

COTE D'IVOIRE - Enquête Permanent auprès des Ménages (1986-1987)

Institut National de la Statistique - Ministère du Plan

Rapport généré le: October 8, 2015

Vous pouvez consulter notre catalogue de données sur: <http://www.ins.ci/n/nada/index.php>

Aperu

Identification

ID
CIV-INS-EPAM-1986

Version

DESCRIPTION DE LA VERSION

Il s'agit de la version après traitement et analyse des données.

DATE DE PRODUCTION

1986

Aperu

RSUM

L'Enquête permanente auprès des ménages (EPAM), connue l'étranger sous le nom "Enquête sur les niveaux de vie en Côte d'Ivoire (CILSS)" est la première enquête LSMS qui a conduit un test de la méthodologie et du questionnaire développés par l'équipe en question. Cette enquête est composée de trois volets complémentaires : l'enquête auprès des ménages, l'enquête communautaire et l'enquête sur les prix. L'enquête auprès des ménages collecte des informations détaillées sur les dépenses, le revenu, les avoirs, les besoins alimentaires ainsi que des informations sur d'autres caractéristiques socio-économiques des ménages. L'enquête communautaire collecte des informations sur les caractéristiques économiques et démographiques des communautés rurales auxquelles appartient chaque grappe de ménages. Ceci vise à établir un lien entre les données collectées au niveau des ménages et celles collectées au niveau communautaire. La composante prix collecte les prix des biens sur le marché le plus rapproché des grappes de ménages retenus, de manière ce qu'il soit possible de construire des indices de prix régionaux pour l'enquête auprès des ménages.

TYPE DE DONNÉES

Donnéeschantillonnées [ssd]

UNITS D'ANALYSE

Ménages et individus de type africains résidant sur le territoire national.

Champ

NOTES

- Santé
- Éducation
- Emploi
- Logement
- Activités agro-pastorales

TOPICS

Sujet	Taxonomie	URI
SANTÉ		
ÉDUCATION		
EMPLOI		

Sujet	Taxonomie	URI
AGRICULTURE		
LOGEMENT		

KEYWORDS

SANTE, EDUCATION, EMPLOI, AGRICULTURE, LOGEMENT

Couverture

COUVERTURE GOGRAPHIQUE

L'enquete est representative pour le milieu urbain et rural des 10 ples de dveloppement et de la ville d'Abidjan.

GEOGRAPHIC UNIT

Ple de dveloppement et le milieu urbain et rural

UNIVERS

Ensemble des mnages de type africain rsidant en Cte d'Ivoire.

Producteurs et sponsors

INVESTIGATEUR PRINCIPAL

Nom	Affiliation
Institut National de la Statistique	Ministre du Plan

FINANCEMENT

Nom	Abbrviation	Rle
ETAT DE COTE D'IVOIRE	CIV	
BANQUE MONDIALE	BM	

Production des mtadonnes

MTADONNES PRODUITES PAR

Nom	Abbrviation	Affiliation	Rle
INSTITUT NATIONAL DE LA STATISTIQUE	EPAM 86	MINISTERE DU PLAN	CONCEPTION ET EXECUTION

DATE DE LA PRODUCTION DES MTADONNES

2009-12-26

VERSION DU DOCUMENT DDI

EPAM 1986 v1.0

ID DU DOCUMENT DDI

CIV-EPAM86-DDI-1.0

Echantillonnage

Méthode d'échantillonnage

La taille de l'échantillon retenue chaque année était de 1600 ménages et le plan de sondage, un plan de rotation. C'est-à-dire la moitié de l'échantillon total des ménages retenus était revisitée l'année suivante, alors que l'autre moitié était remplacée par un nouvel échantillon des ménages. L'enquête ainsi constituée produisait une séquence de quatre bandes de données en coupe transversale et trois bandes de données de panels superposées de 800 ménages chacune (1985-86, 1986-87, 1987-88).

Plan d'échantillonnage et sélection de l'échantillon

L'objectif principal de la sélection de l'échantillon de l'enquête auprès des ménages était d'obtenir une coupe transversale de ménages africains ayant une représentation nationale, et de permettre certains de ces ménages d'être interviewés les années suivantes comme ménage de panel. Le tirage deux degrés a été adopté comme procédure d'échantillonnage. Au premier niveau du tirage de l'échantillon, 100 unités primaires de sondage (UPS) ont été sélectionnées dans le pays à partir de la liste de toutes les UPS de la base de sondage. Au second niveau du tirage, une grappe de 16 ménages a été sélectionnée dans chaque UPS. Cette approche a conduit à un échantillon de 1600 ménages l'année, constitué à partir de 100 grappes de 16 ménages chacune. La moitié des ménages retenus dans l'échantillon était remplacée chaque année, alors que l'autre moitié (celle constituée des ménages formés à partir des données de panel en 1986, 1987 et 1988) était interviewée pour une seconde fois.

Il est important de noter qu'une nouvelle base de sondage ainsi que de nouvelles procédures de sélection des UPS et d'établissement des listes de ménages ont été utilisées pour sélectionner la moitié des grappes de ménages interviewés en 1987 (l'autre moitié de ménages était sélectionnée à partir des données de panels de l'échantillon de 1986) et toutes les grappes de ménages interviewés en 1988. Ainsi, les ménages sélectionnés dans la première base constituent le Bloc 1 et ceux sélectionnés à partir de la seconde base le Bloc 2.

Procédure d'échantillonnage pour le Bloc 1

La base de sondage.

La base de sondage des enquêtes de 1985, 1986 et la moitié de l'échantillon de 1987 (sauf pour Abidjan et Bouak) était une liste des localités sur la base du recensement de 1975, mise à jour jusqu'en 1983 par les démographes de la Direction de la Statistique et basée sur une population totale estimée à 9,4 millions.

La base de sondage du Bloc 1 d'Abidjan et de Bouak a été constituée à partir des données du recensement électoral de 1979-1980 conduit dans ces deux villes. L'enquête électorale avait produit des cartes détaillées avec un découpage de chaque quartier en sous-sections plus petites (lots). Les sous-sections de chaque ville ont été regroupées en fonction du type de logement par les statisticiens de la Direction de la Statistique pour former les Unités Primaires de Sondage (UPS). À partir de la liste des UPS de chaque ville et de la population de chaque UPS, le nombre de UPS requis était sélectionné en utilisant la procédure de tirage systématique. La procédure de tirage systématique utilise comme "pas" initial la population totale de la ville divisée par le nombre de UPS retenus dans chaque ville. Un problème rencontré dans le processus de sélection provient du fait qu'un quartier de la ville d'Abidjan (Yopougon), qui était relativement sous-peuplé au moment du recensement électoral en 1980, est devenue l'une des plus grandes agglomérations de Côte d'Ivoire. Cette difficulté était toutefois inévitable puisque les données exactes de la population de Yopougon n'étaient pas disponibles au moment de la sélection des UPS.

Tirage des UPS

Une stratification géographique n'était pas nécessaire puisque la procédure de sondage systématique utilisée dans le tirage des UPS garantissait un échantillon représentatif et équilibré dans les régions et leurs types d'habitat. Les principales régions retenues dans le plan de sondage étaient : Fort Est, Fort Ouest, et Savane. Les types d'habitat retenus étaient les suivantes : grandes villes, petites villes, villages (grands et petits) rattachés aux villes, villages centres et les villages rattachés à ceux-ci. Les 100 UPS ont été sélectionnées à partir d'une liste des UPS classées suivant les régions et les types d'habitats, et dont la probabilité de sélection était proportionnelle à la taille de la population.

Tirage des ménages dans chaque UPS.

Une pré-enquête a été conduite en juin-juillet 1984 afin d'établir une base de sondage pour le second degré de l'échantillonnage, c'est-à-dire d'établir la liste des ménages dans chaque UPS à partir de laquelle 16 ménages ont été sélectionnés. Le dénombrement des ménages devait être appliqué aux enquêtes de 1985 et 1986, pour éviter un second dénombrement coûteux en 1986. Alors, il fallait que la pré-enquête de 1984 dénombre suffisamment de ménages pour pouvoir retenir deux grappes dans chaque UPS et remplacer les ménages qui ne seraient pas en mesure de participer. Sur la base de ce critère, 64 ménages ont été dénombrés au sein de chaque UPS. Pour les UPS ayant 64 ménages ou moins, chaque ménage était dénombré. Pour tirer les ménages, le pas de sondage utilisé a été calculé en divisant par 64 le nombre de ménages estimés dans chaque UPS. Par exemple, dans une UPS comptant environ 640 ménages, chaque dixième ménage a été inclus dans la liste. Les enquêteurs ont commencé à partir d'un point de l'UPS choisi au hasard. Pour des raisons opérationnelles, le pas de sondage maximum a été fixé à 30. En pratique, il paraît que, en

appliquant le pas de sondage fixe, les énumérateurs ont compté les portes, plutôt que les logements. Malgré le fait que les énumérateurs devraient commencer l'énumération à partir du point choisi au hasard, il semble que parfois dans le milieu rural et dans les petites villes ils commencent au centre de l'UPS.

Procédure d'échantillonnage pour le Bloc 2

La base de sondage

La base de sondage pour le Bloc 2 a été élaborée à partir d'une liste de localités provenant des résultats du Recensement de sites habités (RSH) préparatoire au recensement général de la population de 1988. À partir des travaux cartographiques du RSH, des UPS ont été constituées avec une estimation de leur population.

Tirage des UPS

Les UPS ont été sélectionnées avec une probabilité proportionnelle à la taille. Cependant, pour éviter les coûts élevés liés à l'énumération de chaque ménage de l'UPS lors de la pré-enquête, la Direction de la Statistique a décidé d'énumérer une plus petite unité au sein de chaque UPS. Les UPS ont été divisées en petits blocs appelés "lots". Les 16 ménages ont été sélectionnés à partir d'un lot choisi au hasard. Les lots au sein des UPS ont été sélectionnés avec une probabilité égale, ce qui veut dire que la taille de l'lot n'a pas influé sur la sélection. (Les lots sont petits par rapport aux UPS sélectionnées dans le cadre du Bloc 1 mais on ne dispose pas d'autres renseignements sur la situation géographique des lots au sein des UPS).

Tirage de ménages au sein des UPS

Tous les ménages dans chaque lot tiré du Bloc 2 ont été énumérés. Seize ménages étaient choisis au hasard à partir de la liste des ménages énumérés pour chaque lot. Biais dans la sélection des ménages au sein des UPS, données du Bloc 1.

Une analyse des données collectées sur les quatre ans, a révélé une réduction de la taille du ménage (non-pondéré) de l'ordre de 24% entre 1985 et 1988. Trois hypothèses de ce changement étaient : (1) la variation de la taille du ménage reflétait un changement démographique réel, (2) elle était due aux erreurs de mesures non liées au sondage, ou (3) biais dans le plan de sondage. Les études visant à expliquer la variation observée ont d'emblée éliminé les deux premières explications (pour plus de détails sur la question, voir Coulombe et Demery 1993, et Demery et Grootaert 1993). La troisième hypothèse a été étudiée avec beaucoup plus d'attention, puisqu'une modification de la base de sondage intervenue entre deux enquêtes pouvait avoir des effets sur les résultats. En effet, les études conduites par Coulombe et Demery (1993) démontrent que la réduction considérable de la taille des ménages enregistrée au cours des années est apparue entre les premières et les secondes données de panels en 1987, c'est-à-dire entre la fin des données du Bloc 1 et le début des données du Bloc 2. Sur cette base, il est raisonnable d'attribuer les variations de 26% de la taille du ménage observées aux modifications de la base de sondage et des modalités de sondages survenues entre les deux blocs. En effet, les procédures d'énumération des données du Bloc 1 indiquent un biais vers la sélection des ménages de grande taille. À la suite d'une discussion avec le consultant Christopher Scott, Grootaert et Demery (1993, p.269) ont proposé les explications suivantes : "Dans l'unité primaire de sondage où l'énumération des ménages devait avoir lieu, les enquêteurs étaient tenus de choisir le point de départ de l'énumération de façon aléatoire, à partir de ce point, ils devaient collecter des informations dans le même ménage situé au pas de sondage défini, jusqu'à l'énumération de 64 ménages. Cette procédure d'énumération comporte deux sources possibles de biais. D'une part, il est possible que la sélection du premier ménage de départ n'ait pas été aléatoire, mais sujette à un biais attribuable à l'énumérateur (telle que la sélection d'un point de départ où il y a beaucoup de ménages ou une location accessible). D'autre part, en pratique les enquêteurs comptaient le pas en fonction des portes d'entrée des maisons et non les résidences elles-mêmes. Cette approche conduit à un échantillon biaisé si le nombre de portes d'entrée varie suivant les maisons. La probabilité de sélection d'une résidence à deux portes est deux fois plus élevée que celle d'une résidence à une porte. Étant donné que les grandes résidences ont généralement plus de portes que les logements de taille moyenne, le fait de compter les logements à partir du nombre de portes d'entrée pourrait être l'origine du biais de l'échantillon, conduisant à une sur-représentation des logements de grande taille et par conséquent de grands ménages. En effet, la moyenne de la superficie des logements enregistrés était substantiellement plus élevée en 1985 et 1986 que celle enregistrée en 1987 et 1988, corroborant cette interprétation. Comme explique Christopher Scott (1994), le biais lié à la sélection de grands ménages dans l'échantillonnage est renforcé car "là où plus de deux ménages partageaient une porte, l'enquête voulait que seulement un ménage soit sélectionné. En l'absence des instructions spécifiques, il y aurait une tendance naturelle de choisir le ménage principal, renforçant ainsi le biais pour de larges ménages." Dans le Chapitre VI sur "L'utilisation des Données", un ensemble de coefficients de pondération est proposé pour corriger ce biais (coefficients de pondération des tailles de ménages).

Écarts au plan de sondage

Problèmes liés à l'estimation imprécise de la population des UPS

Scott et Amenuvegbe (1991) ont étudié le problème du fait que les estimations de population utilisées pour sélectionner les UPS avec une probabilité proportionnelle à la taille sont souvent incorrectes. L'erreur augmente avec le nombre d'années qui se sont passées depuis le recensement précédent. L'étendue de l'erreur devient apparente quand les listes des UPS sont mises au point et la taille "correcte" de la population est connue. Cet écart entre les estimations des tailles des UPS et la taille "correcte" de

la population doit être corrigée soit en variant le nombre de ménages dans l'échantillon dans chaque UPS ou soit en appliquant des coefficients de pondération chaque UPS, m'/n ou m'/m , où m' est le nombre de ménages que l'on retrouve au sein de l'UPS "i" au moment du dénombrement et n ou m est la mesure de taille (n = population, m = nombre de ménages) utilisée lors du tirage des UPS.

Comme cela a été mentionné plus tôt, l'enquête EPAM a fixé le nombre de ménages 16 par grappe, et de grappes une par UPS. Ainsi, toute tentative de remédier ce problème doit faire intervenir des coefficients de pondération m'/n . Cependant, il n'est pas possible de calculer ces coefficients que ce soit pour les données du Bloc 1 (m' n'est pas disponible puisque seulement 64 ménages par grappe étaient dénombrés), ou pour les données du Bloc 2 (m' n'est pas non plus disponible puisque le lieu de dénombrement était un lot et non pas une UPS).

Taux de réponse

Afin de maintenir la taille de l'échantillon, l'enquête EPAM de 1985 a institué une procédure qui permet de remplacer les ménages de l'échantillon initial, qu'on n'a pas pu enquêter pour des raisons de refus ou de déplacement. Le ménage de substitution est choisi parmi les 48 ménages supplémentaires de la grappe retenue durant le dénombrement. Durant la pré-enquête, l'enquêteur collectait des informations sommaires relatives à la taille et aux caractéristiques socio-économiques du ménage. Ainsi, durant la première année de l'enquête les ménages ont été remplacés (selon nécessité) par un des 48 ménages de remplacement ayant les caractéristiques et taille similaires. Durant la première année de l'enquête, 124 ménages parmi les 1600 initialement retenus (7,8 % de l'échantillon total) n'ont pas été interviewés et par conséquent ont été remplacés par d'autres ménages. La raison majeure de non-réponse des ménages durant l'enquête provient de l'incapacité de l'enquêteur à localiser le logement ou l'adresse du ménage.

Pondération

Cf méthode d'échantillonnage.

Questionnaires

Aperu

Le questionnaire ménage

Le questionnaire auprès des ménages a été appliqué la totalité des ménages en deux visites ou "passages" qui se sont faites deux semaines d'intervalle. La collecte des données en deux passages offre quelques avantages. Un seuil de rappel peut être tablé pour la collecte des revenus et dépenses des ménages lors du deuxième passage, ce seuil tant la première visite de l'enquêteur. De même la correction et la vérification des données erronées peuvent se faire lors du deuxième passage de l'enquêteur dans le même ménage. La collecte des données en deux tapes offre l'avantage supplémentaire d'éviter une longue interview en faveur de deux interviews plus courtes.

Le document de travail de Grootaert (1986) donne une présentation détaillée du questionnaire ménage de la première année d'enquête (1985). La description suivante du questionnaire ménage EPAM présente les principaux thèmes abordés dans chaque section, tout en mentionnant les modifications apportées au questionnaire lors des enquêtes ultérieures. Les sections 16 (données anthropométriques) et 17 (identification des ménages de panels) qui ne figuraient pas dans le questionnaire initial, ni dans la présentation de Grootaert (1986), mais qui apparaissent dans les enquêtes subséquentes, sont décrites dans ce document. Ce questionnaire était presque entièrement précodé. Toutefois, les réponses à certaines questions n'étaient pas précodées.

Le questionnaire communautaire

Le questionnaire communautaire mesure l'accès de la collectivité à l'infrastructure économique, aux services sociaux de base. Ce questionnaire communautaire collecte des données dans chaque grappe rurale. L'objectif de ce questionnaire était de mesurer les caractéristiques communes à tous les ménages de la grappe considérée. Les questions sont adressées à un groupe de personnes les mieux informées et ayant une connaissance détaillée du village. En général, le groupe de personnes interrogées est composé du chef du village et de quelques notables de la collectivité.

Le questionnaire sur les prix

L'enquête sur les prix et l'enquête sur les ménages d'EPAM faisaient partie du même projet et ont été conduites conjointement. L'objectif principal de l'enquête sur les prix était de mettre à disposition des utilisateurs des données de prix à partir desquelles les indices de prix régionaux peuvent être construits. Les composantes du questionnaire de l'enquête sur les prix sont les suivantes :

1. Liste des biens couramment achetés suivant les résultats de l'Enquête Budget-Consommation de 1979 en Côte d'Ivoire.
2. Liste des biens dont les données étaient déjà collectées par la division de l'indice des prix de la Direction de la Statistique.
3. Liste des biens de consommation de la section du questionnaire ménage sur l'auto-consommation. (Ceci permettrait une conversion de la valeur des biens de consommation produits par le ménage en quantité).

L'enquête sur les prix a collecté les informations sur les prix alimentaires et non-alimentaires vendus sur le principal marché de chaque grappe où l'enquête a été conduite. Les données sur les prix ont été enregistrées au même moment où se réalisait la collecte des données auprès des ménages. Quand il y avait au moins trois commerçants vendant le même produit sur le même marché, mais des points de vente différents, trois prix étaient enregistrés pour chaque bien. La marchandise était pesée et le poids enregistré en grammes. Aucun des articles mesurés pendant l'enquête n'était acheté par les enquêteurs. Si un vendeur refusait que ses articles soient mesurés, il fallait attendre un client pour que l'information sur la vente soit enregistrée.

Durant les deux premières années de l'enquête 1985-1986, l'information collectée portait sur 22 biens dont 18 biens de consommation alimentaire et quatre biens de consommation non-alimentaire. Dans l'enquête sur les prix, un numéro était associé à chaque bien, ces numéros ont été utilisés pour identifier les variables des données de l'enquête sur les prix. Le chiffre qui précède chaque bien dans la liste suivante correspond à la numérotation des biens dans le questionnaire de l'enquête sur les prix.

Collecte des données

Dates de la collecte des données

Dbut	Fin	Cycle
1986-02-22	1987-01-24	12 MOIS

Période

Dbut	Fin	Cycle
1985-02-22		12 MOIS

Mode de collecte de données

Interview face face [f2f]

NOTES SUR LA COLLECTE DES DONNES

Organisation et réalisation de l'enquête sur le terrain

Enquête pilote des questionnaires

Plus de 80 ménages ont été contactés pendant l'enquête pilote des questionnaires sur le terrain, qui a duré deux semaines et demi. L'équipe a contacté les individus et les ménages engagés dans diverses activités et résidant dans les quartiers pauvres et riches des deux plus grandes agglomérations, et résidant dans des petites villes, dans les villages, dans la cte, dans les régions du sud-est spécialisées dans la culture du café et du cacao, et dans les villages de la savane du nord spécialisés dans l'agriculture de subsistance.

Procédures pour la réalisation de l'enquête sur le terrain

Les données de l'enquête EPAM ont été collectées par cinq équipes de collecte, dont deux équipes étaient stationnées à Abidjan et trois opéraient dans les antennes régionales de la Direction de la Statistique, Abengourou, Bouaké, et Man. Chaque équipe était composée d'un superviseur et de deux enquêteurs, d'un anthropomètre, d'un opérateur de saisie et d'un chauffeur.

Les modalités de l'organisation de l'enquête sont décrites par Ainsworth et Munoz (1986). L'équipe chargée de collecter les données couvrait deux grappes de ménages en quatre semaines. Durant la première semaine, les enquêteurs interrogeaient les ménages de la grappe A. La deuxième semaine ils collectaient des informations relatives au premier passage du questionnaire dans la grappe B. La troisième et la quatrième semaines étaient consacrées à la collecte des informations du deuxième passage des grappes A et B respectivement.

La réalisation de l'enquête sur le terrain et dans chaque grappe se faisait de la façon suivante (Ainsworth et Munoz, 1986, p.22). Après avoir collecté les informations relatives du premier passage de l'enquête dans une grappe donnée (Sections 1-8 et quand nécessaire Sections 16 et 17), chaque équipe rapportait les questionnaires moitié remplis dans l'antenne régionale. Les opérateurs de saisie enregistraient les données en utilisant un programme de saisie spécialement adapté pour détecter les erreurs et les incohérences flagrantes. Le programme de saisie produisait pour chaque ménage, une liste de toutes les données saisies, ainsi qu'une liste des erreurs dans chaque section et entre les sections. Dans l'antenne régionale, le superviseur procédait à une revue du questionnaire et des listings dans l'objectif de détecter des inconsistances liées soit à la collecte ou à la saisie. Les erreurs liées à la collecte étaient encadrées sur le questionnaire, et les erreurs de saisie étaient notées sur les listings. Durant le deuxième passage, les enquêteurs repositionnaient les questions du premier passage dont les réponses erronées ou inconsistantes avaient été notées par le superviseur avant de procéder aux sections du deuxième passage du questionnaire (Sections 9-15). À l'antenne régionale, l'opérateur de saisie faisait les corrections nécessaires des erreurs de collecte du premier passage avant de procéder à la saisie des données collectées durant le deuxième passage. L'opérateur de saisie devait enfin produire un listing des données collectées durant les deux passages de l'enquête. Afin de s'assurer de la compatibilité entre les données collectées par les enquêteurs et le listing généré par l'antenne régionale, le superviseur vérifiait le listing produit par l'opérateur de saisie. Toute erreur était corrigée sur l'imprimé et modifiée plus tard par les opérateurs.

Durant le premier passage, l'anthropomètre enregistrait le poids et la taille de tous les membres du ménage. Durant le deuxième passage, les membres du ménage dont le poids et la taille relevés, paraissaient erronés par le programme de saisie étaient mesurés à nouveau et pesés. De plus, le programme de saisie sélectionnait au hasard 20% des individus dont les mesures de la taille et le poids devaient être collectés lors du deuxième passage de l'enquête. Dans le processus de collecte des données, l'une des responsabilités importantes du superviseur était de contrôler la qualité des données. Cette responsabilité du superviseur comprenait : codification des réponses qui n'étaient pas pré-codées dans le questionnaire, conduite des interviews de contrôle dans 25% des ménages sélectionnés de manière aléatoire dans chaque grappe, s'assurer que toutes les parties du questionnaire étaient correctement remplies avant de les envoyer à l'antenne régionale, et révision des listings des deux passages du questionnaire afin de détecter les erreurs de collecte et de saisie et superviser la correction de toutes les erreurs sur le terrain et l'antenne régionale.

Les modifications des procédures pour la réalisation de l'enquête étaient minimales au cours des quatre années de l'enquête et ont été apportées dans le but d'améliorer la supervision (Daho, 1992). Certains enquêteurs ont été affectés à d'autres fonctions, mais ces

changements n'ont apparemment pas affecté la qualité des données. La définition du ménage n'a pas changé au cours des quatre années d'enquête.

Enquêteurs

Nom	Abbréviation	Affiliation
INSTITUT NATIONAL DE LA STATISTIQUE	INS	MINISTÈRE DU PLAN

SUPERVISION

Dans le processus de collecte des données, l'une des responsabilités importantes du superviseur était de contrôler la qualité des données. Cette responsabilité du superviseur comprenait : codification des réponses qui n'étaient pas pré-codées dans le questionnaire, conduite des interviews de contrôle dans 25% des ménages sélectionnés de manière aléatoire dans chaque grappe, s'assurer que toutes les parties du questionnaire étaient correctement remplies avant de les envoyer à l'antenne régionale et révision des listings des deux passages du questionnaire afin de détecter les erreurs de collecte et de saisie et superviser la correction de toutes les erreurs sur le terrain et l'antenne régionale.

Les modifications des procédures pour la réalisation de l'enquête étaient minimales au cours des quatre années de l'enquête et ont été apportées dans le but d'améliorer la supervision (Daho, 1992). Certains enquêteurs ont été affectés à d'autres fonctions, mais ces changements n'ont apparemment pas affecté la qualité des données. La définition du ménage n'a pas changé au cours des quatre années d'enquête.

Traitements des données

Edition des données

Il y a un consensus général que la qualité des données de l'enquête auprès des ménages est assez bonne. Une étude informelle menée par Ainsworth et Mehra (1988), et relative à la qualité des données collectées lors des enquêtes EPAM en 1985 et 1986, évalue les données suivant trois critères : leur exactitude, leur cohérence interne et leur état au complet. Sur la base de ces trois critères les données EPAM auprès des ménages sont de très bonne qualité.

Exactitude

. La qualité des données est meilleure quand les réponses aux questions proviennent des membres du ménage directement concernés et non des personnes tierces. L'étude rétrospective des données EPAM auprès des ménages de 1985 et 1986 montre que 93% des femmes interviewées avaient répondu pour elles-mêmes la section consacré la fécondité et que 79-80% de tous les membres du ménage adultes avaient répondu eux-mêmes aux questions sur leur situation d'emploi. Le pourcentage d'enfants répondant directement aux questions qui leur étaient posées variait entre 43 et 45%. Toutefois, ces taux demeurent élevés par rapport aux taux enregistrés pour l'Enquête auprès des Niveaux de Vie du Prou (soit 29 pourcent).

Ensemble des Données

. Une évaluation des diverses variables et modules d'EPAM dont sexe, âge, caractéristiques des parents, éducation, santé, emploi, migration, fécondité, agriculture et entreprises familiales non agricoles, montre que le pourcentage des données manquantes était relativement faible. Dans la plupart des cas, les taux fluctuaient entre (0,01 et 0,05%) et dans tous les cas n'excédaient pas 0,76 %. Cohérence interne des Données.

. La vérification de la cohérence interne des données au Tableau XI montre qu'il y avait très peu d'erreurs. Le questionnaire des ménages était presque entièrement précodé, limitant ainsi les erreurs de codification. De même, un processus décentralisé de saisie des données permet une correction immédiate des erreurs détectées par le programme de saisie. Les codes d'identification personnels et ceux des ménages étaient enregistrés dans chaque section, facilitant ainsi la liaison des données travers les sections. Une comparaison des données collectées dans le cadre du recensement de la population de 1988 et des données de l'enquête EPAM de la même année montre une grande similarité dans la distribution de la taille des ménages. Pour une évaluation brève de la qualité des données EPAM voir Grootaert (1992). "Le contrôle de la qualité durant la collecte et la saisie des données EPAM ont été extraordinaires par rapport à la pratique courante (voir Ainsworth et Munoz 1986; Grootaert, 1986; Daho, 1992). La plupart des études analytiques conduites à partir de ces données montrent un degré élevé de cohérence interne. Les variables démographiques d'EPAM sont en accord avec les résultats de l'Enquête Ivoirienne sur la Fécondité (Ainsworth, 1989). De même, une comparaison des données agricoles EPAM et celles obtenues à partir de sources différentes confirme la qualité des données (Deaton et Benjamin, 1988). Il y avait une cohérence entre l'évolution des dépenses des ménages en Côte d'Ivoire et celle des principaux agrégats macro-économiques affichés par les services de la comptabilité nationale. Les tendances à la hausse en 1986 et la tendance à la baisse enregistrée en 1987-1988 sont mises en évidence par les données EPAM. En résumé, les données EPAM sont une source fiable dans l'étude et l'évolution de bien-être et de la pauvreté..." (p.36).

Autres traitements

Une autre évaluation de la qualité des données EPAM, "The CILSS : A Preliminary Assessment of Data Quality" par James Daniel et Lionel Demery, examine les variables démographiques à partir de la fiche de composition du ménage et trouve que les erreurs estimées à partir des observations totales sont très faibles. La qualité des données s'est améliorée au fil des années. L'étude a aussi examiné la consistance des données des ménages panels au cours des enquêtes consécutives et trouve que la proportion des erreurs était relativement faible : 1%, quand la question portait sur le sexe des membres du ménage, et relativement plus élevée quand l'âge des membres était l'information recherchée. (L'information était considérée erronée si la différence d'âge excédait un an). Toutefois, compte tenu qu'une grande proportion d'Africains vivant dans les zones rurales ne connaissent pas leur âge exact et que la collecte des informations sur l'âge devenait une devinette, ce résultat probablement n'est pas lié à la collecte. Les études consacrées à l'évaluation détaillée de la qualité des données des enquêtes communautaires et des enquêtes sur les prix ne sont pas disponibles.

Evaluation des donnes

Estimation des erreurs d'échantillonnage

Problèmes liés à l'estimation imprécise de la population des UPS

Scott et Amenuvegbe (1991) ont étudié le problème dû au fait que les estimations de population utilisées pour sélectionner les UPS avec une probabilité proportionnelle à la taille sont souvent incorrectes. L'erreur augmente avec le nombre d'années qui se sont passées depuis le recensement précédent. L'étendue de l'erreur devient apparente quand les listes des UPS sont mises au point et la taille "correcte" de la population est connue. Cet écart entre les estimations des tailles des UPS et la taille "correcte" de la population doit être corrigé soit en variant le nombre de ménages dans l'échantillon dans chaque UPS ou soit en appliquant des coefficients de pondération à chaque UPS, m_i/n ou m_i/m , où m_i est le nombre de ménages que l'on retrouve au sein de l'UPS "i" au moment du dénombrement et n ou m est la mesure de taille (n = population, m = nombre de ménages) utilisée lors du tirage des UPS.

Comme cela a été mentionné plus tôt, l'enquête EPAM a fixé le nombre de ménages à 16 par grappe, et de grappes une par UPS. Ainsi, toute tentative de remédier à ce problème doit faire intervenir des coefficients de pondération m_i/n . Cependant, il n'est pas possible de calculer ces coefficients que ce soit pour les données du Bloc 1 (m_i n'est pas disponible puisque seulement 64 ménages par grappe étaient dénombrés), ou pour les données du Bloc 2 (m_i n'est pas non plus disponible puisque le lieu de dénombrement était un lot et non pas une UPS).