

REPUBLIQUE DU MALI
Un Peuple - Un But – Une Foi

MINISTERE DE L'ECONOMIE ET DES FINANCES
INSTITUT NATIONAL DE LA STATISTIQUE



ENQUETE MODULAIRE ET PERMANENTE AUPRES DES MENAGES

Projet de Document Méthodologique

Août 2010

Sommaire

INTRODUCTION.....	1
1. CADRE INSTITUTIONNEL.....	2
2. OBJECTIFS.....	2
2.1. Objectif global.....	2
2.2. Objectifs spécifiques.....	2
3. METHODOLOGIE.....	3
3.1. Le questionnaire.....	3
3.1.1. Le module de base.....	3
3.1.2. Les modules supplémentaires.....	5
3.2. Échantillon.....	5
3.2.1. Observations préparatoires.....	5
a) Contraintes de précision.....	5
b) Plan de sondage.....	5
c) Les strates d'enquête.....	6
d) La taille de l'échantillon en ménages.....	6
e) Critères de choix du nombre de ménages par SE.....	7
3.2.2. Répartition de l'échantillon entre les strates.....	8
3.3. Coefficients d'extrapolation.....	9
3.4. Organisation de l'enquête.....	10
3.4.1. Le personnel.....	10
3.4.2. Recrutement et formation des agents.....	11
3.4.3. Contrôle de la qualité du travail sur le terrain.....	11
3.4.4. Gestion des questionnaires.....	12
3.5. Saisie des données.....	12
3.6. Analyse des résultats.....	12

INTRODUCTION

Le Gouvernement du Mali a marqué sa volonté de répondre aux besoins en informations crédibles pour formuler, mettre en oeuvre et évaluer le CSLP, les OMD et les divers programmes et politiques sectoriels, en adoptant en 2006 le Schéma Directeur de la Statistique (SDS). La mise en oeuvre de ce Schéma a conduit à certaines reformes institutionnelles aussi bien au niveau de la structure centrale de production statistique qu'au niveau des services sectoriels.

Malgré ces efforts de reforme, le Système Statistique National reste confronté à des difficultés dont l'insuffisance de coordination des activités statistiques au niveau national. L'INSTAT et les autres structures sectorielles continuent de réaliser des enquêtes ad hoc auprès des ménages à des coûts souvent très élevés sans une réelle prise en compte des besoins des utilisateurs. Aussi ces enquêtes produisent souvent les mêmes indicateurs avec des résultats contradictoires.

Par ailleurs, le SSN souffre d'une insuffisance de ressources tant humaines que financières dont il est nécessaire de mettre en synergie afin d'en tirer le maximum de profit.

C'est pourquoi l'INSTAT, en collaboration avec Statistique Suède, a initié les réflexions sur la mise en place d'un système de collecte permettant de fédérer l'ensemble des enquêtes auprès des ménages.

Dans ce cadre, deux consultants suédois de court terme ont séjourné à Bamako pour discuter avec l'ensemble des structures chargées de la production statistique des conditions de mise en oeuvre de ce dispositif de collecte. La première mission a eu lieu du 26 au 30 octobre 2009 et la seconde du 1^{er} au 12 mars 2010.

Pour s'enquérir de l'expérience d'un pays ayant déjà implémenté une enquête similaire, une mission de l'INSTAT accompagnée des assistants suédois s'est rendue au Lesotho du 13 au 19 juin 2010.

Cette opération qui sera réalisée sous forme modulaire présente un triple avantage :

- (i). réduction des coûts des opérations à travers une capitalisation des acquisitions ;
- (ii). amélioration de la qualité des données ;
- (iii). Production d'un flux continu de statistiques pertinentes en fonction des besoins des utilisateurs, en particulier pour suivre le développement du marché du travail, les données de consommation, pour les comptes nationaux, régionaux et trimestriels et des informations pour le suivi de la pauvreté.

Elle doit être mise en œuvre avec l'implication de l'ensemble des acteurs concernés pour la planification des différents modules.

1. CADRE INSTITUTIONNEL

L'enquête modulaire et permanente auprès des ménages dénommée **EMOP** sera réalisée par l'Institut National de la Statistique (INSTAT) avec l'assistance technique du projet d'assistance de Statistique Suède. L'enquête est mise en œuvre par le Gouvernement Malien avec l'appui technique et financier de l'Agence Suédoise pour la Coopération au Développement International (Asdi). Pour la suite, le financement serait assuré principalement par le budget national.

Il sera institué un cadre de dialogue permanent, à travers le sous comité de statistiques démographiques et sociales, afin d'assurer l'implication de toutes les structures du SSN. Aussi, l'enquête reste ouverte à tous les utilisateurs désireux d'inclure un module supplémentaire. Le coût additionnel sera donc à la charge du demandeur.

2. OBJECTIFS

2.1. Objectif global

L'objectif principal de l'EMOP est de produire de façon régulière et permanente des indicateurs pertinents sur la situation socio-économique des ménages, données nécessaires au suivi du CSCRP, des OMD et à la formulation de différentes politiques sectorielles.

Les indicateurs utilisés dans l'étude répondent à des normes de mesure internationales, ce qui facilitera la comparaison des résultats avec ceux des autres pays. Les résultats seront représentatifs au niveau national, régional et par milieu résidence en privilégiant les analyses selon le sexe afin de faire ressortir les questions liées au genre.

De plus, l'EMOP cherche à renforcer la capacité de l'INSTAT et des autres structures du SSN à observer et à interpréter les données sur les conditions de vie dans la perspective d'organiser cette opération de façon continue.

2.2. Objectifs spécifiques

L'EMOP vise les objectifs spécifiques suivants :

- Fournir des informations sur la taille et la composition des ménages ;
- Identifier les mouvements de migration interne et externe et les tendances à l'urbanisation ;

- Disposer des informations régulières sur le logement, la santé, l'emploi, l'éducation, les revenus etc.;
- Mesurer l'accès aux principaux services d'infrastructure.
- Quantifier l'activité productive des ménages et mesurer la situation économique des différents ménages ; apprécier les stratégies choisies par les ménages maliens pour faire face aux besoins économiques. Evaluer les dotations économiques à l'intérieur du ménage.

D'autres thématiques peuvent être incluses en fonction des besoins des utilisateurs ou à la demande d'autres structures sectorielles.

3. METHODOLOGIE

La mise en œuvre du projet commence par l'élaboration de certains outils pour le module de base. Il s'agit : du support de collecte (les questionnaires), des manuels (manuel du superviseur, manuel du contrôleur, manuel de l'enquêteur) pour le suivi des travaux de collecte, des applications informatiques pour la saisie des données collectées, des plans de tabulation et d'analyse)

3.1. Le questionnaire

Le questionnaire de l'EMOP comporte un module de base et un ou plusieurs modules supplémentaires qui sont insérés à la demande des utilisateurs ou d'autres structures sectorielles de statistique.

3.1.1. Le module de base

Le module de base sera conçu pour prendre en compte les préoccupations du CSCRP et des OMD ainsi que les exigences des comptes nationaux. Il comprend neuf principales sections et aborde les thématiques suivantes:

- une section fondamentale sur les caractéristiques des ménages fournit les informations sur l'âge et le sexe des membres, leur lien avec le chef de ménage, l'état matrimonial, le statut migratoire;
- une section sur l'éducation informe sur la scolarité des membres du ménage, le niveau d'études atteint, les diplômes et l'alphabétisation;
- une section sur la santé renseigne sur la morbidité de certaines maladies et la fréquentation des infrastructures sanitaires;

- une section sur l'emploi informe sur les indicateurs de l'emploi dont la durée de travail au cours des 12 derniers mois, le taux de chômage, etc.;
- une section sur le logement et les possessions du ménage renseigne sur les caractéristiques du logement tels que le type de logement, le nombre de pièces, la nature du toit, le type d'aisance et les biens possédés par le ménage;
- une section sur la migration et les transferts donne des informations sur les mouvements migratoires des membres du ménage et les transferts monétaires reçus;
- une section sur la sécurité alimentaire renseigne sur les difficultés éprouvées par les ménages pour se nourrir;
- une section sur la perception de la pauvreté capte l'opinion de la population par rapport à la pauvreté;
- une section sur la consommation des ménages cerne le niveau de vie ou de bien-être des ménages à travers ses dépenses de consommation.

À part la section Consommation, qui est renseignée tous les trimestres, les autres sections n'apparaissent qu'une ou deux fois pendant l'année. Le tableau ci-dessus donne la répartition des sections entre les trimestres de collecte.

Sections	Trimestre de collecte			
	1	2	3	4
M Membres	1			
M Membres nouveaux		1	1	1
E Education	1			
S Santé 3-7		1		
S Santé (Morbidité) 1-2	1	1	1	1
W Emploi 1-13	1		1	
W Emploi 14-36			1	
L Logement	1			
PM Possessions	1			
MT Migration et transferts		1		
SA Sécurité alimentaire	1		1	
P Pauvreté subjective		1		
C Consommation	1	1	1	1
Grappe	1			

Le module de base peut également inclure un petit nombre de questions qui, n'appartenant pas logiquement aux neuf (9) principales sections, ne sont pas nombreuses pour constituer une section à part. Il peut s'agir des questions liées à la consommation de tabac ou à l'utilisation de téléphones portables.

Il faut préciser qu'il est facile d'inclure ou d'exclure quelques questions d'une année à l'autre en fonction de la variabilité de l'indicateur à calculer.

3.1.2. Les modules supplémentaires

A la demande des utilisateurs, l'EMOP peut intégrer un ou plusieurs modules supplémentaires au cours d'une phase donnée. Ces modules peuvent consister à approfondir l'une des sections du module de base ou à introduire un domaine nouveau.

Le coût additionnel de ce module supplémentaire sera pris en charge par le demandeur.

3.2. Échantillon

3.2.1. Observations préparatoires

a) Contraintes de précision

Traditionnellement, les résultats des enquêtes d'envergure nationale réalisées au Mali sont représentatifs au niveau national et régional mais aussi au niveau de chaque milieu (Urbain et Rural) avec un niveau de précision acceptable. Pour satisfaire ces exigences, la méthodologie de l'EMOP propose une marge d'erreur relative admissible au seuil de 10% et de 15% respectivement pour les indicateurs nationaux et ceux régionaux.

b) Plan de sondage

Comme la plupart des enquêtes auprès des ménages réalisées au Mali, le plan de sondage proposé est un plan classique stratifié à deux degrés :

- ✓ Au premier degré, on tire les unités primaires (UP=SE) d'échantillonnage avec une probabilité proportionnelle à leur taille en ménages.
- ✓ Au second degré, on tire un nombre constant d'unités secondaires (US = Ménages) de façon aléatoire après avoir dressé la liste de tous les US à la suite d'un dénombrement exhaustif de l'UP.

c) Les strates d'enquête

La stratification dans les enquêtes au Mali est basée sur le milieu (urbain ou rural) et les neuf régions, ce qui donne 17 strates (Bamako, la capitale, est constituée uniquement d'une strate urbaine).

d) La taille de l'échantillon en ménages

En général, la détermination de la taille de l'échantillon exige de faire un arbitrage entre les contraintes de coûts et les exigences de précisions. Le compromis serait alors de répondre à la question suivante:

Quelle taille d'échantillon doit-on s'attendre avec un coût minimum et une précision suffisante ?

Comme variable d'intérêt pour dimensionner l'échantillon, on pourrait choisir l'incidence de pauvreté. Selon les résultats de l'ELIM 2006, la pauvreté monétaire (mesurée comme la capacité de satisfaire les besoins de base) varie entre les régions de 0,29 à 0,81, sauf à Bamako où elle est de 0,08. Cela indique que la précision requise doit être obtenue pour un intervalle assez large couvrant presque toutes les proportions. Pour cette raison, et en tenant compte de tous les autres indicateurs qui seront obtenus, l'approche de la détermination de la taille de l'échantillon en utilisant la proportion d'une variable d'intérêt n'a pas été utilisée.

Comme alternative, une série de calculs sont présentés pour démontrer comment une taille d'échantillon fixée va influencer la gamme des proportions où la précision souhaitée est obtenue.

Par exemple, pour une taille totale d'échantillon de 3600 ménages, on constate qu'il correspond une précision relative désirée de 0,10 au niveau national pour toutes les proportions de plus de 10%.

Les calculs sont basés sur l'intervalle de confiance de 95% pour une estimation de la proportion p :

$$p^* \pm 2 \sqrt{\frac{p(1-p)DEFF}{n(1-t)}}$$

Où :

n : représente la taille de l'échantillon dans n'importe quel domaine examiné.

t : est la proportion anticipée de non-réponse.

DEFF: représente l'effet de sondage.

La précision *relative* est obtenue en divisant la partie de la droite par p .

C'est cette quantité qui doit être tout au plus égale à 0,10 et à 0,15 respectivement pour les indicateurs présentés au niveau national et régional:

$$2 \sqrt{\frac{p(1-p)DEFF}{N(1-t)}} * \frac{1}{p} \leq 0,10 \quad (1)$$

$$2 \sqrt{\frac{p(1-p) DEFF}{n(1-t)}} * \frac{1}{p} \leq 0,15 \quad (2)$$

Où N et n désignent respectivement la taille de l'échantillon au niveau national et régional.

L'analyse des résultats obtenus à la suite des calculs présentés à l'Annexe B nous amène à conclure qu'un échantillon de $N = 3600$ ménages est amplement suffisant si les indicateurs ne sont requis qu'au niveau national.

Pour les indicateurs régionaux, il est nécessaire d'augmenter la taille de l'échantillon à $N = 6300$ ménages ou plus afin d'obtenir la précision requise au niveau régional.

Une raison supplémentaire pour un échantillon relativement grand est qu'une partie importante de l'EMOP est constituée de variables de budget des ménages (et non des proportions), pour lesquels aucun calcul de précision n'a été réalisé jusqu'à présent.

Il est donc recommandé de commencer l'EMOP avec un échantillon de **$N = 7200$** ménages, mais d'analyser les résultats dès que possible afin d'établir, dans les prochains cycles d'EMOP, les effets de la réduction de la taille de l'échantillon à $N = 6300$ par exemple.

e) Critères de choix du nombre de ménages par SE

Une fois la taille de l'échantillon déterminée, on se penche sur la question du nombre de ménages à enquêter dans chaque SE. La pratique des enquêtes ménages au Mali est telle que dans chaque SE, un nombre constant de ménages soit (**a**) est tiré puis interviewés par la suite. Dans les enquêtes antérieures réalisées au Mali, il y a des cas de 6, 10, 12, 15, 30 ménages enquêtés par unité primaire d'échantillonnage.

La manière habituelle de déterminer (**a**) s'appuie sur l'effet de sondage DEFF et les coûts c_1 et c_2 associés à l'introduction d'une unité de première étape et une unité de deuxième étape dans l'échantillon. Aucune détermination de type n'a été constatée dans de précédents documents méthodologiques. Toutefois, il est assez courant que (**a**) dans la gamme de 6 - 12 unités est suffisant pour capter la variation dans une SE. Évidemment, le nombre de ménages par grappe dépend aussi de la capacité de chaque équipe de collecte.

La décision préliminaire est donc d'interviewer 6 ménages dans chaque SE. Ce choix devrait être évalué lorsque les données du premier cycle de l'EMOP vont être disponibles.

3.2.2. Répartition de l'échantillon entre les strates

Les données du RGPH 2009 donnent une indication de la structure de la base de sondage en fonction de l'effectif des ménages et des SE par région et par strate. Ces informations sont résumées dans le tableau ci-après:

Tableau 1 Nombre de ménages et nombre de SE par région et milieu, RGPH 2009

Région	Nombre de ménages			Nombre de SE		
	Urbain	Rural	Total	Urbain	Rural	Total
Kayes	42 711	266 083	308 794	290	2 630	2 920
Koulikoro	25 212	341 599	366 811	181	3 224	3 405
Sikasso	66 819	343 630	410 449	404	3 164	3 568
Ségou	31 348	359 768	391 116	175	3 150	3 325
Mopti	38 313	341 641	379 954	222	3 055	3 277
Tombouctou	16 608	104 946	121 554	108	989	1 097
Gao	18 521	71 752	90 273	147	669	816
Kidal	4 906	7 833	12 739	47	96	143
Bamako	288 176	0	288 176	1 539	0	1 539
Total	532 614	1 837 252	2 369 866	3 113	16 977	20 090

Les possibilités méthodologiques utilisées dans la répartition de l'échantillon sur les strates se réfèrent aux procédures qui n'emploient que les tailles des strates (où d'autres caractéristiques des strates qui pourraient être disponibles). La taille d'une strate correspond à l'effectif des ménages qui le composent. Elle est notée N_h où h désigne l'indice de strate.

Pour palier à la variabilité entre les tailles des strates et à la sous représentativité des strates urbaines (à l'exception de Bamako) par rapport à celles rurales, l'EMOP a utilisé l'approche qui consiste à répartir l'échantillon proportionnellement à la racine cubique de N_h . (Le détail de cette procédure figure à l'**ANNEXE C**)

Pour une taille d'échantillon fixée à sept mille deux cent (7200) ménages pour l'ensemble du pays, sachant que six (6) ménages vont être enquêtés par SE, la taille correspondante de l'échantillon en SE serait donc de mille deux cent (1200).

Pour assurer une bonne représentativité du district de Bamako et de la région de Kidal où Bamako s'est considérablement réduit pendant que Kidal est resté faible (Annexe C: Tableau 2a) après l'application de la procédure citée plus haut, il a été envisagé de fixer en avance le nombre de SE dans ces deux régions.

Ainsi, en fixant préalablement les tailles en SE de Bamako et de Kidal respectivement à cent quarante quatre (144) et à quatre vingt (80), on obtient la répartition suivante consignée dans le tableau ci-après:

Tableau 2: Répartition de 1200 SE sur les strates

Région	Strate	Ménages Nh	Échantillon, SE Prop. à $\sqrt[3]{Nh}$	
Kayes	Urbain	42 711	52	
	Rural	266 083	96	148
Koulikoro	Urbain	25 212	44	
	Rural	341 599	104	148
Sikasso	Urbain	66 819	60	
	Rural	343 630	104	164
Ségou	Urbain	31 348	47	
	Rural	359 768	106	153
Mopti	Urbain	38 313	50	
	Rural	341 641	104	154
Tombouctou	Urbain	16 608	38	
	Rural	104 946	70	108
Gao	Urbain	18 521	39	
	Rural	71 752	62	101
Kidal *	Urbain	4 906	39	
	Rural	7 833	41	80
Bamako	Urbain	288 176	144	144
		2 369 866	1200	1200
Urbain		532 614	513	
Rural		1 837 252	687	
		2 369 866	1200	

* La répartition entre Kidal urbain et Kidal rural a été faite sur la base de la population

3.3. Coefficients d'extrapolation

L'unité d'échantillonnage étant le ménage, pour chaque ménage, il faut calculer un coefficient d'extrapolation qui est l'inverse de la probabilité de tirer ce ménage.

Soit les notations suivantes:

- N_h : effectif des ménages dans la strate selon le RGPH 2009 (à Kidal, effectif de la population)
- N_{hi} : effectif des ménages dans la SE i de la strate h selon le RGPH de 2009 (à Kidal, effectif de la population)
- a_h : le nombre de SE tirées dans la strate h
- t_{ihj} : la proportion estimée du segment j choisi dans la SE i de la strate h ¹

¹ Afin de limiter le travail de dénombrement à l'intérieur des SE dont les populations ne répondent plus aux normes fixées, les grandes SE peuvent être segmentées et seul un segment est retenu dans l'échantillon.

- L_{hi} : le nombre de ménages dénombrés par l'EMOP 2011 dans la SE i de la strate h
- n_{hi} : le nombre de ménages tirés dans la SE i de la strate h
- P_{1hi} : la probabilité de sondage au premier degré de la i -ème SE de la strate h
- P_{2hi} : la probabilité de sondage au deuxième degré du i -ème ménage dans la SE i

La probabilité de tirer la SE i au premier degré est $P_{1hi} = \frac{a_h \times N_{hi}}{N_h} \times t_{ihj}$.

La probabilité de tirer un ménage quelconque au second degré est $P_{2hi} = \frac{n_{hi}}{L_{hi}}$.

La probabilité finale de tirer un ménage quelconque de la SE i est donc

$$P_i = P_{1hi} * P_{2hi} = \frac{a_h \times N_{hi}}{N_h} \times t_{ihj} \times \frac{n_{hi}}{L_{hi}}$$

Et le coefficient d'extrapolation, le poids à affecter à chaque ménage sera $W_i = \frac{1}{P_i}$

3.4. Organisation de l'enquête

L'approche qui sera utilisée pour la collecte des informations est l'approche de l'interview directe au cours d'une série de quatre visites dont une visite par trimestre. Pour se faire, les enquêteurs à travers le support de collecte (le questionnaire), auront la charge d'interviewer et de recueillir les informations nécessaires à la réalisation de cette opération.

L'organisation pratique des travaux de l'EMOP exige en moyenne une affectation de dix huit (18) grappes par enquêteur au cours d'un trimestre de collecte sachant que la charge de travail journalière d'un enquêteur est estimée à deux (2) ménages.

Pour un suivi régulier et un contrôle efficace des travaux de terrain, trois (3) agents enquêteurs seront placés sous la responsabilité directe d'un (1) contrôleur. Il s'efforcera de visiter chacun au moins quatre (4) fois par mois. La répartition des visites dans le mois entre ses agents est laissée à son initiative. Cependant, dès les premiers jours, il doit visiter tous les enquêteurs afin de détecter rapidement celui (ou ceux) qui a (ont) le plus de problèmes afin de l'aider (les aider) à les surmonter.

3.4.1. Le personnel

La gestion et la supervision de l'opération sont assurées par la Direction Générale de l'INSTAT qui veillera au respect du calendrier prévisionnel des activités. Elle pourra dans certains cas prendre des décisions, trouver des solutions et même trancher en cas de problèmes.

Une équipe technique formée des cadres des différents départements de l'INSTAT est chargée de l'élaboration des outils et de la méthodologie de l'enquête. Cette équipe travaille en collaboration avec une assistante à long terme de Statistiques Suède. Elle bénéficie également de l'appui technique ponctuel des assistants suédois de court terme.

Au niveau régional, la coordination de l'opération sera assurée par le Directeur Régional de la Planification du Développement, de la Statistique, de l'Informatique, de la Population et de l'Aménagement du Territoire en attendant la mise en place des Délégations Régionales de l'INSTAT.

Le personnel d'exécution viendra compléter cette équipe. Il sera constitué de contrôleurs (22), d'enquêteurs (66), d'agents de dénombrement (35), d'agents de réception de dossier (5), et d'agents de saisie (15).

Le personnel d'appui se compose de : un (1) comptable, un (1) agent administratif.

3.4.2. Recrutement et formation des agents

Il sera procédé au recrutement des enquêteurs au niveau de chaque région. La proximité de la zone d'activité est fondamentale pour la réussite d'une telle opération. Après la définition du profil des enquêteurs, on lancera un avis de recrutement.

a) Les candidats subiront des tests préalables : analyse des dossiers, test de connaissance

b) Ceux qui seront présélectionnés, suivront une formation.

c) La liste définitive des agents de collecte sera obtenue après le test qui sanctionnera la formation. Afin de pallier à la défection de certains, une liste servant de réserve sera constituée.

Un document portant sur le recrutement décrira les procédures à suivre pour cette activité. La formation des personnes recrutées sera dispensée par les membres de l'équipe technique. Elle est axée sur le manuel d'instruction des enquêteurs et contrôleurs et une partie pratique qui comportera des sorties de terrain visant à tester la connaissance et l'habileté des candidats.

Les superviseurs recevront au préalable une formation sur le contenu, l'administration du questionnaire, le contrôle et la supervision des opérations de collecte.

3.4.3. Contrôle de la qualité du travail sur le terrain

L'enquête ou le travail de terrain prendra neuf (9) semaines au cours d'un trimestre. Des procédures strictes seront établies pour contrôler la qualité de l'enquête. Le Contrôleur aura la responsabilité directe du contrôle de la qualité des données recueillies dans le questionnaire et de la performance des enquêteurs. Il vérifiera régulièrement le travail et corrigera s'il y a lieu des erreurs commises par l'enquêteur. D'ailleurs cette correction se fera avec l'enquêteur pour qu'il en tire profit pour la suite du travail. De plus, les responsables régionaux et les membres de l'équipe technique coordonneront les opérations et contribueront aussi à revoir les résultats après l'évaluation de la

qualité des données communiquées. Ils fourniront également une assistance continue et directe et veilleront à la structure de l'organisation du travail de terrain.

3.4.4. Gestion des questionnaires

Chaque questionnaire rempli sera soumis à un contrôle immédiat et sera acheminé à l'INSTAT. Tous les questionnaires subiront un dernier contrôle avant l'opération de saisie. Les questionnaires critiqués seront saisis par des opérateurs préalablement entraînés à cette fin. Les données une fois saisies, seront tabulées et évaluées. Ainsi le résultat final de l'opération sera disponible pour son exploitation.

3.5. Saisie des données

L'activité vérification et de codification précèdera l'activité de saisie ; des agents seront recrutés pour la réalisation de ces activités. L'activité de saisie elle-même va être gérée par des personnes d'encadrement et il y aura des agents également recrutés et sélectionnés. Les personnes sélectionnées suivront une formation qui durera au plus une semaine.

Pour la saisie des données le logiciel choisi est Csprow. Le programme de saisie et les tests de cohérence seront réalisés à la phase de la définition et de la préparation des outils de l'enquête. La tabulation se fera sur SPSS.

3.6. Analyse des résultats

Pour analyser les résultats de l'enquête, on fera recours aux logiciels SPSS et STATA. Une équipe d'analyse sera mise en place pendant que le travail de terrain se poursuit. L'analyse consistera à décrire les résultats préliminaires selon le plan d'analyse préalablement conçu et à rédiger un rapport dans lequel les principaux résultats de l'enquête seront présentés. Cette approche vise à pouvoir présenter les résultats préliminaires dans le plus court délai possible, après la collecte des informations sur le terrain. La rédaction du rapport final prendra plus de temps, car elle implique un travail concerté entre les experts des différentes structures.

ANNEXES

ANNEXE A

TABLEAU DE BORD DES INDICATEURS A MESURER					
Intitule	Périodicité	Désagrégation géographique	Désagrégation selon le genre	Trimestre de collecte	Variables de calcul
1. Structure de la population	1 an	nationale, régionale, milieu	Sexe, groupe d'âge et GSE	1	M3, M4, M6, M10a
2. Taille moyenne des ménages	1 an	nationale, régionale, milieu		1	M1
3. Age moyen des chefs de ménage	1 an	nationale, régionale, milieu	Sexe, groupe d'âge et GSE	1	M4, M5
EDUCATION					
4. Taux d'alphabétisation des 15 à 24 ans	1 an	Nationale, région, milieu	Sexe, groupe d'âge et GSE	1	M4, E9
Taux brut de scolarisation	1 an	Nationale région, milieu		1	M4, E4, E7
Taux net de scolarisation	1 an	Nationale, région milieu		1	M4, E4, E7
5. Taux d'alphabétisation des 15 ans et plus	1 an	Nationale, région milieu	Sexe, groupe d'âge et GSE	1	M4, E9
6. Pourcentage de ménages vivant à moins de 30mn d'une école fondamentale 1er cycle et 2nd cycle	2 ans	Nationale, milieu, Régional		1	Module Grappe
7. Pourcentage de ménages vivant à moins de 5 km d'une école fondamentale 1er cycle et 2nd cycle	2 ans	Nationale, milieu, Régional		1	Module Grappe
EMPLOI					
8. Pourcentage des femmes salariées dans le secteur non agricole	1 an	Nationale, milieu, Régional	groupe d'âge et GSE	3	M3, W14, W15
9. Pourcentage des femmes salariées dans le secteur privé	annuelle	Nationale, milieu, Régional	groupe d'âge et GSE	3	M3, W15, W16
10. Ratio emploi/population*	annuelle	Nationale	Sexe, groupe d'âge et GSE	3	M4, W01 à W12

TABLEAU DE BORD DES INDICATEURS A MESURER

Intitule	Périodicité	Désagrégation géographique	Désagrégation selon le genre	Trimestre de collecte	Variables de calcul
11. Taux de chômage*	annuelle	Nationale, milieu	Sexe, groupe d'âge et GSE	1, 3	W01 à W12
12. Effectif des chômeurs	annuelle	Nationale, milieu	Sexe, groupe d'âge et GSE	1, 3	W01 à W12
13. Effectif des individus sous-employés selon la durée du travail	annuelle	Nationale, milieu	Sexe, groupe d'âge et GSE	3	W20
14. Effectif des individus dont les compétences sont sous-utilisées	annuelle	Nationale, milieu	Sexe, groupe d'âge et GSE	3	E8, W01, W13, W15
15. Effectif des travailleurs à temps complet avec des gains mensuels faibles (entre 40 et 48 h/semaine)	annuelle	Nationale, milieu	Sexe, groupe d'âge et GSE	3	W13, W20, W26, W39
16. Effectif des travailleurs à temps partiel avec des gains faibles horaires (moins de 40h/semaine)	annuelle	Nationale, milieu	Sexe, groupe d'âge et GSE	3	W13, W20, W26, W39
17. Taux d'activité de la population active (15-64 ans)**	annuelle	Nationale	Sexe, groupe d'âge et GSE	3	M4, W1-W12
18. Taux de chômage des jeunes (15-24 ans)**	annuelle	Nationale	Sexe, groupe d'âge et GSE	1,3	M4, W1-W12
19. Taux de travailleurs pauvres*	annuelle	Nationale	Sexe, groupe d'âge et GSE	3	W1, W26, W39
20. Proportion des travailleurs dont la durée du travail excède 48 h par semaine)*	annuelle	Nationale	Sexe, groupe d'âge et GSE	3	W1, W20
21. Proportion des travailleurs dans les emplois précaires*	annuelle		Sexe, groupe d'âge et GSE	3	Pas pris en compte (AP 7 Enq 1, 2,3)
22. Taux d'inactivité	annuelle	Nationale, régionale, milieu	Sexe, groupe d'âge et GSE	3	M4, W4, W11

TABLEAU DE BORD DES INDICATEURS A MESURER

Intitule	Périodicité	Désagrégation géographique	Désagrégation selon le genre	Trimestre de collecte	Variables de calcul
23. Taux de pluriactivité - cf. 27.	annuelle	Nationale, régionale, milieu	Sexe, groupe d'âge et GSE	3	W24,
24. Proportion d'enfants de 6 à 14 ans exerçant une activité économique	annuelle	Nationale, régionale, milieu	Sexe, groupe d'âge et GSE	3	M4, W01
25. Pourcentage de la population de 65 ans ou plus bénéficiant d'une pension*	annuelle	Nationale, région, milieu	Sexe, groupe d'âge et GSE	3	M4, W13
26. Pourcentage des actifs occupés pratiquant au moins une activité secondaire	annuelle	Nationale, régionale, milieu	Sexe, groupe d'âge et GSE	3	W30
27. Revenu mensuel tiré de l'Emploi principal	annuelle	Nationale, régionale, milieu	Sexe, groupe d'âge et GSE	3	W29
28. Revenu mensuel tiré de l'Emploi secondaire	annuelle	Nationale, régionale, milieu	Sexe, groupe d'âge et GSE	3	W45
29. Durée dans le chômage pour les 15 à 24 ans	annuelle	Nationale, régionale, milieu	Sexe, groupe d'âge et GSE	3	M4, W09

SANTE

31. Pourcentage des ménages vivant dans un rayon de moins de 15 km d'un établissement sanitaire	2 ans	Nationale, régionale, milieu		2	Module Grappe, S5
32. Pourcentage des ménages vivant dans un rayon de moins de 5 km d'un établissement sanitaire	2 ans	Nationale, régionale, milieu		2	Module Grappe, S5
33. Pourcentage de la population malade ou blessée au cours des 3 derniers mois	annuelle	Nationale, régionale, milieu	Sexe, groupe d'âge et GSE	1, 2, 3, 4	S1
34. Taux de fréquentation des infrastructures sanitaires	2 ans	Nationale, régionale, milieu	Sexe, GSE	2	S4
35. Proportion de la population morbide ayant un recours thérapeutique moderne			Sexe, GSE	2	S4
36. Proportion des ménages ayant accès à une source d'eau potable	2 ans	Nationale, régionale, milieu			L7(à modifier les modalités)

CARACTERISTIQUES DU LOGEMENT

TABLEAU DE BORD DES INDICATEURS A MESURER

Intitule	Périodicité	Désagrégation géographique	Désagrégation selon le genre	Trimestre de collecte	Variables de calcul
37. Proportion des ménages ayant accès à l'électricité	annuel	Nationale, régionale, milieu		1	L6
38. Pourcentage de la population ayant accès aux crédits logements	annuel	Nationale	Homme, Femme	1	PM5 PM3
39. Proportion des ménages propriétaires de leur logement	annuel	Nationale	Homme, Femme	1	L12
40. Proportion des ménages utilisant des méthodes appropriées pour évacuer les eaux usées	annuel	régionale, milieu		1	L9
41. Proportion des ménages utilisant le gaz comme principale source d'énergie pour la cuisson	annuel	régionale, milieu		1	L11
42. proportion des ménages situés à moins d'1 km d'une source d'eau potable	annuel	régionale, milieu		4	Module Grappe
43. proportion des ménages situés à moins 30 mn d'une source d'eau potable	annuel	régionale, milieu		4	Module Grappe
44. Proportion des ménages utilisant des toilettes appropriées	annuel	régionale, milieu		4	L10
45. Proportion des ménages selon la distance et le temps nécessaire pour atteindre le marché de produits alimentaires le plus proche	annuel	régionale, milieu		4	Module Grappe
46. Proportion des ménages selon la distance et le temps nécessaire pour atteindre les transports publics les plus proches	annuel	régionale, milieu		4	Module Grappe
CONSOMMATION DES MENAGES					
47. Coefficient budgétaire des 12 fonctions de dépense	1 an	nationale, milieu, régionale		1, 2, 3,4	Module Consommation
48. Dépenses de Consommation annuelle moyenne par ménage	1 an	nationale, milieu, régionale		1, 2, 3,4	Module Consommation
49. Dépenses de Consommation annuelle par tête	1 an	nationale, milieu, régionale		1, 2, 3,4	Module Consommation
PAUVRETE MONETAIRE ET SUBJECTIVE					

TABLEAU DE BORD DES INDICATEURS A MESURER

Intitule	Périodicité	Désagrégation géographique	Désagrégation selon le genre	Trimestre de collecte	Variables de calcul
52. Pourcentage des ménages selon l'appréciation de leur situation financière	annuelle	Nationale, régionale, milieu		2	P1 P5
53. Pourcentage des ménages selon l'appréciation de leur niveau de vie par rapport à la situation actuelle de la communauté	annuelle	Nationale, régionale, milieu		2	P7, P8 x
54. Pourcentage des ménages selon leur perception de la pauvreté	annuelle	Nationale, régionale, milieu		2	P2 P4
55. Opinions de la population sur l'action des autorités en matière de lutte contre la pauvreté	annuelle	Nationale, régionale, milieu	Sexe	2	P5 P9, P10
56. Pourcentage de la population ayant connu de difficulté alimentaire au cours de l'année	annuelle	Nationale, régionale, milieu		2	SA1
57. Proportion des ménages selon les stratégies adoptées pour faire face aux difficultés alimentaires	annuelle	Nationale, régionale, milieu		2	SA2
58. Proportion des ménages ayant reçu de l'assistance alimentaire	annuelle	Nationale, régionale, milieu		2	SA2
MIGRATION ET TRANSFERTS					
59. Proportion des ménages ayant au moins un membre vivant à l'étranger	biennal	Nationale, régionale, milieu		2	MT1,
60. Nombre d'émigrés moyen par ménage	biennal	Nationale, régionale, milieu		2	MT2
61. Montant moyen des transferts reçus par les ménages	biennal	Nationale, régionale, milieu		2	MT4, MT5

* = Indicateurs stratégiques de l'OIT

** = Indicateurs additionnels de l'OIT

1. Synthèse rapport de suivi des OMD au Mali

2. Annexe Etude "analyse des déficits en travail décent au Sénégal" -Honoré T Djerma

3. http://www.uis.unesco.org/i_pages/fr/indspec/tecspe_ger.htm - 7k

4. http://www.uis.unesco.org/i_pages/fr/indspec/tecspe_ner.htm - 8k

5. http://www.humanitarianinfo.org/niger/uploads/assessments/PAM_EFSA%20resume%20fr_Nov%2005.pdf

ANNEXE B

Les calculs pour déterminer la taille de l'échantillon

Les calculs évalueront tout simplement la partie d'à gauche dans les formules (1) et (2) pour certaines tailles d'échantillons sélectionnés n , dans chaque cas pour une gamme de valeurs de p , afin de déterminer où la précision relative est assez petite.

Avant que les calculs ne soient effectués, il est nécessaire de déterminer t et DEFF. Pour la non-réponse, $t = 0,03$ est utilisé ce taux correspond bien aux résultats de l'ELIM 2006.

DEFF prend une valeur distincte pour chaque variable et chaque domaine d'études. D'habitude, on l'estime pour la variable d'intérêt à partir d'enquêtes antérieures. Dans les calculs ici, où on n'utilise pas une variable d'intérêt, les résultats sont affichés pour DEFF = 1,0; 1,2; 1,5; 2,0. Ces valeurs sont comparables aux seules valeurs calculées de DEFF qu'on a trouvé au Mali, c'est-à-dire, dans le rapport de l'EDS.

Le nombre total de ménages dans l'échantillon est désigné par N . On fait l'approximation que chaque région a la même taille d'échantillon, qui sera ensuite $N/9$. Les calculs dans les tableaux A - D montrent le côté gauche de la formule (1) et la formule (2) avec $n = N/9$, comparativement aux cibles de 0,10 et 0,15 respectivement.

Dans tous les tableaux, la zone marquée en couleur montre là où la précision n'est pas assez bonne par rapport à la cible ("Target").

Tableau A

National		Target	DEFF =			
N	p		1,00	1,20	1,50	2,00
3600	0,50	0,10	0,03	0,04	0,04	0,05
	0,40	0,10	0,04	0,05	0,05	0,06
	0,30	0,10	0,05	0,06	0,06	0,07
	0,25	0,10	0,06	0,06	0,07	0,08
	0,20	0,10	0,07	0,07	0,08	0,10
	0,15	0,10	0,08	0,09	0,10	0,11
	0,10	0,10	0,10	0,11	0,12	0,14

Un échantillon total aussi petit que $N = 3600$ offrira une très bonne précision pour les indicateurs au niveau national, pour toutes les valeurs de p , sauf celles aussi petites que 10% ou moins et celles de 15% pour un DEFF égal 2. Des tailles d'échantillons plus grandes seront encore plus performantes.

Tableau B

Regional		Target	DEFF =			
N	p		1,00	1,20	1,50	2,00
4500	0,50	0,15	0,09	0,10	0,11	0,13
	0,40	0,15	0,11	0,12	0,14	0,16
	0,30	0,15	0,14	0,15	0,17	0,20
	0,25	0,15	0,16	0,17	0,19	0,22
	0,20	0,15	0,18	0,20	0,22	0,26
	0,15	0,15	0,22	0,24	0,26	0,31
	0,10	0,15	0,27	0,30	0,33	0,39

Au niveau régional, un échantillon total de N = 4500 (500 ménages par région), donnera une précision satisfaisante seulement pour des indicateurs de plus de 30 %, si DEFF est aussi faible que 1,0 ou 1,2, et seulement pour les indicateurs de plus de 40% si DEFF = 1,5 ou plus.

Tableau C

Regional		Target	DEFF =			
N	p		1,00	1,20	1,50	2,00
6300	0,50	0,15	0,08	0,08	0,09	0,11
	0,40	0,15	0,09	0,10	0,12	0,13
	0,30	0,15	0,12	0,13	0,14	0,17
	0,25	0,15	0,13	0,15	0,16	0,19
	0,20	0,15	0,15	0,17	0,19	0,22
	0,15	0,15	0,18	0,20	0,22	0,26
	0,10	0,15	0,23	0,25	0,28	0,33

Lorsque l'échantillon est porté à N = 6300 (700 ménages par région), les indicateurs de l'ordre de 20 – 25 % et plus auront la précision requise.

Tableau D

Regional		Target	DEFF =			
N	p		1,00	1,20	1,50	2,00
7200	0,50	0,15	0,07	0,08	0,09	0,10
	0,40	0,15	0,09	0,10	0,11	0,12
	0,30	0,15	0,11	0,12	0,13	0,16
	0,25	0,15	0,12	0,14	0,15	0,18
	0,20	0,15	0,14	0,16	0,18	0,20
	0,15	0,15	0,17	0,19	0,21	0,24
	0,10	0,15	0,22	0,24	0,26	0,30

Une nouvelle augmentation de la taille de l'échantillon à $N = 7200$ améliore quelque peu la situation pour des valeurs élevées de DEFF, mais l'amélioration est marginale.

Quelques calculs additionnels ont été effectués pour évaluer les précisions obtenues dans les cellules des tableaux basés sur une séparation de plusieurs catégories (par exemple, l'état matrimonial des chefs de ménage, séparés par sexe et zone). Puisque la base de chaque résultat ne consiste que d'une partie de l'échantillon, les résultats de ce type ne peuvent pas atteindre une précision comparable à la précision des taux nationaux ou régionaux.

Les conclusions sont qu'un échantillon de $N = 3600$ est amplement suffisant si les indicateurs ne sont requis qu'au niveau national. Pour les indicateurs régionaux, il est nécessaire d'augmenter la taille de l'échantillon à $N = 6300$ ou plus afin d'obtenir la précision requise. L'influence de l'effet de sondage (DEFF), peut être très importante dans cette situation - mais les valeurs de DEFF n'ont pas été disponibles, et donc une certaine prudence est recommandée. En outre, lorsque les entrées dans les tableaux importants sont considérées, le nombre d'observations qui sont à la base pour les cellules individuelles sera réduit de manière substantielle. Même avec $N = 7200$, beaucoup de proportions ne seront pas estimées avec la précision requise. Une raison supplémentaire pour un échantillon relativement grand est qu'une partie importante de l'EMOP est constituée de variables de budget des ménages (et non des proportions), pour lesquels aucun calcul de précision n'a été réalisé jusqu'à présent.

Il est recommandé de commencer l'EMOP avec un échantillon de $N = 7\ 200$ ménages, mais d'analyser les résultats dès que possible afin d'établir, dans les prochains cycles d'EMOP, les effets de la réduction de la taille de l'échantillon à $N = 6\ 300$ par exemple.

ANNEXE C

L'échantillon de l'EMOP

Préalables

Après les considérations et les décisions du chapitre 3.2 - c'est-à-dire, en bref,

Précisions relatives de 0,10 (national), 0,15 (régional)

Sondage en deux étapes, avec 17 strates dans la première étape

Un échantillon total de 7200 ménages, avec 6 ménages par SE -

Il faut déterminer la répartition de l'échantillon par les strates.

Le tirage des SE dans la première étape sera basée en le recensement de 2009, où on a les nombres de ménages et de SE dans les strates qui sont indiqués dans le Tableau 1.

Tableau 1 Nombre de ménages et nombre de SE par région et milieu, RGPH 2009

Région	Nombre de ménages			Nombre de SE		
	Urbain	Rural	Total	Urbain	Rural	Total
Kayes	42 711	266 083	308 794	290	2 630	2 920
Koulikoro	25 212	341 599	366 811	181	3 224	3 405
Sikasso	66 819	343 630	410 449	404	3 164	3 568
Ségou	31 348	359 768	391 116	175	3 150	3 325
Mopti	38 313	341 641	379 954	222	3 055	3 277
Tombouctou	16 608	104 946	121 554	108	989	1 097
Gao	18 521	71 752	90 273	147	669	816
Kidal	4 906	7 833	12 739	47	96	143
Bamako	288 176	0	288 176	1 539	0	1 539
Total	532 614	1 837 252	2 369 866	3 113	16 977	20 090

La répartition sera effectuée sans référence à une variable d'intérêt, comme on a discuté en déterminant la taille de l'échantillon. Les possibilités méthodologiques sont donc les procédures qui n'emploient que les tailles des strates (où bien des autres caractéristiques des strates qui pourraient être disponibles). La taille d'une strate est son nombre de ménages, N_h .

Il est nécessaire de trouver une répartition qui résout les problèmes suivants:

1) il y a une très grande variation entre les tailles des strates

2) les strates urbaines son petites, à l'exception de Bamako qui est une strate entièrement urbaine.

La méthode standard pour la répartition basée en les tailles des strates est la répartition proportionnelle, où l'échantillon de chaque strate est directement proportionnel à la taille N_h de la strate. Une extension utile de cette méthode est la répartition par puissances, qui augmente la participation des petites strates dans l'échantillon sans perdre beaucoup de précision². Dans la situation sans variable d'intérêt, l'échantillon est réparti proportionnellement à N_h^α . La puissance α est un nombre entre 0 et 1 - le plus souvent, on emploie $\alpha=1/2$ ou $\alpha=1/3$. En autres mots, la répartition est effectuée proportionnellement à la racine carrée de N_h , ou à la troisième racine de N_h . Avec $\alpha=1$, on obtient la répartition proportionnelle classique.

Si on suppose que l'échantillon aura 12 ménages par SE, il faut choisir 600 SE; avec 6 ménages par SE il en faut 1200, pour arriver au total de 7200 ménages. Dans les tableaux 2a et 2b, les répartitions obtenues pour $\alpha = 1, 1/2, 1/3$, pour 600 SE et pour 1200 SE, sont présentées.

Tableau 2a - Répartition de 600 SE

Région	Strate	Ménages		Répartition de l'échantillon, SE				
		Nh	Prop. à Nh		Prop. à $\sqrt{N_h}$		Prop. à $\sqrt[3]{N_h}$	
Kayes	Urbain	42 711	11		23		28	
	Rural	266 083	67	78	57	80	51	79
Koulikoro	Urbain	25 212	6		18		23	
	Rural	341 599	86	92	65	83	55	78
Sikasso	Urbain	66 819	17		29		32	
	Rural	343 630	87	104	65	94	55	87
Ségou	Urbain	31 348	8		20		25	
	Rural	359 768	91	99	66	86	56	81
Mopti	Urbain	38 313	10		22		27	
	Rural	341 641	86	96	65	87	55	82
Tombouctou	Urbain	16 608	4		14		20	
	Rural	104 946	27	31	36	50	37	57
Gao	Urbain	18 521	5		15		21	
	Rural	71 752	18	23	30	45	33	54
Kidal	Urbain	4 906	1		8		13	
	Rural	7 833	2	3	10	18	16	29
Bamako	Urbain	288 176	73	73	59	59	52	52
		2 369 866	599	599	602	602	599	599
Urbain		532 614	135		208		241	
Rural		1 837 252	464		394		358	
		2 369 866	599		602		599	

² Voir par exemple Särndal, Swensson & Wretman, Model assisted survey sampling (Springer, 2003).

Tableau 2b Répartition de 1200 SE

Région	Strate	Ménages	Répartition de l'échantillon, SE					
		Nh	Prop. à Nh		Prop. à $\sqrt{N_h}$		Prop. à $\sqrt[3]{N_h}$	
Kayes	Urbain	42 711	22		46		55	
	Rural	266 083	135	157	114	160	102	157
Koulikoro	Urbain	25 212	13		35		46	
	Rural	341 599	173	186	129	164	111	157
Sikasso	Urbain	66 819	34		57		64	
	Rural	343 630	174	208	130	187	111	175
Ségou	Urbain	31 348	16		39		50	
	Rural	359 768	182	198	133	172	112	162
Mopti	Urbain	38 313	19		43		53	
	Rural	341 641	173	192	129	172	111	164
Tombouctou	Urbain	16 608	8		29		40	
	Rural	104 946	53	61	72	101	75	115
Gao	Urbain	18 521	9		30		42	
	Rural	71 752	36	45	59	89	66	108
Kidal	Urbain	4 906	2		16		27	
	Rural	7 833	4	6	20	36	31	58
Bamako	Urbain	288 176	146	146	119	119	104	104
		2 369 866	1199	1199	1200	1200	1200	1200
Urbain		532 614	269		414		481	
Rural		1 837 252	930		786		719	
		2 369 866	1199		1200		1200	

Exemple: Dans le Tableau 2a, le nombre de SE dans la strate Koulikoro rural, avec répartition proportionnellement à $\sqrt{N_h}$, est calculé comme

$$\sqrt{(341599) / (\sum \sqrt{N_h}) * 600} = 584,5 / 5419,7 * 600 = 64,7$$

La somme est calculée sur toutes les 17 strates.

Si on veut changer la taille total de l'échantillon (en SE), il suffit de changer le nombre 600 dans cette formule.

L'effet d'augmentation des petites strates est atteint, surtout dans le cas $\alpha = 1/3$. L'effet se manifeste pour les régions moins peuplées (Tombouctou, Gao) et pour toutes les strates urbaines (avec l'exception de Bamako). Cette augmentation garantie des précisions meilleures dans les régions mentionnées, ainsi que dans le milieu urbain. Quand même, la région de Kidal reçoit toujours un échantillon très petit, et il sera difficile d'obtenir la précision relative de 0,15 pour les indicateurs qui ont une valeur basse à Kidal (moins de 30 %, environ) si l'échantillon ne contient que 29 ou 58 SE, ou bien 348 ménages.

Dans le district de Bamako, comme dans toutes les strates de population nombreuse, l'échantillon a été réduit - jusqu'à un niveau qui ne permet plus d'obtenir la précision qu'on requière pour les indicateurs régionaux, surtout si l'on considère l'hétérogénéité qu'on peut s'attendre à Bamako. Les calculs généraux présentés à l'Annexe B indiquent qu'on requière environ $7200/9 = 800$ ménages (11 %) dans chaque région; et il sera préférable si une région très hétérogène puisse avoir un échantillon suffisant. Il y a peu de calculs de variances disponibles des enquêtes précédentes pour guider la discussion. On note que Bamako a reçu 16 % de l'échantillon en MICS-ELIM 2009, et 9 % de l'échantillon en ELIM 2006, toutefois sans motivations documentées.

Une solution pour Bamako et Kidal est de fixer en avance le nombre de SE dans ces deux régions, avant de répartir les autres SE. Il est proposé de inclure 480 ménages (ou plus) à Kidal, et au moins 864 ménages à Bamako. Avec 480 ménages, on devrait achever la précision relative pour un intervalle plus ample de proportions (les proportions d'environ 25 % et plus). À Bamako, cet échantillon signifie 12 % de l'échantillon total, et on peut appliquer les résultats de la discussion générale sur la précision qu'on trouve à l'Annexe B.

Il vaut bien souligner que les nombres 480 et 864 (en termes de SE, 40 et 72 SE avec 12 ménages chacune, ou 80 et 144 SE avec 6 ménages par SE) sont choisis après des considérations des précisions relatives qui doivent résulter - mais ces nombres peuvent être changés si on donne des bonnes motivations (qui doivent être documentées). Par exemple, on trouvera peut-être, en analysant des données collectées, que les variables économiques montrent une variance plus élevée à Bamako, ce qui pourrait mener à une augmentation correspondante de l'échantillon de Bamako.

Les résultats obtenus quand on établie d'abord les échantillons en Kidal et Bamako (480 et 864 ménages) et quand on répartie après le reste de l'échantillon ($7200-480-864= 5856$ ménages) aux autres strates, sont présentés dans les Tableaux 3a et 3b. À Kidal, la répartition de l'échantillon fixe de 480 ménages sur les deux strates (urbaine et rurale) est basée sur la racine carrée et la troisième racine de l'effectif (révisé) de la population. La raison est que la base de sondage actualisée (RGPH 2009) ne contient pas le nombre de ménages par SE à Kidal. Cela requiert une modification des coefficients d'extrapolation pour Kidal (N_h et N_{hi} doivent se référer à la population totale, non au nombre de ménages), et on aura une perte de précision, probablement petite, pour les résultats qui se réfèrent aux ménages à Kidal.

Tableau 3a. Répartition de 600 SE avec échantillon de taille fixée a Kidal et Bamako

Région	Strate	Ménages	Échantillon, SE			
		Nh	Prop. à \sqrt{Nh}		Prop. à $^3\sqrt{Nh}$	
Kayes	Urbain	42 711	21		26	
	Rural	266 083	53	74	48	74
Koulikoro	Urbain	25 212	16		22	
	Rural	341 599	60	76	52	74
Sikasso	Urbain	66 819	27		30	
	Rural	343 630	61	88	52	82
Ségou	Urbain	31 348	18		23	
	Rural	359 768	62	80	53	76
Mopti	Urbain	38 313	20		25	
	Rural	341 641	60	80	52	77
Tombouctou	Urbain	16 608	13		19	
	Rural	104 946	34	47	35	54
Gao	Urbain	18 521	14		20	
	Rural	71 752	28	42	31	51
Kidal	Urbain	4 906	19		19	
	Rural	7 833	21	40	21	40
Bamako	Urbain	288 176	72	72	72	72
		2 369 866	599	599	600	600
Urbain		532 614	220		256	
Rural		1 837 252	379		344	
		2 369 866	599		600	

* La répartition entre Kidal urbain et Kidal rural a été faite sur la base de la population

Tableau 3b. Répartition de 1200 SE avec échantillon de taille fixée a Kidal et Bamako

Région	Strate	Ménages	Échantillon, SE			
		Nh	Prop. à \sqrt{Nh}		Prop. à $^3\sqrt{Nh}$	
Kayes	Urbain	42 711	43		52	
	Rural	266 083	107	150	96	148
Koulikoro	Urbain	25 212	33		44	
	Rural	341 599	121	154	104	148
Sikasso	Urbain	66 819	53		60	
	Rural	343 630	121	174	104	164
Ségou	Urbain	31 348	37		47	
	Rural	359 768	124	161	106	153
Mopti	Urbain	38 313	40		50	
	Rural	341 641	121	161	104	154
Tombouctou	Urbain	16 608	27		38	
	Rural	104 946	67	94	70	108
Gao	Urbain	18 521	28		39	
	Rural	71 752	55	83	62	101
Kidal *	Urbain	4 906	38		39	
	Rural	7 833	42	80	41	80
Bamako	Urbain	288 176	144	144	144	144
		2 369 866	1201	1201	1200	1200
Urbain		532 614	443		513	
Rural		1 837 252	758		687	
		2 369 866	1201		1200	

* La répartition entre Kidal urbain et Kidal rural a été faite sur la base de la population

L'échantillon urbain, qui pourrait paraître grand, est de 37 % et de 42 % dans les deux cas, comparable aux 39 % de MICS-ELIM. Les autres différences entre les deux alternatives présentées sont petites. On note que les différences entre les régions sont moins prononcées quand $\alpha = 1/3$. C'est un avantage, car on gagne de précision dans les petites régions, sans perdre beaucoup dans les grandes régions.

Quant à la choix entre 6 ou 12 ménages par SE (ou bien un nombre intermédiaire), c'est une expérience bien établie que les ménages d'une (SE) se ressemblent souvent entre eux, ce qui indique qu'on aura une grande partie des informations pertinentes après avoir collecté les données d'un nombre petit de ménages. Au Mali aussi, on confirme cette expérience. La variation entre les SE est plus importante que la variation interne des SE. Il est donc raisonnable de supposer que les meilleures précisions seront obtenues avec un nombre assez petit de ménages par SE.

Sur la base des calculs et de la discussion ici, il est recommandé que l'échantillon pour l'EMOP sera répartie selon la seconde alternative dans le Tableau 3b, 1200 SE réparties proportionnellement à la troisième racine de la taille de la strate.

En préparant le budget de l'enquête et la logistique de la collecte de données, il est possible qu'on soit obligé de considérer des changements.

La distribution sur les cercles

Quand on tire l'échantillon du premier degré en utilisant le tirage systématique dans chaque strate, avec probabilités proportionnelles à la taille (définie comme le nombre de ménages), on peut facilement prévoir le nombre de SE qu'on aura dans chaque cercle, séparément pour la partie urbaine et la partie rurale du cercle. La condition est seulement que toutes les SE d'un même cercle soient localisées ensemble dans la base du tirage.

Pour faciliter la planification des travaux avant que l'échantillon officiel soit disponible, les nombres de SE prévus par cercle ont été calculés. Pour chaque strate séparément, les nombres son proportionnels au total de ménages dans les cercles du même milieu et de la même région. Il faut noter que les nombres calculés peuvent différer d'une unité (ou rarement deux unités) des nombres achevés dans l'échantillon réalisé.

Il est important de souligner que les calculs ici sont faits seulement pour aider en la planification des travaux de terrain. Les résultats de l'enquête ne seront pas présentables sur le niveau de cercle.

Le Tableau 4 présente les nombres prévus dans la région de Kayes, avec deux strates (urbain et rural) où on avait obtenu 52 et 96 SE dans la répartition décrite ci-dessus.

Tableau 4. Répartition des SE prévue, par cercle, dans la région de Kayes, avec un échantillon ppt systématique

		<i>SE urbain</i>		<i>SE rural</i>					
		148		52		96			
		Pop		Ménages		SE		SE par cercle	
						Prop mén x 52		Prop mén x 96	
Kayes	Kayes	Urbain	150 389	23 569	145	25,83		26	
		Rural	362 973	57 840	551		21,25	21	47
	Bafoulabe	Urbain	2 938	507	7	0,56		1	
		Rural	230 988	34 973	392		12,85	13	14
	Diema	Urbain	17 507	2 709	23	2,97		3	
		Rural	194 555	30 346	271		11,15	11	14
	Kenieba	Urbain	11 282	2 010	20	2,20		2	
		Rural	182 871	32 281	326		11,86	12	14
	Kita	Urbain	54 282	8 475	58	9,29		9	
		Rural	380 097	54 466	566		20,01	20	29
	Nioro	Urbain	51 619	7 745	62	8,49		8	
		Rural	178 869	27 635	261		10,15	10	18
	Yelimane	Urbain	17 209	2 442	19	2,68		3	
		Rural	161 233	23 796	219		8,74	9	12
Kayes	Urbain		47 457	334			52	148	
	Rural		261 337	2 586			96		

Exemple: Dans la partie rurale du cercle de Kita, on prévoit $54466/261337*96 = 20$ SE.

Les calculs ont été réalisés pour toutes les régions. Le nombre prévu par cercle varie entre 7 et 58. Il y a quelques cercles où on ne prévoit que 0, 1 ou 2 SE dans la partie urbaine.

Le problème des grandes SE

Dans le sondage pour MICS-ELIM, on a observé quelques SE très grandes dans la base de la région de Kidal - si grandes qu'elle serait inclus plus qu'une fois dans l'échantillon. On a proposé alors de refaire la cartographie pour identifier des nouvelles SE d'une taille moyenne, non seulement à Kidal, mais partout où ça serait nécessaire. Si on a en effet réalisé cette proposition, il n'y aura pas de problèmes de ce type pour le sondage l'EMOP. Si non, il faudra résoudre les problèmes quand ils arrivent. La solution plus naturelle est de segmenter une SE qui se trouve dans l'échantillon si elle est trop grande. Il est nécessaire alors de se rendre compte des proportions des parties maintenues de la SE, en comparaison avec la taille total, en nombre de ménages. Si on n'a pas la possibilité d'arriver à des proportions exactes, on peut continuer les travaux avec des proportions estimées. Ce qui comporte une certaine perte de précision. Les proportions des segments utilisés doivent être incluses dans les coefficients d'extrapolation, voir chapitre 3.3.