

REPUBLIQUE DE GUINEE
Travail-Justice-Solidarité

MINISTERE DU PLAN

DIRECTION NATIONALE DE LA STATISTIQUE

**ENQUETE INTEGREE DE BASE POUR
L'EVALUATION DE LA PAUVRETE
(AVEC MODULE BUDGET ET CONSOMMATION)**

QUIBB/EIBEP 2002-2003

PLAN DE SONDAGE

Conakry, Décembre 2001

Méthodologie de l'enquête

Introduction

Le QUIBB et l'EIBEP sont administrés sur un même échantillon. De ce fait, ils ont un plan d'échantillonnage commun.

Le document contient cinq paragraphes :

- le premier paragraphe traite du contexte et des objectifs de l'Enquête Intégrée de Base pour l'Évaluation de la Pauvreté (EIBEP) ;
- le second a trait à l'organisation de la collecte des données;
- le troisième présente la base de sondage;
- le quatrième traite de la distribution de l'échantillon entre les différentes strates,
- enfin, le dernier paragraphe fournit les principes de tirage de l'échantillon et les formules de calcul des pondérations à utiliser dans l'extrapolation des données

1 CONTEXTE ET OBJECTIFS DE L'ENQUETE INTEGREE DE BASE POUR L'EVALUATION DE LA PAUVRETE (EIBEP)

L'EIBEP est une enquête auprès des ménages à objectifs multiples et s'inscrit dans le cadre de la mise en place du système d'informations statistiques pour la réduction de la pauvreté. L'EIBEP est l'élément central du dispositif statistique du suivi-évaluation de la mise en œuvre de la stratégie de réduction de la pauvreté telle que définie dans le Document final de Stratégie de la Réduction de la Pauvreté (DSRP) de la Guinée

L'objectif primordial de l'EIBEP demeure l'actualisation du profil de pauvreté élaboré à partir des données de l'Enquête Intégrale Budget Consommation (EIBC) de 1995 avec une possibilité d'analyse des déterminants au niveau des régions administratives.

L'EIBEP permettra également de mieux connaître la distribution des dépenses annuelles selon certaines caractéristiques des ménages et les inégalités entre les différents groupes socio-économiques. Elle fournira des informations utiles à l'élaboration des comptes nationaux en particulier l'estimation de l'autoconsommation qui est égale à la valeur des biens produits et consommés directement par les ménages.

Les résultats de l'EIBEP permettront de construire un nouvel indice de prix à la consommation dont la méthodologie sera harmonisée avec celle adoptée par la plupart des pays de la CEDEAO.

Enfin l'EIBEP fournira des indicateurs de base pour le suivi des indicateurs au point d'achèvement tels que définis dans les stratégies sectorielles (données du Questionnaire Unifié des Indicateurs de Développement et données sur l'éducation, la santé, la nutrition, l'emploi.....).

2 LA BASE DE SONDAGE

Pour sélectionner un échantillon probabiliste de ménages, il faudra disposer d'une base de sondage aréolaire et de celle des ménages.

La liste des ZD issues des travaux de cartographie du Recensement Général de la Population et de l'Habitat de 1996 (RGPH96) présente les caractéristiques d'une bonne base aréolaire. C'est elle qui est utilisée comme base de sondage de l'ÉIBEP.

En effet, elle couvre la totalité du territoire national sans omission, ni double compte ou chevauchement. Les limites des ZD sont bien connues ainsi que les tailles qui sont uniformes avec une moyenne de 1450 personnes environ, soit 207 ménages.

Les cartes d'une qualité adéquate sont disponibles et permettent de repérer chaque ZD.

La liste de ces unités comporte des données descriptives qui portent sur la situation géographique (préfecture, sous-préfecture), le nombre de ménages, la population résidente de 1996 et la zone d'emplacement (urbaine ou rurale). Chaque ZD est identifié par un numéro de 7 chiffres composé des codes de la préfecture (2 premiers chiffres), de ceux de la sous-préfecture (2 chiffres) et ceux de la ZD (3 chiffres).

Les ZD formées uniquement des camps réfugiés du Libéria sont marquées de deux étoiles. Toutes ces données sont stockées dans un fichier Excel.

3 ORGANISATION DE LA COLLECTE DES DONNEES

L'organisation de la collecte a été pour l'essentiel basée sur celle de l'EIBC. A la différence de l'EIBC qui était un échantillon rotatif, l'ÉIBEP utilise un échantillon fixe avec trois passages répétées de 4 mois chacun (la collecte des données dure 12 mois). Seulement 3 mois seront consacrés à la collecte effective sur le terrain à l'exception de Conakry. Le 4^{ème} mois est consacré au déplacement des équipes et à la correction des questionnaires sur le terrain.

Un volet nutritionnel a été introduit (pesée des préparations) et ne concernera que la moitié de l'échantillon.

Le relevé des données sur l'alimentation consiste à peser de façon régulière les produits qui entrent dans la préparation des repas du ménage et d'enregistrer leurs poids. Ces pesées peuvent être effectuées soit pendant sept (7) jours consécutifs dans 2 ménages jusqu'à la couverture complète du nombre de ménages concernés dans la ZD en plus de la charge normale d'un enquêteur, soit tous les 3 jours (en milieu urbain) ou 2 jours (en milieu rural) pendant 7 visites. Un test méthodologique a été introduit au niveau de l'enquête pilote. La première méthode semble mieux appropriée. C'est elle qui sera adoptée sur le terrain.

3.1 Milieu urbain :

3.1.1 En milieu urbain, les dépenses fréquentes des ménages subissent un cycle mensuel en fonction du jour de paie. Il devient important que le relevé des dépenses couvre un mois entier pour chaque ménage. Vu l'imprécision de la mémoire des enquêtés, il sera nécessaire d'effectuer des visites fréquentes afin de compléter ces relevés. Des visites à un intervalle de 3 jours semblent répondre à ce souci. On prévoit 11 visites à intervalle de 3 jours.

3.1.2 On estime qu'un enquêteur pourra visiter 4 ménages par jour. A chaque visite, il s'occupera d'une section du questionnaire ménage, des relevés de dépenses alimentaires et non alimentaires fréquentes, de l'autoconsommation, de l'auto prélèvement et des opérations de troc. De même, la pesée des produits alimentaires pour les ménages concernés sera effectuée.

3.1.3 Les 12 ménages sélectionnés dans une même ZD ont été répartis en 3 lots de 4 ménages. Un enquêteur s'occupera d'une ZD et travaillera dans un lot par jour c'est à dire aura une charge de travail de 4 ménages à enquêter par jour : En plus, l'enquêteur doit effectuer des pesées successives pour les produits entrant dans la préparation des repas au niveau de 2 ménages pendant une semaine. Il enquêtera donc 12 ménages au cours d'un cycle de 33 jours dont 6 ménages seront concernés pour les pesées des produits entrant dans la préparation des repas. Le programme de travail d'un enquêteur au cours d'un cycle de 33 jours est le suivant :

1^{er} Lot :	1 ^{er} jour,	4 ^{ème} jour,	7 ^{ème} jour	28 ^{ème} jour,	31 ^{ème} jour
2^{ème} Lot :	2 ^{ème} jour,	5 ^{ème} jour,	8 ^{ème} jour,	29 ^{ème} jour,	32 ^{ème} jour
3^{ème} Lot :	3 ^{ème} jour,	6 ^{ème} jour,	9 ^{ème} jour,	30 ^{ème} jour,	33 ^{ème} jour

3.1.3 Afin de permettre à l'enquêteur de prendre des jours de repos, on a prévu un enquêteur de réserve par équipe.

3.1.4 Les équipes de terrain en milieu urbain (Conakry et Autres Villes) sont constitués comme suit :

- Un Chef d'équipe
- Six enquêteurs (dont un enquêteur de réserve)

Les équipes de terrain en milieu urbain (Autres Villes) sont constitués comme suit :

- Un Chef d'équipe
- Cinq enquêteurs (dont un enquêteur de réserve)

3.1.5 Les enquêteurs d'une même équipe en milieu urbain (Conakry et Autres Villes) travailleront selon un calendrier identique dans des grappes différentes. Les grappes seront choisies aussi proches les unes des autres afin de faciliter la mise en place des enquêteurs et du contrôle.

3.1.6 Au total, au cours d'un passage, une équipe à Conakry fait 15 ZD à Conakry et 12 ZD au niveau de Autre Urbain (si elle ne fait que de l'urbain).

3.2 Milieu rural

3.2.1 En milieu rural, l'influence du jour de paie est minime alors que l'autoconsommation devient beaucoup plus significative. Par conséquent, on propose de diminuer la période de mémorisation de 3 à 2 jours et réduire le cycle de 33 à 16 jours par ZD.

3.2.2 Les équipes de terrain en milieu rural sont constituées comme suit :

- Un Chef d'équipe
- Cinq enquêteurs (dont un enquêteur de réserve)

3.2.3 En milieu rural, les enquêteurs d'une même équipe travailleront ensemble dans une même grappe. Chaque équipe disposera d'un échantillon de 20 ménages par grappe.

3.2.3 Un enquêteur aura donc à traiter 5 ménages dans une grappe. Ces ménages seront répartis en 2 lots de 2 ménages ou de 3 ménages. L'enquêteur visitera 2 ou 3 ménages par jour à intervalle de 2 jours, selon le programme de travail suivant :

1^{er} Lot (2 ménages): 1^{er} jour, 3^{ème} jour, 5^{ème} jour, 15^{ème} jour, 16^{ème} jour
2^{ème} Lot (3 ménages): 2^{ème} jour, 4^{ème} jour, 6^{ème} jour, 14^{ème} jour, 15^{ème} jour

3.2.4 Au total, au cours d'un passage, une équipe en milieu rural enquêtera 6 ZD.

4 RESSOURCES

La DNS déploiera 26 véhicules tout terrain pour réaliser les opérations de terrain (5 seront destinés à la supervision).

5 OBJECTIFS

L'objectif du plan de sondage est de fournir à l'analyste un échantillon d'une taille suffisante pour lui permettre de décrire et analyser tous les indicateurs de pauvreté au niveau national et au niveau plus réduit des Régions Administratives aussi bien en milieu urbain qu'en milieu rural.

6 LA STRUCTURE DE L'ECHANTILLON

Les éléments déterminants dans le choix d'un plan de sondage sont les objectifs de l'enquête et l'organisation de la collecte qui, elle-même dépend fortement des moyens du projet.

L'échantillon de l'EIBEP comme on l'a souligné plus haut, doit permettre une analyse adéquate des résultats au niveau national et dans les régions administratives du pays aussi bien en milieu urbain qu'en milieu rural. Cela demeure valable que si chaque strate considérée contient environ 400 ménages dans l'échantillon¹.

¹ Ce résultat est utilisé par Chris SCOTT dans plusieurs plans de sondage d'enquêtes réalisées en Afrique.

On se fixe d'avoir un échantillon autour de 7500 ménages, Les ménages de l'échantillon sont répartis entre les différentes strates qui sont constituées des milieux urbains et ruraux des 7 régions administratives et de la Zone Spéciale de Conakry., ce qui fournit un total de 15 strates.

La variable discriminante utilisée pour répartir les ménages est la dépense totale des ménages (la dépense par tête étant l'indicateur de niveau de vie utilisée dans les études sur la pauvreté).

On se propose alors de trouver la taille des strates qui rend minimum la variance inter-strate de la dépense des ménages sous la contrainte que la somme des tailles n_h de toutes les 15 strates fournissent un échantillon de 7612 ménages.

En supposant les rapports Sh/Yb stables, la règle de NEYMAN² conduit à :

$$n_h/(N_h * Yb) = \text{constante, c'est à dire}$$

$$n_h/Y = \text{constante d'où } n_h = cY$$

n_h la taille de l' échantillon dans la strate h

N_h le nombre total de ménages dans la strate h

Yb =dépense moyenne dans la strate h

Y = la dépense totale

Les données sur les dépenses des ménages proviennent de l'EIBC de 1994.

Le tableau 1 fournit la distribution de l'échantillon par strates. Il a été donc réparti la taille de l'échantillon dans les différentes strates proportionnellement à la dépense totale. Des ajustements ont été opérés. Il faut rappeler qu'il est proposé de tirer 12 ménages en milieu urbain et 20 en milieu rural et que les 4 ou 5 enquêteurs des équipes urbaines travaillent chacun dans un ZD pendant un cycle, tandis ceux des équipes rurales évoluent ensemble dans une seule ZD.

Ces ajustements consistent à obtenir des tailles qui sont multiples de 48 (1 ZD/enquêteur * 4 enquêteurs/ équipe * 12 ménages/ZD) dans les centres urbains de l'intérieur du pays et divisibles par 20 (1 ZD/équipe * 20 ménages/ZD) en milieu rural.

On a également tenu compte de la nécessité d'avoir une taille dans chaque strate qui puisse permettre de mener des analyses adéquates. Ce traduit naturellement les ajustements opérés au niveau de certaines tailles.

La répartition des ZD par cycles est donnée dans les tableaux ci-dessous. Elle a été faite en équilibrant la charge de travail au niveau de chaque passage.

² Il s'agit de trouver les tailles des strates qui rendent minimum la variance sous la contrainte ? $n_h = n$, on a le résultat suivant : $n_h/n = N_h * S_h / \sum N_h * S_h$ (n_h taille de l'échantillon dans la strate h et n taille de l'échantillon total).

Tableau 1 : Répartition du nombre de ménages et de ZD à tirer selon la strate

STRATE	MILEU URBAIN		MILIEU RURAL		TOTAL	
	Nombre de ménage	Nombre de ZD	Nombre de ménage	Nombre de ZD	Nombre de ménage	Nombre de ZD
Conakry	1200	100	0	0	1200	100
Kindia	432	36	560	28	992	64
Boké	480	40	420	21	900	61
Mamou	336	28	400	20	736	48
Labé	336	28	540	27	876	55
Faranah	432	36	400	20	832	56
Kankan	336	28	460	23	796	51
Nzérékoré	480	40	800	40	1280	80
TOTAL	4032	336	3580	179	7612	515

La répartition des ZD par passage est donnée dans les tableaux ci-dessous.

Elle a été faite en équilibrant la charge de travail dans les 3 ou 6 cycles que dure le passage.

Tableau 2 : Répartition des ZD de l'échantillon du milieu urbain par cycles de 33 jours

STRATES	TOTAL		NOMBRE DE ZD PAR CYCLE		
	Nombre de ménages	Nombre de ZD	1	2	3
Conakry	1200	100	35	35	30
Kindia	432	36	12	12	12
Boké	480	40	16	12	12
Mamou	336	28	8	12	8
Labé	336	28	8	8	12
Faranah	432	36	12	12	12
Kankan	336	28	12	8	8
Nzérékoré	480	40	12	16	12
TOTAL	4032	336	115	115	106

Tableau 3 : Répartition des ZD de l'échantillon du milieu rural par cycles de 16 jours

STRATES	TOTAL		NOMBRE DE ZD PAR CYCLE					
	Nombre de ménages	Nombre de ZD	1	2	3	4	5	6
Conakry	0	0	0	0	0	0	0	0
Kindia	560	28	5	5	4	4	5	5
Boké	420	21	4	3	4	4	3	3
Mamou	400	20	3	4	4	3	3	3
Labé	540	27	4	4	4	5	5	5
Faranah	400	20	4	4	3	3	3	3
Kankan	460	23	4	3	4	4	4	4
Nzérékoré	800	40	6	7	7	6	7	7
TOTAL	3580	179	30	30	30	29	30	30

Les tableaux 2 et 3 montrent qu'il faut prévoir la constitution de 56 équipes dont 7 à Conakry, 19 dans Autre Urbain et 30 en milieu rural. La combinaison des tâches <urbain /rural> pour une même équipe à l'intérieur du pays devrait améliorer l'organisation du travail sur le terrain.

7 LES MODALITES DU TIRAGE DE L'ECHANTILLON

La répartition de l'échantillon ci-dessus indique 15 strates dont 8 en milieu urbain et 7 en milieu rural. Le tirage de l'échantillon des ZD qui sont les unités primaires (UP) se fera indépendamment dans chaque strate. L'échantillon d'UP sera à probabilité proportionnelle à la taille, c'est-à-dire au nombre de ménages recensés dans chaque ZD. L'équipe autonome fera le dénombrement des ménages dans la ZD choisie.

Avant le calcul des probabilités de tirage des unités, il convient de préciser que la contrainte de l'organisation qui consiste à faire travailler un enquêteur par ZD en milieu urbain introduit un nouveau degré de tirage supplémentaire dans ce milieu. Les centres urbains qui moins de 4 ZD peuvent être regroupés avec ceux qui proches ou contigus avant d'effectuer le tirage.

L'échantillon est donc à 3 degrés en milieu urbain et à 2 en milieu rural :

Urbain		Rural	
1 ^{er} degré	ZD	Ville	
2 ^{ème} degré	Ménage	ZD	
3 ^{ème} degré	-	Ménage	

Pour les Villes et les ZD, la base de sondage est le Recensement Général de la Population et de l'Habitation de 1996. Au niveau des ménages, la base de sondage sera fournie par le listing des ménages issu du dénombrement.

L'échantillon comme on l'a vu plus haut est stratifié au 1^{er} degré. Les strates sont constituées des 7 Régions Administratives subdivisées en Urbain et Rural, la Ville de Conakry, constituant à elle seule une strate. Ce qui fait un total de 15 strates.

7.1 Milieu urbain

1^{er} degré de tirage :

On tire un centre urbain avec probabilité proportionnelle à la taille. La probabilité pour qu'un centre urbain k soit sélectionné est de :

$$P_{1hk} = a_h * M_{hk}/M_h$$

a_h = nombre de ZD tirés dans la strate h

M_{hk} = nombre de ménages dans la ville k de la strate h dans la base de sondage (RGPH-96)

M_h = nombre de ménage dans la strate h dans la base de sondage (RGPH-96)

2^{ème} degré de tirage :

Chaque fois que la Ville est tiré, on tire indépendamment et de façon systématique au niveau de la Ville 4 ZD.

Soient :

M_{hki} le nombre de ménages recensés en 1996 dans la ZD i de la ville k de la strate h,

M_{hki}' l'effectif obtenu par l'équipe autonome lors du dénombrement des ménages de cette ZD en 2002.

Les deux résultats sont souvent différents du fait du mouvement des ménages ou des erreurs dans les deux opérations.

La probabilité P_{2hki} de tirage relative de la ZD i au niveau de la Ville k de la strate h est :

$$P_{2hki} = M_{hki} / M_{hk} = M_{hki}' / M_{hk}$$

i = 1 à m_{hk} , m_{hk} étant le nombre total de ZD dans la ville k

3^{ème} degré de tirage

On tire dans la ZD échantillon, 12 ménages sur la liste des ménages dénombrés M_{hki}'

La probabilité pour qu'un ménage j soit tiré dans la **ZD i** est :

$$P_{3hki} = 12/M_{hki}'$$

j = 1 à n_i n_i étant le nombre total de ménages dénombrés dans la ZD i

Probabilité totale

La probabilité globale pour qu'un ménage de la **strate h** appartienne à l'échantillon est le produit de ces trois probabilités, soit :

$$P_{hkij} = P_{1hki} * P_{2hki} * P_{3hki} = (a_h \times M_{hk}/M_h) \times (M_{hki}/M_{hk}) \times (12/M_{hki}')$$

$$P_{hkij} = (12a_h/M_h) \times (M_{hki}/M_{hki}')$$

Les coefficients de pondération à utiliser lors de la sortie de résultats sont obtenus en inversant la probabilité P_{hkij} :

$$Whki = 1/P_{hkij} = (M_h/12a_h) (M_{hki}'/M_{hki})$$

$$Whki = 1/P_{hkij} = (M_h/12a_h) \times (M_{hki}'/M_{hki})$$

Milieu rural

Le tirage est à 2 degrés et faite proportionnellement à la taille :

Tirage au 1^{er} degré :

Soit M_{hi} le nombre de ménages recensés en 1996 dans la ZD i de la strate h, et soit M_{hi}' l'effectif obtenu par l'équipe autonome lors du dénombrement dans la même ZD.

La probabilité P_{1hi} pour que la ZD i soit choisie est donnée par la formule :

$$P_{1hi} = a_h * M_{hi}' / M_{hi} = a_h * M_{hi}' / M_h$$

i = 1 à m_h m_h étant le nombre total de ZD dans la strate h

a_h = le nombre de ZD tiré dans la strate h.

Tirage au deuxième degré :

On tire exactement 20 ménages dans l'UP choisie à partir des **Mhi'** ménages qui y sont dénombrés.

La probabilité P_{2hij} pour qu'un ménage soit tiré dans la ZD i est alors égale à :

$$P_{2hij} = 20/M_{hi}'$$

$j = 1$ à n_i n_i étant le nombre total de ménages dénombrés dans la ZD i

Probabilité totale

La probabilité globale pour qu'un ménage de la strate h appartienne à l'échantillon est le produit de ces deux probabilités, soit :

$$P_{hij} = P_{1hi} \times P_{2hij} = 20 a_h \times (M_{hi} / M_{hi}') \times 1/M_h$$

Les coefficients de pondération à utiliser lors la sortie de résultats sont obtenus en inversant la probabilité P_{hij} :

$$W_{hi} = 1/P_{hij} = (M_h / 20 a_h) \times (M_{hi}' / M_{hi})$$

A l'aide d'une feuille Excel on peut saisir et garder toutes ces données indispensables au calcul de ces coefficients qu'il faudra inscrire dans le fichier des données de l'enquête. Cette feuille existe déjà, il suffit d'extraire les ZD-échantillon et de compléter avec les M_{hi}' qui vont provenir régulièrement de l'opération de dénombrement des ménages de ces ZD. Les résultats de ces calculs seront transférés facilement dans un fichier SPSS en prenant la précaution de les sauvegarder dans la version Excel 4.0.