périodique des données établies par le programme permet de découvrir les causes d'erreurs et d'améliorer continuellement l'ensemble des méthodes utilisées, puisqu'il aura des incidences sur divers éléments de la conception d'une enquête tels que la composition du questionnaire, les techniques d'entrevue, la formation des interviewers et d'autres encore.

Par contre, la variance de réponse corrélative ne peut être évaluée à partir d'un simple programme de réinterviews. Il faut mettre au point un plan particulier fondé sur les échantillons superposés. À ce stade-ci de l'enquête, il est prévu de mener une étude périodique afin de fournir des estimations de la "variance de réponse corrélative". La diversité et l'importance de ces programmes sont très souvent déterminées par les dépenses budgétaires prévues.

13.4 Projets d'évaluation et d'amélioration

On ne peut trop insister sur la nécessité de créer des programmes courants d'évaluation et de vérification des enquêtes continues à grande échelle comme l'E.P.A. Cette

programs to monitor the quality of data collected on a monthly basis and in addition some ongoing developmental programs with a view to improving the quality of survey.

These programs may be broadly classified into two groups, each serving a somewhat different purpose. One group of programs, which are related to quality assurance, such as observation, re-interview, error analysis etc., is intended to monitor on a monthly basis the quality of data collected from the LFS. Also, programs such as interviewers' performance standards can be used as a tool for improving data collection operations over a period of time. The other group of programs which includes variance-covariance analysis, stratification index, sample up-dating variance components etc. aim at improving the quality of survey design.

enquête suppose des programmes visant à contrôler la qualité des données recueillies mensuellement et, en outre, des programmes actuels d'amélioration dont l'objet est d'élever la qualité de l'enquête.

Ces programmes peuvent être classés en deux grandes catégories, dont les buts sont bien différents. Une de ces catégories qui vise à assurer la qualité des données par l'observation, les réinterviews, l'analyse des erreurs, etc., sert à contrôler mensuellement la qualité des données recueillies dans l'E.P.A. Des programmes tels que l'étude des normes de rendement des interviewers peuvent également contribuer à améliorer les méthodes de collecte pendant une certaine période. L'autre catégorie de programmes, qui comprend l'analyse de la variance et de la covariance, l'indice de stratification, la mise à jour de l'échantillon, les composantes de la variance, etc., vise à améliorer la qualité du plan de l'enquête.

CHAPTER – 14 – CHAPITRE

SAMPLE ALLOCATION

The final major step in the revision of the survey was to increase the sample size from approximately 35,000 - 55,000 households. This was done to enable provision of more detailed and more accurate estimates, especially for provincial data. The redesigned sample was put in the field towards the end of 1974 and it ran in parallel with the old LFS sample during the whole of 1975. The old LFS was phased out in December 1975, and from January 1976 the regular estimates have been derived from the redesigned sample. The expansion of the sample however, was planned for 1976 - 1977 (except in Prince Edward Island where the expansion took place in February 1976). This difference in timing for the implementation of the redesigned sample and the expanded sample, in general, restricts the choice of allocation of the total sample to various provinces and the choice of sample designs. However, the original design was made flexible in order to provide for a sample increase in an efficient manner without disturbing the original sample.

Various allocation strategies were investigated [13] and discussed with users. As a result the allocation of 35,000 households to different provinces, as used in the redesigned sample, remained unchanged and the additional sample (approximately 20,000 households) is allocated to various provinces following certain criteria. The basis for allocation of the original sample is given in section 14.1 and for the expanded sample in section 14.2.

14.1 Allocation of the Original Sample

The allocation used in the redesigned sample was the same as that of the old LFS. This allocation, obtained at the time of the 1961 redesign, was based on (i) the minimum variance allocation of unemployed estimated at the national level, and (ii) equal coefficient of variation of unemployed for each province. A compromise allocation between these two extremes provided reasonably good estimates of unemployed for each province except Prince Edward Island, and provided a high degree of reliability for national estimates of unemployed, permitting additional cross-classification of data such as unemployed by age-sex at the national level. The following Table 14.1, gives the allocation of the redesigned sample as for March 1977 for each province by SRUs and NSRUs along with the corresponding coefficient of variation (CV) for unemployed. It will be noticed that in addition to Prince Edward Island, the CVs for Saskatchewan and Manitoba are also more than 10%, demanding special consideration in the allocation of the expanded sample.

RÉPARTITION DE L'ÉCHANTILLON

Le dernier point important étudié lors de la réorganisation de l'enquête a été l'éclargissement de l'échantillon qui est passé de 35,000 - 55,000 ménages. Cette modification a permis d'obtenir un plus grand nombre d'estimations plus détaillées et plus exactes, surtout dans le domaine des données concernant les provinces. Le nouveau plan de sondage a été mis à l'épreuve sur le terrain vers la fin de 1974 et il a été appliqué simultanément avec l'ancien échantillon de l'E.P.A., pendant toute l'année 1975. l'ancien échantillon a été abandonné progressivement et mis de côté en décembre 1975; les estimations courantes tirées du nouvel échantillon ont été calculées à partir du mois de janvier 1976. L'élargissement de l'échantillon toutefois a été prévu pour l'année 1976 - 1977 (sauf dans l'Île-du-Prince-Édouard où l'agrandissement a été effectué en février 1976). Ce délai entre la réorganisation de l'échantillon et l'élargissement de l'échantillon limite restreint, de façon générale, la répartition de l'ensemble de l'échantillon entre diverses provinces et le choix de plans de sondage. Cependant, le plan initial était assez souple pour que l'élargissement de l'échantillon se fasse efficacement sans que l'échantillon initial soit perturbé.

Un certain nombre de méthodes de répartition ont été étudiées [13] et analysées avec les utilisateurs; en conséquence, la répartition des 35,000 ménages entre diverses provinces, dans le nouvel échantillon, n'est pas modifiée et l'échantillon supplémentaire (environ 20,000 ménages) est réparti suivant certains critères précis. Les principes de la répartition de l'échantillon initial sont expliqués au paragraphe 14.1 et ceux concernant l'échantillon élargi, au paragraphe 14.2.

14.1 Répartition de l'échantillon initial

La répartition adoptée pour le nouvel échantillon est la même que celle qui a été appliquée à l'ancien échantillon de l'E.P.A. Cette répartition, établie au moment de la refonte en 1961, était fondée (i) sur la répartition de la variance minimale des chômeurs calculée à l'échelle nationale, et (ii) sur le coefficient de variation égal des chômeurs dans chaque province. Un compromis entre ces deux extrêmes a permis d'obtenir des estimations assez justes du nombre de chômeurs dans chaque province, sauf dans l'Île-du-Prince-Édouard, et des estimations nationales du taux de chômage dont le degré de fiabilité était très élevé, qui ont permis de faire des classements recoupés de données, par exemple chômeurs selon l'âge et le sexe à l'échelle nationale. Le tableau ci-dessous (14.1) illustre bien la répartition de l'échantillon remanié au mois de mars 1977, pour chaque province, par U.A.R. et U.N.A.R., et indique le coefficient de variation (C.V.) correspondant des chômeurs. Il convient de souligner qu'outre l'Île-du-Prince-Édouard, les C.V. du Manitoba et de la Saskatchewan sont supérieurs à 10 %, ce qui signifie qu'il faudra leur porter une attention particulière lors de la répartition de l'échantillon élargi.

TABLE 14.1 Allocation of the Original Sample

TABLEAU 14.1 Répartition de l'échantillon initial

Province	Number of households by — Nombre de ménages par		Total sample size — Taille globale de l'échantillon	CV for unemployed — C.V. des chômeurs
	NSRUs — U.N.A.R.	SRUs — U.A.R.		
Newfoundland – Terre-Neuve	784	1,031	1,815	8.64
Prince Edward Island – Île-du-Prince-Édouard	363	116	479	15.54
Nova Scotia – Nouvelle-Écosse	1,945	1,317	3,262	7.79
New Brunswick – Nouveau-Brunswick	1,633	906	2,539	9.12
Québec	2,495	4,813	7,308	4.22
Ontario	2,719	6,825	9,544	4.76
Manitoba	769	1,091	1,860	11.38
Saskatchewan	1,262	735	1,997	13.56
Alberta	1,529	1,875	3,404	8.91
British Columbia – Colombie-Britannique	1,403	2,794	4,197	6.36
Canada	14,902	21,503	36,405	2.36

The following points should be noted in relation to the figures in this table: (i) the reference period for this table is March 1977; (ii) the actual sample sizes for a particular month would be slightly different owing to natural growth of the sample; (iii) CVs would also differ from month to month; (iv) although the sample for Prince Edward Island was expanded in February 1976, this table relates to the original sample only; (v) the figures should be treated as approximations only as they are based on certain assumptions as outlined in the following section.

Il faut tenir compte des points suivants concernant les chiffres figurant dans ce tableau: (i) la période de référence est le mois de mars 1977; (ii) la taille réelle de l'échantillon établie pour un mois particulier peut varier légèrement en raison de la croissance naturelle de l'échantillon; (iii) les C.V. peuvent également varier d'un mois à un autre; (iv) même si l'échantillon de l'Île-du-Prince-Édouard a été élargi en février 1976, ce tableau correspond à l'échantillon initial seulement; (v) les chiffres doivent être considérés uniquement comme des approximations puisqu'ils sont fondés sur certaines hypothèses, comme on l'explique dans la section suivante.

14.2 Allocation of the Expanded Sample

As a result of consultation with users, the proportion of unemployed in each province was considered as the main characteristic for the examination of various alternatives [17]. The following parameters have been used in obtaining the allocations and the corresponding coefficients of variation (CVs):

(i) The proportion of persons unemployed in a type of area within a province was based on averages of unemployed and total population estimates over the five-year period January 1970 to December 1974.

(ii) Binomial factors were approximated by the average of binomial factors over the same five-year period using the old LFS.

14.2 Répartition de l'échantillon élargi

Par suite des consultations menées avec les utilisateurs, la proportion de chômeurs dans chaque province a été considérée comme la caractéristique principale requise pour l'étude des diverses solutions de rechange [17]. On a utilisé les paramètres suivants pour établir les répartitions et les coefficients de variation appropriés (C.V.):

(i) La proportion de personnes en chômage dans un type de région à l'intérieur d'une province a été calculée en fonction des moyennes de chômeurs et des estimations globales de la population couvrant une période de cinq années, c'est-à-dire de janvier 1970 à décembre 1974.

(ii) Les facteurs binomiaux ont été évalués d'après la moyenne des facteurs binomiaux établie pour la même période de cinq années, à partir de l'E.P.A. précédente.

(iii) The effective time period was March 1977. Population estimates were extrapolated to determine the population by type of area within each province for March 1977.

(iv) Average household sizes by type of area within provinces were determined by projecting the average household size for the March surveys of 1970-1975, to obtain an estimate for the March 1977 survey.

(iii) La période de référence est le mois de mars 1977. Les estimations de la population ont été extrapolées afin de déterminer la population selon le type de région dans chaque province, en mars 1977.

(iv) La taille moyenne du ménage par type de région dans les provinces a été établie par une projection de la taille moyenne du ménage lors des enquêtes effectuées aux mois de mars de 1970-1975, afin d'obtenir une estimation en vue de l'enquête du mois de mars 1977.

TABLE 14.2. Allocation of the Expanded Sample[1]

TABLEAU 14.2. Répartition de l'échantillon élargi[1]

Province	Non-self-representing — Non autoreprésentatif			Self-representing — Autoreprésentatif			Total sample size — Taille globale de l'échantillon	S.R. — F.S.	CV for unemployed — C.V. des chômeurs
	Number of households — Nombre de ménages	S.R. — F.S.	Increase — Majoration	Number of households — Nombre de ménages	S.R. — F.S.	Increase — Majoration			
		%			%			%	
Newfoundland – Terre-Neuve	1,568	3.20	100	1,523	2.13[2] 1.06 1.60	100[3]	3,091	2.34	6.57
Prince Edward Island – Île-du-Prince-Édouard	1,089	4.80	200	347	3.19	200	1,436	4.23	8.94
Nova Scotia – Nouvelle-Écosse	2,917	2.40	50	1,317	1.06	0	4,234	1.73	7.01
New Brunswick – Nouveau-Brunswick	2,449	2.40	50	1,811	2.13	100	4,260	2.31	7.20
Québec	3,742	0.75	50	4,813	0.33	0	8,555	0.45	4.01
Ontario	4,078	0.75	50	6,825	0.33	0	10,903	0.42	4.60
Manitoba	2,307	2.40	200	2,182	1.11	100	4,489	1.52	7.86
Saskatchewan	3,785	2.40	200	2,204	1.77	200	5,989	2.09	7.62
Alberta	3,057	1.60	100	3,749	1.08	100	6,806	1.26	6.24
British Columbia – Colombie-Britannique	2,805	1.40	100	3,163	0.95[4] 0.48	100[3]	5,968	0.83	5.77
Canada	**27,797**	**1.40**	**85.2**	**27,934**	**0.56**	**29.9**	**55,731**	**0.80**	**2.18**

[1] Excluding special areas. – Sans les régions spéciales.

[2] 2.13 for St. John's; 1.06 for Cornerbrook, Wabanna-Belle Island and Windsor-Grand Falls; 1.60 for all other SRUs in Newfoundland. – 2.13 pour St-John's; 1.06 pour Cornerbrook, Îles Wabana-Belle et Windsor-Grand Falls; 1.60 pour toutes les autres U.A.R. de Terre-Neuve.

[3] 100% increase is only for the SRUs St. John's and Victoria, and no increase for other SRUs in these provinces. – La majoration de 100 % ne se retrouve que dans les U.A.R. de St.John's et de Victoria; il y a absence de majoration dans les autres U.A.R. de ces provinces.

[4] 0.95 for Victoria and 0.48 for all other SRUs in British Columbia. – 0.95 pour Victoria et 0.48 pour toutes les autres U.A.R. de la Colombie-Britannique.

Using these constraints, approximately 20,000 households were allocated to various provinces under the criterion of **equal CV** for estimates of unemployed for all provinces. The sample size so obtained for a particular province was then allocated between SRUs and NSRUs within a province taking account of the variance and cost per household by type of area. Some adjustments in the allocation were finally made to suit the design of the survey. Details about the allocation principle and comparison of alternatives are given in [13], [17]. The total sample size for a province is thus obtained as the sum of this allocation of the additional sample and the allocation as given in Table 14.1. The resultant total sample size, the sampling ratio (SR), and the increase as percent of the redesigned sample by SRUs and NSRUs, and corresponding expected CVs are given in the following Table 14.2.

In addition, there are about 500 households (at the national level) in the original sample belonging to the special areas. These are allocated to different provinces using the original sampling ratio for NSR areas. Due to operational problems, the increase in the sample size for special areas is not contemplated. The apartment sample is also increased, depending upon the increase in the corresponding SRU.

As mentioned earlier, the figures given in the table are only approximate because of certain projections and assumptions used in calculations.

Une fois ces conditions satisfaites, un nombre d'environ 20,000 ménages a été réparti entre les diverses provinces selon le principe du **C.V. égal** des estimations des chômeurs dans l'ensemble des provinces. Une fois la taille de l'échantillon établie pour une province particulière, elle était ensuite répartie entre les U.A.R. et les U.N.A.R. dans la province, en tenant compte de la variance et du coût par ménage selon le type de région. Il a fallu apporter certaines rectifications pour se conformer au plan de l'enquête. Les détails concernant le principe de la répartition et les comparaisons relatives aux solutions de rechange sont données en [13] et [17]. On calcule la taille globale de l'échantillon dans une province en additionnant la répartition de l'échantillon supplémentaire et la répartition qui figure au tableau 14.1. Le résultat de cette addition, la fraction de sondage (F.S.) et la majoration en pourcentage de l'échantillon remainié par U.A.R. et par U.N.A.R. ainsi que les C.V. prévus figurent au tableau 14.2 ci-dessous.

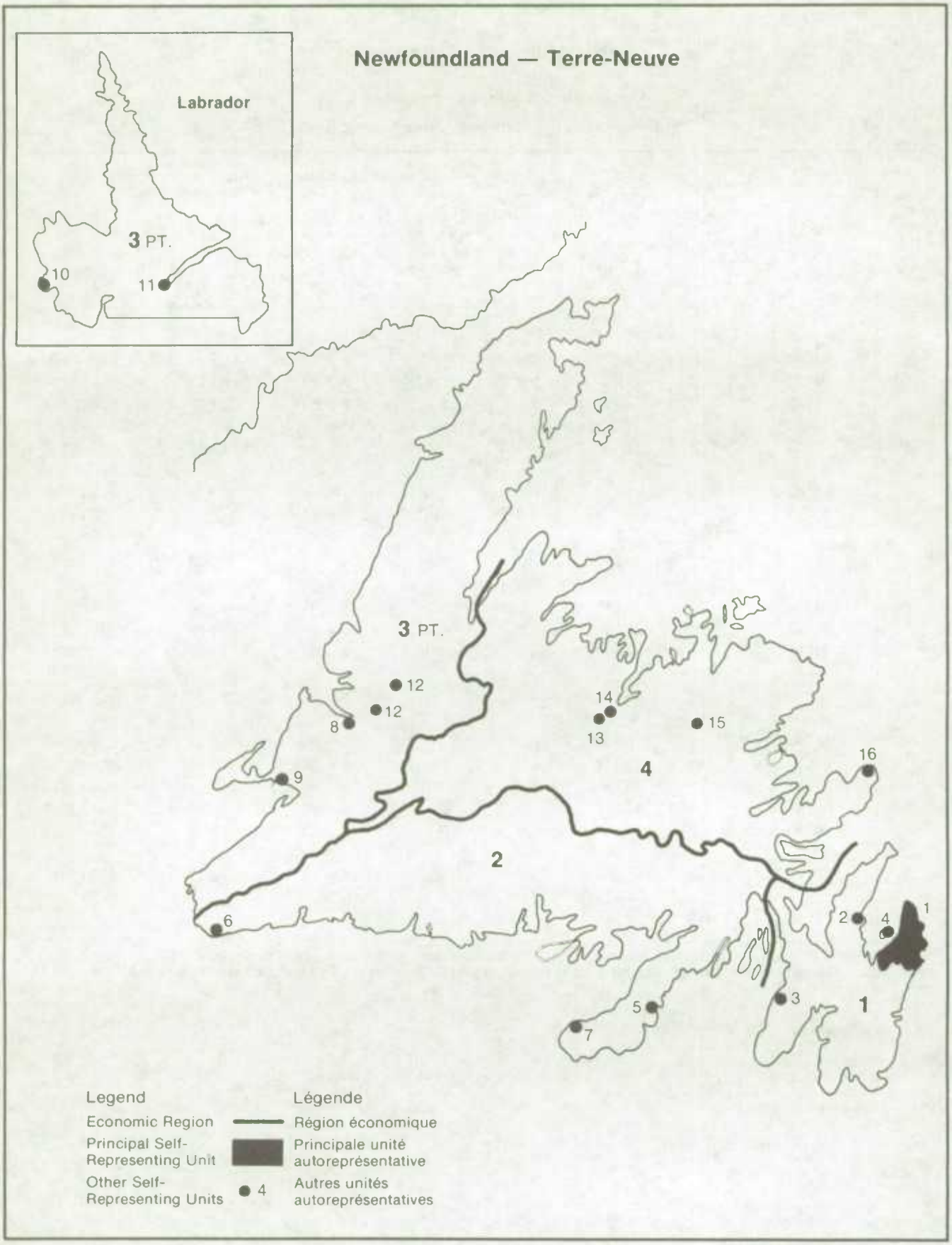
En outre, quelque 500 ménages (à l'échelle nationale) compris dans l'échantillon initial sont situés dans des régions spéciales. Ces ménages sont répartis entre les différentes provinces suivant la fraction de sondage initiale des régions N.A.R. À la suite de difficultés techniques, l'augmentation de la taille de l'échantillon dans les régions spéciales n'a pas été envisagée. L'échantillon des appartements a également été agrandi selon l'élargisement du l'U.A.R. correspondante.

Comme il a été mentionné précédemment, les chiffres fournis dans ce tableau ne sont que des approximations en raison des projections et des hypothèses utilisées pour effectuer les calculs.

APPENDICES

ANNEXES

Appendix 0.1 **Annexe 0.1**



APPENDIX – 0.2 – ANNEXE

NEWFOUNDLAND – TERRE-NEUVE

Composition of Non-self-representing Units
Composition des unités non autoreprésentatives

ER R.É.	Stratum Strate	Number of PSUs Nombre d'U.P.É.	NSRU population Population des U.N.A.R.	Population		LF characteristics[1] used in stratification and % of labour force Caractéristiques de la P.A.[1] pour la stratification et % de la population active		
				Rural Rurale	Urban Urbaine	(1)	(2)	(3)
				%				
01	01000	12	24,996	100	0	S 41	C – BT 17	Ma – IM 21
	01020	10	19,469	100	0	S 42	C – BT 20	Ma – IM 14
	01040	8	14,202	0	100	S 45	C – BT 17	Ma – IM 18
T	3	30	58,667	76	24	S 43	C – BT 18	Ma – IM 18
02	02000	9	19,214	100	0	Ma – IM 18	S 26	F – FP 23
	02020	8	10,163	0	100	Ma – IM 42	S 30	F – FP 12
T	2	17	29,377	65	35	MA 36	S 28	F – FP 19
03	03000	10	20,141	100	0	S 45	F – FP 14	Ma – IM 8
	03020	11	22,944	100	0	S 35	F – FP 35	Ma – IM 13
	03040	4	6,579	0	100	S 55	F – FP 12	Ma – IM 14
T	3	25	49,664	87	13	S 41	F 26	Ma – IM 12
04	04000	10	20,546	100	0	S 34	F – FP 32	C – BT 13
	04020	12	23,543	100	0	S 38	F – FP 14	C – BT 22
	04040	9	19,396	100	0	S 36	F – FP 19	C – BT 12
	04060	8	15,214	0	100	S 38	F – FP 10	C – BT 7
	04080	8	16,334	0	100	S 52	F – FP 9	C – BT 15
T	5	47	95,033	67	33	S 40	F – FP 17	C – BT 15
Prov.	13	119	232,741	73	27	–	–	–

[1] Services (S), manufacturing (Ma), construction (C), forestry and fisheries (F), Mining (Mi), transportation (T), agriculture (A). – Services (S), industries manufacturières (IM), bâtiment et travaux publics (BT), forêts et pêche (FP), industrie minière (Mi) transport (T), agriculture (A).

ER	Composition
01	Division 1
02	Divisions 2, 3
03	Divisions 4, 5, 9
04	Divisions 6, 7, 8

R.É.	Composition
01	Division n° 1
02	Divisions n°s 2, 3
03	Divisions n°s 4, 5, 9
04	Divisions n°s 6, 7, 8

APPENDIX – 0.3 – ANNEXE

NEWFOUNDLAND – TERRE-NEUVE

1971 List, by ER and Population

Liste des U.A.R. de 1971, par R.É. et selon la population

Serial No.[1] – N° de série[1]	Name of SRU – Nom de l'U.A.R.	1971 population – Population en 1971	Number of subunits – Nombre de sous-unités	Numerical identification – Code d'identification	Economic region – Région économique
1	St. John's	131,814	12	01101	01
2	Carbonear, town – ville	11,205	01	01201	01
3	Placentia Bay – Baie de Plaisance	6,576	01	01301	01
4	Wabana, town – ville	6,079	01	01401	01
5	Marystown	9,719	01	02101	02
6	Channel – Port aux Basques	7,100	01	02201	02
7	Fortune, town – ville, Grand Banks	5,640	01	02301	02
8	Corner Brook, city – cité	26,309	02	03101	03
9	Stephenville, town – ville	14,906	01	03201	03
10	Labrador city – ville (L.I.D.) – (D.A.L.) and – et Wabush (L.I.D.) – (D.A.L.)	11,009	03	03301	03
11	Happy Valley, town – ville and – et Goose Bay	8,528	02	03401	03
12	Deer Lake, Pasadena town – ville	5,385	01	04101	04
13	Windsor, town – ville and – et Grand Falls	14,321	01	04101	04
14	Bishop's Falls and – et Botwood	8,248	01	04201	04
15	Gander, town – ville	7,748	01	04301	04
16	Bonavista, town – ville	5,924	01	04401	04
	Total	280,511	–	–	–

[1] As on the map. – Voir la carte.

Prince Edward Island — Île-du-Prince-Édouard

Legend
Principal Self-Representing Unit
Other Self-Representing Units

Légende
Principale unité autoreprésentative
Autres unités autoreprésentatives

APPENDIX – 1.2 – ANNEXE

PRINCE EDWARD ISLAND – ÎLE-DU-PRINCE-ÉDOUARD

Composition of Non-self-representing Units
Composition des unités non autoreprésentatives

ER R.É.	Stratum Strate	Number of PSUs Nombre d'U.P.É.	NSRU population Population des U.N.A.R.	Population		LF characteristics[1] used in stratification and % of labour force Caractéristiques de la P.A.[1] pour la stratification et % de la population		
				Rural Rurale	Urban Urbaine	(1)	(2)	(3)
				%				
10	10000	10	19,315	81	19	S 45	A 25	IM 7
	10020	14	29,315	79	21	S 42	A 15	IM 16
	10040	12	23,081	80	20	S 58	A 16	IM 8
Prov.	3	36	71,704	79	21	S 48	A 18	IM 11

[1] Services (S), manufacturing (Ma), construction (C), forestry and fisheries (F), mining (Mi), transportation (T), agriculture (A). – Services (S), industries manufacturières (IM), bâtiment et travaux publics (BT), forêts et pêche (FP), industrie minière (Mi), transports (T), agriculture (A).

ER	Composition
10	All of Prince Edward Island

R.É.	Composition
10	Totalité de l'Île-du-Prince-Édouard

APPENDIX – 1.3 – ANNEXE

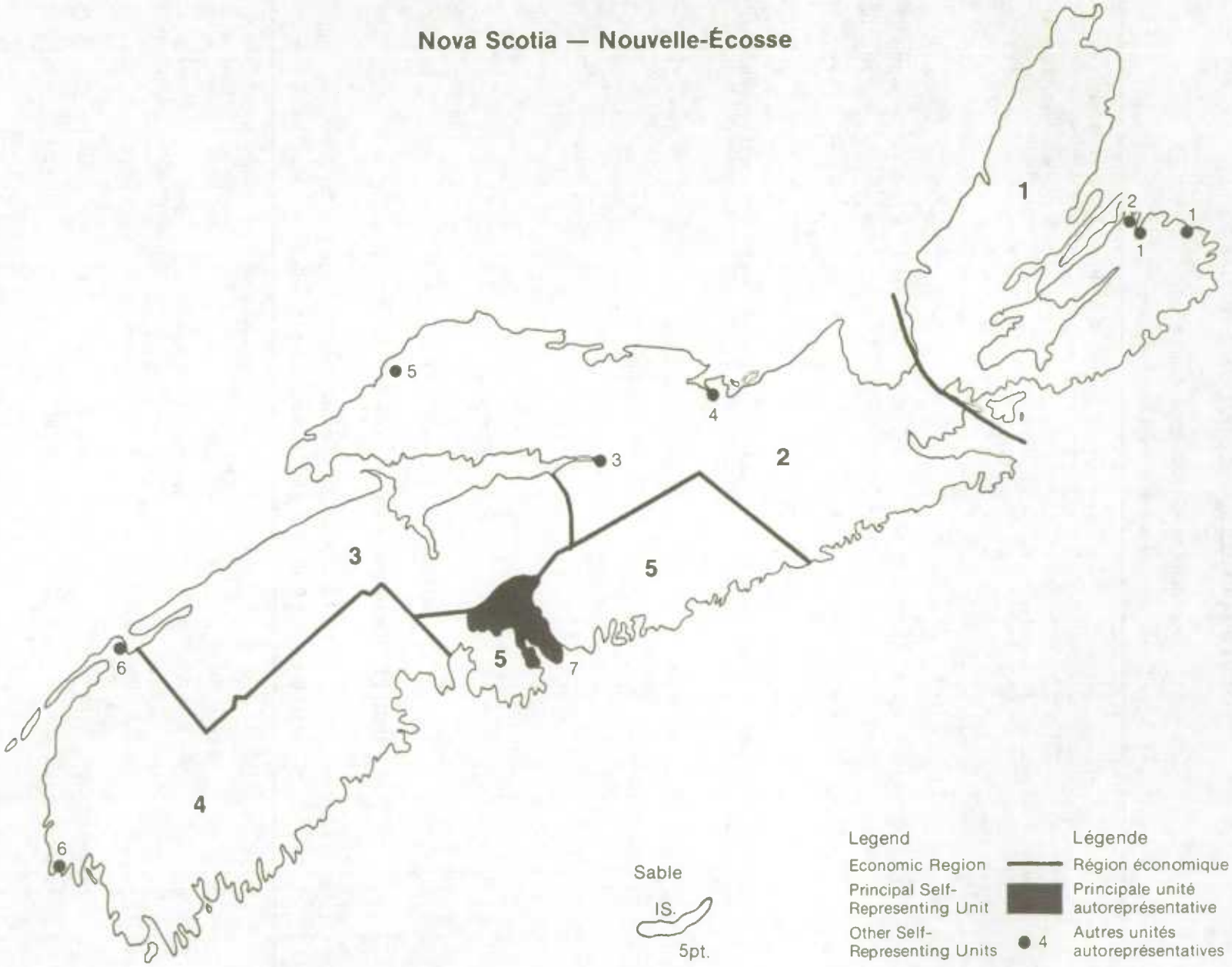
PRINCE EDWARD ISLAND – ÎLE-DU-PRINCE-ÉDOUARD

1971 SRU List, by ER and Population
Liste des U.A.R. de 1971 par R.É. et selon la population

Serial No.[1] N° de série[1]	Name of SRU Nom de l'U.A.R.	1971 population Population en 1971	Number of subunits Nombre de sous-unités	Numerical identification Code d'identification	Economic region Région économique
1	Charlottetown	25,253	03	10101	10
2	Summerside	14,004	02	10201	10
	Total	39,257	–	–	–

[1] As on the map. – Voir la carte.

Nova Scotia — Nouvelle-Écosse



APPENDIX – 2.2 – ANNEXE

NOVA SCOTIA – NOUVELLE-ÉCOSSE

Composition of Non-self-representing Units

Composition des unités non autoreprésentatives

ER – R.É.	Stratum – Strate	Number of PSUs – Nombre d'U.P.É.	NSRU population – Population des U.N.A.R.	Population		LF characteristic[1] used in stratification and % of labour force – Caractéristiques de la P.A.[1] pour la stratification et % de la population active		
				Rural – Rurale	Urban – Urbaine	(1)	(2)	(3)
				%				
21	21000	10	20,098	78	22	S 44	Ma – IM 24	F – FP 5
	21020	12	24,034	81	19	S 52	Ma – IM 12	F – FP 13
T	2	22	44,132	80	20	S 49	Ma – IM 17	F – FP 10
22	22000	12	24,671	83	17	S 45	Ma – IM 19	C – BT 10
	22020	10	19,704	83	17	S 42	Ma – IM 21	C – BT 10
	22040	12	24,872	75	25	S 57	Ma – IM 11	C – BT 13
	22060	10	20,089	59	41	S 54	Ma – IM 15	C – BT 8
T	4	44	89,336	75	25	S 50	Ma – IM 17	C – BT 10
23	23000	15	29,663	62	38	S 58	Ma – IM 14	A 11
	23020	14	28,395	82	18	S 49	Ma – IM 15	A 8
	23040	16	29,974	75	25	S 57	Ma – IM 13	A 10
T	3	45	88,032	73	27	S 55	Ma – IM 14	A 10
24								
	24000	11	23,746	59	41	S 50	Ma – IM 28	F – FP 4
	24020	12	27,188	83	17	S 44	Ma – IM 26	F – FP 7
	24040	13	24,920	80	20	S 42	Ma – IM 25	F – FP 23
	24060	13	25,581	100	–	S 48	Ma – IM 21	F – FP 10
T	4	49	101,435	81	19	S 46	Ma – IM 25	F – FP 11
25	25000	10	21,515	100	–	S 49	Ma – IM 17	C – BT 14
	25020	9	17,281	75	25	S 62	Ma – IM 12	C – BT 12
T	2	19	38,796	88	12	S 55	Ma – IM 15	C – BT 13
Prov.	15	179	361,731	78	22	–	–	–

[1] Services (S), manufacturing (Ma), construction (C), forestry and fisheries (F), mining (Mi), transportation (T), agriculture (A). – Services (S), industries manufacturières (IM), bâtiment et travaux publics (BT), forêts et pêche (FP), industrie minière (Mi), transports (T), agriculture (A).

ER	Composition
21	Counties of Cape Breton, Inverness, Richmond, Victoria
22	Counties of Antigonish, Guysborough, Pictou, Colchester, Cumberland
23	Counties of Hants, Kings, Annapolis
24	Counties of Lunenburg, Queens, Shelburne, Digby, Yarmouth
25	County of Halifax

R.É.	Composition
21	Comtés de Cap Breton, Inverness, Richmond, Victoria
22	Comtés de Antigonish, Guysborough, Pictou, Colchester, Cumberland
23	Comtés de Hants, Kings, Annapolis
24	Comtés de Lunenburg, Queens, Sherburne, Digby, Yarmouth
25	Comté de Halifax

APPENDIX – 2.3 – ANNEXE

NOVA SCOTIA – NOUVELLE-ÉCOSSE

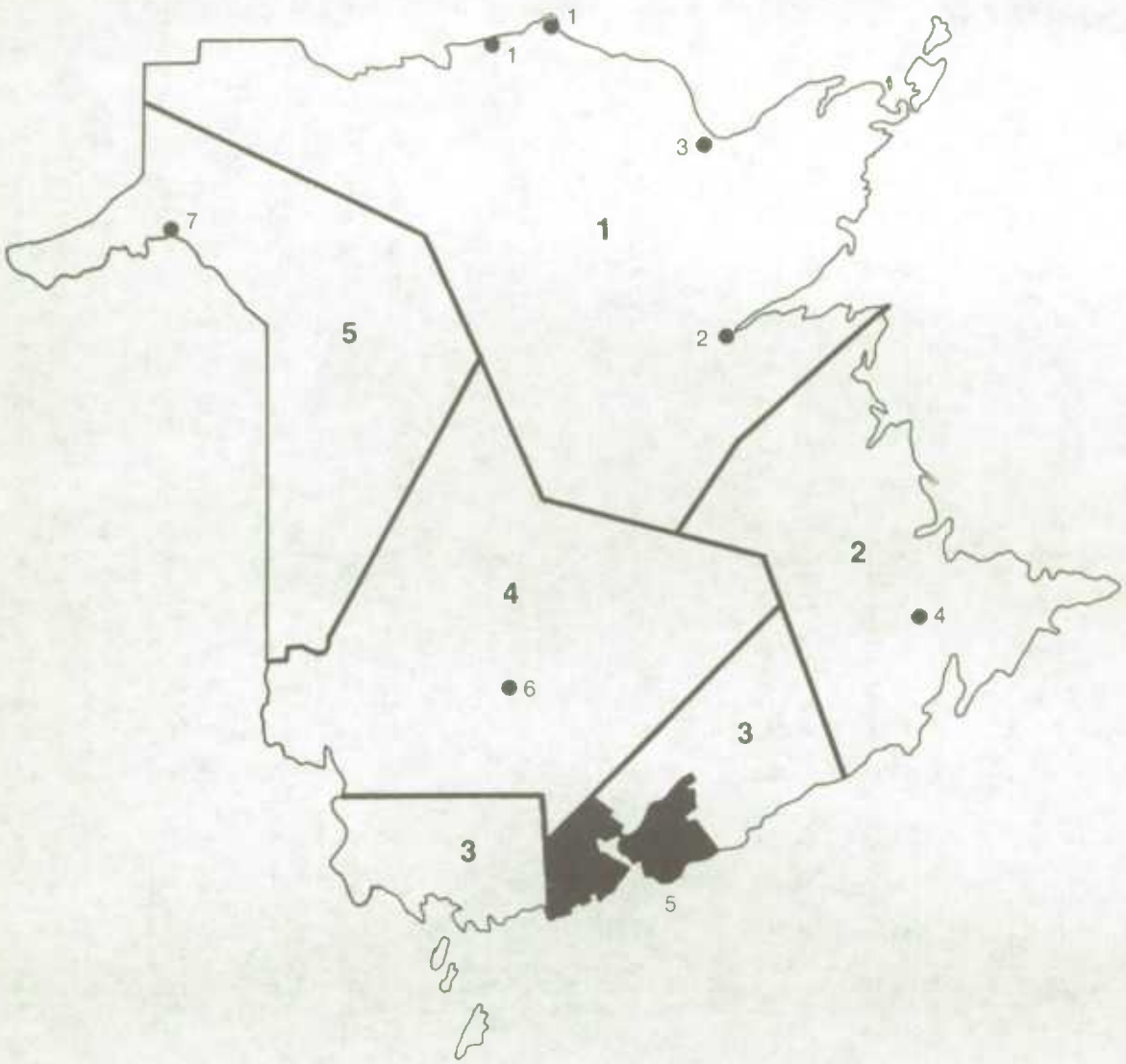
1971 SRU List, by ER and Population
Liste des U.A.R. de 1971 par R.É. et selon la population

Serial No.[1] – N° de série[1]	Name of SRU – Nom de l'U.A.R.	1971 population – Population en 1971	Number of subunits – Nombre de sous-unités	Numerical identification – Code d'identification	Economic region – Région économique
1	Sidney, city and Glace Bay – Sidney, ville et Glace Bay	90,879	11	21101	21
2	Sidney Mines[1]	32,760	05	21201	21
3	Truro	24,693	04	22101	22
4	New Glasgow	23,435	03	22201	22
5	Amherst	9,966	01	22301	22
6	Yarmouth and Digby	10,879	01	24101	24
7	Halifax	222,635	24	25101	25
	Total	**415,247**	**01 Apt. – App.**	–	–

[1] As on the map. – Voir la carte.

Appendix 3.1 Annexe 3.1

New Brunswick — Nouveau-Brunswick



Legend	Légende
Economic Region	Région économique
Principal Self-Representing Unit	Principale unité autoreprésentative
Other Self-Representing Units ● 4	Autres unités autoreprésentatives

APPENDIX – 3.2 – ANNEXE

NEW BRUNSWICK – NOUVEAU-BRUNSWICK

Composition of Non-self-representing Units

Composition des unités non autoreprésentatives

ER – R.É.	Stratum – Strate	Number of PSUs – Nombre d'U.P.É.	NSRU population – Population des U.N.A.R.	Population		LF characteristics[1] used in stratification and % of labour force – Caractéristiques de la P.A.[1] pour la stratification et % de la population active		
				Rural – Rurale	Urban – Urbaine	(1)	(2)	(3)
				%				
31	31000	15	30,982	72	28	S 43	Ma – IM 20	F – Fp 15
	31020	12	24,137	73	27	S 36	Ma – IM 37	F – Fp 12
	31040	14	28,111	74	26	S 48	Ma – IM 14	F – Fp 18
	31060	13	26,363	72	28	S 46	Ma – IM 20	F – Fp 4
T	4	54	109,593	73	27	S 44	Ma – IM 22	F – Fp 12
32	32000	12	22,890	70	30	S 43	Ma – IM 3	T 9
	32020	12	22,381	69	31	S 54	Ma – IM 13	T 10
	32040	11	21,328	60	40	S 62	Ma – IM 15	T 11
T	3	35	66,599	66	34	S 53	Ma – IM 20	T 10
33	33000	11	24,516	57	43	S 48	Ma – IM 27	F – Fp 10
	33020	9	18,655	69	31	S 47	Ma – IM 17	F – Fp 9
T	2	20	43,171	62	38	S 49	Ma – IM 20	F – Fp 10
34	34000	13	26,446	59	41	S 43	Ma – IM 20	T 12
	34020	11	24,106	77	23	S 60	Ma – IM 9	T 11
T	2	24	50,552	68	32	S 51	Ma – IM 15	T 12
35	35000	10	20,942	68	32	S 37	Ma – IM 27	A 14
	35020	9	20,824	54	46	S 53	Ma – IM 16	A 16
	35040	11	24,271	49	51	S 53	Ma – IM 16	A 7
T	3	30	66,037	57	43	S 48	Ma – IM 19	A 12
Prov.	14	163	335,952	66	34	–	–	–

[1] Services (S), manufacturing (Ma), construction (C), forestry and fisheries (F), mining (Mi), transportation (T), agriculture (A). – Services (S), industries manufacturières (IM), bâtiment et travaux publics (BT), forêts et pêche (FP), industrie minière (Mi), transports (T), agriculture (A).

ER	Composition
31	Counties of Northumberland, Restigouche, Gloucester
32	Counties of Alberta, Kent, Westmorland
33	Counties of Charlotte, Kings, Saint John
34	Counties of Queen's, Sudbury, York
35	Counties of Carleton, Madawaska, Victoria.

R.É.	Composition
31	Comtés de Northumberland, Restigouche, Gloucester
32	Comtés de Alberta, Kent Westmorland
33	Comtés de Charlotte, Kings, Saint John
34	Comtés de Queen's, Sudbury, York
35	Comtés de Carleton, Madawaska, Victoria.

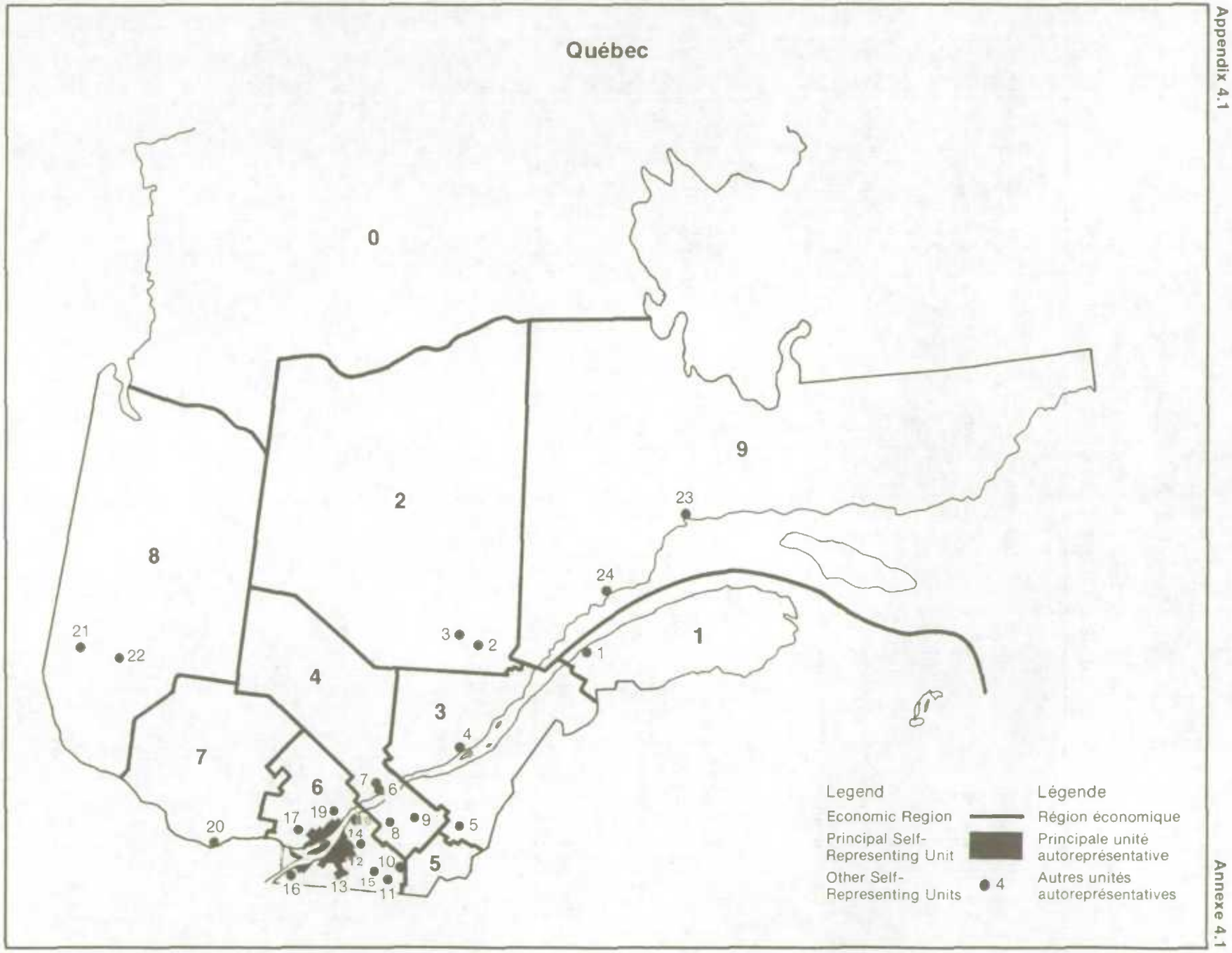
APPENDIX – 3.3 – ANNEXE

NEW BRUNSWICK – NOUVEAU-BRUNSWICK

1971 SRU List, by ER and Population
Liste des U.A.R. de 1971 par R.É. et selon la population

Serial No.[1] – N° de série[1]	Name of SRU – Nom de l'U.A.R.	1971 population – Population en 1971	Number of subunits – Nombre de sous-unités	Numerical identification – Code d'identification	Economic region – Région économique
1	Dalhousie and – et Campbellton	19,773	02	31101	31
2	Chatham and – et Newcastle	19,762	03	31201	31
3	Bathurst	16,674	02	31301	31
4	Moncton	71,416	09	32101	32
5	Saint John	106,744	14	33101	33
6	Fredericton	37,361	04	34101	34
7	Edmundston	12,365	01	35101	35
	Total	**284,095**	–	–	–

[1] As on the map. – Voir la carte.


Québec
0
2
9
8
4
3
7
6
5
1
23
24
21
22
3
2
1
4
7
6
17
19
20
14
9
8
5
12
10
16
13
15
11
Legend
Economic Region
Principal Self-Representing Unit
Other Self-Representing Units
4
Légende
Région économique
Principale unité autoreprésentative
Autres unités autoreprésentatives

APPENDIX – 4.2 – ANNEXE

QUÉBEC

Composition of Non-self-representing Units

Composition des unités non autoreprésentatives

ER – R.É.	Stratum – Strate	Number of PSUs – Nombre d'U.P.É.	NSRU population – Population des U.N.A.R.	Population		LF characteristics[1] used in stratification and % of labour force – Caractéristiques de la P.A.[1] pour la stratification et % de la population active		
				Rural – Rurale	Urban – Urbaine	(1)	(2)	(3)
				%				
41	41000	13	70,292	76	24	S 55	Ma – IM 16	F – FP 7
	41020	12	64,326	47	53	S 58	Ma – IM 11	F – FP 8
	41040	12	65,770	71	29	S 53	Ma – IM 18	F – FP 7
T	3	37	200,388	65	35	S 55	Ma – IM 15	F – FP 7
42	42000	19	105,924	54	46	S 52	Ma – IM 17	A 11
T	1	19	105,924	54	46	S 52	Ma – IM 17	A 11
43	43000	16	88,009	63	37	S 53	Ma – IM 19	A 12
	43020	16	82,254	53	47	S 46	Ma – IM 23	A 13
	43040	18	95,876	54	46	S 57	Ma – IM 13	A 11
	43060	14	75,228	48	52	S 50	Ma – IM 21	A 12
	43080	17	88,719	66	34	S 52	Ma – IM 17	A 11
T	5	81	430,086	57	43	S 52	Ma – IM 18	A 12
44	44000	17	91,228	59	41	S 42	Ma – IM 31	A 15
	44020	17	92,693	62	38	S 46	Ma – IM 26	A 14
T	2	34	183,921	61	39	S 44	Ma – IM 28	A 14
45	45000	12	62,063	56	44	S 47	Ma – IM 31	Mi 1
	45020	12	62,605	47	53	S 40	Ma – IM 27	Mi 12
T	2	24	124,668	51	49	S 43	Ma – IM 29	Mi 7
46	46000	13	69,255	41	59	S 42	Ma – IM 37	A 11
	46020	13	72,527	72	28	S 38	Ma – IM 29	A 21
	46040	18	94,158	64	36	S 48	Ma – IM 19	A 17
	46060	15	83,655	58	42	S 55	Ma – IM 22	A 7
	46080	17	93,938	62	38	S 50	Ma – IM 26	A 9
T	5	76	413,533	60	40	S 47	Ma – IM 26	A 13
47	47000	17	90,648	59	41	S 58	Ma – IM 13	A 7
T	1	17	90,648	59	41	S 58	Ma – IM 13	A 7
48	48000	17	87,915	67	33	S 53	Ma – IM 16	Mi 7
T	1	17	87,915	67	33	S 53	Ma – IM 16	Mi 7
49	49000	14	37,601	49	51	S 49	T 12	Mi 11
T	1	14	37,601	49	51	S 49	T 12	Mi 11
Prov.	21	319	674,684	59	41	–	–	–

[1] Services (S), manufacturing (Ma), construction (C), forestry and fisheries (F), mining (Mi), transportation (T), agriculture (A). – Services (S), Industries manufacturières (IM), bâtiment et travaux publics (BTO), forêts et pêche (FP), industrie minière (Mi), transports (T), agriculture (A).

ER	Composition
41	Counties of Rimouski (part), Matane, Matapedia, Bonaventure, Gaspé-est, Gaspé-ouest, Îles-de-la-Madeleine.
42	Counties of Chicoutimi, Lac Saint-Jean est, Lac Saint-Jean ouest, Abitibi (part).
43	Counties of Frontenac (part), Beauce, Dorchester, Bellechasse, Montmagny, Montmorency, Lévis, Lotbinière (part), Mégantic (part), Portneuf (part), Québec (part), Rivière-du-Loup, Rimouski (part), Témiscouata, Kamouraska, L'Islet, Charlevoix-ouest, Charlevoix-est, Saguenay (part), Wolfe (part).
44	Counties of Drummond, Arthabaska, Nicolet, Mégantic (part), Yamaska, Lotbinière (part), Champlain, Saint-Maurice, Maskinongé (part), Portneuf (part), Abitibi (part), Québec (part).
45	Counties of Stanstead, Sherbrooke, Richmond, Wolfe (part), Compton, Frontenac (part).
46	Counties of Shefford, Brome, Bagot, Yamaska (part), Iberville, Papineau (part), Rouville, Saint-Hyacinthe, Richelieu, Verchères, Berthier (part), Maskinongé (part), Labelle (part), Joliette (part), Montcalm (part), Terrebonne, St-Jean, Napierville, Laprairie, Huntingdon, Châteauguay, Beauharnois, Soulanges, Vaudreuil, Deux-Montagnes, Argenteuil, L'Assomption, Missisquoi, Chambly, Île-de-Montréal and Ile-Jésus.
47	Counties of Pontiac, Gatineau, Labelle (part), Papineau (part), Hull, Montcalm (part), Berthier (part), Joliette (part), Maskinongé (part).
48	Counties of Abitibi (part), Témiscamingue.
49	Counties of Saguenay (part).
40	Counties of Abitibi (part), Saguenay (part).

R.É.	Composition
41	Comtés de Rimouski (partie), Matane, Matapédia, Bonaventure, Gaspé-est, Gaspé-ouest, des Îles-de-la-Madeleine.
42	Comtés de Chicoutimi, du Lac Saint-Jean est, du Lac Saint-Jean ouest, d'Abitibi (partie).
43	Comtés de Frontenac (partie), Beauce, Dorchester, Bellechasse, Montmagny, Montmorency, Lévis, Lotbinière (partie), Mégantic (partie), Portneuf, (partie), Québec (partie), Rivière-du-Loup, Rimouski (partie), Témiscouata, Kamouraska, L'Islet, Charlevoix-ouest, Charlevoix-est, du Saguenay (partie), de Wolfe (partie).
44	Comtés de Drummond, Arthabaska, Nicolet, Mégantic (partie), Yamaska, Lotbinière (partie), Champlain, Saint-Maurice, Maskinongé (partie), Portneuf (partie), Abitibi (partie), Québec (partie).
45	Comtés de Stanstead, Sherbrooke, Richmond, Wolfe (partie), Compton, Frontenac (partie).
46	Comtés de Shefford, Brome, Bagot Yamaska (partie), Iberville, Papineau (partie), Rouville, Saint-Hyacinthe, Richelieu, Verchères, Berthier (partie), Maskinongé (partie), Labelle (partie), Joliette (partie), Montcalm (partie), Terrebonne, St-Jean, Napierville, Laprairie, Huntingdon, Châteauguay, Beauharnois, Soulanges, Vaudreuil, Deux-Montagnes, Argenteuil, L'Assomption, Missisquoi, Chambly, Île-de-Montréal, et Île-Jésus.
47	Comtés de Pontiac, Gatineau, Labelle (partie), Papineau (partie), Hull, Montcalm (partie), Berthier (partie), Joliette (partie), Maskinongé, (partie).
48	Comtés de Abitibi (partie), Témiscamingue.
49	Comtés du Saguenay (partie).
40	Comtés de Abitibi (partie), Saguenay (partie).

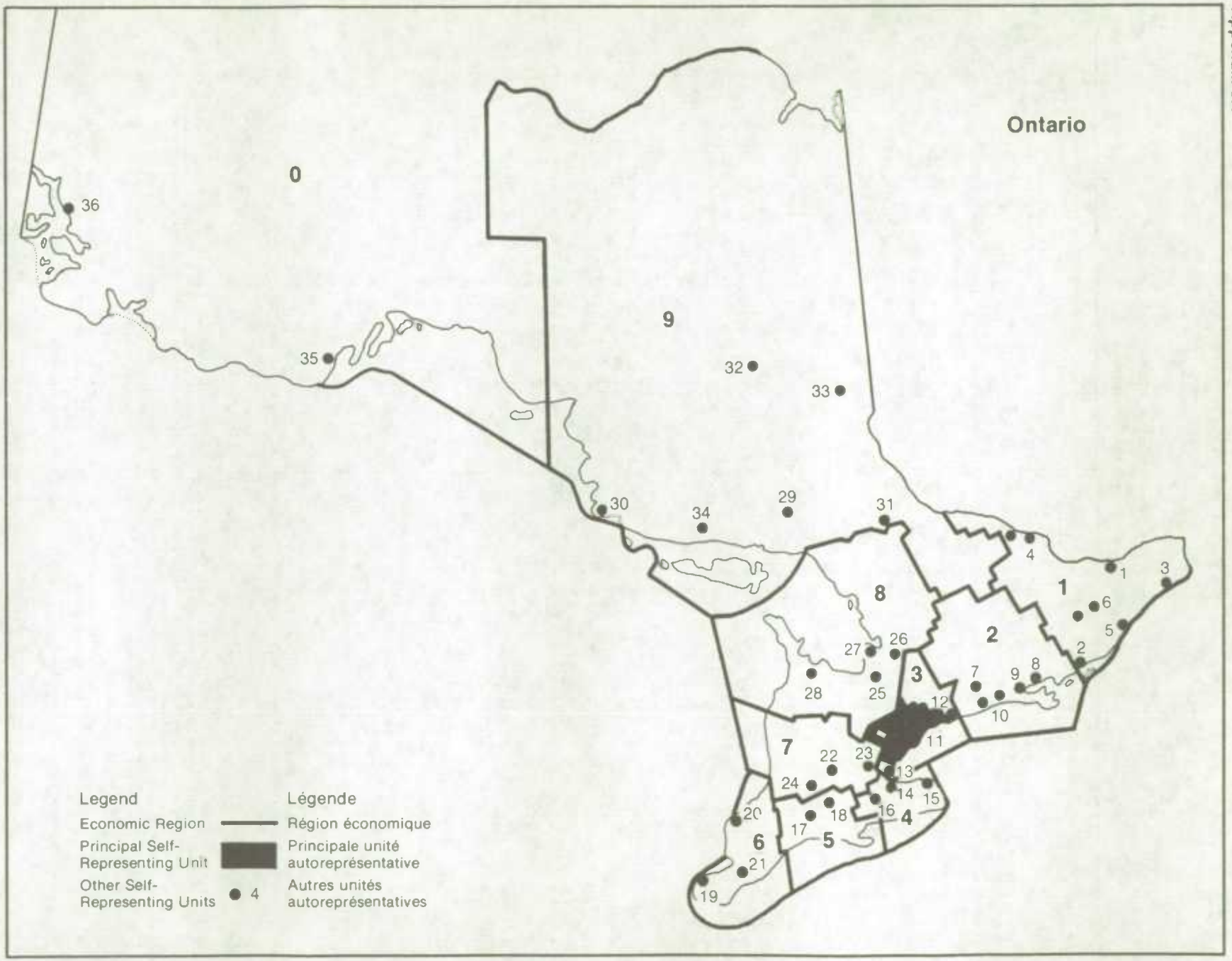
APPENDIX – 4.3 – ANNEXE

QUÉBEC

1971 SRU List, by ER and Population
Liste des U.A.R. de 1971 par R.É. et selon la population

Serial No.[1] — N° de série[1]	Name of SRU — Nom de l'U.A.R.	1971 Population — Population en 1971	Number of subunits — Nombre de sous-unités	Numerical identification — Code d'identification	Economic region — Région économique
1	Rimouski	28,956	01	41101	41
2	Chicoutimi - Jonquiere	133,703	05	42101	42
3	Alma	22,622	01	42201	42
4	Québec City – Ville de Québec	479,604	22	43101	43
			01 Apt. – App.	43123	
5	Thetford Mines	26,126	01	43201	43
6	Trois-Rivières	97,930	03	44101	44
7	Shawinigan	59,307	02	44201	44
8	Drummondville	46,524	02	44301	44
9	Victoriaville	26,526	01	44401	44
10	Sherbrooke	84,570	04	45101	45
11	Magog	14,383	01	45201	45
			90	46101	
12	Montréal	2,739,226	19	46201	46
			04 Apt. – App.	46191	
13	St-Jean	47,044	02	46301	46
14	Saint-Hyacinthe	39,693	02	46401	46
15	Granby	39,307	02	46811	46
16	Valleyfield	37,430	01	46501	46
17	St-Jérome	35,335	01	46601	46
18	Sorel	34,479	01	46701	46
19	Joliette	29,350	01	46801	46
20	Hull	149,230	07	47101	47
21	Rouyn-Noranda	28,562	01	48101	48
22	Val-d'Or	24,512	01	48201	48
23	Sept-Îles	24,320	01	49101	49
24	Baie-Comeau	27,220	01	49201	49
	Total	4,275,959	–	–	–

[1] As on the map. – Voir la carte.


Ontario
0
9
8
1
2
3
7
6
5
4
36
35
32
33
30
34
29
31
27
26
28
25
12
11
10
22
23
24
13
14
15
20
18
16
17
21
19
Legend
Economic Region
Principal Self-Representing Unit
Other Self-Representing Units
4
Légende
Région économique
Principale unité autoreprésentative
Autres unités autoreprésentatives

APPENDIX – 5.2 – ANNEXE

ONTARIO

Composition of Non-self-representing Units

Composition des unités non autoreprésentatives

ER R.É.	Stratum Strate	Number of PSUs Nombre d'U.P.É.	NSRU population Population des U.N.A.R.	Population		LF characteristics[1] used in stratification and % of labour force Caractéristiques de la P.A.[1] pour la stratification et % de la population active		
				Rural Rurale	Urban Urbaine	(1)	(2)	(3)
				%				
51	51000	17	87,692	72	28	S 51	Ma – IM 25	A 11
	51020	17	89,494	68	32	S 55	Ma – IM 14	A 18
	51040	15	80,525	65	35	S 58	Ma – IM 16	A 9
T	3	49	257,711	68	32	S 55	Ma – IM 18	A 13
52	52000	13	69,928	60	40	S 49	Ma – IM 31	A 8
	52020	11	67,279	76	24	S 57	Ma – IM 14	A 10
	52040	15	81,259	69	31	S 53	Ma – IM 22	A 12
T	3	39	218,466	68	32	S 53	Ma – IM 22	A 10
53	53000	13	62,271	85	15	S 53	Ma – IM 21	A 11
T	1	13	62,271	85	15	S 53	Ma – IM 21	A 11
54	54000	15	78,132	62	38	S 44	Ma – IM 24	A 18
T	1	15	78,132	62	38	S 44	Ma – IM 24	A 18
55	55000	17	90,851	60	40	S 40	A 30	Ma – IM 18
	55020	14	77,098	64	36	S 44	A 23	Ma – IM 20
T	2	31	167,949	62	38	S 45	A 27	Ma – IM 19
56	56000	14	71,396	58	42	S 45	Ma – IM 25	A 17
	56020	13	74,292	66	34	S 45	Ma – IM 25	A 18
T	2	27	145,688	62	38	S 45	Ma – IM 25	A 18
57	57000	14	74,394	66	34	S 44	A 29	Ma – IM 17
	57020	17	89,045	56	44	S 48	A 14	Ma – IM 25
T	2	31	163,439	61	39	S 46	A 21	Ma – IM 21
58	58000	18	93,804	53	47	S 49	Ma – IM 19	A 18
	58020	17	95,018	62	38	S 51	Ma – IM 19	A 13
	58040	13	66,349	48	52	S 63	Ma – IM 13	A 3
T	3	48	255,171	55	45	S 53	Ma – IM 17	A 15
59	59000	14	73,539	57	43	S 45	Ma – IM 11	Mi 26
	59020	19	105,146	44	56	S 50	Ma – IM 22	Mi 1
T	2	33	178,685	49	51	S 48	Ma – IM 17	Mi 11
50	50000	14	74,556	32	68	S 48	Ma – IM 17	Mi 11
T	1	14	74,556	32	68	S 48	Ma – IM 17	Mi 11
Prov.	20	300	1,602,068	61	39	–	–	–

[1] Services (S), manufacturing (Ma), construction (C), forestry and fisheries (F), Mining (Mi), transportation (T), agriculture (A). – Services (S), industries manufacturières (IM), bâtiment et travaux publics (BT), forêts et pêche (FP), industrie minière (Mi), transport (T), agriculture (A).

ER	Composition
51	Counties of Frontenac, Leeds, Renfrew, Lanark, Grenville, Ottawa-Carleton, Dundas, Stormont, Glengarry, Prescott, Russell.
52	Counties of Haliburton, Victoria, Durham, Hastings, Lennox and Addington, Prince Edward, Northumberland, Peterborough.
53	Counties of Halton, Peel, York, Ontario, Toronto.
54	Counties of Brant, Haldimand, Wentworth, Niagara.
55	Counties of Elgin, Oxford, Norfolk, Middlesex.
56	Counties of Lambton, Kent, Essex.
57	Counties of Huron, Perth, Waterloo, Wellington.
58	Counties of Bruce, Grey, Dufferin, Muskoka, Simcoe, Parry Sound.
59	Counties of Algoma, Cochrane, Sudbury, Nipissing, Manitoulin, Timiskaming.
50	Counties of Kenora, Rainy River, Thunder Bay.

R.É.	Composition
51	Comtés de Frontenac, Leeds, Renfrew, Lanark, Grenville, Ottawa-Carleton, Dundas, Stormont, Glengarry, Prescott, Russell.
52	Comtés de Haliburton, Victoria, Durham, Hastings, Lennox et Addington, Prince Edward, Northumberland, Peterborough.
53	Comtés de Halton, Peel, York, Ontario, Toronto.
54	Comtés de Brant, Haldimand, Wentworth, Niagara.
55	Comtés de Elgin, Oxford, Norfolk, Middlesex.
56	Comtés de Lambton, Kent, Essex.
57	Comtés de Huron, Perth, Waterloo, Wellington.
58	Comtés de Bruce, Grey, Dufferin, Muskoka, Simcoe, Parry Sound.
59	Comtés de Algoma, Cochrane, Sudbury, Nipissing, Manitoulin, Timiskaming.
50	Comtés de Kenora, Rainy River, Thunder Bay.

APPENDIX – 5.3 – ANNEXE

ONTARIO

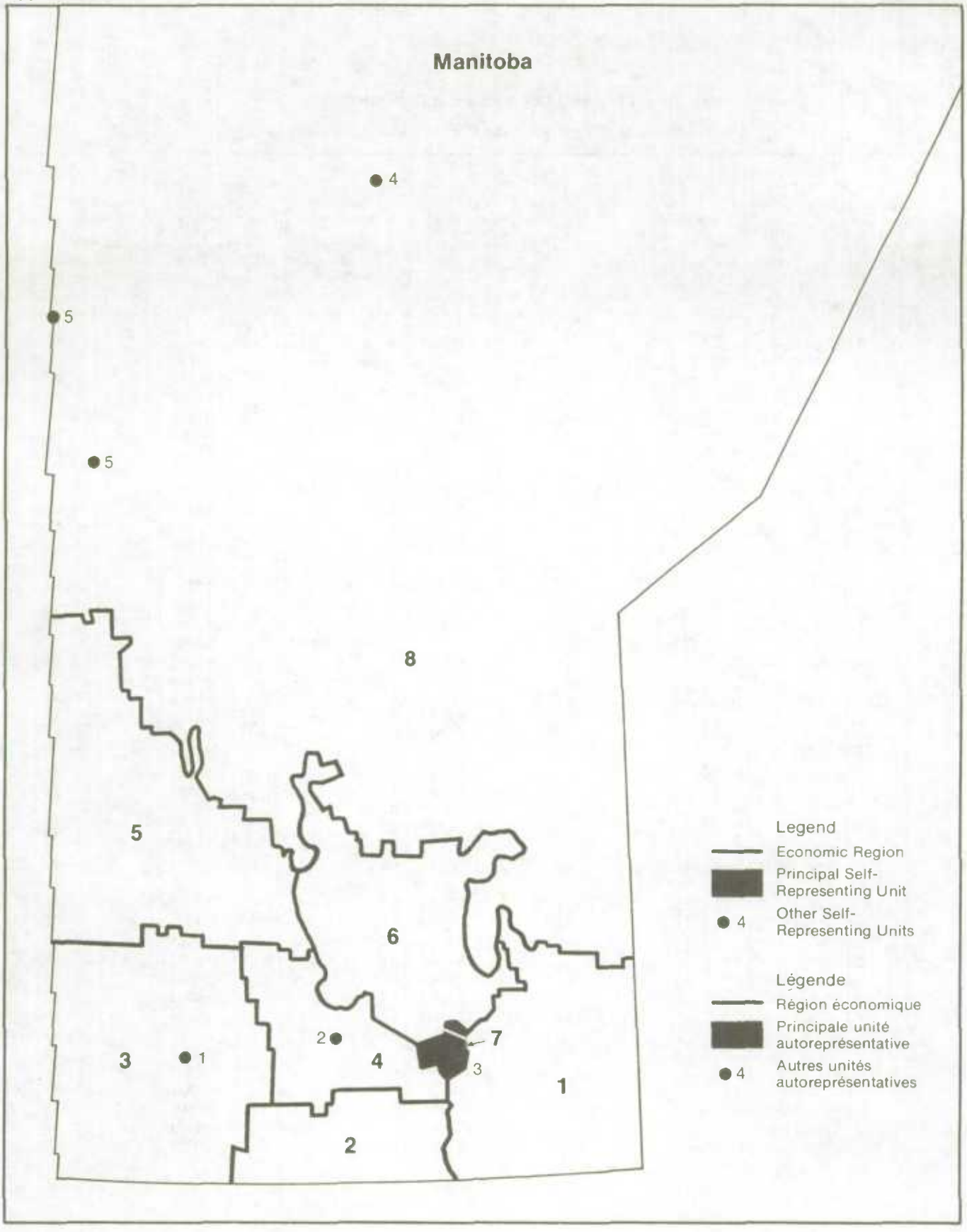
1971 SRU List, by ER and Population

Liste des U.A.R. de 1971 par R.É. et selon la population

Serial No.[1] – Nº de série[1]	Name of SRU – Nom de l'U.A.R.	1971 population – Population en 1971	Number of subunits – Nombre de sous-unités	Numerical identification – Code d'identification	Economic region – Région économique
1	Ottawa	453,280	16	51101	51
			02 Apt. – App.	51117	
2	Kingston	85,877	04	51201	51
3	Cornwall	47,116	02	51301	51
4	Pembroke and – et Petawawa	26,083	01	51401	51
5	Brockville and – et Prescott	24,930	01	51501	51
6	Smith Falls and – et Perth	19,581	01	51601	51
7	Peterborough	63,531	03	52101	52
8	Belleville	35,125	02	52201	52
9	Trenton	28,650	01	52301	52
10	Port Hope and – et Cobourg	27,188	01	52401	52
11	Toronto	2,628,043	79	53101	53
			14 Apt. – App.	53180	
12	Oshawa	120,318	05	53201	53
			01 Apt. – App.	53206	
13	Burlington	87,023	03	53301	53
			01 Apt. – App.	53304	
14	Hamilton	411,500	17	54101	54
			02 Apt. – App.	54118	
15	St. Catharines – Niagara	303,429	14	54201	54
16	Brantford	80,284	03	54301	54
17	London	286,011	13	55101	55
			02 Apt. – App.	55114	
18	Woodstock	26,173	01	55201	55
19	Windsor	258,643	10	56101	56
			01 Apt. – App.	56111	
20	Sarnia	77,917	03	56201	56
21	Chatham	35,317	02	56301	56
22	Kitchener and – et Waterloo	226,846	08	57101	57
			01 Apt. – App.	57109	
23	Guelph	62,659	03	57201	57
24	Stratford	24,508	01	57301	57
25	Barrie	38,176	02	58101	58
26	Orillia	24,040	01	58201	58
27	Midland and – et Penetanguishene	23,503	01	58301	58
28	Owen Sound	18,469	01	58401	58
29	Sudbury	155,424	07	59101	59
30	Sault Ste. Marie	80,332	03	59201	59
31	North Bay	49,187	02	59301	59
32	Timmins	41,473	02	59401	59
33	Kirkland Lake	14,472	01	59501	59
34	Elliot Lake	9,093	01	59601	59
35	Thunder Bay	112,093	04	50101	50
36	Kenora and – et Keewatin	13,064	01	50201	50
	Total	6,019,358	–	–	–

[1] As on the map. – Voir la carte.

Appendix 6.1 Annexe 6.1



APPENDIX – 6.2 – ANNEXE

MANITOBA

Composition of Non-self-representing Units

Composition des unités non autoreprésentatives

ER R.É.	Stratum Strate	Number of PSUs Nombre d'U.P.E.	NSRU population Population des U.N.A.R.	Population		LF characteristics[1] used in stratification and % of labour force Caractéristiques de la P.A.[1] pour la stratification et % de la population active		
				Rural Rurale	Urban Urbaine	(1)	(2)	(3)
				%				
61	61000	18	58,102	74	26	CHS – BPM 35	SG – CS 12	W – B 8
T	1	18	58,102	74	26	CHS – BPM 35	SG – CS 12	W – B 8
62	62000	14	45,705	64	36	CHS – BPM 35	SG – CS 31	W – B 10
T	1	14	45,705	64	36	CHS – BPM 35	SG – CS 31	W – B 10
63	63000	11	36,293	64	36	CHS – BPM 28	SG – CS 43	W – B 16
	63020	11	35,841	69	31	CHS – BPM 47	SG – CS 22	W – B 15
T	2	22	72,134	66	34	CHS – BPM 37	SG – CS 33	W – B 16
64	64000	9	29,806	88	12	CHS – BPM 34	SG – CS 34	W – B 11
T	1	9	29,806	88	12	CHS – BPM 34	SG – CS 34	W – B 11
65	65000	17	55,394	57	43	CHS – BPM 45	SG – CS 32	W – B 9
T	1	17	55,394	57	43	CHS – BPM 45	SG – CS 32	W – B 9
66	66000	16	52,791	69	31	CHS – BPM 39	SG – CS 23	W – B 7
67	All SRU – Ensemble des U.A.R.							
68	Remote – Région éloignée							
T	7	96	313,932	68	32	–	–	–

[1] Cattle, hog, sheep (CHS), small grain (SG), wheat (W). – Bovins, porcs, moutons (BPM), céréales secondaires (CS), blé (B).

ER	Composition
61	Divisions 1, 5 pt. 19 pt.; nos. 1, 5 (part.), 19 (part.)
62	Divisions 2, 3 pt.; nos. 2, 3 (part.)
63	Divisions 3 pt. 4, 7 pt. 8, 10 pt. 11 pt. 13 pt.; nos. 3 (part.), 4, 7 (part.), 8, 10 (part.), 11 (part.), 13 (part.)
64	Divisions 6, 7 pt. 10 pt. 18 pt.; nos. 6, 7 (part.), 10 (part.), 18 (part.)
65	Divisions 10 pt. 11 pt. 13 pt. 14 pt. 15 pt. 17, 18 pt.; nos. 10 (part.), 11 (part.), 13 (part.), 14, 15, 17, 18 (part.)
66	Divisions 5 pt. 9, 12 pt.; nos. 5 (part.), 9, 12 (part.)
67	Winnipeg CMA, 20
68	Divisions 12 pt. 15 pt. 16, 18 pt. 19 pt.; nos. 12 (part.), 15 (part.), 16, 18 (part.), 19 (part.)

R.É.	Composition
61	Divisions nos 1, 5 (part.), 19 (part.)
62	Divisions nos 2, 3 (part.)
63	Divisions nos 3 (part.), 4, 7 (part.), 8, 10 (part.), 11 (part.), 13 (part.)
64	Divisions nos 6, 7 (part.), 10 (part.), 18 (part.)
65	Divisions nos 10 (part.), 11 (part.), 13 (part.), 14 (part.), 15 (part.), 17, 18 (part.)
66	Divisions nos 5 (part.), 9, 12 (part.)
67	R.M.R. de Winnipeg, 20
68	Divisions nos 12 (part.), 15 (part.), 16, 18 (part.), 19 (part.)

APPENDIX – 6.3 – ANNEXE

MANITOBA

1971 SRU List, by ER and Population

Liste des U.A.R. de 1971 par R.É. et selon la population

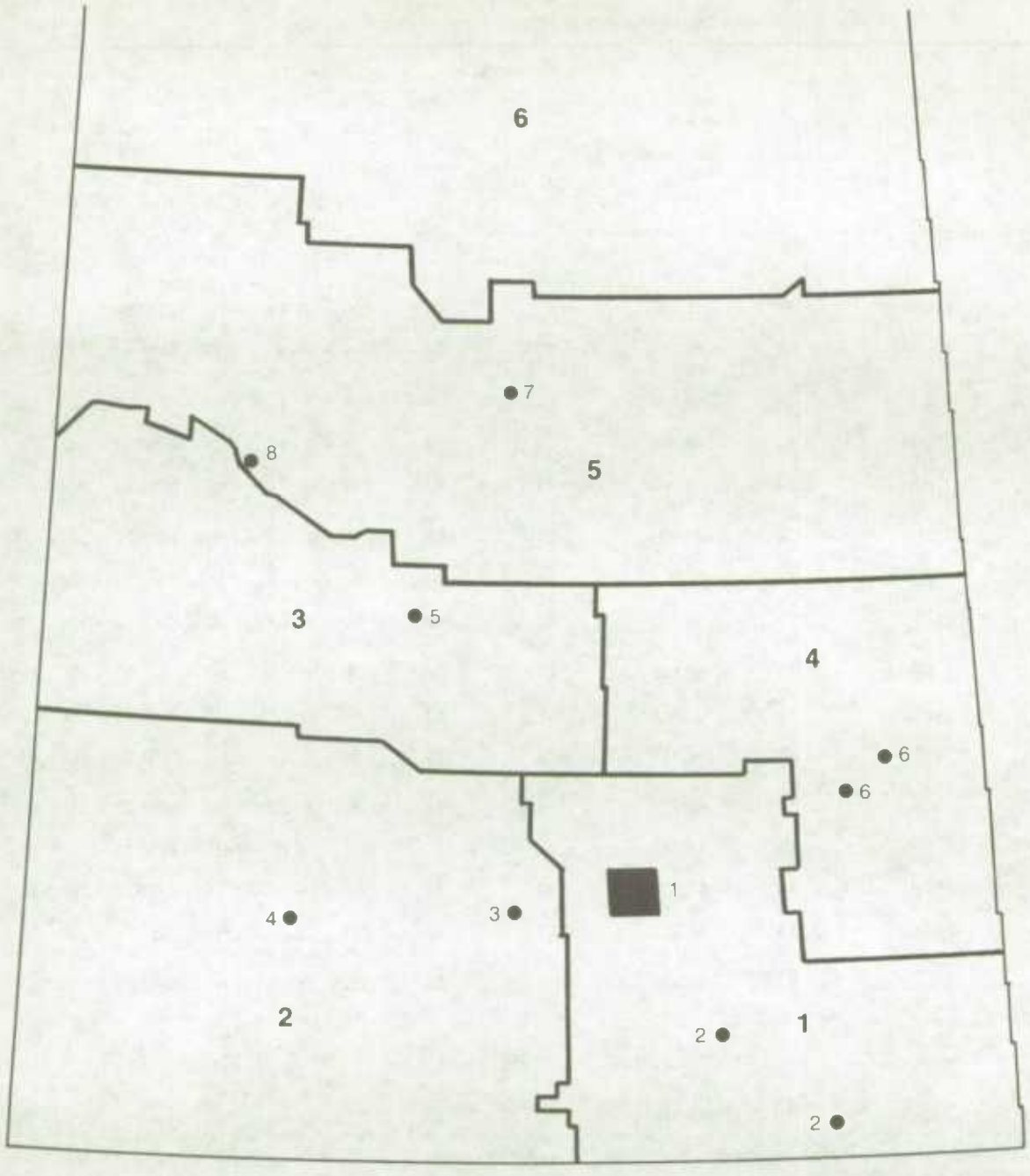
Serial No.[1] – N° de série[1]	Name of SRU – Nom de l'U.A.R.	1971 population – Population en 1971	Number of subunits – Nombre de sous-unités	Numerical identification – Code d'identification	Economic region – Région économique
1	Brandon	31,150	02	63101	63
2	Portage la Prairie[1]	12,950	01	64101	64
3	Winnipeg	540,262	01	66101	
			29	67101	66
			02 Apt. – App.	67130	67
4	Thompson[1]	19,001	01	68101	68
5	Flin Flon[1] and – et The Pas[2]	15,406	01	68201	68
	Total	618,769	–	–	–

[1] As on the map. – Voir la carte.

[2] SRU crosses Saskatchewan boundary but is considered in Manitoba (Flin Flon has 8073 in Manitoba and 471 in Saskatchewan. The Pas (6062) is in Manitoba). – L'U.A.R. dépasse les frontières de la province mais est considérée comme faisant partie du Manitoba (Flin Flon a la partie 8073 au Manitoba et la partie 471 en Saskatchewan. Le Pas (6062) se trouve au Manitoba).

Appendix 7.1 **Annexe 7.1**

Saskatchewan



Legend	Légende
Economic Region	Région économique
Principal Self-Representing Unit	Principale unité autoreprésentative
Other Self-Representing Units ● 4	Autres unités autoreprésentatives

APPENDIX – 7.2 – ANNEXE

SASKATCHEWAN

Composition of Non-self-representing Units
Composition des unités non autoreprésentatives

ER – R.É.	Stratum – Strate	Number of PSUs – Nombre d'U.P.E.	NSRU Population – Population des U.N.A.R.	Population		LF characteristics[1] used in stratification and % of labour force – Caractéristiques de la P.A.[1] pour la stratification et % de la population active		
				Rural – Rurale	Urban – Urbaine	(1)	(2)	(3)
				%				
71	71000	12	40,484	73	27	W – B 49	CHS – BPM 30	SG – CS 11
	71020	12	40,451	70	30	W – B 65	CHS – BPM 14	SG – CS 12
T	2	24	80,935	72	28	W – B 57	CHS – BPM 22	SG – CS 11
72	72000	12	40,550	69	31	W – B 56	CHS – BPM 32	SG – CS 6
	72020	14	42,954	75	25	W – B 74	CHS – BPM 13	SG – CS 9
T	2	26	83,504	72	28	W – B 65	CHS – BPM 22	SG – CS 8
73	73000	12	41,021	66	34	W – B 39	SG – CS 25	CHS – BPM 26
	73020	12	39,807	60	40	W – B 63	SG – CS 18	CHS – BPM 14
T	2	24	80,828	63	36	W – B 51	SG – CS 22	CHS – BPM 20
74	74000	16	55,019	58	42	W – B 25	CHS – BPM 32	SG – CS 26
	74020	13	41,775	71	29	W – B 41	CHS – BPM 20	SG – CS 24
T	2	29	96,794	64	36	W – B 32	CHS – BPM 27	SG – CS 25
75	75000	14	46,836	71	29	SG – CS 30	CHS – BPM 46	W – B 13
	75020	16	53,276	72	28	SG – CS 35	CHS – BPM 32	W – B 17
	75040	13	46,585	59	41	SG – CS 62	CHS – BPM 21	W – B 10
T	3	43	146,697	68	32	SG – CS 42	CHS – BPM 33	W – B 14
76	Remote – Région éloignée							
T	11	146	488,758	68	32	–	–	–

[1] Cattle, hog, sheep (CHS), small grain (SG), wheat (W). – Bovins, porcs, moutons (BPM), céréales secondaires (CS), blé (B).

ER	Composition
71	Divisions 1, 2, 6
72	Divisions 3, 4, 7, 8
73	Divisions 11, 12, 13
74	Divisions 5, 9, 10
75	Divisions 14, 15, 16, 17
76	Division 18

R.É.	Composition
71	Divisions n°s 1, 2, 6
72	Divisions n°s 3, 4, 7, 8
73	Divisions n°s 11, 12, 13
74	Divisions n°s 5, 9, 10
75	Divisions n°s 14, 15, 16, 17
76	Division n° 18

APPENDIX – 7.3 – ANNEXE

SASKATCHEWAN

1971 SRU List, by ER and Population

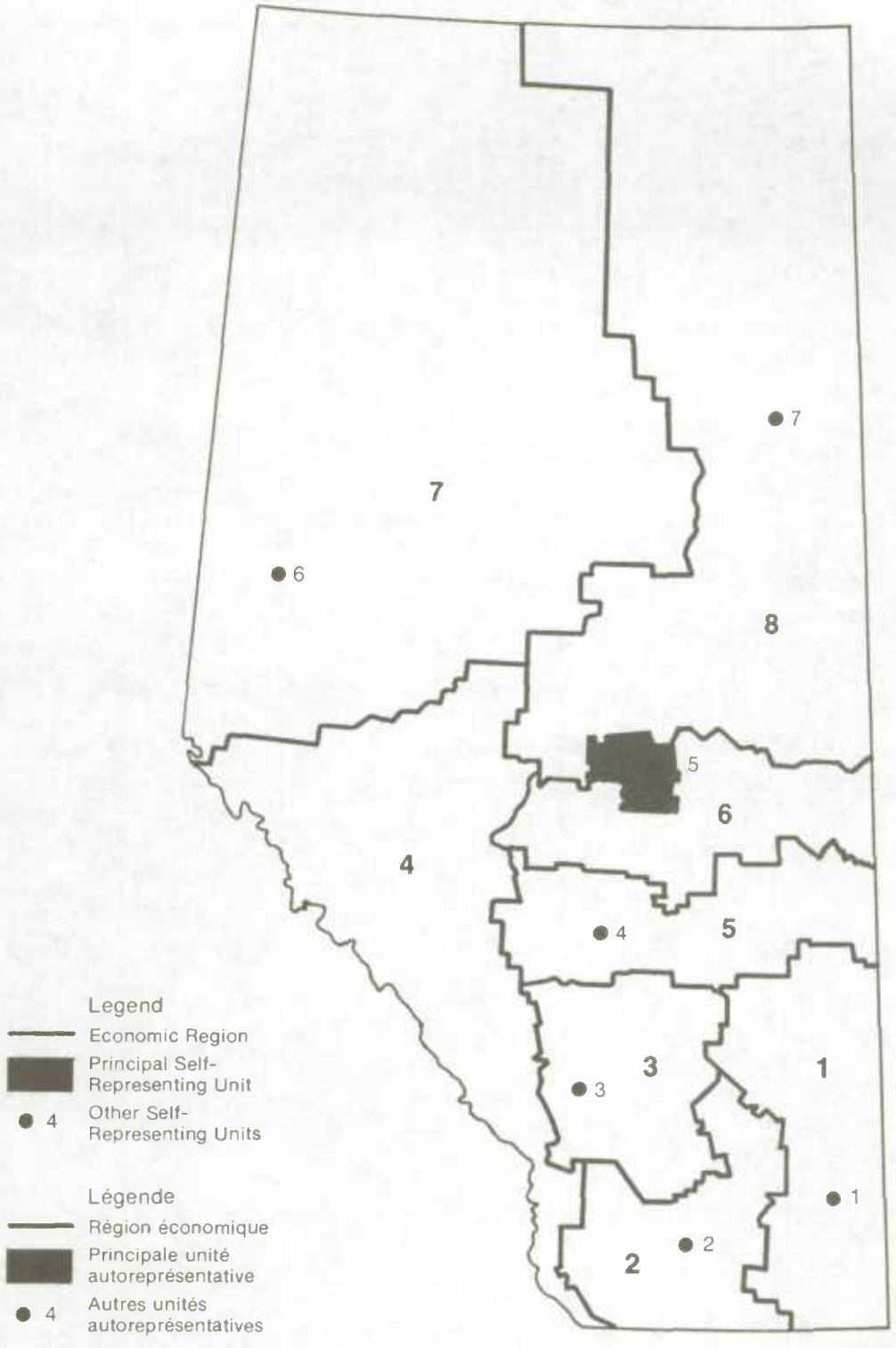
Liste des U.A.R. de 1971 par R.É. et selon la population

Serial No.[1] — Nº de série[1]	Name of SRU — Nom de l'U.A.R.	1971 population — Population en 1971	Number of subunits — Nombre de sous-unités	Numerical identification — Code d'identification	Economic region — Région économique
1	Regina	140,734	10	71101	71
2	Estevan - Weyburn[1]	17,965	01	71201	71
3	Moose Jaw	31,854	02	72101	72
4	Swift Current[1]	15,415	01	72201	72
5	Saskatoon	126,449	07 01 Apt. – App.	73101 73108	73
6	Yorkton and – et Melville[1]	18,805	01	74101	74
7	Prince Albert	28,464	02	75101	75
8	North Battleford[1]	13,345	01	75201	75
	Total	**393,031**	–	–	–

[1] As on the map. – Voir la carte.

Appendix 8.1 **Annexe 8.1**

Alberta



APPENDIX – 8.2 – ANNEXE

ALBERTA

Composition of Non-self-representing Units

Composition des unités non autoreprésentatives

ER – R.É.	Stratum – Strate	Number of PSUs – Nombre d'U.P.É.	NSRU population – Population des U.N.A.R.	Population		LF characteristics[1] used in stratification and % of labour force – Caractéristiques de la P.A.[1] pour la stratification et % de la population active		
				Rural – Rurale	Urban – Urbaine	(1)	(2)	(3)
				%				
81	81000	10	23,367	75	25	CHS – BPM 53	W – B 29	SG – CS 9
T	1	10	23,367	75	25	CHS – BPM 53	W – B 29	SG – CS 9
82	82000	9	31,282	53	47	CHS – BPM 46	SG – CS 16	W – B 17
	82020	12	39,687	56	44	CHS – BPM 59	SG – CS 16	W – B 9
T	2	21	70,969	55	45	CHS – BPM 53	SG – CS 16	W – B 13
83	83000	10	31,682	70	30	CHS – BPM 36	SG – CS 35	W – B 21
	83020	13	42,550	63	37	CHS – BPM 67	SG – CS 16	W – B 2
T	2	23	74,232	66	34	CHS – BPM 54	SG – CS 24	W – B 10
84	84000	12	39,297	36	64	CHS – BPM 77	SG – CS 4	W – B -
T	1	12	39,297	36	64	CHS – BPM 77	SG – CS 4	W – B -
85	85000	16	52,935	59	41	CHS – BPM 62	SG – CS 16	W – B 10
	85020	12	39,531	65	35	CHS – BPM 73	SG – CS 10	W – B 1
T	2	28	92,466	62	38	CHS – BPM 67	SG – CS 13	W – B 6
86	86000	18	60,343	60	40	CHS – BPM 59	SG – CS 16	W – B 6
	86020	18	59,457	56	44	CHS – BPM 58	SG – CS 15	W – B 3
T	2	36	119,800	58	42	CHS – BPM 59	SG – CS 16	W – B 5
87	87000	12	40,277	54	46	SG – CS 40	CHS – BPM 34	W – B 5
	87020	9	29,427	68	32	SG – CS 59	CHS – BPM 14	W – B 11
T	2	21	69,704	60	40	SG 48	CHS – BPM 26	W – B 8
88	88000	14	45,474	73	27	CHS – BPM 63	SG – CS 19	W – B 2
	88020	10	33,228	62	38	CHS – BPM 76	SG – CS 8	W – B 2
T	2	24	78,702	68	32	CHS – BPM 68	SG – CS 14	W – B 2
Prov.	14	175	568,537	60	40	–	–	–

[1] Cattle, hogs, sheep (CHS), small grain (SG), wheat (W). – Bovins, porcs, moutons (BPM), céréales secondaires (CS), blé (B).

ER	Composition
81	Divisions 1, 4
82	Divisions 2, 3
83	Divisions 5, 6
84	Divisions 9, 14
85	Divisions 7, 8
86	Divisions 10, 11
87	Division 15
88	Divisions 12, 13

R.É.	Composition
81	Divisions nos 1, 4
82	Divisions nos 2, 3
83	Divisions nos 5, 6
84	Divisions nos 9, 14
85	Divisions nos 7, 8
86	Divisions nos 10, 11
87	Division no 15
88	Divisions nos 12, 13

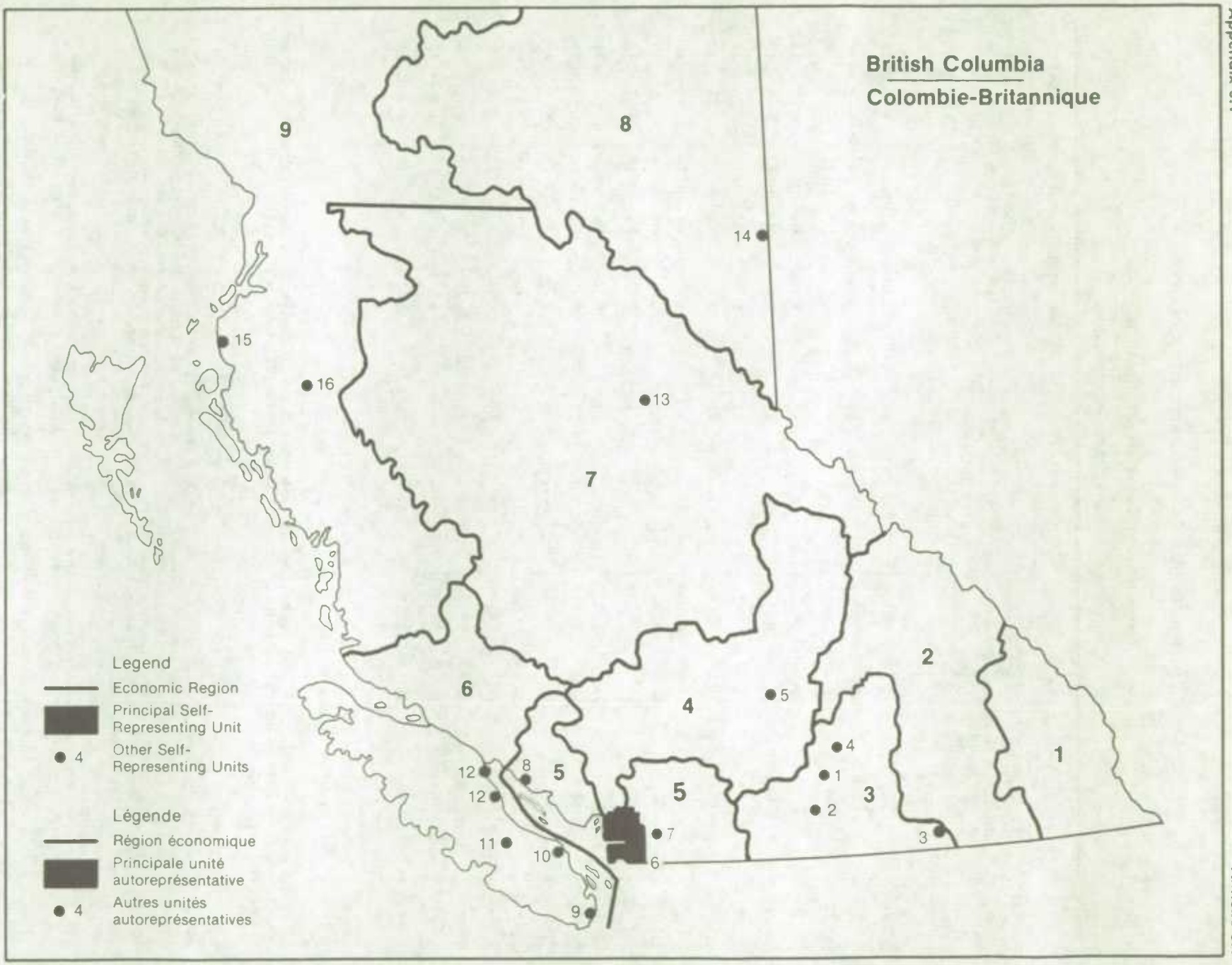
APPENDIX – 8.3 – ANNEXE

ALBERTA

1971 SRU List, by ER Population
Liste des U.A.R. de 1971 par R.É. selon la population

Serial No.[1] – N° de série[1]	Name of SRU – Nom de l'U.A.R.	1971 population – Population en 1971	Number of subunits – Nombre de sous-unités	Numerical identification – Code d'identification	Economic region – Région économique
1	Medicine Hat	28,773	02	81101	81
2	Lethbridge	41,217	03	82101	82
3	Calgary	403,319	25 02 Apt. – App.	83101 83126	83
4	Red Deer	27,674	02	85101	85
5	Edmonton	495,310	28 02 Apt. – App.	86101 86129	86
6	Grande Prairie[1]	13,079	01	87101	87
7	Fort McMurray[1]	6,847	03	88101	88
	Total	**1,016,219**	–	–	–

[1] As on the map. – Voir la carte.


British Columbia
Colombie-Britannique
9
8
7
6
5
4
3
2
1
14
15
16
13
12
12
8
11
10
9
7
6
5
4
1
2
3
Legend
Economic Region
Principal Self-Representing Unit
4
Other Self-Representing Units
Légende
Région économique
Principale unité autoreprésentative
4
Autres unités autoreprésentatives

APPENDIX – 9.2 – ANNEXE

BRITISH COLUMBIA – COLOMBIE-BRITANNIQUE

Composition of Non-self-representing Units

Composition des unités non autoreprésentatives

ER – R.É.	Stratum – Strate	Number of PSUs – Nombre D'U.P.É.	NSRU population – Population des U.N.A.R.	Population		LF characteristics[1] used in stratification and % of labour force – Caractéristiques de la P.A.[1] pour la stratification et % de la population active		
				Rural – Rurale	Urban – Urbaine	(1)	(2)	(3)
				%				
91	91000	11	39,404	31	69	S 48	Mi 14	Ma – IM 15
T	1	11	39,404	31	69	S 48	Mi 14	Ma – IM 15
92	92000	19	74,957	56	44	S 51	Ma – IM 16	T 14
T	1	19	74,957	56	44	S 51	Ma – IM 16	T 14
93	93000	18	67,453	75	25	S 51	Ma – IM 15	A 13
T	1	18	67,453	75	25	S 51	Ma – IM 15	A 13
94	94000	11	40,048	65	35	S 42	Ma – IM 16	T 10
T	1	11	40,048	65	35	S 42	Ma – IM 16	T 10
95	95000	10	37,757	84	16	S 45	Ma – IM 14	A 10
	95020	11	40,993	76	24	S 55	Ma – IM 11	A 11
T	2	21	78,750	80	20	S 50	Ma – IM 12	A 11
96	96000	12	45,587	70	30	S 47	F – FP 14	Ma – IM 22
	96020	12	45,653	72	28	S 38	F – FP 29	Ma – IM 18
T	2	24	91,240	71	29	S 42	F – FP 22	20
97	97000	19	77,198	64	36	S 43	F – FP 22	T 9
T	1	19	77,198	64	36	S 43	Ma – IM 22	T 9
98	98000	12	23,187	54	46	S 47	T 14	C – BT 13
T	1	12	23,187	54	46	S 47	T 14	C – BT 13
99	Remote – Régions éloignée							
T	10	135	492,237	65	35	–	–	–

[1] Services (S), manufacturing (Ma), construction (C), forestry and fisheries (F), mining (Mi), transportation (T), agriculture (A). – Services (S), industries manufacturières (IM), bâtiment et travaux publics (BT), forêts et pêche (FP), industrie minière (Mi), transports (T), agriculture (A).

ER	Composition
91	Division of East Kootenay
92	Divisions of Central Kootenay, Columbia - Shuswap
93	Divisions of North Okanagan, Okanagan - Similkameen, Kootenay Boundary, Central - Okanagan
94	Divisions of Thompson - Nicola, Squamish - Lillooet
95	Divisions of Central Fraser Valley, Dewdney - Alouette, Power River, Sunshine Coast, Fraser - Cheam, Greater Vancouver
96	Divisions of Capital, Cowichan - Valley, Alberni - Clayoquot, Comox - Strathcona, Mount Waddington, and Nanaimo
97	Bulkley - Nechako, Cariboo and Fraser - Fort George
98	Divisions of Peace River - Liard
99	Divisions of Kitimat - Stikine, Ocean Falls, Skeena A, Stikine

R.É.	Composition
91	Division de East Kootenay
92	Divisions de Central Kootenay, Columbia - Shuswap
93	Divisions de North Okanagan, Okanagan - Similkameen, Kootenay Boundary, Central - Okanagan
94	Divisions de Thompson - Nicola, Squamish - Lillooet
95	Divisions de Central Fraser Valley, Dewdney - Alouette, Power River, Sunshine Coast, Fraser - Cheam, Greater Vancouver
96	Divisions de Capital, Cowichan - Valley, Alberni - Clayoquot, Comox, Strathcona, Mount Waddington, et Nanaimo
97	Bulkley - Nechko, Cariboo et Fraser Fort George
98	Divisions de Peace River - Liard
99	Divisions de Kitimat - Stikine, Ocean Falls, Skeena A, Stikine

APPENDIX – 9.3 – ANNEXE

BRITISH COLUMBIA – COLOMBIE-BRITANNIQUE

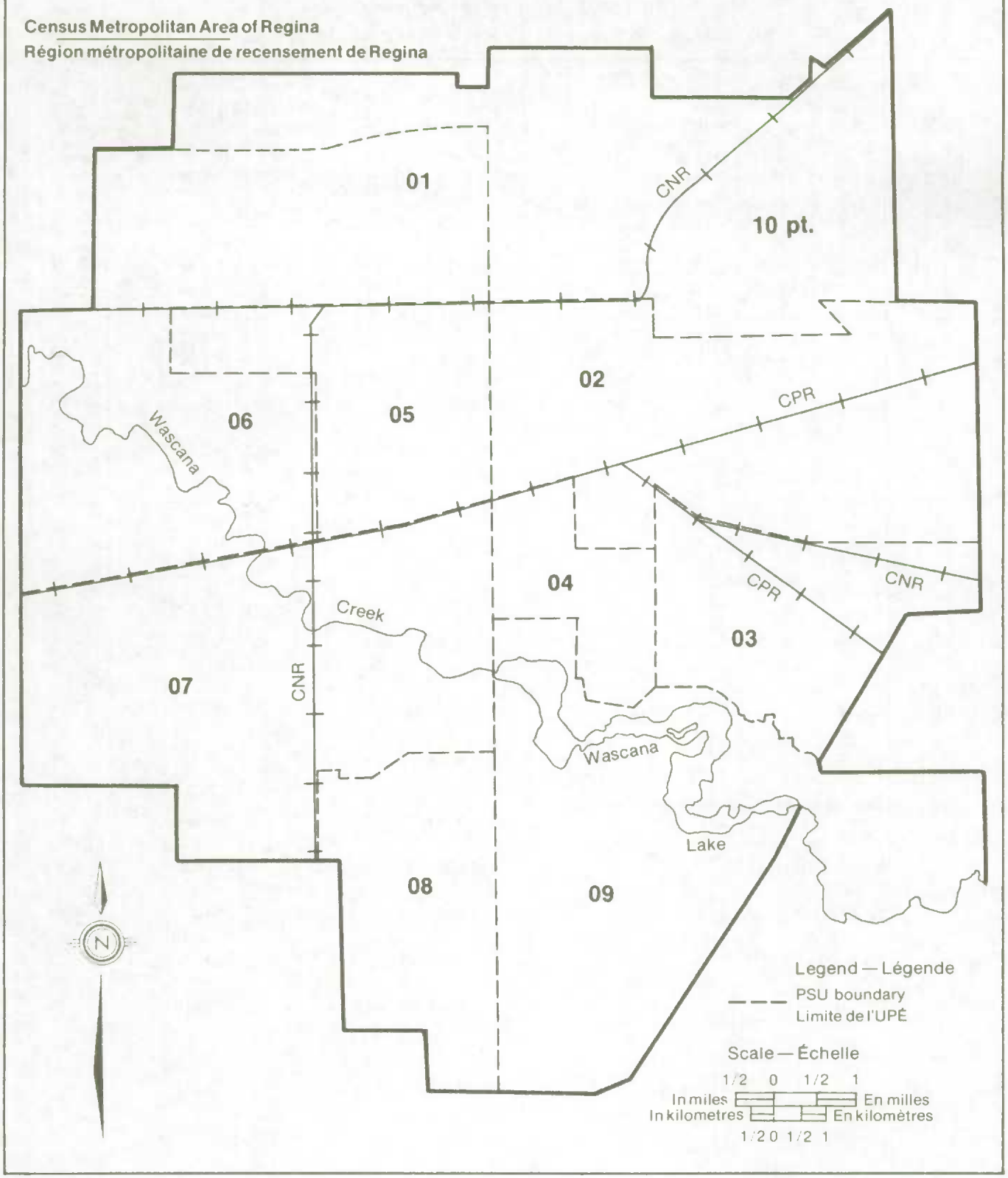
1971 SRU List, by ER and Population
Liste des U.A.R. de 1971 par R.É. et selon la population

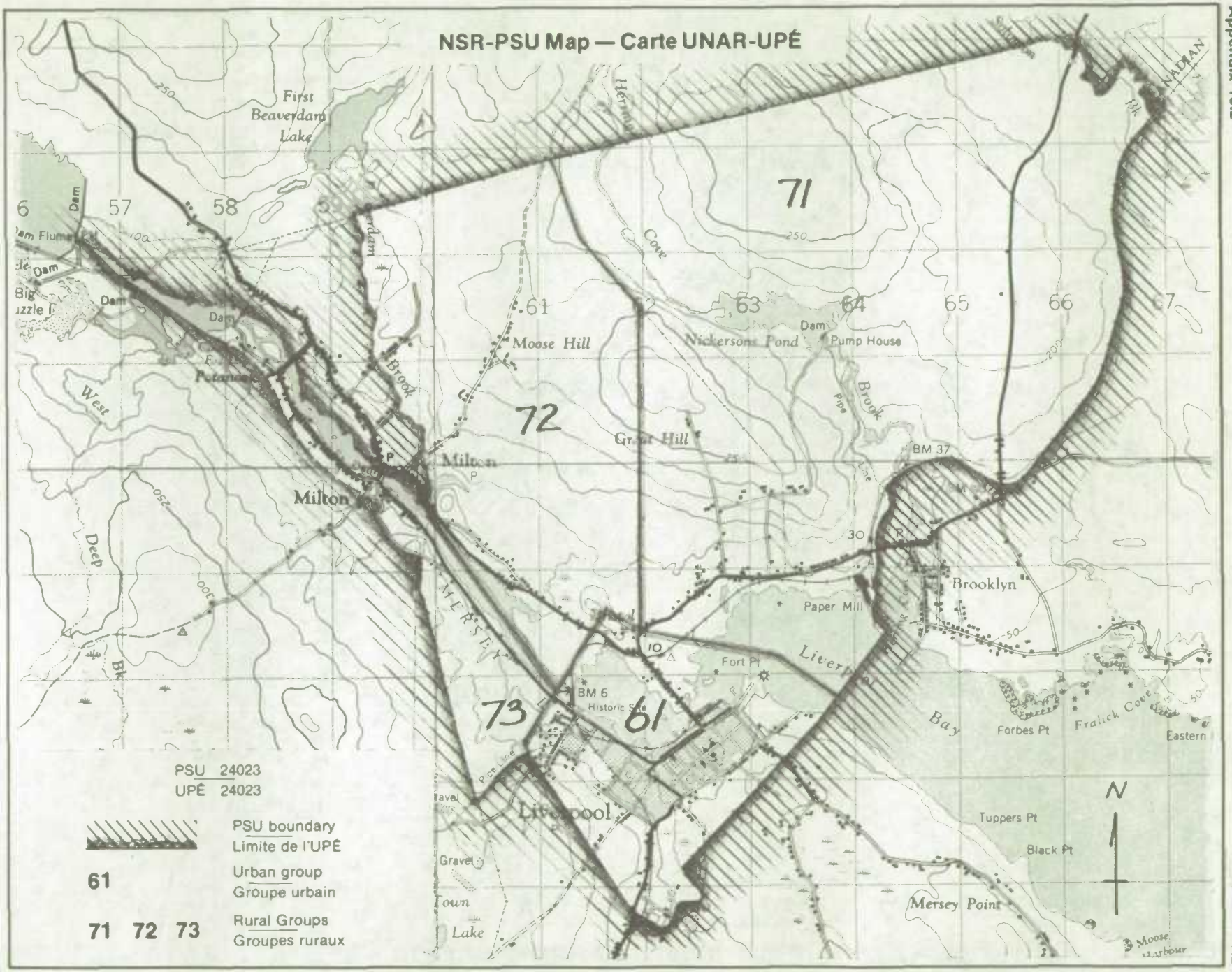
Serial No.[1] – Nº de série[1]	Name of SRU – Nom de l'U.A.R.	1971 Population – Population en 1971	Number of subunits – Nombre de sous-unités	Numerical identification – Code d'identification	Economic Region – Région économique
1	Kelowna	36,920	03	93101	93
2	Penticton	18,146	01	93201	93
3	Rossland and – et Trial	17,177	01	93301	93
4	Vernon	16,885	01	93401	93
5	Kamloops	43,663	03	94101	94
6	Vancouver	1,080,602	61 02 Apt. – App.	95101 95163	95
7	Chilliwack and et Mission City	19,355	01	95201	95
8	Powell River	15,228	01	95301	95
9	Victoria	194,703	12 01 Apt. – App.	96101 96113	96
10	Nanaimo	38,384	03	96201	96
11	Port Alberni	26,427	02	96301	96
12	Courtenay and – et Comox[1], Campbell River	20,944	02	96,401	96
13	Prince George	49,100	03	97101	97
14	Dawson Creek	11,885	01	98101	98
15	Prince Rupert	16,766	01	99101	99
16	Kitimat	11,803	01	99201	99
	Total	1,617,988			

[1] As on the map. – Voir la carte.

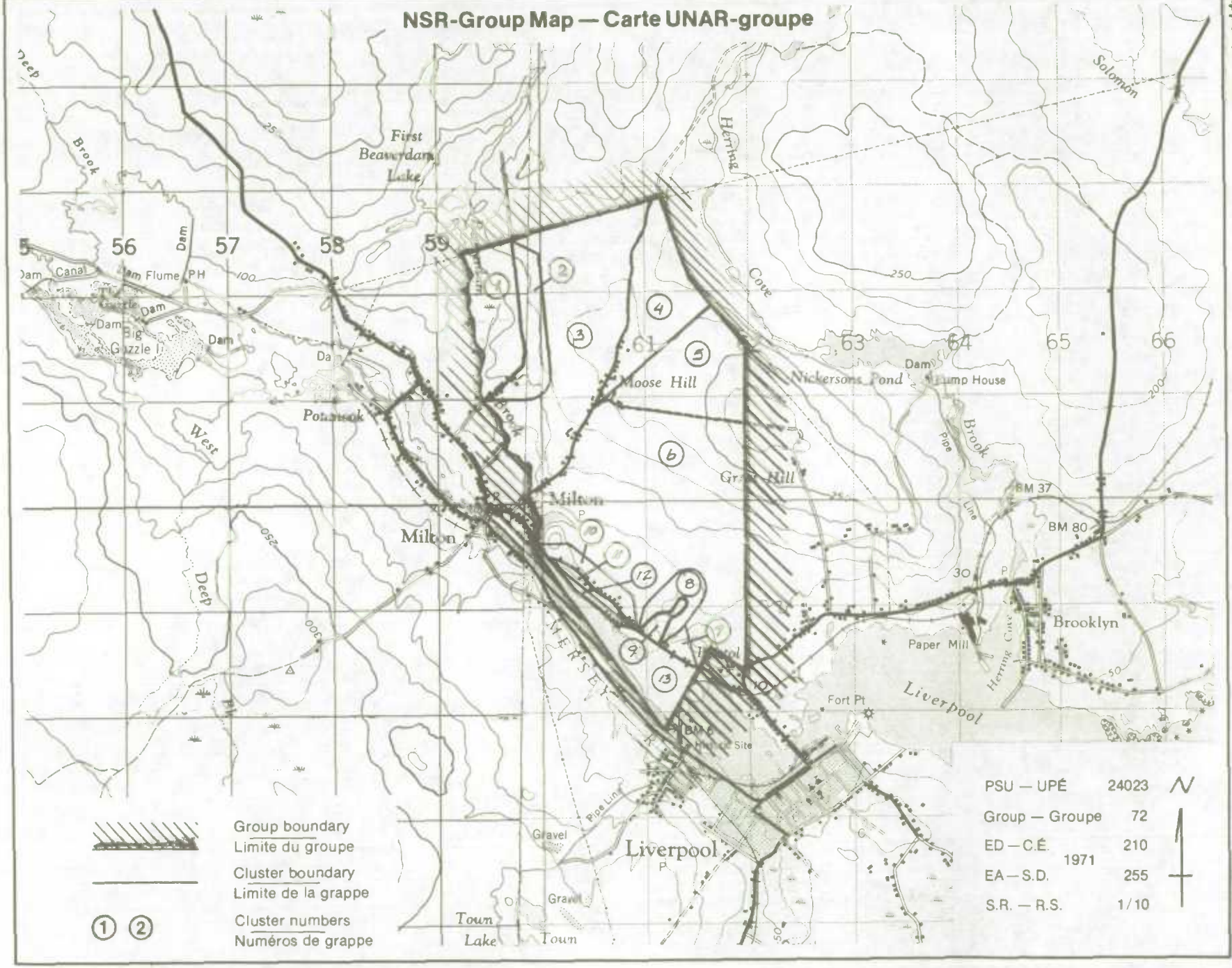
Appendix 10.1 Annexe 10.1

Map for Regina SRU — Carte de l'UAR de Regina




NSR-PSU Map — Carte UNAR-UPÉ
First Beaverdam Lake
Moose Hill
Nickersons Pond
Pump House
Great Hill
Milton
Brooklyn
Paper Mill
Fort Pt
Liverpool
Bay
Forbes Pt
Fralick Cove
Eastern
Tuppers Pt
Black Pt
Mersey Point
Moose Harbour
Town Lake
Gravel
Deep
West
Brook
Cove
MERSEY
BM 6
BM 37
Historic Site
Pipe Line
Dam
Flume
Big Puzzle I
71
72
73
61
57
58
63
64
65
66
N
PSU 24023
UPÉ 24023
PSU boundary
Limite de l'UPÉ
61
Urban group
Groupe urbain
71 72 73
Rural Groups
Groupes ruraux

NSR-Group Map — Carte UNAR-groupe



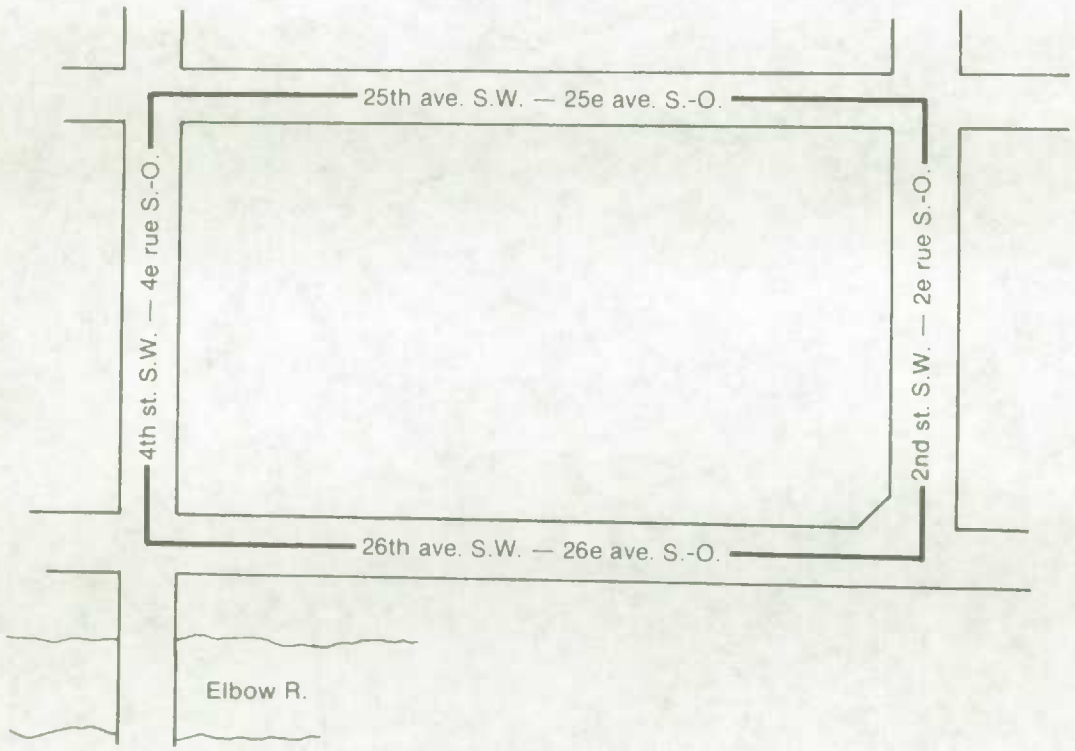
Appendix 10.4 **Annexe 10.4**

Statistics Canada / Statistique Canada

CLUSTER DIAGRAM
SCHÉMA DE LA GRAPPE

1 FORM NO. FORMULE 01

PSU UPE 2	83101	GROUP GROUPE 3	04	CLUSTER GRAPPE 4	095
ROTATION NO. Nº DE RENOUVELLEMENT 5	2	COUNT COMPTE 6	036		
SURVEY IDENTIFICATION IDENTIFICATION DE L'ENQUÊTE 7	1	RANDOM START ORIGINE CHOISIE AU HASARD 8	006	SAMPLING RATIO FRACTION DE SONDAGE 9	008
DATE OF INTRODUCTION DATE D'INTRODUCTION 10	10 (MO. MO.) 76 (YR. AN.)				



LOCATION ENDROIT: CALGARY C

6-4000-471
12/10/76

Appendix 10.5

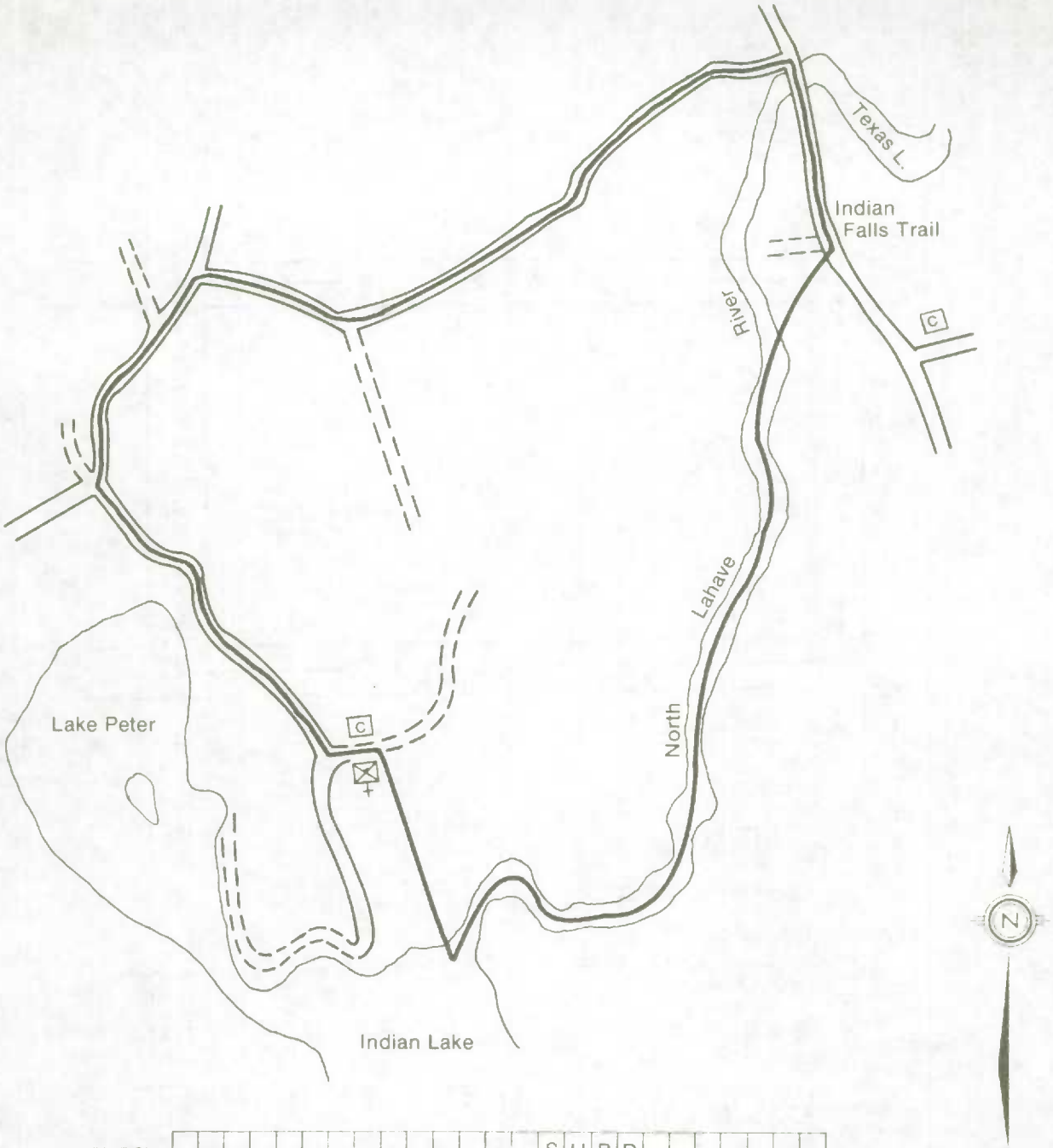
Annexe 10.5

Statistics Canada / Statistique Canada

CLUSTER DIAGRAM
SCHÉMA DE LA GRAPPE

1 FORM NO / FORMULE 01

PSU / UPE	2	2 4 0 2 8	GROUP / GROUPE	3	7 3	CLUSTER / GRAPPE	4	0 1 7	ROTATION NO / Nº DE RENOUVELLEMENT	5	3
COUNT / COMPTE	6	0 1 2	SURVEY IDENTIFICATION / IDENTIFICATION DE L'ENQUÊTE	7	1	RANDOM START / ORIGINE CHOISIE AU HASARD	8	0 0 1	SAMPLING RATIO / FRACTION DE SONDAGE	9	0 0 4
DATE OF INTRODUCTION / DATE D'INTRODUCTION	10	0 9 7 6 (MO. MO. / YR. AN)									



LOCATION / ENDROIT: LUNENBURG MUN SUBD LOT A

6–4000–471
12/10/76

APPENDIX – 11.1 – ANNEXE

Frequency Distribution of Clusters by their Size in SRUs (Built-up)

Distribution de la fréquence des grappes selon leur taille dans les U.A.R. (zones bâties)

Provinces / Sizes – Taille	0	1	2	3	4	5	6	7	8	9	Total
0- 9	0	0	0	0	0	0	0	0	0	0	0
10-19	145	69	147	181	991	1,596	311	328	509	824	5,101
20-29	488	116	741	498	4,195	8,128	1,112	1,331	2,598	4,035	23,242
30-39	353	49	553	431	3,456	6,770	1,007	792	1,741	2,898	18,050
40-49	182	28	324	270	2,377	3,947	878	344	1,045	1,498	10,893
50-59	136	14	197	193	1,768	2,657	539	166	700	970	7,340
60-69	85	4	137	68	1,636	1,902	359	97	363	729	5,380
70-79	23	1	50	20	1,375	1,316	88	49	149	327	3,398
80-89	10	0	14	4	1,131	1,021	20	16	73	151	2,440
90+	12	0	14	8	2,538	2,086	25	29	136	215	5,063
Total	1,434	281	2,177	1,673	19,467	29,423	4,339	3,152	7,314	11,647	80,907

APPENDIX – 11.2 – ANNEXE

Frequency Distribution of Clusters by their Size in SRUs (Fringe)
Distribution de la fréquence des grappes selon leur taille dans les U.A.R. (banlieue)

Provinces / Sizes – Taille	0	1	2	3	4	5	6	7	8	9	Total
0- 9	4	1	25	6	105	183	8	19	22	35	408
10-19	160	57	343	327	2,860	3,774	252	285	576	1,052	9,686
20-29	105	24	247	218	1,783	2,304	140	192	209	687	5,909
30-39	60	13	163	123	942	1,180	43	63	118	401	3,106
40-49	35	14	96	79	558	701	42	43	93	274	1,935
50-59	27	3	94	64	345	442	23	31	70	178	1,277
60-69	17	1	62	15	272	340	14	14	45	148	928
70-79	8	1	26	11	212	246	5	6	20	85	620
80-89	0	1	9	5	154	196	4	3	15	48	435
90+	1	0	18	8	384	528	6	4	23	79	1,051
Total	417	115	1,083	856	7,615	9,894	537	660	1,191	2,987	25,355

APPENDIX – 12.1 – ANNEXE

Population Summary by Province
Aperçu de la population par province

Province	Population				
	SRU – U.A.R.	NSRU – U.N.A.R.	Special areas – Régions spéciales	Indian Reserves – Réserves indiennes	Total
Newfoundland – Terre-Neuve	280,551	232,741	10,779	–	524,071
Prince Edward Island – Île-du-Prince-Édouard	39,257	71,704	–	326	111,287
Nova Scotia – Nouvelle-Écosse	415,247	361,731	8,576	3,406	788,960
New Brunswick – Nouveau-Brunswick	284,095	335,952	11,129	3,381	634,557
Québec	4,275,959	1,674,684	54,433	22,688	6,027,764
Ontario	6,019,358	1,602,068	46,251	35,429	7,703,106
Manitoba	618,769	313,932	31,019	24,998	988,718
Saskatchewan	393,031	488,758	17,474	23,802	923,065
Alberta	1,016,219	568,537	28,002	21,963	1,634,721
British Columbia – Colombie-Britannique	1,617,988	492,237	42,692	31,704	2,184,621
Canada[1]	14,960,474	6,142,344	250,355	167,697	21,520,870

[1] Excluding Yukon and Northwest Territories. – Sans le Yukon et les Territoires du Nord-Ouest.

APPENDIX – 13.1 – ANNEXE

Statistics Canada Statistique Canada

HOUSEHOLD RECORD DOCKET

CONFIDENTIAL when completed

1 FORM NO. 03

2 Docket No. | 3 Survey date | 4 Assignment No. | 5 Designated Interviewer no. (No change 1 / or) | Your interviewer no.

6 P.S.U. | Group | Cluster | Rot. No. | 7 Listing | Mult | 8 Type of dwelling (Enter code)

9 *Record time of every call on this household.*

Mon.	:	:	:	:	:	
Tues.	:	:	:	:	:	
Wed.	:	:	:	:	:	
Thur.	:	:	:	:	:	
Fri.	:	:	:	:	:	
Sat.	:	:	:	:	:	

10 Listing address

11 Mailing address

12 INTERVIEWER CHECK ITEM

If unable to make contact with a responsible household member, specify reason in NOTES and enter code in 41.

13 Telephone no. –

OR No telephone 1 ○ Go to 15

OR Telephone no. refused 2 ○ Go to 15

14 INTERVIEWER CHECK ITEM

1 ☐

- *If "T" in above box, request permission to interview by phone*

Permission granted 2 ○ Go to 15

Permission denied 3 ○ Go to 15

- *Otherwise* 4 ○ *go to 15*

15 *Determine and record the best time to call on this household.*

20 INTERVIEWER CHECK ITEM

- *If first interview at this dwelling or new household since last interview*

1 ○ *go to 21*

- *If a subsequent interview with the same household*

2 ○ *go to 25*

21 WHAT ARE THE NAMES OF ALL PERSONS NOW LIVING OR STAYING AT THIS ADDRESS WHO HAVE NO USUAL PLACE OF RESIDENCE ELSEWHERE?

Enter names in 32

22 ARE THERE ANY PERSONS AWAY FROM THIS HOUSEHOLD ATTENDING SCHOOL, VISITING, TRAVELLING OR IN HOSPITAL WHO USUALLY LIVE HERE?

1 Yes ○ *Enter names in 32* 2 No ○

23 DOES ANYONE ELSE LIVE AT THIS ADDRESS SUCH AS OTHER RELATIVES, ROOMERS, BOARDERS, OR EMPLOYEES?

1 Yes ○ *Enter names in 32* 2 No ○

24 *Enter answers for 33 through 40 for each person recorded in 32, then go to 41*

25 DO THE FOLLOWING PERSONS STILL LIVE OR STAY HERE?

- *Read names of all persons recorded in 32 and*
- *Enter appropriate code in 40*

26 DOES ANYONE ELSE NOW LIVE OR STAY AT THIS ADDRESS SUCH AS OTHER RELATIVES, ROOMERS, BOARDERS, ETC.?

1 Yes ○ *Enter answers for 32 through 40 for all new household members.*

2 No ○ *Go to 41*

31 HRD Pg	Ln	32 Names of household members	33 Age	34 Sex	35 Ms	36 Fe mld.	37 Rel to H	38 Educ. 1	38 Educ. 2	40 Membership	50 Supplementary questions Ln	a	b	c	d	e	f	g	h	i	j	k	l	m
	1	given name / surname									1													
	2	given name / surname									2													
	3	given name / surname									3													
	4	given name / surname									4													
	5	given name / surname									5													
	6	given name / surname									6													
	7	given name / surname									7													
	8	given name / surname									8													

41 Household Response

month

response

42 IS THIS DWELLING OWNED BY A MEMBER OF THIS HOUSEHOLD?

1 Yes ○ *Go directly to* FORMS NO. 05 & 06.

2 No ○ *Complete* FORM NO. 04

43 Forms Control

	04	05	06
form			
printed			
completed			

44 NOTES

Ln	Retain	Item no.	See over for additional NOTES.
1	1 ○		
2	2 ○		
3	3 ○		
4	4 ○		

6-4000-473: 18-10-76

FRANÇAIS AU VERSO

"Authority – Statistics Act, Chapter 15 Statutes of Canada 1970 - 71 - 72"

APPENDIX 13.1 – Concluded – ANNEXE 13.1 – fin

Statistics Canada Statistique Canada

DOSSIER DU MÉNAGE DOCUMENT CONFIDENTIEL une fois rempli

1 FORMULE 03

2 N° du dossier | 3 Date d'enquête | 4 N° de tâche | 5 N° de l'interviewer désigné (Aucun changement 1 ou) | Votre n° d'interviewer

6 U.P.É. | Groupe | Grappe | N° de renouv. | 7 N° de liste | Mult | 8 Genre de logement — *Inscrivez le code*

9 *Notez l'heure de chaque visite ou appel à ce ménage*

Lun.	:	:	:	:	:	:
Mar.	:	:	:	:	:	:
Mer.	:	:	:	:	:	:
Jeu.	:	:	:	:	:	:
Ven.	:	:	:	:	:	:
Sam.	:	:	:	:	:	:

10 Adresse de lieu

11 Adresse postale

12 *À L'INTERVIEWER*

S'il vous est impossible de rejoindre un membre responsable du ménage, indiquez la raison dans les NOTES et inscrivez le code à 41

13 N° de téléphone ( ) –

ou Pas de téléphone 1 ○ } *Passez à 15*

ou N° de téléphone refusé 2 ○ } *Passez à 15*

14 *À L'INTERVIEWER*

1 ☐

• *Si "T" dans la case ci-dessus, demandez la permission d'interviewer par téléphone.*

Permission accordée 2 ○ *Passez à 15*

Permission refusée 3 ○ *Passez à 15*

• *Autrement* 4 ○ *passez à 15*

15 *Fixer le meilleur moment pour les visites ou appels futurs et notez-le ici.*

20 *À L'INTERVIEWER*

• *S'il s'agit de la première interview à ce logement ou d'un nouveau ménage depuis l'interview précédente*

1 ○ *passez à 21*

• *S'il s'agit d'une interview subséquente du même ménage*

2 ○ *passez à 25*

21 VEUILLEZ NOMMER TOUTES LES PERSONNES QUI DEMEURENT ACTUELLEMENT À CETTE ADRESSE ET QUI N'ONT PAS DE RÉSIDENCE HABITUELLE AILLEURS.

Inscrivez les noms au poste 32

23 Y-A-T-IL QUELQU'UN D'AUTRE QUI DEMEURE À CETTE ADRESSE, PAR EXEMPLE, UN PARENT, CHAMBREUR, PENSIONNAIRE OU EMPLOYÉ?

Oui 1 ○ *Inscrivez les noms à 32* Non 2 ○

22 Y A-T-IL DES PERSONNES QUI SONT ABSENTES DU MÉNAGE PARCE QU'ELLES SONT AUX ÉTUDES, EN VISITE, EN VOYAGE OU À L'HÔPITAL, MAIS QUI DEMEURENT HABITUELLEMENT ICI?

Oui 1 ○ *Inscrivez les noms à 32* Non 2 ○

24 *Inscrivez les réponses de 33 à 40 pour chaque personne indiquée au poste 32, puis passez à 41.*

25 LES PERSONNES SUIVANTES DEMEURENT-ELLES ENCORE ICI?

• *Lisez les noms de chaque personne inscrite à 32, et*

• *Inscrivez le code approprié au poste 40*

26 Y A-T-IL QUELQU'UN D'AUTRE QUI DEMEURE ACTUELLEMENT À CETTE ADRESSE, PAR EXEMPLE, UN PARENT, CHAMBREUR OU PENSIONNAIRE?

Oui 1 ○ *Inscrivez les réponses de 32 à 40 pour tout nouveau membre du ménage*

Non 2 ○ *Passez à 41.*

31 OM P	L	32 Noms des membres du ménage	33 Âge	34 Sexe	35 É.M.	36 U.F.a.m.	37 L.a.-vec le C.	38 Niv. d'inst. [1] [2]	40 Composition	50 Questions supplémentaires L	a	b	c	d	e	f	g	h	i	j	k	l	m
	1	Prénom / Nom de famille								1													
	2	Prénom / Nom de famille								2													
	3	Prénom / Nom de famille								3													
	4	Prénom / Nom de famille								4													
	5	Prénom / Nom de famille								5													
	6	Prénom / Nom de famille								6													
	7	Prénom / Nom de famille								7													
	8	Prénom / Nom de famille								8													

41 Code-réponse

mois

code-réponse

42 CE LOGEMENT APPARTIENT-IL À UN MEMBRE DE CE MÉNAGE?

Oui 1 ○ Non 2 ○

Passez directement aux FORMULES 05 & 06 — *Remplissez la* FORMULE 04

43 Contrôle des formules

formule	04	05	06
imprimées			
remplies			

44 NOTES

L	À	N° du poste
1	1 ○	
2	2 ○	
3	3 ○	
4	4 ○	

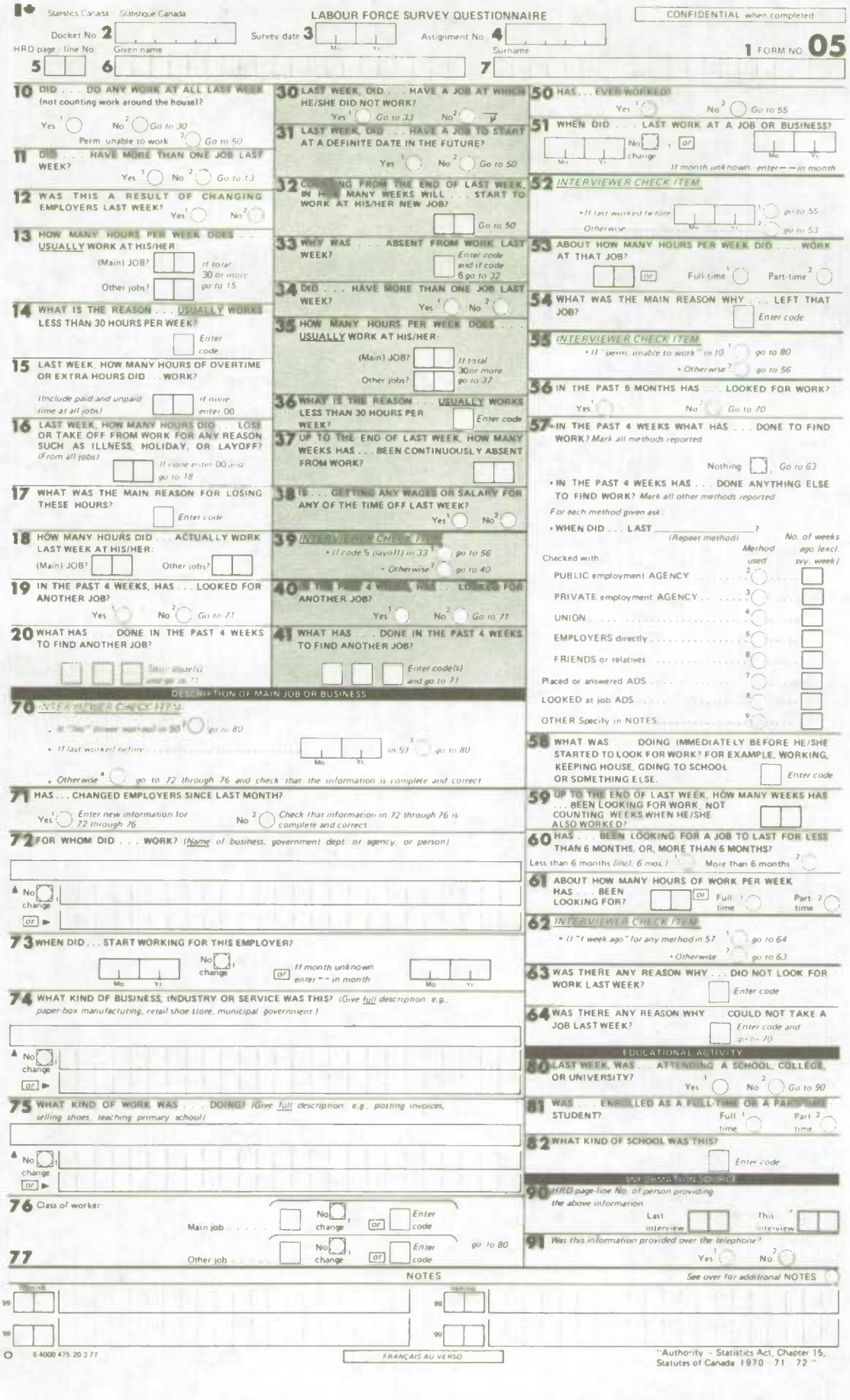
Passez au verso pour NOTES additionnelles. ○

6-4000-473 18-10-76

SEE REVERSE FOR ENGLISH

"Déclaration exigée en vertu de la Loi sur la statistique, chapitre 15, Statuts du Canada de 1970-71-72."

APPENDIX – 13.2 – ANNEXE



1 Statistics Canada – Statistique Canada

LABOUR FORCE SURVEY QUESTIONNAIRE

CONFIDENTIAL when completed

Docket No. 2 | Survey date 3 Mo Yr | Assignment No. 4

HRD page – line No. 5 | Given name 6 | Surname 7

1 FORM NO. 05

10 DID . . . DO ANY WORK AT ALL LAST WEEK (not counting work around the house)?
Yes 1 ◯ No 2 ◯ Go to 30 Perm. unable to work 3 ◯ Go to 50

11 DID . . . HAVE MORE THAN ONE JOB LAST WEEK?
Yes 1 ◯ No 2 ◯ Go to 13

12 WAS THIS A RESULT OF CHANGING EMPLOYERS LAST WEEK?
Yes 1 ◯ No 2 ◯

13 HOW MANY HOURS PER WEEK DOES . . . USUALLY WORK AT HIS/HER:
(Main) JOB? ☐☐ If total 30 or more go to 15
Other jobs? ☐☐

14 WHAT IS THE REASON . . . USUALLY WORKS LESS THAN 30 HOURS PER WEEK?
☐ Enter code

15 LAST WEEK, HOW MANY HOURS OF OVERTIME OR EXTRA HOURS DID . . . WORK?
(Include paid and unpaid time at all jobs) ☐☐ If none enter 00

16 LAST WEEK, HOW MANY HOURS DID . . . LOSE OR TAKE OFF FROM WORK FOR ANY REASON SUCH AS ILLNESS, HOLIDAY, OR LAYOFF? (From all jobs)
☐☐ If none enter 00 and go to 18

17 WHAT WAS THE MAIN REASON FOR LOSING THESE HOURS?
☐ Enter code

18 HOW MANY HOURS DID . . . ACTUALLY WORK LAST WEEK AT HIS/HER:
(Main) JOB? ☐☐ Other jobs? ☐☐

19 IN THE PAST 4 WEEKS, HAS . . . LOOKED FOR ANOTHER JOB?
Yes 1 ◯ No 2 ◯ Go to 71

20 WHAT HAS . . . DONE IN THE PAST 4 WEEKS TO FIND ANOTHER JOB?
☐☐☐ Enter code(s) and go to 71

30 LAST WEEK, DID . . . HAVE A JOB AT WHICH HE/SHE DID NOT WORK?
Yes 1 ◯ Go to 33 No 2 ◯

31 LAST WEEK, DID . . . HAVE A JOB TO START AT A DEFINITE DATE IN THE FUTURE?
Yes 1 ◯ No 2 ◯ Go to 50

32 COUNTING FROM THE END OF LAST WEEK, IN HOW MANY WEEKS WILL . . . START TO WORK AT HIS/HER NEW JOB?
☐☐ Go to 50

33 WHY WAS . . . ABSENT FROM WORK LAST WEEK?
☐ Enter code and if code 6 go to 32

34 DID . . . HAVE MORE THAN ONE JOB LAST WEEK?
Yes 1 ◯ No 2 ◯

35 HOW MANY HOURS PER WEEK DOES . . . USUALLY WORK AT HIS/HER:
(Main) JOB? ☐☐ If total 30 or more go to 37
Other jobs? ☐☐

36 WHAT IS THE REASON . . . USUALLY WORKS LESS THAN 30 HOURS PER WEEK? ☐ Enter code

37 UP TO THE END OF LAST WEEK, HOW MANY WEEKS HAS . . . BEEN CONTINUOUSLY ABSENT FROM WORK? ☐☐

38 IS . . . GETTING ANY WAGES OR SALARY FOR ANY OF THE TIME OFF LAST WEEK?
Yes 1 ◯ No 2 ◯

39 INTERVIEWER CHECK ITEM
• If code 5 (layoff) in 33 1 ◯ go to 56
• Otherwise 2 ◯ go to 40

40 IN THE PAST 4 WEEKS, HAS . . . LOOKED FOR ANOTHER JOB?
Yes 1 ◯ No 2 ◯ Go to 71

41 WHAT HAS . . . DONE IN THE PAST 4 WEEKS TO FIND ANOTHER JOB?
☐☐☐ Enter code(s) and go to 71

50 HAS . . . EVER WORKED?
Yes 1 ◯ No 2 ◯ Go to 55

51 WHEN DID . . . LAST WORK AT A JOB OR BUSINESS?
Mo Yr | No change 1 | or | Mo Yr
If month unknown, enter – – in month

52 INTERVIEWER CHECK ITEM
• If last worked before Mo Yr 1 ◯ go to 55
Otherwise 2 ◯ go to 53

53 ABOUT HOW MANY HOURS PER WEEK DID . . . WORK AT THAT JOB?
☐☐ or | Full-time 1 ◯ Part-time 2 ◯

54 WHAT WAS THE MAIN REASON WHY . . . LEFT THAT JOB?
☐ Enter code

55 INTERVIEWER CHECK ITEM
• If "perm. unable to work" in 10 1 ◯ go to 80
• Otherwise 2 ◯ go to 56

56 IN THE PAST 6 MONTHS HAS . . . LOOKED FOR WORK?
Yes 1 ◯ No 2 ◯ Go to 70

57 IN THE PAST 4 WEEKS WHAT HAS . . . DONE TO FIND WORK? Mark all methods reported
Nothing ☐ Go to 63

• IN THE PAST 4 WEEKS HAS . . . DONE ANYTHING ELSE TO FIND WORK? Mark all other methods reported
For each method given ask:
• WHEN DID . . . LAST _____? (Repeat method)

Checked with:	Method used	No. of weeks ago (excl. any week)
PUBLIC employment AGENCY	2 ◯	☐
PRIVATE employment AGENCY	3 ◯	☐
UNION	4 ◯	☐
EMPLOYERS directly	5 ◯	☐
FRIENDS or relatives	6 ◯	☐
Placed or answered ADS	7 ◯	☐
LOOKED at job ADS	8 ◯	☐
OTHER Specify in NOTES	9 ◯	☐

58 WHAT WAS . . . DOING IMMEDIATELY BEFORE HE/SHE STARTED TO LOOK FOR WORK? FOR EXAMPLE, WORKING, KEEPING HOUSE, GOING TO SCHOOL OR SOMETHING ELSE.
☐ Enter code

59 UP TO THE END OF LAST WEEK, HOW MANY WEEKS HAS . . . BEEN LOOKING FOR WORK, NOT COUNTING WEEKS WHEN HE/SHE ALSO WORKED? ☐☐

60 HAS . . . BEEN LOOKING FOR A JOB TO LAST FOR LESS THAN 6 MONTHS, OR, MORE THAN 6 MONTHS?
Less than 6 months (incl. 6 mos.) 1 ◯ More than 6 months 2 ◯

61 ABOUT HOW MANY HOURS OF WORK PER WEEK HAS . . . BEEN LOOKING FOR? ☐☐ or | Full time 1 ◯ Part time 2 ◯

62 INTERVIEWER CHECK ITEM
• If "1 week ago" for any method in 57 1 ◯ go to 64
• Otherwise 2 ◯ go to 63

63 WAS THERE ANY REASON WHY . . . DID NOT LOOK FOR WORK LAST WEEK?
☐ Enter code

64 WAS THERE ANY REASON WHY . . . COULD NOT TAKE A JOB LAST WEEK?
☐ Enter code and go to 70

70 INTERVIEWER CHECK ITEM
• If "Yes" please read in 50 1 ◯ go to 80
• If last worked before Mo Yr in 51 3 ◯ go to 80
• Otherwise 4 ◯ go to 72 through 76 and check that the information is complete and correct

DESCRIPTION OF MAIN JOB OR BUSINESS

71 HAS . . . CHANGED EMPLOYERS SINCE LAST MONTH?
Yes 1 ◯ Enter new information for 72 through 76
No 2 ◯ Check that information in 72 through 76 is complete and correct

72 FOR WHOM DID . . . WORK? (Name of business, government dept. or agency, or person)
No change 1 | or ►

73 WHEN DID . . . START WORKING FOR THIS EMPLOYER?
Mo Yr | No change 1 | or If month unknown enter – – in month | Mo Yr

74 WHAT KIND OF BUSINESS, INDUSTRY OR SERVICE WAS THIS? (Give full description: e.g., paper-box manufacturing, retail shoe store, municipal government)
No change 1 | or ►

75 WHAT KIND OF WORK WAS . . . DOING? (Give full description: e.g., posting invoices, selling shoes, teaching primary school)
No change 1 | or ►

76 Class of worker:
Main job . . . . . ☐ No change 1 | or ☐ Enter code
Other job . . . . . ☐ No change 1 | or ☐ Enter code

77 go to 80

EDUCATIONAL ACTIVITY

80 LAST WEEK, WAS . . . ATTENDING A SCHOOL, COLLEGE, OR UNIVERSITY?
Yes 1 ◯ No 2 ◯ Go to 90

81 WAS . . . ENROLLED AS A FULL-TIME OR A PART-TIME STUDENT?
Full time 1 ◯ Part time 2 ◯

82 WHAT KIND OF SCHOOL WAS THIS?
☐ Enter code

INFORMATION SOURCE

90 HRD page-line No. of person providing the above information
Last interview ☐☐ | This interview ☐☐

91 Was this information provided over the telephone?
Yes 1 ◯ No 2 ◯

NOTES — See over for additional NOTES

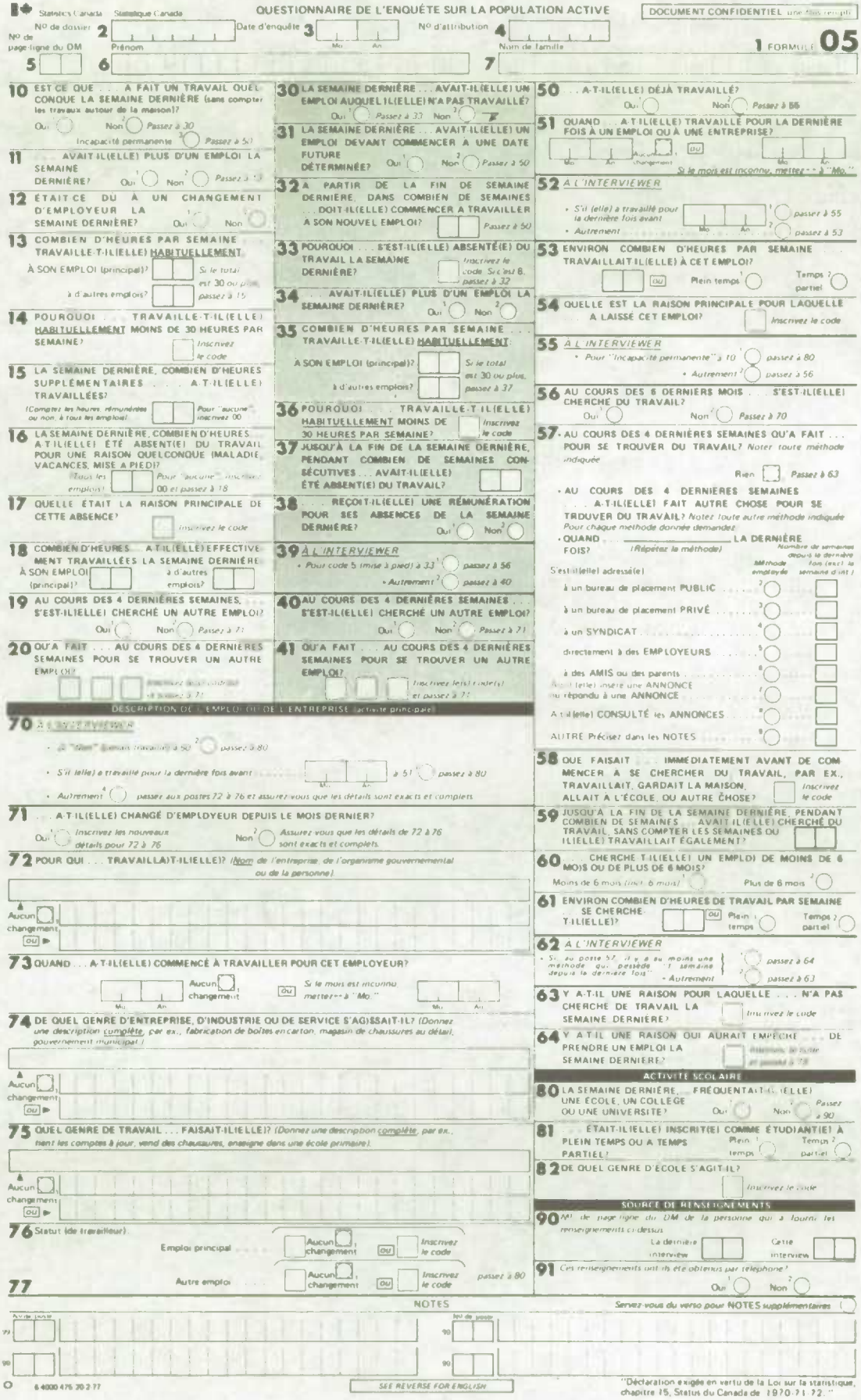
Item no. 99 ☐☐ | Item no. 99 ☐☐
Item no. 99 ☐☐ | Item no. 99 ☐☐

O 6 4000 475 20 2 77

FRANÇAIS AU VERSO

"Authority – Statistics Act, Chapter 15, Statutes of Canada 1970 – 71 – 72"

APPENDIX 13.2 – Concluded – ANNEXE 13.2 – fin



Statistics Canada / Statistique Canada

QUESTIONNAIRE DE L'ENQUÊTE SUR LA POPULATION ACTIVE

DOCUMENT CONFIDENTIEL une fois rempli

1 FORMULE 05

2 N° de dossier — 3 Date d'enquête (Mo, An) — 4 N° d'attribution — 5 N° de page-ligne du DM — 6 Prénom — 7 Nom de famille

10 EST-CE QUE ... A FAIT UN TRAVAIL QUELCONQUE LA SEMAINE DERNIÈRE (sans compter les travaux autour de la maison)? Oui 1 / Non 2 *Passez à 30* / Incapacité permanente 3 *Passez à 50*

11 ... AVAIT-IL(ELLE) PLUS D'UN EMPLOI LA SEMAINE DERNIÈRE? Oui 1 / Non 2 *Passez à 13*

12 ÉTAIT-CE DU À UN CHANGEMENT D'EMPLOYEUR LA SEMAINE DERNIÈRE? Oui 1 / Non 2

13 COMBIEN D'HEURES PAR SEMAINE ... TRAVAILLE-T-IL(ELLE) HABITUELLEMENT À SON EMPLOI (principal)? *Si le total est 30 ou plus, passez à 15*; à d'autres emplois? *passez à 15*

14 POURQUOI ... TRAVAILLE-T-IL(ELLE) HABITUELLEMENT MOINS DE 30 HEURES PAR SEMAINE? *Inscrivez le code*

15 LA SEMAINE DERNIÈRE, COMBIEN D'HEURES SUPPLÉMENTAIRES ... A-T-IL(ELLE) TRAVAILLÉES? *(Comptez les heures rémunérées ou non, à tous les emplois)* *Pour "aucune" inscrivez 00*

16 LA SEMAINE DERNIÈRE, COMBIEN D'HEURES A-T-IL(ELLE) ÉTÉ ABSENT(E) DU TRAVAIL POUR UNE RAISON QUELCONQUE (MALADIE, VACANCES, MISE À PIED)? *(Tous les emplois)* *Pour "aucune" inscrivez 00 et passez à 18*

17 QUELLE ÉTAIT LA RAISON PRINCIPALE DE CETTE ABSENCE? *Inscrivez le code*

18 COMBIEN D'HEURES ... A-T-IL(ELLE) EFFECTIVEMENT TRAVAILLÉES LA SEMAINE DERNIÈRE À SON EMPLOI (principal)? ; à d'autres emplois?

19 AU COURS DES 4 DERNIÈRES SEMAINES, S'EST-IL(ELLE) CHERCHÉ UN AUTRE EMPLOI? Oui 1 / Non 2 *Passez à 71*

20 QU'A FAIT ... AU COURS DES 4 DERNIÈRES SEMAINES POUR SE TROUVER UN AUTRE EMPLOI? *Inscrivez le(s) code(s) et passez à 71*

30 LA SEMAINE DERNIÈRE ... AVAIT-IL(ELLE) UN EMPLOI AUQUEL IL(ELLE) N'A PAS TRAVAILLÉ? Oui 1 *Passez à 33* / Non 2

31 LA SEMAINE DERNIÈRE ... AVAIT-IL(ELLE) UN EMPLOI DEVANT COMMENCER À UNE DATE FUTURE DÉTERMINÉE? Oui 1 / Non 2 *Passez à 50*

32 À PARTIR DE LA FIN DE SEMAINE DERNIÈRE, DANS COMBIEN DE SEMAINES ... DOIT-IL(ELLE) COMMENCER À TRAVAILLER À SON NOUVEL EMPLOI? *Passez à 50*

33 POURQUOI ... S'EST-IL(ELLE) ABSENTÉ(E) DU TRAVAIL LA SEMAINE DERNIÈRE? *Inscrivez le code. Si c'est 6, passez à 37*

34 ... AVAIT-IL(ELLE) PLUS D'UN EMPLOI LA SEMAINE DERNIÈRE? Oui 1 / Non 2

35 COMBIEN D'HEURES PAR SEMAINE ... TRAVAILLE-T-IL(ELLE) HABITUELLEMENT: À SON EMPLOI (principal)? *Si le total est 30 ou plus, passez à 37*; à d'autres emplois? *passez à 37*

36 POURQUOI ... TRAVAILLE-T-IL(ELLE) HABITUELLEMENT MOINS DE 30 HEURES PAR SEMAINE? *Inscrivez le code*

37 JUSQU'À LA FIN DE LA SEMAINE DERNIÈRE, PENDANT COMBIEN DE SEMAINES CONSÉCUTIVES ... AVAIT-IL(ELLE) ÉTÉ ABSENT(E) DU TRAVAIL?

38 ... REÇOIT-IL(ELLE) UNE RÉMUNÉRATION POUR SES ABSENCES DE LA SEMAINE DERNIÈRE? Oui 1 / Non 2

39 *À L'INTERVIEWER*
- *Pour code 5 (mise à pied) à 33* 1 *passez à 56*
- *Autrement* 2 *passez à 40*

40 AU COURS DES 4 DERNIÈRES SEMAINES ... S'EST-IL(ELLE) CHERCHÉ UN AUTRE EMPLOI? Oui 1 / Non 2 *Passez à 71*

41 QU'A FAIT ... AU COURS DES 4 DERNIÈRES SEMAINES POUR SE TROUVER UN AUTRE EMPLOI? *Inscrivez le(s) code(s) et passez à 71*

DESCRIPTION DE L'EMPLOI OU DE L'ENTREPRISE (activité principale)

70 *À L'INTERVIEWER*
- *À "Non" (aucun travail) à 50* 2 *passez à 80*
- *S'il (elle) a travaillé pour la dernière fois avant* (Mo, An) *à 61* 1 *passez à 80*
- *Autrement* 4 *passez aux postes 72 à 76 et assurez-vous que les détails sont exacts et complets*

71 ... A-T-IL(ELLE) CHANGÉ D'EMPLOYEUR DEPUIS LE MOIS DERNIER? Oui 1 *Inscrivez les nouveaux détails pour 72 à 76* / Non 2 *Assurez-vous que les détails de 72 à 76 sont exacts et complets*

72 POUR QUI ... TRAVAILLA(T)-IL(ELLE)? *(Nom de l'entreprise, de l'organisme gouvernemental ou de la personne)* Aucun changement ou ▶

73 QUAND ... A-T-IL(ELLE) COMMENCÉ À TRAVAILLER POUR CET EMPLOYEUR? (Mo, An) Aucun changement ou *Si le mois est inconnu, mettez-y "Mo."* (Mo, An)

74 DE QUEL GENRE D'ENTREPRISE, D'INDUSTRIE OU DE SERVICE S'AGISSAIT-IL? *(Donnez une description complète, par ex., fabrication de boîtes en carton, magasin de chaussures au détail, gouvernement municipal.)* Aucun changement ou ▶

75 QUEL GENRE DE TRAVAIL ... FAISAIT-IL(ELLE)? *(Donnez une description complète, par ex., tient les comptes à jour, vend des chaussures, enseigne dans une école primaire).* Aucun changement ou ▶

76 Statut (de travailleur). Emploi principal ... Aucun changement ou *Inscrivez le code*

77 Autre emploi ... Aucun changement ou *Inscrivez le code* *passez à 80*

50 ... A-T-IL(ELLE) DÉJÀ TRAVAILLÉ? Oui 1 / Non 2 *Passez à 55*

51 QUAND ... A-T-IL(ELLE) TRAVAILLÉ POUR LA DERNIÈRE FOIS À UN EMPLOI OU À UNE ENTREPRISE? (Mo, An) ou (Mo, An) *Si le mois est inconnu, mettez-y "Mo."*

52 *À L'INTERVIEWER*
- *S'il (elle) a travaillé pour la dernière fois avant* (Mo, An) 1 *passez à 55* / 2 *passez à 53*
- *Autrement* 2 *passez à 53*

53 ENVIRON COMBIEN D'HEURES PAR SEMAINE ... TRAVAILLAIT-IL(ELLE) À CET EMPLOI? ou Plein temps 1 / Temps partiel 2

54 QUELLE EST LA RAISON PRINCIPALE POUR LAQUELLE ... A LAISSÉ CET EMPLOI? *Inscrivez le code*

55 *À L'INTERVIEWER*
- *Pour "Incapacité permanente" à 10* 1 *passez à 80*
- *Autrement* 2 *passez à 56*

56 AU COURS DES 6 DERNIERS MOIS ... S'EST-IL(ELLE) CHERCHÉ DU TRAVAIL? Oui 1 / Non 2 *Passez à 70*

57 AU COURS DES 4 DERNIÈRES SEMAINES QU'A FAIT ... POUR SE TROUVER DU TRAVAIL? *Notez toute méthode indiquée*

Rien *Passez à 63*

- AU COURS DES 4 DERNIÈRES SEMAINES ... A-T-IL(ELLE) FAIT AUTRE CHOSE POUR SE TROUVER DU TRAVAIL? *Notez toute autre méthode indiquée*

Pour chaque méthode donnée demandez:

- QUAND ... _____ LA DERNIÈRE FOIS? *(Répétez la méthode)*

S'est-il(elle) adressé(e)	Méthode employée	Nombre de semaines depuis la dernière fois (excl. la semaine d'int.)
à un bureau de placement PUBLIC	2	
à un bureau de placement PRIVÉ	3	
à un SYNDICAT	4	
directement à des EMPLOYEURS	5	
à des AMIS ou des parents	6	
A-t-il(elle) inséré une ANNONCE ou répondu à une ANNONCE	7	
A-t-il(elle) CONSULTÉ les ANNONCES	8	
AUTRE Précisez dans les NOTES	9	

58 QUE FAISAIT ... IMMÉDIATEMENT AVANT DE COMMENCER À SE CHERCHER DU TRAVAIL, PAR EX., TRAVAILLAIT, GARDAIT LA MAISON, ALLAIT À L'ÉCOLE, OU AUTRE CHOSE? *Inscrivez le code*

59 JUSQU'À LA FIN DE LA SEMAINE DERNIÈRE, PENDANT COMBIEN DE SEMAINES ... AVAIT-IL(ELLE) CHERCHÉ DU TRAVAIL, SANS COMPTER LES SEMAINES OÙ IL(ELLE) TRAVAILLAIT ÉGALEMENT?

60 ... CHERCHE T-IL(ELLE) UN EMPLOI DE MOINS DE 6 MOIS OU DE PLUS DE 6 MOIS? Moins de 6 mois *(incl. 6 mois)* 1 / Plus de 6 mois 2

61 ENVIRON COMBIEN D'HEURES DE TRAVAIL PAR SEMAINE ... SE CHERCHE-T-IL(ELLE)? ou Plein temps 1 / Temps partiel 2

62 *À L'INTERVIEWER*
- *Si, au poste 57, il y a eu au moins une méthode qui possède "1 semaine depuis la dernière fois"* 1 *passez à 64*
- *Autrement* 2 *passez à 63*

63 Y A-T-IL UNE RAISON POUR LAQUELLE ... N'A PAS CHERCHÉ DE TRAVAIL LA SEMAINE DERNIÈRE? *Inscrivez le code*

64 Y A-T-IL UNE RAISON QUI AURAIT EMPÊCHÉ ... DE PRENDRE UN EMPLOI LA SEMAINE DERNIÈRE? *Inscrivez le code et passez à 73*

ACTIVITÉ SCOLAIRE

80 LA SEMAINE DERNIÈRE, ... FRÉQUENTAIT-IL(ELLE) UNE ÉCOLE, UN COLLÈGE OU UNE UNIVERSITÉ? Oui 1 / Non 2 *Passez à 90*

81 ... ÉTAIT-IL(ELLE) INSCRIT(E) COMME ÉTUDIANT(E) À PLEIN TEMPS OU À TEMPS PARTIEL? Plein temps 1 / Temps partiel 2

82 DE QUEL GENRE D'ÉCOLE S'AGIT-IL? *Inscrivez le code*

SOURCE DE RENSEIGNEMENTS

90 N° de page-ligne du DM de la personne qui a fourni les renseignements ci-dessus. La dernière interview ; Cette interview

91 Ces renseignements ont-ils été obtenus par téléphone? Oui 1 / Non 2

NOTES — *Servez-vous du verso pour NOTES supplémentaires*

Au de poste 79 ; 90 — Au de poste 90 ; 90

6 4000 476 30-2-77

SEE REVERSE FOR ENGLISH

"Déclaration exigée en vertu de la Loi sur la statistique, chapitre 15, Statuts du Canada de 1970-71-72."

APPENDIX – 13.3 – ANNEXE

DEFINITIONS AND EXPLANATIONS

Labour Force

The labour force is composed of that portion of the civilian non-institutional population 15 years of age and over who, during the reference week, were employed or unemployed.

Employed

The employed includes all persons who, during the reference week:

(a) did any work[1] at all;

(b) had a job but were not at work due to:

own illness or disability,
personal or family responsibilities,
bad weather,
labour dispute,
vacation,
other reason not specified above (excludes persons on layoff and persons whose job attachment was to a job to start at a definite date in the future).

Unemployed

The unemployed includes those persons who during the reference week:

(a) were without work, had actively looked for work in the past four weeks (ending with reference week), and were available[2] for work;

[1] Work includes any work for pay or profit. That is, paid work in the context of an employer-employee relationship, or self-employment. It also includes unpaid family work where unpaid family work is defined as unpaid work which contributed directly to the operation of a farm, business or professional practice owned or operated by a related member of the household.

[2] Persons in this group meeting the following criteria are regarded as available:
(i) were full-time students seeking part-time work who also met condition (ii) below. (Full-time students looking for full-time work are classified as not available for work in the reference week);

(ii) reported that there was no reason why they could not take a job in reference week, or, if they could not take a job it was because of "own illness or disability" or "personal or family responsibilities".

DÉFINITIONS ET EXPLICATIONS

Population active

Toutes les personnes civiles non pensionnaires d'institution âgées de 15 ans et plus qui, durant la semaine de référence, avaient un emploi ou étaient à la recherche d'un emploi.

Personnes occupées

Toutes les personnes qui, durant la semaine de référence,

a) avaient un emploi[1],

b) avaient un emploi, mais n'ont pas travaillé pour des raisons:

de maladie ou d'incapacité,
de charges familiales ou personnelles,
de mauvaises conditions atmosphériques,
de conflit de travail,
de vacances,
autres que celles qui précèdent (ne sont pas comprises les personnes en congédiement et les personnes dont la date d'entrée en fonction est encore à venir).

Personnes en chômage

Toutes les personnes qui, durant la semaine de référence:

a) étaient sans emploi, étaient à la recherche d'un emploi depuis quatre semaines (se terminant la semaine de référence) et étaient disponibles[2];

[1] Un emploi comprend tout travail effectué contre rémunération ou en vue d'un bénéfice, c'est-à-dire tout travail rémunéré, qu'il s'agisse d'un travail pour un employeur ou d'un emploi autonome. Cela comprend également le travail familial non rémunéré, c'est-à-dire l'exécution de tâches non rémunérées contribuant à l'exploitation d'une entreprise, d'une ferme ou d'un bureau professionnel appartenant à un parent du ménage ou exploité par celui-ci.

[2] Sont disponibles les personnes répondant aux critères suivants:
(i) les étudiants à plein temps à la recherche d'un emploi à temps partiel et qui répondent au critère ci-dessous. (Les étudiants à plein temps à la recherche d'un emploi à plein temps ne sont pas considérés comme étant disponibles dans la semaine de référence;
(ii) les personnes qui ont déclaré qu'il n'y avait pas de raison qui les empêchait d'accepter un emploi pendant la semaine de référence ou, si elles n'ont pas accepté d'emploi, que c'était pour des raisons de maladie ou d'incapacité ou de charges familiales ou personnelles.

(b) had not actively looked for work in the past four weeks but had been on layoff[3] for 26 weeks or less and were available[4] for work;

(c) had not actively looked for work in the past four weeks but had a new job to start in four weeks or less from reference week, and were available[4] for work.

Not in the Labour Force

That portion of the civilian non-institutional population 15 years of age and over who, during the reference week were neither employed nor unemployed.

Unemployment Rate

The unemployment rate represents the number of unemployed persons as a per cent of the labour force. The unemployment rate for a particular group (age, sex, marital status, etc.) is the unemployment in that group expressed as a per cent of the labour force for that group.

Participation Rate

The participation rate for a particular group (age, sex, marital status, etc.) is the labour force in that group expressed as a percentage of the population for that group.

Seasonal Adjustment

Fluctuations in economic time series are caused by trend, seasonal, cyclical and irregular movements. A seasonally-adjusted series is one from which movements occurring on a regular annual basis are eliminated.

Movements which usually occur in a statistical time series are separated out from those which have actually occurred; thereby leaving as a residual, the underlying trend of the series plus any erratic movements in the data. By the seasonal adjustment process, one is therefore better able to observe the general direction in which a series of statistics, such as those produced on labour force statistics, is moving.

[3] Persons are classified as being on layoff only when they expect to return to the job from which they were laid off.

[4] Persons in this group meeting the following criteria are regarded as available:
(i) were not full-time students;
(ii) reported that there was no reason why they could not take a job in reference week, or, if they could not take a job it was because of "own illness or disability" or "personal or family responsibilities" or "already had a job".

b) n'avaient pas cherché de travail au cours des quatre semaines précédentes mais se trouvaient en congédiement[3] depuis 26 semaines ou moins et étaient disponibles[4];

c) n'avaient pas cherché de travail au cours des quatre semaines précédentes, mais devaient occuper un nouvel emploi dans les quatre semaines suivant la semaine de référence et étaient disponibles[4].

Inactifs

Toutes les personnes civiles non pensionnaires d'institution âgées de 15 ans et plus qui, pendant la semaine de référence, n'étaient ni occupées ni en chômage.

Taux de chômage

Le taux de chômage est le pourcentage de personnes en chômage dans la population active. Le taux de chômage d'une catégorie particulière (âge, sexe, état matrimonial, etc.) représente la proportion de la population active comprise dans cette catégorie qui est en chômage.

Taux d'activité

Le taux d'activité dans une catégorie donnée (âge, sexe, état matrimonial, etc.) est le pourcentage de la population active par rapport à la population de cette catégorie.

Désaisonnalisation

Les variations dans les séries chronologiques économiques sont causées par des mouvements tendanciels, saisonniers, cycliques et irréguliers. Une série est désaisonnalisée lorsque les mouvements qui se présentent régulièrement chaque année sont éliminés.

Les mouvements qui touchent habituellement les séries chronologiques statistiques sont différenciés de ceux qui surviennent effectivement; par conséquent, il reste la tendance principale des séries et tous les mouvements erratiques. En appliquant la désaisonnalisation, on peut mieux analyser l'orientation générale d'une série statistique telle que celle qui est produite concernant les statistiques de la population active.

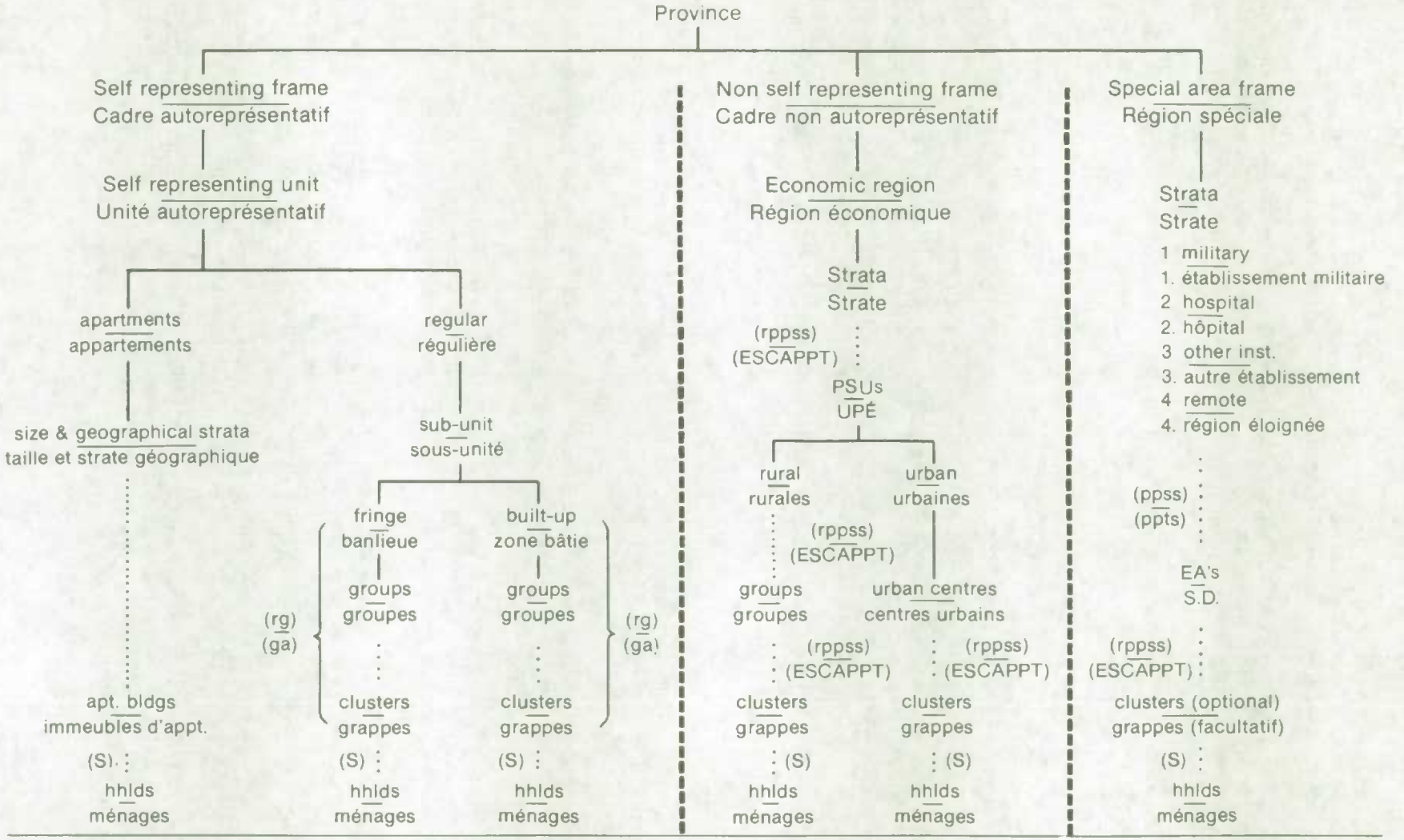
[3] Des personnes sont considérées comme étant en congédiement uniquement lorsqu'elles prévoient retourner à l'emploi qu'elles occupaient précédemment.

[4] Sont disponibles les personnes qui répondent aux critères suivants:
(i) ne sont pas étudiants à plein temps;
(ii) ont déclaré qu'il n'y avait pas de raisons qui les empêchaient d'accepter un emploi pendant la semaine de référence ou, si elles n'ont pas pu accepter d'emploi, que c'était pour des raisons de maladie ou d'incapacité, de charges familiales ou personnelles ou parce qu'elles avaient déjà un emploi.

While seasonal adjustment allows a much better look at the underlying trend of a series, the adjusted series still contain irregular fluctuations. Small changes between months in the adjusted series may simply be erratic movements and, to obtain a clearer idea of the underlying trend, users should examine the adjusted series over a number of months.

Bien que la désaisonnalisation permettre de mieux saisir la tendance principale d'une série, les séries désaisonnalisées comprennent quand même des variations irrégulières. De faibles modifications d'un mois à un autre, qu'on a observées dans les séries désaisonnalisées, peuvent être simplement des mouvements erratiques; pour obtenir une idée claire et précise de la tendance principale, les utilisateurs doivent étudier les séries désaisonnalisées qui couvrent un certain nombre de mois.

Diagramatic Representation of LFS Frame
Diagramme du cadre de l'EPA



* abbreviations for methods of selection

pps — probability proportional to size
ppss — pps systematic
rppss — randomized pps systematic
s — systematic
rg — random group

* abréviations pour les méthodes d'échantillonnage

ppt — probabilité proportionnelle à la taille
ppts — ppt et tirage systématique
rppts — ppt avec classement aléatoire et tirage systématique
s — systématique
ga — groupe aléatoire

——— level of stratification
niveau de stratification

..... stage of sampling *
étape de l'échantillonnage *

REFERENCES

[1] Drew, D., and Lindeyer, J., "Sample Design System for the Canadian Labour Force Survey", Household Surveys Development Division, 1976.

[2] Fellegi, I.P., "Sampling With Varying Probabilities Without Replacement: Rotating and Non-rotating Samples", Journal of the American Statistical Association, 58, 1963, pages 183 - 201.

[3] Fellegi, I.P., Gray, G.B., and Platek, R., "The New Design of the Canadian Labour Force Survey", Journal of American Statistical Association, 62, 1967, pages 421 - 453.

[4] Gray, G.B., "Variance Components and Variance Functions", Proceedings of Canadian Conference in Applied Statistics, 1971, pages 119 - 126.

[5] Gray, G.B., "Components of Variance Model in Multi-stage Stratified Samples", Survey Methodology Journal, Household Surveys Development Division, 1(1), 1975, pages 27 - 43.

[6] Hartley, H.O., and Rao, J.N.K., "Sampling With Unequal Probabilities Without Replacement", Annals of Mathematical Statistics, 33, 1962, pages 350 - 374.

[7] Keyfitz, N., "Estimates of Sampling Variance Where Two Units are Selected from Each Stratum", Journal of the American Statistical Association, 52, 1957, pages 503 - 510.

[8] Labour Force Survey Division, Statistics Canada, "Conceptual, Definitional and Methodological Changes in the Labour Force Survey", 1976.

[9] Petrie, B., "Project Review: Assessment and Revision of the Canadian Labour Force Survey", discussion paper presented at Provincial Conference, 1973.

[10] Platek, R., and Singh, M.P., "Some Aspects of Redesign of the Canadian Labour Force Survey", Proceedings of the American Statistical Association, 1972, pages 379 - 402.

[11] Platek, R., and Singh, M.P., "A Strategy for Updating Continuous Surveys", Metrika, 1977.

[12] Platek, R., and O'Haver, H., "Analysis, Control and Measurement of Survey Errors due to Interviewers", presented at the 40th session of International Statistical Institution in Warsaw, 1975.

BIBLIOGRAPHIE

[1] Drew, D., et Lindeyer, J., *Sample Design for the Canadian Labour Survey*, Division de l'élaboration d'enquêtes auprès des ménages, 1976.

[2] Fellegi, I.P., *Sampling With Varying Probabilities Without Replacement: Samples, Journal of the American Statistical Association,* 58, 1963, pages 183 - 201.

[3] Fellegi, I.P., Gray, G.B. et Platek, R., *The New Design of the Canadian Labour Force Survey, Journal of the American Statistical Association*, 62, 1967, pages 421 - 453.

[4] Gray, G.B., *Variance Components and Variance Functions*, compte rendu de la Conférence canadienne de la statistique appliquée, 1971, pages 119 - 126.

[5] Gray, G.B., *Components of Variance Model in Multistage Stratified Samples, Survey Methodology Journal*, Division de l'élaboration d'enquêtes auprès des ménages, 1(1), 1975, pages 27 - 43.

[6] Hartley, H.O., Rao, J.N.K., *Sampling with Unequal Probabilities Without Replacement*, Annales de la statistique mathématique, 33, 1962, pages 350 - 374.

[7] Keyfitz, N., *Estimates of Sampling Variance Where Two Units Are Selected from Each Stratum, Journal of the American Statistical Association*, 52 (1957), pages 503 - 510.

[8] Division de l'enquête sur la population active, Statistique Canada, 1976.

[9] Petrie, B., *Project Review: Assessment and Revision of the Canadian Labour Force Survey*, document de travail présenté au cours de la Conférence provinciale, 1973.

[10] Platek, R., et Singh, M.P., *Some Aspects of Redesign of the Canadian Labour Force Survey*, compte rendu de réunions de l'*American Statistical Association*, 1972, pages 379 - 402.

[11] Platek, R., Singh, M.P., *A Strategy for Updating Continuous Surveys*, Metrika, 1976.

[12] Platek, R. et O'Haver, H., *Analysis, Control and Measurement of Survey Errors due to Interviewers,* document présenté lors de la quarantième séance de l'Institut international de statistique, à Varsovie 1973.

REFERENCES – Concluded

[13] Platek, R., Singh, M.P., and Gray, G.B., "Strategy for the Increase in Sample Size of the Canadian Labour Force Survey", Household Surveys Development Division, discussion paper for Provincial Conference, 1973.

[14] Platek, R., and Petrie, B., "Development of the Revised Labour Force Survey", paper prepared for Eight Conference of Commonwealth Statisticians, Bermudas, 1975.

[15] Rao, J.N.K., Hartley, H.O. and Cochran, W.G., "On a Simple Procedure of Unequal Probability Sampling Without Replacement", Journal of Royal Statistical Society, 24(2), 1962, pages 482 - 490.

[16] Sampling and Survey Research Staff, Statistics Canada, *Methodology, Canadian Labour Force Survey*, Catalogue 71-504, 1965.

[17] Singh, M.P., and Lawes, M., "Further Studies on Sample Size Increase for the Canadian Labour Force Survey", Household Surveys Development Division, discussion paper for Provincial Conference, 1975.

[18] Tremblay, V., Singh, M.P., and Clavel, L., "Methodology for the New Re-interview Program", Household Surveys Development Division report, 1976.

BIBLIOGRAPHIE – fin

[13] Platek, R., Singh, M.P. et Gray, G.B., *Strategy for the Increase in Sample Size of the Canadian Labour Force Survey*, Division de l'élaboration d'enquêtes auprès des ménages, document de travail présenté lors de la conférence provinciale en 1973.

[14] Platek, R. et Petrie, B., *Development of the Revised Labour Force Survey*, document rédigé en prévision de la Huitième conférence des statisticiens des pays membres du Commonwealth, Bermudes, 1975.

[15] Rao, J.N.K., Hartley, H.O. et Cochran, W.G., *On a Simple Procedure of Unequal Probability Sampling Without Replacement*, *Journal of Royal Statistical Society*, 24(2), 1962, pages 482 - 490.

[16] Service de recherche en sondage et enquêtes, Statistique Canada, nº 71-504 au catalogue, 1965, *Méthodologie de l'enquête sur la population active du Canada*.

[17] Singh, M.P. et Lawes, M. *Further Studies on Sample Size Increase for the Canadian Labour Force Survey*, Division de l'élaboration d'enquêtes auprès des ménages, document de travail pour la conférence provinciale, 1975.

[18] Tremblay, V., Singh, M.P., Clavel, L. *Methodology for the New Re-interview Program*, Division de l'élaboration d'enquêtes auprès des ménages, rapport, 1976.

STATISTICS CANADA LIBRARY
BIBLIOTHÈQUE STATISTIQUE CANADA
1010529539