

DN 1355  
SAJ

MINISTERE DU PLAN ET DE  
L'AMENAGEMENT DU TERRITOIRE

REPUBLIQUE DU CAMEROUN  
Paix - Travail - Patrie

-----  
DIRECTION DE LA STATISTIQUE ET  
DE LA COMPTABILITE NATIONALE  
-----

ENQUETE CAMEROUNAISE AUPRES DES MENAGES  
(ECAM)

Ière phase

PLAN DE SONDAGE  
-----

Par : - Christopher SCOTT  
- André NGASSAM

Yaoundé, octobre 1992.

## 1. CADRE

L'Enquête Camerounaise Auprès des Ménages (ECAM) s'insère dans le cadre du projet Dimensions Sociales de l'Ajustement (DSA) parrainé par la Banque Mondiale et qui est en cours d'exécution dans des nombreux pays africains. L'ECAM se divise en deux éléments. Le premier, appelé Enquête sur les priorités, est constitué d'une enquête de suivie annuelle d'un contenu assez léger et basée sur un échantillon relativement important (généralement de 8 à 10 mille ménages). Cette série d'enquêtes, répétées en principe chaque année, sera interrompue après la première année par le deuxième élément, une enquête d'une nature plus approfondie sur un échantillon plus restreint intitulée Enquête intégrale et qui constitue la phase II de l'ECAM.

L'ensemble de ce programme d'enquêtes sera exécuté par la Direction de la Statistique et de la Comptabilité Nationale du Ministère du Plan et de l'Aménagement du Territoire, basée à Yaoundé.

Le présent rapport a pour objectif de proposer un plan de sondage en vue de l'ECAM, phase I. Il prend comme point de départ le rapport préliminaire rédigé en mai 1992 par la Direction de la Statistique traitant du même problème.

## 2. FACTEURS DETERMINANTS DU PLAN DE SONDAGE.

Deux types de contraintes déterminent le choix du plan de sondage de l'ECAM phase I : celles qui traduisent les objectifs spécifiques de l'enquête et celles qui résultent des moyens logistiques choisis pour son exécution.

En ce qui concerne les objectifs, il y en a un qui est déterminant quant au plan de sondage : la taille ainsi que la répartition de l'échantillon doit permettre une analyse adéquate de l'impact de la politique de restructuration sur chacun d'une dizaine de groupes socio-économiques bien définis. Or l'expérience des autres enquêtes africaines indique qu'un échantillon d'environ 400 ménages est nécessaire pour chaque sous-population ainsi visée. En préparant le plan de sondage, la première étape consistera donc à spécifier ces sous-populations visées pour l'analyse. Celles-ci seront

appelées groupes socio-economiques ou GSE. On déterminera par la suite les taux de sondage par strate de façon à assurer l'échantillon minimal requis pour chaque GSE.

Quant aux moyens logistiques, les principales contraintes sont les suivantes : L'enquête sera effectuée par 11 équipes itinérantes, chacune composée de 4 enquêteurs et un contrôleur et munie d'une voiture (sauf à Douala, où deux équipes partageront un véhicule). Un dénombrement des ménages sera organisé dans chaque aire de sondage sélectionnée. Cette opération sera complétée dans l'ensemble de l'échantillon avant que l'enquête proprement dite ne commence. A partir des listes ainsi dressées on tirera un échantillon de 20 ménages à enquêter dans chaque aire. L'enquête proprement dite durera environ 3 mois.

Ces contraintes impliquent un échantillon d'environ 8 000 ménages répartis dans environ 400 unités aréolaires.

### 3. BASE DE SONDAGE.

Il convient de distinguer les deux cas : base de sondage aréolaire et base de sondage pour les ménages.

Au niveau aréolaire on cherchera un ensemble d'unités qui couvrent la totalité de la surface camerounaise sans lacune et sans chevauchement. Les limites de chaque unité doivent être bien définies et faciles à repérer. Les unités ne devraient pas varier excessivement de taille et de préférence, la taille de chaque unité devrait être connue à l'avance afin qu'on puisse en tenir compte lors du tirage de l'échantillon. Enfin il serait bon que la liste qui constitue la base de sondage comporte des données descriptives de chaque unité permettant la stratification de la base en fonction des caractéristiques significatives.

Si on met de côté les villes de Yaoundé et Douala, on peut dire qu'au Cameroun l'ensemble des zones de dénombrement (ZD) du recensement le plus récent constitue une approche assez satisfaisante de cet idéal. Les ZD sont d'une taille assez uniforme avec une moyenne autour de 1 000 personnes, soit 200 ménages. Le Recensement Général de la Population et de l'Habitat exécuté en avril 1987, donne la taille de chaque ZD en fonction de sa population. Des cartes, datant de cette époque et d'une qualité adéquate, sont disponibles pour chaque ZD.

Quant à la stratification, les données sont limitées mais on sait au moins si chaque ZD est urbaine ou rurale ainsi que sa situation géographique au sein du Cameroun.

Malheureusement, il en est autrement en ce qui concerne les deux grandes villes. Un manque de temps a fait que la vaste opération de la cartographie censitaire n'a pu être complétée pour ces villes. Par conséquent les ZD ont dû être délimitées à partir des photographies aériennes vieilles d'une vingtaine d'années et d'une qualité médiocre. Pire encore dans bien des cas ce qui a été délimité sur les cartes est un groupe de 2 à 5 ZD sans la moindre indication des limites entre ces ZD.

Le problème du tirage d'un échantillon valable dans ces conditions a déjà été confronté lors des préparatifs pour l'Enquête Démographique et Sanitaire (DHS) en 1989. Les responsables de la DHS ont essayé de "se débrouiller" avec les ZD telles qu'elles se présentaient mais ont finalement avoué que les résultats étaient très loin d'être satisfaisants. Il est évident que, vue les mouvements des populations et l'aménagement continu de l'habitat, la situation s'aggrave avec chaque année qui passe. On est enfin arrivé au point où la constitution d'une nouvelle base de sondage dans ces deux villes doit être considérée comme une condition essentielle pour l'exécution d'une enquête valable

Heureusement, c'est justement à ce moment que s'est présentée une ouverture. En décembre 1991 la Direction du Cadastre du Ministère de l'Urbanisme et de l'Habitat a organisé une prise de photos aériennes sur l'ensemble de Yaoundé. La même direction prévoit d'ailleurs une opération analogue sur la ville de Douala en novembre 1992. Il sera facile pour la Direction de la Statistique d'obtenir un jeu de ces photos. Nous proposons donc qu'on dessine, en travaillant sur les photos, un ensemble d'îlots qui couvrent exhaustivement les deux métropoles. L'objectif sera de constituer des îlots d'une taille constante et nous proposons 100 ménages pour celle-ci. Comme le seul moyen qui permettra de juger des tailles sera de compter les toits vus dans les photos on ne peut espérer atteindre une grande précision dans ces estimations. Nous proposons que les agents travaillant sur le

découpage effectuent des visites sur le terrain de temps en temps afin de vérifier la nature de l'habitat et les taux d'occupation dans les différents quartiers, mais même avec l'aide de ces renseignements, il semble réaliste de prévoir un abattement de + 40 % au moins, en ce qui concerne les tailles obtenues. Un tel niveau d'imprécision sera néanmoins supportable car chaque îlot tiré dans l'échantillon sera ensuite soumis au dénombrement de ses ménages tout comme dans les autres secteurs de l'échantillon. C'est ici que l'exactitude nécessaire sera introduite.

#### 4. AU NIVEAU AREOLAIRE, UN SEUL DEGRE DE SONDAGE OU DEUX ?

Si nous admettons le choix de la ZD comme l'unité principale de sondage aréolaire (ou bien, de l'îlot tracé sur les photos dans le cas des métropoles), on peut toutefois se demander si on n'aurait pas intérêt à introduire un degré de sondage supplémentaire soit avant, soit après la ZD (ou îlot). Il convient d'examiner séparément le cas de la ZD et celui de l'îlot.

En ce qui concerne les ZD, l'introduction d'une unité supérieure à la ZD aurait deux conséquences significatives : on réduirait quelque peu la distance moyenne entre les ZD voisines dans l'échantillon et on augmenterait légèrement l'erreur de sondage. Or en sélectionnant plus de 300 ZD réparties sur l'ensemble du Cameroun on trouvera que la distance entre les ZD-échantillons est déjà très réduite et le gain en la réduisant davantage serait négligeable. Il semble donc préférable de maintenir la ZD comme unité primaire de sondage. Quant à l'introduction d'une unité inférieure à la ZD, il s'agirait du découpage de chaque ZD en segments dont on tirerait un ou deux au hasard. L'effet principal en serait une diminution de la taille de l'opération du dénombrement. Cela présente un avantage certain mais il serait contrebalancé par le travail supplémentaire qu'exigerait la délimitation des segments : nouvelles cartes à préparer et reproduire, nouvelle technique à apprendre, nouveau manuel à rédiger, ... Nous estimons que la solution la plus simple serait de maintenir le principe de la ZD comme la seule unité aréolaire de sondage en admettant toutefois que quelques ZD d'une taille

.../...

exceptionnelle pourront être subdivisées en deux ou trois segments avant le tirage, alors que des rares ZD excessivement petites pourront être regroupées avec des ZD voisines. L'unité de sondage serait alors normalement la ZD mais exceptionnellement le segment ou bien le groupe de deux ZD. On maintiendra ainsi rigoureusement le principe d'un seul degré de sondage aréolaire partout.

Quant à l'échantillonnage dans les métropoles, la situation est quelque peu différente car ici les unités aréolaires seront construites sur mesure, c'est à dire qu'on choisira la taille moyenne à donner aux îlots selon le besoin de l'enquête. Plus ceux-ci seront petits, moins sera le travail de dénombrement mais plus on aura des problèmes à les délimiter avec confiance sur les cartes. On peut estimer qu'une taille d'environ 100 ménages par îlot serait optimale. Si on suppose un échantillon total de l'ordre de 2 000 ménages pour les deux grandes villes ensemble (voir le tableau 2) il faudra prévoir 100 îlots pour l'échantillon. Si on devait tirer cet échantillon à un seul degré de sondage on aurait environ 3 000 îlots à délimiter, car le nombre de ménages dans les deux métropoles s'élève à environ 300 000. Si par contre on introduisait un degré de sondage préalable on pourrait envisager de diviser les deux villes en un ensemble d'unités primaires (UP) qui seraient bien plus grandes que les ZD mais dont les limites pourraient être tracées avec certitude sur les cartes et dont les populations pourraient être déterminées à partir du recensement (après extrapolation).

Quelle serait la taille minimale des telles UP dans l'hypothèse qu'on puisse obtenir leurs populations à partir du recensement ? Vue l'incertitude des opérations censitaires dans ces villes il nous semble que dans l'hypothèse la plus optimiste une taille moyenne de 30 000 habitants (6 000 ménages) par UP serait minimale. Le nombre total d'UP constituées serait alors de l'ordre de 50. Afin que le 1er degré de sondage soit utile on ne peut guère supposer un taux de sondage supérieur à 1/2. Un tel taux diminuerait de moitié le travail du découpage des UP en îlots, ce qui présente un avantage net. Par contre l'effectif

..../...

total de l'échantillon primaire s'élèverait à 25 unités seulement : or c'est ce paramètre surtout qui détermine la variance-sondage et il faut dire que le chiffre paraît bien trop modeste pour bien représenter un domaine aussi significatif dans l'enquête que celui des deux grandes métropoles.

Enfin un dernier argument en faveur d'un sondage à un seul degré au niveau aréolaire pour les deux métropoles est le suivant. La délimitation d'un ensemble d'îlots à titre de base de sondage aréolaire pour les deux grandes villes représente un investissement dans l'infrastructure statistique nationale. La présente enquête ne sera pas la dernière à en profiter. Il est évident que l'outil qui en résultera si ce travail est accompli sur une base exhaustive servira bien mieux les besoins des enquêtes futures que s'il est confiné au sein d'un échantillon primaire.

En somme, pour ces diverses raisons nous proposons partout un échantillon tiré à un seul degré aréolaire.

## 5. REPRESENTATION DES GROUPES SOCIO-ECONOMIQUES

### 5.1. Définition des groupes socio-économiques : principes généraux.

Le choix des GSE résulte d'une négociation entre analyste et sondeur. Rappelons d'abord qu'il s'agit de trouver une répartition de l'échantillon qui assurera un minimum de 400 ménages pour chaque GSE, ce chiffre étant considéré comme le minimum qui permette à l'analyste d'étudier avec une profondeur suffisante l'impact de la politique de restructuration sur un GSE donné. La contrainte fondamentale étant la taille globale de l'échantillon (dans notre cas 8 000 ménages) l'analyste souhaiterait étudier, dans cette enveloppe, le nombre maximum de GSE qu'on puisse inclure.

De prime abord on pourrait imaginer que ce nombre serait obtenu en divisant 8 000 par 400, soit 20. Mais ce serait faux car il y a une autre contrainte. Le sondeur ne sera pas

.../...

en mesure d'identifier les GSE avec précision au moment du sondage.

On verra dans le chapitre 5.2 que les GSE qui intéressent les analystes seront définis chacun en fonction de deux critères, dont l'un est géographique et l'autre socio-économique. Par exemple, on voudrait étudier les ménages résidant à Yaoundé/ Douala et rémunérés par le secteur public ; ou bien les ménages de la zone forestière qui cultivent un produit d'exportation. Or au moment du tirage de l'échantillon le sondeur connaîtra la situation géographique de chaque unité mais pas son classement socio-économique. Il devra donc fixer le taux de sondage en fonction du seul critère géographique. C'est ainsi que chaque strate géographique sera dotée d'un taux unique qu'on devra appliquer quelle que soit la classe socio-économique considérée au sein de la strate. Le sondeur cherchera alors la classe la plus rare dans la strate, calculera le taux de sondage qui donnera 400 ménages pour celle-ci et appliquera ce taux à l'ensemble des ménages de la strate. Il en résulte qu'il tirera bien plus de 400 ménages pour les classes autres que la plus rare dans la strate. C'est ainsi que, en exigeant un minimum de 400 ménages pour tous les GSE on sera amené à en tirer beaucoup plus que cela pour certains GSE.

## 5.2. Choix des GSE : Essai préliminaire

Pour se fixer une base de travail l'analyste choisira d'abord un découpage géographique et un classement socio-économique. Dans le cas du Cameroun il a été proposé les catégories suivantes :

### Découpage géographique

URBAIN :            Yaoundé + Douala  
                      Autres villes

RURAL :            Savane  
                      Hauts plateaux  
                      Zone côtière  
                      Zone forestière

.../...

Classement socio-économique

- Pour le secteur urbain :   Public  
                                  Para-public et privé formel  
                                  Informel  
                                  Autres
- Pour le secteur rural :   Cultivateurs, culture d'exportation  
                                  Cultivateurs, autres  
                                  Eleveurs  
                                  Autres

En croisant ces deux classifications on aurait en principe  $2 \times 4 = 8$  groupes urbains et  $4 \times 4 = 16$  groupes ruraux, soit un total de 24. Entre parenthèses, nous savons déjà que nous serons obligés de réduire ce nombre en regroupant certaines catégories mais avant d'aborder ce problème cherchons à raffiner un peu les définitions.

Classement urbain/rural. Le recensement applique une définition rigoureusement administrative. Nous estimons qu'un critère plutôt sociologique serait plus approprié pour cette enquête, à condition toutefois qu'il soit strictement objectif. Nous proposons donc la définition URBAIN = localité de plus de 10 000 habitants, en nous basant sur le recensement le plus récent.

Zones agro-climatiques. On définira les 4 citées plus haut autant que possible comme des groupes de provinces. Le réalisme nous oblige cependant de descendre au niveau du département dans deux cas particuliers. Les définitions proposées sont les suivantes :

SAVANE : Provinces de l'Extrême-Nord, du Nord et de l'Adamaoua.

HAUTS PLATEAUX : Provinces de l'Ouest et du Nord-Ouest plus les départements de Moungo et Mémé.

ZONR COTIERE : Provinces du Littoral (sauf Moungo) et du Sud-Ouest (sauf Mémé).

ZONE FORESTIERE : Provinces du Centre, du Sud et de l'Est.

.../...

Classes socio-économiques. Celles-ci seront définies en fonction de l'emploi du chef du ménage. On adoptera les définitions suivantes :

**SECTEUR PUBLIC :** L'ensemble des agents de l'Etat : Décisionnaires, Fonctionnaires et Contractuels.

**SECTEURS PARA-PUBLIC ET PRIVE FORMEL :** Tous les autres salariés. Nous avouons que l'inclusion des domestiques dans ce groupe peut être considérée comme une anomalie mais nous sommes contraints par l'absence des données statistiques.

**SECTEUR INFORMEL :** Tous les autres actifs occupés y compris les employeurs. Certains de ces derniers seront mal classés ici mais les données manquent pour les distinguer. La catégorie des employeurs est de toute façon très peu nombreuse.

**AUTRES :** Il s'agit de tous ceux qui n'ont pas d'emploi rémunéré : femmes au foyer, élèves/étudiants, retraités/vieillards, autres inactifs, apprentis, aides familiaux non rémunérés.

Classement des cultivateurs. Les cultures d'exportation sont le café, le cacao, le coton et le tabac. Il y a chevauchement presque total entre les cultivateurs des cultures vivrières et celles destinées à l'exportation : 94,5 % de ceux qui pratiquent une culture d'exportation pratiquent également une culture vivrière. D'autre part, 52,6 % de ceux qui pratiquent des vivrières pratiquent également une culture d'exportation. Dans cette situation notre souhait de distinguer une catégorie de cultivateurs du secteur "exportation" nous oblige d'adopter une définition un peu compliquée :

**CULTIVATEURS, SECTEUR EXPORTATION :** Ceux qui pratiquent une culture d'exportation quelle que soit leur situation par rapport à la pratique d'autres cultures.

**CULTIVATEURS, AUTRES :** Les cultivateurs qui ne pratiquent aucune culture d'exportation.

Éleveurs. Le chevauchement entre les éleveurs et les cultivateurs est encore plus extrême : moins de 1 % des éleveurs ne cultivent pas. Il est donc impossible de définir une catégorie utile d'éleveurs

.../...

qui les distinguerait des cultivateurs. Pour cette raison nous allons abandonner tout effort en vue du ciblage des éleveurs dans le plan de sondage. Nous verrons plus loin cependant qu'en ciblant les deux catégories de cultivateurs, on assurera en effet un nombre suffisant d'éleveurs pour les besoins des analystes.

Ménages ruraux non cultivateurs. Il s'agit du petit groupe de ménages qui ne pratiquent pas l'agriculture malgré qu'ils habitent le milieu rural. (On y comptera le groupe infime d'éleveurs "purs".).

### 5.3. Effectifs des chefs de ménage par GSE : méthodes d'estimation et sources de données

Une fois ces différentes définitions admises, on devra estimer l'importance numérique de chaque catégorie. Nous avons consulté les sources suivantes :

Recensement. (RGPH). Il s'agit en général des résultats publiés dans Démo 87, Vol. II, Tome 1 - Population de droit, dans les ménages, non redressée. Quelques données proviennent d'un tableau que la Direction nationale du Recensement a pu tirer sur notre demande mais il y a actuellement des obstacles d'ordre physique à la production des tableaux par cette direction. Nous avons extrapolé les données brutes sur 6 ans en appliquant des taux d'accroissement différentiels selon les milieux et les provinces.

Secteur public. Des données relatives à la mi-1992 nous ont été fournies par le Ministère de la Fonction Publique et de la Réforme Administrative.

Statistiques agricoles. Les données sur les cultivateurs et les éleveurs ont été fournies par le Ministère de l'Agriculture, Direction des Enquêtes Agro-économiques. Nous avons extrapolé celles-ci sur 6 ans de la même manière.

.../...

Hypothèses. Il y a un écart considérable entre les données disponibles et les données qu'on aurait souhaité consulter. Par conséquent il a fallu faire de nombreuses hypothèses. On ne peut citer ici que les plus significatives.

1). Les principales données sur Yaoundé-Douala proviennent d'un tableau tiré par la Direction du Recensement sur notre demande. Malheureusement l'arrondissement de Douala I y manquaient. Celui-ci constituait 20 % de la population de cette ville. Nous l'avons restitué par une simple extrapolation, en multipliant toutes les données pour Douala dans ce tableau par le facteur 1,25.

2). Le tableau qu'on vient de citer concerne les chefs de ménage alors que les données sur le secteur public fournies par la Fonction publique concernent les individus. Afin de convertir ces dernières en ménages équivalents nous avons fait l'hypothèse que, parmi les salariés des métropoles, le rapport entre le nombre de chefs de ménages et les autres adultes est le même dans le secteur public que dans le secteur para-public/privé formel.

3). Nous avons fait l'hypothèse analogue au sein du milieu "Autres villes".

4). Les autres données relatives à ce milieu ont été obtenues par soustraction. A cette fin il a fallu consulter le recensement pour obtenir des données sur les chefs de ménages en milieu urbain. Mais en empruntant ces données il a fallu d'abord les ajuster car le milieu urbain dans le recensement a été défini selon un critère plus large (on y a admis les villes en dessous de 10 000 habitants). Afin de tenir compte de cette différence nous avons diminué les chiffres du recensement pour le secteur urbain du facteur uniforme de 13 %, ce qui représente l'écart entre les populations recensées pour ces deux domaines.

5). Un problème s'est présenté au sujet des données fournies par la Fonction publique. On nous a communiqué les effectifs pour Yaoundé, pour Douala et pour chacune des chefs lieux de province, plus le total national. Il nous a fallu convertir cela en un chiffre pour les villes de plus de 10 000 habitants. Nous avons

.../...

fait l'hypothèse que les effectifs des agents de l'état par rapport à la population représenterait le même taux dans l'ensemble des villes de 10 000 h. et plus que dans les chefs lieux de province.

6). En milieu rural il a fallu estimer la catégorie des non cultivateurs. Le seul moyen était de soustraire le total des cultivateurs du total des ménages. Le premier vient d'une enquête agricole, le deuxième du recensement national. L'estimation est excessivement hasardeuse.

#### 7). REPARTITION DE L'ECHANTILLON

On est maintenant en mesure de commencer une série d'essais en vue d'une répartition de l'échantillon qui soit conforme aux contraintes imposées. Chaque essai se répercutera sur le choix des GSE ainsi que sur la taille de l'échantillon. Dans la mesure où la taille calculée dépasse la limite fixée, on sera obligé de revoir le choix des GSE. Des modifications ici permettront un deuxième essai, on verra la nouvelle taille d'échantillon, et le processus se répètera jusqu'à ce qu'on ait atteint un échantillon d'une taille admissible.

Avant d'aborder notre premier essai formel, il convient d'introduire déjà deux petites modifications du schéma des GSE dans le but de réduire le nombre total de GSE. Nous proposons que les deux groupes "Autres ménages", l'un dans le milieu urbain l'autre dans le rural, sois considérés seulement au niveau de l'ensemble de chaque milieu, sans qu'on essaye de les ventiler entre les deux strates urbaines, ni entre les 4 strates rurales. Leurs effectifs sont faibles et dans le cas rural, très mal connu. Si on insistait sur leur ventilation détaillée par strate on serait obligé d'allouer un grand échantillon rien que pour satisfaire à cette exigence. Avec cette modification on réduit le nombre de GSE DE 8 à 7 en milieu urbain et de 12 à 9 en milieu rural. Nous aurons ainsi un ensemble de 16 GSE pour notre premier essai.

Passons aux calculs détaillés. Il sont faits dans le tableau 1. La structure de ce tableau est conçue pour permettre au sondeur

.../...

de voir l'impact précis sur l'échantillon des décisions prises relatives au choix des GSE. Les strates géographiques sont disposées en colonnes : les lignes correspondent aux classes socio-économiques. Chaque case du tableau représente un GSE potentiel mais on pourra regrouper les cases, soit horizontalement soit verticalement, pour constituer les GSE plus convenables. L'indice h désigne la strate et l'indice g la caractéristique socio-économique.

Dans chacun des cases de ce tableau à deux entrées, on cherche à inscrire trois quantités :

$N_{gh}$  : le nombre de ménages de la population

$n_{gh}$  : le nombre de ménages de l'échantillon

$$F_{gh} = n_{gh}/N_{gh}$$

On commence avec les N. Ces quantités proviendront des diverses sources citées plus haut : recensement, enquête ...

Après avoir inscrit les  $N_{gh}$  on passe aux  $n_{gh}$ . On inscrira 400 pour le total de chaque GSE. Pour ce premier essai la plupart des GSE sont constitués d'une case unique du tableau : dans ces cas on inscrira 400 dans la case. Certains GSE peuvent être constitués de plusieurs cases : dans ces cas on inscrira les 400 dans la colonne du total correspondant à cet ensemble. (Il faudra prévoir les colonnes de totaux nécessaires à cette fin).

On passe ensuite aux  $F_{gh}$ . Pour chacun des cases où on voit  $n = 400$  on calculera  $F_{gh} = n_{gh}/N_{gh}$  et on inscrira cette quantité en dessous du n correspondant. Là où les 400 figurent dans une colonne total (c'est-à-dire là où le GSE s'étend sur plusieurs cases) on recopiera la valeur  $F_{gh}$  calculé sur le total, dans chacun des cases constituant le GSE. Par exemple, dans la ligne des non cultivateurs le taux  $F_{gh} = 5,49$ , calculé initialement sur le total des 4 zones doit être porté dans chacune des colonnes des 4 zones. En multipliant cette valeur de F dans chaque colonne par les N correspondant on complètera les n pour ces cases ( $n=134$  :  $n = 172$  ;  $n = 80$  ;  $n = 14$  à la ligne des non cultivateurs), chiffres qui totalisent évidemment à 400.

Enfin on fait le total  $N_h = \sum_g N_{gh}$  pour chaque colonne en l'inscrivant en bas. On cherche alors la plus grande fraction de sondage,  $F_{\max, h}$ , dans chaque colonne  $h$  et on l'inscrit en dessous du  $N_h$ . C'est ce taux de sondage qu'il faudra appliquer à l'ensemble de la strate  $h$ , car il assure l'échantillon voulu pour le plus petit des éléments trouvés dans la colonne.

En multipliant  $N_h$  par  $F_{\max, h}$  on obtiendra  $n_h$ , l'échantillon total par colonne, et en faisant la somme à travers les colonnes on obtiendra la taille globale de l'échantillon.

En étudiant le tableau 1 on voit immédiatement que la taille globale de l'échantillon, 14 053 ménages, dépasse de loin la limite qu'on s'est fixée. Comment peut-on modifier les GSE pour réduire l'échantillon ?

Examinons d'abord le milieu urbain. On y trouve un échantillon de 6 371, ce qui représente déjà 80 % de la limite imposée. En examinant les  $F_{\max, h}$  on voit que le problème provient entièrement du secteur public. Les effectifs dans ce secteur sont tellement faibles qu'on s'est trouvé obligé d'adopter des taux de sondage excessivement élevés et la conséquence en est un échantillon trop important. Cette observation suggère comme solution qu'on regroupe les deux strates urbaines pour ce secteur : c'est-à-dire qu'on réduise les deux GSE du secteur public en un seul, qui se rapportera alors au niveau du milieu urbain entier. Cette modification est introduite dans le tableau 2, deuxième essai. Le nombre de ménages dans l'échantillon urbain se trouvera ainsi réduit à 4 548.

Passons ensuite au milieu rural. Nous y trouvons un échantillon de 7 682 ménages. Comment peut-on le réduire ?

Nous remarquons d'abord que trois des quatre  $F_{\max, h}$  proviennent de la ligne des non cultivateurs. Or cette catégorie est assez hétérogène. Ce qui est plus grave, les chiffres dans cette ligne sont très mal estimés. On peut se demander si ces données seront utiles pour l'analyste. Nous proposons donc qu'on abandonne cette catégorie en tant que facteur déterminant de la

.../...

Tableau 1 : REPARTITION DE L'ECHANTILLON

1er essai

SECTEUR	URBAIN		
	Yaoundé/Douala	Autres Villes	Total Urbain
Public	N = 35 167 n = 400 F = 11,37	60 433 400 6,62	95 600
Parapublic et privé	N = 182 154 n = 400 F = 2,20	81 760 400 4,89	263 914
Indépendants	N = 98 400 n = 400 F = 4,07	196 143 400 2,04	294 543
Autres	N = 25 943 n = 165 F = 6,34	37 134 235 6,34	63 077 400 6,34
	$N_h = 341\ 664$ $F_{max,h} = 11,37$ $n_n = 3885$	375 470 6,62 2 486	717 134  6 371

Notes: Les valeurs de F représentent 1 000 fois les taux de sondage.

Tableau 1 (suite) REPARTITION DE L'ECHANTILLON

1er essai

	RURAL				Total Rural
	Savane	Hauts plateaux	Zone côtière	Forêt	
Cultivateurs (secteur exportation)	N = 138 859 n = 400 F = 2,88	258 341 400 1,55	23 436 400 17,07	177 356 400 2,26	597 992
Autres cultivateurs	N = 296 050 n = 400 F = 1,35	74 463 400 5,37	57 218 400 6,99	100 044 400 4,00	527 775
Non cultivateurs	N = 24474 n = 134 F = 5,49	31 337 172 5,49	14 498 80 5,49	2 534 14 5,49	72 843
	$N_h = 459 383$	364 141	95 152	279 934	11998610
	$F_{max,h} = 5,49$	5,49	17,07	5,49	
	$n_h = 2 522$	1 999	1 624	1 537	7 682

Note : les valeurs de F représentent 1000 fois les taux de sondage.

répartition de l'échantillon. On laissera en place les N afin de pouvoir donner la meilleure estimation possible du total rural, mais on ne tiendra aucun compte de cette ligne quant au choix des Fh.

Le sous-tableau qui suit montre l'effet de cette modification.

---

	Savane	: Hauts Plateaux	: Côtière	: Forêt	: Total
Nh	: 459 383	: 364 141	: 95 152	: 279 934	: 1 198 610
Fmax	b 2, 88	5,37	17,07	4,00	
n <sub>n</sub>	1 323	1 965	1 624	1 120	6 032

---

L'échantillon total rural s'élève maintenant à 6 032. Malheureusement ce chiffre reste encore trop élevé. Nous devons chercher encore une modification.

On remarquera que la population de la zone côtière est d'une très faible importance. Ne serait-il pas possible de regrouper cette zone avec des autres ? En effet un examen de la constitution des zones agro-écologiques montre qu'un remaniement serait justifié : une partie de la zone côtière pourrait très bien passer dans les "Hauts Plateaux" (il s'agit du restant de la province du Sud-Ouest) alors que le reliquat du Littoral pourrait être regroupé avec la zone forestière. On réduit ainsi le nombre de zones à 3. Cette modification est introduite dans le tableau 2 ainsi que celles citées plus haut.

.../...

Tableau 2. : REPARTITION DE L'ECHANTILLON

1er essai

SECTEUR	URBAIN		
	YAOUNDE/ DOUALA	AUTRES VILLES	TOTAL URBAIN
Public	N = 35 167 n = 147 F = 4,18	60 433 253 4,18	95 600 400 4,18
Parapublic et privé	N = 182 154 n = 400 F = 2,20	81 760 400 4,89	263 914
Indépendants	N = 98 400 n = 400 F = 4,07	196 143 400 2,04	294 543
Autres	N = 25 943 n = 165 F = 6,34	37 134 235 6,34	63 077 400 6,34
	N <sub>h</sub> = 341 664 F <sub>max,h</sub> = 6,34 n <sub>h</sub> = 2167	375 470 6,34 2381	717 134  4 548

Tableau 2(suite) : REPARTITION DE L'ECHANTILLON  
1er essai

	RURAL			
	Savane	Hauts Plateaux	Forêt	Total Rural
Cultivateurs sec- teur d'exportation	:N = 138 859 n = 400 F = 2,88	: 283 790 400 1,41	: 175 343 400 2,28	: 597 992
Autres cultivateurs	:N = 296 050 n = 400 F = 1 1,35	: 111 554 400 3,59	: 120 171 400 3,33	: 527 775
Non cultivateurs	:N = 24 474	: 48 161	: 208	: 72 843
	:N <sub>h</sub> : 459 383 F <sub>max,h</sub> = 2,88 n <sub>h</sub> = 1 323	: 443 505 3,59 1 590	: 295 722 3,33 984	: 1 198 610

(1) Les valeurs de F représentent 1000 fois les taux de sondage

(2) Hauts plateaux : Provinces de l'Ouest, du Nord-Ouest et du Sud-Ouest plus de département du Moungo

Forêt : Provinces du Centre, du Sud, de l'Est et du Littoral  
(moins le département du Moungo)

Nous arrivons ainsi à un échantillon total de 8 445 ménages. Ce total est suffisamment près à notre chiffre ciblé ("environ 8 000") pour qu'on puisse l'accepter.

Quelques petits ajustements restent à faire. Rappelons qu'il s'agit de tirer exactement 20 ménages dans chaque ZD (ou îlot). Il faudra donc ajuster les  $n_h$  au multiple de 20 le plus près. Ceci changera à leur tour les  $F_h$  à appliquer. Le dernier cadre du tableau 2 se trouve ainsi modifié légèrement comme il est montré au tableau 3.

TABLEAU 3 : REPARTITION DEFINITIF DE L'ECHANTILLON

URBAIN		RURAL		
Yaoundé-Douala	Autres villes	Savane	Hauts Plateaux (2)	Forêt
$N_h$ 341 664	375 470	459 383	443 505	295 722
$F_h$ 6,322	6,339	2,873	3,562	3,314
$n_h$ 2 160	2 380	1 320	1 580	890

Nous sommes maintenant en mesure de donner (tableau 4) la liste définitive des GSE avec le nombre estimé des ménages qui seront titrés dans l'échantillon dans chaque GSE, calculé à partir des fractions  $F_h$  qui figurent dans le tableau 3. (Certains de ces effectifs son légèrement inférieurs à 400 à cause de l'ajustement que nous venons de signaler.) Rappelons que certaines de ces estimations sont calculées d'une manière assez grossière.

Il importe enfin de noter qu'en visant certains GSE explicitement on n'exclut pas les analyses sur d'autres catégories. En effet toute catégorie dont l'échantillon dépasse 400 (approximativement) peut être considérée comme analysable avec à peu près la même précision que celle que nous avons exigée pour les GSE visés. Nous pouvons ainsi calculer en particulier le nombre de ménages éleveurs à prévoir dans l'échantillon avec les résultats suivants :

Ménages éleveurs prévus

Savane : 902 Hauts plateaux (2) : 950 Forêt : 816.

Rappelons que la quasi-totalité de ceux-ci sont déjà inclus dans l'un ou l'autre des GSE cultivateurs. Ces chiffres montrent que l'analyste trouvera un nombre largement suffisant d'éleveurs dans chacune des trois strates rurales sans que cette catégorie ait été ciblée explicitement.

Tableau 4 : LISTE DES GSE RETENUS

Définition du GSE en fonction de l'activité du chef du ménage	Nombre de ménages prévus dans l'échantillon
<u>GSE URBAINS</u>	
1. Secteur public	605
Secteurs para-public et privé:	
2. Yaoundé + Douala	1 152
3. Autres villes	518
Indépendants :	
4. Yaoundé + Douala	622
5. Autres villes	1 243
6. Autres (inactifs)	400
<u>GSE RURAUX</u>	
Cultivateurs du secteur exportation :	
7. Savane	399
8. Hauts plateaux (2)	1 011
9. Zone forestière	581
Autres cultivateurs :	
10. Savane	851
11. Hauts plateaux (2)	397
12. Zone forestière	398

## 7. MODALITES DU TIRAGE DE L'ECHANTILLON

### 7.1. Principes

La répartition de l'échantillon définie dans le tableau 2 implique une stratification des bases aréolaires de sondage en 5 strates, dont 2 en milieu urbain et 3 en milieu rural. Le tirage de l'échantillon aréolaire se fera indépendamment dans chaque strate. La méthode est la même partout sauf dans la première strate urbaine, où la base de sondage ainsi que la méthode de sondage est différente.

Nous décrirons par la suite les deux cas en termes théoriques. L'application en suivra au paragraphe 7.2.

#### 7.1.1 Strates autres villes et rurales.

La description qui suit concerne l'échantillonnage au sein d'une strate donnée.

Dans chaque strate on tirera un échantillon de ZD avec probabilité proportionnelle à la taille, c'est-à-dire au nombre de ménages recensés dans chaque ZD. Dans chacune des ZD tirée on fera un dénombrement des ménages habitant la ZD. Soit  $M_i$  le nombre de ménages recensés dans la ZD  $i$  et soit  $M_i'$  l'effectif de ménages enregistrés lors du dénombrement dans cette ZD. Ces deux quantités seront souvent différentes, que ce soit à cause des mouvements réels des populations ou des erreurs dans l'une ou l'autre opération.

On montre facilement que la probabilité pour que la ZD  $i$  soit dans l'échantillon est donnée par la formule :

$$p_{1i} = a m_i / \sum M_i$$

dans une strate donnée, où  $a$  = le nombre de ZD tirées dans la strate et la somme  $\sum M_i$  s'étend sur la totalité des ZD dans la strate (tirées ou non). La quantité  $p_{1i}$  représente la probabilité relative au 1er degré de sondage et à l'unité  $i$ .

.../...

Au 2ème degré de sondage on tirera exactement 20 ménages dans chaque ZD à partir des  $M_i'$  ménages dénombrés dans cette ZD. La probabilité pour qu'un ménage donné soit tiré dans la ZD  $i$  tirée est donc :

$$p_{2i} = 20 / M_i'$$

La probabilité globale pour qu'un ménage paraisse dans l'échantillon est le produit de ces deux probabilités, soit :

$$F_i = p_{1i} P_{2i} = 20 \left( a / \sum M_i \right) \cdot \left( M_i / M_i' \right)$$

Rappelons qu'il s'agit ici des quantités relatives à une seule strate. D'une manière plus générale on peut écrire, pour la strate  $h$  :

$$F_{hi} = 20 \left( ah / \sum_i M_{hi} \right) / \left( M_{hi} / M_{hi}' \right)$$

Remarquons que dans le cas où  $M_{hi}$  et  $M_{hi}'$  sont toujours identiques cette probabilité est constante pour tous les ménages de la strate. Un tel échantillon est appelé "autopondéré" à l'intérieur des strates. Mais puisqu'en fait  $M_{hi}$  et  $M_{hi}'$  sont différents il faudra dans notre cas pondérer les données lors de l'analyse. Les coefficients de pondération sont l'inverse des probabilités globales données par la formule précédente, soit :

$$W_{hi} = 1 / F_{hi} = \left( \sum_i M_{hi} / ah \right) \cdot \left( M_{hi}' / M_{hi} \right) / 20.$$

Ces poids devront être calculés pour chaque ZD de l'échantillon et inscrits dans le fichier des données de l'enquête. Il est donc indispensable d'enregistrer et de garder les valeurs  $M_i$  (qui proviennent du recensement) dès que le tirage des ZD est au point, et d'enregistrer et garder les  $M_i'$  à partir du moment où le dénombrement des ménages en vue de l'enquête ait été completé.

Enfin rappelons que dans toutes les formules citées plus haut les sommes  $\sum$  s'étendent sur la totalité des ZD de la strate, qu'elles soient tirées ou non dans l'échantillon.

.../...

### 7.1.2 Strate de Yaoundé et Douala.

Dans cette strate les unités aréolaires seront définies par des traits tracés sur des photos aériennes et on ne disposera pas d'estimation de la taille de chaque unité. On appellera ces unités les ilôts. Puisqu'il n'y a pas de mesure de taille on doit les tirer avec un taux constant. Ce sera :

$$p_1 = ah / Ah$$

où ah = le nombre d'ilôts tirés dans la strate (ce sera la quantité nh = 2 160 qui figure en bas de la première colonne du tableau 3, divisé par 20, d'où ah = 108) et Ah = le nombre total d'ilôts créés dans les deux villes ensemble. (A titre indicatif on peut prévoir un chiffre entre 3 et 4 mille pour ah.)

Au 2ème degré de sondage on tirera 20 ménages dans chaque ilôt, à partir des  $M_1'$  ménages dénombrés dans cet ilôt. La probabilité pour qu'un ménage donné soit tiré dans l'ilôt i de l'échantillon est donc :

$$p_{21} = 20 / M_1'$$

Enfin la probabilité globale pour qu'un ménage de cette strate paraisse dans l'échantillon est le produit de ces deux probabilités, soit :

$$F_1 = p_1 p_{21} = 20 ah / (Ah M_1')$$

Ces quantités étant variables entre ilôts, il faudra prévoir des coefficients de pondération lors de l'analyse. Les coefficients à utiliser sont

$$W_i = A_h M_1' / (20 ah)$$

Comme dans le cas des autres strates, il importe d'enregistrer et de conserver les paramètres  $A_h$  et  $M_1$  dès qu'ils sont connus.

## 7.2. Application

En appliquant les différentes formules citées plus haut il importe de noter que le tirage de l'échantillon du 1er degré (sauf dans la strate Yaoundé-Douala) se fera sur la base du recensement. Les unités sont les ZD et les mesures de taille  $M_1$  sont les effectifs de ménages recensés dans chaque ZD. Il convient d'écrire  $M_h = \sum_i M_{hi}$ , le total des ménages recensés dans la strate  $h$  (somme sur la totalité des ZD dans la strate et non pas sur l'échantillon de ZD seulement).

Il importe de bien distinguer les  $M_h$  des quantités  $N_h$  citées dans les tableaux 1, 2, 3 et 4. La différence en est que les  $M_h$  se réfèrent aux données brutes du recensement tandis que les  $N_h$  sont les données extrapolées jusqu'en 1993.

Les quantités  $n_h$  par contre dans ces tableaux concernent directement l'échantillon car en les divisant par 20 on obtient les valeurs  $a_h$  citées aux paragraphes 7.1.1 et 7.1.2.

Afin d'exécuter le tirage de l'échantillon aréolaire il faudra calculer pour chaque strate le rapport :

$$I_h = M_h / a_h \quad \text{- strates autres que Yaoundé-Douala}$$

où :  $I_h = A_h / a_h \quad \text{- strates Yaoundé-Douala}$

qui donne l'intervalle du tirage systématique (le "pas" de tirage) propre à chaque strate. Au moment de la rédaction de ce rapport ni les  $M_h$  ni la quantité  $A_h$  ne sont connus. Les  $M_h$  pourront être obtenus de la Direction nationale du Recensement et la valeur de  $A_h$  sera connue dès qu'on aura complété la délimitation et l'énumération des îlots à Yaoundé et Douala.

Pour les strates autres que Yaoundé-Douala, le tirage des ZD se fera avec probabilité proportionnelle à la taille suivant la méthode classique. On obtiendra d'abord la liste des ZD de chaque strate, avec la "taille" (nombre de ménages recensés) de

.../...

chaque ZD. La liste doit être ordonnée par arrondissement, en prenant les arrondissements dans un ordre "serpentin" - ordre qui assure que les unités qui sont voisines sur la liste sont voisines sur la carte. Les tailles devront être cumulées et les cumuls devront figurer dans une colonne de manière à ce qu'on trouve un cumul en face de chaque ZD dans la liste. On obtiendra alors pour chaque strate une série de sondage à partir de l'intervalle  $I_h$  déjà calculé, en commençant avec un nombre aléatoire choisi entre 0001 et  $I_h$  et en ajoutant systématiquement l'intervalle  $I_h$ . Pour chaque terme de cette série on cherchera dans la colonne ou qui dépasse ce terme. La ZD correspondant à ce cumul est tirée. Le processus se répètera de façon indépendante avec chaque strate.

En ce qui concerne la strate Yaoundé-Douala, la méthode est plus simple car les unités (ilôts) seront tirées avec probabilités égales. Il faudra d'abord numéroter les ilôts créés de 0001 à la fin. La numérotation sera faite par département. Au sein de chaque département, on suivra l'ordre serpentin des ilôts sur la carte. Ensuite, on tirera un nombre aléatoire entre 0001 et l'intervalle  $I_h$ . A partir de ce nombre, on ajoutera systématiquement l'intervalle  $I_h$  pour obtenir la série de sondage. Les termes de cette série donnent les numéros des ilôts tirés.

Quant à l'échantillonnage au 2ème degré, dans chaque ZD ou ilôt tiré, on devra organiser un dénombrement des ménages sur le terrain. Bien qu'on vise un échantillon de 24 ménages par ZD ou ilôt, on en tirera 24 afin de disposer d'une réserve pour le remplacement éventuel des ménages défaillants. Le tirage se fera encore par la méthode systématique. L'intervalle de tirage (le "pas") sera calculé dans chaque ZD ou ilôt par la formule  $M_i / 24$ , où  $M_i$  est l'effectif de ménages dénombrés dans la ZD ou l'ilôt  $i$ . Les 4 ménages de réserve seront choisis par un deuxième sous-échantillonnage systématique parmi les 24, à l'intervalle de 1 sur 6. Leurs noms seront alors extraits de la liste et donnés sur une liste séparée au contrôleur qui seul est habilité à autoriser l'utilisation d'un ménage remplaçant.

.../...

### 7.3 Extrapolation

Comme il a été indiqué au paragraphe 7.1, on doit calculer un coefficient de pondération pour chaque ZD et chaque îlot sur la base des formules qui figurent à la fin des paragraphes 7.1.1 et 7.1.2 respectivement. A noter qu'on doit utiliser 20 et non 24 dans ces formules, mais dans le cas exceptionnel où un ménage défaillant n'est pas remplacé ce chiffre doit être corrigé de manière à ce qu'il corresponde au nombre de ménages effectivement enquêtés dans la ZD ou îlot. Le coefficient de pondération doit être saisi dans le fichier de l'enquête pour chaque ménage (ou au moins pour chaque ZD/îlot car il sera le même pour tous les ménages d'un même ZD/îlot).

### 7.4 Traitement des ZD/îlots d'une taille aberrante

Il convient de distinguer deux cas : les ZD dont la taille est anormale selon le recensement et les ZD (ou îlots) qui se révèlent avoir la taille aberrante seulement au moment du dénombrement. "Taille aberrante" d'ailleurs peut être taille trop grande ou trop petite. Traitons séparément donc des 4 cas.

#### ZD trop grandes selon le recensement

Celles-ci ne présentent aucun problème de fond. Le seul inconvénient vient au moment du dénombrement où elles exigent un volume de travail excessif. Au lieu d'engager les équipes dans ce travail les responsables préféreront peut-être l'option qui consiste à subdiviser les grandes ZD en des segments d'une taille plus convenable. Si cette solution est adoptée on fixera un seuil au-delà duquel les ZD seront divisées. Nous proposerons 500 ménages recensés. Comme la taille moyenne des ZD dans le recensement est de l'ordre de 200 on pourrait adopter la règle suivante : au cas où une ZD dépasse 500 ménages recensés, diviser l'effectif de ménages recensés par 200 pour déterminer le nombre de segments. Ce travail ne sera entrepris cependant que si la ZD est tirée dans l'échantillon. Pour chacune des ZD concernées, on enverra une équipe pour diviser la ZD dans le nombre de segments prévu, des tailles approximativement égales.

.../...

Dans certains cas ce travail pourrait être fait entièrement au bureau, sur la base des cartes du recensement. Les segments seront alors insérés dans la liste des ZD qui constitue la base de sondage en leur affectant une taille estimative ajustée pour que la taille totale de la ZD ne change pas. Ce procédé permettra de maintenir la série de tirage sans modification et de voir sur lequel des segments le point d'échantillonnage tombe. On maintiendra de cette manière le principe d'un sondage à un seul degré aréolaire. c'est ainsi que toutes les formules restent valables : il suffit de reconnaître que l'unité primaire de sondage sera parfois un segment de ZD au lieu d'une ZD entière.

#### ZD/îlots trouvés trop grands lors du dénombrement

Même si on n'a pas segmenté la ZD dans le cadre de la méthode précédente on peut toujours résoudre le problème en créant des segments. Ceux-ci devraient être de tailles approximativement égales et on devra prendre note du nombre créés, soit  $s_i$  dans la ZD/îlot  $i$ . On tirera un des segments de façon aléatoire à probabilité égale, disons le segment  $ij$ . On effectuera le dénombrement des ménages dans ce segment, pour obtenir  $M_{hij}$  ménages et on en tirera 20 comme d'habitude. Dans ce cas la formule d'extrapolation change car il y a un degré supplémentaire de sondage. Dans l'expression pour le coefficient de pondération  $W_{hi}$  (fin 7.1.1 pour les ZD ou fin 7.1.2 pour les îlots) la quantité  $M_{hi}$  sera remplacé par  $s_i M_{hij}$ .

Ce procédé peut augmenter l'erreur de sondage et on aurait tort de l'appliquer souvent et systématiquement dans un domaine donné. Mais à condition qu'il ne soit utilisé que rarement il est admissible.

#### ZD trop petites dans le recensement

Le seul problème se pose si les ZD ont moins de 20 ménages, de sorte qu'on ne pourra tirer l'échantillon secondaire souhaité. Comme nous ne pouvons juger de la population actuelle à partir

.../...

des données du recensement il serait prudent de se donner une marge de sécurité. On fixera le seuil donc à 30 ménages recensés.

Des tels ZD seront sûrement rare dans notre échantillon. Au cas où on en trouve il faudra regrouper chaque ZD en dessous du seuil avec une ZD voisine. En principe ce regroupement devrait être effectué avant le tirage mais comme ces cas seront très rares on peut se permettre de l'effectuer seulement après le tirage et seulement pour les ZD tirées. Le groupe de deux ZD qui en résulte deviendra la nouvelle unité primaire choisie et les formules restent inchangées, avec  $M_{hi}$  égale au nombre de ménages recensés dans les deux ZD ensemble.

#### ZD/ilôts trouvés trop petits lors du dénombrement

Au cas où une ZD ou îlot tiré normalement se révèle avoir moins de 20 ménages lors du dénombrement, la solution est simple : on tirera pour l'échantillon tous les ménages qu'on y trouve. Si ce nombre est  $m_{hi}$  dans la ZD/ilôt  $hi$ , on remplacera en principe 20 par  $m_{hi}$  dans les formules de pondération, mais comme on aura aussi  $M_{hi}' = m_{hi}$  les deux  $m_{hi}$  s'annuleront et on obtiendra :

$$W_{hi} = \sum_i M_{hi} / (ahM_{hi}) \quad - \text{strates autres que Yaoundé - Douala}$$

$$\text{et } W_{hi} = A_h / ah \quad - \text{strate Yaoundé - Douala}$$

Enfin dans le cas particulier d'un ZD/ilôt qui s'est entièrement vidé depuis le recensement, cette unité s'annule automatiquement sans même qu'on ait besoin de la classer comme "non réponse" car les données y sont égales à zéro. On aurait tort de la remplacer.

#### 7.5 Ilôts impossibles à subdiviser

Enfin nous traitons d'une anomalie concernant la délimitation des îlots qui se présentera presque certainement une ou deux fois.

Dans quelques rares endroits sur les cartes de Yaoundé et Douala on trouve des ensembles d'habitations qui sont tellement entassées qu'on ne voit aucun trait naturel permettant de les

.../...

séparer en îlots. Dans des tels cas nous avons autorisé aux agents chargés de la création des îlots de constituer des "îlots doubles". Ceux-ci auraient une population cible de 200 ménages. Dans des cas très exceptionnels les agents peuvent même créer des îlots triples.

Dans les îlots doubles on doit affecter deux numéros d'îlot. Au cas où, lors du dénombrement des ménages, un tel îlot est rencontré on devra établir la liste des ménages dans l'ensemble de l'îlot double. Au cours de cette opération les ménages seront numérotés, comme c'est l'habitude dans tous les îlots. Par la suite l'îlot double sera divisé en deux îlots distincts en fonction de la numérotation des ménages : les ménages portant un numéro impair constitueront le premier îlot alors que ceux qui portent un numéro pair feront le deuxième îlot. Comme dans le cas normal, un seul des îlots sera tiré. La valeur de  $M_{hi}$  sera évidemment le nombre de ménages énumérés dans l'îlot tiré : par exemple, si c'est le deuxième îlot qui est tiré  $M_{hi}$  sera le nombre des ménages dénombrés dans l'îlot double. L'échantillon de ménages sera tiré parmi les ménages pairs seulement.

Au cas où on a créé un îlot triple on pourra allouer les ménages dans le dénombrement successivement aux îlots 1,2 et 3 en rotation. C'est-à-dire, le premier îlot sera constitué des ménages numérotés 1, 4, 7 ..., le deuxième îlot des ménages 2,5,8 ... et le troisième îlot des ménages 3,6,9 ...