



Organisation des Nations Unies
pour l'alimentation et l'agriculture

BASE DE SONDAGE PRINCIPALE (BSP) POUR LES STATISTIQUES AGRICOLES

Module 3:

**Plan d'échantillonnage et estimation de la
BSP lors de l'utilisation de bases de type
liste**

Objectifs de la présentation

- Présenter différents plans d'échantillonnage et diverses techniques d'estimation lorsque la BSP est une base de sondage de type liste.
- Fournir des lignes directrices relatives à l'utilisation d'informations auxiliaires en vue d'améliorer le plan d'échantillonnage.

Plan

- 1) Techniques d'échantillonnage et estimation lorsque la BSP est une base de sondage de type liste.
- 2) Utilisation d'informations auxiliaires pour améliorer le plan d'échantillonnage

Introduction

- Le plan d'échantillonnage probabiliste est très important dans les statistiques agricoles:
 - Les estimations peuvent être calculées avec un **intervalle de confiance** et **une précision**.
 - Ce type de plan d'échantillonnage est largement adopté dans les enquêtes agricoles.
- Incidence d'un plan d'échantillonnage inadapté (erreurs d'échantillonnage)
 - Pourrait avoir une incidence sur la portée de l'étude (couverture, population échantillonnée)
 - Peut être à l'origine de biais de sélection (mauvaise procédure d'échantillonnage)
 - Cela pourrait avoir une incidence sur les estimations



Plan
d'échantillonnage
lorsque la BSP est
une base de
sondage de type
liste

1.0 Plan d'échantillonnage lorsque la BSP est une base de sondage de type liste

**Le plan d'échantillonnage dépend du type de base liste.
Par exemple:**

- Liste des ménages → échantillonnage à deux degrés
- Liste des exploitations ou des fermes → échantillonnage à un seul degré ou à deux degrés
- Liste des entreprises agricoles commerciales ou des grandes exploitations foncières → inventaire complet

Dans le secteur des ménages, **des procédures d'échantillonnage complexes** sont généralement utilisées; impliquant l'échantillonnage à deux degrés associé à la stratification, à la PPT, à l'échantillonnage aléatoire, etc.

1.1 Quelles procédures d'échantillonnage pour quel type de BSP?

1. Échantillon stratifié vs échantillon non stratifié

- Procédure d'échantillonnage aléatoire
 - Échantillonnage aléatoire simple
 - Échantillonnage systématique
 - Échantillonnage stratifié
 - Grappes et échantillonnage à plusieurs degrés: échantillonnage en grappes
- Pour les enquêtes agricoles à grande échelle, l'échantillonnage stratifié est le plus communément utilisé
 - Réduire la dispersion de l'échantillon
 - Produire de meilleures estimations
 - Prendre en compte la structure de la population
 - Arguments naturels en faveur de la stratification
 - Climat
 - Saison
 - Utilisation des terres / couverture des sols

Quelles procédures d'échantillonnage pour quel type de BSP?

2. Importance des informations auxiliaires

L'utilisation d'informations auxiliaires pour définir:

- Stratification
- Plan d'échantillonnage
- Répartition de la taille de l'échantillon (p. ex. par strates)
- Guider de la procédure d'échantillonnage (PPT ou PPTM)
- Choix de la méthode de sélection

Quelles procédures d'échantillonnage pour quel type de BSP?

2. Importance des informations auxiliaires

Exemples d'informations auxiliaires

- Taille totale de la population agricole par ZD
- Superficie totale cultivée d'une région spécifique
- Production totale d'une récolte spécifique dans une région donnée
- Écart-type pour une variable spécifique
- Moyenne de la variable (âge, production, superficie,...)
- Utilisation des terres ou couverture des sols

1.2. Échantillonnage stratifié

1. Définir des strates dans une base de sondage de type liste

Variables communes utilisées comme des strates:

- Zone agro-écologique
 - Climat
 - Niveau des précipitations
 - Saison

- Régions ou districts
 - Structure géographique, ventilation (Expérience du Népal)

- Potentiel productif des exploitations agricoles / des ménages
 - Niveau de mécanisation
 - Pratiques d'agriculture intensive

1.2. Échantillonnage stratifié

2. Unités d'échantillonnage

- L'unité d'échantillonnage dépendra des unités dans la base de sondage et des informations auxiliaires disponibles pour chaque unité de la base.
- Ci-dessous quelques exemples d'unités d'échantillonnage:
 - Ménages, Exploitants, Exploitations;
 - Exploitant individuel au sein des ménages;
 - Coopératives de production;
 - Conditionneurs modernes et traditionnels (produits semi-finis, produits finis);
 - Les personnes impliquées dans les activités de commercialisation (acheteurs/commerçants);
 - Centres de recherche et centres de semences;
 - Services de vulgarisation.

1.3. Échantillonnage stratifié à un seul degré ou à deux degrés

Définitions

- **Échantillonnage à plusieurs degrés:** Méthode qui, pour l'agriculture, utilise de grandes zones ou grappes géographiques comme premier degré.
- Dans un **plan d'échantillonnage à deux degrés**, les grappes (**UPE**) sont sous-échantillonnées et les unités secondaires échantillonnées sont les unités déclarantes (**USE**).
- **Échantillonnage à un seul degré:** schéma d'échantillonnage dans lequel l'échantillon est sélectionné directement à partir d'une liste des unités couvertes par l'enquête.

1.3. Échantillonnage stratifié à un seul degré ou à deux degrés

1. Unité primaire d'échantillonnage

UPE Approche de la zone de dénombrement (ZD) vs. Approche du village

- Le village est généralement une unité administrative de base composée de quartiers, eux-mêmes formés de groupes de membres d'une ou de plusieurs familles et éventuellement des hameaux qui y sont rattachés.
- Il est dirigé par un chef qui est désigné ou intronisé selon les règles en vigueur dans la société concernée. le village s'intègre, par conséquent, dans le découpage administratif d'un pays.

1.3. Échantillonnage stratifié à un seul degré ou à deux degrés

1. Unité primaire d'échantillonnage

UPE Approche Zone de dénombrement (ZD) vs. Approche Village

- Les ZD ont été définies pour faciliter le travail sur le terrain (recensement et enquêtes). Les ZD, en règle générale, correspondent à la charge de travail du recenseur: environ 800 habitants en zone rurale et 1500 habitants en zone urbaine.
- Étant donné l'homogénéité relative de la ZD par rapport au village du point de vue de la taille, de l'identification, etc., l'ensemble de ZD répond mieux aux caractéristiques de la base d'échantillonnage.
- La Zone de dénombrement est la plupart du temps utilisée comme Unité primaire d'échantillonnage (UPE) dans un échantillonnage à plusieurs degrés.

1.3. Échantillonnage stratifié à un seul degré ou à deux degrés

1. Unité primaire d'échantillonnage

UPE: Approche de la zone de dénombrement (ZD) vs. Approche du village

- ZD comme UPE, exemple:
 - Namibie
 - Côte d'Ivoire
 - Ghana
 - Éthiopie
 - Népal
 - Géorgie
 - Bulgarie
 - Etc.

1.3. Échantillonnage stratifié à un seul degré ou à deux degrés

1. Unité primaire d'échantillonnage

UPE: Approche axée sur la zone de dénombrement (ZD) vs. Approche axée sur le village

- Le village comme Unité primaire d'échantillonnage (les unités secondaires d'échantillonnage seront sélectionnées dans le village). Cela peut être efficace lorsque le recensement est récent.
- Toutefois, cette approche comporte de nombreux désavantages. Le nombre et la taille des villages varient fréquemment. L'approche du village entraîne un biais d'échantillonnage dans la mesure où:
 - Seuls les exploitants qui sont dans les principaux villages pendant l'enquête sont inclus dans l'échantillon
 - Ceux qui se trouvent dans les hameaux rattachés au village en sont exclus
 - Ce risque est élevé en raison de la forte dispersion des zones cultivées et également souvent de la grande taille des villages

1.3. Échantillonnage stratifié à un seul degré ou à deux degrés

Unité primaire d'échantillonnage

UPE: Approche de la zone de dénombrement (ZD) vs. Approche du village

- Village comme UPE:
 - Évaluation de l'impact des projets communautaires:
 - Côte d'Ivoire: Projet V4C «pour une cacaoculture durable en Côte d'Ivoire»
 - Nigéria: Projet WAAPP, évaluation des bénéficiaires
 - Ghana: Projet V4C «pour une cacaoculture durable», Bonne pratique agricole: Élagage, désherbage, écussonnage.

1.3. Échantillonnage stratifié à un seul degré ou à deux degrés

1. Unité primaire d'échantillonnage

UPE: Approche de la zone de dénombrement (ZD) vs. Approche du village

- Village comme UPE:
 - Enquêtes agricoles nationales:
 - Chine
 - Burkina Faso

1.3. Échantillonnage stratifié à un seul degré ou à deux degrés

2. Unité secondaire d'échantillonnage (USE)

- Les types d'USE sont généralement:
 - Ménages (Éthiopie, Lesotho)
 - Exploitations (Géorgie)

Expérience de l'Éthiopie

- La liste des ZD est utilisée comme UPE et la liste des ménages dans les ZD échantillonnées est utilisée comme unité secondaire d'échantillonnage.
- Les ZD sont sélectionnées par échantillonnage avec PTT systématique, la taille étant le nombre de ménages tiré du RPH.
- Les ménages sont ensuite répertoriés et les ménages de l'échantillon sont sélectionnés à partir de la liste.

L'expérience du Népal

- Le plan d'échantillonnage adopté était un échantillonnage stratifié à deux degrés avec le district comme strate.
- Les unités primaires d'échantillonnage (UPE) étaient soit une circonscription individuelle soit une sous-circonscription ou un groupe de circonscriptions contiguës et les unités secondaires d'échantillonnage (USE) étaient les exploitations agricoles avec une publication des données au niveau district.
- Chacun des 75 districts était considéré comme une strate.
- La première étape de la sélection a été faite à l'aide d'un échantillonnage avec probabilité proportionnelle à la taille systématique (PPT systématique) avec le nombre d'exploitations dans les USE comme mesures de taille.
- La sélection des exploitations agricoles au cours de la deuxième étape a été faite à l'aide d'un tirage systématique à probabilité égale.

1.4. Estimation

- Les échantillons aléatoires sont toujours fortement privilégiés, car seuls les échantillons aléatoires permettent l'**inférence statistique**:
 - Utiliser des données d'enquête par sondage pour calculer une statistique à partir de laquelle il est possible d'estimer les paramètres de la population.
- **Une statistique** est une quantité calculée à partir d'un échantillon qui est utilisée pour estimer les paramètres de population et décrire l'échantillon.
- **Un estimateur** d'un paramètre θ est une statistique T_n qui prend une valeur dans l'ensemble des valeurs possibles de θ .
- Avec un échantillonnage aléatoire nous pouvons aussi calculer la **variance** du paramètre estimé.

1.4. Estimation

1. Estimateur de Horvitz-Thompson:

- Les estimateurs de Horvitz-Thompson (1952) sont sans biais pour les paramètres tels que la moyenne, le total et les pourcentages.
- Concentrons-nous sur la difficulté à estimer un total de population (paramètre):

$$Y = \sum_{k \in U} y_k$$

L'estimateur de Horvitz-Thompson pour Y est donné par:

$$\hat{Y} = \sum_{k \in S} \frac{y_k}{\pi_k}$$

- Où π_k est la **probabilité d'inclusion**. L'inverse de π_k est le **poids d'échantillonnage** qui est également appelé le **facteur d'expansion**.

1.4. Estimation

2. Poids d'échantillonnage en appliquant la PPT dans un échantillonnage à deux degrés

- 1) **Premier degré:** Échantillonnage PPT → les grandes UPE ont une probabilité plus élevée d'être échantillonnées

$$P_{1i} = n_1 \frac{M_i}{N}$$

M_i = Taille de l'UPE (nombre d'USE dans l'UPE)

N = Taille de la population d'UPE (nombre d'USE dans la population)

n_1 = Nombre d'UPE qui doivent être sélectionnées au premier degré

P_{1i} = Probabilité que l'UPE i soit échantillonnée dans le premier degré

1.4. Estimation

2. Poids d'échantillonnage appliquant la PPT dans un échantillonnage à deux degrés

2) Deuxième degré: Échantillonner exactement le même nombre d'individus par UPE → les individus dans une grande UPE ont une probabilité plus petite d'être échantillonnés

$$P_{2j} = \frac{n_{2i}}{M_i}$$

M_i = Taille de l'UPE (nombre d'USE dans l'UPE)

n_2 = Nombre d'USE à échantillonner dans chaque UPE au deuxième degré

P_{2ij} = Probabilité de chaque individu j soit échantillonné dans UPE i

1.4. Estimation

2. Poids d'échantillonnage appliquant la PPT dans un échantillonnage à deux degrés

3) Finalement: Le deuxième degré compense le premier degré, de telle sorte que chaque individu dans la population a la même probabilité d'être échantillonné

$$P_{ij} = P_{1i} * P_{2ij} = n_1 \frac{M_i}{N} * \frac{n_{2i}}{M_i} = n_1 \frac{n_2}{N}$$

- Cela signifie que tous les USE dans la population ont la même probabilité de sélection, quelle que soit la taille de leur UPE.
- Le poids de sondage est l'inverse de la probabilité de sélection.

→ Poids de sondage = $1/P_{ij}$

1.5 L'estimation de la superficie totale des terres consacrées à une culture spécifique (Éthiopie)

- Dans les formules suivantes \hat{A}_h évalue la superficie totale consacrée à une culture spécifique de la strate h:

Où:

- h représente la strate
- n_h est le nombre total de ZD tirés dans la hème strate
- M_h est la mesure de la taille de la hème strate obtenue à partir de la base d'échantillonnage
- m_{hi} est la mesure de la taille de la ième ZD de l'échantillon obtenues à partir de la base d'échantillonnage
- H_{hi} est la taille de la ième ZD échantillonnée dans la hème strate
- h_{hi} est le nombre de ménages agricoles tirés dans la ième ZD de l'échantillon dans la hème strate
- a_{hij} est la valeur de la superficie totales des parcelles pour les ménages agricoles j, dans la ième ZD dans la hème strate consacrée à une culture spécifique
- a_{hi} est la superficie totale des parcelles consacrées à une culture spécifique, évaluée dans l'échantillon pour la ZD i de la strate h

1.5 L'estimation de la superficie totale des terres consacrées à une culture spécifique (Éthiopie)

La variance d'échantillonnage pour l'estimation du total de la superficie de la strate

$$Var(\hat{A}_h) = (1 - f_h) \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\hat{A}_{hi} - \frac{\hat{A}_h}{n_h} \right)^2 + f_h \sum_{i=1}^{n_h} (1 - f_{hi}) \left(\frac{h_{hi}}{h_{hi} - 1} \right) \sum_{j=1}^{h_{hi}} \left(\hat{A}_{hij} - \frac{\hat{A}_{hi}}{h_{hi}} \right)^2$$

Où

- f_h est la probabilité moyenne du premier degré de sélection de ZD dans la strate h.
- $f_{hi} = \frac{h_{hi}}{H_{hi}}$ probabilité moyenne du deuxième degré de sélection dans le ième échantillon de ZD dans la strate h.
- \hat{A}_{hi} est la superficie totale pondérée d'une culture spécifique dans la ième ZD et la hème strate.
- \hat{A}_{hij} sont les valeurs pondérées de la superficie du jème ménage agricole dans la ième ZD et les hèmes strates consacrées à une culture spécifique.



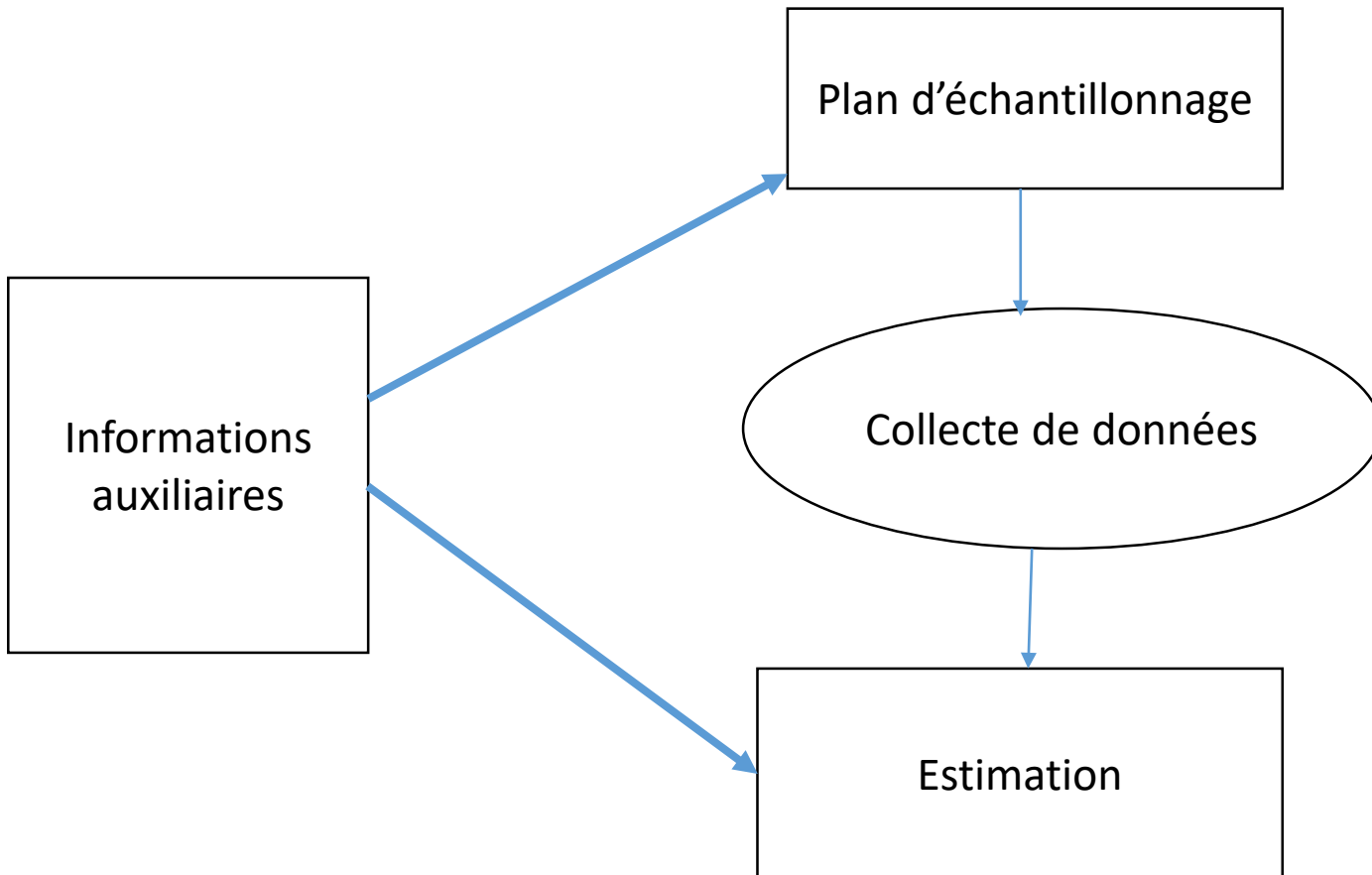
Informations auxiliaires pour améliorer le plan d'échantillonnage

2.1 Concept

- La notion d'information auxiliaire regroupe toutes les informations externes à l'enquête elle-même qui permettent d'augmenter la précision des résultats d'une enquête.
- De façon générale, toutes les informations connues sur la population sont appelées informations auxiliaires.
- Ces informations peuvent être la connaissance des valeurs d'une ou de plusieurs variables sur toutes les unités de la population ou simplement une fonction de ces valeurs.

2.1 Concept

- Les informations auxiliaires peuvent être utilisées à deux moments: Au stade de la conception et au stade de l'estimation des paramètres.



2.2. Utiliser des informations auxiliaires dans l'échantillonnage

- L'introduction d'informations auxiliaires peut avoir deux objectifs:
 - Améliorer la précision des estimations pour un coût donné
 - Améliorer l'organisation de l'enquête
- Utiliser des informations auxiliaires peut conduire à transformer la probabilité d'inclusion du premier degré, du second degré, ou les deux en même temps.

2.2. Utiliser des informations auxiliaires dans l'échantillonnage

Échantillonnage à probabilité inégale

- Il est particulièrement intéressant lorsque les variables sont liées par un effet de taille.
- Si une variable auxiliaire X permet de mesurer approximativement cet effet, il est recommandé de sélectionner les unités d'observation avec une probabilité proportionnelle à cette variable auxiliaire.
- Le gain en précision sera alors très important.

2.2. Utiliser des informations auxiliaires dans l'échantillonnage

Stratification

- La technique de stratification classique améliore presque toujours la précision des estimateurs.
- Il est nécessaire de disposer d'informations auxiliaires qui permettent d'assigner chaque unité à une strate.
- La probabilité d'inclusion ne doit pas nécessairement être identique d'une strate à l'autre.
- En rappel, avec l'allocation de Neyman, plus la dispersion de la variable d'intérêt est forte dans une strate, plus la partie de l'échantillon qui doit être attribuée à la strate est importante.
- Il est très important que les variables auxiliaires permettent non seulement de stratifier les unités statistiques, mais également de mesurer la dispersion au sein de chaque strate.

2.2. Utiliser des informations auxiliaires dans l'échantillonnage

Échantillonnage à plusieurs degrés

- Un plan d'échantillonnage typique consiste à sélectionner les unités primaires avec probabilités inégales proportionnelles au nombre d'unités secondaires.
- Ensuite, un nombre fixe d'unités secondaires est sélectionné à partir des unités primaires sélectionnées.
- Au premier degré de l'échantillonnage, nous avons besoin d'informations auxiliaires pour appliquer le tirage à PPT.
- Parfois, les UPE ou les USE sont stratifiées pour une plus grande précision.

2.2. Utiliser des informations auxiliaires dans l'échantillonnage

Expériences des pays

- Utilisation le nombre de ménages du RPH comme la taille pour la PTT:
 - Éthiopie
 - Lesotho
 - Géorgie
 - Népal
 - Etc.
- Utiliser des informations auxiliaires pour la stratification:
 - Bulgarie
 - Éthiopie
 - Géorgie
 - Chine
 - Népal
 - Etc.

Résumé des expériences des pays

Pays	Étapes	Stratification et variables auxiliaires	UPE	USE ou unités statistiques	Sélection UPE	Échantillonnage pour des exploitations agricoles spéciales
Bulgarie	Un seul degré	Localisation géographique, taille et spécialisation des exploitations agricoles		Exploitations	NA	Prendre toutes les unités
Éthiopie	Deux degrés:	géographique	ZD	Ménage	PPT- taille (nombre de ménages)	Prendre toutes les unités
Géorgie	Deux degrés	Géographique Région Taille de l'exploitation Type d'exploitation	Grappes des strates	Exploitations	PPT- taille (nombre de ménages)	Prendre toutes les unités
Népal	Deux degrés	district Taille des exploitations	Circonscription ou sous-circonscription ou un groupe de circonscriptions contiguës	Exploitations	PPT- taille (nombre de ménages)	Prendre toutes les unités

Discussion Plan d'échantillonnage dans votre pays

Résumé

- **L'échantillonnage à deux degrés** est approprié pour les enquêtes agricoles, en particulier pour le secteur des ménages.
- Plus souvent, les UPE (généralement des ZD) sont d'abord sélectionnées avec PPT.
- Et les listes d'exploitations et/des ménages sont tirées de ces unités échantillonnées.
- Il est recommandé qu'un échantillonnage à deux degrés soit associé à **la stratification fondée sur les informations auxiliaires**.
- Faire l'inventaire intégral est la conduite préconisée pour les exploitations agricoles modernes ou spéciales.

Références

- Le Département des affaires économiques et sociales du secrétariat des Nations Unies; 2005. *Designing Household Survey Samples: Lignes directrices pratiques*. New York.
- Le Département des affaires économiques et sociales du secrétariat des Nations Unies; 2005. *Household Sample Surveys in Developing and Transition Countries*. New York.
- UNSD, Banque Mondiale, FAO. 2010 *Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales*. Rome.
- Voir: <http://www.gsars.org/>

Références

- FAO, 2017. *Master Sampling Frames for Agriculture- Supplement on selected Country Experiences* . Rome.
- FAO. 2015. *Programme mondial pour le recensement de l'agriculture 2020*, Rome
- Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales. 2015. *Manuel sur les Bases de sondage principales pour les statistiques agricoles. Publication Global strategy, Rome*
- Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales. 2016. *Guidelines for the integrated survey framework*. Rome.

MERCI